ESSAYS IN INDUSTRIAL ORGANIZATION AND APPLIED MICROECONOMICS

Vera Sharunova

A dissertation
submitted to the Faculty of
the Department of Economics
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

February 2023

ESSAYS IN INDUSTRIAL ORGANIZATION AND APPLIED MICROECONOMICS

Vera Sharunova

Advised by:
Julie Holland Mortimer, Ph.D.
Michael D. Grubb, Ph.D.
Charles T. Murry, Ph.D.

Abstract

This dissertation consists of three self-contained essays that explore industrial organization of the advertising agency and airline markets, as well as the role of local political environment in identification of treatment effects in minimum wage studies.

In the first chapter, “Competitor Avoidance and the Structure of the Advertising Agency Industry in the United States,” co-authored with Sylvia Hristakeva, we develop an applied theory model to show that the tendency of advertisers to avoid sharing their agencies with product-market competitors may have led to creation of a unique organizational structure in the advertising agency market, known as a holding company (HC). HCs control multiple agencies and coordinate their bidding choices when competing for new clients. Although many other professional service markets, such as markets for legal and accounting services, feature competitor avoidance, HCs are forbiddingly costly in these markets due to restrictions on outside ownership.

Using a theoretical model, we show that HC structure helps agencies manage client conflicts by allowing them to choose an unconflicted agency to bid for a client. We collect a novel dataset on identities of bidding agencies and estimate that serving an advertiser’s competitor reduces an agency’s odds to compete for the advertiser by 91.6 percent. We predict that the market concentration would increase by 35 percent if competitor avoidance was not a factor in this market. We also predict

that banning bid coordination within HCs would increase the average number of bidders in an account review from four to nine. Auction theory predicts that an increase in the number of bidders would create a downward pressure on the mark-ups charged by agencies, however some of this pressure may be counteracted by increased costs of winning a client due to entering multiple bids.

In the second chapter, "'Use It or Lose It", or "Cheat and Keep?" The Effects of Slot Restrictions on Airline Incentives," co-authored with Ratib Ali, we investigate the impact of slot control on competition in the domestic airline market. The Federal Aviation Administration manages congestion in high-density airports by capping the number of flights permitted in any given hour and allocating the rights (or slots) to a take-off or landing among airlines. Airlines must use their slots at least 80% of the time to keep them for the next season. This rule creates a perverse incentive for airlines to hold on to underutilized slots by operating unprofitable flights instead of forfeiting these slots to a rival. Using exogenous removal of slot control at the Newark Airport in 2016, we investigate the lengths at which airlines go to meet the minimum requirements that let them keep the slots while violating what a neutral observer might call the "spirit" of the regulation.

In the third chapter, "Political Trends in Minimum Wage Policy Evaluation Studies," co-authored with Andrew Copland and Jean-François Gauthier, we explore the role of local political environment in identification of treatment effects in minimum wage studies. The effects of minimum wage on employment in low-wage sectors have long been debated in the literature. Some economists find small disemployment effects, whereas others argue that these effects are close to zero and statistically insignificant. The core of the debate lies in establishing adequate control groups for areas that experience minimum wage changes. At the same time, minimum wage changes are almost always a consequence of a political vote. Our paper adds to the debate surrounding control group identification by highlighting the importance of accounting for underlying political trends. Failure to do so may result in a violation of the standard "parallel trends" assumption maintained in most of the literature. We illustrate this possibility by re-estimating Dube et al. (2010) on a sample of politically aligned and unaligned counties and controlling for state expenditures that may be used to finance confounding policies. We document that the sample of never politically aligned county pairs produces a positive and significant estimate of elasticity of employment (0.245), suggesting that the restaurant industry labor market may be non-competitive. In contrast, when we restrict the Dube et al. (2010) sample to perfectly politically aligned counties, we obtain a marginally significant estimate of employment elasticity of -0.145. These two estimates explain the seminal result in Dube et al. (2010) that the elasticity of employment with respect to minimum wage is zero.

Contents

List of Tables

List of Figures

ACKNOWLEDGMENTS

This work has only been possible thanks to the unwavering support from my advisors. Thank you to Michael Grubb, for your exceptional guidance and thorough feedback on my numerous research ideas and projects from day one. Thank you to Julie Mortimer, for your insightful advice and being a role model and a living example of a woman who can have it all. Thank you to Charlie Murry, for the many helpful suggestions that improved my work immensely.

In addition to my committee, I have also benefited greatly from the excellent instruction and support of Mehmet Ekmekçi, Uzi Segal, Tayfun Sönmez, Utku Ünver, and Bumin Yenmez in my early graduate years. An extra thank you to Uzi Segal for providing me with various entertaining puzzles and non-economics literature throughout the years. I would also like to thank Hanno Foerster and Joanna Venator for their feedback at the finish line of my Ph.D. For my development as a teacher, I am forever indebted to Çan Erbil, Tracy Reagan, and Kalpana Seshadri of the English department.

I have also been fortunate to work with an incredible group of co-authors—Ratib Ali, Andrew Copland, Jean-François Gauthier, and Sylvia Hristakeva. Thank you for all the brainstorming, discussions, insight, and creativity. I would not have been able to pull the weight of research on my own without your generous support. I am especially grateful for Sylvia's calming and inspiring presence in our joint research project, which propagated throughout all of my work.

While at Boston College and in Boston more generally, I have been incredibly lucky to make meaningful and (hopefully) lasting friendships with Sara Adelsberg, Gian Caspari, Jacob Caudell, Andrew Copland, Vito Cormun, Bogdan Genchev, Marco Errico, Ethan Farber, Valeria Ferraro, Dalton Fung, Alexey Khazanov, Luwen Mai, Gage Martin, Shane McMiken, Sam and Tina Patti, Simone Pesce, Luigi Pollio, Laura Seaberg, Mariia Tiurina, Tess Torregrosa, and Pietro Visaggio. From solving problem sets with me to making excellent Italian food, you all added joy to my life in different, yet integral, ways. An extra special thank you goes to Andrew Copland and Sam and Tina Patti for the many baked goods that we made and shared together, for going the extra mile to decorate the office for my birthday, for taking me to my first baseball game, drive-thru Dunkin' Donuts, Maine Lobster Festival, and many more.

Thank you to my family in Russia—mom Elena, aunt Natalya, and cousin Artyom—for being there for me my entire life and putting up with my occasional grumpiness and rare visits back home. The biggest thank you of all is reserved for my mother Elena. More than any other cause, this dissertation was written because of all the hard work you had done raising me, which was not easy as a single parent. I am eternally grateful to you for instilling your values of integrity, empathy, and diligence in me. All these qualities were indispensable to navigating a new life abroad. I would also like to thank my new family in Switzerland—Andreas, Katharina, and Flavia Huber for enthusiastically rooting for me at the end of my Ph.D. And finally, thank you to my husband, Marius Huber, for being a constant source of happiness, love, fun, and support.

Chapter 1

Competitor Avoidance and the Structure of the Advertising Agency Industry in the United States

with Sylvia Hristakeva

1.1 Introduction

Intermediaries play an important role in the production and provision of the goods and services offered to the final consumer. Business-to-business interactions are pervasive, and intermediaries often work with firms that directly compete with each other. For example, retail stores carry hundreds of competing products and airlines employ one of the three leading global distribution systems for reservations, inventory, and fare management (Altexsoft, 2019). At the same time, firms are often reluctant to work with intermediaries that are employed by their direct competitors. For example, Asker and Ljungqvist (2010) show that firms are concerned about disclosure of confidential information to strategic rivals in the equity and debt issuing market, and thus, prefer not to share a bank with a competitor.[1] We refer to such behavior as competitor avoidance, which can be rationalized by real or perceived conflicts of interest.

Advertisers' desire to avoid hiring an agency that is used by a competitor is an oft-cited concern in the advertising agency industry. Silk (2012) points out that perspectives of agencies and advertisers on client conflicts are deeply divided. Whereas clients are concerned about intentional or inadvertent information sharing, "divided loyalties" and "who is getting the

[1] Aobdia (2015) and Chang et al. (2016) document similar patterns for firms' choices of auditors and M&A advisors. Dai (2016) documents competitor avoidance in the market for pharmaceutical advertising agencies, a market separate from the one we study here.

best resources", agencies maintain that "the problem is often more illusory than real" and that "[c]lients are very restrictive about their agencies accepting competitive business—narrowness that often goes to extreme lengths." (Jones (1998); cited on p.87 in Silk, 2012). In our interviews with two industry practitioners, we confirm that attitudes towards client conflicts are just as strong as they were two decades ago.[2]

In this paper, we analyze how competitor avoidance may shape the organizational structure of advertising agencies in the United States and impact the industry's concentration. One may expect that the presence of competitor avoidance leads to a highly fragmented market with many firms, each serving a limited number of clients from the same industry. This intuition, however, imposes that each ad agency operates independently and that agencies are not able to adapt their organizational structure to alleviate the effects of competitor avoidance. In reality, many advertising agencies operate under the umbrella ownership of a conglomerate, which we call a holding company (HC). For example, The Interpublic Group of Companies (IPG) owns five agency brands—Deutsch, FCB, Hill Holliday, McCann, and The Martin Agency. According to our industry contacts, agencies that belong to the same HC coordinate with each other when evaluating whether to bid for a new client. Consistent with bid coordination, we observe that typically only one agency per HC bids for a new client's account. For example, in 2011 when BMW's account came up for a review, one of IPG's five agencies, Deutsch, was already working with Volkswagen. Even though Deutsch could not compete for BMW, the other four agencies had no conflicts, and IPG chose The Martin Agency to compete in the review. The HC structure coupled with coordinated bidding mitigates the 'constraints' imposed by competitor avoidance. As a result, it allows HCs to capture a larger share of clients at the HC level, even though client conflicts have just as much bite at the agency level. Thus, holding all else equal, a HC is likely

[2] One of our industry contacts works for a major agency in a top-4 holding company, while the other works for a mid-size independent agency with prior experience in a HC agency. Both sources wished to remain anonymous.

to be better positioned to compete for accounts than a standalone agency.

Interestingly, the holding company structure appears to be distinctive to the advertising agency industry and is not used in other professional service industries that may experience competitor avoidance. Laws prohibiting outside ownership essentially preclude holding company organization for accounting and law firms von Nordenflycht (2011).[3] This makes the advertising agency industry a unique case study that can be relevant for other professional service markets.

The relationships between advertisers and advertising agencies result from a formal review process, where standalone agencies and HCs compete for an advertiser's account. During this process, HCs select one of the agencies under their control to compete in the review. Holding companies effectively form advertisers' choice sets through coordinating their agencies' bidding decisions. In particular, HCs base their choices of bidding agencies on potential client conflicts that may arise at each of their agencies should one of them win a prospective account. Thus, a study of competitor avoidance and its connection to the HC structure requires a dataset that records not only winning, but also bidding agencies in account reviews.

The existing literature has not studied competitor avoidance and holding company organization extensively due to the lack of comprehensive data sources. To our knowledge, there is a single data vendor, Winmo, that only collects information on relationships between advertisers and their agencies, i.e. only the data on winning agencies.[4] To this end, we hand-collect a novel dataset on relationships between a set of major US advertisers and creative agencies of record (AORs), as well as the identities of

[3]As described on p.2 of Solomon et al. (2020): "The American Bar Association's (ABA) Model Rule of Professional Conduct 5.4, Professional Independence of a Lawyer, includes several provisions: "A lawyer or law firm shall not share legal fees with a nonlawyer;" a"lawyer shall not form a partnership with a nonlawyer;" and a lawyer shall not practice law for profit if 'a nonlawyer owns any interest therein.'" Similar restrictions are in place for accounting firms. Licensed CPA firms must be at least majority-owned, and in some states, wholly-owned by licensed individual CPAs (Bosher, 2021).

[4]Winmo was formerly known as the *Advertising Red Books*.

bidding agencies that were competing for these accounts for the period between 2005 and 2018.[5] These detailed data allow us to adequately model HCs' choices of bidding agencies and identify competitor avoidance from the observed bidding patterns.

In this paper, we develop a theoretical model of HCs' bidding choices and take it to the data. We estimate that the competitor avoidance effect is extremely strong at the agency level—holding all else equal, serving an additional competitor of an account in review decreases an agency's odds of bidding for this account by 91.6%. At the same time, client conflicts have no effect on a HC's likelihood to compete for an account. These results suggest that HC structure may be used to provide a valid separation for competing clients' accounts. We then use our theoretical model and a series of simulations to formally illustrate how competitor avoidance rewards the creation of a holding company structure with coordinated bidding. We also explore two alternative explanations that could rationalize the existence of HC structure—separation of agencies based on industry specialization and a more general avoidance of large advertisers due to competition for internal agency resources, such as talent. We find no evidence for the alternative hypotheses in our data.

Our theoretical model combines four main ingredients: (i) existence of competitor avoidance at the agency level; (ii) no competitor avoidance at the HC level; (iii) bid coordination within HCs; and (iv) economies of scale at the agency level. In this setting, economies of scale, which we also call the scale effect, refer to the fact that larger agencies are able to serve their clients at a lower cost due to the possibility of maintaining an in-house production of ads. Silk and Berndt (1993) document that economies of scale are highly significant in the operations of advertising agencies. In addition to that, scale economies act as a countervailing force to competitor avoidance. Mechanically, the odds of sharing an agency with a competitor increase with

[5] Even though advertisers may contract with multiple agencies for smaller projects, agency of record is the one that is responsible for developing and overseeing an entire advertising campaign.

agency size, and consequently, with its economies of scale. Thus, the presence of both competitor avoidance and economies of scale creates a trade-off for the incentives to maintain a HC structure. Our model predicts that HCs arise only if the competitor avoidance effect is strong relative to scale.

We model agency reviews as second-score auctions, where HCs submit a two-dimensional bid that specifies the proposed level of 'quality' and compensation. Here, 'quality' describes the match surplus between an account and its potential agency, which depends on the agency's economies of scale (measured by its logged billings) and the number of potential client conflicts. A crucial assumption in our model is that HCs make their bidding choices myopically, treating each review independently. This assumption shuts down any dynamic considerations which may be important for explaining HCs' bidding choices, e.g. protecting their highest-surplus agencies from acquiring a conflict if they anticipate that a larger account may come up for review in the future.

In the empirical section, we estimate a discrete choice model based on the relationship that follows directly from the model. Our empirical findings show that both scale and competitor avoidance play a role in HCs' decisions on whether to bid for a new client. First, we find that increasing agency billings by \$100M (from unconflicted accounts) from the mean, increases the odds of the agency being chosen by its HC to compete for an account by 10.3%. Second, we find that serving an additional competitor of an account in review decreases the odds of an agency to bid for this account by about 91.6%. The magnitudes of estimates suggest that competitor avoidance is indeed relatively more important than economies of scale, confirming our theoretical prediction that HCs arise only if the competitor avoidance effect is strong.

We conduct two counterfactual exercises based on the estimates. In the first counterfactual, we investigate how the market shares of major HCs and standalone agencies would change if competitor avoidance was no longer a concern in this industry. This counterfactual is motivated by continuing proposals from the agency side of the market to resolve client conflicts by assigning clients to different offices or teams within the same agency and maintaining strict firewalls between them. If successful, this approach would effectively remove competitor avoidance at the agency level. We find that competitor avoidance benefits standalone agencies and most of the holding companies. The only two HCs that would benefit from its removal are Omnicom and IPG, who own the two largest agencies in the market, BBDO and The Martin Agency. We predict that removing avoidance would shift clients away from smaller agencies and towards BBDO and The Martin Agency. As a result, the market HHI would increase by 35%, from 2,400 to 3,250.

In the second counterfactual, we examine the consequences of banning bid coordination within HCs. According to our industry contacts, HCs coordinate bids not only to manage conflicts, but also to save costs of creating an original sample work in preparation for reviews. If bid coordination were banned, we predict that an average holding company would increase the number of bids that it enters by 150%, without any meaningful increase in the number of accounts that it wins. This implies that bid coordination at the HC level is indeed effective at saving the bidding costs. In addition, we find that banning bid coordination increases the average number of bidders in a review from four to nine. Even though we are not able to quantify the impact on agency compensation, our theoretical model predicts that an increase in the number of bidders would decrease agency compensation in cases when a single HC enters both its best and second-best agencies. However, these predictions should be interpreted with caution because dynamic incentives are likely to be important in this framework. Increased bidding costs may put some HC agencies out of business, offsetting the predicted increase in the number of bidders. We leave a more definitive answer to this question for future research.

Our main contribution lies in formally connecting the existence of competitor avoidance and HC organizational structure in the advertising agency industry. Many other professional services are subject to competitor avoidance, therefore our results may be re levant for other industries. In particular, The American Bar Association is currently discussing the possibility of loosening restrictions on non-lawyer ownership of law firms, which Arizona and Utah have already implemented to an extent (Skolnik, 2022). Allowing non-lawyers to own law firms creates a potential for a holding company structure to arise in the market for legal services. Our paper sheds light on what the effects of such a change in legislation might be. Moreover, the advertising agency industry is already highly concentrated, so antitrust authorities may need to consider structural remedies for any potential mergers. Our paper suggests that requiring HCs to give up some of their agencies or large accounts may be a viable option.

The rest of the paper is structured as follows: Sections 2.2 and 1.3 describe the related literature and industry background. Sections 1.4 and 1.5 introduce the model and a series of simple simulations that illustrate the interplay between competitor avoidance and scale effects, and how a strong competitor avoidance effect incentivizes agencies to organize as holding companies. Section 1.6 describes our data sources and variables. Section 1.7 presents the estimation strategy and results. Section 1.8 reports the results of the counterfactual simulations. Section 1.9 concludes.

1.2 Related Literature

This paper contributes to three strands of literature: the theoretical and empirical literature on competitor avoidance, the literature on the purpose of holding company structure in the advertising agency market, and more generally, to the literature on industrial organization of advertising markets.

First, our paper relates to the theoretical and empirical literature on whether firms should avoid sharing intermediaries or use them as a credible mechanism of facilitating collusion. In an early work, Bernheim and Whinston (1985) show that, under complete information, there exists an equilibrium where two competitors hire a common marketing agency and choose its compensation scheme in a way that achieves price collusion through the agency's choice of marketing intensity. In contrast, Villas-Boas (1994) points out that incomplete information is paramount in interactions with intermediaries, and that information transfers through a common advertising agency can be detrimental. Villas-Boas shows that the incentive to share an agency depends on how access to rival's information affects the competitive environment. Relying on this result, Villas-Boas rationalizes the observed differences in attitudes towards competitive conflicts in the US and Japan. Japanese firms are more inclined to defend their market share against competitors by launching a new product, as opposed to using price and non-price sales promotions, like their American counterparts. Sharing information about a launch of a new product might preempt the competitor, whereas sharing information about sales promotions might prompt the competitor to retaliate, creating differential incentives to share agencies.

Gal-Or (1991) is another example of how incomplete information may affect the incentive to share a common agency. Gal-Or shows that if in Bernheim and Whinston (1985)'s setting agencies have private information about their costs and the costs are positively correlated, the benefits from price or quantity coordination may be offset by the agency's extraction of information rents through misreporting of its costs to both competitors. Instead, hiring individual agencies helps advertisers to limit the agencies' information rents by getting a signal about their true costs from observing the competitor's revenue.

More recently, Decarolis et al. show that advertisers can facilitate collusive bidding in online ad auctions by delegating their bids to a common digital agency. Instead of analyzing costs and benefits of hiring a common digital agency, the paper focuses on comparing performance of the two auction mechanisms used in practice (the generalized second-price auction and the Vickrey-Clarke-Groves mechanism) under collusive bidding. The authors illustrate a theoretical possibility of bid collusion in a relatively simple setting with complete information and a single digital agency serving all advertisers. In practice, multiple digital agencies participate in online ad auctions, so Decarolis et al. potentially overstate the benefits of such cartels. Similarly to the intuition from Villas-Boas (1994) and Gal-Or (1991), allowing for incomplete information in this setting may diminish the benefits from collusion even further.

The main conclusion from the described theoretical studies is that firms' reluctance or willingness to share an intermediary is an empirical question. The answer to this question depends on firms' perceptions of how information flows may affect their market outcomes. Indeed, advertisers often cite concerns over possibility of leakage of confidential information, such as advertising costs, as a reason for not sharing an advertising agency with their competitors (Silk, 2012). Using the variation in dissolution of client-agency relationships following mergers of advertising agencies, Rogan (2014) studies whether longer relationships tend to be more stable in a reduced-form setting. Intuition suggests that longer relationships reflect a better match and should thus be more stable. However, the longer the relationship lasts, the more sensitive information an agency learns about the client. In this case, even a threat of competitive conflict may be enough to dissolve the relationship. Rogan finds that longer relationships persist if a merger brings only a few competitors under the same roof and that the likelihood of dissolution increases with the number of competitors—the more so, the longer the relationship.

Fear of information leakage creates reluctance to share a common intermediary in other professional service markets as well. Aobdia (2015) finds that concerns about information spillovers affect firms' choice of auditors and diminish auditors' benefits from industry specialization. These concerns are likely justified. Using a quasi-natural experiment of the 2002 collapse of Arthur Andersen, the author shows that discretionary decisions (e.g. investment, R&D, advertising, and tax policies) of former Arthur Anderson's clients diverged immediately after the shock. In a similar vein, Asker and Ljungqvist (2010) find that product-market competitors avoid sharing the same equity-issuing investment banks out of fear of commercially sensitive information leaking to rivals. More importantly, the authors show that competitor avoidance reduces the pool of unconflicted banks that firms can hire from and allows them to charge higher fees.[6]

Our paper complements the existing literature by documenting competitor avoidance in the market for creative advertising agencies of record (AORs).[7] To the best of our knowledge, we are the first to formally show that competitor avoidance can rationalize the existing holding-company organizational structure. We find that the use of HC structure is associated with an increased market concentration.

Second, this paper contributes to literature studying the purpose of the holding companies in the context of advertising agencies. Silk and Berndt (2004) investigate whether holding companies' cost functions are subject to economies of scale and scope. If so, HCs could be organized for the purposes of leveraging the associated cost savings. The authors find that the long-run cost function is subject to only slight economies of scale and small economies of scope of one to two percent.

Another study by von Nordenflycht (2011) suggests that the value of holding companies stems from financial expertise and diversification of client risk. Financial expertise allows holding companies to implement better financial and management practices and increase profitability of agencies under their umbrellas. In addition, a holding company, as a publicly-traded entity, is less likely to experience major changes in stock

[6] Even though Dai (2016) remains agnostic about the reasons for competitor avoidance, it corroborates that competitor avoidance has 'real effects'. Focusing on the pharmaceutical industry, where manufacturers hire multiple advertising agencies, Dai finds that 14% of all matches that could have formed in a non-conflict environment cannot form if conflicts of interest are present.

[7] AORs are responsible for planning and creating advertising campaigns, as well as coordinating the work of specialty agencies that may be hired for small, specific projects. In the vast majority of cases, advertisers hire a single AOR and establish close and long-lasting relationships with them.

price as a large client departs. Here, von Nordenflycht (2011) directly credits competitor avoidance for producing said risk diversification: "By mitigating client conflict the holding companies offer some diversification across large clients." (p.148). However, the author also points out that "the conflict-mitigation argument [for HC structure] leaves questions unanswered... we still need an explanation for how conflict mitigation confers an advantage on holding companies..." (p.147). We believe that this paper provides the missing explanation. We show that holding companies are able to mitigate competitor avoidance by selecting which of their agencies (if any) should bid for new accounts. Contrary to a brief remark in King et al. (2003), our data suggest that agencies from the same holding company rarely compete for clients. Many instances of 'co-bidding' by sibling agencies reflect joint pitches, where HCs draft the staff from both agencies to deliver a better pitch. It is also not uncommon for HCs to create single-client dedicated shops by transferring existing staff to a new agency.[8,9]

Finally, our paper contributes to a growing research on industrial organization of advertising markets and its impact on downstream market outcomes. Combining data on national TV ad placements and their average prices, Hristakeva and Mortimer (2021) reveal that 'legacy' advertisers that have longer contractual relationships with TV networks tend to pay lower prices than newer firms for equivalent advertising inventory. This may benefit incumbents and potentially soften price competition from newcomers. Decarolis and Rovigatti (2021) document that advertisers outsource bidding in online sponsored searched auctions to a few concentrated intermediaries, causing a substantial decline in the search platform's revenues.

[8]For example, Team One (Publicis) was founded as a dedicated hub for Lexus and Garage Team Mazda — for Mazda. We Are Unlimited (Omnicom) was founded for McDonald's in 2016 and folded into DDB in 2019 shortly after losing the account (Adweek, 2019a; 2019b; 2019c).

[9]Bid coordination by HCs has been confirmed to us in a personal interview by an industry insider.

1.3 Industry Background

1.3.1 Structure of the Industry and Nature of Conflicts

Creative agencies provide several types of services to their clients. They are responsible for planning, creating, and producing advertising campaigns, working with media agencies to place ad copies across distribution outlets, and advising clients on marketing strategies, more generally. Most of large advertisers outsource to these tasks to a creative agency. [10] Smaller and local advertisers tend to rely on in-house creative agencies (Horsky, 2006).

The creative agency industry consists of two types of firms: independent, or standalone, agencies and holding companies (HCs). Holding companies control multiple agencies under their umbrella and select which agency in their portfolio get to compete for new accounts. In our sample, 44% of agencies are part of a major holding company—Dentsu, Havas, The Interpublic Group (IPG), MDC Partners (MDC), Omnicom, Publicis, or WPP. At the same time, these 44% of agencies control 82% of the market. Even though this market has a large competitive fringe due to independents, there are a few HC agencies that dominate the market, e.g. BBDO (Omnicom), The Martin Agency (IPG), and Leo Burnett (Publicis) by controlling on average 7.92% (about $1.5B), 5.97% (about $1.1B), and 5.50% (about $1B) of advertising expenditures in our sample. Holding companies themselves vary by their market share (about 3%-23%) and the number of agencies that they own. See Table 1.2 for more details.

Anecdotally, holding company structure was invented by Marion Harper, the CEO of future IPG, as a way of "circumvent[ing] barriers to growth" imposed by clients' avoidance of competitive conflicts (Silk and King III, 2013). Harper's idea was that a holding company could control

[10] Albeit this statistic is dated, Horsky (2006) documents that 86% of large US advertisers (with over $1M in annual advertising expenditures) used creative agency services in 2003.

multiple independently-operated agencies in order to satisfy clients' demands for exclusive relationships.[11]

What are the reasons for avoiding sharing an agency with a competitor? Advertisers often cite concerns over possibility of leakage of confidential information to justify their reluctance. Silk (2012) recounts an incident that happened in 2008 when Levi Strauss conducted an agency review for its media account and requested from candidate agencies that they share "supporting documentation" in the form of vendor invoice data that would let them gauge the scope of services that the agencies offered. Although not damaging by itself, together with other industry sources of data, Levi Strauss could identify the clients' brands and learn their advertising costs, which are considered a trade secret. The American Association of Advertising Agencies said that this request violated the organization's best-practice guidelines for maintaining confidentiality of a client's proprietary information. However, agencies have access not only to clients' advertising costs, but they also learn sensitive information about product performance, consumer and market research that clients share with them in order to produce effective advertising campaigns (Rogan, 2014).

Many advertisers are concerned that they will get a second-rate treatment if they share an agency with their competitor (Wooley, 2018). Woolley explains that this concern stems from fear that "competitor will get the better creative and strategy team and therefore will get the better ideas and better campaign." As the director of advertising services of General Motors put it: "I don't think it's so much an issue of sharing information. It's a question of resources and who is getting the best resources" (Silk, 2012, p.87). These concerns are especially strong for creative agencies, where the success of an advertising campaign could be subjective and difficult to measure.

Regardless of whether concerns over competitive conflicts are justified,

[11] As mentioned above, our data suggest that independent operation may not hold in practice entirely. Only about 4% of bidding instances in account reviews feature co-bidding by agencies from the same HC.

the creative agency industry in the United States clearly exhibits patterns of competitor avoidance. In Appendix A, we show that the observed share of agency-account relationships in conflict (only 12.7%) is highly unlikely to arise in this market by chance (26.5%). In the next subsection, we explain how the review and agency bidding process works and allows HCs to alleviate competitor conflicts.

1.3.2 Review Process and Agency Bidding

The two associations that represent each side of the market for advertising agency services in the United States, the Association of National Advertisers (ANA) and the American Association of Advertising Agencies (4A's), coauthored two white papers on best practices in agency search and selection process (ANA/4A's, 2011, 2013). In this section, we describe the review and agency bidding process, drawing on the insights from these two sources.

For a smaller advertising account, agency search typically consists of three phases:

- *Request for information:* Advertisers send out a questionnaire to 10-15 candidate agencies asking to provide information on ownership (holding company or independent), locations of offices, headquarters, years in business, number of employees, areas of expertise, recent account wins and losses, services offered, agency leadership bios, awards and recognitions, and brief summary of relevant case studies (ANA/4A's, 2013). At this stage, the advertiser discloses other agencies that might be working on its account and requests that the agencies share their active client lists, mentioning the scope and duration of relationships, in order to evaluate if potential conflicts of interest disqualify an agency from pursuing the account.

- *Request for proposal:* Remaining 6-8 agencies receive further details from advertisers about their goals for prospective advertising campaigns. Advertisers may solicit concrete examples of agencies' past work that are relevant to their vision for the ad campaigns. The two sides may discuss expected advertising budget and compensation method in broad strokes. Exact compensation is finalized only after advertisers pick the winning agencies.

Large and established accounts, who are aware of the industry landscape, tend to skip the formal RFI/RFP process. Instead, they reach out directly to HCs in order to invite them to pitch for their account in the finalist stage (Adweek, 2016; Adweek, 2017; Provokemedia, 2016).

- *Finalist stage:* At this point, HCs would typically organize an internal meeting to decide which agency (if any) is best positioned to compete for the account.[12] Up until this stage the review process is relatively costless for agencies. However, in the final stage, advertisers post a 'spec work' assignment that could range anywhere from a hypothetical case study to a sample of a creative product (e.g., developing an ad copy). This requires diverting staff resources away from existing clients. As mentioned in King et al. (2003), "[t]he process of agency selection is typically one of an advertiser inviting several agencies to make 'speculative' presentations, which require competing agencies to invest substantial amounts for which they receive only minimal compensation (Rothenberg, 1994)." Indeed, the process is so costly that typically only two to four agencies submit a formal bid.

Spec work presentations help advertisers reveal information about unobserved match quality with candidate agencies and affect their choices of winning agencies. In practice, it is hard to identify and measure the characteristics that affect agency winning independently from bidding. Quality

[12] According to statements of agency insiders in personal interviews.

of spec work is subjective and unobservable to the econometrician; and the two parties finalize compensation schemes only after advertisers announce the winners. In the next section, we model agency search and review process as a second-score auction, where agencies have private information about the quality of their spec work (match quality) prior to submitting the formal bid and incurring participation costs. We show that in equilibrium the agency with the highest match surplus wins the review. For our simulations in Section 1.5, we assume that advertisers draw a second error shock that is added to the match surplus determined in the bidding stage. After that, advertisers choose agencies with largest total match surplus as winners.

1.4 Model

This section presents a stylized model of agency search and review process that illustrates how the scale and avoidance effects can be identified from bidding data. The interaction is modeled as a second-score auction. In a nutshell, a buyer (advertiser) solicits bids from suppliers (agencies), chooses the highest scoring bid according to a scoring rule, and determines the final payment to the winner in a way that makes the highest score equal to the second-highest score.

Suppose that account m exogenously initiates a review and solicits bids from holding companies (HCs). The account chooses agency j that maximizes its utility function u_{jm} as the winner. In the context of the scoring auction, u_{jm} is the scoring rule used by the account. For simplicity, we assume that it is linear.

$$u_{jm}(q_{jm}, p_{jm}) = q_{jm} - p_{jm}, \tag{1}$$

where q_{jm} is the quality of match between j and m and p_{jm} is the compensation bid submitted by agency j.[13] Agency j gets a different match quality

[13] As suggested in the previous section, q_{jm} could be interpreted as the quality of spec

draw for different accounts.[14] We assume that q_{jm} are private information to HCs and treat them as random i.i.d. draws from some continuous distribution $F(\mu_q, \sigma_q)$ with an infinite support. In what follows, we assume that HCs bid myopically, which allows us to treat each review as an independent market. For this reason, we simplify our equations by omitting the m subscript.

A set of potential holding company bidders $h \in \mathcal{H}$ simultaneously choose whether or not to compete (bid) for the account, which agency j to choose from its portfolio of agencies A_h, and what the compensation bid p_j should be. Even though HCs do not explicitly control quality of their bidding agencies, they effectively choose the quality of their bid when choosing the bidding agency based on q_j draws. So, with a slight abuse of notation, we refer to contracts offered by a holding company h as (q_{hj}, p_{hj}). If h chooses not to bid, we normalize its payoff to zero and refer to such a bid as $(q_h = -\infty, p_h = \infty)$.

Each HC tenders a bid (q_{hj}, p_{hj}) to maximize its expected static profit from winning the account

$$\max_{(q_{hj}, p_{hj})} \pi_h(q_{hj}, p_{hj}; \boldsymbol{q}_{-h}, \boldsymbol{p}_{-h}) = (P_{hj}(\boldsymbol{q}, \boldsymbol{p}) - c_{hj}) \Pr[\text{win} \,|(\boldsymbol{q}, \boldsymbol{p})] - \kappa_h, \quad (2)$$

where the probability of winning and the final payment P_{hj} depend on the score of agency j from HC h and the scores of its competitors, given the vector of all submitted bids $(\boldsymbol{q}, \boldsymbol{p}) = (q_{hj}, p_{hj}, \boldsymbol{q}_{-h}, \boldsymbol{p}_{-h})$. c_{hj} is the agency's marginal cost of serving the account and $\kappa_h > 0$ is the fixed cost of bidding that is assumed to be the same across all agencies within the same HC.

More specifically, the probability of j winning among other bidders B_{-hj} is given by

$$\Pr[\text{win} \,|(\boldsymbol{q}, \boldsymbol{p})] = \Pr(u_j \geq \max_{k \in B_{-hj}} u_k) = \Pr(q_{hj} - p_{hj} \geq \max_{k \in B_{-hj}} q_{h'k} - p_{h'k}). \quad (3)$$

work produced by j for m.

[14]We could allow for correlation for between q_j draws for accounts in the same industry to reflect specialization, although there is no evidence of industry specialization in the data.

Note that the subscript h in the bid and the cost function only reflects agency j's affiliation with HC h, but neither of them varies with h. The costs are assumed to be common knowledge because client rosters of agencies are publicly available.

We also assume the following functional form for the cost c_{hj} to reflect the scale and competitor avoidance effects prevalent in the industry:

$$c_{hj}(M_{hj}, n_{hj}) = \tilde{c}\left(M_{hj}\right) + \kappa_c n_{hj}, \tag{4}$$

where $\tilde{c}(M_{hj})$ is the marginal cost of serving an additional account (regardless of the market it operates in) as a function of the total billings of agency j, M_{hj}. We assume that $\tilde{c}\,'(M_{hj}) < 0$ in order to capture scale effect. Parameter $\kappa_c > 0$ captures the cost of conflict that the agency has to incur if it serves one competitor of the account it bids for. Note that we choose to introduce competitor avoidance on the cost side because agencies often incur additional costs if they wish to serve competing accounts, e.g. hire separate staff and maintain separate offices. We could just as well introduce competitor avoidance as a downgrade in the match quality $q_{hj} - \kappa_c n_{hj}$. In fact, our data cannot identify the effects of match quality and costs separately. So, we could present the bidding model in terms of agency surplus (quality less costs) and mark-up (compensation less costs). We decide to keep the more standard notation because it maps more readily to the agency review practices described in Section 1.3.

1.4.1 Equilibrium

We now turn to the equilibrium analysis. We first consider the equilibrium choices of HCs, conditional on participating in the auction. Thus, the cost of bidding κ_h is sunk. Later in the section, we analyze the optimal participation choices as well. The analysis builds on Che (1993)'s treatment of scoring auctions.

Proposition 1. Conditional on participation, the optimal bidding contract

(q^*_{hj}, p^*_{hj}) of holding company h is given by

$$
\begin{aligned}
q^*_{hj} &= \{q_{hj} \mid j = \underset{k \in A_h}{\operatorname{argmax}} (q_{hk} - c_{hk})\}, \\
p^*_{hj} &= c_{hj}.
\end{aligned}
\tag{5}
$$

In equilibrium, the highest-surplus agency wins the account and receives the final payment

$$
P^*_{hj} = q^*_{hj} - \max_{k \in B_{-hj}} (q^*_{h'k} - c^*_{h'k}). \tag{6}
$$

Proof. Optimal agency choice. Let $v_{hj} = q_{hj} - c_{hj}$ stand for the surplus generated by agency j of holding company h. First, we will show that it is weakly dominant for the HC to choose the agency with the highest surplus to bid for the account. Let $v^*_h = \max_{j \in A_h} v_{hj}$ stand for the highest surplus generated by the agency in h, with the corresponding match quality and cost denoted by q^*_h and c^*_h. Suppose that the HC submits a bid (q_h, p_h) with surplus v_h, where $q_h \neq q^*_h$. We will show that the HC can increase its expected profit by submitting an alternative bid (p'_h, q^*_h), where $p'_h = p_h + q^*_h - q_h$. First, notice that bids (q_h, p_h) and (q^*_h, p'_h) achieve the same score. Therefore, $\Pr[\text{win} \mid (q^*_h, p'_h, \boldsymbol{q}_{-h}, \boldsymbol{p}_{-h})] = \Pr[\text{win} \mid (q_h, p_h, \boldsymbol{q}_{-h}, \boldsymbol{p}_{-h})]$.

According to the rules of the auction, the final payment to the winner matches its score to the score of the second-scoring bidder. Let P_h and P'_h denote the final payments received by the HC, conditional on winning with bids (q_h, p_h) and (q^*_h, p'_h) respectively. Suppose also that the second-scoring bid is given by $(\widehat{q}, \widehat{p})$. Therefore, the comparison of expected profits boils down to

$$
(P'_h - c^*_h) - (P_h - c_h) = (q^*_h - \widehat{q} + \widehat{p} - c^*_h) - (q_h - \widehat{q} + \widehat{p} - c_h) = v^*_h - v_h > 0, \tag{7}
$$

The statement in equation 7 is true by definition of v^*_h. Notice that we made no assumptions on the value of p_h, therefore it is optimal for the HC to pick the highest surplus agency, regardless of the compensation bid.

Optimal compensation bid. Now we can find the optimal compensation bid submitted by the highest surplus agency. Together with v_{hj}, we consider an additional change of variable $b_{hj} = q_{hj} - p_{hj}$, which can be interpreted as the price-adjusted quality 'bid'. This way, we can interpret our second-score auction as a standard second-price auction where the highest bidder (in terms of b_{hj}) wins and pays the second-highest bid. The previously defined surplus $v_{hj} = q_{hj} - c_{hj}$ can be interpreted as the private 'value' of winning the auction. As in a standard IPV setting, v_{hj} follows the distribution of q_{hj}, but its mean is shifted depending on c_{hj}. More formally, $v_{hj} \sim F_{hj}(\mu_q - c(M_{hj}) - \kappa_c(n_{hj}), \sigma_q)$.

Applying the change in variables, we can rewrite the profit in equation 2 as a standard second-price auction profit function that depends on own bid b_{hj} and bids of competitors $k \in B_{-hj}$, subsumed in $\boldsymbol{b}_{-h}$:

$$\pi_h(b_{hj}; \boldsymbol{b}_{-h}) = (v_{hj} - \max_{k \in B_{-hj}} b_{h'k}) \Pr[b_{hj} \geq \max_{k \in B_{-hj}} b_{h'k}]. \tag{8}$$

Note that relative to equation 2, we omit κ_h because it is sunk, conditional on entry.

From the original result in Vickrey (1961), we know that is a weakly dominant strategy for each bidder to bid its value $b^*_{hj} = v^*_{hj}$ for all hj. After we revert the change of variables, we can derive the equilibrium (q^*_{hj}, p^*_{hj}) for each bidding agency j in the set of bidders B

$$q^*_{hj} = \{q_{hj} \,|\, j = \operatorname*{argmax}_{k \in A_h} (q_{hk} - c_{hk})\},$$
$$p^*_{hj} = c_{hj}.$$

Final payment to the winner. Assume that the highest and second-highest score agencies have surplus v^*_{hj} and $\max_{k \in B_{-hj}} v^*_{h'k}$. The final payment received by the winning agency is P^* such that the winning agency's score matches the score of the second-scoring agency

$$P^*_{hj} = q^*_{hj} - \max_{k \in B_{-hj}} (q^*_{h'k} - c^*_{h'k}). \tag{9}$$

$\square$

Proposition 2. The equilibrium participation strategies of holding companies $h \in \mathcal{H}$ are given by

$$b_h = \begin{cases} v_h^* & \text{if} \quad v_h^* \geq v_{h0}^*, \\ \text{no entry} & \text{if} \quad v_h^* < v_{h0}^*. \end{cases} \tag{10}$$

The threshold v_{h0}^* is described by

$$\left(v_{h0}^* - \int_{-\infty}^{v_{h0}^*} v \prod_{h' \in \mathcal{H} \setminus h} F_{h'}'^{*}(v) \mathrm{d}v \right) \prod_{h' \in \mathcal{H} \setminus h} F_{h'}^*(v_{h0}^*) = \kappa_h, \tag{11}$$

where F_h^* is the distribution of the maximum surplus of all agencies of the holding company h.

Proof. The optimal participation decision depends on the expected profit of the HC, given the optimal bid, and its bidding cost κ_h. The optimal choice of a bidding agency can be regarded as an internal auction. At the end of this auction, the marginal distributions of agency surpluses get aggregated to the joint distribution of surplus created by the HC. We refer to the HC surplus as v_h^* and its distribution as F_h^*. Once v_h^* and F_h^* are known, the scoring auction is equivalent to a standard IPV setting with v_h^* and F_h^* as the primitives.

Let N_h denote the number of agencies in holding company h. In order to determine the distribution of the surplus generated by h v_h^*, we need to derive the distribution of the N_hth order statistic on independent and non-identically distributed (inid) random variables with distributions $F_{h1}, \ldots, F_{hN_h}$, where $v_{hj} \sim F_{hj}(\mu_q - c(M_{hj}) - \kappa_c(n_{hj}), \sigma_q)$.

Theorem 4.1 in Bapat and Beg (1989, p.83) shows that the distribution of the N_hth order statistic Y_{N_h} is given by

$$F_h^*(y) = \Pr(Y_{N_h} \leq y) = \frac{1}{N_h!} \operatorname{Per} \underbrace{\begin{pmatrix} F_{h1}(y) \\ \vdots \\ F_{hN_h}(y) \end{pmatrix}}_{N_h \text{ times}}, \tag{12}$$

where $\operatorname{Per}(\cdot)$ denotes the permanent of the $N_h \times N_h$ matrix of marginal distributions of v_{hj}.[15]

Proposition 4 in Tan and Yilankaya (2006) analyzes equilibria in second-price auctions with participation costs and asymmetric bidders. It shows that an intuitive equilibrium from equation 10, where a stronger bidder has a lower participation threshold, always exists.[16]

Let v_{h0}^* stand for the surplus that makes the HC indifferent between participating in the auction or not. Therefore, the expected profit of the HC that wins the account with surplus v_{h0}^* should be equal to the fixed cost of participation. If the surplus of the HC is greater than the threshold value, then the HC would prefer to participate in the review. Otherwise, it would prefer to stay out. Equation 10 summarizes this simple logic.

We now characterize the threshold surplus v_{h0}^*. First, we plug in equation 6 in equation 2 and conditioning the expected second-score surplus on the winning surplus v_{h0}^*, we get

$$(v_{h0}^* - \mathbb{E}[\max_{k \in B_{-hj}} v_{h'k}^* \mid v_{h0}^* \geq \max_{k \in B_{-hj}} v_{h'k}^*]) \prod_{h' \in \mathcal{H} \setminus h} F_{h'}^*(v_{h0}^*) = \kappa_h, \tag{13}$$

where the last left-hand-side term is the equilibrium probability of HC h winning the review with surplus v_{h0}^*. Using the fact that F_h^* distributions describe inid variables with infinite support, we can express the expected

[15]The permanent of a matrix is a function similar to the determinant, but it does not take into account the sign of permutations in its calculation.

[16]Strong in the sense of first-order stochastic dominance of relevant valuation distributions F_h^*.

second-score surplus as

$$\left(v_{h0}^* - \int_{-\infty}^{v_{h0}^*} v \prod_{h' \in \mathcal{H} \setminus h} F_{h'}'^*(v) \mathrm{d}v\right) \prod_{h' \in \mathcal{H} \setminus h} F_{h'}^*(v_{h0}^*) = \kappa_h.$$

□

1.4.2 Estimation Equation

In this subsection, we describe the estimation equation and how the scale and competitor avoidance effects can be identified from HC bidding choices.

Suppose that HC h chooses whether to enter one of its agencies into a review or choose the outside option. For brevity of notation, we treat the outside option as a $\emptyset$-agency. The agency surpluses are denoted by $v_{h\emptyset}, v_{h1}, \ldots, v_{hN_h}$. Let F_{-hj}^* be the distribution of the maximum order statistic of all agency surpluses except for agency j. Thus, the probability of agency j being chosen by its holding company h to compete for the account is given by

$$\Pr(v_{hj} \geq \max_{i \neq j \in A_h \cup \emptyset} v_{hi} \,|v_{hj}) = F_{-hj}^*(v_{hj}). \tag{14}$$

Proposition 3. Under the assumptions on the cost function c_{hj} in equation 4, i.e. $\tilde{c}\,'(M_{hj}) < 0$ and $\kappa_c > 0$, the probability of j being selected to bid for an account F_{-hj}^* is increasing in M_{hj} and decreasing in n_{hj}.

Proof. Trivially, the surplus of the agency v_{hj} is increasing in M_{hj} and decreasing in n_{hj}. F_{-hj}^* is increasing in v_{hj} because it is a cdf. □

The comparative static in Proposition 3 motivates our empirical specification. In Section 1.7, we estimate equation 14, assuming that F_{-hj}^* is EV type 1.

1.4.3 Assumptions and Discussion

Throughout the model, we maintain several identifying assumptions. First, we assume that holding companies bid myopically and non-strategically. This assumption rules out situations in which HCs may choose not to enter their best agency for a given review if they anticipate a larger competitor's account to come up for review in a later period. Similarly, HCs will not tailor their choice of agencies based on anticipated bidding agencies of other HCs.

Identification of scale and avoidance effects depends on the assumption that total agency billings M_{hj} and the number of competitors served by the agency n_{hj} only affect the marginal costs of serving the account, but do not affect the match quality q_{hj}. For example, agency billings are likely indicative of the agency's quality. We alleviate this issue by including agency fixed effects in our specification.

Similarly, n_{hj} is likely correlated with potential cost benefits that could arise from industry specialization. This suggest that the magnitude of the competitor avoidance effect that we identify in our model should be interpreted as the net of costs and benefits.

1.5 Intuition and Motivating Simulations

This section illustrates the connection between competitor avoidance, economies of scale, and agencies' decisions to organize as holding companies. Due to a short sample period, we are not able to track the ad agency market back to the creation of first holding companies. For this reason, we use simple simulations to showcase how competitor avoidance and economies of scale may create an incentive to introduce the holding company structure. In what follows, we first show that competitor avoidance has an effect on the market only if the scale effect is present. Second, we show that HCs develop only if the avoidance effect is strong relative to the scale effect. Finally, we show that, when economies of scale are present, competitor avoidance may benefit some holding companies and hurt others, depending on the sizes of agencies they control.

For these simulations, we use the model of agency bidding choices developed in Section 1.4. Each bidder is either a standalone agency (SA) or a holding company (HC). The model predicts how a bidder's willingness to enter an agency to compete in a review depends on that agency's billings (scale effect) and the number of existing client conflicts (competitor avoidance effect). We use these predictions to simulate how the market for advertising agencies evolves under different magnitudes of the scale and avoidance effects. We apply the model to a dynamic setting, while still preserving the myopic bidding assumption.

Each agency starts with an initial set of relationships that determines its billings, market share, and the number of accounts served across different product categories. For simplicity, we assume that each industry has an equal number of identical advertiser accounts. We then draw a random sequence of accounts that come up for review. After each review is completed, we update the relationship assignments and recalculate agency billings, number of conflicts, and market shares, based on the new relationship matrix.

Consider a single account in review. Each bidder decides whether to bid for the account and incur the bidding cost or stay out. If a bidder is a holding company with multiple agencies, it chooses which of its agencies from the set A_h bids for the account. If a bidder is standalone, then $|A_h| = 1$. For ease of notation, we treat a standalone agency j as its own bidding entity, so $h = j$ in this case.

The surplus of agency j from HC h when bidding for account m is described by

$$v_{hjm} = \gamma - \alpha \text{ num. of } m\text{'s competitors at } j_{hjm} + \beta \log \text{billings from other categories}_{hjm} + \varepsilon_{hjm},$$

where ε_{hjm} is an EV type 1 i.i.d. pair-specific match-quality shock. If a bidder decides not to participate, we assume that it chooses the outside option, or the $\emptyset$-agency. In that case, the bidder saves the bidding cost κ_c. Since κ_c is not separately identified from γ, we normalize κ_c to zero. Thus, the surplus of h from not bidding on an account is simply $v_{h\emptyset m} = \varepsilon_{h\emptyset m}$.

Bidder h picks agency j if and only if

$$v_{hjm} \geq v_{hkm} \text{ for all } k \in A_h \cup \{\emptyset\}.$$

Once agencies submit their bids, the account draws another set ν_{hjm} of i.i.d. EV type 1 shocks for each bidder and selects the one that delivers the highest total utility $v_{hjm} + \nu_{hjm}$. The second logit shock captures the information that gets revealed during the review and bidding process.

Note that presence of both competitor avoidance and economies of scale creates a trade-off in the advertiser's choice of an agency. Mechanically, the odds of sharing an agency with a competitor weakly increase with agency billings, and thus, with its economies of scale. In the simulations, we illustrate how this trade-off affects agency incentives to organize as holding companies. Intuitively, if the scale effect is strong relative to avoidance, agencies would prefer to remain standalone since lost economies of scale outweigh the benefits from decreased competitor avoidance. Once the avoidance effect becomes relatively more important, agencies would be more willing to organize as HCs. That way, they are able to benefit from separating competitors across different agencies within the holding company, despite foregoing benefits from economies of scale.

In our simulations, we focus on a simple market with two initially identical bidders that, depending on the simulation, operate as standalone agencies or as holding companies with two agencies. For simplicity, all accounts are identical, and each industry has an equal number of accounts. While we hold the initial market shares of the two bidders equal, we repeat the simulation 50 times by randomizing the starting number of competitor conflicts at each agency. We simulate bidding choices over 500 consecutive reviews, assuming exogenous dissolution of existing relationships and arrival of these accounts for review. We choose to simulate over a relatively large horizon to approximate the long-run distribution of agency market shares. We simulate each block of 500 reviews over 50 paths of logit error draws to average out the logit shocks. Therefore, the limit market shares at the end of 500 reviews are simulated 2,500 times. Even though asymmetric market shares can arise depending on particular starting points of competitor conflicts and logit error draws, averaging out over 2,500 simulations smoothes them out. The figures to follow report these averages.

First, the figures below illustrate the relationship between economies of scale and competitor avoidance in a simple market with only standalone agencies. Here, we see the expected result that increased competitor avoidance leads to a market where clients are more spread out over the existing agencies, therefore decreasing market concentration. Next, we show our main result that a strong competitor avoidance effect gives rise to HC organization, observed in the advertising agency industry. Finally, we show that, when economies of scale are present, competitor avoidance may benefit some holding companies and hurt others, depending on the sizes of agencies these HCs control.

1.5.1 Competitor Avoidance and Economies of Scale

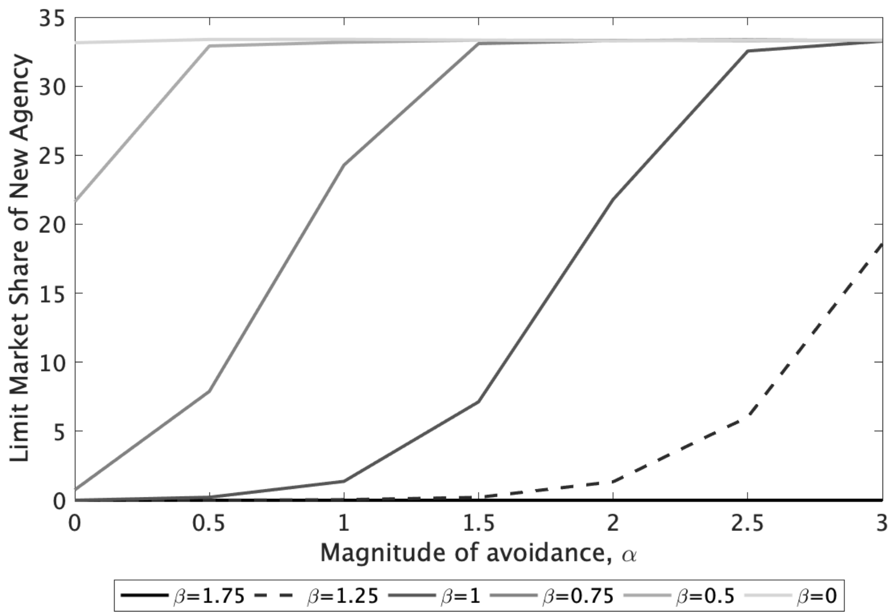
Competitor avoidance and economies of scale are important factors for advertisers' choice of agencies and, due to this, for agencies' decisions on whether to compete for a particular advertiser's account. At the same time, these two factors push bidding choices in opposite directions as the likelihood of competitor conflict increases with agency size. In order to demonstrate this tension more formally, we consider a simple market with two identical standalone agencies, each controlling half of the market in the long run. In this exercise, we do not allow firms to form holding companies, aiming to isolate the interplay between competitor avoidance and scale without the added complication of a holding company structure.

We disturb the long-run market shares by allowing a third standalone agency to enter the market at no cost. The entrant has the same surplus parameters (α, β, γ) as the incumbents, but has zero market share at the start of the simulations. We expect the long-run distribution of market shares to differ depending on relative importance of competitor avoidance α and economies of scale β for agencies' bidding choices. As described above, we approximate the long run by simulating agency bidding choices over 500 consecutive reviews, averaging over logit error draws and possible initial distribution of competitor conflicts between the incumbents.

Figure 1.1 depicts the predicted 'limit' market share for the entering agency at the end of 500 reviews as we increase the importance of competitor avoidance along the x-axis. Greater values of α imply that an agency is less likely to bid for an account if it already serves clients from that industry. The solid light gray line shows the predicted market shares for the entrant if there are no benefits to scale ($\beta = 0$). As the color of the lines gets darker, the value of β increases, meaning that an agency is more likely to bid for an account the larger the agency gets. Here, we clearly see that if there are no benefits to scale, the entrant captures a third of the market at any level of α. If $\beta = 0$, then competitor conflicts do not have any effect on the market shares, since both the randomness[17] and competitor avoidance push the market to maximum fragmentation. Thus, unless economies of scale are present competitor avoidance by itself does not affect the long-run distribution of agency market shares.

[17] If both α and β are zero, then the bidding and winning choices only depend on the logit error draws.

FIGURE 1.1 – SIMULATED LIMIT MARKET SHARE OF THE ENTRANT AGENCY



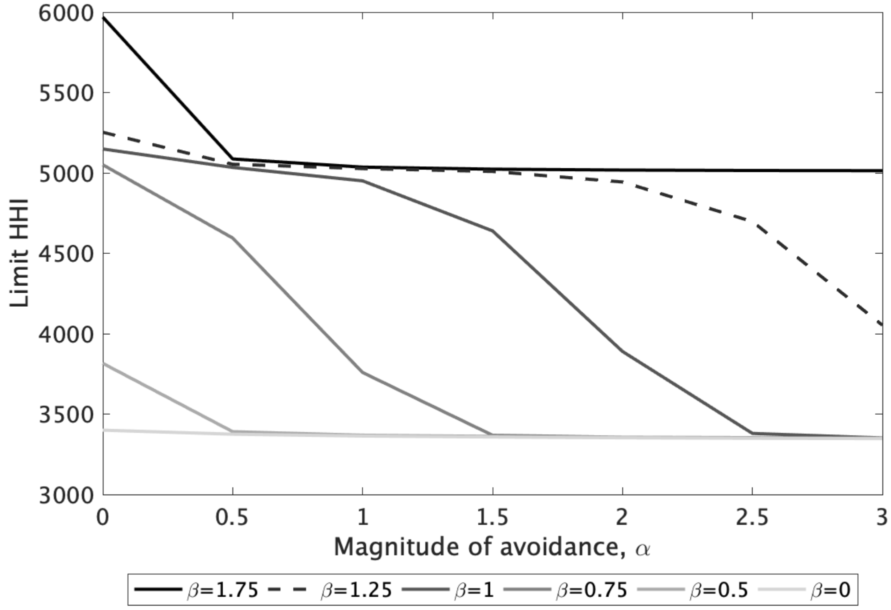
Notes: Limit market share of the entrant agency after 500 reviews averaged over 2,500 simulation paths. The shares are simulated at discrete 0.5 increments of α and interpolated linearly.

Once we allow for scale effect, we see that its interaction with competitor avoidance affects the long-run market shares and concentration. Consider the case of $\beta = 0.5$. Absent competitor avoidance at $\alpha = 0$, the entrant agency gains a market share of about 21%.[18] As competitor avoidance becomes more prominent, the entrant gets an increasingly larger market share. Competitor avoidance effectively 'dampens' the benefit from economies of scale and allows the entrant to bid more aggressively. Naturally, as benefits from economies of scale grow, the larger the dampening effect of competitor avoidance should be to make the entrant competitive. When $\beta = 1.75$, the scale effect is too important for the entrant to gain any market share.

Figure 1.2 illustrates the changes in market concentration associated with changes in the long-run market shares shown in Figure 1.1. When

[18]The long-run market share of the entrant is not zero due to the logit error term, i.e. there will be simulation draws in which the new agency gets a large enough logit draw to choose to bid for an account and large enough second draw to win its first account.

FIGURE 1.2 – SIMULATED LIMIT HHI



Notes: Limit HHI after 500 reviews averaged over 2,500 simulation paths. The HHI is simulated at discrete 0.5 increments of α and interpolated linearly, creating the drastic slope change at $\alpha = 0.5$ for $\beta = 1.75$.

the scale effect is zero, each agency gets an equal market share, resulting in the HHI of 3,333. When $\alpha = 0$, the limit HHI grows with β as the entrant agency captures increasingly smaller share of the market. In particular, when $\beta = 1.75$, the entrant never gains any market share; one of the existing agencies that initially receives a more favorable logit shock starts taking over, pushing the concentration above 5,000. Note that we simulate HHI at discrete 0.5 increments of α and use linear interpolation for the points in-between two increments. This explains the apparent change in the slope of the solid black line when we move from $\alpha = 0.5$ to $\alpha = 0$.

More generally, for all levels of $\beta > 0$, the plots show the expected result that competitor avoidance leads to a more fragmented market and lower HHI.[19] When $\beta = 0$, competitor avoidance has no effect on market concen-

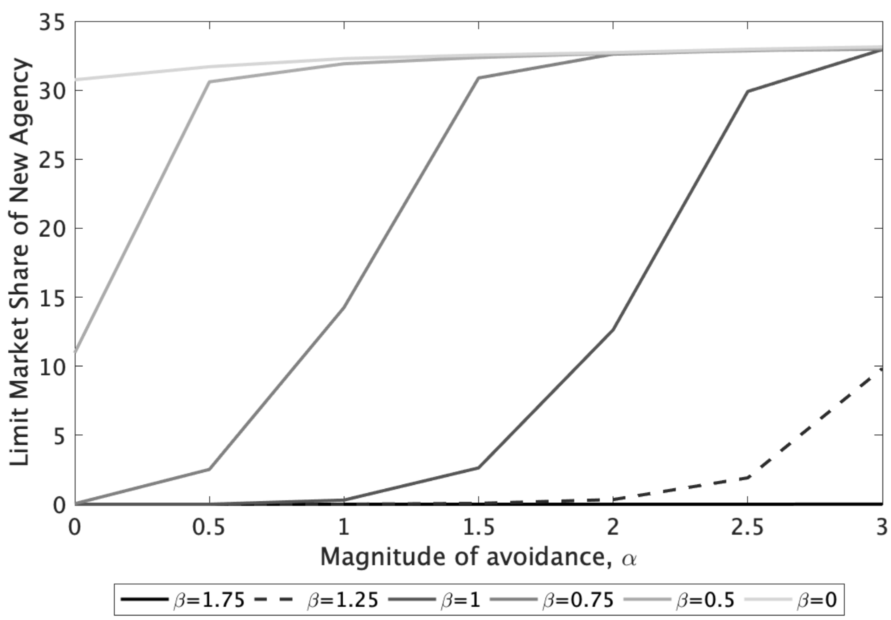
[19]Note that competitor avoidance need not lead to lower prices in the market. In fact, Asker and Ljungqvist (2010) find that existence of client conflicts leads to higher prices in the market for underwriting investment banks. Unfortunately, the advertising agency industry is highly protective of pricing data; thus, we are not able to analyze any price considerations.

tration.

1.5.2 *Incentive to Create Holding Companies*

In this subsection, we show that agencies have an incentive to introduce a holding company structure when the competitor avoidance effect is strong relative to scale. We return to the simple market setting from above, where we start off with two identical standalone agencies that share the market equally. Now suppose that one of them has the option of introducing a sibling agency with zero market share at no cost. If it does, then it turns into a two-agency HC and is able to coordinate its bidding decisions with the sibling. As before, we simulate the evolution of the market over 500 consecutive reviews and plot the limit market shares and HHI for different values of α and β.

FIGURE 1.3 – SIMULATED LIMIT MARKET SHARE OF THE NEW HC AGENCY



Notes: Limit market share of the sibling agency after 500 reviews averaged over 2,500 simulation paths. The shares are simulated at discrete 0.5 increments of α and interpolated linearly.

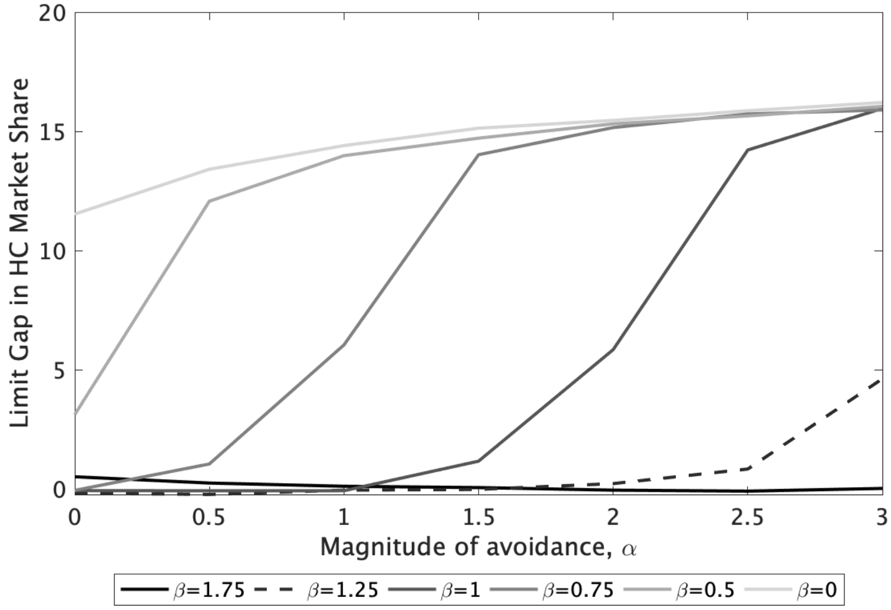
Figure 1.3 plots the market share captured by the new HC agency. Unsurprisingly, Figure 1.3 is similar to Figure 1.1 where the new agency is independent. When $\beta = 0$, the new HC agency captures one third of the market regardless of the value of α. Notice that when $\alpha = 0$, the light gray line dips below 33 percent. In this extreme case, where both α and β are zero, the long-run market shares are determined exclusively by the realizations of logit shocks. Also, due to bid coordination, the new HC agency bids less frequently than its independent counterpart. Therefore, 500 reviews may not be enough to reach the long run at that particular point.

More importantly, the interaction of scale and avoidance effects shapes the bidding firms' structure—the new HC agency gains market share and HC structure emerges if the importance of competitor avoidance is large enough relative to economies of scale. For example, when $\beta = 1.25$, the new agency starts developing only for $\alpha > 1.5$. If $\beta = 1.75$, the new HC agency does not develop at any value in the range of α.

Figure 1.4 plots the average difference between the market share of the bidding firm in case it operates as a HC or as a standalone agency. We focus on the difference between market shares rather than plotting the levels because, when β is large, it is possible for one incumbent agency to capture the entire market or no market share at all, regardless of its structure. The outcome depends on the initial amount of client conflicts at each agency and the logit draws. In order to make the effect of HC organization more transparent, we fix the initial amount of conflict and logit draws and simulate the long-run market shares twice—under HC and SA structure. We then average their difference across all such simulations.

Figure 1.4 shows that the bidding firm which has the option to operate as a HC is weakly better off under all combinations of α and β, and the benefits from the HC structure increase with competitor avoidance. For example, with $\beta = 1$, the market share of the HC increases by about 6 percentage points if $\alpha = 2$ and by about 14pp if $\alpha = 2.5$. Figure 1.5 plots the associated changes in market concentration. Mechanically, HHI increases in

FIGURE 1.4 – SIMULATED LIMIT GAP IN MARKET SHARE OF HC



Notes: Limit gap in market share when the standalone agency turns into a HC. The results are presented after 500 reviews and averaged over 2,500 simulation paths. The shares are simulated at discrete 0.5 increments of α and interpolated linearly. The light gray line ($\beta = 0$) does not flatten out after 500 reviews because the bidding and winning process is increasingly guided by random logit draws as α decreases.

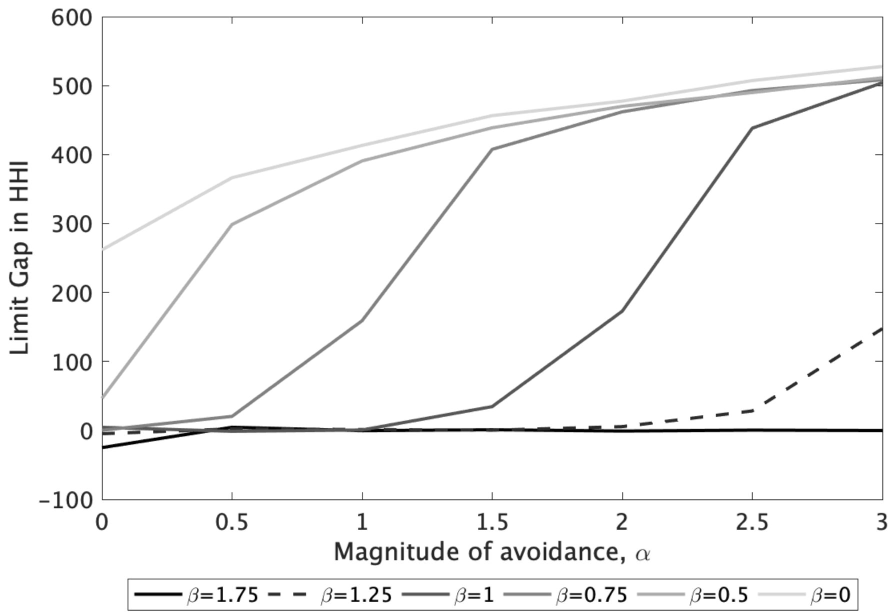
all scenarios when the HC structure emerges. More importantly, the comparative statics with respect to avoidance show that if one agency is able to organize as a HC, increased avoidance can lead to a higher market concentration (moving from the SW to the NE corner of Figure 1.5).

1.5.3 Holding Company Competition

The example above shows that in presence of competitor avoidance, there is an incentive for agencies to organize as holding companies. In that example, the HC benefits from its structure and gains market share at the expense of the standalone agency. We now consider a simulation where both bidders operate as HCs and investigate whether the initial allocation of clients across the agencies affects the HCs' long-run market shares.

As before, the two bidding firms start with equal market shares.

FIGURE 1.5 – SIMULATED LIMIT GAP IN HHI

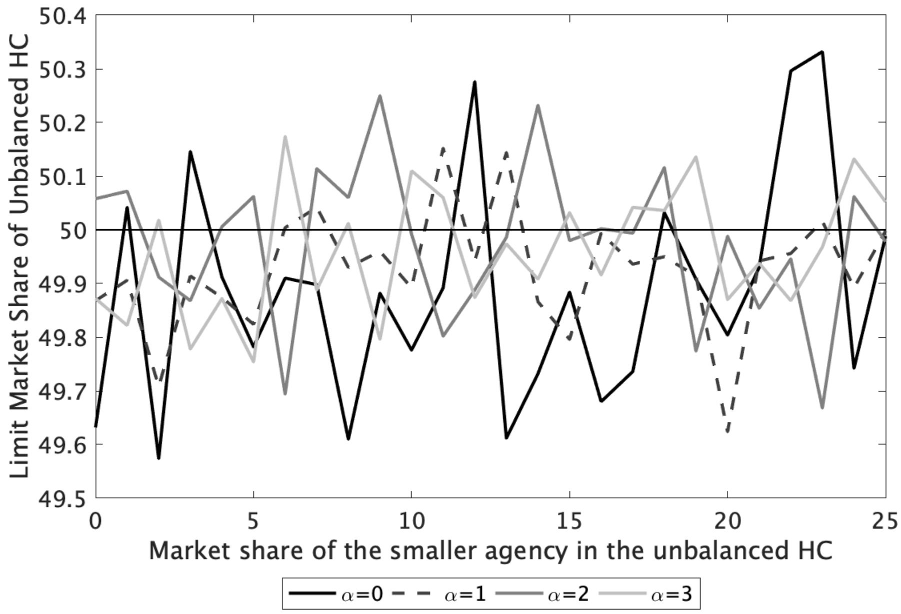


Notes: Limit gap in HHI when the standalone agency turns into a HC. The results are presented after 500 reviews and averaged over 2,500 simulation paths. The HHI is simulated at discrete 0.5 increments of α and interpolated linearly. The light gray line ($\beta = 0$) does not flatten out after 500 reviews because the bidding and winning process is increasingly guided by random logit draws as α decreases.

However, if they are initially symmetric and both allowed to operate as HCs, then on average their long-run market shares remain equal. We therefore allow the two HCs to start with different allocations of market shares across the agencies they control. Asymmetric agency sizes could be rationalized through agency-specific bidding constants γ_j (perhaps, due to different quality or bidding costs) or an early entry advantage, although our simulations abstract away from either.

Suppose that one of the HCs in the simulation starts with a balanced profile of market shares, (25%, 25%). The profile of the other HC is denoted by (x%, $50 - x$%). It varies along the x-axis from 0 to 25%, representing the starting market share of the smaller agency in the unbalanced HC. Based on the initial market shares, the four agencies are endowed with different logged billings and have a different number of competitor conflicts at the start of the simulation. We are interested in examining how the competition between the two differently balanced HCs evolves under various parameters of the scale and avoidance effects.

FIGURE 1.6 – SIMULATED LIMIT MARKET SHARE OF THE UNBALANCED HC, $\beta = 0$

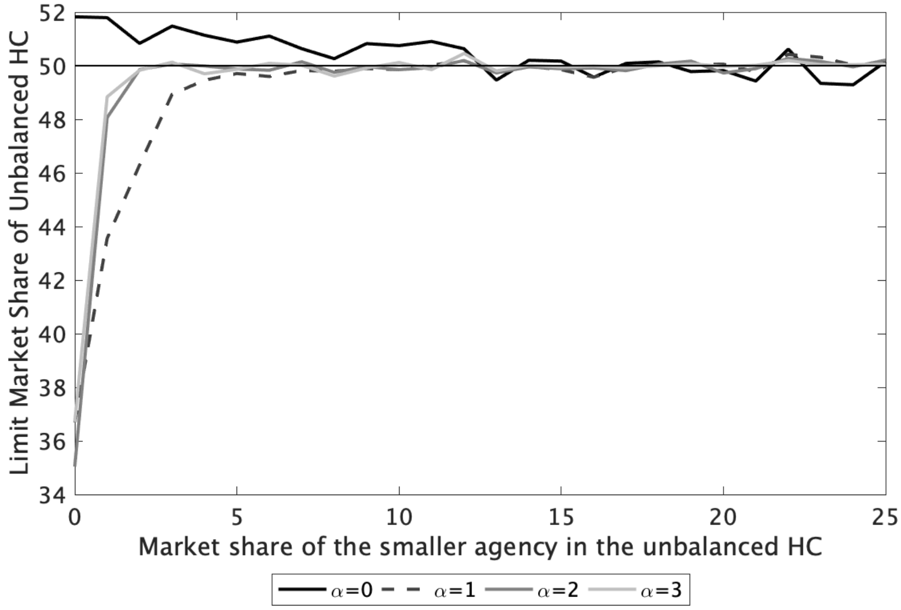


Notes: Limit market share of the unbalanced HC after 500 reviews averaged over 2,500 simulation paths. The shares are simulated as a function of the starting market share of the smaller agency x for $\beta = 0$ and $\alpha \in \{0, 1, 2, 3\}$.

Figure 1.6 shows the simulated market share of the unbalanced HC when it competes against the balanced HC in the market where the scale parameter is $\beta = 0$. We conduct simulations for integer values of α between 0 and 3. Once again we observe the expected result that competitor avoidance has no effect on the market if the scale effect is zero. As with symmetric standalone agencies, both the randomness and avoidance push the market to maximum fragmentation in the long run. So ultimately, the market shares across the four HC agencies equalize.

In Figure 1.7, we increase the value of scale to $\beta = 1$. If there is no avoidance, the unbalanced HC benefits from keeping a relatively large agency when $x < 12$, implying that the market share of the larger agency is between

FIGURE 1.7 – SIMULATED LIMIT MARKET SHARE OF THE UNBALANCED HC, $\beta = 1$

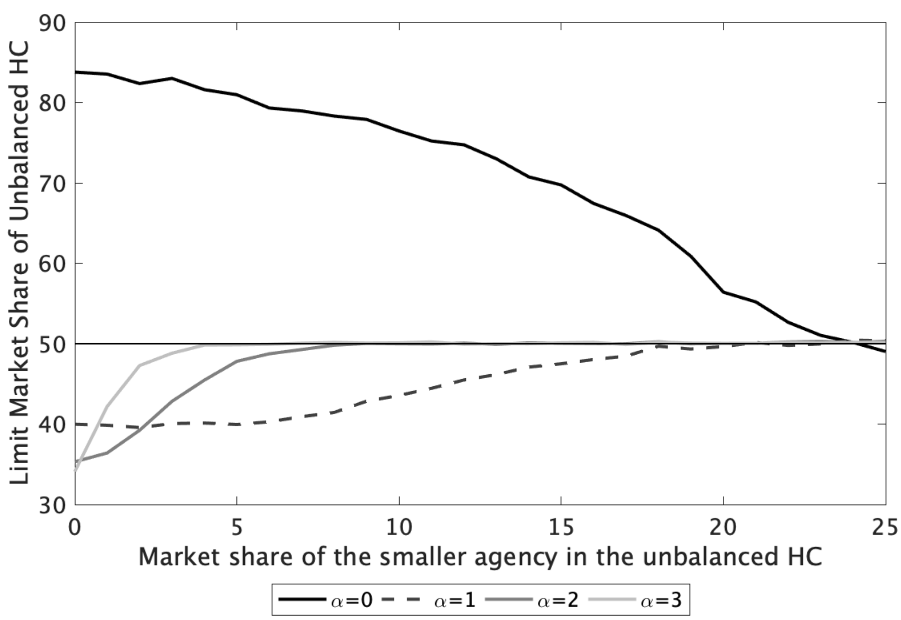


Notes: Limit market share of the unbalanced HC after 500 reviews averaged over 2,500 simulation paths. The shares are simulated as a function of the starting market share of the smaller agency x for $\beta = 1$ and $\alpha \in \{0, 1, 2, 3\}$.

50 and 38. For $x > 12$, the initial market shares of the four HC agencies are similar enough to produce comparable logged billings, and thus, economies of scale. Therefore, the benefit from having a larger agency vanishes. In contrast, under competitor avoidance ($\alpha > 0$), the unbalanced HC loses from having a large agency with numerous conflicts. Somewhat counterintuitively, the unbalanced HC is worst off if $\alpha = 1$. Due to myopic bidding, when α is relatively small, the unbalanced HC may choose the larger agency to bid on some accounts due to its economies of scale. If α is large, the benefits of scale economies are diminished—the larger agency never bids and rapidly loses its market share. On the other hand, the smaller agency benefits from fewer competitor conflicts and rapidly gains market share. The sooner the market shares of the four agencies equalize, the sooner the two HCs become symmetric, producing the observed 50-50 market share split.

Note that if we allowed our simulations to run for an arbitrarily large number of reviews, ultimately the limit market shares of the four agencies would equalize, regardless of their starting values and the magnitude of α. In order to sustain different market shares in the long run, we need to allow for different bidding constants γ_j. However, our simulations are still useful for illustrating one fact: depending on the importance of the scale effect, agencies which are already large (whatever the reason might be) would become even larger if advertisers stopped regarding competitor conflicts as an issue. This intuition is reinforced in Figure 1.8, where the unbalanced HC takes over most of the market if $\alpha = 0$ for a wide range of x. Having a larger agency only stops being an advantage for the unbalanced HC at $x = 24$. In fact, in an unreported simulation, we show that the large agency completely takes over at $\beta = 4$, effectively suggesting that the HC structure collapses if there is no competitor avoidance.

FIGURE 1.8 – SIMULATED LIMIT MARKET SHARE OF THE UNBALANCED HC, $\beta = 1.75$



Notes: Limit market share of the unbalanced HC after 500 reviews averaged over 2,500 simulation paths. The shares are simulated as a function of the starting market share of the smaller agency x for $\beta = 1.75$ and $\alpha \in \{0, 1, 2, 3\}$.

All in all, these simulations show that, when economies of scale are present, competitor avoidance benefits some holding companies and hurts others, depending on the initial sizes of agencies within each HC. To the extent that agency size depends on agency fixed effects and timing of entry, we are interested in examining which of the observed holding companies in the advertising agency market benefit and lose from competitor avoidance. We are also interested in determining how competitor avoidance affects the actual market concentration. We report our findings in Section 1.8.

1.6 Data and Descriptive Statistics

This section introduces our data sources, defines the main variables of interest, and describes how we construct the sample used in the empirical analysis.

1.6.1 Data Sources and Variables

Our sample contains information on relationships between major US advertisers and their creative agencies of record (AORs) for the period between 2005 and 2018. AORs are responsible for planning and creating advertising campaigns, as well as coordinating the work of specialty agencies that may be hired for small, specific projects. Advertisers tend to hire separate AORs for each brand, although this is not always the case. For this reason, we often use the term 'account' to refer to the appropriate unit of relationship.[20] An account has a single AOR at any given period of time.

We hand-collect data on advertiser and AOR matches from the leading trade publications. We primarily rely on articles from Ad Age (published since 1930), Adweek (1979), and Accounts in Review (2001). In what follows, we refer to this sample of matches as the relationship sample. Occasionally, articles that report relationship changes include identities of other

[20] For example, even though AT&T Inc. owns both AT&T wireless and DirecTV, it hires separate AORs for each brand.

agencies that competed for the account. We refer to the subset of relationships for which we observe identities of competing agencies as the bidding sample. We supplement the samples with the data on expenditures of US advertisers from AdSpender for the period between 2004 and 2017.

We use the information on agency bidding patterns in order to elicit advertisers' preferences for agency characteristics. If a holding company chooses an agency to bid for an account in review, it anticipates that its chances of winning the account are higher than those of any other agencies controlled by the HC. Therefore, whether a HC chooses an agency with a given set of characteristics to bid on some accounts, but not others, implicitly reveals how strongly advertisers value (or avoid) said characteristics. We use the relationship and advertising expenditure data to calculate them.

We define the main variables used in our empirical analyses below. The ANA and 4A's suggest that these variables are informative in judging the match quality between advertisers and their agencies, and thus, can reveal advertisers' preferences for agency size and competitor avoidance (ANA/4A's, 2011, 2013).

- *log billings in other categories:* An agency's total client billings, or the sum of advertising budgets of all of its clients, is a common measure of agency size used in the industry. It serves as a good proxy for agency economies of scale; however, it may also be capturing a confounding quality effect.[21] In order to minimize the potential confound, we measure an agency's economies of scale for account from product category c by excluding the billings of clients from category c from total agency billings. Additionally, we find evidence for diminishing scale effects by including the square of log billings in other categories. For this reason, we impose diminishing returns by choosing the log, rather than quadratic, specification.

[21] For example, endogenous sorting of more talented creatives into larger agencies may produce such a confounding effect.

- # *competitors:* The number of account's competitors currently served by a given agency or a holding company. Following the industry standard, we define conflicts at the category-parent level.[22]

- # *years category experience* measures an agency's category experience from recent (within the last three years) dissolved relationships. For each dissolved relationship in a given category, we calculate how many of the last three years the agency spent serving the account. We then sum up the number of years across all such relationships to get a cumulative category experience.

- *log account budget:* Logged advertising budget of the account.

- *large account* is a dummy variable that takes on the value of one if the account is within the top-25th percentile by advertising expenditure in that year.

- # *large accounts* counts the number of large accounts among an agency's clients.

Note that we lag all the variables of interest by one year in our estimation. In practice, an agency may participate in multiple reviews at the same time, and advertisers may not know their exact advertising budgets at the time of the review. Therefore, the best approximation of current agency size, account size, and number of potential conflicts is the previous year's values.

The variables introduced above map to the effects that can rationalize existence of the HC structure. In particular, log billings in other categories capture potential (dis)economies of scale of an agency. The number of competitors variable captures the potential competitor avoidance (if negative) or specialization (if positive) effect. The interaction between the large account dummy and the number of large accounts at the agency captures avoidance of large advertisers (regardless of their product market) due to capacity constraints, or potential competition for agency's internal resources.

[22] If an agency works with two brands of the same parent company and considers adding another competing parent, the number of potential competitors in this case is one, not two.

1.6.2 *Sample Coverage and Descriptive Statistics*

Our relationship sample covers nine product categories: autos and light trucks, beer, liquor, soft drinks and bottled waters, finance, insurance, telecom, retail stores, and casual restaurant chains. Within each product category, we collect information on relationships and advertising expenditures of relatively large advertisers that spend at least $50M per year on advertising activities. We then add their smaller competitors in order to accurately account for potential amount of conflict. Overall, we observe 89 unique parents that have 167 advertising accounts. Table 1.1 summarizes the number of advertisers at the parent-company and account level. Some accounts come in and out of the sample, sometimes due to temporarily switching to in-house agencies and sometimes due to lack of information in our data sources. The standard errors capture the variation in the number of parents and accounts observed over the duration of the panel. Table 1.1 suggests that potential for client conflict is high, so any suggestive evidence for competitor avoidance is not due to lack of data on competitors' agency relationships. Indeed, in Appendix A we reject the null hypothesis that the observed lack of competitor overlap reflects random matching between a large number of agencies and a small number of advertisers within a category.

In the relationship sample, we observe 123 unique agencies, 69 of which are independent and 54 are part of one of the major holding companies—Dentsu, Havas, The Interpublic Group (IPG), MDC Partners (MDC), Omnicom, Publicis, or WPP. Even though the market for creative advertising agencies has a large competitive fringe, there are a few agencies that dominate the market, e.g. BBDO (Omnicom), The Martin Agency (IPG), and Leo Burnett (Publicis) by controlling on average 7.92% (about $1.5B), 5.97% (about $1.1B), and 5.50% (about $1B) of advertising expenditures in our sample.

Holding companies vary by their market share and the number of agen-

TABLE 1.1 – NUMBER OF PARENTS AND ACCOUNTS IN EACH CATEGORY

Product category	Parents		Accounts	
	Mean	Std.Dev.	Mean	Std.Dev.
Autos & light trucks	17.00	0.00	31.21	1.05
Beer	4.93	0.99	18.86	3.63
Finance	13.21	1.05	14.07	1.38
Insurance	11.00	0.00	11.00	0.00
Liquor	2.00	0.00	12.21	2.61
Restaurants	16.50	0.52	24.57	1.16
Retail stores	10.50	1.61	12.57	1.65
Soft drinks & bottled waters	5.28	0.82	20.64	3.71
Telecom	3.78	0.58	4.78	0.58

Notes: The table reports the mean and standard deviation over the sample period of 2005-2018.

cies that they own, see Table 1.2. For simplicity, our empirical analyses focus on the largest 27 agencies controlled by the HCs. Several HCs control multiple smaller agencies that typically serve only one client. These are aggregated into the 'HC other creative' agency, capturing the fact that these HCs are always able to find a small unconflicted agency in their portfolio.

TABLE 1.2 – COMPOSITION OF HOLDING COMPANIES

Holding company	Market share	Number of agencies
Dentsu	4.59 [2.40]	2
Havas	2.82 [0.92]	2
IPG	20.51 [3.28]	5 and 1 other
MDC	4.94 [1.71]	2 and 1 other
Omnicom	22.95 [4.69]	6 and 1 other
Publicis	15.48 [1.54]	6
WPP	11.12 [1.24]	4 and 1 other
Independent	17.55 [3.45]	–

Notes: The table reports the mean and standard deviation of market shares over the duration of the panel. Market shares are reported in percents.

For the empirical analysis, we also keep four independent, or standalone, agencies (Barkley, Droga5, McKinney, and Wieden & Kennedy) that consistently participate in account reviews throughout the duration of the panel. All in all, we focus on top-35 agencies that on average account for 79% of advertising expenditures observed on our sample and for 49% of all advertising budgets recorded by AdSpender in the nine categories of interest. The discrepancy can be explained by some advertisers using in-house creative agencies and smaller advertisers not hiring creative AORs at all. Table 1.3 provides additional details on the sample coverage.

TABLE 1.3 – SAMPLE COVERAGE

Sample base	All agencies	All HC agencies	Bidding agencies
AdSpender	61.66% [3.86%]	50.42 % [3.04%]	49.01% [3.94%]
Relationship sample	100% (by definition)	82.08% [3.19%]	79.44% [2.03%]

Lastly, we tabulate the variables of interest in the relationship and bidding samples to check for potential selection of observations where we observe identities of bidders. As expected, trade publications are more likely to report on account reviews that involve either large accounts, or large agencies. As a result, the mean advertising budget of an account and agency billings in the bidding sample are higher than those in the relationship sample. There appears to be no selection in the number of competitors served by the agencies from the two samples. This leads us to believe that the bidding sample does not suffer from major selection concerns.

In the next section, we introduce our estimation strategy and discuss the results of the empirical analyses.

1.7 Estimation and Results

Section 1.5 explains how competitor avoidance effect may lead to rise of the holding-company firm structure if the scale effect is not too large. In this section, we evaluate the magnitudes of scale and avoidance effects revealed from HC bidding choices.

TABLE 1.4 – SUMMARY STATISTICS

	Bidding sample HC	Bidding sample All	Relationship sample HC	Relationship sample All
Account ad budget	148.33	143.64	143.48	128.30
	[186.22]	[180.59]	[184.24]	[169.05]
Agency billings	657.59	640.41	594.94	497.42
	[506.67]	[493.45]	[507.96]	[481.31]
Agency billings in other categories	462.28	450.78	409.17	332.12
	[453.31]	[440.85]	[449.48]	[419.58]
Accounts served by agency	4.71	4.71	4.32	3.82
	[2.51]	[2.49]	[2.59]	[2.57]
Parents served by agency	3.71	3.71	3.42	3.05
	[1.75]	[1.74]	[1.48]	[1.85]
Competitors at agency	0.15	0.15	0.14	0.12
	[0.39]	[0.38]	[0.37]	[0.35]
Categories served by agency	3.30	3.32	3.05	2.75
	[1.40]	[1.43]	[1.48]	[1.54]
Years of category experience within the last 3 years	0.92	0.89	0.82	0.71
	[1.87]	[1.83]	[1.78]	[1.67]
Large account	0.28	0.27	0.27	0.25
	[0.45]	[0.45]	[0.44]	[0.43]
Large accounts served by agency	1.22	1.18	1.09	0.93
	[1.15]	[1.13]	[1.14]	[1.06]
Agencies	31	35	54	123
Parents	78	80	81	89
Accounts	138	144	144	167
Observations	1,341	1,476	1,524	2,081

Notes: The table reports the mean and standard deviation over the panel duration for different samples we use. Advertising budgets of accounts and agency billings are reported in millions of 2016 dollars.

Competitor avoidance is not the only explanation for why HCs may choose to maintain separate agencies. Therefore, we test two alternative explanations: separation by industry specialization at the agency level (the opposite of avoidance effect) and due to capacity constraints. If an agency works with two large clients and its resources are limited, it may be forced to prioritize one client's account over the other. Similarly to competitor avoidance, this can create a tendency for avoidance of large advertisers regardless of their product market affiliation.

We base our estimation specification on equation 14 from Section 1.4. The equation relates the agency-account match surplus to the probability of the agency being chosen by its HC to bid for that account. We focus on bidding rather than winning patterns for two reasons. First, the sample of winners is likely selected on unobservables, e.g. quality of staff assigned to the account, interpersonal connections between agency and account executives, etc. Second, HCs are endogenously forming accounts' choice sets through coordinating their agencies' bidding decisions. Failing to account for endogenous choice sets may bias the estimates in our empirical model (Draganska et al., 2009). We also omit incumbent HC's bidding decisions for account reviews with a known incumbent agency. If an account wants to replace the existing agency with a new one, it is likely that the incumbent HC's choices will be influenced by the past relationship in an unobserved way.

The surplus of agency j that belongs to HC h and competes for account m in year t is given by

$$v_{hjmt} = X_{hjmt}\theta + \varepsilon_{hjmt},$$

where X_{hjmt} contains the observable characteristics, such as agency billings and number of account's competitors already served by an agency, θ is the vector of parameters, ε_{hjmt} is an EV type 1 i.i.d. pair-specific match-quality shock. The payoff from not bidding on an account is normalized to zero.

Integrating the payoff shocks out, we obtain the standard logit probability expression

$$P(h \text{ chooses } j \text{ to bid for } m \text{ at time } t) = \frac{\exp(X_{hjmt}\theta)}{1 + \sum_{k \in A_h} \exp(X_{khit}\theta)}.$$

We estimate the parameters of the surplus function by maximizing the log-likelihood function

$$\mathcal{L}(\theta) = \sum_{mt \in R} \sum_{h \in \mathcal{H}} \sum_{j \in A_h \cup \{\varnothing\}} y_{hjmt} \ln \left(\frac{\exp(X_{hjmt}\theta)}{1 + \sum_{k \in A_h} \exp(X_{khit}\theta)} \right) + (1 - y_{hjmt}) \ln \left(\frac{1}{1 + \sum_{k \in A_h} \exp(X_{khit}\theta)} \right),$$

$mt \in R$ stands for a review of account m at time t and y_{hjmt} is equal to one if agency j of HC h bids on account mt and is zero otherwise.

Table 1.5 presents the estimates of the logit model. The first three specifications test for competitor avoidance, industry specialization, and capacity constraints at the agency level. Specifications (1) and (2) differ in our measure of account size— log budget or a dummy if the budget of the account is in the top-25th percentile by advertising expenditures in that year. In specification (3), we test for avoidance of large accounts by including the interaction with the large account dummy. The last specification tests for competitor avoidance at the HC level.

First, we find that HCs are more likely to select their larger agencies to compete for accounts in review in all specifications. As mentioned before, we measure agency size for account from product category c by including the log of agency billings from all other categories except for c. We exclude the category in question to minimize the potential for confounding effects of agency quality in the category. We impose the log specification following previous research by Silk and Berndt (1994) that found that agencies with that bill \$42M-\$52M in 2016 dollars annually fully realize size-related economies. Their sample contains agencies with mean and maximum billings of \$21M and of \$414M in 2016 dollars, respectively. Admittedly, these estimates may no longer reflect modern-day agency cost functions since the advertising and marketing industries had changed rapidly with the wide-spread use of the Internet. However, it is likely that diminishing returns to scale still hold true. Based on specification (1), we find that

TABLE 1.5 – PROBABILITY OF AN AGENCY TO BE SELECTED BY ITS HC TO BID FOR AN ACCOUNT

	(1)	(2)	(3)	(HC level)
log billings in other categories	0.170***	0.170***	0.120**	0.529***
	(0.043)	(0.043)	(0.051)	(0.137)
# years category experience	-0.033	-0.032	-0.024	
	(0.110)	(0.110)	(0.109)	
# competitors	-2.481**	-2.482**	-2.521**	0.053
	(1.224)	(1.226)	(1.232)	(0.116)
log account budget	0.155			0.158*
	(0.101)			(0.083)
large account		0.702*	0.461	
		(0.406)	(0.429)	
# large clients at ag $\times$ large account			0.132*	
			(0.080)	
const	-3.299***	-3.360***	-3.175***	-5.431***
	(1.022)	(1.006)	(1.007)	(1.052)
Agency f.e.	Yes	Yes	Yes	
Category f.e.	Yes	Yes	Yes	
Time f.e.	Yes	Yes	Yes	
N	2,331	2,331	2,331	477
N markets	540	540	540	477
N reviews	75	75	75	75

Notes: Standard errors are in parentheses, *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively. The specification includes top-27 HC agencies, four aggregated HC agencies, and frequently bidding standalone agencies: Barkley, Droga5, McKinney, Wieden & Kennedy.

increasing agency billings by $100M (from unconflicted categories) from the mean, increases the odds of the agency being chosen by its HC to compete for an account by 10.3%, holding all else equal.

Second, we find that serving an additional competitor of an account in review decreases the odds of the agency being chosen to bid for this account by about 91.6%.[23] The negative sign of the competitor avoidance coefficient suggests that if there are any benefits from industry specialization, they are

[23]Such a high estimate is certainly in line with the statements made by our industry contacts. One of them recounted a discussion that their team had about whether they could work with a non-profit that worked in the same category as an existing for-profit client.

likely outweighed by the avoidance effect. Thus, agency separation by industry specialization is unlikely to rationalize the HC structure on its own.

Next, we turn to specification (3). The interaction term between the number of large accounts served by agency and the large account dummy test for avoidance of large advertisers due to capacity constraints. We find that adding an additional large client increases the agency's odds of bidding for a large account by 14.1%. It is likely that the interaction term picks up some of the scale effect since the coefficient on log billings in other categories drops to 0.120. Therefore, the avoidance of large advertisers explanation for existence of HCs is not supported in our data.

Finally, competitor conflicts do not appear to be binding at the HC level as the coefficient on the number of competitors is close to zero and insignificant.

Combining the evidence from all specifications, we can conclude that the competitor avoidance effect is extremely strong at the agency level and can rationalize existence of holding companies in our model. The magnitude of the avoidance effect suggests that its impact on firm structure and concentration is economically significant. In the next section, we simulate a counterfactual evolution of the industry between 2005-2018 assuming that competitor avoidance effect is zero. We also simulate how the industry would evolve if the coordinated bidding within HCs were banned.

1.8 Counterfactuals

In this section, we conduct two counterfactual exercises that study the impact of competitor avoidance and bid coordination on various market outcomes. In the first counterfactual, we study the impact of competitor avoidance on the market shares of major holding companies and market concentration. In the second counterfactual, we examine how bid coordination within HCs affects the number of bidders competing in account reviews.

1.8.1 *Competitor Avoidance and Firewall Policies*

In 2000, the American Association of Advertising Agencies (4A's) published a position paper on conflict policy guidelines, explaining its perspective on appropriate conflict policies. The 4A's argued that the industry lacked a common standard for evaluating what constituted a conflict. Instead, many conflicts were determined on a case-by-case, sometimes arbitrary, basis and were rarely referenced in client-agency contracts. In particular, the 4A's suggested that agencies should be able to serve competitors as long as their accounts were handled out of different offices of that agency, citing practices used in other professional service industries that assign different staff to handle competing accounts.

Over two decades later, this suggestion does not seem to have taken on. Even though agencies claim that they maintain strict firewalls between competing accounts, agencies to this day steer clear of new accounts if there is a potential conflict with an existing client. Moreover, accounts frequently leave their agencies if such conflicts emerge.[24] Motivated by the lasting divide between agencies and advertisers in their view of conflicts of interest and the prediction in Section 1.5.3 that some holding companies benefit from competitor avoidance, while others are hurt by it, we use our estimates and model of agency bidding in order to simulate what would happen to the market shares of major holding companies (and the market HHI) in the world without competitor avoidance.

We first simulate the predicted (factual) market shares of major HCs using the sample of all relationship changes observed in 2005-2018. We then

[24] Some examples include: Havas Worldwide sought to move Charles Schwab to a different agency within Havas to take on TD Ameritrade (Adweek, 2014a); Omnicom is likely to steer clear of the pitch due to its close ties with Apple, a major competitor of Microsoft (Adweek, 2014b).

simulate the counterfactual market shares by setting the coefficient of competitor avoidance α to zero, approximating the situation where advertisers would not mind sharing agencies with competitors. Both simulations are based on market observables. In particular, HC market shares are calculated based on agency billings, which are, in turn, calculated from the advertising budgets observed in the market. We assume that the HC structures remain fixed even though competitor avoidance is removed.

We simulate counterfactual market outcomes over 5,000 paths of error draws. We draw three types logit shocks: (i) ε_{hjm}, the agency-account-specific shock that affects the bidding surplus of agency j of HC h when bidding for account m; (ii) $\varepsilon_{h\emptyset m}$, the shock to the HC's value of outside option, which we interpret as the shock to HC's costs of participating in m's review; and (iii) ν_{hjm}, the shock that affects the match surplus between j and m after j enters the review.

Figures 1.9 and 1.10 illustrate the predicted (solid line) and counterfactual (dashed line) evolution of market shares of major HCs with and without avoidance. Table 1.6 summarizes the market shares of HCs and standalone agencies at the end of the simulation period. Standalone agencies, which tend to be smaller than HC agencies, benefit from avoidance because advertisers are willing to hire a smaller agency, hoping to avoid competitors. The majority of the holding companies also benefit from avoidance. WPP benefits the most, increasing its market share by 3.29% relative to the counterfactual without avoidance.

Omnicom and IPG, the two largest holding companies, are hurt by competitor avoidance. Omnicom's counterfactual market share would go up by about 9.5pp if competitor avoidance were removed. For IPG, the counterfactual market share would go up by about 3.3pp. In the counterfactual, the two HCs would dominate the market, capturing 44% and 35% of the market. This result is consistent with our predictions in Section 1.5.3. As the scale effect becomes increasingly more important, the largest agencies in the sample keep growing even larger if there is no competitor avoid-

TABLE 1.6 – PREDICTED AND COUNTERFACTUAL MARKET SHARES WHEN AVOIDANCE IS REMOVED

HC or SA	Predicted	Counterfactual	ΔP−C
Dentsu	1.35	0.50	0.85
Havas	3.34	2.10	1.24
IPG	31.91	35.19	-3.28
MDC Partners	4.16	2.23	1.92
Omnicom	34.59	44.05	-9.46
Publicis	7.22	4.77	2.45
WPP	9.27	5.98	3.29
Barkley	2.83	1.43	1.40
Droga5	0.61	0.24	0.36
McKinney	1.73	1.29	0.43
Wieden & Kennedy	3.00	2.21	0.80

Notes: The market shares are reported in percents and rounded to two decimals. Barkley, Droga5, McKinney, Wieden & Kennedy are the four standalone agencies in our sample.

ance. Indeed, Omnicom and IPG own the two largest agencies in our sample – BBDO and The Martin Agency. In our simulations, the market shares of BBDO and The Martin Agency experience the largest (8%) and second-largest (4.6%) increase in market shares.

Even though there are winners and losers from competitor avoidance, the fact that the largest HC (Omnicom) is predicted to gain 9.5% in the counterfactual produces an HHI increase of 35% by the end of the simulation period. In fact, the evolution of HHI closely follows the evolution of Omnicom's market share.[25] More specifically, the HHI would increase from 2,400 to 3,250. Therefore, removing avoidance would push the market from moderately to highly concentrated classification, according to the horizontal merger guidelines provided by the DOJ (2010).

Considering the importance of starting market shares for the counterfactual predictions, we investigate how BBDO happened to grow so large.

[25]The market shares with and without avoidance only start to visibly diverge in 2008 because we observe 3 and 4 smaller account moves between 2005-2006 and 2006-2007. Later in the sample we observe on average 14 account moves each year.

Firstly, BBDO had the early entry advantage—its roots go back to 1891—making it the oldest agency in our sample. Moreover, BBDO won the Network of The Year title from World Advertising Research Center each year from 2006 to 2019 among other awards for their specific creative campaigns. Therefore, the combination of first-mover advantage and consistently high agency quality have likely contributed to BBDO's size.

1.8.2 Independent Bidding

We now turn to examining how bid coordination within holding companies affects the number of bidders competing in account reviews. Unfortunately, due to lack of pricing data, we are not able to analyze how bid coordination impacts agency compensation directly. However, we are able to use our theory model to qualitatively predict how bid coordination may impact prices through the change in the number of bidders. Moreover, our industry contacts suggested that bid coordination is used not only as a way of managing client conflicts, but also as a way of saving the review participation costs associated with developing spec work. Therefore, change in the number of bidders entering from the same HC is interesting in its own right.

In this exercise, we remove the restriction that each HC can submit only one bid for a given account in review. In order to facilitate the comparison between the factual and counterfactual outcomes, we maintain the assumption that agencies from the same HC receive identical outside-option shocks, i.e. identical shocks to bidding costs. For standalone agencies, the ban on bid coordination will only impact the number of submitted bids and account wins through changes in bidding behavior of holding companies.

Table 1.7 presents the results of this counterfactual exercise. We report the total number of bids entered by a HC or SA for the 176 account changes observed in the sample. The reported numbers are averaged across 5,000 simulations. For this reason, the total number of wins sums up to 176 for both joint and independent bidding.

TABLE 1.7 – PREDICTED AND COUNTERFACTUAL MARKET SHARES IF BID COORDINATION IS BANNED

HC or Standalone	Number of entered bids			Number of wins		
	Joint	Independent	ΔJ−I	Joint	Independent	ΔJ−I
Dentsu	51.76	84.53	-32.77	2.64	2.54	0.10
Havas	53.19	85.00	-31.81	5.82	5.55	0.27
IPG	101.04	317.98	-216.93	47.17	51.07	-3.90
MDC Partners	70.33	137.62	-67.29	6.67	6.17	0.50
Omnicom	88.56	292.98	-204.42	78.46	77.04	1.42
Publicis	68.99	247.37	-178.39	7.65	7.30	0.35
WPP	78.61	219.47	-140.86	16.92	16.21	0.72
Barkley	52.79	51.70	1.10	5.44	5.16	0.28
Droga5	38.23	37.00	1.23	1.26	1.19	0.08
McKinney	43.74	42.60	1.15	1.72	1.66	0.06
Wieden & Kennedy	52.40	51.33	1.07	2.25	2.11	0.14

Notes: The reported number of entered bids and wins are the averages over 5,000 simulations, rounded to two decimals. The total number of wins sums up to 176, the number of relationship changes observed in the sample. Barkley, Droga5, McKinney, Wieden & Kennedy are the four standalone agencies in our sample.

As expected, the change in bidding and winning patterns of the standalone agencies is minimal. Mechanically, SAs win more often under bid coordination because HCs enter fewer bidders. The increased number of wins changes the characteristics of SAs, which makes them more likely to bid in future account reviews. This intuition explains the slight discrepancy in the simulated number of entered bids for standalone agencies under joint and independent bidding.

For the holding companies, the number of entered bids increases dramatically—on average, by 150%. The number of account wins barely changes, suggesting that any additional agencies entered due to ban on bid coordination are rarely chosen by advertisers. Any discrepancies in the number of wins are due to additional draws of ν_{hjm} winning shocks that HCs get by entering multiple agencies.

The only HC that potentially stands to benefit from ban on bid coordination is IPG. Back of the envelope calculations suggest that this prediction is unlikely to hold true in reality. Based on the average account budget from Table 1.4, IPG could attract an additional $578.5M in agency billings. Using an, albeit outdated, compensation practice of agencies working on a flat 15% commission over an account's advertising budget, IPG would raise an additional $86.78M in revenues (Denford and Indo, 2010). IPG's trailing 12-month operating margin in 2018 was around 10.3% (Macrotrends, 2022). Thus, the implied $8.7M increase in operating profit should justify the costs associated with submitting nearly 217 additional bids. Thus, the average break-even bidding cost should be around $41,000. The actual bidding costs are likely much higher. The only benchmark estimate of bidding costs we were able to find is from a survey of 120 Australian agencies that a single review cost at around $30,000 in US dollars (Duval, 2021). However, this survey is based not only on creative, but also on PR and communications agencies, whose costs rarely include costs of production of ad copies. Thus, we expect that the surveyed costs are a lot lower compared to the bidding costs of creative AORs.

Overall, the number of bidders in an average account review goes up from 3.97 to 8.91 as agencies switch to independent bidding. According to our theoretical model, an increase in the number of bidders weakly increases the surplus of the second-scoring agency. By equation 9 from Section 1.4, this translates into a weak decrease in the winning agency's mark-up. Note that with the ban on bid coordination, HCs may enter not only their highest-surplus agencies, but potentially those that were second-best in their 'internal' auctions. Therefore, leaving the truthful revelation issues aside, any additional competitive pressure comes from HCs' own agencies, not from competing HCs. Thus, it is possible for a HC to cannibalize its own mark-up if it submits a runner-up in an account review. At the same time, higher fixed costs of winning an account, due to submitting more bids at the holding company level, may counteract the downward pricing pressure associated with increased competition.

Our prediction of the impact of bid coordination on agency compensation should be interpreted with caution. In a more realistic setting, the increase in costs associated with a more frequent bidding may push some HC agencies out of the market, offsetting the increase in the number of bidders. A definitive answer to this question is left for future research and is contingent on availability of data on agency compensation.

1.9 Conclusion

In this paper, we study the market for advertising agencies that is characterized by competitor avoidance, economies of scale, and the holding company organizational structure. We use a theoretical model and a series of supporting simulations to show that the HC structure can be viewed as an endogenous response to competitor avoidance. It helps agencies to mitigate client conflicts and save bidding costs through bid coordination. We show that the HC structure arises when the competitor avoidance effect is strong relative to the scale effect. Our estimates show that this is indeed in the case. Increasing an agency's billings by $100M (from unconflicted accounts) from the mean increases the odds of the agency being chosen by its HC to compete for an account by 10.3%. At the same time, we estimate that serving an additional competitor of an account in review decreases the odds of an agency to bid for this account by about 91.6%.

We conduct two counterfactual exercises based on these estimates. In the first counterfactual, we investigate how the market shares of major HCs and standalone agencies would change if competitor avoidance was no longer a concern in this industry. We show that standalone agencies and most of the HCs benefit from competitor avoidance. The only two HCs whose market shares are hurt by it are Omnicom and IPG, who own the two largest agencies in the market, BBDO and The Martin Agency. We predict that removing avoidance would shift clients away from smaller agencies and towards BBDO and The Martin Agency and increase the market HHI by 35%.

In the second counterfactual, we examine the consequences of banning bid coordination within HCs. According to our industry contacts, HCs coordinate bids not only to manage conflicts, but also to save costs of creating an original sample work in preparation for reviews. If bid coordination were banned, we predict that an average holding company would increase the number of bids that it enters by 150% without any meaningful increase in the number of accounts it wins. This implies that bid coordination at the HC level is indeed effective at saving the bidding costs and may be responsible for lower compensation mark-ups depending on the pass-through rate in this market. In addition, we find that banning bid coordination increases the average number of bidders in a review from four to nine. Even though we are not able to quantify the impact on agency compensation, our theoretical model predicts that an increase in the number of bidders would decrease agency compensation. These predictions should be interpreted with caution because dynamic incentives are likely to be important in this framework. Increased bidding costs may put some HC agencies out of business, offsetting the predicted increase in the number of bidders. We leave a more definitive answer to this question for future research.

FIGURE 1.9 – PREDICTED AND COUNTERFACTUAL MARKET SHARES WHEN AVOIDANCE IS REMOVED

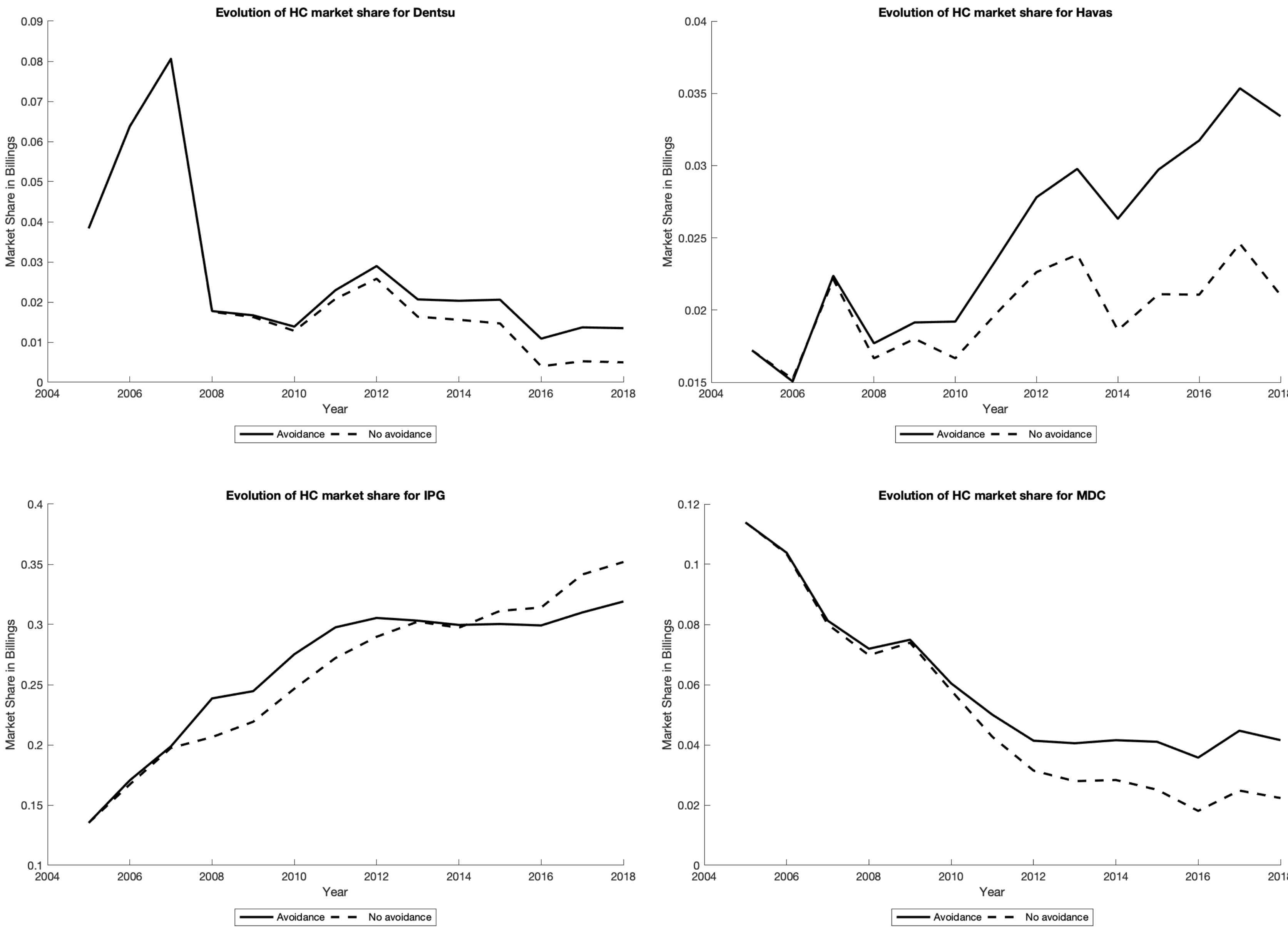
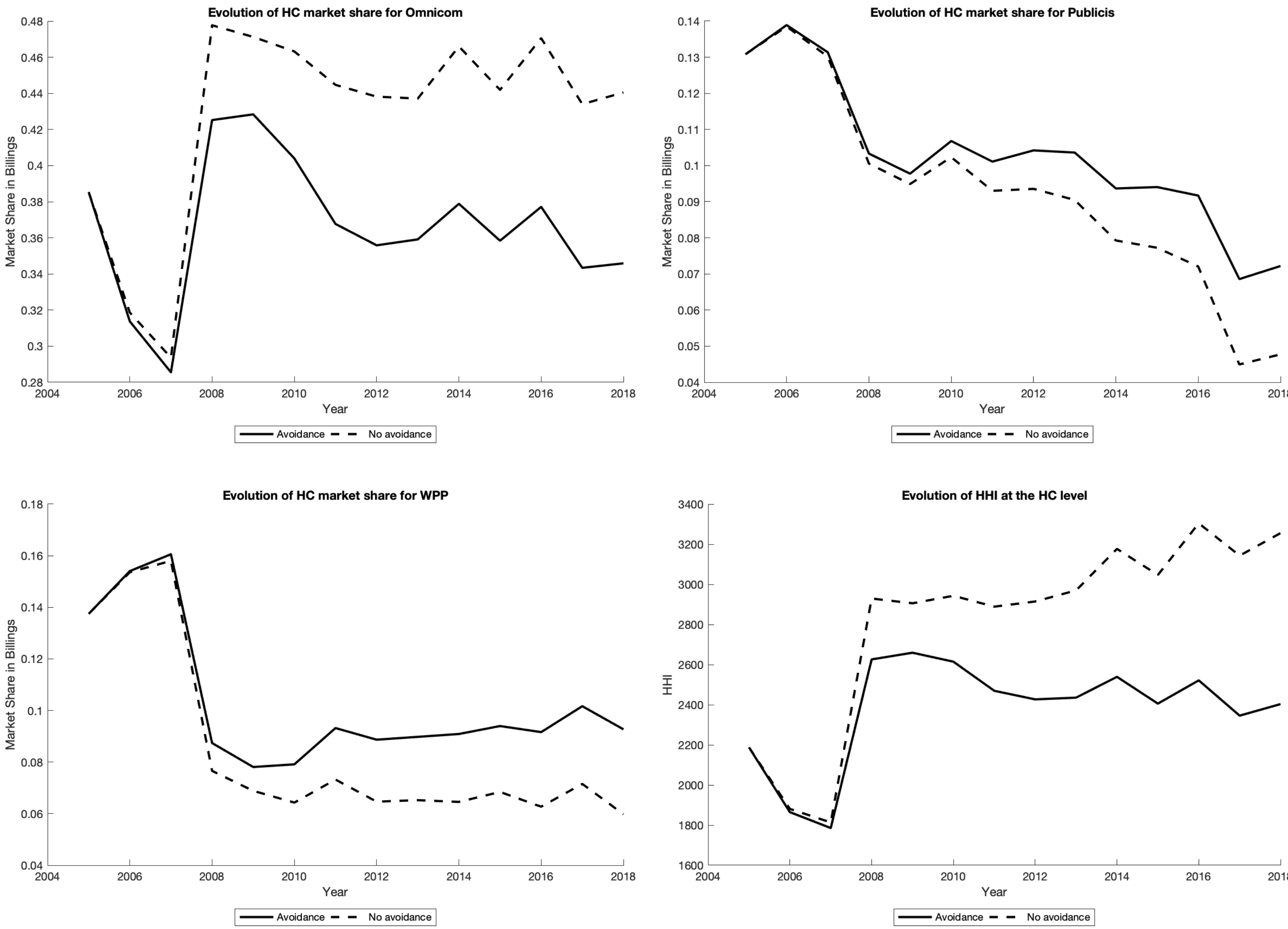


Figure 1.10 – Predicted and Counterfactual Market Shares When Avoidance Is Removed (Continued)



Chapter 2

"Use It or Lose It," or "Cheat and Keep?" The Effects of Slot Restrictions on Airline Incentives

with Ratib Ali

2.1 Introduction

Air traffic congestion has been responsible for a significant welfare loss in the US economy. Peterson et al. (2013) estimate that a 10% reduction in flight delays[26] would increase net total welfare by $17.6 billion. The congestion problem has been a major concern for all stakeholders since the 1960s and is exacerbated by limited airport capacities.

Regulators all over the world, including the Federal Aviation Administration (FAA) in the US, manage congestion by capping the number of hourly flights and rationing the rights to perform a take-off or a landing within a given timeframe (also known as "slots") among airlines. Slot controls are currently in place in most major airports worldwide and in three US airports – JFK and LaGuardia Airports in New York City, and Reagan National Airport near Washington, DC.

Following the International Air Transport Association (IATA) procedures, the FAA revisits slot allocations each year at the start of the winter and summer seasons. Airlines are allowed to keep their slot holdings for the next season provided that they comply with the use-it-or-lose-it rule by using their slots at least 80% of the time during the current season. This approach biases the allocation process in favor of legacy carriers, who were

[26]Flight delays are measured as a fraction flights that are delayed by 15 minutes or more.

incumbents when slot control was first introduced in 1969 and still hold the majority of slots in slot-controlled airports.[27]

While this regulation has been successful in managing congestion and delays, it may also create an incentive for airlines to hold onto unprofitable slots in order to keep their competitors out of highly-demanded airports (GAO, 2012).[28] Specifically, when slots at an airport are limited, airlines not only want a slot to operate a flight on a given route, but also to prevent a rival from controlling said slot and competing directly against the incumbent, possibly on other, more profitable, routes. The incentive to foreclose manifests in slot burning – using slots at a loss instead of forfeiting them to a rival and incurring a greater loss in profits.

Given the prevalence of slot control, it is important to study its effects on consumers and investigate the alleged anticompetitive incentives it creates for airlines. This paper uses reduced-form analyses to assess evidence of slot burning, relying on a natural experiment created by the removal of slot control at Newark Airport in November of 2016. Removal of slot control eliminates historical precedence created by the use-it-or-lose-it rule and allows entry of new airlines who did not previously hold slots at the airport, thereby eliminating slot-burning incentives.

We measure the extent of slot burning by using the empirical probability of observing a small flight on a given route as a proxy variable. Anecdotally, we have heard of airlines operating frequent flights using smaller aircraft to use up multiple slots as opposed to carrying the same number of passengers in fewer flights using larger aircraft (GAO, 2012). Slot incumbents, owing to their large slot endowments, may be more prone to slot burning. In line with the anecdotes, we find circumstantial evidence for airlines burning slots: frequency of small flights between Newark and Philadelphia

[27]Calculated from the slot holder reports for the winter season of 2018 published by the FAA.

[28]Due to the disruptive effects of the COVID-19 pandemic, the FAA suspended the use-it-or-lose-it rule on March 11, 2020 in order to relieve airlines from the need of flying 'ghost planes', in other words, slot burning (Pallini, 2020).

Airports, only 80 miles apart, decreases by nearly 77% following removal of Newark's slot control in 2016, with slot incumbents accounting for 35% of the drop.

Indeed, we find that slot incumbents are twice as likely to operate smaller aircraft if slot restrictions are in place. We note that slot burning occurs most often during relatively offpeak hours of 10am to 1pm – a timeframe where there is enough demand to support some flights, but not enough demand to warrant usage of larger aircraft.

Airlines, however, contend that using multiple smaller flights is needed to meet demand for schedule-sensitive passengers. We address this argument by looking at the probability of flying small aircraft to the same destination within a 30-minute timeframe. We find that slot incumbents are 75% more likely to fly consecutive small flights when Newark Airport is slot-controlled.

We also find that slot restrictions are associated with increased airfares and decreased delays, suggesting that, while slot controls are effective in decreasing congestion, they do so at the expense of higher airfares due to restricted entry. In addition, we analyze changes in several quantity metrics, namely, the number of seats(-miles) offered on a route (as a proxy for quantity supplied) and the number of passengers(-miles) transported (as a proxy for quantity demanded). We find that slot incumbents offer more seats and transport more passengers under slot restrictions. Together with the increase in airfares, this observation suggests that both demand and supply shift under slot control. We discuss a potential rationale for these shifts and their effects on consumer surplus in Section 2.5.

Lastly, we investigate the patterns of entry to and exit from markets, finding that slot liberalization resulted in entry by low-cost carriers to certain routes. We highlight the competing interests that must be met when aviation authorities consider mitigating congestion through slot restrictions; namely, effects on price, delays, entry, and the incentives of slot incumbents to burn slots. An evaluation of consumer welfare, as a function of frequency, prices, and delays, is left for future work.

The rest of the paper is organized as follows: Section 2.2 surveys the relevant literature. Section 2.3 reviews the industry background and introduces the data. Section 2.4 develops a theoretical model that informs our reduced-form analyses. Section 2.5 presents our hypotheses and discusses the results. Section 2.6 provides an overview of the entry and exit decisions following Newark's reclassification. Section 2.7 concludes.

2.2 Related Literature

Our research contributes to the literature on airlines' access to airport facilities and its effect on downstream market outcomes, specifically airfares and congestion. Ciliberto and Williams (2010) investigate whether access to airport gates, which is usually determined by long-term exclusive contracts between airlines and airports, allows airlines to charge higher prices on flights in and out of their hubs. In the first stage, Ciliberto and Williams (2010) recover carrier-route fixed effects from a reduced-form pricing equation as a measure of the hub premium, which they later regress on measures of access to airport gates. They find that an increase in the percentage of gates controlled by an airline is associated with an increase in its airfares, especially in more congested airports as defined by the number of departures per gate.

Snider and Williams (2015) study the changes in airfares following the change in access to airport facilities mandated by the AIR-21 Act. They use a regression discontinuity approach to identify changes in airfares at airports that were required to improve access to their facilities, relative to airports that were exempt from this requirement based on two threshold rules. They find that airfares decrease by 13.4% on routes where one airport is covered by the legislation and by 20.2% on routes where both end points are covered.

In a recent paper, Fukui (2019), similarly to us, uses the exogenous change in slot restrictions at Newark Airport in 2016 to study the impact of slot control on average airfares. They use a difference-in-differences approach, where the treatment group consists of routes to or from Newark and the control group consists of routes to or from the two other NYC airports – the JFK and LaGuardia Airports. They find that the average fare on Newark routes decreases by about 2.5% relative to the JFK and LGA routes, with the majority of the effect coming from non-dominant Newark airlines. Our study documents the effect on airfares by using all major airports as the control group, as opposed to just JFK and LGA, although our primary focus is on slot burning.

In another study from Newark Airport, Luttmann (2019) exploits reinstitution of slot control at the JFK and Newark Airports in 2008 to evaluate the effectiveness of slot control in managing delays. Using the 2007-2008 data, Luttmann (2019) finds no evidence of a reduction in delays at both airports. They suggest that these findings are consistent with the internalization hypothesis claiming that delays at an airport decrease with an emergence of a dominant airline. In contrast to Luttmann's results, we find that delays at Newark Airport increase following the removal of slot restrictions in 2016. Figure 2.1 shows that the average length of delays at Newark Airport diverges away from other NYC Airports following the abolition of slot restrictions. Additionally, we find a decrease in delays following the 2008 classification as well (unlike Luttmann, 2019), but we focus on the 2016 (as opposed to the 2008) reclassification to abstract away from the confounding effects of the Great Recession.[29]

A 2012 report on slot-control rules published by the Government Accountability Office is the closest to our paper and has inspired our proxy variable for slot burning. In particular, GAO proposed three indicators of airline slot burning: (i) using smaller aircraft; (ii) flying to the same desti-

[29] We submitted FOIA requests to PANYNJ asking for Newark's time-stamped departure and arrival data from 2007-2009, but, after a diligent search, the agency could not locate any responsive records.

nation at higher daily frequencies; and/or (iii) operating flights with lower average load factors (passenger-to-capacity ratio). In particular, GAO compares the number and proportion of small aircraft flights (under 100 passengers) to and from slot-controlled airports to those from other large domestic hubs, controlling for flight distance and other relevant characteristics. They find that the odds that a flight to and from a slot-controlled airport uses a small aircraft are 75% higher than the odds for a flight to and from other large hub airports that are not slot-controlled (GAO, 2012). The evidence that GAO finds is only suggestive since the estimates are not causal and rely on how well other large domestic airports act as a control group for the slot-controlled airports. We improve upon the GAO's methodology by considering a natural experiment created by exogenous removal of slot restrictions at Newark Airport in 2016.

A related paper looks at the divestment of slots as a structural remedy to the 2013 US Airways-American Airlines merger to back out airline response to the reallocation of slots designed to promote competition and the associated effects on consumers (Ali, 2020). This paper, instead, summarizes airline behavior and its effects on consumers following a wholesale abolition of slot controls.

Lastly, Swaroop et al. (2012) investigate whether more US airports need slot control and if the existing slot levels are optimal. In particular, they use an econometric model to quantify the costs of schedule delay, i.e. costs that a passenger suffers from having to choose the departure time from the set airline schedule as opposed to flying at her preferred time, and the costs of queuing delay, resulting from congestion. They later simulate optimal slot control policies at major US airports that minimize the sum of schedule and queuing delay costs and conclude that slot control should be implemented at 12 additional airports and slot caps decreased at the already slot-controlled airports. However, they do not take into account the impact of likely airfare increases as a result of constrained capacity and do not take into account anti-competitive effects generated by slot burning.

2.3 Industry Background and Data

2.3.1 Slot Control at Newark Airport

The history of slot control in the United States goes back to the introduction of the High Density Rule (HDR) in 1969. The HDR capped the number of hourly arrivals and departures at five major airports – JFK, LaGuardia, Newark, O'Hare, and the Reagan Airport – and was seen at the time as a temporary measure to curb growing delays. The rule was suspended in Newark a year later since the number of flights was well under the cap, even at peak times. However, the HDR proved to be successful in managing congestion in the rest of the airports, so the FAA extended it indefinitely in 1973.

Initially, the slots were allocated by a group of airline representatives, the so-called scheduling committees, on a voluntary concession basis. However, after the deregulation of the airline industry in 1978, scheduling committees had been having difficulties agreeing on slot allocations, and the antitrust immunity which made their existence possible invited scrutiny from the FTC. By 1986, the FAA replaced scheduling committees with the slot-allocation procedures as they are currently known. In particular, the FAA instituted the use-it-ot-lose-it rule that stipulated withdrawal of slots that did not clear the 65% usage threshold. The minimum utilization threshold was increased to its current level of 80% in 1992.[30]

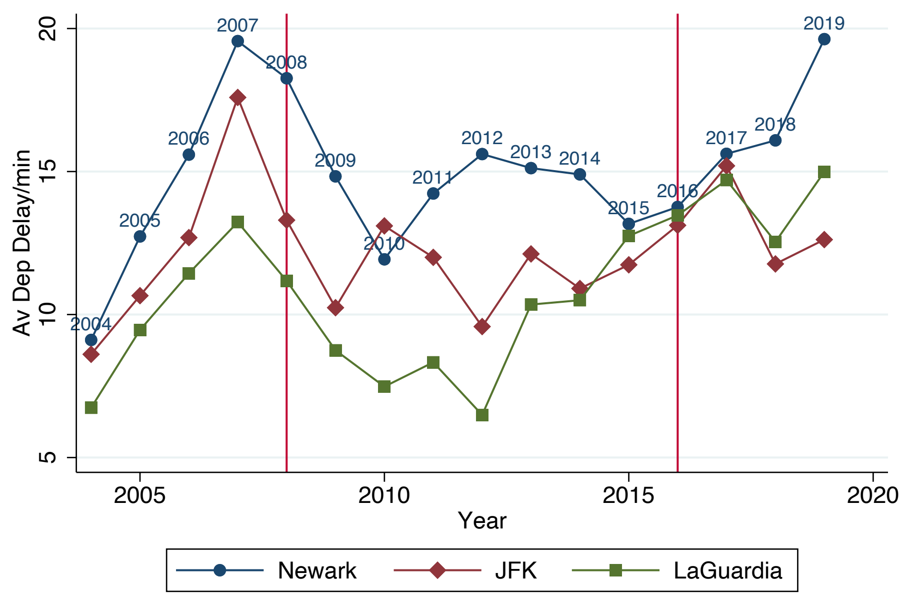
Since then, several government agencies[31] studied the effects of the High Density Rule on the quality of air service and concluded that the HDR was limiting competition and preventing improvements in the quality of

[30] Amdt. 93-65, 57 FR 37315, Aug. 18, 1992.

[31] GAO, Airline Competition: Industry Operating and Marketing Practices Limit Market Entry, GAO/RCED-90-147 (Washington, D.C.: Aug. 29, 1990); National Research Council Transportation Research Board, Entry and Competition in the U.S. Airline Industry: Issues and Opportunities, Special Report 255 (Washington, D.C.: 1999); Department of Transportation, Study of the High Density Rule: Report to Congress (Washington, D.C.: May 1995).

service, partly because new entrants could not enter the slot-controlled airports. Eventually, the AIR-21 Act (2000)[32] called for a phase-out of the HDR at Chicago O'Hare Airport by July 2002 and at JFK and LaGuardia Airports by January 2007. After the expiration of slot control rules, air carriers have promptly increased their operations in JFK, making 2007 one of the worst years in terms of delays. All three of the New York City metropolitan area airports were affected by congestion at JFK (see Figure 2.1), so the FAA temporarily reinstated slot control soon after. Even though Newark remained slot-free from 1970 to 2008, the FAA decided to preemptively institute slot control rules at Newark as well, fearing that air carriers would shift their operations from JFK and LaGuardia and create additional congestion at Newark.[33]

FIGURE 2.1 – AVERAGE DEPARTURE DELAYS AT NYC AIRPORTS



Generally, existence of slot control at an airport entails: (i) caps on the number of arrivals and departures performed within a 30-minute and one-hour timeframe; (ii) minimum usage requirement applied over a pool of slots within a given timeframe over a predefined period (henceforth, re-

[32]The Wendell H.Ford Aviation Investment and Reform Act for the 21st Century. Public Law 106-181.

[33]73 FR 60543.

ferred to as a slot period; differs across slot-controlled airports). For Newark Airport, the limit on the number of operations was set at 44 within each 30-minute window and 81 within each one-hour window from 6am to 10:59pm every day of the week.[34] The compliance with the minimum usage requirement was thus determined for each day of the week within a 30-minute and one-hour time periods, for example, Mondays from 6:00 to 6:30am during the winter season of 2015.

Newark's slot control rules were in place until the winter season of 2016. According to the FAA, the reasons for removal of slot restrictions in 2016 at Newark were three-fold: (i) improved capacity at JFK, following the runway reconstruction scheduled to begin in 2017, was expected to decrease the spillover effect that prompted slot control back in 2008; (ii) improvement in on-time performance and duration of delays from 2007 to 2015, and (iii) the FAA's prediction of future demand and capacity at Newark Airport suggested that the slot restrictions were no longer necessary.

As for the first reason, both JFK and LaGuardia underwent runway reconstructions that temporarily reduced the airports' capacities in 2017. The total number of scheduled flights at JFK and LGA decreased by about 3,000 and 1,200 respectively, relative to the 2016 levels, reversing the historical trends. The number of scheduled flights promptly returned to and exceeded the 2016 levels in 2018, after the construction projects were completed. In light of these events, we investigate if the FAA removed slot control at Newark to allow air carriers to reschedule their operations from JFK and LGA to Newark in anticipation of restricted capacity. We find no evidence of shifts in JFK's operations. However, we do find that the LGA routes that experienced a decrease in the number of scheduled flights in 2017 or 2018 (relative to 2016) tend to experience an increase in scheduled frequency at Newark. Routes that potentially shifted from LGA to Newark are not a part of the sample of airports we use to test for slot burning, therefore we believe that possible shifts in operations did not affect patterns of aircraft usage at

[34] 14 CFR 93.163(b) of January 1, 2009.

Newark in any spurious manner.[35]

The second and third reasons cited by the FAA are both predicated on current improvements in on-time performance, as well as a belief held by the FAA (based off their future demand prediction at Newark Airport) that relaxing slot restrictions will not result in congestion due to entry. However, government and industry reports, as well as this study, show an increase in flight delays following the 2016 reclassification. This implies that any gains in on-time performance between 2009 and 2015 were likely due to effectiveness of slot-control restrictions. Figure 2.1 above shows that the average delay for flights departing from Newark Airport fell in 2008 when the airport was escalated to a slot-restricted airport, but the average delays went back up to its pre-2008 levels following a de-escalation to a slot-facilitated airport. No such reversions occurred for other NYC airports in 2016 that kept slot controls and completed runway reconstructions by 2018. As a result, we believe that the reclassification was not endogenous to any sustained or systemic changes in schedule management at Newark Airport, and therefore, can exploit the reclassification as an exogenous change.

2.3.2 *Suggestive Evidence of Slot Burning at Newark Airport*

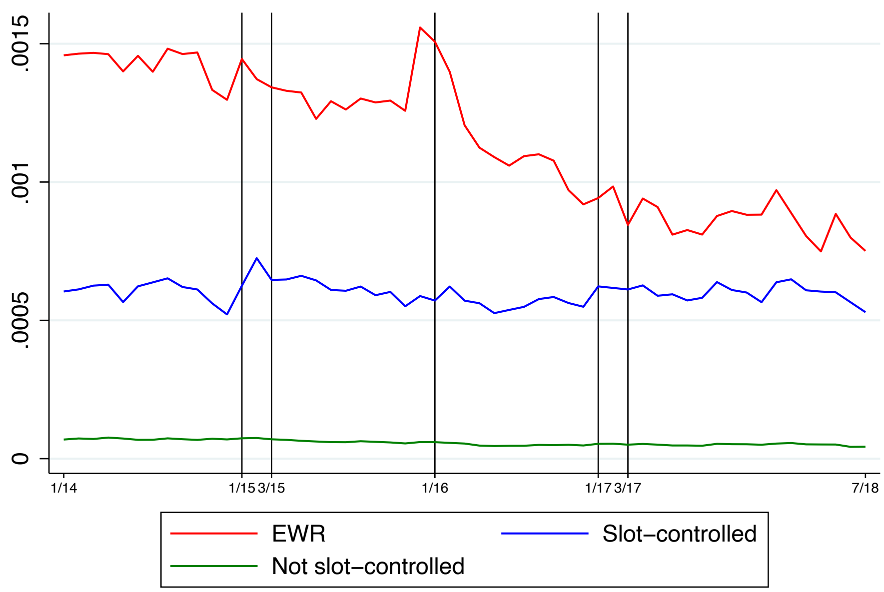
Interestingly, the FAA's review of Newark's operational performance concluded that scheduled demand was consistently below the 81 hourly flight cap, yet the FAA could not accommodate requests for new flights in summer of 2016 as the allocated slots reached the limit.[36] This conclusion suggests that the incumbent carriers might have relied on slot burning to meet the use-it-or-lose-it requirements and keep new entrants out of the airport.

[35]See Appendix B for more details.

[36]From 81 FR 19861: "For example, in the 3 p.m. through 8:59 p.m. local hours, weekday scheduled demand in the May-August period averaged 71 flights per hour in 2011, 74 flights per hour in 2013, and 72 flights per hour in 2015. [...] At the same time, the FAA denied requests for new flights as slots are allocated up to the scheduling limits."

As mentioned previously, our proxy variable for slot burning is usage of small aircraft. If an airline is burning a slot in order to prevent a competitor from acquiring it, the airline would minimize losses associated with operating an unprofitable flight by flying a smaller aircraft. To this end, Figure 2.2 highlights three stylized facts. First, slot-controlled airports use more small aircraft than non-slot-controlled airports. Second, in line with the FAA's evaluation of capacity usage at Newark prior to 2016, the share of small flights in Newark is around three times higher than in the other slot-controlled airports. Third, there is a meaningful change in the usage of small aircraft in Newark Airport around 2016 that cannot be explained by a general time trend.

FIGURE 2.2 – PERCENTAGE OF SMALL FLIGHTS AT NEWARK AND OTHER TOP-28 AIRPORTS



Moreover, airlines with relatively large pools of slots under their control may burn slots more often relative to airlines with smaller slot endowments. We refer to such airlines as slot incumbents, and we split our analysis by whether an airline is a slot incumbent or not. Table 2.1 below summarizes the total number of daily slots held by each airline at Newark in 2015. United, together with its regional partners, like Air Wisconsin and Republic Airways, held 869 daily slots, nearly 80% of all slots available at Newark.

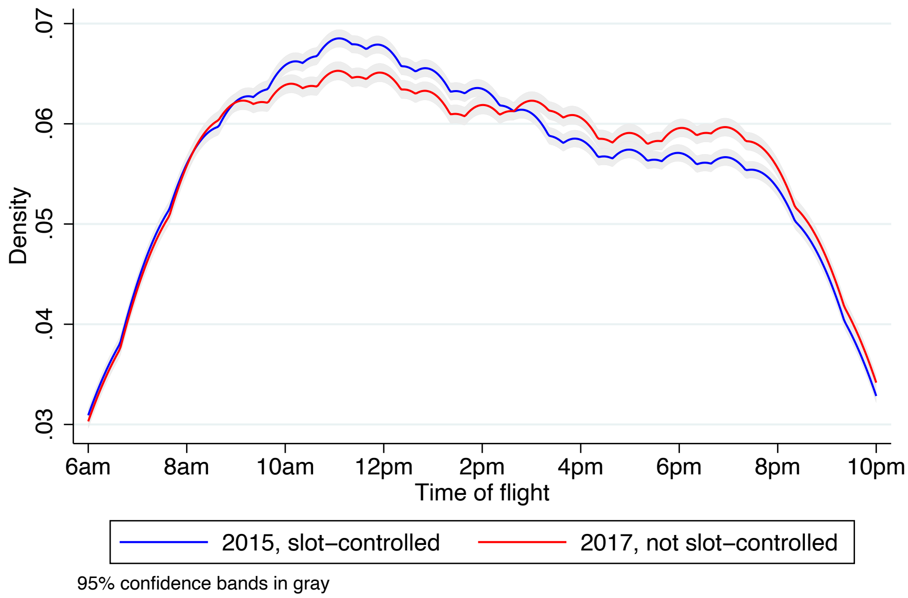
TABLE 2.1 – DAILY SLOT HOLDERS IN 2015

Airline	Daily slots	Percentage
American	64	5.87
Alaska	4	0.37
Delta	65	5.96
JetBlue	40	3.67
Southwest	35	3.21
United	869	79.72
Virgin	13	1.19

In addition to varying daily slot endowments across airlines, the same airline generally holds a different number of slots in each slot period. This is due to the fact that slots granted to airlines are tied to a particular one-hour slot period. Furthermore, compliance with the use-it-or-lose-it requirements is determined based on utilization of slots in the same 30-minute or one-hour slot period during a scheduling season. These two facts imply that some slot periods may experience more slot burning than others due to lack of demand or differing slot-burning incentives of airlines. Taking this observation into account, we revisit patterns of usage of small aircraft by time of arrival to/departure from Newark Airport. Figure 2.3 suggests that, under the slot control regime, the empirical probability of observing a small flight within the 10am-1pm timeframe is higher relative to the 4-8pm timeframe, while it is relatively uniform after slot control rules are lifted. We employ a Kolmogorov-Smirnov test for equality of the two distributions and reject the null hypothesis that the two distributions are the same with a D-value of 0.0468 (p=0.000).

We also investigate in which slot periods United, the slot incumbent, and the rest of Newark's airlines may be burning slots. As discussed more rigorously in section 2.4, we incorporate slot and minimum usage constraints into airlines' profit-maximization problems. Whenever an airline's minimum usage constraint is binding and the slot constraint is slack, we interpret such a phenomenon as slot burning. Figures 2.4 and 2.5 attempt to assess if United and Newark's low-cost carriers complied

FIGURE 2.3 – DISTRIBUTION OF SMALL AIRCRAFT BY TIME OF FLIGHT AT NEWARK AIRPORT



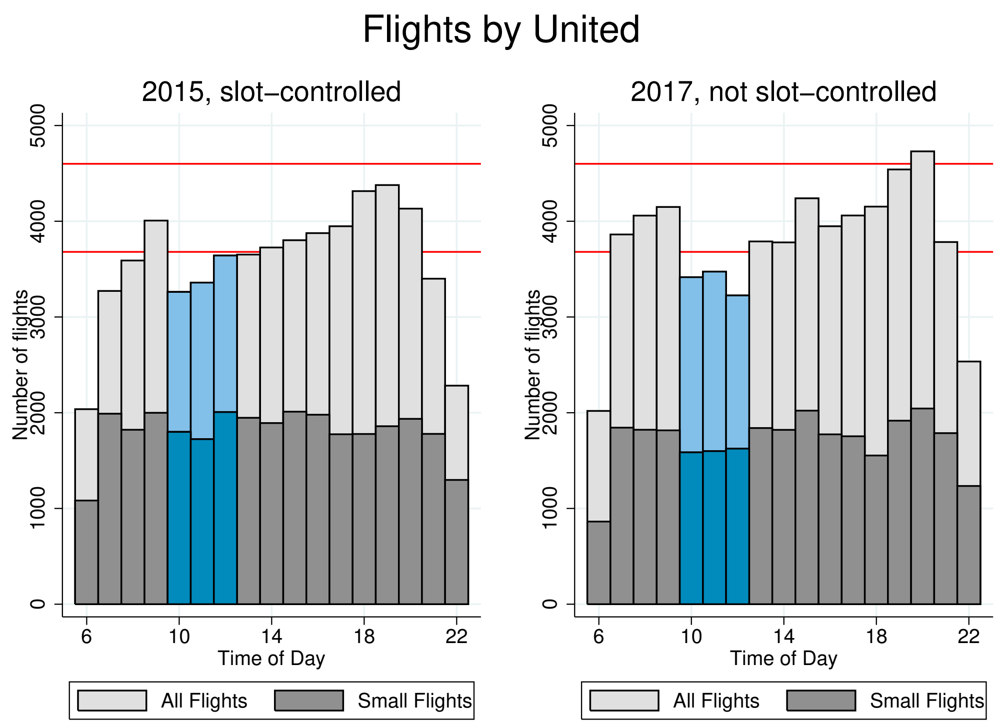
with said constraints. The top red line represents an airline's slot capacity – the average number of slots available to United or the group of low cost carriers in a one-hour slot period during the first quarters of 2015 and 2017.[37] The bottom red line is 80% of the top red line and represents the level of usage required to satisfy the use-it-or-lose-it rule.

These graphs should be interpreted cautiously, and together with other suggestive evidence of slot burning, for two reasons. First, we do not have data on each airlines' hourly slot holdings, which may vary significantly from one hour to another.[38] Second, we only observe actual, as opposed to scheduled, departure and arrival times. Both departing and arriving flights experience significant delays, so mapping of a flight into a slot period based on time of departure or arrival could be inaccurate. These factors contribute to occasional non-compliance with the minimum usage requirements or excess of flights over the slot capacity. However, it is clear that airlines under slot controls do not use *all* their slots *all* the time, indicating slack; this slack

[37] For United, 869 slots $\cdot$ 90 days/17 hours $= 4,600$ possible flights. For low-cost carriers (Alaska, Jet Blue, Southwest, and Virgin), 92 slots $\cdot$ 90 days/17 hours $\approx$ 487 possible flights.

[38] We submitted a FOIA request to PANYNJ to get these data.

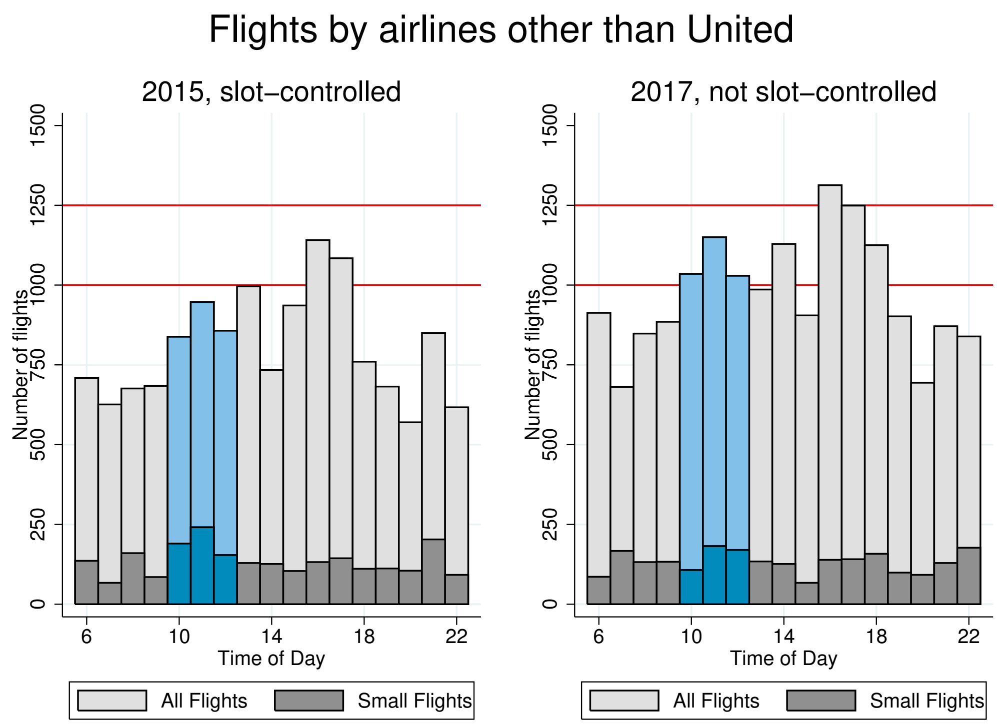
FIGURE 2.4 – UNITED'S COMPLIANCE WITH SLOT AND MINIMUM USAGE CONSTRAINTS



is not due to lack of demand, shown by the increase in frequency following the abolition of slot controls. We provide evidence that this trend is explained by slot burning. In particular, we focus on usage of small aircraft during offpeak slot periods (dark blue bars in Figures 2.4 and 2.5) that seems to decrease when slot controls are removed in 2017, contrary to the overall increase in flight frequencies.

In section 2.5 we refine our descriptive analysis of slot burning, formulate testable hypotheses, and bring them to data. We also investigate the rationing effects of slot control on consumers by looking at the effect of slot restrictions on the number of seats, seat-miles, number of passengers, passenger-miles, and airfare. Since a stated benefit of slot controls is reduced flight delays due to air traffic congestion, we also investigate whether lifting slot restrictions at Newark Airport increases the likelihood of flight delays.

FIGURE 2.5 – OTHER AIRLINES' COMPLIANCE WITH SLOT AND MINIMUM USAGE CONSTRAINTS



2.3.3 *Data and Descriptive Statistics*

This paper uses data from three data sources. The main dataset was obtained from the Port Authority of New York and New Jersey (PANYNJ) using FOIA requests and contains information on exact date and time of arrival and departure of all flights to/from Newark Airport, as well as their operating carriers and aircraft types for the 2015-2017 period. We supplement these data with information on the number of passengers and seats for all domestic origin and destination airports at the monthly level from the Air Carrier Statistics (T-100) database by the Bureau of Transportation. Lastly, we use the Airline Origin and Destination Survey (DB1B) from the Bureau of Transportation for the information on ticketing carriers and airfares for a 10% sample of all tickets issued for all domestic itineraries.

Table 2.2 presents summary statistics of operations at Newark Airport pre- and post-introduction of slot control in 2008 and pre- and post-removal of slot control in 2016. In our empirical analysis, we use the first quarters of each year to control for seasonalities in demand for air travel and scheduling of flights.

TABLE 2.2 – DESCRIPTIVE STATISTICS OF OPERATIONS AT NEWARK AIRPORT IN 2007-2017

Variable	2007		2009		2015		2017	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Small	0.18	0.39	0.26	0.44	0.25	0.43	0.15	0.36
Legacy	0.95	0.21	0.94	0.24	0.81	0.39	0.81	0.39
Delay	35.73	5.16	32.89	5.51	28.19	3.67	31.37	4.54
Load factor	0.78	0.12	0.78	0.11	0.85	0.07	0.87	0.07
Tourist	0.15	0.35	0.14	0.35	0.14	0.36	0.15	0.36
Distance, mi								
< 325	0.20	0.40	0.13	0.33	0.13	0.34	0.12	0.32
325 to 602	0.12	0.32	0.14	0.34	0.11	0.31	0.07	0.26
603 to 998	0.31	0.46	0.30	0.46	0.27	0.44	0.25	0.43
> 998	0.37	0.48	0.43	0.50	0.49	0.50	0.55	0.50
Frequency, daily								
< 6	0.56	0.50	0.81	0.39	0.70	0.46	0.53	0.50
between 6 and 8	0.35	0.48	0.19	0.39	0.22	0.41	0.17	0.38
> 8	0.09	0.29	–	–	0.08	0.28	0.30	0.46
No. routes	24		23		24		24	
No. carriers	17		19		15		16	
No. carrier-routes	59		61		72		75	
No. passengers, mil	1,239.92		1,468.53		1,708.00		2,130.09	
No. flights	17,105,706		12,621,843		12,624,550		16,901,660	

As discussed above, slot burning is best evidenced by more frequent usage of small aircraft in relatively less demanded slot periods when an airport is slot-controlled. Unfortunately, we cannot use rigidly defined one-hour slot periods for our analysis to see which slot periods may experience slot burning. Our data provide information on actual arrival and departure times, as opposed to scheduled ones. Given that both departing and arriving flights experience significant delays, the mapping from time of departure or arrival to a slot period is not straightforward. Therefore, in our analysis, we aggregate slot periods into peak and offpeak categories. We define offpeak periods to be between 10am and 1pm, where the demand is sufficiently high to support some flights (unlike the 1am-4am timeframe), but not enough to warrant usage of multiple larger aircraft (unlike a busy timeframe like 7pm-10pm). We refer to any slot-controlled period that is not offpeak as peak. Our results are robust to changing the definition of offpeak periods to 10am-2pm and 11am-2pm.

One could argue that using small aircraft is needed to meet demand for less dense routes (like, for instance, Newark Airport to Martha's Vineyard, MA), and therefore, using small aircraft for these routes should not constitute slot burning. Such justification is harder to accept for dense routes. Our analysis, therefore, is limited to the 28 largest airports in the US by domestic passenger enplanements.[39] Expanding the analysis to all airports (with a dummy variable for the 28 largest airports, which returns a negative coefficient, corroborating the argument described above) yields no meaningful difference in the results to the variables of interest.

We expect airlines that hold large pools of slots at an airport to be more likely to engage in slot burning. We refer to such airlines as slot incumbents. In Newark Airport, slot incumbents are United and regional airlines that operate flights ticketed by United and using United's slots (see Table 2.1). We incorporate the slot incumbent dummy into our regression analysis in order to control for differential slot-burning incentives of Newark's airlines.

2.4 Theoretical Model

In this section we develop a simple theoretical model to inform our empirical analysis and give a formal definition of slot burning.

2.4.1 Consumers

For simplicity, assume that consumers purchase only direct flights operated by airline j between a given origin o and destination d airport pair. A

[39] See Appendix A for the list of the top-28 domestic airports.

consumer i chooses a flight $jodtk$ between airports o and d during a slot period t on an aircraft of capacity k[40] in order to maximize their utility, u_{ijodtk}, given by

$$u_{ijodtk} = x_{jodtk}\beta - \alpha p_{jodtk} + \beta \sum_k f_{jodtk} + \xi_{jodtk} + \epsilon_{ijodtk},$$

where x_{jodtk} is a vector of product characteristics, p_{jodtk} is the product airfare, f_{jodtk} is the number of flights available on the service route, ξ_{jodtk} is the unobserved product characteristics, and ϵ_{ijodtk} is the i.i.d. logit error term.

A consumer i chooses product $jodtk$ if their utility from $jodtk$ exceeds the utility from any other product, including the outside good. This results in the standard logit share equations, which we use in the airlines' profit-maximization problems.

2.4.2 Airlines

We assume that airlines compete by playing a one-shot game, where they simultaneously choose prices $\mathbf{P}$ and frequencies $\mathbf{F}$ of flights for a given aircraft size k and slot period t[41], subject to the aircraft capacity constraints and slot and use-it-or-lose-it constraints at slot-controlled airports. For simplicity, assume that there are only two aircraft sizes – small and large – so that $k \in \{s, \ell\}$ is a discrete variable that defines capacities of small and large flights. Empirically, we observe that some airlines operate both small and large flights on the same route, e.g. Jet Blue and United, therefore we must allow for this in our theoretical model. To clarify, if an airline does not operate a small flight on route od at slot period t, then $f_{jodts} = 0$.

Let $\boldsymbol{p}_j, \boldsymbol{f}_j$ define the choice variables of airline j for each service route

[40] We differentiate products by aircraft capacity in order to incorporate capacity choice into the airlines' profit-maximization problems. Consumers may not take the aircraft capacity into account when purchasing flights, so we sum up frequencies of flights between an airport pair over all possible capacities.

[41] As a reminder, a slot period at Newark Airport is defined to be a 30-minute window on a particular day of week during a scheduling season, e.g. 6:00-6:30 am on Mondays during the winter season of 2015.

$odtk$ in its set of products $\Omega_j = O_j \times O_j \times T \times K$, where O_j is the set of airports that airline j operates in, T is the set of slot periods, and K is the set of aircraft capacities. Airline j's profit-maximization problem is thus

$$\max_{\boldsymbol{p}_j, \boldsymbol{f}_j} \pi_j \left(\boldsymbol{p}_j, \boldsymbol{f}_j, \boldsymbol{P}_{-j}, \boldsymbol{F}_{-j}\right) = \sum_{odtk} Ms_{jodtk} \left(\boldsymbol{P}, \boldsymbol{F}\right) p_{jodtk} - f_{jodtk} MC_{jodtk} - FC_{jodtk}$$

subject to

$$k f_{jodtk} \geq Ms_{jodtk} \left(\boldsymbol{P}, \boldsymbol{F}\right) \text{ for all } odtk \in \Omega_j \tag{1}$$

$$\sum_{dk} f_{jodtk} + \sum_{dk} f_{jdotk} \leq S_{jot} \text{ for all } ot \in O \times T \tag{2}$$

$$\sum_{dk} f_{jodtk} + \sum_{dk} f_{jdotk} \geq 0.8 S_{jot} \text{ for all } ot \in O \times T \tag{3}$$

$$f_{jodtk} \geq 0 \text{ for all } odtk \in \Omega_j \tag{4}$$

Inequality 1 describes the aircraft capacity constraint a for each product $odtk$ in Ω_j[42], 2 is the slot constraint with $S_{jot} = \infty$ for $ot \in O \times T$ if airport o is not slot-controlled in period t, 3 is the use-it-or-lose-it (UIOLI) constraint, and 4 is the non-negativity constraint on the frequency of flights.

The Lagrangian is then

$$\mathcal{L} = \sum_{odtk} Ms_{jodtk} \left(\boldsymbol{P}, \boldsymbol{F}\right) p_{jodtk} - f_{jodtk} MC_{jodtk} - FC_{jodtk} - \gamma_{jodtk} (Ms_{jodtk}(\boldsymbol{P}, \boldsymbol{F}) - k f_{jodtk})$$
$$+ \lambda^s_{jot} \left(S_{jot} - \sum_{dk} f_{jodtk} + \sum_{dk} f_{jdotk} \right) - \lambda^u_{jot} \left(0.8 S_{jot} - \sum_{dk} f_{jodtk} + \sum_{dk} f_{jdotk} \right)$$

and the FOCs simplify to

$$s_{jodtk} + \sum_{\widetilde{odtk} \in \Omega_j} \frac{\partial s_{j\widetilde{odtk}}}{\partial p_{jodtk}} \left(p_{j\widetilde{odtk}} - \gamma_{j\widetilde{odtk}} \right) = 0$$

$$M \sum_{\widetilde{odtk} \in \Omega_j} \frac{\partial s_{j\widetilde{odtk}}}{\partial f_{jodtk}} \left(p_{j\widetilde{odtk}} - \gamma_{j\widetilde{odtk}} \right) - MC_{jodtk} + \gamma_{jodtk} k + \lambda^u_{jot} - \lambda^s_{jot} = 0$$

[42] We abstract away from potential fleet constraints. It is not clear how to model them since legacy airlines are known to hire regional airlines to operate flights ticketed through their booking systems.

We distinguish two types of slot periods – peak and offpeak periods. Peak periods are periods during which the slot constraint is binding, so $\lambda^s_{jot} > 0$, and the UIOLI constraint is automatically satisfied, so $\lambda^u_{jot} = 0$. In other words, airlines do not need to burn slots in order to satisfy the minimum usage requirements during the peak periods. In contrast, offpeak periods are periods during which the slot constraint is slack, so $\lambda^s_{jot} = 0$, and the UIOLI constraint is binding, so $\lambda^u_{jot} > 0$. Therefore, in order to satisfy the minimum usage requirements during offpeak periods, airlines must burn slots.

Definition 1. Airline j burns slots at airport o during a slot period t if the slot constraint is slack and the UIOLI constraint binds, or $\lambda^s_{jot} = 0$ and $\lambda^u_{jot} > 0$.

Our theoretical model predicts different responses of flight frequencies in peak- and offpeak slot periods after a removal of slot control. In conjunction with Figures 2.4 and 2.5, that establish airlines and slot periods with binding slot and UIOLI constraints, we are able to draw corollaries to test for evidence of slot burning using the reduced-form approach.[43]

2.5 Empirical Analysis and Results

This section introduces testable hypotheses from the theoretical model with the corresponding regression specifications and discusses the results.

2.5.1 Frequency

Consider a counterfactual where slot control is removed, i.e. $\lambda^s_{jot} = 0$ and $\lambda^u_{jot} = 0$. We expect flight frequencies to increase during the peak slot periods and decrease during the offpeak periods, holding all else equal.[44] We also expect the

[43]We do not attempt to recover the values of λ^s and λ^u. This endeavor is left for future work.

[44]Removal of slot control rules decreases entry costs for new airlines, in particular, low-cost carriers. Changes in frequencies and composition of flights in response to entry post slot control are a part of the changes due to removal of the slot and UIOLI constraints per se, since removal of slot control eliminates the foreclosure incentive that causes slot burning in the first place.

effect to be more pronounced for legacy airlines because they hold more slots than low-cost carriers. We explore this hypothesis in regression specification (1).

Hypothesis 1. *After slot control removal, the frequency of flights increases in the peak and decreases in the offpeak slot periods, more so for slot incumbents, holding all else fixed.*

Model 1. The model tests the change in flight frequency due to slot-restrictions,, including any differential changes between peak and offpeak hours, by using flights between Newark Airport and the 28 largest airports in the US for the first quarters of 2015 and 2017. The years were chosen to fall on two sides of the 2016 reclassification of Newark Airport from a slot-controlled (Level 3) airport to a schedule-facilitated (Level 2) airport. The proprietary data comes from FOIA requests to the Port Authority of New York and New Jersey (PANYNJ) to discern whether a flight takes place during peak hours or offpeak hours.

$$\begin{aligned} freq(\text{-miles}) = &\beta_1 slot + \beta_2 incumbent + \beta_3 slot \times incumbent + \beta_4 offpeak \\ &+ \beta_5 offpeak \times incumbent + \beta_6 offpeak \times slot \\ &+ \beta_7 offpeak \times incumbent \times slot + \beta_i(controls) \end{aligned}$$

We include fixed effects for distance (binned at less than 325 miles, between 325 and 602 miles, between 603 and 998 miles, and more than 998 miles), following GAO (2012) specifications, and airport fixed-effects for the non-Newark airport in the origin-destination pair. We use data from the first quarters of 2015 and 2017 to control for seasonality.

Discussion, Columns 1. First, we compare the number of flights before and after slot-controls by carrier incumbency and time of day (peak or offpeak). Following the removal of slot-controls, the incumbent increases the number of flights during peak hours (*p=0.0000*), but operates roughly the same number of flights during offpeak hours (*p=0.2102*). Thus, our prediction for peak hours is validated, but cannot be confirmed for offpeak hours. Controlling for distance and airports, non-incumbents offer more flights during both peak and offpeak hours when Newark Airport is slot-controlled.

Next, we compare the number of flights during peak and offpeak periods by carrier incumbency and slot regime. We observe that non-slot-incumbents fly more flights during peak hours than offpeak hours when Newark Airport is not slot-controlled (*p=0.0000*). However, they fly roughly the same number of flights (*p=0.4513*) when Newark Airport becomes slot-controlled. This conduct can be evidence of slot burning, or explained by the fact that the slot constraint binds during the peak hours, forcing the airlines to reallocate their flights to offpeak hours.

We also observe slot incumbents have more flights during peak hours than offpeak hours when Newark Airport is not slot-controlled (*p=0.0255*). However, the incumbents fly *more* flights during offpeak hours than during peak hours when Newark Airport becomes slot-controlled (*p=0.0030*). A binding slot constraint during peak hours cannot explain why they operate *more* flights during offpeak hours when Newark Airport is slot-constrained. This conduct is indicative of slot burning.

While the number of flights operated by the incumbent during offpeak period is roughly the same regardless of whether Newark Airport is slot-controlled or not (*p=0.2102*), there is a decline in *frequency-miles* when Newark becomes slot-controlled (Column 1b). This confirms our hypothesis that slot-burning happens along shorter routes.

Lastly, we note that frequency of flights is uncorrelated with measures of quantity (either seats or passengers), which are discussed in Subsection 2.5.5.

2.5.2 Use of Small Flights

Slot burning implies that airlines are operating loss-making flights in the offpeak periods. The best approach to minimizing said loss is to (i) fly a smaller aircraft, (ii) across a shorter distance, (iii) on a route with relatively high demand. In our regression analysis, we restrict our sample to the 28 largest airports by passenger enplanements in the contiguous United States and control for the flight distance in order to hold factors (ii) and (iii) fixed. Therefore, usage of small aircraft at slot- and non-slot-controlled airports can be used as a proxy for slot burning. We explore this hypothesis in regression specification (2). As with Hypothesis 1, we expect the effect to be stronger for slot incumbents.

Hypothesis 2. *Under slot control, usage of small aircraft is more prevalent during the offpeak slot periods, more so for slot incumbents, holding all else fixed.*

Model 2. The model tests if the probability of using small aircraft during peak and offpeak periods for the slot incumbent and non-slot incumbents changes when Newark's slot control is removed. The independent variable is whether Newark Airport is slot-controlled (2015)[45] or not (2017).

$$\begin{aligned}\mathbb{L}_{small} = &\beta_1 slot + \beta_2 incumbent + \beta_3 slot \times incumbent + \beta_4 offpeak \\ &+ \beta_5 offpeak \times incumbent + \beta_6 offpeak \times slot \\ &+ \beta_7 offpeak \times incumbent \times slot + \beta_i(controls)\end{aligned}$$

The dependent variable is an indicator for small aircraft, defined as aircraft carrying 100 passengers or less. This definition is to be consistent with the GAO's (2012) model; we try different definitions of "small" (for instance, aircraft carrying fewer than 81 passengers), to no meaningful change in the coefficients. The $\mathbb{L}$ denotes log-odds.

We include the distance and airport fixed effects as before. Additionally, we use fixed effects for daily flight frequency (binned at less than 6, between 6 and 8, and more than 8 flights), following GAO (2012) specifications.

As a robustness check, we also add data from 2007 and 2009, to encompass before and after Newark Airport transitioned to being a slot-controlled airport. We find that the qualitative results remain the same – that all carriers are more likely to use small aircraft in slot-controlled airports. However, given the severe change in air travel demand following the financial crisis, we decided against including the years 2007-2009 into our main (or any other) specification.

Discussion, Column 2. We find that slot incumbents, under no restrictions, were 20% more likely to use small aircrafts during offpeak hours (than during peak hours), a statistic that jumps to 40% under slot restrictions (*p= 0.0213*).[46] That

[45] Only flights between 6am and 10:59am are slot-controlled at Newark Airport, and are coded as such in the data.

[46] From Model 2, the likelihood of slot incumbent to use small aircraft during offpeak hours when Newark Airport is slot-controlled = (Coefficient on $slot_1 \times incumbent_1 \times offpeak_1$)/(Coefficient on $slot_1 \times incumbent_1 \times offpeak_0$) = 7.772/ 5.527 = 1.406. The same exercise for when incumbents fly without slot restrictions yields 1.216.

the slot incumbent is twice as likely to use small aircrafts during offpeak hours (than peak hours) when Newark Airport becomes slot-controlled is indicative of slot burning. Similarly, non-incumbents exhibit an increased reliance on small aircrafts when Newark Airport is slot-controlled (*p= 0.0012*). The difference between the two odds ratios for all carriers shows that the increased usage of small aircraft during offpeak hours cannot only be explained by the type of consumer demand during offpeak hours. Figure 2.3 in Section 2.3 illustrates that usage of small flights increases during offpeak hours (for example, 10am-1pm) and decreases during peak hours (for example, 7pm to 10pm) when Newark Airport is slot-controlled.

2.5.3 Consecutive Flights

A likely argument in defense of airlines' frequent use of small aircraft is catering to the time-sensitivity of demand. We address this concern by looking at the probability of observing consecutive flights to the same destination within a short timeframe (30 minutes in the baseline specification).

Hypothesis 3. *Under slot control, the probability of observing consecutive flights to the same destination is greater in the offpeak slot periods, more so for slot incumbents, holding all else fixed.*

We explore this hypothesis in regression specification (3a). In regression specification (3b), we further refine our test by examining the probability of observing consecutive *small* flights to the same destination within a short timeframe.

Model 3a. The second model investigates the probability of an airline flying multiple flights on a route within a 30-minute window. The dependent variable indicates whether the same airline offers another flight from the same origin to the same destination within the 30-minute window. The qualitative results are the same if the window is changed to 45- or 60-minutes. The same set of fixed effects are used from the previous model.

$$\begin{aligned}\mathbb{L}_{consecutive} =& \beta_1 slot + \beta_2 incumbent + \beta_3 slot \times incumbent + \beta_4 offpeak \\ &+ \beta_5 offpeak \times incumbent + \beta_6 offpeak \times slot \\ &+ \beta_7 offpeak \times incumbent \times slot + \beta_i(controls)\end{aligned}$$

Model 3b. This model investigates the probability of an airline flying multiple *small* flights on a route within a 30-minute window. This specification helps us narrow down the mechanism by which airlines burn slots (that is, whether airlines are indeed flying multiple small flights during offpeak hours).

$$\begin{aligned}\mathbb{L}_{consec\_small} =& \beta_1 slot + \beta_2 legacy + \beta_3 slot \times legacy + \beta_4 offpeak + \beta_5 offpeak \times legacy \\ &+ \beta_6 offpeak \times slot + \beta_7 offpeak \times legacy \times slot + \beta_i (controls)\end{aligned}$$

Discussion, Columns 3. We find that while non-incumbents do not significantly crowd their flights under slot restrictions, slot incumbents are 11% more likely to do so when operating out of Newark Airport under slot controls than without slot controls (*p= 0.0429*). The definition of a consecutive flight in the main specification is another flight within a 30-minute window; the qualitative results are robust to alternate definitions of consecutive (45- or 60-minutes).

We hypothesized that refining our test by examining if the probability of observing consecutive *small* flights to the same destination will yield similar results. In fact, slot incumbents are 75% more likely to fly consecutive small flights when Newark Airport is slot-controlled (*p=0.0000*); non-incumbent carriers are almost twice as likely to fly consecutive small flights when Newark Airport is slot-controlled (*p=0.0110*). These results suggest that all carriers increase their reliance on small consecutive flights when Newark Airport becomes slot-controlled, which is indicative of slot burning.

TABLE 2.3 – REGRESSION COEFFICIENTS

	(1a) Frequency	(1b) Freq-mile	(2) Small	(3a) Consecutive	(3b) SmallConsec
Model	OLS	OLS	Logistic	Logistic	Logistic
$slot_0 \times incumbent_0 \times offpeak_1$	-0.828***	-245.4*	0.874	0.654**	0.265
	(-9.99)	(-1.97)	(-1.70)	(-3.14)	(-1.79)
$slot_0 \times incumbent_1 \times offpeak_0$	5.577***	8554.0***	9.979***	1.358***	3.192***
	(126.28)	(128.61)	(44.99)	(4.57)	(4.58)
$slot_0 \times incumbent_1 \times offpeak_1$	5.424***	7620.1***	12.14***	0.754**	1.166
	(74.36)	(69.37)	(39.60)	(-3.14)	(0.52)
$slot_1 \times incumbent_0 \times offpeak_0$	1.206***	778.2***	1.392***	1.048	0.782
	(22.71)	(9.73)	(5.81)	(0.63)	(-0.76)
$slot_1 \times incumbent_0 \times offpeak_1$	1.272***	1435.3***	1.804***	0.704*	1.681
	(14.31)	(10.72)	(6.70)	(-2.49)	(1.48)
$slot_1 \times incumbent_1 \times offpeak_0$	5.114***	6672.3***	5.527***	1.430***	2.967***
	(111.81)	(96.88)	(34.24)	(5.28)	(4.27)
$slot_1 \times incumbent_1 \times offpeak_1$	5.310***	6636.2***	7.772***	0.955	3.251***
	(71.71)	(59.50)	(33.37)	(-0.54)	(4.41)
Controls:					
Distance	Yes	Yes	Yes	Yes	Yes
Airport	Yes	Yes	Yes	Yes	Yes
Frequency			Yes	Yes	Yes
N	77,776	77,776	77,776	77,776	21,458

Notes: t statistics in parentheses. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.5.4 *Delay*

Although not from the theoretical model, the rationale provided by the FAA for slot restrictions invokes congestion, leading us to believe that slot controls may alleviate delays.

Hypothesis 4. *Slot control improves airlines' on-time performance.*

Model 4. This model investigates the probability that a flight is delayed by longer than 30 minutes. The data comes from the T-100 database, which contains aggregate information on flight schedules. We use airport-specific dummies to account for the possibility that delays could be caused by congestion at the other endpoint airport.

$$\mathbb{L}_{delay} = \beta_1 slot + \beta_2 incumbent + \beta_3 slot \times incumbent + \beta_i(controls)$$

Discussion, Column 4. The relegation of Newark Airport to a Level 2 airport resulted in a worse on-time performance for Newark Airport. Column 4 shows that while slot incumbents fare slightly worse in terms of on-time performances in 2017 (when Newark Airport was not slot-controlled), all airlines experience less delays in 2015. The probability of a legacy carrier being delayed by 30 minutes or more decreases by about 17%,[47] while non-incumbent carriers are 68% less likely to be delayed in 2015.

TABLE 2.4 – REGRESSION COEFFICIENTS (CONTINUED)

	(4) Delayed	(5a) Seats (thousands)	(5b) Seat-mile (millions)	(6a) Passengers (thousands)	(6b) Pax-mile (millions)	(7) Mktfare ($)
Model	Logistic	OLS	OLS	OLS	OLS	OLS
$slot_0 \times incumbent_1$	1.026	-66.42***	-54.56***	-52.83***	-44.39***	41.00***
	(0.90)	(-22.16)	(-12.80)	(-21.04)	(-12.16)	(79.87)
$slot_1 \times incumbent_0$	0.321***	-12.90	-9.028	-9.293	-8.504	13.70***
	(-32.49)	(-1.73)	(-0.85)	(-1.48)	(-0.93)	(16.98)
$slot_1 \times incumbent_1$	0.850***	-50.82 ***	-36.31***	-39.74***	-30.74**	46.58***
	(-6.02)	(-6.60)	(-3.32)	(-6.16)	(-3.28)	(55.97)
Constant		231.0***	89.69***	194.7***	76.55***	291.4***
		(14.55)	(3.97)	(14.64)	(3.96)	(117.30)
Controls:						
Distance	Yes	Yes	Yes	Yes	Yes	Yes
Airport	Yes	Yes	Yes	Yes	Yes	Yes
Frequency	Yes	Yes	Yes	Yes	Yes	Yes
Region	Yes					Yes
N	71,099	1,989	1,989	1,989	1,989	1,419,871

Notes: t statistics in parentheses. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

[47] 1 - 0.850/1.026 = 0.172.

2.5.5 *Quantity and Price*

We have no testable hypotheses for quantity or price. Changes in price, capacity, and quantity of passengers are all ambiguous and dependent on various intertwined factors. For illustration, assume that under no slot restrictions, an airline flies one 150-passenger flight along a route. Following slot restrictions, the airline might choose to fly two or three 60-passenger flights, depending on load factors and passenger sensitivity to frequency. Without knowing the direction of the change in supply, it is not possible to know, a priori, the direction of the change in price. Questions relating to consumer surplus, therefore, can only be answered by the data.

Models 5-6. The subsequent models estimate the impact of slot restrictions on the number of seats available (Model 5) and the number of passengers transported (6). The T-100 database is used for these estimations, since it comes with aggregate numbers for seats and passengers. Since T-100 database does not include time of flight, thus whether a flight is traveling during peak or offpeak hours cannot be included in the model. Controls include fixed effects for frequency, distance, year, region, and airport.

$$\begin{aligned} seat(\text{-miles}) \text{ or } passenger(\text{-miles}) = \beta_0 + \beta_1 slot + \beta_2 incumbent + \\ \beta_3 slot \times incumbent + \beta_i(controls) \end{aligned}$$

Model 7. The seventh model looks at the effect on price at slot-controlled airports. We use any routes within the top-28 airports for this analysis, with the *slot* dummy indicating whether any of the airports within the route were slot-controlled in the period the flights took place. By restricting the analysis to the first quarters of 2015 and 2017, we can exploit the reclassification of Newark Airport as an exogenous variation of the independent variable, *slot*. Since we use the DB1B database for this analysis, we do not compute any measures for frequency of flights or time of flight within the route. Controls include fixed effects for distance, year, region, and airport.

$$price = \beta_0 + \beta_1 slot + \beta_2 incumbent + \beta_3 slot \times incumbent + \beta_i(controls)$$

Discussion, Columns 5-7. We find an increase in the total number of seats flown by the slot incumbent (Column 5a, *p= 0.0426*) in presence of slot control, but we cannot reject the null hypothesis that the incumbent flies different seat-miles at the 5% level. This shows that while the incumbent is offering more seats, the seats are on shorter routes, which alludes to slot-burning. Non-incumbents do not show any statistically significant change in seats or seat-miles.

Similarly, incumbents fly more passengers (Column 6a, *p=0.0422*) following slot restrictions, but exhibit no change in passenger-miles. Non-incumbents fly the same number of passengers and passenger-miles.

Column 7 outlines the effect of slot restrictions on price. We find that the slot incumbent (non-incumbents) charge about $5 more ($13 more) per passenger for one-way travel when they serve a slot-controlled airport, with the median ticket price being around $220. These two values are statistically significant (*p=0.0000* for both incumbent and non-incumbents), and different from one another (*p= 0.0000*).

2.5.6 Discussion on Consumer Surplus

The effect of slot restrictions on consumer surplus is ambiguous. Holding all else equal, more passengers fly under slot control but on average pay a higher price. Such an effect is indicative of an outward shift in the demand curve, likely suggesting that the product quality has increased. We also find that the total number of seats offered under slot control increases, suggesting that there is an outward shift in the supply curve as well. The fact that we observe an increase in the number of passengers transported together with an increase in airfares implies that the shift in the demand curve is larger than that in the supply curve.

We attribute the increase in product quality to the success of slot control at curbing delays. Ceteris paribus, a flight with poor on-time performance is considered inferior to one that reliably follows schedule. However, as we document in the next section, removing slot control does add new direct routes for certain communities, decreasing layover times and acting as a counterveiling force to the increase in product quality associated with slot control. We rationalize this by the fact that a consumer's disutility from delay may be different from their disutility from a long layover or poor match between their desired and actual departure time, since the delay is an uncertainty only realized at the point of use, not the point of purchase. Moreover, frequency of flights on a route is another determinant of product quality. All else constant, a route with frequent flights provides a better match between a time-sensitive consumer's desired departure time and the actual departure time. A structural model is needed to quantify the relative tradeoffs and we leave this for future work.

Not only is the effect on consumer surplus ambiguous, it is also heterogeneous. Our findings show that slot restrictions increase the number of flights along dense business-routes in short distances (the routes that facilitate slot burning), while decreasing the number of flights to tourist destinations. Therefore, we expect differential impact on consumers flying different routes; consumers flying tourist routes experience a reduction in consumer surplus (higher price, smaller number of seats available, lower quality product), whereas the effect on business destinations are ambiguous (higher price indicates lower surplus, but a higher product quality indicates higher surplus).

2.6 Entry and Exit Following Reclassification

In this section, we analyze airlines' entry and exit decisions at the route-level following the Newark's removal of slot controls. We also study entry and exit in 2019 in order to benchmark the magnitude of the industry shake-out in 2016.[48] As before, we limit our analysis to the first quarters of 2015 and 2017-2019 to overstep potential seasonality issues. However, given our interest in understanding the differential impact of slot control on communities of varying size, and especially in small and rural communities' access to air transport, we extend our sample to all domestic airports of the contiguous United States.

Table 2.5 above summarizes airlines' entry and exit decisions. Only seven airlines were operating in Newark in 2015. The reclassification brought in three new airlines: Allegiant Air (a low-cost carrier based in Nevada), Elite Airways (a brand new airline operating out of Portland, ME), and Spirit. Not all airlines entered

[48] We use the T-100 database from the Bureau of Transportation Statistics. As of now, January 2020 is the latest available data, but even once more data become available, we could not use 2020 due to the impacts of the COVID-19 pandemic on air travel.

TABLE 2.5 – SUMMARY OF AIRLINES' ENTRY AND EXIT DECISIONS IN 2015-2017 AND 2018-2019

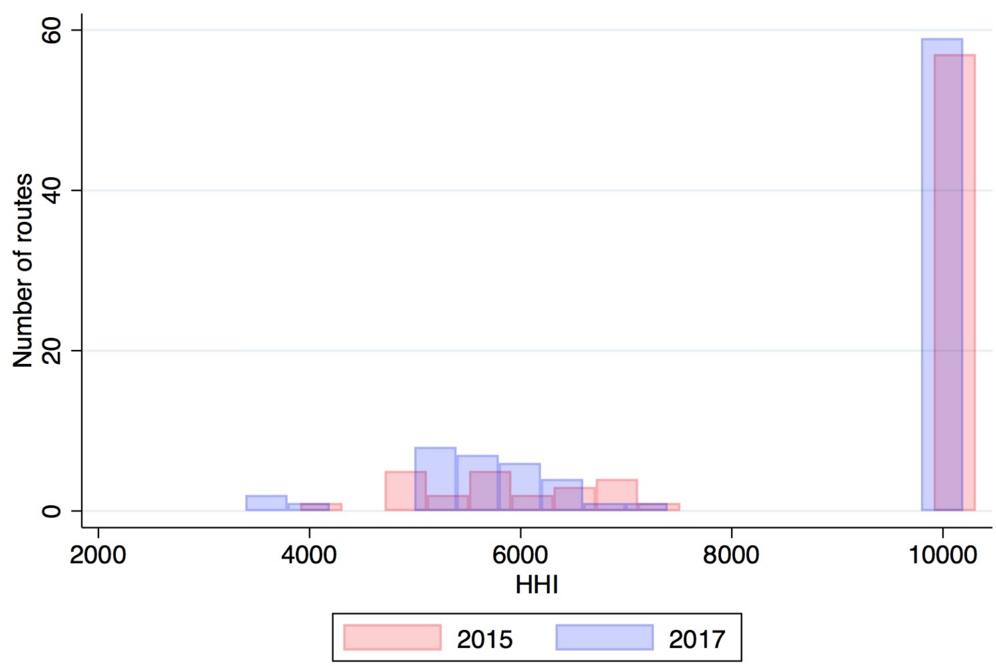
Airline	Number of routes				2017		2019	
	2015	2017	2018	2019	Entry	Exit	Entry	Exit
American	5	6	5	5	1	0	0	0
Allegiant	0	4	1	3	4	0	2	0
Alaska	1	4	4	6	3	0	2	0
Delta	5	5	5	5	0	0	0	0
JetBlue	6	6	7	7	0	0	0	0
Elite	0	1	1	0	1	0	0	1
Spirit	0	3	6	8	3	0	2	0
Southwest	8	7	8	9	2	3	3	2
United	77	83	83	77	9	3	4	10
Virgin	2	2	2	0	0	0	0	2
Total	104	121	122	120	23	6	13	15

new markets; Delta, JetBlue, and Virgin did not change their routes at all. Of the seven existing airlines only two, Southwest and United, dropped routes. Overall, Newark Airport saw 23 entry events on 19 routes and six exits on six routes in 2017. Between 2018 and 2019, the latest available years with no change in slot regime and unrestricted entry, Newark experienced significantly less entry (13 entry events on 13 routes) and more exit (15 exits on 15 routes).

Tables 2.6 and 2.7 below provide detailed information on routes that experienced entry (encoded as 1) and exit (-1) following the reclassification in 2016, and in 2019. Removal of slot control resulted in entry on 19 routes, seven of which are brand new (highlighted in bold in Table 2.6), with six of them due to United. Interestingly, low-cost carriers did not start new routes; instead, they entered mid-sized (Alcoa, TN; Asheville, NC; Hebron, KY; Savannah, GA; Myrtle Beach, SC) and West Coast (Portland, OR; San Diego, CA; San Jose, CA) airports, challenging United on those routes. Another group of airports that experienced entry is tourist destinations in Florida. In this case, an entry by a low-cost carrier was matched by another low-cost carrier and challenged United and Jet Blue, resulting in a four-firm oligopoly (Fort Lauderdale, FL and Orlando, FL).

Airlines exited from six markets in 2017, stopping operations on three routes (Binghamton, NY; Houston, TX; Lake City, FL). Out of all exits, only one occurred on a route from the slot-burning sample – Houston, TX by Southwest. Therefore, we can conclude that the slot incumbent did not operate entirely unprofitable routes just for the sake of burning slots. Moreover, Table 2.6 is suggestive for refuting anecdotal claims that the FAA may be tolerating slot burning if slots are burned on routes providing access to air transport for small and rural communities. Slot controls alone do not appear to create incentives for airlines to operate flights to small airports as evidenced by lack of mass exit from small destinations. Generally, the fact that the reclassification resulted in entry into a variety of destinations, and exit from a handful of destinations, implies that a heterogeneous group of consumers have benefited from the change in slot regime.

FIGURE 2.6 – DISTRIBUTION OF HHI ON ROUTES FROM/TO NEWARK IN 2015 AND 2017



All in all, changes in frequencies of flights on the extensive and the intensive margin contributed to a slight leftward shift of the HHI distribution. See Figure 2.6. The average HHI in 2015 is around 8,211 relative to 8,427 in 2017. However, this seeming increase in concentration is due to United opening six new monopoly routes. Conditional on existence of routes in 2015, the average HHI declines from 7,903 in 2015 to 5,291 in 2017.

Tracking entry patterns over time, we document that a recent entrant, Elite Airways, ceased operations in Newark by 2019, so did Virgin by dropping the Los Angeles and San Francisco routes that were entered by Alaska the same year. United's exit from eight markets (Baltimore, MD; Chattanooga, TN; Des Moines, IA; Fort Wayne, IN; Ithaca, NY; South Bend, IN; Windsor Locks, CT) completely stopped operations on these routes; two of those routes (Chattanooga, TN and Fort Wayne, IN) were entered in 2017, after removal of slot control. Additionally, United exited another two routes that experienced entry in 2017 – Myrtle Beach, SC by Spirit and San Jose, CA by Alaska. The fact that many of the routes with entry in 2017 experienced exit in 2019 implies that the industry was still underway to the long-run equilibrium. The average HHI in 2018 and 2019 were 7,398 and 7,419, respectively.

TABLE 2.6 – ALL ENTRY AND EXIT DECISIONS BY AIRLINE BETWEEN 2015 AND 2017

	Alaska	Allegiant	American	Elite	Southwest	Spirit	United	No. carriers, 2017
Akron, OH	–	–	–	–	–	–	1	1
Alcoa, TN	–	1	–	–	–	–	–	2
Asheville, NC	–	1	–	–	–	–	1	2
Binghamton, NY	–	–	–	–	–	–	-1	0
Chattanooga, TN	–	–	–	–	–	–	1	1
Chicago, IL	–	–	1	–	–	–	–	2
Flint, MI	–	–	–	–	–	–	1	1
Fort Drum, NY[a]	–	–	–	–	–	–	-1	1
Fort Lauderdale, FL	–	–	–	–	1	1	–	4
Fort Wayne, IN	–	–	–	–	–	–	1	1
Hebron, KY	–	1	–	–	–	–	–	3
Houston, TX	–	–	–	–	-1	–	–	0
Kenner, LA	–	–	–	–	-1	–	–	1
Key West, FL	–	–	–	–	–	–	1	1
Lake City, FL	–	–	–	–	–	–	-1	0
Lexington, KY	–	–	–	–	–	–	1	1
Myrtle Beach, SC	–	–	–	–	–	1	–	2
Nashville, TN	–	–	–	–	-1	–	–	1
Orlando, FL	–	–	–	–	1	1	–	4
Portland, OR	1	–	–	–	–	–	–	2
Salt Lake City, UT	–	–	–	–	–	–	1	2
San Diego, CA	1	–	–	–	–	–	–	2
San Jose, CA	1	–	–	–	–	–	1	2
Savannah, GA	–	1	–	–	–	–	–	2
Vero Beach, FL	–	–	–	1	–	–	–	1
No. mkts, 2015	1	0	5	0	8	0	77	
No. mkts entered | exited	3 | 0	4 | 0	1 | 0	1 | 0	2 | 3	3 | 0	9 | 3	

[a]Military airfield.

TABLE 2.7 – ALL ENTRY AND EXIT DECISIONS BY AIRLINE BETWEEN 2018 AND 2019.

	Alaska	Allegiant	Elite	Southwest	Spirit	United	Virgin	No. carriers, 2019
Alcoa, TN	–	1	–	–	–	–	–	2
Asheville, NC	–	1	–	–	–	–	–	2
Atlanta, GA	–	–	–	–	1	–	–	3
Avoca, PA	–	–	–	–	–	-1	–	0
Baltimore, MD	–	–	–	–	–	-1	–	0
Chattanooga, TN	–	–	–	–	–	-1	–	0
Des Moines, IA	–	–	–	–	–	-1	–	0
Fort Wayne, IN	–	–	–	–	–	-1	–	0
Horseheads, NY	–	–	–	–	–	1	–	1
Indianapolis, IN	–	–	–	-1	–	–	–	1
Ithaca, NY	–	–	–	–	–	-1	–	0
Los Angeles, CA	1	–	–	–	–	–	-1	2
Montrose, CO	–	–	–	–	–	1	–	1
Myrtle Beach, SC	–	–	–	–	–	-1	–	1
Nashville, TN	–	–	–	1	–	–	–	2
Oakland, CA	–	–	–	1	–	–	–	1
Orlando, FL	–	–	–	-1	–	–	–	3
Palm Springs, CA	–	–	–	–	–	1	–	1
Presque Isle, ME	–	–	–	–	–	1	–	1
San Diego, CA	–	–	–	1	–	–	–	3
San Francisco, CA	1	–	–	–	–	–	-1	2
San Jose, CA	–	–	–	–	–	-1	–	1
South Bend, IN	–	–	–	–	–	-1	–	0
Tampa, FL	–	–	–	–	1	–	–	3
Vero Beach, FL	–	–	-1	–	–	–	–	0
Windsor Locks, CT	–	–	–	–	–	-1	–	0
No. mkts, 2018	4	1	1	8	6	83	2	
No. mkts entered | exited	2 | 0	2 | 0	0 | 1	3 | 2	2 | 0	4 | 10	0 | 2	

It is possible that the exits in 2019 are delayed decisions due to the change in slot rules in 2016. If it were the case, then the shutdown of some routes altogether is a concerning effect of the slot liberalization. However, the number of routes shut down is still small, and all the towns losing direct service to/from Newark Airport have (i) connecting flights to Newark Airport, and (ii) are within 100 miles of another airport with direct service to Newark Airport. All of this suggests that even if the 2019 exits were due to the 2016 change in slot rules, its effects are minimal.

Comparing the 2015 and 2019 figures, we can conclude that removal of slot control at Newark brought in a competitive low-cost carrier, Spirit, and 16 additional carrier-routes thereby decreasing concentration by nearly 10%. Thus, we can conclude that removal of slot control was favorable to promoting competition at the Newark Airport.

2.7 Conclusion

In this study, we show that firms respond to slot restrictions by using smaller flights to use their allocated slots in order to meet the usage requirements. Eliminating such restrictions results in entry, primarily from newly formed low-cost carriers with no historic footprint at the airport. This entry can lower prices, but also result in flight delays due to congestion. However, more passengers fly when Newark Airport is slot-controlled (along routes used to burn slots), implying an ambiguous change in consumer surplus following slot liberalization (lower price, fewer passengers, more delays).

However, since low-cost entrants offer a different assortment of products than the incumbent (almost always a legacy carrier), any change in the relative balance between the two will have differential impact on passengers flying routes dominated by legacy or low-cost carriers. Due to this heterogenous effect on consumers, policy decisions on slot restrictions to manage congestion at airports must be balanced with an eye on the foreclosure incentive by airlines, and subsequent changes in consumer welfare due to changes in product quality (frequency on a route) and the price paid by passengers.

Chapter 3

Political Trends in Minimum Wage Policy Evaluation Studies

with Andrew Copland and Jean-Fran¸ois Gauthier

3.1 Introduction

Research into the impacts of minimum wage changes has been a growing field in economic research for over a century. Initial work was largely funded and conducted by United States Government economists, examining the implications of minimum wages for federal policy reasons.[49] Over time, the mantle has shifted to academic researchers. Though interest in the topic waxes and wanes with broader economic trends, few subjects inspire the passion and heated debates, among the general population and experts alike, than the impact of the minimum wage on labor market outcomes.

The most recent wave of interest in the minimum wage effect on employment in the low-wage sector has been spurred by the publication of two papers by Dube et al. (2010, hereafter DLR) and Allegretto et al. (2011, hereafter ADR) that challenged the previous consensus in the literature, summarized in Neumark and Wascher (2008*b*, hereafter NW), that the elasticity of low-wage workers' employment with respect to minimum wage is negative, albeit small. DLR and ADR argued that previous studies, which relied on a more traditional panel-data state-level approach, failed to account for local economic conditions responsible for different employment patterns and produced spurious negative elasticities. Instead, DLR and ADR proposed to identify areas with comparable local economic conditions that have different minimum wages due to differences in state-level poli-

[49]The most noteworthy example is the analysis conducted by the Minimum Wage Study Commission between 1977 and 1981, which itself claims to have "conducted the most exhaustive inquiry ever undertaken into the issues surrounding [the Fair Labor Standards Act of 1938] since its inception" (See Commission (1981)).

cies. In particular, focusing on bordering counties that belong to states with different minimum wage levels establishes relevant comparison groups. With that approach, DLR and ADR found that the disemployment effects of minimum wage are close to zero and statistically insignificant. In a follow-up work, NW criticized ADR and DLR for obtaining the insignificant estimate by restricting the sample so much that the identifying variation was compromised. The debate continues to-date, with the most recent working paper by Neumark criticizing the DLR and ADR methodology on the grounds of bordering counties not sharing local economic shocks to the same extent multi-state commuting zones do (Jha et al., 2022).[50]

The divergent predictions in the literature are due to difficulties in establishing adequate counterfactuals for treated units. Both Neumark and Wascher (1992, 2008a, 2008b) and Allegretto et al. (2011) and Dube et al. (2010) rely on difference-in-differences (DiD) approach to identify causal effects of minimum wage changes. The DiD method requires that treated and control units evolve similarly prior to the policy intervention (the so-called "parallel trends" assumption) and that the policy change of interest is not confounded by any other contemporaneous policy change. In particular, NW define the treatment group to include states where the state minimum wage level is above the federal level and the control group to include states where the federal minimum wage is binding. ADR and DLR use geographic proximity of counties that belong to states with different minimum wage levels to establish the control and treatment groups. At the same time, minimum wage changes are political decisions and are more likely to be adopted in areas where political beliefs of the population support such policies. We argue that NW and ADLR differences in employment elasticity estimates could partly be due to failure to account for political trends that can underlie endogenous adoption of minimum wage changes in each state, thereby violating the parallel trends and lack of confounding policies assumptions.

More specifically, even if states with and without binding state minimum wage levels and geographically proximate counties are sufficiently similar along dimensions clearly and immediately related to the dependent variable (such as trends and levels in teen unemployment, as observed and debated in Neumark et al. (2014*b*)

[50]Jha et al. (2022) are aware that a minimum wage change in one of the states in the multi-state commuting zone is likely to spill over within the entire commuting zone. Yet, they argue that their prediction of negative disemployment effects would still survive.

and Allegretto et al. (2017)), it may not be enough to satisfy the conditions required by the statistical model. Geographic proximity might be a good start to identifying "sufficiently similar" county pairs that can then be analyzed, but this should be viewed as a necessary, not sufficient condition for county similarity.[51] For reasons outlined above, political alignment may well be as important to economic shock responses as geographic location.

Moreover, neither of the NW or ADLR studies considers effects of potentially confounding policies that may be implemented by state governments that are inclined to change minimum wage laws. Specifically, this assumption, as defined in Dube et al. (2010), but with very slightly different notation, is $\mathbb{E}\left[\ln\left(w_{it}^{M}\right), ''_{ipt}\right] = 0$, that no residual change in employment effects can be correlated with changes in the minimum wage. While this may hold in many cases, and be weakly satisfied in most, we believe there are some instances and particular sources of residual employment effects that are likely cause for concern. Given minimum wage changes are political decisions, it seems minimally necessary to consider other legislative changes that may also lead to employment effects. For example, if liberal state governments tend to be more likely to both raise the minimum wage and increase discretionary state spending, the fundamental identifying assumption critical to NW and ADLR's findings would fail to hold. While these correlations will not always be sufficiently strong to overturn the results found when ignoring them, assuming they do not exist may well lead to spurious conclusions. Accounting for these other policies, and critically examining the relevant dimensions for these correlations, is a necessary first step that has gone relatively unexamined in the current body of literature.

In this paper, we extend the contiguous-county-across-state-borders methodology of DLR (2010) by explicitly controlling for endogenous political trends that may violate the DiD identification assumptions. First, we use presidential election

[51] There are likely cases where even geographic proximity may not be enough to suggest similarity in physical conditions experienced by neighboring counties. State borders are often non-randomly drawn, and lie along prominent physical features of the terrain. Well known geological phenomena, then, can easily lead to particular county-pairs experiencing consistently different "physical shocks." One prominent example of this is the impact of mountains on temperature and rainfall: a well-known result for residents who live along the North Carolina-Tennessee border where the Tennessee side frequently receives substantially more rain than the corresponding counties just across the ridge-line in North Carolina.

voting results to construct a sample of politically aligned county pairs. Second, we use state spending in education, public welfare, hospitals, health and capital outlay that capture state governments' funding for contemporaneous policies that may have residual impact on employment. Our central result is that DLR's estimates are sensitive to the inclusion of these additional covariates. In particular, we find that the employment elasticity is between -0.12 and -0.18 and statistically significant. Our result is closer to the NW estimates of between -0.1 and -0.3.

First, we use presidential election voting results at the county level between 1988 and 2016 to verify whether contiguous counties are indeed similar and if political similarities are time-invariant. Contiguous counties from adjoined states that have diverging political views are very likely to have significant differences in unobservables, both in levels and in changes. Thus, including them as a treatment-control pair is likely to violate the parallel trend assumption necessary for DLR's identification since impacts of unobservable shocks may not be netted out on average. Then, we use DLR's replication package data on minimum wage changes and restaurant employment and extend their results by using only county pairs that consistently remain similar in terms of political views. Comparing our restricted sample results against the unrestricted county pairs sample allows us to assess the bias introduced when using dissimilar control counties. We also account for other factors likely to be correlated with residual employment and earnings by incorporating changes in state spending. We multiply change in per capita state spending per year by each county's population to generate a rough measure of discretionary spending, which is then included as a regressor in the full model.

Our first finding is that the elasticity of employment with respect to minimum wage is estimated to lie between -0.12 and -0.19 and is statistically significant at 10%. This result is closer to the one found in the traditional literature, represented by Neumark and Wascher. Even though our findings run contrary to the zero effect estimated by DLR, we do not advocate for NW's approach. For instance, NW criticize DLR for excessively constraining the sample of observations (by focusing on contiguous county pairs) to the point that they lose useful identifying variation and estimate insignificant, close to zero effects because of that. We find that despite dropping roughly 60% of DLR's sample observations by focusing on politically aligned counties, standard errors of the estimates of minimum wage effects decrease as a result of truncating the sample.

We also show that politically unaligned counties predict a positive elasticity of employment with respect to minimum wage, which is difficult to rationalize within the standard competitive model of minimum wage labor market. Therefore, the zero effect of minimum wage on employment found in DLR (2010) likely averages out different estimates for politically aligned and unaligned counties. Consistently with that, we find that inclusion of state spending controls does not change DLR's results in a meaningful way—the majority of differences between the two estimates comes from political alignment.

All of these results taken together suggest that accounting for underlying political trends may be important in policy evaluation studies that focus on policies that are implemented as a result of political vote. Although we focus on the minimum wage debate in this paper, we are concerned that a similar issue might permeate other studies in different fields that utilize the county-pairs methodology for identification. Several papers over the past two decades or so have relied on the underlying similarity assumption to motivate identification.[52] We believe that the necessary assumption of no correlation between the variable of interest and the error term may not always hold because of other, unobserved policy changes enacted by different states at different times.

Moreover, to the extent that contiguous-county DiD implicitly assigns the weight of one to geographic proximity to establish a valid control group, it could be viewed as a special case of synthetic controls (SC). In cases where high-frequency data on outcome variables is not readily available, Dube and Zipperer (2015) point out the importance of using relevant predictors for reconstructing counterfactual trends for treated units. Our controls for underlying political trends may be used as relevant predictors for SC methods as well. We leave a more rigorous examination of this question for future research.

[52]Several examples across fields are: Huang (2008) (impact of bank deregulation), Naidu (2012) (mentioned in the previous section), Agrawal and Hoyt (2014) (tax policy's impact on commuting times), Bohn and Santillano (2017) (the impact of "locally enforced immigration programs" on employment in "immigrant intensive industries"), Kahn and Mansur (2013) (electricity prices and unionized labor's effect on the clustering of manufacturing industries in states), Rohlin and Thompson (2013) (the impact of sales taxes on employment at state borders), and others.

The rest of the paper is organized as follows. Section 3.2 reviews the relevant literature on minimum wage, contiguous-counties across state borders DiD, and importance of political preferences and outcomes for identification of treatment effects in policy evaluation studies. Section 3.3 introduces our proposed controls for underlying political trends. Sections 3.4 and 3.5 describe our data sources and summary statistics. Sections 3.6 and 3.7 present our empirical strategy and results. Our current results are preliminary, so we discuss our plans for future work in Section 3.8. Finally, Section 3.9 concludes.

3.2 Related Literature

The motivating use-case of difference-in-difference estimation–paired contiguous county designs from Dube et al. (2010)–has expanded in applications to incorporate a wide variety of policy interventions. Using contiguous counties separated by state borders has been applied to, among many others, estimating the effects of banking regulations (Huang (2008)), disenfranchisement of Black Americans in the antebellum south (Naidu (2012)), and the impact of taxes on commuting times (Agrawal and Hoyt (2014)). While our note is broadly relevant for difference-in-difference and synthetic control estimation in policy-related circumstances, it resonates most clearly for estimation that uses geographical boundaries to delineate treatment and control.

3.2.1 Minimum Wage Research

First and foremost, this paper is both rooted in, and inspired by, the decades-old minimum wage research program. Methodologically, these papers have generally utilized either large panel data to isolate minimum wage changes, or identified specific instances of plausible natural experiments (where otherwise similar geographic areas diverge in their implementation of a minimum wage).[53]

[53]For a thorough review of the minimum wage literature up through the mid 2000s, see Neumark and Wascher (2007).

3.2.1.1 *Early Approaches and Findings*

Panel Data Studies

The earliest rigorous analyses of the impacts of minimum wages on employment and earnings attempted to identify effects through regression performed on large national panels of employment, earnings, and minimum wages by state, and controlling for possible (identified) confounding factors. Their results largely coalesced around the "traditional consensus" view: increasing the minimum wage tends to have a statistically significant impact on unemployment with measured elasticities between 0.1 and 0.3.[54]

Case Studies

The "New Minimum Wage Research Conference," held in November of 1991, reignited debate about the labor market effects of minimum wages. Several of the papers presented evidence directly at odds with the previous "consensus:" by focusing on specific events, Card (1992*b*), Card (1992*a*), and Katz and Krueger (1992) all report statistically insignificant disemployment effects of minimum wage increases. At the expense of the statistical power provided by large panel datasets, the case study approach seek to better control unobserved heterogeneity between disparate geographic and economic regions by comparing only sufficiently alike groups, potentially improving accuracy. Later event studies, such as Card and Krueger (1994), Card and Krueger (2000), and Dube et al. (2007) all arrive at similar conclusions.[55]

Distinct from the geographical-case study approach is an employer-case study method–focusing on the decisions made by a single, large, multi-state employer subjected to differential minimum wage shocks across subsections of its workforce. Coviello et al. (2022) uses this approach to estimate impacts of minimum wage changes on productivity and worker welfare.

[54] Doucouliagos and Stanley (2009) gives an overall "uncorrected" estimate of the employment elasticity of -0.19 (though they suggest this figure is inaccurate due to publication bias). Chapter 6 of Neumark and Wascher (2008*b*) suggests a historical "consensus view" of the elasticity was between -0.1 to -0.3. See Brown et al. (1982), Commission (1981), and Joint Economic Committee Republicans (1995) for additional early surveys of results.

[55] The "city-level event study" approach, used in Dube et al. (2007), remains popular; see, for example Jardim et al. (2022*a*).

3.2.1.2 *Subsequent Advancements*

Some attempts at reconciliation have come in the form of meta analyses and extensive literature reviews (see Doucouliagos and Stanley (2009), Schmitt (2015), and Neumark and Wascher (2008*a*) for examples), yielding little in the way of conclusive results. Others have attempted to improve the underlying econometric methodology. Dube et al. (2010) and Addison et al. (2012) independently sought to extend the intuition present in Card and Krueger (1994) and other quasi-experiments, creating a "panel of quasi-experiments" by matching all "treated" counties in states that had raised their minimum wage to "control" counties located across the state border. This approach has gained traction both across different settings (such as those described earlier in this section), but to study impacts of minimum wages in different national contexts. Muravyev and Oshchepkov (2016) use a similar approach when estimating the impact of an increase in Russia's national minimum wage, while Kong et al. (2021) do the same for China.

3.2.1.3 *Ongoing Debate*

Critiques of Dube et al. (2010) have largely focused on two perceived weaknesses (see Neumark et al. (2014*b*) and Neumark et al. (2014*a*)). First, the county-pair matching methodology unintentionally removes too much useful variation, biasing estimates of effects toward zero. Second, paired counties proved to be poor controls due to differences in observable demographics (see Allegretto et al. (2017) for a fuller accounting of, and response to, these concerns).

Most recently, attention has shifted once again to the appropriateness of Dube et al. (2010)'s geographic controls. Jha et al. (2022) suggests that commuting zones offer a more natural control-treatment specification, and reruns Dube et al. (2010)'s regression focusing only on counties across state borders within a commuting zone. This partially captures some of this paper's underlying motivation–more economically integrated counties are more likely to face similar underlying economic shocks than less integrated counterparts. At the same time, Jardim et al. (2022*b*) have focused on the concern of policy spillovers, demonstrating the importance of accounting for treatment spillover effects in control units.

3.2.1.4 *Geographically-Based Controls*

Another vein of methodological research in political science questions the validity of geographically-based controls in regression discontinuity research designs. Keele and Titiunik (2015*a*) explicitly caution against geographically-based regression discontinuity research designs without taking explicit care justifying an “independent treatment” condition similar to our “no political trends” identifying assumption. (Keele and Titiunik (2015*b*) subsequently demonstrates the sensitivity of estimates to regression specification, and the possibility of geographic neighbors being poor controls for each other.) Additional findings from the field further emphasize the importance of caution. For example, Gerber and Huber (2009) demonstrates the existence of behavioral responses of individuals to political changes, namely that consumption patterns respond to electoral shifts (and not, as one might expect, changes to economic conditions directly). As a result, difference-in-difference estimation in political science tend to emphasize and justify geographically-based research designs more explicitly (Posner (2004) being a notable example).

Some economists have raised similar concerns to their political science colleagues. Roth et al. (2022) directly references the underlying concern we raise with regard to policy trends.[56] Problematically, we find the economics literature tends to refocus on other structural issues with difference in difference estimation. Issues related to “policy bundles,” confounding treatments, and unobserved asymmetries between treatment and control are briefly mentioned, but quickly set aside in favor of other structural causes of estimation bias. While part of the solution is certainly (as Roth et al. (2022) and others suggest) to rely on context-specific knowledge to avoid confounds, we suggest there are, at least some, generally applicable steps that can be taken to ensure proper specification. From this lens, our paper can be viewed as an attempt to synthesize some of the concerns raised in other fields and highlight the practical effects policy trends have on estimation for economists.

We pay particular attention to difference-in-difference estimation methods that

[56]Specifically, Roth et al. (2022) mention “For example, Democratic-leaning states may be more likely to adopt a particular policy (e.g., the minimum wage) and be exposed to different macro-economic shocks” as an example of when a researcher might expect the parallel trend assumption to fail.

rely on geographic proximity for construction of control groups. The observation that neighboring municipal entities can exhibit substantially different political preferences is not new. For example Nall (2015), examining the impact of highway construction on polarization, finds significant political differences between otherwise similar, neighboring counties, noting "[c]ounties often delimit school districts, public services, and other factors relevant to residential sorting, making them units of interest in their own right." Moreover, unobserved divergence in political preferences are compounded once self selection is considered. Political preferences themselves are often an influential factor for individuals choosing where to live at a granular level (McCartney et al. (2021)). Liu et al. (2019) finds specific evidence of partisan sorting at the county level, reinforcing Nall (2015)'s emphasis on the importance of county political control.

3.2.1.5 Heterogeneous Treatment Effects

Our paper is also related to the growing body of research that reexamines DiD estimation of heterogeneous treatment effects. One branch of this literature focuses on heterogeneity in treatment timing, and demonstrates the resulting fragility of "no pretrend" assumptions due to spillovers (see, for example, Goodman-Bacon (2021) and Sun and Abraham (2021)). A second strand has sought to adapt DiD estimation methods for circumstances with continuous (i.e. non-binary) treatment intensities (see de Chaisemartin et al. (2022) and Callaway et al. (2021) for recent approaches). Most closely related are recent works that have focused on bias arising from contamination generally (Goldsmith-Pinkham et al. (2022)). A handful of applied works have sought to individually estimate the effects of multiple related treatments that can each impact outcome variables of interest. Examples include Meinhofer et al. (2021) who directly incorporate the effects of medical and recreational marijuana legalization policies independently, and Chernozhukov et al. (2021) who estimate the effects of a wide array of Covid-related government interventions. Our paper reinforces the importance of this approach, and recommends considering confounding policy trends more broadly.

3.2.1.6 *Identification Assumptions*

Another body of work focuses more directly on conditions required for necessary identification assumptions to be satisfied. As this strand of literature is also closely associated with our approach and findings, we briefly review several of the current methods and associated critiques used to verify the lack of parallel trends specifically, or controlling for confounding treatments generally.

Parallel Trend Analysis

Several recent papers examine biases associated with pre-trend tests. Roth (2022) provides theoretical and simulation results, suggesting "the bias of conventional estimates conditional on passing a pre-test can be worse than the unconditional bias" and that current approaches to pre-tests are under-powered. Rambachan and Roth (2021) argues in favor of an alternative approach to pre-trends entirely: due to the limitations of the standard 'statistical insignificance' test, they propose a pre-trend test that requires pre-intervention deviations between treatment and control groups to be restricted to a pre-defined set Δ. One similarity between our proposal and Rambachan and Roth (2021) is the explicit mentioning of the possibility of confounding policies in Rambachan and Roth. However, in Rambachan and Roth (2021), the consideration of confounding policies is largely confined to a discussion about the possibility and feasibility of monotonicity restrictions for identifying possible candidates for their set Δ, and is not discussed in a broader context.

Treatment-Control Similarity and Asymmetric Shocks

Note that traditional difference in difference regression designs require both control and treatment to follow parallel trends in key variables, which alone does not necessarily impose a strict "similarity" condition on the treatment and control groups themselves. Instead, what is generally required is that differences between treatment and control groups must be orthogonal to variables being estimated. Of course, ensuring orthogonality is itself often a difficult task, leading to debates regarding proper control specification.[57] Minimizing bias is thus most transparently

[57]See, for example, Dube and coauthors' (Dube et al. (2010), Allegretto et al. (2017)) back-and-forth with Neumark and coauthors (Neumark et al. (2014*b*), Neumark and Wascher

achieved by choosing a control group that mirrors treated units as closely as possible.

The importance of controlling for the influence of localized characteristics on outcomes is one of the central motivations behind modeling decisions. Continuing to use the minimum wage literature as an example, controlling for regional policy and structural differences was one of Dube, Lesterer, and Reich's (2010) reasons for their county-matching design. As detailed explicitly in their follow-up paper Allegretto et al. (2017), "High minimum wage states are concentrated on the Pacific Coast, the Northeast, and parts of the Midwest; tend to be Democratic-leaning; and have experienced less de-unionization. These disparities raise the possibility that trends in other policies and economic fundamentals may also differ between these groups of states" (p. 560).

For many longer-term trends, this approach may well sufficiently avoid estimate bias. Our concern here, however, is *policy bundling.* Minimum wage changes are not stochastic events whose probability is determined by a function of political lean and unionization. Instead it is a policy crafted and implemented as the result of particular political forces. Moreover, minimum wage legislation and political outcomes are themselves endogenously linked (Markovich and White (2022)), leading to possible increases in estimation bias over longer time periods.

One of the primary underlying concerns we seek to raise is the possibility of dissimilarities between treatment and control, combined with policy endogeneity and bundling, leading to asymmetric shocks. Crucially, it is not theoretically clear that deviations between treatment and control will be limited to *different responses to common shocks.* Instead, given the possibility of confounding policy influences and non-arbitrary boundaries between units of observation, it is theoretically possible for treatment and control groups to be hit by *entirely different economic shocks.* Common approaches to relax the parallel trend assumptions (for example, Rambachan and Roth (2021)) may themselves be insufficient for alleviating bias.

(2017)) regarding the proper control group specification in contiguous county research on effects of minimum wage changes.

3.2.2 Political Preferences and Policy Outcomes

3.2.2.1 Policy Bundling

While there are some DID studies that discuss the possibility of confounding policy interventions, their inclusion tends to be: 1) restricted to comparatively narrow domains, and 2) included mostly as an afterthought or without sufficient direct consideration. In practice, this has led researchers to narrowly consider circumstances where confounding policies may be present. As such, the existence of "policy bundling" is not an issue that has been entirely ignored, but rather confined to the most blatant applications.

For example, consider the problem of estimating of the impacts of various policies on Covid-19 spread and public health outcomes. It would be difficult to argue, without substantial evidence, that policy-based mitigation control efforts were implemented randomly and independently. Instead, we observe states pass multiple public health measures in relatively rapid succession, the stringency of which is likely to be closely associated with beliefs regarding the outcome variables in question (hospital admissions, confirmed cases, Covid deaths, etc.). It is therefore not surprising that recent examinations into the effects of individual policies consider these policy confounds. To take one recent paper, Chernozhukov et al. (2021) incorporate not only a broad basket of policies passed by different governments, but also behavioral responses by individuals (specifically they account for the degree to which social distancing was observed, as measured through Google Mobility Reports data).

3.2.2.2 Policy Responsiveness

Our work is heavily informed by political science research. The sensitivity of policy outcomes not only to the electorate's underlying preferences, but also the larger political context, is certainly not a new revelation.[58] Numerous papers have identified interactions between the political process and economic outcomes, such

[58] This sentiment is exemplified by Obama's 2009 statement that "elections have consequences."

as Margalit (2013) and Brunner et al. (2011) (political preferences and economic circumstances), and Charles and Stephens (2013) (economic outcomes and voter turnout). In many ways, our findings are an extension of these previous insights, and an application of the motivation behind Besley and Case (2003). Minimum wage increases may track political preferences on the whole, but in practice are determined by elected representatives chosen through variously-democratic electoral maps.

This relationship between electoral representation and policy changes is particularly acute with regards to the minimum wage; Simonovits et al. (2019) find persistent deviations between citizen preferences and legislatively enacted minimum wages across states. The importance of the political process for understanding observed outcomes is described directly in a follow-up study Simonovits and Bor (2021):

> "...considering variation across states, we find a great degree of heterogeneity despite the short-span of our study: in many states, legislative inaction together with increasing support for higher minimum wages has lead to significantly increasing gaps between preferences and policies. In contrast, other states have seen overly responsive policy-making that led to a significant improvement of representation, at least on the short run. Taken together, our findings point to the importance of using measures of representation that can uncover important differences across states."

3.3 Setup and Proposed Solutions

For simplicity of exposition, consider a simple setting where a group of treated units receives treatment at the same time. We can decompose the total change in some variable of interest into change due to treatment (minimum wage) and other changes not related to treatment. Thus, we obtain the following expressions for the treated and untreated groups

$$\Delta Y_T = \Delta Y_T^{\text{MW}} + \Delta Y_T^{\text{Other}}$$

$$\Delta Y_{UT} = \Delta Y_{UT}^{\text{Other}}$$

The idea of DiD is that we can use the observed changes in ΔY_{UT} to recover the unobserved ΔY_T^{MW} as long as $\Delta Y_T^{\text{Other}} = \Delta Y_{UT}^{\text{Other}}$. Namely,

$$\Delta Y_T - \Delta Y_{UT} = \Delta Y_T^{\text{MW}} + \left(\Delta Y_T^{\text{Other}} - \Delta Y_{UT}^{\text{Other}}\right)$$

Thus, if treated and untreated groups respond similarly to shocks that are unrelated to minimum wage changes, the difference in the change in the variable of interest between treated and untreated groups could be attributed entirely to changes in minimum wage. We argue that the $\Delta Y_T^{\text{Other}} = \Delta Y_{UT}^{\text{Other}}$ assumption (also known as parallel trends) is more likely to hold for the sample of treated and untreated counties that are politically similar.

How could political (dis)similarity affect the parallel trends assumption? First, county pairs that have diverging political views are more likely to have unobservable differences that get stronger over time. For example, Allegretto et al. (2017) point to the relative political similarities of different regions, "High minimum wage states are concentrated on the Pacific Coast, the Northeast, and parts of the Midwest; tend to be Democratic-leaning; and have experienced less deunionization."(p.2). Stronger unions are generally associated with stronger labor market rigidities, implying that minimum wage employers would have less room to adjust the demanded quantity of labor ober time. If strength of labor unions is correlated with local political environment, then employment trends of treated and untreated units with different political environments diverge over time, violating parallel trends. In Section 3.5, we provide some preliminary evidence suggesting that politically similar county pairs are also more similar in levels and in trends of some relevant labor market variables than politically unaligned counties.

The second concern we address is the reality and importance of simultaneous policy changes to the minimum wage change. Minimum wage changes are political events, often directly influenced by outside factors, such as which political party controls state legislature.

For example, consider the case of minimum wage and employment in the restaurant industry. For simplicity, we assume all other demographics and characteristics between the two counties are identical, and thus abstract away from our previous concern (though, we believe that, in reality, both issues could easily have compounding effects). Say, there is an election where state Democrats retake the state legislature and Governor's Mansion from Republican control.[59] Following the traditional political realities of many swing states, one agenda item for the new Democratic majorities is to increase the minimum wage, which they do so in a law slated to take effect the next year. At the same time, they pass a spending bill which encourages infrastructure developments across the state (again, for expositional ease, we'll assume these projects are school construction).

If our regression analysis is strictly focused on the employment effects from the minimum wage adjustment in the restaurant sector, it might not be immediately clear that the spending bill should have any adverse effects on our analysis. However, we identify at least two different channels that might lead to spillover that would hurt the estimation. One possibility would be supply-side effects (from the restaurant's perspective, that is): the increase in construction projects would lead to a hiring boom for construction workers, which might affect the availability of potential restaurant employees, but concentrated only in the county that also saw the increase in minimum wage. Of course, spillovers (workers from the matched county, or other nearby counties, choosing to work on construction projects in the "treatment" county) might lead us to believe that this labor market issue would be mitigated, however one crucial aspect of the matched-county pairs methodology is that this county-spillover is not substantially powerful (else, all estimates in the regression become suspect immediately).

The second possible channel can be described as a "demand-side effect." A new school built in the treated county leads to an increase in construction workers in the immediate vicinity (that is, in the treated, but not control, county). If these new workers impact the demand for restaurant services (such as, an increase in the lunch rush of local eateries), then any adverse employment effects that would have resulted from the minimum wage change are simultaneously mitigated by this countervailing force. Even if a restaurant was going to lay off an employee due to the higher wage change, they may choose to keep the employee thanks to

[59] An election, of course, is not strictly necessary; all we seek to do is introduce some political event which causes the state legislature to change sufficiently enough to likely encourage an "exogenous" policy shift, like a minimum wage increase.

the increased demand for their products. Excluding this asymmetric demand shock for restaurant services from the regression leads the analyst to put all the weight on the minimum wage.

3.4 Data

3.4.1 Wage Data

The employment and wage data come from the Quarterly Census of Employment and Wages (QCEW), which provides payroll data at the county and industry level. The data are constructed from payroll tax fillings that each firm is legally required to submit. The data cover all workers enrolled in unemployment insurance, which represents 98% of all workers. The data spans from the first quarter of 1990 to the second quarter of 2006. We only include the mainland states since we are interested that straddle counties on state boundaries. To preserve confidentiality where few workers work in the restaurant industry, full reports for the 66 quarters are available for 1380 of all the 3109 mainland counties. We will concentrate our attention on counties with complete reports for all quarters, but we will include counties with partial reporting as a robustness check.

3.4.2 State Spending Data

Data for the state finances come from the Annual Survey of State Government Finances.[60] This survey provides detailed yearly spending at the state level. An estimation of spending per county is obtained by multiplying the yearly state spending by the fraction of state's population residing in said county.[61]

State spending is divided into broad function categories.[62] As noted in the literature, governments tend to adopt many social policies including minimum wage

[60]The data is freely available at https://www.census.gov//govs/state/historical_data.html

[61]Historical data on population by county comes from Intercensal Estimates provided by Census Bureau `https://www.census.gov/programs-surveys/popest/data/tables.html`.

[62]Education, public welfare, hospitals, health, highways, police protection, correction, natural resources, parks and recreation, government administration, interest on general debt and other and unallocable.

changes in conjunction. Thus, in our analysis, we will control for spending in education, public welfare, hospitals, health and capital outlay. The latter is a broad spending category in physical capital expansion and maintenance.[63] We believe that it is important to control for spending in infrastructure and other social programs. It is possible that an increase in minimum wage indeed leads to a decrease in employment everything else equal, but that other initiatives are put in place at the same time that stimulates the supply of restaurants offsetting on the net, the employment loss due to a minimum wage increase.

3.4.3 Election Data

We use voting data in general presidential elections from CQ Voting and Elections Collection from CQ Press (an imprint of SAGE Publications). Data are combined from a variety of official sources and provides poll results at the county level for all election cycles going back to 1789. In this paper, we use county voting data for 6 presidential elections 1988, 1992, 1996, 2000, 2004 and 2008, coinciding with the elections that generally cover the time-frame of the dataset. The county FIPS codes are not listed, so we match counties manually using their name and state to their FIPS code using the list of FIPS codes from the Census Bureau.[64] This procedure allows us to match election results as a percentage of votes for each party at the county level to the counties in our sample.

3.5 Summary Statistics

DLR (2010) concentrate their attention on the restaurant industry. The reason is relatively straightforward: 30% of all minimum wage workers are employed by this industry, a larger share than any other. Looking at this industry is also interesting from a statistical standpoint since in the vast majority of counties, this

[63]This category is formally defined as "direct expenditure for contract or force account construction of buildings, grounds, and other improvements, and purchase of equipment, land, and existing structures. Includes amounts for additions, replacements, and major alterations to fixed works and structures. However, expenditure for repairs to such works and structures is classified as current operation expenditure".

[64]The fips code list is available at https://www.census.gov/geo/reference/codes/cou.html.

industry is large enough to have published data and thus allows the comparison of a large number of contiguous county pairs. We follow DLR (2010) and focus on this industry as well.

TABLE 3.1 – DESCRIPTIVE STATISTICS

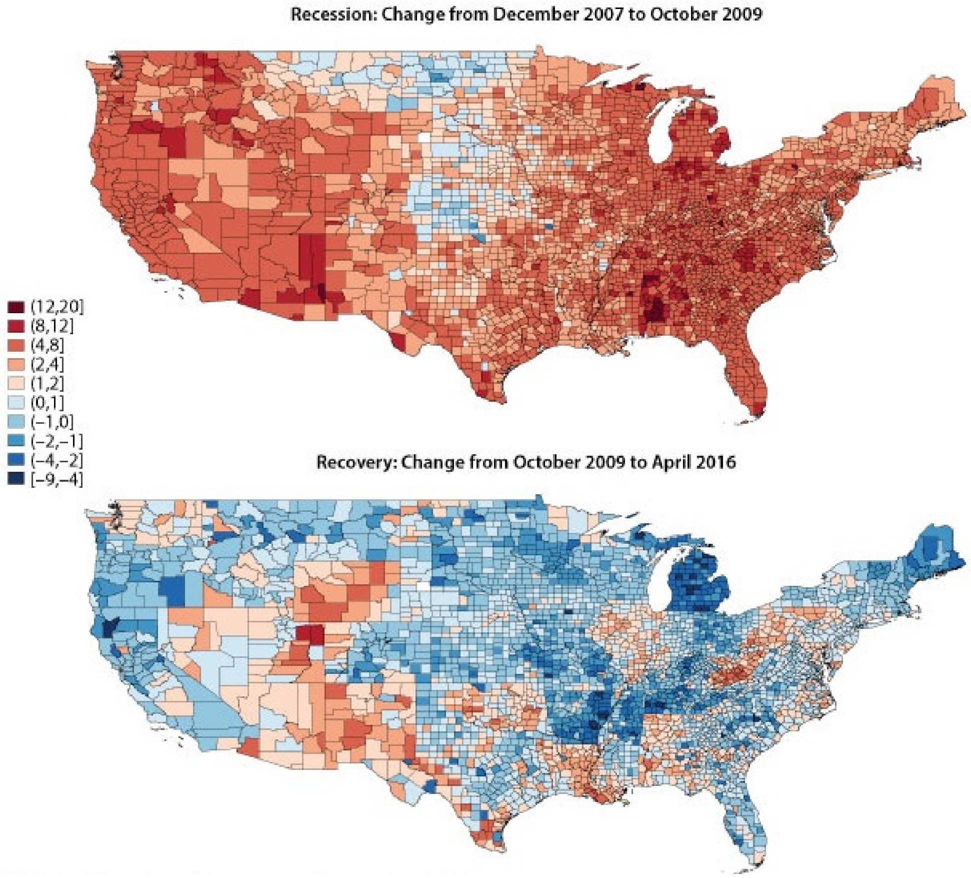
	(1)		(2)	
	DLR Sample		Perfectly Aligned Sample	
	Mean	Std.dev.	Mean	Std.dev.
Population, 2000	167,956	297,750	152,396	335,276
Intercensal population	163,250	290,201	148,317	325,779
Population density, 2000	556	3,335	326	1,033
Land area (square miles)	1,380	2,470	1,178	2,190
Overall private employment	32,185	101,318	28,749	98,854
Restaurant employment	4,185	7,809	3,584	7,684
Restaurant average weekly earnings ($)	172	46	168	43
Minimum wage ($)	4.84	0.67	4.86	0.69
County state spending ($)				
Capital outlay	9,900	1,190	9,793	1,158
Education	11,359	1,138	11,274	1,074
Health	9,054	1,288	8,957	1,243
Hospitals	8,897	1,411	8,818	1,317
Public Welfare	10,947	1,269	10,884	1,219
Observations	70,620		27,720	
Counties	504		281	
County pairs	754		288	

Notes: Sample means are reported for all contiguous border county-pairs with a full balanced panel of observations. Weekly earnings and minimum wages are in nominal dollars. Sources: QCEW; U.S Department of Labor, Employment Standards Administration, Wage and Hour Division. U.S. Bureau of the Census, *2000 Census*, U.S. Bureau of the Census, *Annual Intercensal Population Estimates*.

3.5.1 Differential Response to Shocks

DLR (2010)'s analysis requires that "minimum wage differences within the pair are uncorrelated with the differences in residual employment (or earnings) in either county." This assumption will not hold if, for example, counties in a pair are not affected symmetrically by a given economic shock. In such an event, unless suffi-

FIGURE 3.1 – DIFFERENCES IN UNEMPLOYMENT RATES DURING AND AFTER THE GREAT RECESSION (PERCENTAGE POINTS)



cient covariates are included in the model to account for the differential effect, the effects of an unobservable shock will not be netted out of the regression results.

Of course, this concern is only significant if economic shocks impact counties differently. However, we believe there is sufficient cause to believe this might be the case. Let us first present a particular event (albeit not representative of the "average shock") that shows that some contiguous counties are not affected symmetrically.

Figure 3.1 shows the evolution of employment rates during the Great Recession of 2007-2009 and in the recovery period of 2009 to 2016.[65] It is clear that not all counties recovered in the same way, which is a not a problem per se. However, one can clearly trace out differences in recoveries around some state borders suggesting that contiguous counties on both sides of the border clearly differ in their response

[65]The map comes from a research article from the Federal Reserve Bank of St. Louis (Dvorkin and Shell, 2016).

to the shock. Some of the state-border differences aren't a major concern for Dube et al. (2010)'s analysis: for example, the relatively successful recovery observed in western Texas compared to the persistent unemployment in eastern New Mexico, or the differences between the drop in unemployment in western Nebraska compared to the persistent recession effects in Wyoming might be initially forgiven, as these state-pairs don't appear in DLR's dataset.[66] However, several of the most striking differences across state borders may well be cause for concern. There is a notable line, coinciding with the New York-Pennsylvania border (one of the state-pairs that is included in DLR): counties on the New York side (at least as far as unemployment is concerned, between October 2009 and April 2016) have observed noticeably and consistently larger drops in unemployment than the counties on the Pennsylvania side- a fact that runs contrary to the fundamental identifying assumption necessary in the county-pairs analytical design. Other state borders are even starker- the recovery in Arkansas has been noticeably better than that experienced in Oklahoma and Louisiana (both borders are included in DLR); Michigan's near uniform substantial decrease in unemployment contrasts clearly with Ohio and Indiana, as well as Wisconsin. The two Dakotas have also experienced relatively different paths during the recovery, and the state lines between Colorado and New Mexico, California and Nevada, Oregon and Nevada, and Illinois with its neighbors to the west are all potentially causes for concern, as all these pairs are included in DLR.[67] Although this is one of many shocks that can befall the U.S. economy, it clearly shows that at least some border counties are not equivalently affected by such shock, which would invalidate Dube et al. (2010)'s symmetric response assumption.

3.5.2 Evidence of the Importance of Political Alignment

We now proceed in showing that counties that are consistently similar in their political views are also much more similar in variables that determine local labor

[66]This exclusion is because neither of these state-pairs differentially changed their minimum wages.

[67]Note that we aren't offering these specific pair examples as evidence that the fundamental findings of Dube et al. (2010) are incorrect. It must be noted that the recovery period after the Great Recession does fall outside the time period studied by DLR in 2010. However, we believe that this should be taken as an informative example of how widespread the asymmetric response of neighboring counties to economic shocks can be.

market conditions, but are unlikely to be substantially affected by minimum wage changes. Following Dube et al. (2016), we consider levels and 4-quarter and 12-quarter changes (trends) in log population, log of overall private sector and manufacturing employment, log of average private sector and manufacturing weekly earnings, and private sector employment-to-population ratio (EPOP). For each of the variables, we test for differences in mean absolute values between perfectly aligned contiguous counties and never aligned contiguous counties.

For each election cycle, we define two counties as being politically similar or aligned if both counties in the pair voted 70% or more for one of the two main parties or if the difference in votes for Republicans differs by less than 10%.[68] We then calculate the number of election cycles around the sample period that the two counties were politically similar. Out of the six election cycles in the 1988-2008 period, perfectly and never aligned counties would have voted similarly in each election and none of the elections, respectively. Figure 3.5 in Appendix D displays (in blue) all contiguous county pairs along state borders that saw a minimum wage differential between 1990 and 2006. Red segments represent county pairs that are politically aligned in all 6 elections. Among the 754 county pairs used by DLR (2010), around 60% are not perfectly politically aligned and 13% are never aligned.

Table 3.2 shows the results for the aforementioned variables. For almost all statistics, the mean absolute differences are larger for never aligned county pairs, and the gaps are statistically significant at the 1% level. Thus, we believe that politically aligned county pairs are more likely to produce parallel trends needed for identification.

[68]We use the share of GOP voters instead of Democratic voters to better account for periodic third party challengers and their impact on political alignment. While Ralph Nader captured a relatively large share of total votes cast in the 2000 Presidential Election (likely drawn from mostly Democratically-aligned citizens), the growth of the Libertarian Party, and periodic successes of challengers from the right-side of the political spectrum (most notably Ross Perot in 1992 and 1996), suggest that focusing on Democratic votes might lead to improper inclusion. One prominent example is Loving County, Texas, which regularly favors Republican candidates, but when given a third option from the right, often embraces it wholeheartedly (voting for both Ross Perot in 1992 and George Wallace in 1968). Loving County is likely politically different from other Republican counties in ways not measured by Democratic vote share.

TABLE 3.2 – MEAN ABSOLUTE DIFFERENCES IN LABOR MARKET VARIABLES BETWEEN NEVER AND PERFECTLY ALIGNED COUNTY PAIRS

	Never Aligned	Perfectly Aligned	Gap
Level:			
EPOP	0.113	0.101	0.012***
	(0.001)	(0.000)	(0.001)
Log mfg earnings	0.255	0.230	0.025***
	(0.002)	(0.001)	(0.003)
Log private earnings	0.190	0.183	0.007***
	(0.001)	(0.001)	(0.002)
Log mfg employment	1.281	1.205	0.076***
	(0.011)	(0.006)	(0.012)
Log private employment	1.305	1.137	0.168***
	(0.009)	(0.005)	(0.010)
Log population	1.008	0.879	0.128***
	(0.007)	(0.004)	(0.008)
4-quarter change:			
EPOP	0.014	0.013	0.001***
	(0.000)	(0.000)	(0.000)
Log mfg earnings	0.082	0.077	0.005***
	(0.001)	(0.001)	(0.001)
Log private earnings	0.055	0.048	0.007***
	(0.001)	(0.000)	(0.001)
Log mfg employment	0.122	0.108	0.014***
	(0.002)	(0.001)	(0.002)
Log private employment	0.065	0.053	0.011***
	(0.001)	(0.000)	(0.001)
Log population	0.015	0.011	0.003***
	(0.000)	(0.000)	(0.000)
12-quarter change:			
EPOP	0.003	0.004	-0.001***
	(0.000)	(0.000)	(0.000)
Log mfg earnings	0.115	0.106	0.009***
	(0.001)	(0.001)	(0.002)
Log private earnings	0.100	0.097	0.003***
	(0.001)	(0.000)	(0.001)
Log mfg employment	-0.053	-0.027	-0.026***
	(0.003)	(0.002)	(0.003)
Log private employment	0.025	0.031	-0.006***
	(0.002)	(0.001)	(0.002)
Log population	0.020	0.016	0.004***
	(0.000)	(0.000)	(0.000)

Notes: Significance levels: *10%, **5%, ***1%.

3.6 Empirical Strategy

We now turn our attention to how we estimate the effect of minimum wage on earnings and employment. We rely on the contiguous county pairs along state borders framework proposed by DLR (2010), which assumes that absent changes in minimum wage the outcome variables in paired counties would have evolved in a similar way. The major difference in our approach is that we argue this assumption might not always hold for all county-pairs. Instead, we restrict our analysis to counties that have historically been aligned in their political preferences, making the identifying assumption more credible.

The effect of minimum wage changes can be estimated by the following difference-in-difference model:

$$ln(y_{ipt}) = \alpha + \eta ln(MW_{it}) + \delta ln(y_{it}^{TOT}) + \gamma ln(pop_{it}) + S_{it}\beta + \phi_i + \tau_{pt} + \varepsilon_{ipt}. \quad (1)$$

In equation (1), the unit of observation is county i in pair p in quarter t. Our main dependent variables are log earnings and log employment in the restaurant industry. Employment measures the number of workers at the county level in the restaurant industry. Our coefficient of interest η can be interpreted as elasticity of the outcome variable with respect to the minimum wage in this log-log specification. Except for controlling for levels and lags of state spending, we include the same controls as DLR (2010):

- *Observable initial differences in earnings and employment*, $ln(y_{it}^{TOT})$:

  For the earnings regression, we control for the quarterly average employment level in the private sector and for the employment regression, we include the quarterly employment in the private sector.[69] Failing to control for the baseline level of earnings and employment would lead to identification problems, since we wouldn't be able to distinguish changes in the outcome variables that are due to changes in minimum wage from variations in aggregate earnings and employment.

[69] Ideally one would include earning and employment for both the private and public sector, but data for the public sector is not available at the analysis level.

- *Annual county population, $ln(pop_{it})$*: We control for the yearly population in the county in the employment regression as it may be correlated with the baseline employment level and to account for the fact that the pool of minimum wage worker will differ with population size.
- *State spending, S_{it}*: As we mentioned earlier, it is likely that changes in minimum wage are accompanied by other policies. To reduce the pressure on minimum wage employers, the legislature could decrease the fiscal burden for these employers along with the minimum wage change. If a minimum wage would alone reduce employment, it could be offset by a lighter burden on employers. To identify the true effect of minimum wage changes, it is therefore paramount to take into account other policies that could affect the labor demand and supply in the county. Even when looking only at the restaurant industry, the true effect of a change in minimum wage change could be shadowed simply because minimum wage workers see a change in their outside option. Hence, we control for yearly level of spending in education, public welfare, hospitals, health and capital outlay. To get a measure at the county level, we use *per capita* state spending times the county population level.
- *Fixed observables and unobservables, ϕ_i and τ_{pt}*: We include county fixed effects to sweep out idiosyncratic time invariant observables and unobservables at the county level. We also include a pair flexible trend, thus identifying the minimum wage effect from within county pairs variation.

3.6.1 Identifying assumption

To define our identifying assumption more formally, we require that $E[ln(MW_{it}), \varepsilon_{ipt}] = 0$ which means that the (log) minimum wage in the pair is uncorrelated with the unobservable residual employment and earnings in either county. This is the same identifying assumption used by Dube et al. (2010), as well as all other papers with similar methodology. This assumption is violated when, following a shock, counties diverge in their response absent a change in minimum wage: more precisely, not only the level of the outcome of interest changes, but also its path.

As we have previously shown, there is reason to believe this assumption might not hold. Recovery from the Great Recession is not only "lumpy," but certain state borders show a discontinuous difference in unemployment change since the end of the recent recession. More broadly speaking, if any variable that wasn't included in the original regression is correlated to unemployment or earnings, the identifying assumption is violated. One way states may react differently is by enacting different policies in response to a shock. Hence, controlling for state spending is a way to control for at least one dimension of policy variation. Accounting for political similarity between counties allows us to capture important differences between counties along social and economic dimensions that could be correlated with changes in minimum wage, increasing the likelihood county-pairs will behave the same way to various shocks, thus making the common trend assumption more convincing.

3.7 Results

In this section we present our main results for the restaurant industry. In Table 3.3 , we present our results using all contiguous county pairs as in DLR (2010), effectively replicating their results.

Columns 1 and 2 in Table 3.3 are the well-known findings reported in DLR—increasing the minimum wage by 10% leads to an increase of roughly 2% in earnings, with no statistical change in employment. Column 1 includes county pair-period interaction dummies, and column 2 adds controls for total private sector employment and earnings. Column 3 through 6 add state spending control variables at the county level. We find no statistical or economically significant change in either earnings or unemployment.

In Table 3.4, we restrict DLR's sample to counties that have been historically aligned in their voting patterns in presidential elections. First, we find that our estimates of the elasticity of earnings with respect to minimum wage are similar to the ones documented in DLR and in the minimum wage literature more generally. Second, in column 1, we find that controlling for political alignment produces a marginally significant disemployment effect of -0.12, which more in line with the NW estimates. In column 2, adding total private sector employment decreases

TABLE 3.3 – FULL DLR SAMPLE

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
			ln(earnings)			
ln(MWt)	0.200***	0.189***	0.202***	0.203***	0.203***	0.196***
	(0.065)	(0.060)	(0.067)	(0.067)	(0.066)	(0.064)
			ln(employment)			
ln(MWt)	0.057	0.016	0.023	0.028	0.025	0.011
	(0.115)	(0.099)	(0.102)	(0.102)	(0.108)	(0.127)
lntotprivatesector		0.393***	0.396***	0.399***	0.402***	0.405***
		(0.117)	(0.115)	(0.115)	(0.116)	(0.118)
lnpop	1.116***	0.714***	0.707***	0.723***	0.733***	0.774***
	(0.190)	(0.246)	(0.248)	(0.253)	(0.255)	(0.245)
Controls						
County f.e.	Y	Y	Y	Y	Y	Y
County-pair × Period dummies	Y	Y	Y	Y	Y	Y
State Spending						
L0			Y	Y	Y	Y
L1				Y	Y	Y
L2					Y	Y
L3+L4						Y
Observations	70,620	70,582	70,582	70,582	70,582	70,582
Counties	754	754	754	754	754	754
County pairs	835	835	835	835	835	835

Notes: Columns 1 and 2 replicate the two subcolumns in column 6 in DLR (2010). Columns 3 through 6 add contemporaneous and lagged state spending allocated to a county on per capita basis. Robust standard errors, in parentheses, are clustered at the state and border segment levels. L0 denotes contemporaneous state spending. Significance levels: *10%, **5%, ***1%.

significance below the 10% level. Adding in contemporaneous state spending and its lags increases the disemployment effect and restores its significance. Moreover, controlling for more lags of state spending the disemployment further decreases estimate to -0.177.

It may appear that marginal significance of the disemployment effect in our setting could be noise or a result of a cherry-picked sample. However, despite dropping roughly 60% of the sample, standard errors for log minimum wage decrease across all specifications.

How can we reconcile our estimate of disemployment effect and the one in DLR (2010)? In Table 3.7 we provide evidence that the estimate in DLR likely lumps together heterogeneous effects across county pairs with different levels of political

TABLE 3.4 – PERFECTLY ALIGNED PAIRS (1990-2006)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
			ln(earnings)			
ln(MWt)	0.212***	0.202***	0.207***	0.208***	0.209***	0.203***
	(0.038)	(0.035)	(0.034)	(0.034)	(0.035)	(0.032)
			ln(employment)			
ln(MWt)	-0.116*	-0.139	-0.145*	-0.140*	-0.177*	-0.177*
	(0.063)	(0.085)	(0.081)	(0.083)	(0.095)	(0.095)
lntotprivatesector		0.339***	0.339***	0.340***	0.343***	0.343***
		(0.105)	(0.103)	(0.102)	(0.101)	(0.101)
lnpop	1.098***	0.764***	0.753***	0.765***	0.748***	0.748***
	(0.185)	(0.247)	(0.275)	(0.282)	(0.279)	(0.279)
Controls						
County f.e.	Y	Y	Y	Y	Y	Y
County-pair $\times$ period dummies	Y	Y	Y	Y	Y	Y
State Spending						
L0			Y	Y	Y	Y
L1				Y	Y	Y
L2					Y	Y
L3+L4						Y
Observations	27,720	27,684	27,684	27,684	27,684	27,684
Counties	281	281	281	281	281	281
County pairs	288	288	288	288	288	288

Notes: Sample includes perfectly aligned counties that voted similarly in the six election cycles around the sample period. Columns 3 through 6 add contemporaneous and lagged state spending allocated to a county on per capita basis. Robust standard errors, in parentheses, are clustered at the border segment level. There are not enough politically aligned states to support clustering at the state level. L0 denotes contemporaneous state spending. Significance levels: *10%, **5%, ***1%.

alignment. Namely, we run specification 3 from Tables 3.3 and 3.4 on samples of county pairs that vary by political alignment. In column 1, we consider a sample of counties that is never politically aligned in the six election cycles around sample period. In fact, the elasticity of employment with respect to minimum wage is positive and significant at 5%. Therefore, positive estimate of η would either point to a non-competitive labor market in politically unaligned counties, or to an underlying confounding trend. Note also that the estimate of earnings' elasticity becomes insignificant as well.

Column 2 includes all county-pairs, so the result matches the one in column 3 of Table 3.3 that replicates DLR and adds state spending controls. Columns 3 and 4 split the sample into relatively unaligned and relatively aligned counties, respectively. In column 3, the employment elasticity estimate is still positive, albeit less so than in the never aligned sample, but becomes insignificant. In column 4, it becomes negative and moves closer to the perfectly aligned one in column 5.

TABLE 3.5 – COUNTY-PAIRS ACROSS LEVELS OF ALIGNMENT (1990-2006)

	Politically aligned in				
VARIABLES	0	0 to 6	1 to 3	4 to 6	6
			elections		
	(1)	(2)	(3)	(4)	(5)
			ln(earnings)		
ln(MWt)	0.132	0.194***	0.232**	0.198***	0.207***
	(0.087)	(0.040)	(0.097)	(0.043)	(0.034)
			ln(employment)		
ln(MWt)	0.245**	0.023	0.073	-0.051	-0.145*
	(0.116)	(0.079)	(0.140)	(0.083)	(0.081)
lntotprivatesector	0.447***	0.396***	0.499***	0.356***	0.339***
	(0.123)	(0.079)	(0.172)	(0.098)	(0.103)
lnpop	0.853***	0.707***	0.475	0.681***	0.753***
	(0.280)	(0.174)	(0.304)	(0.202)	(0.275)
Controls					
County f.e	Y	Y	Y	Y	Y
State spending L0	Y	Y	Y	Y	Y
County-pair $\times$ period dummies	Y	Y	Y	Y	Y
Observations	10,560	70,582	18,148	41,874	27,684
County pairs	109	754	201	444	288

Notes: Samples include counties that voted differently or similarly in the six election cycles around the sample period. Robust standard errors, in parentheses, are clustered at the state by border segment level. The number of state-level clusters is less than 40 due to restrictions on political alignment in the perfectly aligned case, therefore we are not using clustering method used by DLR. Significance levels: *10%, **5%, ***1%.

Even though our analyses are still preliminary, the results suggest that politically dissimilar county pairs may be affected by different time-varying unobserved shocks that are correlated with changes in minimum wages, thereby violating the parallel trends assumption. Therefore, the role of local political environment warrants additional investigation. We plan to further investigate their role using an up-to-date sample on minimum wage changes. We discuss our plans for future work in Section 3.8.

3.7.1 *Testing for Pre-Existing Trend*

To further justify our approach and sample selection, we test for pre-existing trends in our dependent variables. To do so we run regression (8) from DLR:

$$ln(y_{it}) = \alpha + \eta_{12}[ln(MW_{it-12}) - ln(MW_{it-4})] + \eta_4[ln(MW_{it-4}) - ln(MW_{it})] \\ + \eta_0 ln(MW_{it}) + \gamma ln(pop_{it}) + S_{it}\beta + \phi_i + \tau_{pt} + \varepsilon_{it}. \quad (1)$$

In equation (1), η_{12} (η_4) captures the level of the outcome variables 12 (4) quarters before the minimum wage change. In this specification if $\eta_4 - \eta_{12}$ is significantly different from 0, it would indicate pre-trend differences in the outcomes. Table 3.6 reports the results for DLR's specification and for our perfectly aligned and never aligned sample. In neither case, we find evidence of pre-trends—despite the fact that the disemployment effect is positive and significant in the never aligned sample. For this reason, we believe that the role of local political trends when choosing relevant control groups should be examined in a greater detail.

TABLE 3.6 – TESTING FOR PRE-EXISTING TREND

	All county pairs (DLR)	Perfectly Aligned	Never Aligned
		ln(earnings)	
η_{t-12}	0.031	0.039	-0.048
	(0.046)	(0.069)	(0.076)
η_{t-4}	0.065	0.064	-0.042
	(0.082)	(0.115)	(0.121)
Trend (η_{t-12}-η_{t-4})	0.034	0.025	0.006
	(0.061)	(0.090)	(0.078)
		ln(employment)	
η_{t-12}	0.010	-0.004	0.071
	(0.083)	(0.071)	(0.115)
η_{t-4}	0.056	-0.047	0.144
	(0.186)	(0.174)	(0.222)
Trend (η_{t-12}-η_{t-4})	0.046	-0.043	0.072
	(0.138)	(0.133)	(0.177)
Controls			
County f.e.	Y	Y	Y
State Spending L0	Y	Y	Y
County-pair $\times$ period dummies	Y	Y	Y
Observations	64,200	25,200	9,600

Notes: Here $t - j$ denotes quarters prior to the minimum wage change. All the employment specifications include log of county-level population. Robust standard errors, in parentheses, are clustered at the state and border segment levels. Significance levels: *10%, **5%, ***1%.

3.8 Future Work

We plan to improve the precision of our estimates of disemployment effect by collecting additional data on changes in state minimum wages to occurred prior to 1990 and post 2006. Adding additional observations to the sample of politically aligned counties will hopefully address the marginal significance of our effect. Additionally, our chosen thresholds in the definition of political similarity of counties require robustness checks.

Moreover, we plan to improve our two-way fixed effects (TWFE) model by including recent insights from the literature on identification of treatment effects in staggered rollout designs (Callaway et al. (2021); Goodman-Bacon (2021); Sun and Abraham (2021)). With staggered rollout, units receive treatment at different points in time. In particular, in some cases groups that are continuously treated throughout the sample period will get used as untreated units. This is not a problem per se, unless treatment effects are not constant over time (dynamic) or vary across treated groups (heterogeneous). For example, if a treatment effect becomes stronger over time, the continuously treated group will be trending upwards and will not be a relevant comparison group to the unit that is just receiving the treatment. Similar differential trends will result if treatment effects are heterogeneous. This is likely the case in our study as evidenced by heterogeneity in estimates across county pairs with different levels of political alignment. Therefore a conclusive estimate of disemployment effect should take into account both violation of parallel trends due to staggered rollout concerns and due to the political confounds discussed here.

3.9 Conclusion

Although this work remains in a relatively preliminary state, we believe it provides suggestive evidence of some concerns in the county-pairs treatment-control methodology. Although there is intuitive appeal in assuming contiguous counties are close enough (politically, geographically and economically) to act as controls for each other, it may not always be the case. Minimum wage changes are fundamentally political events, which can be correlated to other policy changes that almost certainly have economic impacts. If the two counties are not sufficiently politically aligned, it is unlikely that these other policy changes will follow the same pattern, which can introduce unobserved and confounding variation into the regression analysis. In addition, economic responses to unobserved shocks are likely partially driven by political forces; if these aren't controlled for, the identification assumption central to the analysis technique can easily fail. We believe that our initial results present evidence that, once these two forces are accounted for, several of the negative disemployment effect findings from previous research once again become relevant. We provide suggestive evidence that the estimate of disemployment effect found in DLR averages out negative and positive treatment effects estimated in politically aligned and unaligned samples. Any conclusive results regarding the disemployment effect require the use of staggered rollout designs.

Bibliography

Addison, J. T., Blackburn, M. L. and Cotti, C. D. (2012), 'The effect of minimum wages on labour market outcomes: County-level estimates from the restaurant-and-bar sector', *British Journal of Industrial Relations* **50**(3), 412–435.
URL: *http://dx.doi.org/10.1111/j.1467-8543.2010.00819.x*

Adweek (2014a), 'Havas Worldwide wins TD Ameritrade's account'. Accessed on October 29, 2022.
URL: *https://www.adweek.com/brand-marketing/havas-unit-wins-td-ameritrades-creative-account-157621/*

Adweek (2014b), 'Microsoft set to begin massive cross-discipline agency review'. Accessed on October 29, 2022.
URL: *https://adage.com/article/agency-news/microsoft-set-begin-cross-discipline-agency-review/291047*

Adweek (2016), 'McDonald's launches a creative review to consolidate its ad business in the U.S.'. Accessed on October 29, 2022.
URL: *https://www.adweek.com/agencyspy/mcdonalds-launches-a-u-s-creative-review-to-consolidate-its-account/107714/*

Adweek (2017), 'We hear: Ogilvy, other agencies in talks to handle Macy's brand work'. Accessed on October 29, 2022.
URL: *https://www.adweek.com/agencyspy/we-hear-ogilvy-in-talks-to-handle-macys-brand-work/136043/*

Adweek (2019a), 'Team One appoints TBWA/Chiat/Day creative director to global ECD on Expedia'. Accessed on Novermber 26, 2022.
URL: *https://www.adweek.com/agencyspy/team-one-appoints-tbwachiatday-creative-director-to-global-ecd-on-expedia/154481/*

Adweek (2019b), 'WPP's Garage Team Mazda parts with staff as client reduces scope of work'. Accessed on Novermber 26, 2022.
URL: *https://www.adweek.com/agencyspy/wpps-garage-team-mazda-parts-with-staff-as-client-reduces-scope-of-work/157112/*

Adweek (2019c), 'Report: We Are Unlimited will be folded into DDB Chicago after McDonald's taps W+K'. Accessed on Novermber 26, 2022.
URL: *https://www.adweek.com/agencyspy/report-we-are-unlimited-will-be-folded-into-ddb-chicago-after-mcdonalds-taps-wk/159591/*

Agrawal, D. and Hoyt, W. H. (2014), State Tax Differentials, Cross-Border Commuting, and Commuting Times in Multi-State Metropolitan Areas, CESifo Working

Paper Series 4852, CESifo Group Munich.
URL: *https://ideas.repec.org/p/ces/ceswps/4852.html*

Ali, R. (2020), Effect of airport slots in competition and antitrust policy: Evidence from a recent merger, Working paper.
URL: *https://sites.google.com/view/ratibali/job-market-paper?authuser=0*

Allegretto, S. A., Dube, A. and Reich, M. (2011), 'Do minimum wages really reduce teen employment? accounting for heterogeneity and selectivity in state panel data', *Industrial Relations: A Journal of Economy and Society* **50**(2), 205–240.

Allegretto, S., Dube, A., Reich, M. and Zipperer, B. (2017), 'Credible research designs for minimum wage studies', *ILR Review* pp. 559–592.

Altexsoft (2019), 'Airline reservation systems and passenger service systems: Navitaire, Amadeus Altea, SabreSonic and more'. Accessed on October 29, 2022.
URL: *https://www.altexsoft.com/blog/airline-reservation-systems-passenger-service-systems/*

ANA/4A's (2011), 'Guidelines for agency search. ANA/4A'S White Paper'. Accessed on October 29, 2022.
URL: *https://www.aaaa.org/agselbrief_09252013/*

ANA/4A's (2013), 'Agency selection briefing guidance. ANA/4A'S White Paper'. Accessed on October 29, 2022.
URL: *https://www.aaaa.org/agselbrief_09252013/*

Aobdia, D. (2015), 'Proprietary information spillovers and supplier choice: Evidence from auditors', *Review of Accounting Studies* **20**, 1504–1539.

Asker, J. and Ljungqvist, A. (2010), 'Competition and the structure of vertical relationships in capital markets', *Journal of Political Economy* **118**(3), 599–647.

Bapat, R. and Beg, M. (1989), 'Order statistics for nonidentically distributed variables and permanents', *Sankhyā: The Indian Journal of Statistics, Series A* pp. 79–93.

Bernheim, B. D. and Whinston, M. D. (1985), 'Common marketing agency as a device for facilitating collusion', *The RAND Journal of Economics* pp. 269–281.

Besley, T. and Case, A. (2003), 'Political institutions and policy choices: Evidence from the United States', *Journal of Economic Literature* **41**(1), 7–73.
URL: *https://www.aeaweb.org/articles?id=10.1257/002205103321544693*

Bohn, S. and Santillano, R. (2017), 'Local immigration enforcement and local economies', *Industrial Relations: A Journal of Economy and Society* **56**(2), 236–262.
URL: *http://dx.doi.org/10.1111/irel.12172*

Bosher, M. (2021), 'The marriage of private equity and public accounting'. Accessed on February 13, 2023.

URL: *https://www.accountingtoday.com/opinion/navigating-the-marriage-of-private-equity-and-public-accounting*

Brown, C., Gilroy, C. and Kohen, A. (1982), 'The effect of the minimum wage on employment and unemployment', *Journal of Economic Literature* **20**(2), 487–528.
URL: *http://www.jstor.org/stable/2724487*

Brunner, E., Ross, S. L. and Washington, E. (2011), 'Economics and Policy Preferences: Causal Evidence of the Impact of Economic Conditions on Support for Redistribution and Other Ballot Proposals', *The Review of Economics and Statistics* **93**(3), 888–906.
URL: *https://ideas.repec.org/a/tpr/restat/v93y2011i3p888-906.html*

Callaway, B., Goodman-Bacon, A. and Sant'Anna, P. H. C. (2021), 'Difference-in-differences with a continuous treatment'.
URL: *https://arxiv.org/abs/2107.02637*

Card, D. (1992*a*), 'Do minimum wages reduce employment? A case study of california, 1987-89', *Industrial and Labor Relations Review* **46**(1), 38–54.
URL: *http://www.jstor.org/stable/2524737*

Card, D. (1992*b*), 'Using regional variation in wages to measure the effects of the federal minimum wage', *Industrial and Labor Relations Review* **46**(1), 22–37.
URL: *http://www.jstor.org/stable/2524736*

Card, D. and Krueger, A. B. (1994), 'Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania', *The American Economic Review* **84**(4), 772–793.
URL: *http://www.jstor.org/stable/2118030*

Card, D. and Krueger, A. B. (2000), 'Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania: Reply', *The American Economic Review* **90**(5), 1397–1420.
URL: *http://www.jstor.org/stable/2677856*

Chang, X., Shekhar, C., Tam, L. H. and Yao, J. (2016), 'Industry expertise, information leakage and the choice of m&a advisors', *Journal of Business Finance & Accounting* **43**(1-2), 191–225.

Charles, K. K. and Stephens, Melvin, J. (2013), 'Employment, wages, and voter turnout', *American Economic Journal: Applied Economics* **5**(4), 111–43.
URL: *http://www.aeaweb.org/articles?id=10.1257/app.5.4.111*

Che, Y.-K. (1993), 'Design competition through multidimensional auctions', *The RAND Journal of Economics* pp. 668–680.

Chernozhukov, V., Kasahara, H. and Schrimpf, P. (2021), 'Causal impact of masks, policies, behavior on early covid-19 pandemic in the U.S.', *Journal of Econometrics*

220(1), 23–62.
URL: *https://doi.org/10.1016/j.jeconom.2020.09.003*

Ciliberto, F. and Williams, J. W. (2010), 'Limited access to airport facilities and market power in the airline industry', *The Journal of Law and Economics* **53**(3), 467–495.

Commission, U. S. M. W. S. (1981), *Report of the Minimum Wage Study Commission*, number v. 1 *in* 'Report of the Minimum Wage Study Commission', The Commission.
URL: *https://books.google.com/books?id=nSHtAAAAMAAJ*

Coviello, D., Deserranno, E. and Persico, N. (2022), 'Minimum wage and individual worker productivity: Evidence from a large US retailer', *Journal of Political Economy* **130**(9), 2315–2360.

Dai, W. (2016), Matching with conflicts: An application to the advertising industry, Working paper.
URL: *http://daisywdai.com/s/dai_ad_agency_conflict.pdf*

de Chaisemartin, C., D'Haultfoeuille, X., Pasquier, F. and Vazquez-Bare, G. (2022), 'Difference-in-differences estimators for treatments continuously distributed at every period'.
URL: *https://arxiv.org/abs/2201.06898*

Decarolis, F., Goldmanis, M. and Penta, A. (2020), 'Marketing agencies and collusive bidding in online ad auctions', *Management Science* **66**(10), 4433–4454.

Decarolis, F. and Rovigatti, G. (2021), 'From mad men to maths men: Concentration and buyer power in online advertising', *American Economic Review* **111**(10), 3299–3327.

Denford, T. and Indo, D. (2010), Technical report. Accessed on December 14, 2022.
URL: *https://cdn2.hubspot.net/hubfs/2395501/IDComms_Oct2016/Docs/IDComms_Guide_To_Agency_Remuneration.pdf*

DOJ (2010), 'Horizontal merger guidelines'. Accessed on Novermber 26, 2022.
URL: *https://www.justice.gov/atr/horizontal-merger-guidelines-08192010*

Doucouliagos, H. and Stanley, T. D. (2009), 'Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis', *British Journal of Industrial Relations* **47**(2), 406–428.
URL: *https://ideas.repec.org/a/bla/brjirl/v47y2009i2p406-428.html*

Draganska, M., Mazzeo, M. and Seim, K. (2009), 'Beyond plain vanilla: Modeling joint product assortment and pricing decisions', *Quantitative Marketing and Economics* **7**(2), 105–146.

Dube, A., Lester, T. W. and Reich, M. (2010), 'Minimum wage effects across state borders: Estimates using contiguous counties', *The review of economics and statistics* **92**(4), 945–964.

Dube, A., Lester, T. W. and Reich, M. (2016), 'Minimum wage shocks, employment flows, and labor market frictions', *Journal of Labor Economics* **34**(3), 663–704.

Dube, A., Naidu, S. and Reich, M. (2007), 'The economic effects of a citywide minimum wage', *Industrial and Labor Relations Review* **60**(4), 522–543.
URL: *http://www.jstor.org/stable/25249108*

Dube, A. and Zipperer, B. (2015), 'Pooling multiple case studies using synthetic controls: An application to minimum wage policies'.

Duval, M. (2021), 'New data on the whole costs of agency pitching'. Accessed on December 14, 2022.
URL: *https://info.duvalpartnership.com/blog/new-data-on-the-whole-costs-of-agency-pitching*

Dvorkin, M. and Shell, H. (2016), 'The recent evolution of U.S. local labor markets', *Economic Synopses* **15**.

Fukui, H. (2019), 'How do slot restrictions affect airfares? New evidence from the US airline industry', *Economics of Transportation* **17**, 51–71.

Gal-Or, E. (1991), 'A common agency with incomplete information', *The RAND Journal of Economics* pp. 274–286.

GAO (2012), 'Slot-controlled airports: FAA's rules could be improved to enhance competition and use of available capacity. report gao-12-902'. Accessed on February 13, 2023.
URL: *https://www.businessinsider.com/coronavirus-faa-suspending-slot-requirements-at-us-airports-2020-3*

Gerber, A. S. and Huber, G. A. (2009), 'Partisanship and economic behavior: Do partisan differences in economic forecasts predict real economic behavior?', *American Political Science Review* **103**(3), 407–426.
URL: *https://doi.org/10.1017/s0003055409990098*

Goldsmith-Pinkham, P., Hull, P. and Kolesár, M. (2022), Contamination bias in linear regressions, Working paper.
URL: *https://doi.org/10.3386/w30108*

Goodman-Bacon, A. (2021), 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics* **225**(2), 254–277.
URL: *https://doi.org/10.1016/j.jeconom.2021.03.014*

Horsky, S. (2006), 'The changing architecture of advertising agencies', *Marketing Science* **25**(4), 367–383.

Hristakeva, S. and Mortimer, J. H. (2021), 'Price dispersion and legacy discounts in the national television advertising market', *Available at SSRN 3911469* .
URL: *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3911469*

Huang, R. R. (2008), 'Evaluating the real effect of bank branching deregulation: Comparing contiguous counties across US state borders', *Journal of Financial Economics* **87**(3), 678–705.
URL: *https://doi.org/10.1016/j.jfineco.2007.01.004*

Jardim, E., Long, M. C., Plotnick, R., van Inwegen, E., Vigdor, J. and Wething, H. (2022*a*), 'Minimum-wage increases and low-wage employment: Evidence from seattle', *American Economic Journal: Economic Policy* **14**(2), 263–314.
URL: *https://www.aeaweb.org/articles?id=10.1257/pol.20180578*

Jardim, E. S., Long, M. C., Plotnick, R., van Inwegen, E., Vigdor, J. L. and Wething, H. (2022*b*), Boundary discontinuity methods and policy spillovers, Working Paper 30075, National Bureau of Economic Research.

Jha, P., Neumark, D. and Rodriguez-Lopez, A. (2022), What's across the border? re-evaluating the cross-border evidence on minimum wage effects, Working Paper 15282, IZA.
URL: *https://doi.org/10.2139/ssrn.4114866*

Joint Economic Committee Republicans (1995), 50 years of research on the minimum wage, Technical report.
URL: *https://www.jec.senate.gov/public/_cache/files/c876c468-ffca-47ed-9468-7193d734bde9/50-years-of-research-on-the-minimum-wage—february-15-1995.pdf*

Jones, J. P. (1998), *How advertising works: The role of research*, SAGE Publications, Incorporated.

Kahn, M. E. and Mansur, E. T. (2013), 'Do local energy prices and regulation affect the geographic concentration of employment?', *Journal of Public Economics* **101**, 105 – 114.
URL: *http://www.sciencedirect.com/science/article/pii/S0047272713000509*

Katz, L. F. and Krueger, A. B. (1992), 'The effect of the minimum wage on the fast-food industry', *Industrial and Labor Relations Review* **46**(1), 6–21.
URL: *http://www.jstor.org/stable/2524735*

Keele, L. J. and Titiunik, R. (2015*a*), 'Geographic boundaries as regression discontinuities', *Political Analysis* **23**(1), 127–155.
URL: *https://doi.org/10.1093/pan/mpu014*

Keele, L. and Titiunik, R. (2015*b*), 'Natural experiments based on geography', *Political Science Research and Methods* **4**(1), 65–95.
URL: *https://doi.org/10.1017/psrm.2015.4*

King, C., Silk, A. J. and Ketelhöhn, N. (2003), 'Knowledge spillovers and growth in the disagglomeration of the US advertising-agency industry', *Journal of Economics & Management Strategy* **12**(3), 327–362.

Kong, D., Qin, N. and Xiang, J. (2021), 'Minimum wage and entrepreneurship: Evidence from China', *Journal of Economic Behavior & Organization* **189**, 320–336.
URL: *https://www.sciencedirect.com/science/article/pii/S0167268121002882*

Liu, X., Andris, C. and Desmarais, B. A. (2019), 'Migration and political polarization in the U.S.: An analysis of the county-level migration network', *PLOS ONE* **14**(11), e0225405.
URL: *https://doi.org/10.1371/journal.pone.0225405*

Luttmann, A. (2019), 'Are passengers compensated for incurring an airport layover? Estimating the value of layover time in the US airline industry', *Economics of Transportation* **17**, 1–13.

Macrotrends (2022), 'Interpublic Group of profit margin 2010-2022 | IPG'. Accessed on December 14, 2022.
URL: *https://www.macrotrends.net/stocks/charts/IPG/interpublic-group-of/profit-margins*

Margalit, Y. (2013), 'Explaining social policy preferences: Evidence from the Great Recession', *American Political Science Review* **107**(1), 80–103.
URL: *https://doi.org/10.1017/s0003055412000603*

Markovich, Z. and White, A. (2022), 'More money, more turnout? Minimum wage increases and voting', *The Journal of Politics* **84**(3), 1834–1838.
URL: *https://doi.org/10.1086/716291*

McCartney, W. B., Orellana, J. and Zhang, C. (2021), "Sort Selling": Political Polarization and Residential Choice, Working Papers 21-14, Federal Reserve Bank of Philadelphia.
URL: *https://ideas.repec.org/p/fip/fedpwp/93604.html*

Meinhofer, A., Witman, A., Hinde, J. and Simon, K. (2021), Marijuana liberalization policies and perinatal health, Working paper.
URL: *https://doi.org/10.3386/w29296*

Muravyev, A. and Oshchepkov, A. (2016), 'The effect of doubling the minimum wage on employment: Evidence from Russia', *IZA Journal of Labor & Development* **5**(1).
URL: *https://doi.org/10.1186/s40175-016-0051-0*

Naidu, S. (2012), Suffrage, Schooling, and Sorting in the Post-Bellum U.S. South, NBER Working Papers 18129, National Bureau of Economic Research, Inc.
URL: *https://ideas.repec.org/p/nbr/nberwo/18129.html*

Nall, C. (2015), 'The political consequences of spatial policies: How interstate highways facilitated geographic polarization', *The Journal of Politics* **77**(2), 394–406.
URL: *https://doi.org/10.1086/679597*

Neumark, D., Salas, J. I. and Wascher, W. (2014*a*), 'More on recent evidence on the effects of minimum wages in the United States', *IZA Journal of Labor policy* **3**(1), 24.

Neumark, D., Salas, J. M. I. and Wascher, W. (2014*b*), 'Revisiting the minimum wage–employment debate: Throwing out the baby with the bathwater?', *ILR Review* **67**(3 Suppl), 608–648.

Neumark, D. and Wascher, W. (2008*a*), 'The history of the minimum wage in the United States', pp. 9–36.
URL: *http://www.jstor.org/stable/j.ctt5hhh46.5*

Neumark, D. and Wascher, W. (2017), 'Reply to "Credible research designs for minimum wage studies"', *ILR Review* **70**(3), 593–609.
URL: *https://doi.org/10.1177/0019793917698429*

Neumark, D. and Wascher, W. L. (2007), 'Minimum wages and employment', *Foundations and Trends® in Microeconomics* **3**(1-2), 1–182.
URL: *http://dx.doi.org/10.1561/0700000015*

Neumark, D. and Wascher, W. L. (2008*b*), *Minimum Wages*, MIT Press.
URL: *http://www.jstor.org/stable/j.ctt5hhh46*

Pallini, T. (2020), 'FAA moves to prevent 'ghost flights' of near-empty planes in the US by easing rules that put them at risk of losing their spots at airports'. Accessed on February 13, 2023.
URL: *https://www.gao.gov/products/gao-12-902*

Peterson, E. B., Neels, K., Barczi, N. and Graham, T. (2013), 'The economic cost of airline flight delay', *Journal of Transport Economics and Policy (JTEP)* **47**(1), 107–121.

Posner, D. N. (2004), 'The political salience of cultural difference: Why Chewas and Tumbukas are allies in Zambia and adversaries in Malawi', *American Political Science Review* **98**(4), 529–545.
URL: *https://doi.org/10.1017/s0003055404041334*

Provokemedia (2016), 'LG's Ken Hong: 'It's very hard to unseat WPP''. Accessed on October 29, 2022.
URL: *https://www.provokemedia.com/latest/article/lg's-ken-hong-'it's-very-hard-to-unseat-wpp'*

Rambachan, A. and Roth, J. (2021), An honest approach to parallel trends, Working

paper.
URL: *https://scholar.harvard.edu/files/jroth/files/honestparalleltrends_main.pdf*

Rogan, M. (2014), 'Too close for comfort? The effect of embeddedness and competitive overlap on client relationship retention following an acquisition', *Organization Science* **25**(1), 185–203.

Rohlin, S. M. and Thompson, J. P. (2013), 'The effect of state and local sales taxes on employment at state borders', *Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board* .

Roth, J. (2022), 'Pretest with caution: Event-study estimates after testing for parallel trends', *American Economic Review: Insights* **4**(3), 305–22.
URL: *https://www.aeaweb.org/articles?id=10.1257/aeri.20210236*

Roth, J., Sant'Anna, P. H. C., Bilinski, A. and Poe, J. (2022), 'What's trending in difference-in-differences? A synthesis of the recent econometrics literature'.
URL: *https://arxiv.org/abs/2201.01194*

Schmitt, J. (2015), 'Explaining the small employment effects of the minimum wage in the United States', *Industrial Relations: A Journal of Economy and Society* **54**(4), 547–581.
URL: *http://dx.doi.org/10.1111/irel.12106*

Silk, A. J. (2012), 'Conflict policy and advertising agency-client relations: The problem of competing clients sharing a common agency', *Foundations and Trends® in Marketing* **6**(2), 63–149.

Silk, A. J. and Berndt, E. R. (1993), 'Scale and scope effects on advertising agency costs', *Marketing Science* **12**(1), 53–72.

Silk, A. J. and Berndt, E. R. (1994), 'Costs, institutional mobility barriers, and market structure: Advertising agencies as multiproduct firms', *Journal of Economics & Management Strategy* **3**(3), 437–480.

Silk, A. J. and Berndt, E. R. (2004), 'Holding company cost economies in the global advertising and marketing services business', *Review of Marketing Science* **2**(1).

Silk, A. J. and King III, C. (2013), 'How concentrated is the US advertising and marketing services industry? Myth versus reality', *Journal of Current Issues & Research in Advertising* **34**(1), 166–193.

Simonovits, G. and Bor, A. (2021), Stability and change in the opinion-policy relationship: Evidence from state minimum wage laws, Working paper.
URL: *https://files.osf.io/v1/resources/3ye4b/providers/osfstorage/5f22b81a5f705a00586188b8*

Simonovits, G., Guess, A. M. and Nagler, J. (2019), 'Responsiveness without representation: Evidence from minimum wage laws in U.S. states', *American Journal*

of Political Science **63**(2), 401–410.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12412*

Skolnik, S. (2022), 'ABA sides against opening law firms up to new competition'. Accessed on February 13, 2023.
URL: *https://news.bloomberglaw.com/business-and-practice/aba-sides-against-opening-law-firms-up-to-new-competition*

Snider, C. and Williams, J. W. (2015), 'Barriers to entry in the airline industry: A multidimensional regression-discontinuity analysis of AIR-21', *Review of Economics and Statistics* **97**(5), 1002–1022.

Solomon, J., Rhode, D. and Wanless, A. (2020), How reforming rule 5.4 would benefit lawyers and consumers, promote innovation, and increase access to justice, Technical report.
URL: *https://law.stanford.edu/publications/how-reforming-rule-5-4-would-benefit-lawyers-and-consumers-promote-innovation-and-increase-access-to-justice/*

Sun, L. and Abraham, S. (2021), 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics* **225**(2), 175–199.
URL: *https://doi.org/10.1016/j.jeconom.2020.09.006*

Swaroop, P., Zou, B., Ball, M. O. and Hansen, M. (2012), 'Do more US airports need slot controls? A welfare based approach to determine slot levels', *Transportation Research Part B: Methodological* **46**(9), 1239–1259.

Tan, G. and Yilankaya, O. (2006), 'Equilibria in second price auctions with participation costs', *Journal of Economic Theory* **130**(1), 205–219.

Vickrey, W. (1961), 'Counterspeculation, auctions, and competitive sealed tenders', *The Journal of finance* **16**(1), 8–37.

Villas-Boas, J. M. (1994), 'Sleeping with the enemy: Should competitors share the same advertising agency?', *Marketing Science* **13**(2), 190–202.

von Nordenflycht, A. (2011), 'Firm size and industry structure under human capital intensity: Insights from the evolution of the global advertising industry', *Organization Science* **22**(1), 141–157.

Wooley, D. (2018), 'Is it time marketers rethink agency competitive conflicts of interest?'. Accessed on September 12, 2020.
URL: *https://www.trinityp3.com/2018/02/marketers-agency-conflicts-of-interest/*

Appendices

A Monte Carlo Simulations

Our modeling choices rely on three stylized facts about the market for creative agencies of record: (i) presence of competitor avoidance at the agency level; (ii) lack of (or less strong) competitor avoidance at the HC level; (iii) bid coordination by holding companies. In this appendix, we use Monte Carlo simulations to investigate whether these stylized facts are consistent with random matching patterns that would arise in a market with the observed number of agencies, HCs, and accounts in each category.

We measure the overall amount of competitor avoidance in the market by calculating the share of relationships in conflict. For example, consider a simple market with two agencies and two accounts from the same product market. Each account hires a single AOR agency, so the total number of relationships in this market is two. If the two accounts hire the same agency, the share of relationships in conflict is one; if different, then it is zero. The amount of competitor avoidance at the HC level follows a similar logic, but with agency-account relationships aggregated up to the HC level.

We measure bid coordination by HCs by calculating the share of co-bidding instances (pairs) for a given number of reviews and HCs' bidding agencies. For example, suppose we observe two reviews that have three and four bidding agencies participating in them. The bidding agencies in the first review are owned different HCs, whereas the second review has three bidding agencies from the same HC. The total number of agency pairs in each review is 3 and 6. The first review has no co-bidding instances, whereas the second review has two. Therefore, the share of co-bidding instances in this example is 2/9.

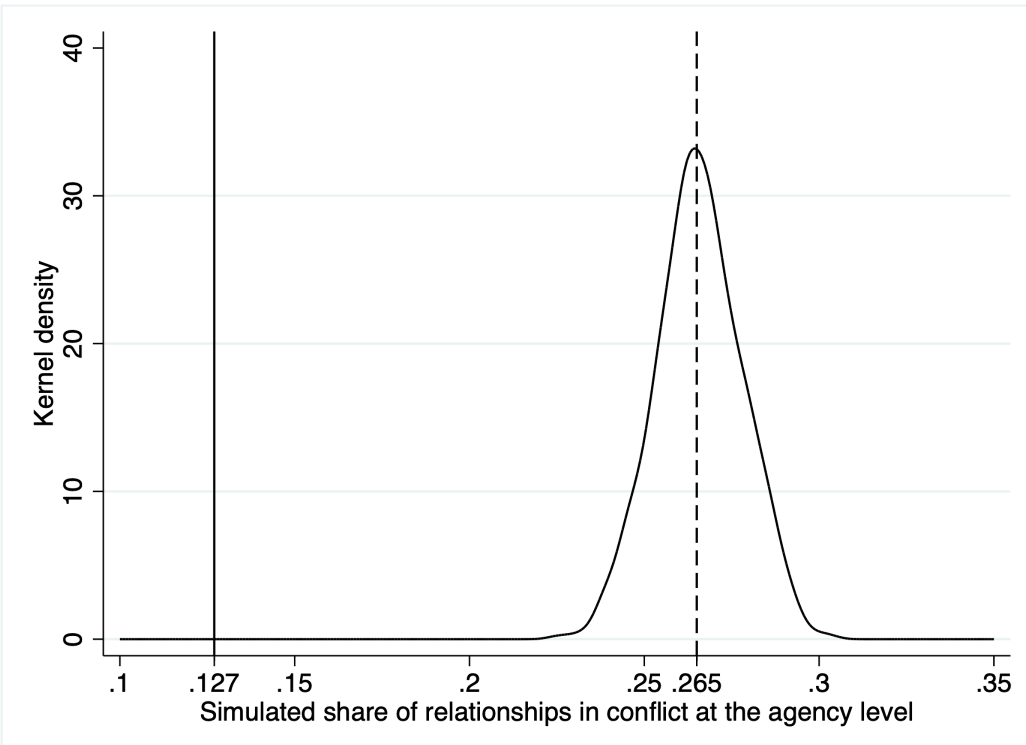
Our null hypotheses are:

1. The observed share of relationships in conflict at the agency level reflects a random matching, i.e. the simulated mean share of relationships in conflict is equal to the observed share of relationships in conflict.

2. The observed share of relationships in conflict at the holding company level reflects a random matching, i.e. the simulated mean share of relationships in conflict is equal to the observed share of relationships in conflict.

3. The observed share of co-bidding instances by agencies from the same holding companies reflects a random matching, i.e. the simulated mean share of co-bidding instances is equal to the observed share of co-bidding instances.

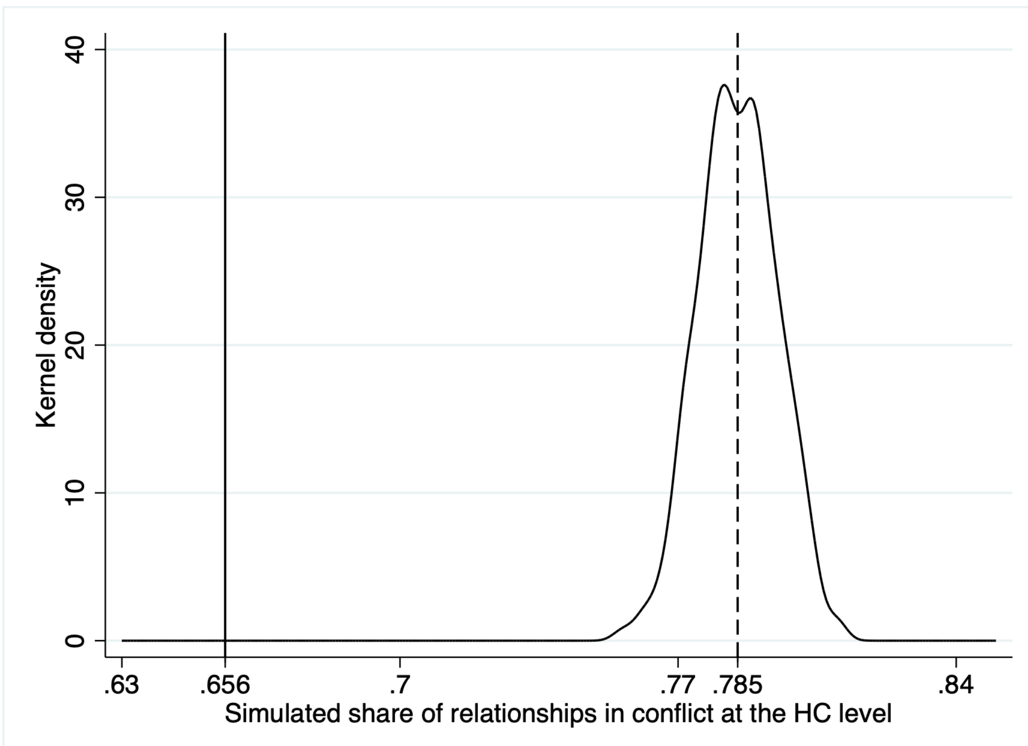
FIGURE 3.2 – ESTIMATED KERNEL DENSITY OF THE SHARE OF RELATIONSHIPS IN CONFLICT AT THE AGENCY LEVEL



In order to test null hypotheses 1 and 2, we randomly assign accounts to agencies or holding companies, respecting the observed number of clients served by each agency or each holding company in every year. We repeat such random assignments 1,000 times and estimate the implied kernel density. Figures 3.2 and 3.3 illustrate the estimated kernel densities at the agency and HC level along with the simulated means (dashed) and the observed values (solid). We then calculate the simulated mean share of relationships in conflict and perform a t-test of equality of the simulated and observed means. For agencies, the null hypothesis about the equality of means is rejected with p-value of 0.0000 ($t = 357.16$). This simulations shows that competitor conflicts at the agency level are less frequent than the randomness would predict, providing suggestive model-free evidence for competitor avoidance.

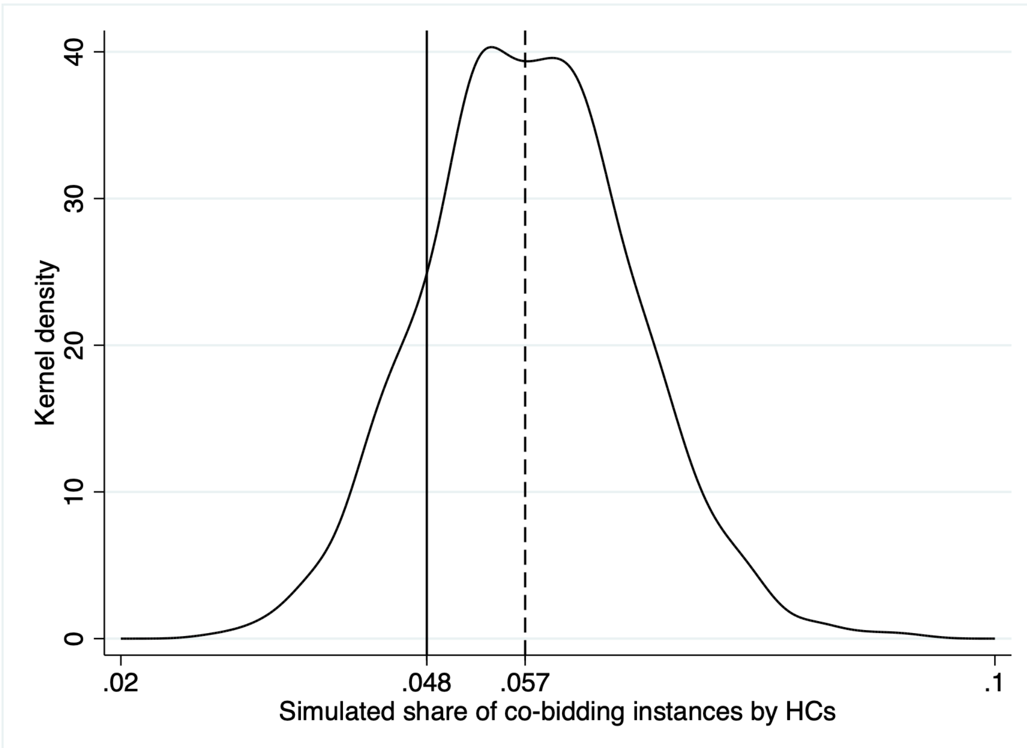
For HCs, the null hypothesis about the equality of means is rejected with p-value of 0.0000 ($t = 413.77$). Even though the share of relationships in conflict at the holding company level is smaller than predicted by the randomness, our estimates in Table 1.5 do not support the implication that this pattern is due to competitor avoidance at the HC level.

FIGURE 3.3 – ESTIMATED KERNEL DENSITY OF THE SHARE OF RELATIONSHIPS IN CONFLICT AT THE HC LEVEL



In order to test null hypothesis 3, we randomly assign bidders to account reviews, respecting the number of holding-company bidders observed in each review. We repeat such random assignments 1,000 times and estimate the implied kernel density. Figure 3.4 illustrates the estimated kernel density. The simulated mean is denoted with a dashed line and the observed mean by a solid. We then perform a t-test of equality of the simulated and observed means and reject the null with p-value of 0.0000 ($t = 30.36$). This simulation shows that co-bidding instances are less frequent in the data than the randomness would predict, providing suggestive model-free evidence for bid coordination within HCs.

FIGURE 3.4 – ESTIMATED KERNEL DENSITY OF THE SHARE OF CO-BIDDING INSTANCES.



B The List of Top-28 Domestic Airports by Passenger Enplanements

In alphabetical order of the airports' three-letter codes: Atlanta, GA (ATL); Boston, MA (BOS); Baltimore, MD (BWI); Charlotte, NC (CLT); Washington, DC (DCA); Denver, CO (DEN); Dallas-Fort Worth, TX (DFW); Detroit, MI (DTW); Newark, NJ (EWR); Fort Lauderdale, FL (FLL); Dulles, VA (IAD); Houston, TX (IAH); Queens, NY (JFK); Las Vegas, NV (LAS); Los Angeles, CA (LAX); Queens, NY (LGA); Orlando, FL (MCO); Chicago, IL (MDW); Miami, FL (MIA); Minneapolis-Saint Paul, MN (MSP); Chicago, IL (ORD); Philadelphia, PA (PHL); Phoenix, AZ (PHX); San Diego, CA (SAN); Seattle-Tacoma, WA (SEA); San Francisco, CA (SFO); Salt Lake City, UT (SLC); Tampa, FL (TPA).

C Shifts in Operations between NYC Airports

In 2017, both JFK and LaGuardia underwent runway reconstructions that temporarily reduced their air traffic capacity. If the FAA preemptively lifted slot control at Newark in order to allow the affected carriers to shift operations from JFK and LGA, our proxy variable for slot burning – usage of small aircraft in peak and off-peak slot periods – could be confounded by patterns of aircraft usage spilt over from JFK and LGA.

In order to test for evidence of spillover operations, we correlate the change in the frequency of scheduled flights by route between 2016 and 2017 and between 2016 and 2018.[70] We find no evidence of shifts in JFK's operations. The pairwise coefficients of correlation between the changes in scheduled flight frequencies at JFK and Newark are insignificant -0.0491 in 2017 and insignificant 0.0495 in 2018. However, we do find that the LGA routes that experienced a decrease in the number of scheduled flights in 2017 or 2018 (relative to 2016) tend to experience an increase in scheduled frequency at Newark, with the correlation coefficients of -0.2140 significant at 5% in 2017 and -0.2754 significant at 5% in 2018.

[70]Figure 2.1 shows reduction in delays in 2018. This could be due to the fact that the reconstructed runways returned to operating at full capacity, or because it takes more than a year to shift operations between airports. For this reason, we study changes in scheduled flight frequencies in 2018 as well.

TABLE 3.7 – CHANGE IN THE NUMBER OF SCHEDULED FLIGHTS ON ROUTES THAT EXPERIENCED A SIGNIFICANT DECREASE AT JFK AND LGA AND AN INCREASE AT NEWARK

Airport	Newark		JFK		LaGuardia	
	Δ2017	Δ2018	Δ2017	Δ2018	Δ2017	Δ2018
Québec City, QC	216	472	-582	-1,156	–	–
Sarasota, FL	122	705	372	344	-685	-643
Jacksonville, FL	-84	220	6	70	-1,090	-1,316
Fort Myers, FL	593	556	-252	-531	-823	-1,093
Nantucket, MA	-34	156	-117	59	-20	262
Indianapolis, IN	395	889	559	1,029	-1,307	-1,054
Grand Rapids, MI	99	202	–	–	-169	348

We further investigate what routes experienced sizeable decrease in scheduled frequency (more than 15%) at JFK and LaGuardia and an increase in scheduled frequency at Newark. We identify seven such routes and document them in Table 3.7 above. None of these routes are a part of the sample of airports we use to test for slot burning, therefore we believe that possible shifts in operations did not affect patterns of aircraft usage at Newark in any spurious manner.

D County Maps

FIGURE 3.5 – CONTIGUOUS COUNTIES ALONG BORDERS THAT HAD AT LEAST ONE MINIMUM WAGE DIFFERENTIAL: DLR VS. POLITICALLY ALIGNED SAMPLE