Forthcoming: *Journal of Urban Economics*

**Agglomeration Economies and the Built Environment:
Evidence from Specialized Buildings and Anchor Tenants**

Crocker H. Liu
Robert A. Beck Professor of Hospitality Financial Management
Nolan School of Hotel Administration
Cornell University
chl62@cornell.edu

Stuart S. Rosenthal
Paul Rubacha Department of Real Estate
Cornell University
sr784@cornell.edu

William C. Strange
SmartCentres Professor of Real Estate
Rotman School of Management
University of Toronto, Toronto Ontario M5S 3E6, Canada
wstrange@rotman.utoronto.ca

February 24, 2024

We thank Kristian Behrens and two anonymous referees for helpful comments and Joaquin Urrego and Maeve Maloney for excellent research assistance. We are also grateful for helpful comments on earlier versions of the paper from participants in sessions at the AREUEA-National Conference, European UEA, Norway Central Bank, Research Institute of Economy, Trade and Industry (RIETI), University of California–Irvine, University of Connecticut, University of North Carolina, University of Pennsylvania, Singapore Management University, Jinan University and University of Wisconsin. Strange acknowledges financial support from the Social Sciences and Humanities Research Council of Canada and the Centre for Real Estate at the Rotman School of Management. Much of this work took place while Rosenthal was at Syracuse University as the Maxwell Advisory Board Professor of Economics in the Department of Economics and Center for Policy Research at Syracuse University. Support from the Center for Policy Research is gratefully acknowledged. Errors are the responsibility of the authors.

Abstract

Previous work on agglomeration economies ignores the built environment. This paper shows that the built environment matters, especially for commercial sectors that dominate city centers. Buildings are specialized beyond random assignment, in part because externality-generating anchor tenants skew a building's other tenants towards the anchor's industry. An anchor elsewhere on the blockface has a much weaker effect, and one that is weaker still if across the street, suggesting rapidly attenuating agglomeration economies. Attenuation is pronounced for retail and information-oriented office industries but is absent for manufacturing. Building managers have incentives and capacities to partly internalize local externalities, contributing to urban productivity.

JEL Codes: O18 (Economic Development: Urban, Rural, Regional), R30 (Real Estate Markets, Spatial Production Analysis, and Firm Location: General)

Key Words: Agglomeration, commercial real estate, anchor tenants, built environment.

I. Introduction

Within a few blocks of the New York Stock Exchange, most buildings do not contain any employment in finance, even among large buildings. This is surprising, since Wall Street is known as a district that specializes in finance, in part because of agglomeration economies. Our resolution of this puzzle relies on a previously unremarked fact: buildings specialize in the same way that neighborhoods do, even within neighborhoods that themselves specialize. This fact, in turn, sheds light on the nature of agglomeration economies, in particular attenuation, and the role of the built environment in urban economic activity.

Our focus on the role of the built environment contrasts with prior studies of agglomeration that have ignored buildings as spatial units into which companies and employment are organized. Early research defined agglomeration as occurring at the regional or metropolitan scale. Glaeser et al. (1992) and Henderson et al. (1995) are two notable examples, while Combes et al (2012), De la Roca and Puga (2017), Eckert et al (2022) and Jales et al (2022) are more recent.[1] Other research has worked at the neighborhood scale. Rosenthal and Strange (2003, 2005, 2020), Arzaghi and Henderson (2008) and Ahlfeldt et al (2015) are examples. Neighborhoods in these papers are typically defined by distance. In our alternative approach, the built environment is central, with agglomeration defined in a way that takes explicit account of activity within individual buildings, along blockfaces, and across streets.[2]

Why consider the built environment? In short, because it matters. Meeting with individuals inside one's own building entails walking down the hall or at most an elevator ride. Meeting with individuals elsewhere may require dressing for weather and navigating city traffic. Individuals operating within the same building are also far more likely to experience spontaneous interactions as they enter and exit

[1] For reviews of the agglomeration literature, see Rosenthal and Strange (2004), Duranton and Puga (2004), Behrens and Robert-Nicoud (2015), and Combes and Gobillon (2015).

[2] Liu et al (2018, 2020) consider spatial patterns of activity within commercial buildings, and in that sense also focus on the built environment. However, the focus in those papers is different. Drawing on roughly 100 offering memorandums for tall commercial buildings around the U.S., Liu et al (2018) model and estimate vertical rent patterns and vertical sorting of different industries. Liu et al (2020) consider analogous patterns for employment density. In contrast, the present paper draws on over 100,000 buildings across five cities and considers a different set of questions using a completely different research design.

through the same doorway, ride the same elevators, and have lunch in the same ground-floor restaurants. The tendency to interact more with others within rather than outside of one's building is well-known in the commercial sector and in information-oriented industries like academia. The importance of the built environment to interactions is also central to architecture and design.[3] For these reasons, it is likely that buildings matter as spatial units and that productivity spillovers decline discretely upon stepping outside of one's building. In this situation, tenants have incentives to choose particular buildings that offer attractive business environments, while building owners have both the incentives and capacities to manage spillovers that take place within buildings. In sum, while abstracting from the built environment makes sense in some situations, the built environment matters to both tenants and building owners, especially for commercial activities in dense downtowns.

To motivate the design of our empirical analysis, we first develop a theoretical model of building employment composition in the presence of spillovers. In the model, tenants are heterogeneous in two ways. First, tenants operate in different industries, which impacts the agglomeration economies that a tenant might enjoy in a particular location. Second, some large tenants, known in the commercial real estate industry as anchors, generate particularly valuable externalities that attract smaller tenants to the building. The importance of anchors is well-understood in the industry.[4] Anchors receive substantial rent discounts, while small tenants pay premia. Furthermore, anchor tenants frequently sign leases with "go dark" clauses that require the firm to operate, while small tenants are sometimes able to break a lease when the anchor departs. The building owner's objective is to select a mix of tenants, anchor and small, that maximizes aggregate rent, which depends on tenant profit. The key predictions of the model are that productivity spillovers encourage building-level specialization and that the presence of an externality-

[3] This is widely recognized in architecture where building design is often thought to affect the quality of social interactions. See, for example, "The Architecture of Social Interaction," in ArchDaily, https://www.archdaily.com/945172/the-architecture-of-social-interaction. It is also found in the writings of Jane Jacobs and her successful opposition of a proposed highway through Greenwich Village in the 1950s (see https://www.theguardian.com/cities/2016/apr/28/story-cities-32-new-york-jane-jacobs-robert-moses).

[4] A Google search on "Anchor Tenant" yields over 8.5 million hits. The vast majority of these are from industry sources commenting on the nature and benefits of anchor tenants, and strategies to attract and retain them.

generating anchor in a building will lead the building to specialize in the anchor's industry to a greater degree.

The model suggests an empirical design that allows us to estimate the equilibrium relationships that govern a building's composition. Central to our strategy, we evaluate how building composition differs when an anchor establishment is present in the building, elsewhere on the blockface, and/or across the street. In each case, anchors are also distinguished by industry. Given the assumption that anchors have disproportionate appeal for smaller tenants in their own building, the patterns we document reveal important features of agglomeration economies that contribute to productivity in city centers. We use establishment-level data from Dun and Bradstreet for five large cities to estimate our models, including New York (the five boroughs), Los Angeles (Los Angeles County), Chicago (Cook County), San Francisco (San Francisco County), and Washington DC (District of Columbia). The data allow us to determine the scale and composition of employment in every commercial building in each of these cities. We focus on three types of industries.

Our first industry is Retail, which relies heavily on foot traffic and is sensitive to shopping externalities. Given previous evidence that retailers profit from close proximity to other retail outlets, retailers are likely to be especially sensitive to the presence of a retail anchor tenant, as in related work on shopping malls (e.g. Brueckner, 1993; Konishi and Sandfort, 2003; Pashigian and Gould (1998); Gould et al, 2005).

The nature of local productivity spillovers is different for information-oriented office industries. Here, learning from others may be facilitated by opportunities for face-to-face contact that are enhanced by location in the same building or at least on the same street (e.g., Rosenthal and Strange, 2020). To consider this part of the economy, we include the more information-oriented parts of FIRE (Finance, Insurance and Real Estate), which we refer to as "Finance". Also included in this group are Advertising, Law, Software and Data Processing, and Engineering and Management. Previous evidence from Rosenthal and Strange (2001), Arzaghi and Henderson (2008), Bosquet and Combes (2017), Liu et al (2020), and Sandvik et al (2020) suggests that clustering in information-oriented industries is likely to be valued.[5] But there is no evidence to date of whether this actually occurs at the building and/or block level, or how the magnitude of such effects might compare to spillovers from shopping externalities in Retail.

We also focus on Manufacturing, for which exporting to distant markets is common. Although ideas certainly spread across manufacturers, manufacturing does not typically require the sort of face-to-face interactions of more information-intensive industries or have reason to benefit from shopping externalities associated with foot traffic (see Rosenthal and Strange, 2001). For these reasons, Manufacturing provides a valuable comparison to Retail and the information-oriented office sector.

Drawing on these three industry groups, several core results emerge. The first is that *buildings are specialized beyond what would occur with random assignment*. This is established through three related exercises. We first show that the composition of employment is correlated among nearby buildings, but the correlations are low. This suggests that some degree of specialization within buildings is common but does not shed light on why. Focusing next on New York's financial district, Monte Carlo evidence, arising from a research design that respects the size distribution of tenants and buildings, confirms that building-level specialization is nonrandom. A more systematic analysis of building-level employment composition follows, drawing on our full five-city sample. Estimates from block fixed effect models, building fixed effect models, and industry-stratified Tobit models all further confirm the nonrandom nature of building specialization.

The second core finding is that *the presence of an anchor establishment skews the rest of the building towards the anchor's industry by 5-25%*. Additional estimates indicate that this relationship is much weaker between industries in the same building. This echoes previous findings (e.g. Rosenthal and Strange, 2003) that agglomeration spillovers are stronger within a given industry as compared to across industries.

The third main finding is that *anchor effects attenuate sharply upon leaving a building*, consistent with the rapid attenuation of agglomeration economies. Within building effects are bigger than effects on

[5] Catalini (2018) is also relevant. It shows that innovative matches are more likely to arise when both partners are located in the same building.

the same blockface. For Retail and the information-oriented office sector, the latter is roughly an order of magnitude smaller. The effects across the street are smaller still. It is clear that the draw between an own-industry anchor tenant and smaller tenants is much weaker upon leaving the target building. This suggests that business owners care about the specific building they are in, even relative to other buildings on the same city block.

The fourth major finding is that *anchor effects differ across industries*. Own-industry anchor effects in the target building are especially strong for Retail, very strong for information-oriented office industries, and important but less strong for Manufacturing. Attenuation of these effects also varies across industries. Anchor effects attenuate sharply for Retail and the information-oriented office industries upon moving outside of the target building. Attenuation is absent, however, for Manufacturing. For this industry, anchor effects are similar regardless of whether the anchor is in the target building, elsewhere on the blockface, or across the street.

Our core results are robust to a variety of model specifications. The primary obstacle to identifying the causal effect of an anchor on a building's composition is that both anchors and smaller tenants may be attracted by characteristics of the building or neighborhood that are unobserved to the econometrician. To address this, we control extensively for both building and neighborhood characteristics, differentiating between characteristics of the building, the blockface, across the street, and the neighborhood. We also estimate using, variously, zipcode, block and building fixed effects. The qualitative patterns and even the magnitudes of our estimates are nearly identical across specifications, suggesting that the controls are effective. Finally, the key results hold for a model with sales per worker as the dependent variable, consistent with the model's predictions.

The different patterns for our three industry groupings have implications for microfoundations. For Retail, it appears that localized shopping externalities, as would be associated with foot traffic, are important and decline notably upon leaving a building. This is of course part of the motivation for shopping malls (e.g. Bruckner, 1993). For the information-oriented segment of the office sector, evidence suggests that face-to-face interactions, as would be facilitated by location in a common building, are also important, and that the tendency for such interactions also drops off sharply upon leaving a building. The absence of attenuation for Manufacturing is different. In this instance, the pattern does not support face-to-face interaction across tenants as an important driver of productivity. For this industry, locating within a given neighborhood appears to be sufficient to achieve location-specific goals.

As a group, our results show that agglomeration economies in dense, commercial districts are mediated by the built environment. Within a building is different than being on the same blockface. And on the same blockface is different than across the street. These patterns and the microfoundations they illuminate are hidden in approaches that consider only colocation in a metropolitan area or even crow-flies distance (e.g. Rosenthal and Strange, 2003 and 2008; Arzaghi and Henderson, 2008). This paper's attenuation result suggests that building owners can manage the part of the agglomeration externalities that is within-building.

The rest of the paper is organized as follows. Section II describes the theoretical model, Section III discusses the data, Section IV presents evidence that buildings are specialized. Section V establishes the importance of the built environment for the nature of agglomeration economies. Section VI concludes.

II. Model

A. Primitives

This section presents a model of the tenant composition of a commercial building. The implications of the model will guide the empirical work to follow. There are two classes of agents in the model, the owners of commercial buildings and the tenants who locate in them. The tenants are differentiated by industry. They are also differentiated according to their impacts on the building's business environment. Those with large impact are referred to as "anchors," while ordinary tenants are referred to as "nonanchors." Building owners choose how to allocate the building's fixed amount of space among the tenants to maximize aggregate rent subject to participation constraints that determine the rent that tenants are willing to pay. This requires that space be allocated to equalize marginal revenue across tenant types. The key implications of the model are that buildings tend to specialize and anchor tenants in a building skew the composition of the building's smaller establishments towards the anchor's industry. These are examined in the paper's empirical sections.

We work with three types of tenants in two industries in a partial equilibrium model. We discuss generalizations later. Industries are denoted $i = \{1, 2\}$ and, without loss of generality, we suppose that anchors belong only to industry 1. Suppose for now that the anchor does not occupy any space. This allows us to focus on the division of space between nonanchor tenants in the two industries. Let s_1 and s_2 denote the space allocated to the nonanchor tenants in industries 1 and 2 and suppose that the total space occupied by the two types of nonanchors is normalized to one. The model will consider four features of a building that impact the building manager's profit: the composition of a building as determined by s_1 and s_2, the quality of the building, the quality of its neighborhood, and the presence of an anchor.

The profits of an industry-1 nonanchor tenant depend on the composition of the building and on the characteristics of the neighborhood. Specifically, an industry-1 nonanchor earns profit

$$\pi_1 = f(s_1) + \alpha g + q + \phi - r_1. \qquad \text{(II.1)}$$

The first term, $f(s_1)$, captures the impact of within-building agglomeration of industry-1 tenants on an industry-1 firm's profits. As in the agglomeration literature, we suppose that $f(\text{-})$ is increasing for low values of s_1 and decreasing for high values of s_1. There are several forces that would give an inverted U-shaped $f(\text{-})$. Agglomeration effects among tenants would lead $f(\text{-})$ to be increasing and concave, where the agglomeration effects decrease as the share of industry-1 tenants increases. Furthermore, these tenants compete against each other in both product and input markets, which can result in $f(\text{-})$ decreasing in s_1 for large enough s_1 (e.g. as when multiple restaurants are present in the same building). In the absence of these effects, $f(\text{-})$ would be globally increasing in s_1, which would lead to the complete specialization of buildings. As will be seen later, however, complete specialization is rare .[6]

[6] Another force that could lead to the absence of complete specialization is heterogeneity in tenant attraction to particular buildings. For example, some law firms might visit the courthouse more frequently, giving them attractions to buildings that are near the courthouse.

The second term in (II.1), αg, captures the impact of the presence of an anchor on the profits of tenants in the same industry. The variable α is a 1-0 dummy variable for the presence of an anchor, while g captures the extra profit associated with an anchor. The third term, q, captures the quality of the building. The fourth term, ϕ, captures location-specific contributions to profit, including neighborhood factors. The final term, r_1, denotes rent.

The profits of an industry-2 tenant are given by

$$\pi_2 = h(s_2) - r_2. \tag{II.2}$$

$h(\text{-})$ is initially increasing in s_2 and later decreasing, as with the parallel expression for the type-1 firms. We suppose that there is no impact of building quality or the building's neighborhood. This allows us to consider the impact of changes in building quality or neighborhood that positively impact only the profits of type-1 tenants. It is straightforward to consider how these factors might impact type-2 tenants.

In this situation, the building owner has a fixed amount of space to allocate between the two types of nonanchor tenants. Normalizing the available space to unity means that (II.2) can be rewritten as

$$\pi_2 = h(1\text{-}s_1) - r_2. \tag{II.3}$$

This indicates that type-2 nonanchor profit arises from agglomeration effects (the h function) and building and neighborhood characteristics that enter into industry-1 tenant profits in (II.1) that determine s_1.

B. Equilibrium building composition

The building owner cannot command that nonanchor tenants occupy a particular building at an arbitrary rent. The tenants must be willing to pay the rent, and this will depend on the building's composition, its characteristics, and the characteristics of the building's neighborhood. Solving for this bid rent for nonanchor tenants is straightforward. Setting profit equal to zero and rearranging gives rent as a function of building composition for a nonanchor tenant for the two industries:

$$r_1(s_1, \alpha, q, \phi) = f(s_1) + \alpha g + q + \phi \tag{II.4a}$$

$$r_2(s_2) = h(1\text{-}s_1) \tag{II.4b}$$

The building owner chooses how to allocate space in the building between the nonanchor tenants in the two industries to maximize aggregate rent, R, defined as:

$$R = s_1 r_1(s_{1,}\alpha, q, \phi) + (1\text{-}s_1)\, r_2(1\text{-}s_1). \tag{II.5}$$

The first-order condition for an interior solution to the building owner's choice problem is:

$$r_1(s_{1,}\alpha, q, \phi) + s_1 f\,'(s_1) - r_2(1\text{-}s_1) - (1\text{-}s_1)\, h'(1\text{-}s_1) = 0. \tag{II.6}$$

(II.6) means that the shares of the building allocated to the two types of nonanchor tenant must set marginal revenues equal for the two tenant types. The second-order condition is that

$$2f\,'(s_1) + s_1 f\,''(s_1) + h'(1\text{-}s_1) - (1\text{-}s_1)\, h''(1\text{-}s_1) < 0. \tag{II.7}$$

We assume that this condition holds which is necessary for a stable interior equilibrium.

Consider now the effect of the presence of an anchor on the optimal allocation of space between nonanchor tenant types. Let s_1^N denote the solution when there is no anchor and s_1^A denote the solution when there is an anchor:

$$r_1(s_1^{N}, 0, q, \phi) + s_1 f\,'(s_1^N) - (1\text{-}s_1^N)\, r_2 - (1\text{-}s_1)\, h'(1\text{-}s_1^N) = 0. \tag{II.8a}$$

$$r_1(s_1^{A}, 1, q, \phi) + s_1 f\,'(s_1^A) - (1\text{-}s_1^A)\, r_2 - (1\text{-}s_1)\, h'(1\text{-}s_1^A) = 0. \tag{II.8b}$$

The left-sides of (II.8a) and (II.8b) are continuous and decreasing in s_1, the latter by (II.7). Since $r_1(s_{1,}1, q, \phi) > r_1(s_{1,}\, 0, q, \phi)$, we must have $s_1^N < s_1^A$. The presence of an anchor makes the building more attractive to industry-1 tenants, raising the rents for these nonanchors, leading the owner to include more in the building. We thus anticipate a positive equilibrium relationship between α and s_1. This relationship will be central to the paper's empirical analysis.

Thus far, we have treated the anchor as consuming no space. One could instead treat anchor size and the space it occupies as part of the composition problem, with the building owner allocating space between the two nonanchor types and the anchor. The solution would set the marginal revenue of space allocated to an anchor equal to marginal revenue from space allocated to the nonanchor tenants as described in (II.6). Intuitively, the anchor's marginal revenue equals rent paid by the anchor plus revenue gained from the anchor's positive effect on type-1 nonanchor tenant rent. Since building owners can charge higher rents when there is an anchor present, they will compete for such anchors by offering discount rents. Gould et al (2005) show very significant rent reductions for anchors.[7] In many cases, the anchor pays zero rent. The exact discount that an anchor receives will depend on the intensity of the competition among building owners. Since our data do not allow us to consider this, we will not present a formal model of this aspect of the building owner's choice problem.

C. Empirical implications

The model has demonstrated that agglomeration spillovers between an anchor and smaller firms result in an anchor skewing a building's composition towards its own industry. This is the primary relationship we estimate later in the paper. As a robustness check, we also estimate a model of building level sales, which amounts to an examination of the paper's assumptions on firm profit.

There are, of course, other potential explanations for what may draw smaller tenants to a given building. Buildings that are attractive to small firms may also be attractive to anchors. So an estimated relationship between an anchor and building composition could be driven by unobserved forces. Or the anchor may attract tenants because the anchor, in industry language, "validates" the quality of the building and its management. According to this reasoning, the anchor attracts small tenants in and outside of its industry by signaling the desirability of the location.

In the empirical work to follow, we address these issues in a number of ways. Our regressions include extensive controls for the level and composition of employment close to and even within a building. We also include, variously, zipcode, block, and building fixed effects. In addition, we compare the effect of an own-industry anchor to anchors from outside the own industry. According to the validation idea, both would attract tenants, while evidence on the importance of localization economies means that agglomeration effects will largely manifest themselves in the own-industry anchor effect.

---

[7] See Gould et al (2005) for an empirical analysis of the relationship of anchor tenants and non-anchor rent. They cannot reject the hypothesis that the retail malls they consider are "optimally" composed in the sense that the share of non-anchors and anchors is a local maximum in total rent.

III. Data and summary statistics

A. Data sources

The primary data for the analysis is a cross-section of establishment level records from Dun and Bradstreet. The data were obtained through Mergent Intellect and were downloaded in 2019 from the Syracuse University library which has a site license with Mergent.[8] D&B markets its data primarily for use by commercial companies seeking to learn about potential business partners and clients. As such, the emphasis at D&B is on meeting the current information needs of the business community. This information includes the street address of an establishment, which is present for all records. The D&B file is vast, providing near comprehensive coverage of establishments in urban areas across the U.S. (details are noted shortly). These features make it possible to identify establishments located at the same address and from there to determine the level and composition of commercial activity at each address. This makes the D&B data uniquely suited to the building-level analysis that is central to our investigation.

We work with 2019 data, and our geographic coverage focuses on five urban areas. We obtained the full set of establishments for the core counties of New York (the five boroughs), Chicago (Cook County), San Francisco (San Francisco County), Los Angeles (Los Angeles County), and Washington, DC. For each establishment, in addition to street address (and latitude/longitude), we observe the establishment's level of employment, primary industry classification based on SIC code, and many other attributes. Sales are also reported at the establishment level for single-site firms. Similar D&B establishment-level data have recently been used by Liu et al (2020), Rosenthal and Strange (2020), Rosenthal and Urrego (2022), and Bruckner and Rosenthal (2023).[9]

[8] Terms of the Syracuse University site license do not allow us to share the D&B data. However, we are able to share our code which can be downloaded from the Mendeley Data repository at 10.17632/2ddr8wbnpb.1 . See Liu, Rosenthal and Strange (2024) for the full reference to the posted files.

[9] Other academic studies have used a panel version of the D&B data developed by Walls and Associates (Walls, 2007) referred to as the National Establishment Time-Series or NETS. While the NETS data are not precisely the same as data used in this paper, both the NETS and Mergent Intellect files draw from the same root Dun and Bradstreet source. Comparison of NETS data to census data has been conducted by Barnatchez, Crane and Decker (2017) and Crane and Decker (2019). Barnatchez et al (2017) conclude that the NETS "can be useful and convenient for studying static business activity in high detail" as we use the Mergent Intellect files for in this study. The largest difference between NETS and Census data noted in their study is that the NETS often imputes data for small establishments and that imputation is especially prevalent in the sales data. Crane and Decker (2019) also emphasize

As indicated above, the D&B database is vast. The full data file from Mergent Intellect provides information on roughly 20 million establishments across the U.S. This compares to roughly 6.1 million employer businesses in 2019 based on U.S. Census data and an additional 26.5 million nonemployer (one-person) businesses also reported by the U.S. Census.[10] With such comprehensive coverage, and knowing the address of each establishment, we are able to determine the level and composition of activity at individual addresses as noted above. In the discussion that follows, we treat the street address of an establishment as synonymous with the building in which it is situated. While some very large structures could encompass multiple street addresses, we believe that workers are more likely to interact when their establishments are situated at the same street address. In part that is because of common entry way into the building and ease of interior travel between establishments at the same address.[11]

Coding establishments to their street addresses was done in several steps. We first sorted establishments by zipcode. This controls for instances in which the same street name is used in multiple locations (e.g. "Main Street"). Data were then further sorted by street address within each zipcode having broken the address field into separate words. Because the same street name is sometimes written differently (e.g. Street versus St. versus Rd.), an iterative matching algorithm was used. Following each iteration, addresses that could not be reliably matched were saved to a separate file and examined. Alternate ways to reference or spell a street name were then coded directly into our programming and the programming was run again. This resulted in code that became many lines long but with a near 100% match rate.[12]

---

that the NETS data often misreports establishment births and deaths which introduces mismeasurement when studying business dynamics. These limitations are mitigated in our study because (i) we aggregate employment (and sales) to various levels of geography (e.g. neighborhood, block, building) which helps to average away errors at the establishment level, and (ii) our focus is on static patterns and not dynamics.

[10] For additional details, on the D&B data, see: https://www.dnb.com/content/dam/english/dnb-data-insight/global-data-collection/dnb_global_and_us_business_data.pdf . For details on Census data see the small business and entrepreneurship website at: https://sbecouncil.org/about-us/facts-and-data/ . Further information is available directly from the U.S. Census at: https://www.census.gov/data/tables/2019/econ/susb/2019-susb-annual.html and https://www.census.gov/programs-surveys/nonemployer-statistics/data/tables.html.

[11] In very large buildings that encompass multiple addresses, interior walls, distance or other barriers can make interior travel between addresses difficult.

[12] Additional checks based on the reported address latitude and longitude for each establishment were used to confirm results.

Our theory posits that anchor tenants generate positive externalities that attract smaller tenants in the anchor's industry. For this reason, when coding locations, we focus on tenants as the entity that matters as opposed to establishments, *per se*. We begin by using the DUNS number provided by Dun and Bradstreet to identify individual establishments: the DUNS number is a unique identifier that D&B assigns to each establishment.[13] The great majority of firms are single-site companies, and for these the establishment is the tenant. The same is true for most large multi-site firms that have multiple offices or facilities in different locations, but not within the same building. In a small number of instances, multiple establishments belonging to the same firm are present at the same address.[14] This occurred for 6,646 establishments out of a total of 1,102,584 (or 0.6%). In such instances, we treat the firm as the tenant and set tenant employment equal to the sum of all of its establishments within the building.[15] Other attributes of the tenant are set equal to information reported for the tenant's largest establishment in the building, as with industry type, as an example. This ensures that within each building, each tenant is associated with just one record that includes total employment for that tenant within the building and for which industry-type is based on the primary function of the tenant.[16] In total, we observe the level and composition of employment for 105,982 commercial buildings spread across the five cities noted above. Only the D&B data allow us to develop such detailed building-specific information.

Next, we code whether an anchor tenant is present in each building, and if so, the primary industry to which it belongs. In the commercial real estate industry, references to anchor tenants as large, noteworthy companies are common, both on an absolute scale and relative to other tenants in a building. But there is no standard definition of what constitutes an anchor tenant. Economic theory only provides

[13] Of the roughly 1.1 million establishments, there were only ten instances in which a DUNS number was repeated for different establishments. For those records, only the first instance was retained.
[14] This was determined by comparing establishment names; those with the same name were treated as belonging to the same firm. In principle one could instead compare parent firm DUNS numbers which are also included in the data. But that data item is missing in most instances whereas firm name is always provided.
[15] When aggregating establishments, we used firm name which is always reported in the data.
[16] We also ran all of our models without making the adjustment above. In those runs, multiple records for the same firm within a given building were treated as separate establishments. Estimates were qualitatively identical while the magnitudes of the estimated patterns were very close to those reported later in the paper.

the qualitative guidance that anchors are, in some sense, large. As discussed earlier, crucial for this paper, anchor tenants are thought to generate within-building spillovers that draw other tenants to the building.

Bearing the above in mind, for a tenant to be coded as an anchor in the analysis to follow, we adopt a three-part definition. It must be the largest employer in the building, it must account for at least 20 percent of employment in the building, and it must reside in a building with at least 50 workers present. Coded in this way, buildings with fewer than 50 workers do not have anchors. Neither do larger buildings in which the largest tenant contains less than 20 percent of a building's employment. We also experimented with alternate criteria for anchor tenants. For two such trials, alternate tables of results are in the Online Appendix D for review. These include models for which anchors are defined as above but with 15% and 25% as the within-building employment threshold. Results are robust.

In order to evaluate spatial attenuation patterns in a way that respects the built environment, we next coded the location of each building in our sample to the city block and blockface (side of the street) where the building is located. This required matching the location of buildings identified in D&B with detailed publicly available GIS shapefiles of the street maps for our five urban areas. The manner in which this was done and sources for the different shapefiles is described in the Online Appendix A.

Having coded the spatial features of the data as above, it was straightforward to create several geographic measures that are central to our analysis. For each building, we determine the set of other buildings in the sample that are within 0.1 miles, approximately two city blocks. Aggregating employment across buildings, we code the level and composition of employment within 0.1 miles outside of each target building. For each building, we also determine the set of buildings on the same blockface – defined as the same side of its street block – and buildings across the street. This allows us to determine the level and composition of employment on the blockface outside of the target building and the same for buildings across the street. Following an analogous procedure, we code the presence of other anchor tenants on the blockface and across the street, and the primary industry of each of those anchors. Comparing the effect of anchor tenants within a target building, elsewhere on its blockface, and across the street is a key part of our empirical strategy, as will become apparent.

As described in the Introduction, we focus on three broad types of industries with a total of 7 subgroups. The first is Retail, which relies heavily on foot traffic and is sensitive to shopping externalities. Using the SIC classification, we include shopping-oriented Retail (SIC 52-54, 56, 57, and 59).[17] A second group of industries are information-oriented office companies for which productivity is likely sensitive to face-to-face interactions. This includes the more information-oriented parts of FIRE (Finance, Insurance and Real Estate), which we refer to as "Finance" hereafter (SIC 62, 63, 6512-6519, 67). Also included in this group are Advertising (SIC 7311-7313, 7319), Law (81), Software & Data Processing (SIC 7370-7379), and Engineering & Management (SIC 8711-8748). Manufacturing (SIC 20-39) is our third and final industry. As suggested in the Introduction, productivity spillovers in Manufacturing are likely sensitive to shipping distances that are relevant to sharing of intermediate inputs and to the commuting distance that defines the geographic scope of a local labor market.

Together, the seven core industries just noted account for roughly 29% of employment within 0.1 miles of our target buildings (Table 1a, Panel B) and 35% of establishments (Table A-1 of the Appendix). The remaining 65% of establishments are grouped together in an all-other category (denoted "Other Industry"). Details on the 2-digit SIC composition of establishments (not employment) for each of the eight industry groupings are in Table A-1 of the appendix. As would be expected, summary measures in Table A-1 make clear that the all-other industry grouping is extremely heterogenous, including notable representation from construction SIC (15-17), transportation and communications (SIC 47 and 48), wholesale trade (SIC 50 and 51), restaurants (SIC 58), and portions of FIRE (SIC 60-67) and the service sectors (SIC 70-89) not included in the core industries defined above.

It is worth emphasizing that when characterizing the level and composition of employment within 0.1 miles outside of a target building, all employment in all buildings in the D&B database is used; this includes nearby buildings with as few as just one worker present. The samples used to estimate our regression models, however, are restricted to target buildings with at least 10 workers present.

[17]Results follow the same pattern when we include Restaurants (SIC 58).

B. Summary statistics

Tables 1a and 1b present summary measures for the target buildings used in the regression analysis. In Table 1a, Panel A, for buildings with 10 or more workers, 39.2% of the buildings are in New York, 36.5% are in Los Angeles, 19.7% are in Chicago, 3.6% are in San Francisco, and 1.0% are in Washington DC. Panel A also divides buildings into small and large based on whether there are 10 to 100 workers present, and 100 or more workers present. Similar sample shares are present for both size categories.

Panel B of Table 1a reports summary measures on the composition of employment within 0.1 mile from buildings in the sample, once again grouping all size buildings together and stratifying into the small and large size categories. Employment shares are broken out by the seven target industries noted earlier, along with the eighth Other Industry category. Focusing on the first pair of columns, which includes all buildings in the sample, among the seven target industries, Retail accounts for the largest share of employment at 38.1%, while Manufacturing is next with 24.1%. The remaining 37.8% of employment is spread between Engineering and Management (15.6%), Finance (10.9%), Software and Data Processing (5.4%), Law (3.7%), and Advertising (1.7%). In the larger buildings (the far-right pair of columns), Retail is less present, accounting for 27.7% of employment among the seven target industries, with most of the difference shifted towards the more information-oriented industries. A last point to note from Panel B is that the seven target industries account for 29.3% of nearby employment with the Other Industry category making up the rest. Table A-1 of the Appendix provides further details on the 2-digit SIC composition of establishments for each of the eight industry groups.

Table 1b, Panel A, summarizes the frequency with which anchors are present for the different industries. The panel also reports the share of buildings with fewer than 50 workers for which, by definition, anchors are never present. As reported towards the bottom of the panel, 80.1% of buildings have fewer than 50 workers. For our sample of buildings overall, therefore, most are too small to have an anchor. The pattern is different for buildings with 100 or more workers (in the far-right columns), of which there are 10,580 in the sample. For this group, anchors are present in 67% of buildings. Of those anchors, 68.2% are in the Other Industry category, 11.2% are in Manufacturing, 7.8% are in Retail, 4.8% are in Engineering, 3.7% are in Finance, and Advertising and Law both account for roughly 1% of anchors.

In Table 1b, Panel B shows that there are many instances in which an industry is not present in a building. The censoring issue is especially pronounced for Advertising, for which only 2.9% of buildings include one or more advertising establishments. Retail is more ubiquitous and is present in 34.4% of the buildings in the sample. Finance is present in 14.7% of buildings while Law appears in 6.7% of buildings. Manufacturing and the Other Industry category are present in 17.1% and 90.6% of buildings, respectively. Among buildings with 100 or more workers (the far-right column), the frequency with which industries are present is notably higher for each of the eight groupings. Finance, for example, is present in 40.2% of larger buildings. There is nothing surprising about the nature of these patterns, but it is important to be aware of the large number of zeros in the data when estimating the building level industry-employment share models that follow. We explain how we address this issue when we introduce the regression models later in the paper.

Panel C of Table 1b summarizes the frequency with which an industry occupies an entire building. As would be expected, this is uncommon and occurs primarily only for Retail and Manufacturing. For buildings with 10 to 100 workers (column 2), 4.2% are fully occupied by Retail and 2.43% by Manufacturing. For buildings with more than 100 workers (column 3), both industries fully occupy 2.85% of buildings. For the other core industries (not including the Other Industry category), the frequency with which an industry fully occupies a building is much smaller. It is partly for this reason that our conceptual model discussed earlier focuses on conditions associated with interior solutions in which a mix of industries occupy space in the typical building.

IV. Building Specialization

A. Employment composition

In this section, we present evidence that commercial buildings tend to be specialized by industry. We begin by summarizing correlations in the composition of employment in nearby commercial buildings. If there are many buildings in a neighborhood and activity across nearby buildings is highly correlated, that would indicate the absence of specialization. If instead activity is only weakly correlated across buildings, then buildings specialize and our focus in later sections shifts to establishing why, including whether random or systematic forces drive observed patterns, and the nature of systematic forces when present.

Bearing the above in mind, Table 2 reports correlations in industry employment shares between target buildings and other nearby buildings. We consider three classes of nearby buildings: buildings elsewhere on the same blockface, buildings on the same block but across the street, and buildings within 0.1 miles not including the target building itself. In Panel A, target buildings include all buildings with 10 or more workers while in Panel B target buildings are restricted to those with 100 or more workers. In both cases, correlations are based on comparisons of employment in the target buildings with total nearby employment for the designated location outside of the target building. As noted in the prior section on data features, in all cases total nearby employment includes employment from all buildings with at least 1 worker in the D&B data. Several patterns are apparent.

First, for both building size categories (Panels A and B) the correlations are nearly all positive, indicating that greater industry presence at a broader level of geography tends to be mirrored at narrower levels of geography. Neighborhoods that specialize are comprised of buildings that tend to specialize as well. Second, the correlations are larger when focusing on employment within 0.1 mile as compared to employment elsewhere on the blockface or across the street. Not all buildings in a specialized neighborhood are specialized in the same activity to the same degree. Third, comparing Panels A and B, the correlations between the composition of large buildings and their neighbors are clearly larger than the corresponding correlations based on smaller target buildings. In part, this may arise because large buildings have more capacity to mirror the composition of nearby activity as compared to small buildings that can only house a limited number of establishments. Fourth, the correlations differ substantially across industries, especially for larger buildings. Focusing on correlations based on employment within 0.1 mile (column 1), for buildings with 100 or more workers (Panel B), correlations for our seven target industries from largest to smallest are Manufacturing (0.372), Law (0.293), Retail (0.297), Finance (0.190), Software/Data (0.158), Engineering/Management (0.112), and Advertising (0.088).

The patterns above confirm that commercial buildings show some degree of specialization, the degree to which differs across industries. As discussed above, this could arise from random assignment of establishments across buildings within a given neighborhood. The pattern could also reflect systematic forces that pull similar-type companies to a common city block or building. The next section provides evidence that both random and non-random forces likely contribute to building-level specialization.

B. Building specialization: random versus systematic forces

This section uses a Monte Carlo design to provide initial evidence that there are instances in which systematic forces contribute to building-level specialization. We begin by first repeatedly measuring the degree of building-level specialization that would arise solely from random assignment of tenants across buildings. This is then compared to the actual level of specialization. A strength of this approach is that it provides compelling evidence of whether observed building-level specialization is a statistically significant departure from randomness. A downside, however, is that the Monte Carlo approach is computationally intensive. For this reason, in this section we focus on a single highly specialized neighborhood, New York’s downtown financial district with the New York Stock Exchange (NYSE) at its center. This area is famous for its specialization in finance, and it is possibly the best-known instance of neighborhood-level specialization in the commercial sector. If there is systematic building-level specialization, the Wall Street neighborhood is a good place to look for it.[18]

[18] The following section uses a completely different regression-based approach that is not computationally intensive. In that section, we draw on the full sample of buildings.

Before introducing the Monte Carlo design, we begin with some summary statistics. Table 3a, Panel A describes the area within 0.25 miles of the NYSE while Panel B provides summary measures out to 1.0 mile of the NYSE. Within one-quarter mile, 27.6% of employment is in Finance (in column 4) and 45.8% of buildings include some Finance employment (in column 5). Within 1.0 mile, the corresponding values are 17.3% and 31.9%, respectively, indicating a decline in financial activity with distance to the NYSE. Not surprisingly, these measures confirm the dominant presence of Finance in the neighborhood close to the NYSE. Nevertheless, even in this famously specialized district, most commercial buildings contain no finance at all, suggestive of specialization at the building level.

To assess the degree of building-level specialization, we use a spatial "G" statistic, defined as $G_i \equiv \sum_b (S_{bi}^N - S_b^N)^2$, where S_b^N is the share of total employment in the neighborhood (N) that is present in building b (b = 1, …, B^N) from among B^N buildings, and S_{bi}^N is the share of industry i employment in the neighborhood located in building b. This spatial G_i measure has been used by Krugman (1991) and Audretsch and Feldman (1996) to summarize the degree of spatial correlation in industry-specific activity across states. The measure addresses the tendency for activity in a given industry to be more (or less) spatially concentrated relative to total employment. At one extreme, if employment in industry i is distributed across space in a manner that mirrors total employment, then $S_{bi}^N = S_b^N$ for all b and G_i equals 0. At the other extreme, as employment in industry i becomes concentrated in a single location, G_i tends towards 1 from below.

Measuring the actual value of G_i for each of our eight industries is straightforward and simply requires that we measure G_i using the expression described above. To assess the degree of specialization that would arise from a purely random assignment of tenants across buildings requires that we take building capacity constraints into account as well as the size of individual tenants. A neighborhood filled with many small buildings, for example, will tend to exhibit more specialization at the building level because each building can only accommodate a small number of establishments, limiting the diversity of activity in any of them. Analogously, building-level specialization is also sensitive to the size distribution of firms within a given industry. An industry dominated by a small number of very large establishments, for example, will appear spatially concentrated. This was emphasized by Ellison and Glaeser (1997), although in a very different context.[19] We address these concerns in the Monte Carlo trials as follows.

For a given sample of buildings (e.g. within 0.25 miles of the NYSE), we conducted 1,000 Monte Carlo simulations. In each simulation, all establishments in the sample were randomly reassigned to buildings within the target area. To allow for building specific space constraints, we assumed that each building could accommodate up to 10% more employment than is actually observed in the data. We then randomly drew a first establishment from among all establishments in the neighborhood. This establishment was assigned to a randomly drawn building in the neighborhood from among the subset of buildings with more capacity than the number of workers in the establishment. A second establishment was next drawn and the exercise repeated, taking into account only the remaining space available in each building. This procedure was repeated until all establishments were assigned to a building. If an establishment is drawn for which no building has sufficient capacity to accommodate the establishment's employment, we split the employment across up to four different buildings as needed.[20] For each randomized sample, we re-compute the spatial G measures for each industry. Repeating this procedure 1,000 times yields an empirical distribution of simulated values for G, with sample size of 1,000. Comparing the actual spatial G for an industry to its simulated values allows us to determine the degree to which the actual G differs from what would be expected with random assignment.

Results are reported in Table 3b for four sets of estimates. Panel A shows results for the 0.25-mile neighborhood with separate estimates for buildings with 10 or more workers in the left set of columns and for buildings with 100 or more workers in the right set of columns. Panel B does the same for the one-mile neighborhood. For each group of estimates, each row reports values for a specified industry for all

---

[19] Ellison and Glaeser (1997) examined the spatial concentration of employment among manufacturing industries at the metropolitan level.

[20] In doing so, we first assign establishment employment to the building with the largest amount of open employment capacity. Any remaining employment in the establishment is then allocated to the next building with the largest amount of open capacity and so on up to four buildings. Employment remaining after spreading the establishment across four buildings was discarded but this rarely happens. We also experimented with discarding the entire establishment if it could not fit into a single building. This yielded results quite similar to the procedure just described.

eight industries. Within each of the four blocks of values, reading left to right, estimates indicate the actual value of G, the mean simulated value, 1 minus the probability that actual G exceeds the mean simulated value (based on a 1-tailed test), and the 90th and 95th percentiles of the 1,000 simulated values for G. Note also that one star indicates that actual G exceeds the 90th percentile simulated value while two stars indicates it exceeds the 95th percentile.

Reviewing the table overall, a dominant pattern is clear: in most cases, buildings are significantly more specialized than would occur from random assignment of establishments. This result holds even after allowing for the influence of the size distributions of both buildings and establishments.

A second compelling pattern in Table 3b is that financial activity is especially spatially concentrated. For Finance, the actual value for G in Panel A (within 0.25 miles; buildings with 10 or more workers) is 0.130, well above the simulated sample mean value of 0.073 as well as the 95th percentile value of 0.10. We see a similar situation for the other three cases (0.25 miles for buildings with 100 or more workers; 0.5 miles for buildings with 10 or more employees; 0.5 miles for buildings with 100 or more employees). For each of the four exercises, the actual value of G for financial establishments is notably larger than the mean value of G across the 1,000 simulations.

Table 3b shows that other industries also exhibit non-random spatial concentration although not to the degree as appears to be the case for Finance. Advertising displays evidence of nonrandom spatial concentration, and especially within 1 mile of the NYSE (reported in Panel B). This echoes Arzaghi and Henderson (2008) who consider spatial concentration of advertisers in Manhattan in areas close to Grand Central Station. In Law, there is a significant but more modest difference between actual G and mean simulated G for the area within 0.25 miles of the NYSE (Panel A). But those differences disappear when the geographic focus is expanded to a 1.0 mile area (Panel B). Manufacturing also displays little evidence of nonrandom building-level specialization. The sharp differences in these patterns across industries is a theme we will return to later in the paper.

In conclusion, this section provides compelling evidence that in a neighborhood famous for its spatial concentration, we see further systematic specialization at the building level. In the next section, we will take a different approach to understanding building specialization by examining systematic patterns between spatial concentration and the attributes of neighborhoods, buildings, and tenant establishments.

V. Spillovers, attenuation, and anchors

A. Agglomeration and anchors

The previous section established that commercial buildings in New York's financial district are specialized beyond what would be expected from a random assignment of establishments. Drawing on our full five-city sample, this section estimates a series of regression models that confirm that non-random specialization of buildings is common while also shedding light on the nature of the agglomeration economies that contribute to specialization. As discussed earlier, we focus on whether an anchor is present in the target building, elsewhere on the blockface, and/or across the street, distinguishing anchors in and outside of a target industry. Controlling for other attributes of a building and its immediate neighborhood, the differences between the effects of an anchor at different locations shed light on the attenuation of agglomeration economies in a specification that incorporates the built environment.

In Part B below, we provide evidence of within-building spillovers that attenuate rapidly outside of the building. Part C considers heterogeneity by industry type, including Retail, information-oriented office industries, and Manufacturing. In Part D, focusing just on Retail and Law, we consider a complementary model that uses building level sale per worker as the dependent variable. Across all of these models, results are robust.

B. Rapid attenuation of anchor effects

We begin by specifying a dependent variable that captures building-level specialization. In the regression models below, the primary dependent variable is the within-building share of employment in industry *i* outside of employment in an industry-*i* anchor if one is present. We denote the share of industry *i* employment in building *b* as $S_{bi} = \frac{E_{bi}-E_{bi}^{*}}{E_{bi}}$, where E_{bi} denotes total employment in industry *i* in building

b, E_{bi}^{*} represents employment at an industry-i anchor in the building, and E_b gives total employment in the building. Section II's theory predicts that S_{bi} will depend on the presence of an anchor and on various neighborhood and building characteristics.

Given this dependent variable, all of the models in this section are of the general form of expression (V.1) below, where i = 1, …, 8 denotes the eight industry groups. Each specification is based on linear models that pool data across industries, with each observation a building-by-industry record:[21]

$$\begin{aligned} S_{bi} = {} & c_{0,i} + c_1 S_{bi}^{N} + c_2 log(E_b^{N}) + c_3 log(E_b\ ) + c_4 log(E_b^{BF}) + c_5 log(E_b^{AS}) \\ & + c_6 Anchor_{bi} + c_7 Anchor_{bi}^{BF} + c_8 Anchor_{bi}^{AS} \\ & + c_9 \textstyle\sum_{k\neq i} Anchor_k + c_{10} \sum_{k\neq i} Anchor_k^{BF} + c_{11} \sum_{k\neq i} Anchor_k^{AS} \\ & + c_{12} X_b + c_{13} X_b^{BF} + c_{14} X_b^{AS} + c_{15} X_b^{N}\ . \end{aligned} \qquad \text{(V.1)}$$

Controls in (V.1) are as follows. In the first line, $c_{0,i}$ is an industry fixed effect that captures any overall industry-specific attributes that might affect the tendency of an industry to concentrate at the building level. Other controls in the first line include the share of employment in industry i in the neighborhood (defined as within 0.1 mile) outside of the target building (S_{bi}^{N}) and the log level of employment in the neighborhood (E_b^{N}, also defined to be outside the target building). Continuing with the first line of (V.1), we also control for the log level of employment in the building (E_b), on the blockface outside of the target building (E_b^{BF}), and across the street (E_b^{AS}). Throughout, we use superscripts for level of geography. *BF* refers to the same blockface as the target building *b*, while *AS* refers to across the street. *N* refers to the 0.1 mile neighborhood of the target building, and absence of a superscript refers to the building itself. The second line includes 1-0 indicators of whether an industry *i* anchor is present in the target building, in at least one other building elsewhere on the blockface, and/or in at least one building across the street. The third line repeats these controls for the presence of an anchor outside of industry *i*, summing across

[21] Although the specification in V.1 does not allow for heterogeneity across industries, pooling data across industries allows us to include block and building fixed effects that capture many possible confounders. We consider industry-stratified models in the following subsection.

the $k \neq i$ industries. The last line includes indicators for additional attributes at the building, blockface, across-the-street, and neighborhood, denoted respectively by X_b , X_b^{BF}, X_b^{AS} , and X_b^N.

In some models, we estimate (V.1) including block fixed effects. This addresses possible effects of unobserved elements in X_b^{BF}, X_b^{AS}, and X_b^N, all of which are captured by block fixed effects. Other controls vary within a street block and remain. In principle, this includes the log level employment measures $log(E_b^N)$, $log(E_b)$, $log(E_b^{BF})$, $log(E_b^{AS})$. In practice, however, estimates were nearly unchanged when these variables were omitted with block fixed effects included in the model, indicating that the block fixed effects largely capture variation in the scale of nearby economic activity. Bearing this in mind, the model is written below with the employment log level variables included but we drop those measures in the block fixed effect tables presented later:

$$\begin{aligned} S_{bi} = \ & c_{0,i} + c_1 S_{bi}^N + c_2 log(E_b^N) + c_3 log(E_b) + c_4 log(E_b^{BF}) + c_5 log(E_b^{AS}) \\ & + c_6 Anchor_{bi} + c_7 Anchor_{bi}^{BF} + c_8 Anchor_{bi}^{AS} \\ & + c_9 \textstyle\sum_{k \neq i} Anchor_k + c_{10} \sum_{k \neq i} Anchor_k^{BF} + c_{11} \sum_{k \neq i} Anchor_k^{AS} + c_{12} X_b \ . \end{aligned} \tag{V.2}$$

In other models, we include building fixed effects which fully capture the influence of $log(E_b^N)$, $log(E_b)$, $log(E_b^{BF})$, $log(E_b^{AS})$, and X_b , simplifying the model to:

$$\begin{aligned} S_{bi} = \ & c_{0,i} + c_1 S_{bi}^N + c_6 Anchor_{bi} + c_7 Anchor_{bi}^{BF} + c_8 Anchor_{bi}^{AS} \\ & + c_9 \textstyle\sum_{k \neq i} Anchor_k + c_{10} \sum_{k \neq i} Anchor_k^{BF} + c_{11} \sum_{k \neq i} Anchor_k^{AS}. \end{aligned} \tag{V.3}$$

It is worth noting that X_b includes both the physical attributes and the management practices of the building. These may have a common draw (or deterrent) for certain industries. Finance and Law, for example, are more likely to engage with employees and clients who place greater value on meeting in high quality, Class A buildings. Manufacturing, in contrast, is more likely to seek out lower quality, less expensive buildings.

In Table 4a, data is pooled across the eight industry groups. The first three columns control for block fixed effects while the second three columns control for building fixed effects. For each group of columns, the first includes all buildings with 10 or more workers, the second restricts the sample to buildings with 10 to 100 workers, and the third includes only buildings with 100 or more workers. Notice that for the block fixed effect models, controls are included for anchors outside of industry-*i*. Those controls drop out in the building fixed effect models as they are fully captured by the combination of building fixed effects and industry fixed effects.

Focus now on the top row, which reports the coefficient on own-industry employment share within 0.1 miles of a target building. Reading across columns, the coefficient is extremely stable and always highly significant. Increasing the industry-*i* neighborhood employment share by 10 percentage points increases industry-*i* employment share in a target building by roughly 3%. Not surprisingly, the composition of nearby activity is strongly positively related to the composition of activity within individual buildings.

Consider now our primary controls of interest, the anchor controls for target building, blockface, and across the street. Reading across columns there is always a sharp attenuation pattern that is especially dramatic for larger buildings. In the block fixed effect model in column 3, for example, the coefficient on own-industry anchor in the target building is 0.143. This indicates that when an own-industry anchor is present in the building, employment share in the anchor's industry among other tenants is 14.3% higher. That effect shrinks to just 1.2% if the anchor is elsewhere on the same blockface and to just 0.25% if across the street. Moreover, these estimates are nearly identical if we replace block fixed effects with building fixed effects in column 6.

Two conclusions follow from these patterns. First, as a measurement issue, conditional on the other controls in the model, block fixed effects appear to capture the same possible confounders as building fixed effects, resulting in very similar magnitude coefficients. Second, and fundamental for our characterization of agglomeration economies, the intensity with which smaller tenants cluster around an anchor in their industry diminishes sharply upon leaving the building. This combination of estimates is consistent with our conceptual model and anecdotes in the commercial real estate industry. It suggests that buildings matter and in part because of within-building productivity spillovers among firms within an industry.

Table 4b repeats the analysis omitting Retail and Manufacturing. Instead, we focus just on the five information-oriented office industries: Finance, Advertising, Law, Software/Data Processing, and Engineering/Management. As suggested earlier, information exchange is important for these industries and likely in ways that benefit from face-to-face interactions between firms. Estimates in Table 4b allow us to test that prior. The structure of the table and specification is as before apart from dropping Retail and Manufacturing.

The qualitative patterns in Table 4b are the same as in Table 4a. The one exception is that in columns 1-3 with block fixed effects, there now appear several negative coefficients on own-industry anchors on the blockface and across the street as opposed to small positive coefficients in Table 4a. Nevertheless, the negative coefficients are quite small and the core patterns are the same as for the previous table, both in columns 1-3 (block fixed effects) and columns 4-6 (building fixed effects). In both sets of models, there is a comparatively large, positive, and highly significant coefficient on own-industry anchor in the target building, and that effect attenuates sharply upon leaving the building. Estimates are also once again quite similar with block fixed effects versus building fixed effects. For all smaller buildings (columns 2 and 5), adding an own-industry anchor to the building increases employment share in the target industry (outside of the anchor) by roughly 7% to 8%. For the larger buildings, the corresponding estimate is roughly 10.4%.

This subsection has demonstrated that building specialization is systematically related to attributes of the immediate neighborhood and building, and in this sense, is non-random. We have also confirmed that there is a strong tendency for smaller tenants to co-locate in the same building as an anchor in their own industry, but that the presence of an anchor elsewhere on the city block has little draw. Conditional on our other model controls, this sharp attenuation pattern suggests rapid attenuation of localized agglomeration economies. This evidence is based on models that pool data across industries, however, which masks heterogenous patterns across industry type. The next subsection addresses this issue by estimating models stratified by industry. As will become clear, this provides evidence on the nature of agglomeration economies in city centers.

C. Heterogeneity and microfoundations

Table 5 presents alternate estimates from building-level regressions that are stratified by industry for each of the eight industries. The dependent variable is as in Tables 4a and 4b. Stratifying by industry, however, precludes the use of building fixed effects which require building-by-industry records. The same is approximately true for block fixed effects given the many thousands of blocks in the data. A further complication is that censoring is more pronounced at the individual industry level (recall the discussion of Table 1b, Panel A, for example).

To address these issues, we use a Tobit specification in Table 5 that takes account of zeros in the dependent variable. All of the Tobit models include a rich set of controls for the level and composition of employment within 0.1 miles outside of a target building. In addition, our primary models replace the block (or building) fixed effects with fixed effects for each of the five cities in the sample. In a robustness check, we instead control for 958 zipcode fixed effects. An advantage of the zipcode fixed effect models is that they do more to address possible unobserved attributes of a building's immediate environment. However, a small number of buildings that are their own zipcode effectively drop out of the sample (e.g., the Empire State Building). In addition, and as would be expected, zipcode fixed effects capture some of the information in our neighborhood employment variables, making it more difficult to discern neighborhood-level patterns apart from the effect of anchors.

In the models that follow, the anchor coefficients are nearly identical regardless of whether we include city or zipcode level fixed effects in the Tobit models. The qualitative patterns from the Tobit models also affirm the patterns from the linear models in Tables 4a and 4b. This robustness suggests that our neighborhood level controls for the composition and level of employment do an effective job of summarizing relevant information about the nearby environment. For reasons above, therefore, we adopt the city fixed effect Tobit model as the preferred specification, results from which are in Table 5. Alternate estimates from the zipcode fixed effect Tobit models are in Online Appendix C.

Note that all of the Tobit models control for each industry's neighborhood employment share S_{bi}^{N} outside of the target building, treating the Other Industry category (industry "8") as the omitted group. For the anchor measures, we include separate 1-0 controls for each of the eight industries in each location: in the target building, elsewhere on the blockface, and across the street. A 1-0 control is also included for whether the target building has fewer than 50 workers, making it impossible to have an anchor, $E^{Bld} < 50$. The industry-specific models in Table 5 are then estimated in the form:

$$S_{bi} = c_{0,i} + \sum_{i=1}^{7} c_{1,i} S_{bi}^{N} \tag{V.4}$$

$$+ c_2 \log(E_b^N) + c_3 log(E_b\ ) + c_4 log(E_b^{BF}) + c_5 log(E_b^{AS})$$

$$+ c_6(E_{bi} < 50) + \sum_{i=1}^{8} c_{7,i} A_i + \sum_{i=1}^{8} c_{8,i} A_i^{BF} + \sum_{i=1}^{8} c_{9,i} A_i^{AS}\ .$$

In (V.4), $c_{0,i}$ should now be interpreted as a vector of location-specific fixed effects. In our primary models, this includes separate constants for each of the five cities discussed earlier. In online Appendix C, these are replaced with zipcode fixed effects.

The large number of controls in (V.4) makes comparison of estimates difficult, both across controls for a given industry and across industries. For this reason, in Table 5, we present only the summary statistics for each regression in the top panel. This includes the number of observations, including censored and uncensored, city fixed effects and a pseudo R-square measure. That panel is followed by a set of bar charts that plot estimates from the eight regressions, including error bars on each bar segment for its 95 percent confidence ban. The graphical representation allows for a quick and clear assessment of the dominant patterns.

Consider now the Group 1 panel of estimates. These include employment level coefficients in Panel A and neighborhood employment share coefficients in Panel B. Each panel reports separate clusters of bar segments for each of the eight industries.

Focusing first on Panel A, for the Retail regression, the effect of own-building employment is significant and much larger than the coefficients on blockface and across-the-street employment, which are much smaller and insignificant. This is the pattern that we see for other industries. However, the estimated effect of neighborhood-level employment for retail is about twice the magnitude of employment in the target building. This pattern may reflect that location choice for brick-and-mortar retail stores is sensitive to pedestrian travel cost that affects visits from potential shoppers. These costs include a fixed component associated with travel outside of your building – possibly because of the need to dress for inclement weather – and travel time. Framed in this manner, it is intuitive that within-building employment is valuable for retail outlets, as travel does not require going outside. It is also intuitive that the scale of nearby employment outside of the target building is especially valuable since that has potential to be an important source of customers. The spatial distribution of potential shoppers outside of the target building, however, is not so important because all such shoppers must travel outside and their walk to the store is always quite short given the 0.1 mile focus. This likely accounts for why the coefficients on employment elsewhere on the blockface and across the street are small.

The pattern is different for the other target industries (aside from the Other Industry category). For these target industries, employment in the target building has a much larger magnitude coefficient than employment in any of the other locations. This suggests that larger buildings with more employment contain a heavier representation of these industries. This is not true, however, for the Other Industry category at the far right, for which the level of employment has little effect regardless of location.

The patterns in Panel B are striking. In this panel, target-industry employment share in the neighborhood (within 0.1 miles) always has the largest effect of any industry employment share. Notice also that in the Retail regression, higher employment share in non-Retail industries always has a negative effect. Mirroring this pattern, for the non-Retail industry regressions, higher Retail share of neighborhood employment always has a negative effect on employment share in the target industry. This indicates that there is a tendency for Retail to be concentrated in different neighborhoods than the other seven industries.

There is significant heterogeneity in how industry employment share in a target building is related to the composition of nearby employment. Target building Finance share of employment increases with neighborhood employment share in each of the other information-oriented office industries, but it is negatively associated with higher presence of Manufacturing. Law employment in a target building is especially strongly associated with higher neighborhood Law share and notably negatively associated with Manufacturing share of neighborhood employment. The reverse is true for Manufacturing.

The pattern is sharply different for the influence of anchor establishments. Group 2 plots the target building anchor coefficients from each of the eight regressions. There are eight panels in the exhibit, each of which displays estimates for a single industry-specific regression. Each panel also includes eight bar segments, one for each of the eight industry anchor coefficients. Reading down the exhibit, the main diagonal (blue) reports estimates for the own-industry anchor coefficient while off-diagonal estimates (gold) are for the other ($k \neq i$) anchor estimates. Two patterns are evident. First, the main diagonal estimates are all positive and significant. Second, and in comparison, the off-diagonal measures are mostly not significant, often very small in magnitude, and negative with just a few exceptions. The dominant main-diagonal pattern confirms yet again that the presence of an own-industry anchor contributes to specialization in that industry among other tenants within the building.

The patterns in the Group 2 exhibit also shed light on the role of anchors as validators of a building's quality. As noted previously, this is perceived to be one of the channels by which anchors attract smaller firms. In this explanation, however, an anchor would affect tenants across industries. This, in turn, would tend to produce diverse buildings. The results in Group 2 are not consistent with this. We see much larger effects of an anchor on small tenants in its own industry than on small tenants in other industries. Furthermore, the small size of the Other Industry anchor effect on this catch-all group is also unfavorable to the validation hypothesis.

The patterns in the Group 3 exhibit offer a different perspective on the role of anchors. The exhibit plots just the own-industry specific anchor coefficients from the Group 2 exhibit for the different industries and locations. This information is organized in two different ways. In Panel A, three clusters of bar segments are reported for the three different locations, target building, elsewhere on the blockface, and across the street. In Panel B, this same information is reconfigured into eight clusters of bar segments, one for each of the industries, and each of which includes estimates for the three locations. Several interesting patterns appear in these plots.

In Panel A, the estimates are much larger for anchors in the target building than for anchors elsewhere on the blockface or across the street. There is, however, an exception, one that is most evident in Panel B. For Manufacturing, own-industry anchor effects are nearly identical regardless of whether a Manufacturing anchor is found in the target building, elsewhere on the blockface, or across the street. Also in Panel B, the magnitude of the attenuation pattern differs across the other target industries. For Retail, the attenuation pattern is especially sharp. For the information-oriented office industries the attenuation pattern is very strong but not as pronounced as for Retail.

As discussed earlier, it is widely perceived in the commercial real estate industry, and confirmed in prior research, that Retail benefits from shopping externalities. The targeted office industries are ones that engage in information exchange and likely benefit from face-to-face interaction for this reason. The rapid attenuation patterns in these sectors suggest that both underlying mechanisms – shopping externalities and information exchange – attenuate sharply upon leaving a building. This is intuitive given the nature of the anticipated interactions.

The absence of attenuation for Manufacturing is different and also intuitive. Manufacturing has absentee customers, and so it would not benefit from any sort of shopping externality. Manufacturing is also not typically engaged in information exchange, reducing potential benefits from face-to-face interaction with neighbors. Instead, Manufacturing relies heavily on truck transport for delivery of physical inputs and product, and opportunities to seek industry-specific workers from a common pool (see Rosenthal and Strange (2001, 2004) for related discussion). Those forces are unlikely to create the same sort of intensity to co-locate in a common building as appears to be present for Retail and the Office-Information sector. Instead, spillovers in Manufacturing are more likely to foster agglomeration over a larger geographic area, consistent with evidence in Rosenthal and Strange (2001, 2003, 2006, 2020).

D. Robustness: Sales per worker

Table 6 presents a robustness check using building-level sales per worker as the dependent variable, focusing separately on Retail and Law. Section II's theory predicts that sales will be related to anchors in the same qualitative way that building composition is, and this is what we consider here. For both industries, the model specifications are as in Table 5, with results presented in tabular form. Because establishment level sales are most reliably measured for single-site firms, only single-site firms are used to measure sales per worker. This differs from the employment share models for which all establishments were used to measure building composition. We do not rule out that unusually productive single-site firms may sort into buildings where valued anchors are present. This type of sorting would increase the positive relationship between the presence of an anchor and sales per worker among smaller tenants but does not apply to the employment share models. It is for these reasons we treat the models in Table 6 as a robustness check.

Bearing this in mind, the patterns in Table 6 mirror those for the employment share model in Table 5. Most important the coefficients on own-industry anchors in the target building are positive and significant. For Retail the coefficient is $12,140 with a t-ratio of 3.54; for Law the coefficient is $19,247 with a t-ratio of 1.66. For both industries, the corresponding coefficients for own-industry anchors elsewhere on the blockface are modestly smaller, while the coefficients on own-industry anchors across the street are much smaller. Once again, the pattern suggests attenuation consistent with priors that concentrating Retail and Law at the building level enhances labor productivity.

VI. Conclusions

This paper considers agglomeration in a novel way. Instead of working with metropolitan-level, city-level or distance-based spatial units, this paper uses individual buildings as the spatial unit, focusing on the understudied commercial sector that dominates downtowns. Drawing on establishment-level data from Dun and Bradstreet and detailed geocoding of establishments at the building level, the paper suggests several new principles that govern the nature and intensity of urban productivity. In establishing our primary findings, we pay special attention to the role of anchor establishments that are perceived as fostering positive spillover effects by the commercial real estate industry. The paper reaches three broad conclusions.

The first is that the built environment matters: companies care about the composition of tenants in their building and are drawn to buildings in which an anchor in their industry is present. The effect of an own-industry anchor attenuates rapidly and is substantially smaller elsewhere on the blockface and even more so across the street. These results are consistent with rapidly attenuating agglomeration economies.

A second conclusion is that attenuation varies by industry, again with implications for the nature of agglomeration economies. Attenuation is most pronounced in the Retail sector, suggesting valuable within-building shopping externalities. Attenuation is also rapid in information-oriented office industries that rely on information exchange and knowledge sharing, activity that is facilitated by face-to-face interaction. The pattern for Manufacturing is different. Although manufacturing anchors attract additional manufacturing tenants, the effect does not attenuate between buildings on the same city block. This suggests a lesser reliance on in-person interaction between nearby firms.

These patterns are extremely robust. They are present in estimates from linear models with block fixed effects in some instances and building fixed effects in others. They are also present in Tobit models that include a rich specification for the level and composition of employment within a few blocks of a building, across the street, and elsewhere on the blockface, along with city fixed effects in our primary models and hundreds of zipcode fixed effects in other specifications.

The final conclusion concerns efficiency: private sector building owners have both incentives and capacities to profit from externalities associated with tenant mix. The incentive is clear. Creating a complementary set of tenants makes the building more productive, increasing rent. The capacity depends on the geographic scale at which the externalities operate. Externalities that operate beyond the building level are outside of the building owner’s scope to manage (as with the limited ability of developers to control for externalities in Helsley and Strange, 1997). However, the finding that there are strong spillover effects within a building means that the building owner has the ability to manage some of the externalities in ways that will enhance productivity in cities.

References

Ahlfeldt, Gabriel, Stephen J. Redding, Daniel M. Sturm and N. Wolf (2015). "The economics of density: evidence from the Berlin Wall." *Econometrica*, 83(6), 2127-2189.

Arzaghi, Muhammad and J.V. Henderson (2008), "Networking off Madison Avenue," *Review of Economic Studies*, 75(4), 1011-1038.

Audretsch, David and Maryann Feldman (1996), "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86(3), 630-640.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker (2017). "An Assessment of the National Establishment Time Series (NETS) Database," Finance and Economics Discussion Series 2017-110. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.110.

Behrens, K., and F. Robert-Nicoud (2014). "Agglomeration Theory with Heterogeneous Agents", G. Duranton, J. V. Henderson and W. Strange (eds), *Handbook in Regional and Urban Economics*, Amsterdam (Holland), Elsevier Press.

Brueckner, Jan K. (1993). Inter-store externalities and space allocation in shopping centers. The Journal of Real Estate Finance and Economics, 7(1), 5-16.

Brueckner, Jan K. and Stuart S. Rosenthal (2022). "Do Risky commercial Tenants Get Short-Term Leases?" Working paper.

Bosquet, Clement and Pierre-Philippe Combes (2017). "Sorting and Agglomeration Economies in French Economics Departments," *Journal of Urban Economics*, 101, 27-44.

Catalini, Christian (2018). "Microgeography and the Direction of Inventive Activity". *Management Science*, 64 (9), 4348–4364.

Combes, P.P., and L. Gobillon (2015). "The Empirics of Agglomeration Economies" *in* G. Duranton, J. V. Henderson and W. Strange (eds), *Handbook in Regional and Urban Economics*, Volume 5, Amsterdam (Holland), Elsevier Press.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, Sebastian Roux (2012). "The productivity advantages of large cities: Distinguishing agglomeration from firm selection." *Econometrica* 80.6: 2543-2594.

Crane, Leland D. and Ryan A. Decker (2019). "Business Dynamics in the National Establishment Time Series (NETS), Finance and Economics Discussion Series 2019-034. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2019.034.

De la Roca, Jorge, and Diego Puga (2017). "Learning by Working in Big Cities." *The Review of Economic Studies,* 84.1: 106-142.

Duranton, Gilles and Henry Overmann (2005), "Testing for Localisation Using Micro-Geographic Data," *Review of Economic Studies*, 72(4), 1077-1106.

Duranton, Gilles and Diego Puga (2004), "Micro-Foundations of Urban Agglomeration Economies", *in* the Handbook of Urban and Regional Economics, Volume 4, Chapter 48, 2119-2172, Elsevier, eds. Vernon Henderson and Jacques Thisse.

Eckert, Fabian, Mads Hejlesen and Conor Walsh (2022), "The Return to Bib-City Experience: Evidence from Refugees in Denmark," *Journal of Urban Economics*, Article num: 103454.

Ellison, G. and E. Glaeser , "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy* 105, 889-927 (1997).

Ellison, Glenn, Edward Glaeser, and William Kerr (2010), "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 100(3), 1195-1213.

Glaeser, Edward, Hedi Kallal, José Scheinkman, and Andrei Shleifer (1992), "Growth in Cities." *Journal of Political Economy* 100: 1126-1152.

Gould, Eric D., B. Peter Pashigian, and Canice J. Prendergast (2005). "Contracts, externalities, and incentives in shopping malls." *Review of Economics and Statistics* 87.3, 411-422.

Helsley, Robert W. and William C. Strange (1997), "Limited Developers," *Canadian Journal of Economics*, May, 329-348.

Henderson, J. Vernon, Ari Kuncoro and Matthew Turner (1995). "Industrial Development in Cities." *Journal of Political Economy*, 103, 1067-1090.

Jales, Hugo, Boqian Jiang and Stuart S. Rosenthal (2022), "Separating Selection From Spillover Effects: Using the Mode to Estimate the Return to City Size." Working paper.

Konishi, Hideo, and Michael T. Sandfort (2003). "Anchor stores." Journal of Urban Economics 53.3, 413-435.

Krugman, Paul, "Localization," Chapter 2 in Geography and Trade, MIT Press, 33-67, (1991).

Liu, Crocker, Stuart S. Rosenthal and William C. Strange (2018). "The Vertical City: Rent Gradients, Spatial Structure, and Agglomeration Economies", *Journal of Urban Economics*, 106, 101-122.

Liu, Crocker, Stuart S. Rosenthal and William C. Strange (2020). "Employment Density and Agglomeration Economies in Tall Buildings," *Regional Science and Urban Economics*, Vol 84. Paper number: 103555.

Liu, Crocker, Stuart S. Rosenthal and Strange C. William (2024), "Code for JUE paper Agglomeration Economies and the Built Environment: Evidence from Specialized Buildings and Anchor Tenants", Mendeley Data, V1, doi: 10.17632/2ddr8wbnpb.1.

Pashigian, B. and Gould, E. (1998). Internalizing Externalities: The Pricing of Space in Shopping Malls. *The Journal of Law and Economics*, 41(1), 115–142.

Rosenthal, Stuart S. and William C. Strange (2001). "The Determinants of Agglomeration," *Journal of Urban Economics*, 50(2): 191-229.

Rosenthal, Stuart and William Strange (2003), "Geography, Industrial Organization, and Agglomeration," *The Review of Economics and Statistics*, 85 (2): 377-393.

Rosenthal, Stuart S. and William Strange (2004), "Evidence on the Nature and Sources of Agglomeration Economies", *in* the Handbook of Urban and Regional Economics, Volume 4, pg. 2119-2172, Elsevier, eds. Vernon Henderson and Jacques Thisse.

Rosenthal, Stuart and William Strange (2005), "The Geography of Entrepreneurship in the New York Metropolitan Area," *Economic Policy Review*, (Special Issue on "Urban Dynamics in New York City"). New York Federal Reserve Bank, December, 11(2), 29-54.

Rosenthal, Stuart S. and William Strange (2006), "The Micro-Empirics of Agglomeration Economies", *in* Companion to Urban Economics, Blackwell, Richard Arnott and Dan McMillen (eds.), pp 7-23, Blackwell.

Rosenthal, Stuart, and William Strange (2008), "The Attenuation of Human Capital Spillovers," *Journal of Urban Economics*, 64(2), 373-389.

Rosenthal, S.S. and Strange, W.C. (2020). How Close Is Close? The Spatial Reach of Agglomeration Economies. *Journal of Economic Perspectives*, 34 (3): 27–49.

Rosenthal, Stuart S. and Joaquin A. Urrego (2022), "Eyes on the Street, Spatial Concentration of Retail Activity and Crime," Working paper.

Sandvik, Jason J, Richard E. Saouma, Nathan T. Seegert, and Christopher T. Stanton (2020). "Workplace Knowledge Flows," *Quarterly Journal of Economics*, 135(3), 1635-1680.

Walls, Donald W. (October 2007). "National Establishment Time-Series Database: Data Overview," SSRN working paper 1022962, Walls & Associates.

Table 1a: Frequency and Composition of Buildings

Panel A: Frequency of buildings by city and building size[a]	All Buildings With 10 or More Workers		Buildings With 10 to 100 Workers		Buildings With 100 or More Workers	
	Obs	**Mean**	**Obs**	**Mean**	**Obs**	**Mean**
Chicago	20,915	0.197	18,832	0.197	2,083	0.197
Los Angeles	38,636	0.365	34,574	0.362	4,062	0.384
New York	41,517	0.392	37,623	0.394	3,894	0.368
San Francisco	3,839	0.036	3,492	0.037	347	0.033
Washington DC	1,075	0.010	881	0.009	194	0.018
Total	105,982	1.00	95,402	1.00	10,580	1.00

Panel B: Employment shares within 0.1 mile by industry	All Buildings With 10 or More Workers		Buildings With 10 to 100 Workers		Buildings With 100 or More Workers	
	Incl Other	Without Other	Incl Other	Without Other	Incl Other	Without Other
Retail	0.112	0.381	0.114	0.394	0.093	0.277
Finance	0.032	0.109	0.031	0.107	0.045	0.134
Advertising	0.005	0.017	0.005	0.017	0.009	0.027
Law	0.011	0.037	0.010	0.035	0.019	0.057
Software/Data Processing	0.016	0.054	0.015	0.052	0.023	0.068
Engineering/Management	0.046	0.156	0.045	0.156	0.057	0.170
Manufacturing	0.071	0.241	0.069	0.239	0.089	0.265
Target industries employment share	0.293	1.000	0.289	1.000	0.335	1.000
Other-Industry employment share	0.707	-	0.711	-	0.665	-

Table 1b: Share of Buildings in Which Anchors are Present (Panel A), Industries are Present (Panel B) And for Which an Industry Occupies the Entire Building(Panel C)[a]

	Buildings With 10 or More Workers		Buildings With 10 to 100 Workers		Buildings With 100 or More Workers	
Panel A: Anchor Establishment Present[a]	**All Buildings**	**Anchor Present**	**All Buildings**	**Anchor Present**	**All Buildings**	**Anchor Present**
Retail	0.011	0.082	0.006	0.081	0.052	0.078
Finance	0.004	0.030	0.002	0.027	0.025	0.037
Advertising	0.001	0.007	0.000	0.000	0.005	0.007
Law	0.001	0.007	0.000	0.000	0.007	0.010
Software/Data Processing	0.003	0.022	0.001	0.014	0.017	0.025
Engineering/Management	0.006	0.045	0.003	0.041	0.032	0.048
Manufacturing	0.014	0.104	0.008	0.108	0.075	0.112
Other Industries	0.094	0.701	0.054	0.730	0.457	0.682
Total	0.134	1.000	0.074	1.000	0.670	1.000
Share of bldgs. emp > 50 & no anchor	0.065		0.035		0.330	
Share of bldgs emp < 50	0.801		0.890		0.000	
Buildings	105,982		95,402		10,580	

Panel B: Industry Present (Not Including Anchor Establishments)	**All Buildings With 10 or More Workers**	**Buildings With 10 to 100 Workers**	**Buildings With 100 or More Workers**
Retail	0.344	0.326	0.493
Finance	0.147	0.119	0.402
Advertising	0.029	0.015	0.152
Law	0.067	0.045	0.261
Software/Data Processing	0.074	0.052	0.269
Engineering/Management	0.208	0.177	0.482
Manufacturing	0.171	0.146	0.401
Other Industries	0.906	0.904	0.920
Buildings	105,982	95,402	10,580

Panel C: Industry Occupies Entire Building	**All Buildings With 10 or More Workers**	**Buildings With 10 to 100 Workers**	**Buildings With 100 or More Workers**
Retail	0.0410	0.0424	0.0285
Finance	0.0036	0.0034	0.0054
Advertising	0.0006	0.0006	0.0005
Law	0.0008	0.0008	0.0007
Software/Data Processing	0.0282	0.0029	0.0017
Engineering/Management	0.0071	0.0071	0.0067
Manufacturing	0.0248	0.0243	0.0285
Other Industries	0.4107	0.4256	0.2761
Buildings	105,982	95,402	10,580

[a] Building level employment shares are reported for buildings with 50 or more workers. Employment shares for employment within 0.1 miles are based on all employment outside of the target building including employment in buildings with less than 50 workers.

Table 2: Correlation Between Industry Employment Shares in Target Building Relative to Nearby Buildings[a]

Panel A: Target buildings with 10 or more workers (105,982 buildings)	Within 0.1 Mile	On the Blockface	Across the Street
Retail	0.092	0.067	0.036
Finance	0.094	0.037	0.038
Advertising	0.045	0.022	0.023
Law	0.160	0.093	0.068
Software/Data Processing	0.092	0.047	0.035
Engineering/Management	0.064	0.033	0.024
Manufacturing	0.210	0.152	0.148
Other Industries	-0.026	0.004	-0.007

Panel B: Target buildings with 100 or more workers (10,580 buildings)	Within 0.1 Mile	On the Blockface	Across the Street
Retail	0.297	0.207	0.118
Finance	0.190	0.086	0.092
Advertising	0.088	0.061	0.067
Law	0.293	0.193	0.155
Software/Data Processing	0.158	0.086	0.085
Engineering/Management	0.112	0.053	0.046
Manufacturing	0.372	0.303	0.289
All-Other Industries	0.249	0.177	0.145

[a] Building level employment shares are based on buildings with 50 or more workers. Employment shares for employment within 0.1 miles are based on all employment outside of the target building including employment in buildings with less than 50 workers. The sample of target buildings is restricted to buildings with 10 or more workers.

Table 3a: Industry Composition Near the New York Stock Exchange[a]

Panel A: Within 0.25 Miles (321 bldgs)	Number of Establishments	Share of Establishments	Employment	Share of Employment	Share of Buildings with Industry Present
Retail	714	0.047	7,442	0.029	0.548
Finance	2,200	0.144	64,617	0.255	0.458
Advertising	112	0.007	1,750	0.007	0.156
Law	1,757	0.115	13,181	0.052	0.380
Software/Data Processing	666	0.044	10,897	0.043	0.386
Engineering/Management	1,875	0.123	18,416	0.073	0.526
Manufacturing	370	0.024	5,994	0.024	0.333
Other Industries	7,577	0.496	131,421	0.518	0.944
All Industries	15,272	1.000	253,718	1.000	-
Panel B: Within 1.0 Miles (1,714 bldgs)	**Number of Establishments**	**Share of Establishments**	**Employment**	**Share of Employment**	**Share of Buildings with Industry Present**
Retail	2,503	0.068	16,557	0.028	0.527
Finance	4,100	0.112	103,222	0.173	0.319
Advertising	222	0.006	4,951	0.008	0.072
Law	3,890	0.106	28,334	0.048	0.221
Software/Data Processing	1,277	0.035	17,835	0.030	0.205
Engineering/Management	3,887	0.106	32,962	0.055	0.402
Manufacturing	953	0.026	18,985	0.032	0.246
Other Industries	19,824	0.541	374,217	0.627	0.961
All Industries	36,656	1.000	597,063	1.000	-

[a] The samples for both panels include only buildings with 10 or more workers.

Table 3b: Spatial G of Industry Employment Across Buildings Close to the New York Stock Exchange[a,b]

	Buildings with 10 or more workers					Buildings with 100 or more workers				
		Simulated Spatial G (1,000 Repetitions)					Simulated Spatial G (1,000 Repetitions)			
Panel A: Within 0.25 Miles (321/143 Buildings)	**Actual G[c]**	**Mean**	**1-Pvalue Actual G > Mean**	**90th Pctl**	**95th Pctl**	**Actual G[c]**	**Mean**	**1-Pvalue Actual G > Mean**	**90th Pctl**	**95th Pctl**
Retail	0.118**	0.092	0.937	0.108	0.118	0.131**	0.091	0.991	0.107	0.116
Finance	0.130**	0.073	1.000	0.094	0.102	0.130**	0.073	1.000	0.094	0.103
Advertising	0.087**	0.060	0.973	0.077	0.084	0.090**	0.061	0.987	0.077	0.082
Law	0.046**	0.033	0.958	0.042	0.045	0.048**	0.033	0.984	0.042	0.044
Software/Data Processing	0.065**	0.040	0.996	0.051	0.054	0.067**	0.041	0.998	0.052	0.056
Engineering/Management	0.033**	0.021	0.984	0.028	0.031	0.034**	0.021	0.995	0.028	0.030
Manufacturing	0.054	0.048	0.709	0.062	0.068	0.055	0.047	0.747	0.062	0.067
Other Industries	0.017**	0.010	0.977	0.014	0.016	0.018**	0.010	0.996	0.014	0.016
Panel B: Within 1.0 Miles (1,713/412 Buildings)	**Actual G[c]**	**Mean**	**1-Pvalue Actual G > Mean**	**90th Pctl**	**95th Pctl**	**Actual G[c]**	**Mean**	**1-Pvalue Actual G > Mean**	**90th Pctl**	**95th Pctl**
Retail	0.045	0.045	0.517	0.054	0.057	0.070**	0.044	1.000	0.054	0.056
Finance	0.087**	0.038	0.981	0.064	0.068	0.090**	0.038	0.987	0.063	0.069
Advertising	0.111**	0.071	1.000	0.082	0.085	0.123**	0.070	1.000	0.080	0.083
Law	0.035	0.035	0.505	0.044	0.046	0.039	0.035	0.743	0.044	0.046
Software/Data Processing	0.044	0.039	0.743	0.049	0.051	0.049*	0.039	0.914	0.048	0.051
Engineering/Management	0.024	0.030	0.810	0.039	0.041	0.028	0.030	0.599	0.039	0.041
Manufacturing	0.058	0.066	0.720	0.078	0.081	0.079*	0.065	0.873	0.077	0.080
Other Industries	0.009*	0.006	0.857	0.009	0.010	0.011**	0.006	0.956	0.009	0.010

[a] Empirical distributions were calculated based on 1,000 simulations in which establishments were randomly reassigned to buildings allowing for space constraints.
[b] The spatial G is calculated as $G_i \equiv \sum_b (S_{bi}^N - S_b^N)^2$, where S_{bi}^N and S_b^N are respectively building *b*'s share of industry-*i* employment in neighborhood *N* and the building's share of total employment in the community. Specified in this manner G_i ranges equals zero when industry *i* employment is distributed across buildings as for total employment and converges to one as industry *i* employment becomes concentrated in a single building.
[c] One star indicates that actual G (column 1 of each panel) exceeds the 90th percentile of the simulated distribution (column 4 of each panel) and two stars indicates that it exceeds the 95th percentile (column 5 of each panel). The likelihood that actual G exceeds the mean of the simulated distribution (column 3) is given by 1 – Pvalue based on a t-distribution with 1,000 degrees of freedom and a 1-tailed test. As an example, a value of 0.981 in the Retail row in the upper right quadrant indicates that actual G exceeds the simulated mean with 98.1% confidence.

Table 4a: Fixed Effect Models Including All Industries
For Industry Employment Share Outside of Own-Industry Anchor[a]

	Block Fixed Effects			Building Fixed Effects		
	Buildings With 10 or More Workers	Buildings With 10 to 100 Workers	Buildings With 100 or More Workers	Buildings With 10 or More Workers	Buildings With 10 to 100 Workers	Buildings With 100 or More Workers
Own Ind Emp share < 0.1 miles	0.3115	0.2996	0.3458	0.3115	0.2996	0.3458
	(38.00)	(33.03)	(24.25)	(57.99)	(51.22)	(25.47)
Own Ind Anchor in Target Bldg[b]	0.0839	0.0693	0.1431	0.0959	0.0792	0.1636
	(28.77)	(17.62)	(28.00)	(30.12)	(18.02)	(28.67)
Own Ind Anchor on Blockface[b]	0.0240	0.0226	0.0118	0.0274	0.0258	0.0135
	(24.38)	(22.69)	(3.42)	(30.26)	(27.56)	(3.63)
Own Ind Anchor Across the Street[b]	0.0161	0.0148	0.0025	0.0184	0.0170	0.0028
	(16.55)	(14.71)	(0.69)	(18.94)	(16.86)	(0.74)
Outside-Own Ind Anchor in Target Bldg[b]	-0.0120	-0.0099	-0.0204	-	-	-
	(-28.77)	(-17.62)	(-28.00)	-	-	-
Outside-Own Ind Anchor on Blockface[b]	-0.0034	-0.0032	-0.0017	-	-	-
	(-24.38)	(-22.69)	(-3.42)	-	-	-
Outside-Own Ind Anchor Across the Street[b]	-0.0023	-0.0021	-0.0004	-	-	-
	(-16.55)	(-14.71)	(-0.69)	-	-	-
Industry Group Fixed Effects	8	8	8	8	8	8
Block Fixed Effects	27,990	27,052	5,821	-	-	-
Building Fixed Effects	-	-	-	106,743	96,058	10,685
Observations	853,944	768,464	85,480	853,944	768,464	85,480
Within R-squared	0.625	0.625	0.639	0.625	0.625	0.639

[a] Sample includes buildings with 10 or more workers in Manhattan, Cook County (Chicago), Los Angeles, San Francisco, Washington DC. T-ratios based on robust standard errors in parentheses.
[b] An establishment is coded as an anchor (1 if yes, 0 if no) if all three of the following conditions hold: (i) it is the largest establishment in the building; (ii) it includes at least 20% of building employment; (iii) it has 10 or more workers.

Table 4b: Fixed Effect Models Including Only Office Industries
For Industry Employment Share Outside of Own-Industry Anchor[a]

	Block Fixed Effects			**Building Fixed Effects**		
	Buildings With 10 or More Workers	Buildings With 10 to 100 Workers	Buildings With 100 or More Workers	Buildings With 10 or More Workers	Buildings With 10 to 100 Workers	Buildings With 100 or More Workers
Own Ind Emp share < 0.1 miles	0.0867	0.0709	0.1581	0.1012	0.0853	0.1659
	(12.55)	(11.31)	(8.81)	(18.64)	(14.52)	(11.84)
Own Ind Anchor in Target Bldg[b]	0.0992	0.0771	0.1038	0.0912	0.0681	0.1046
	(17.35)	(9.46)	(13.34)	(15.20)	(8.00)	(12.57)
Own Ind Anchor on Blockface[b]	-0.0008	0.0003	-0.0020	0.0036	0.0042	0.0021
	(-1.13)	(0.41)	(-0.70)	(5.32)	(6.08)	(0.67)
Own Ind Anchor Across the Street[b]	-0.0018	-0.0011	0.0015	0.0025	0.0027	0.0051
	(-2.67)	(-1.49)	(0.51)	(3.52)	(3.84)	(1.57)
Outside-Own Ind Anchor in Target Bldg[b]	0.0011	0.0028	-0.0054	-	-	-
	(2.18)	(4.38)	(-3.92)	-	-	-
Outside-Own Ind Anchor on Blockface[b]	-0.0020	-0.0015	-0.0030	-	-	-
	(-14.13)	(-9.81)	(-4.55)	-	-	-
Outside-Own Ind Anchor Across the Street[b]	-0.0020	-0.0015	-0.0024	-	-	-
	(-14.36)	(-9.77)	(-3.35)	-	-	-
Industry Group Fixed Effects	5	5	5	5	5	5
Block Fixed Effects	27,990	27,052	5,821	-	-	-
Building Fixed Effects	-	-	-	106,743	96,058	10,685
Observations	533,715	480,290	53,425	533,715	480,290	53,425
Within R-squared	0.032	0.026	0.078	0.034	0.029	0.081

[a] Sample includes buildings with 10 or more workers in Manhattan, Cook County (Chicago), Los Angeles, San Francisco, Washington DC. T-ratios based on robust standard errors in parentheses.
[b] An establishment is coded as an anchor (1 if yes, 0 if no) if all three of the following conditions hold: (i) it is the largest establishment in the building; (ii) it includes at least 20% of building employment; (iii) it has 10 or more workers.

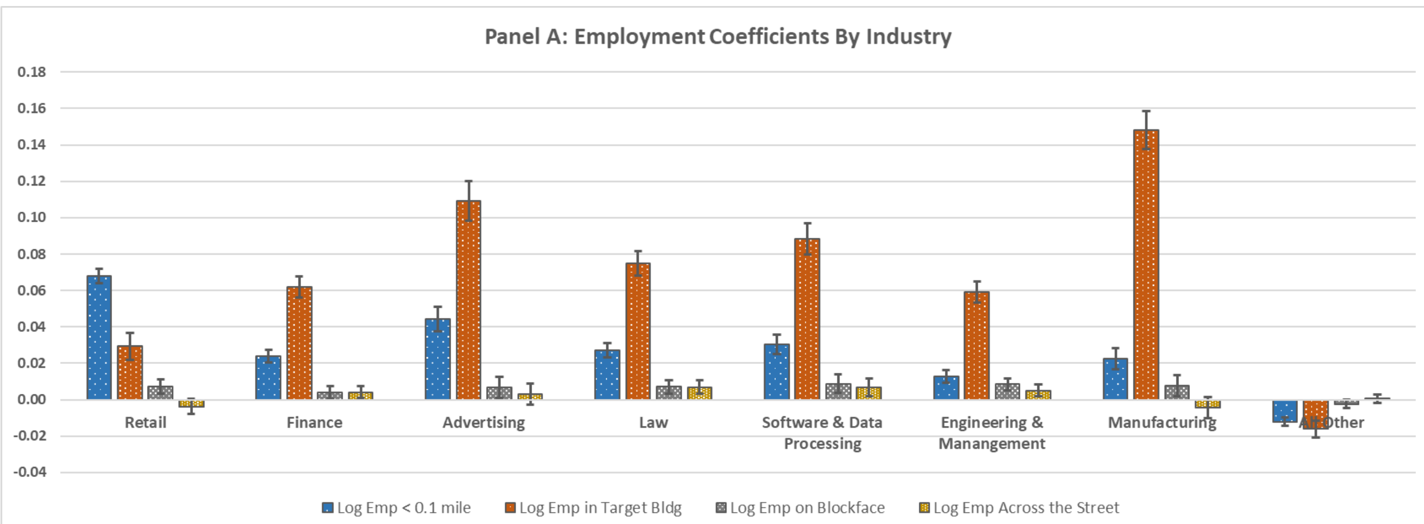
Table 5: Tobit Model Employment Share Within Target Building (not including own-industry anchor)[a]

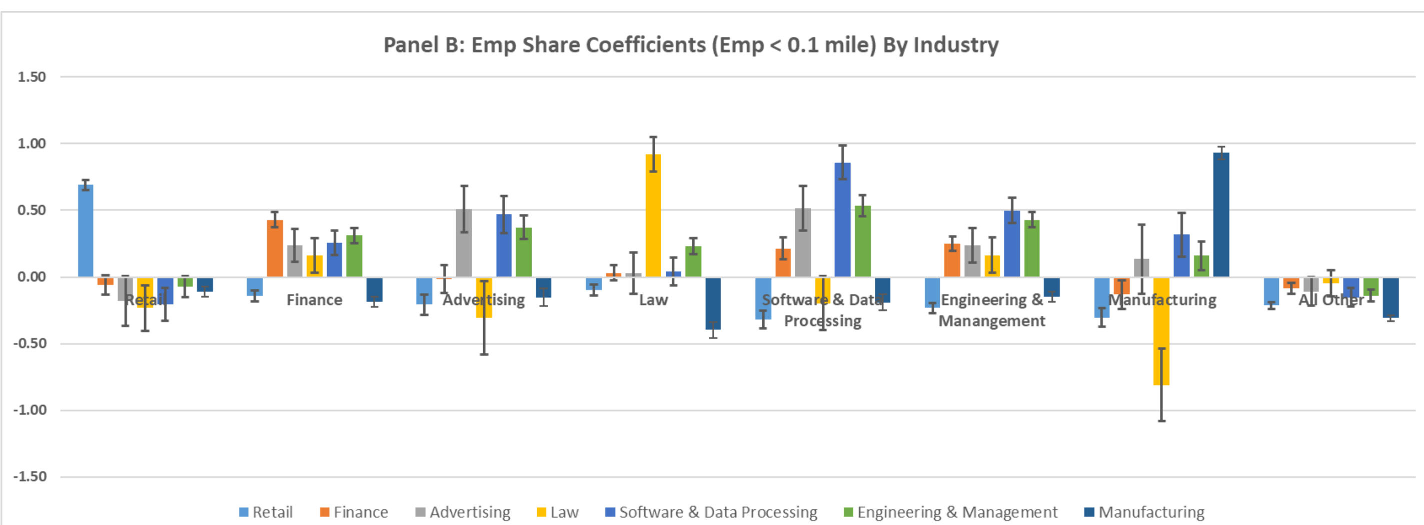
Regression Model Summary Statistics and Notes[a]

	Retail	Finance	Advertising	Law	Software & Data Proc	Eng & Mgmt	Manf	All Other
City FE	5	5	5	5	5	5	5	5
Observations	105,982	105,982	105,982	105,982	105,982	105,982	105,982	105,982
Censored	69,615	90,399	102,958	98,945	98,154	84,014	87,834	9,996
Uncensored	36,367	15,583	3,024	7,037	7,828	21,968	18,148	95,986
Pseudo R-sq	0.0564	0.103	0.227	0.189	0.141	0.0790	0.114	0.0358

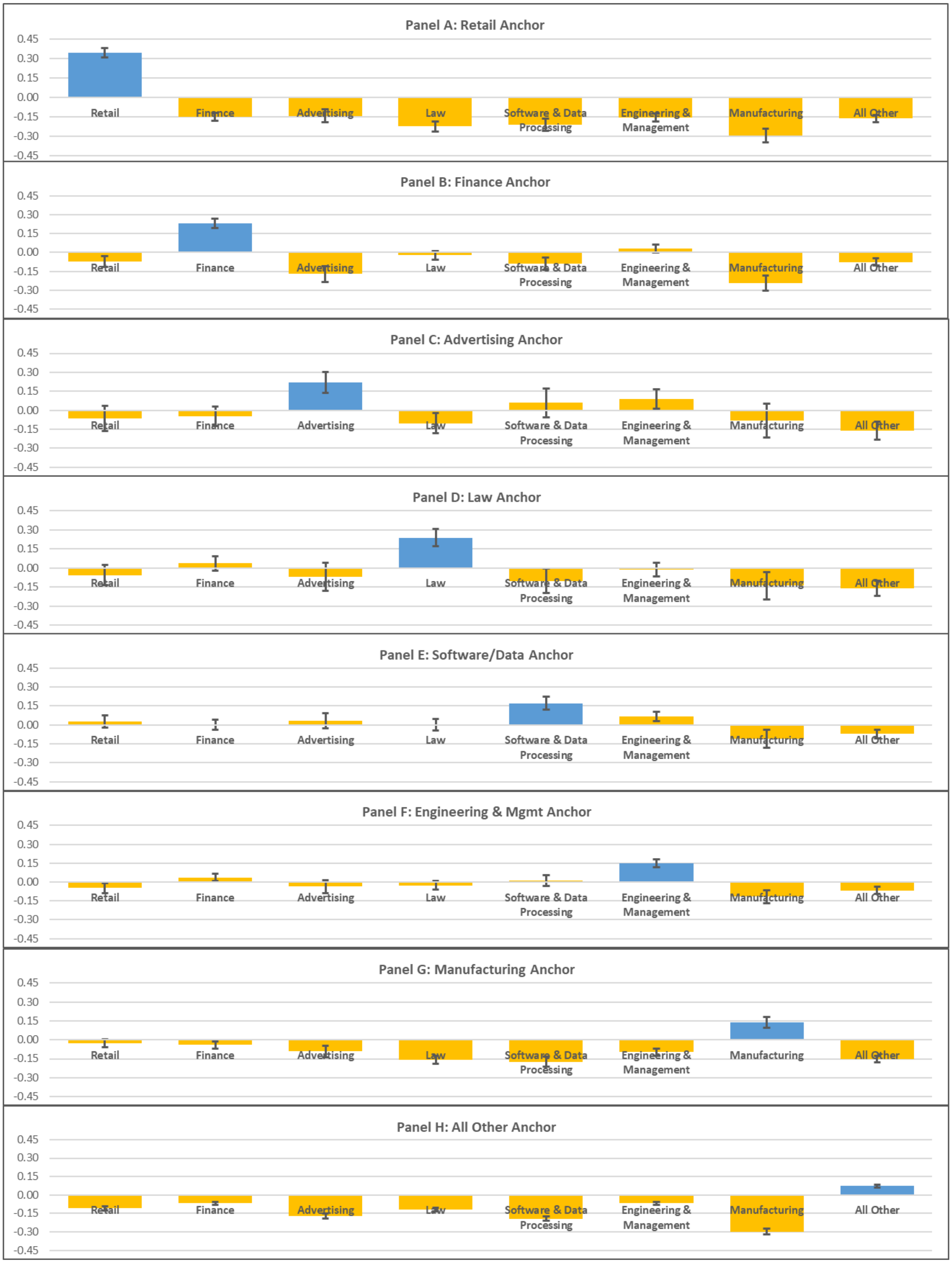
[a] Sample includes buildings with 10 or more workers in the following counties: New York (Manhattan), Cook (Chicago) (Chicago), Los Angeles, San Francisco, Washington DC. Confidence bands based on robust standard errors.
[b] An establishment is coded as an anchor (1 if yes, 0 if no) if all three of the following conditions hold: (i) it is the largest establishment in the building; (ii) it includes at least 20% of building employment; (iii) it has 10 or more workers.

Group 1 Exhibit: Employment and Employment Share Coefficients





Group 2 Exhibit: Target Building Anchor Coefficients



Group 3 Exhibit: Own-Industry Anchor Coefficients Organized by Location and Industry

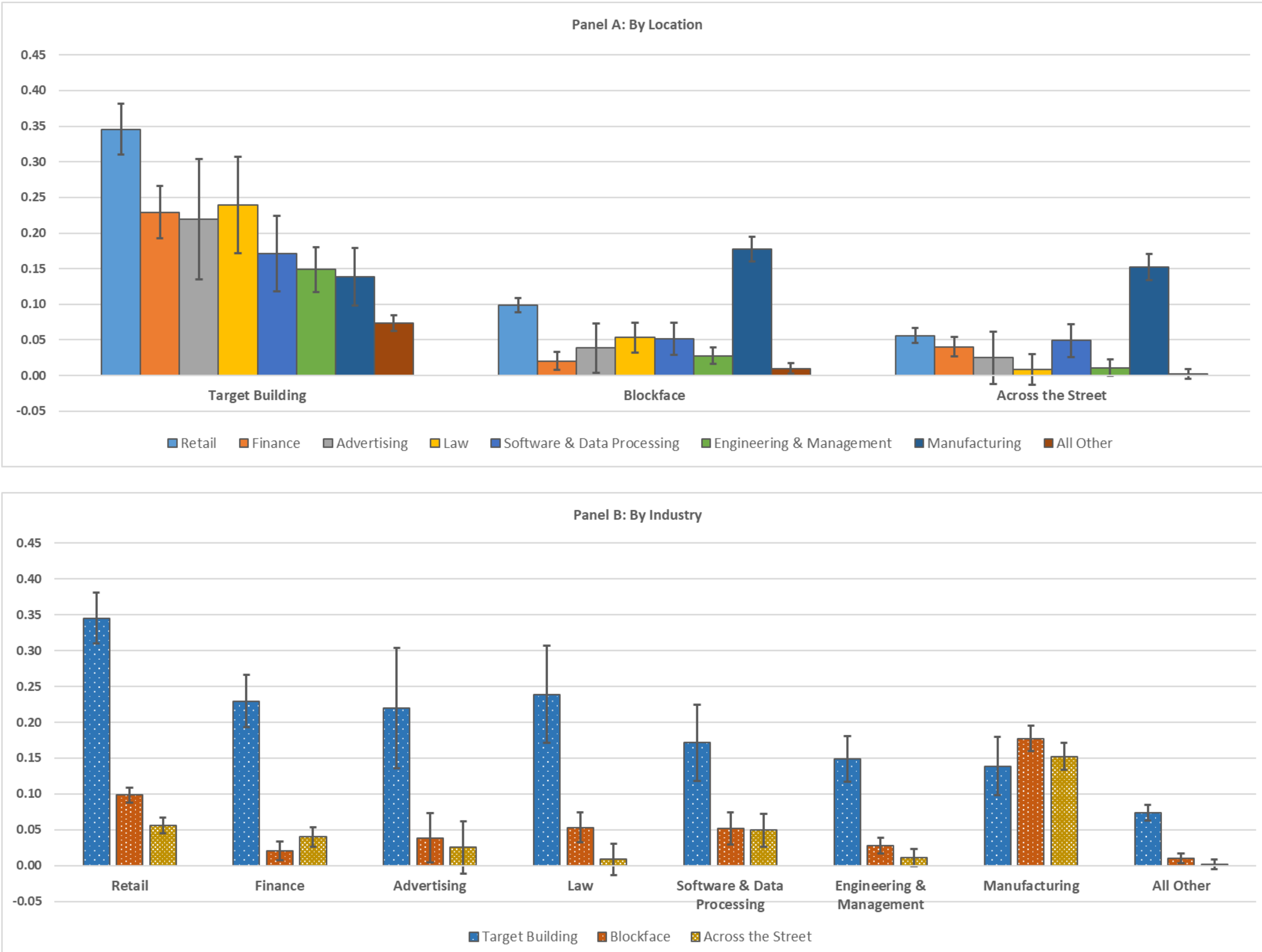


Table 6: Tobit Models of Building Level Average Sales/Worker Not Including Own-Industry Anchors in the Target Building[a]

		Retail			Law	
Log emp < 0.1 miles		11,916			11,108	
		(28.76)			(15.35)	
Log emp in target building		12,522			32,834	
		(15.58)			(24.58)	
Log emp on blockface		1,128			4,079	
		(2.76)			(5.57)	
Log emp across the street		-107.32			2,942	
		(-0.27)			(4.07)	
Emp share < 0.1 Miles						
Retail[b]		43,505			-34,810	
		(11.67)			(-4.39)	
Finance[b]		-3,931			25,316	
		(-0.53)			(2.08)	
Advertising[b]		-28,596			14,056	
		(-1.47)			(0.48)	
Law[b]		-32,965			284,437	
		(-1.92)			(11.95)	
Comp/Software[b]		-15,041			26,441	
		(-1.21)			(1.31)	
Eng/Mgmt[b]		5,397			91,278	
		(0.72)			(8.25)	
Manufacturing[b]		-9,849			-155121	
		(-2.50)			(-13.76)	
Anchor Establishments	**Target Building**	**Blockface**	**Across Street**	**Target Building**	**Blockface**	**Across Street**
Retail[c]	12,140	11,238	7,421	-93,532	-16,155	-5,974
	(3.54)	(10.71)	(7.05)	(-12.26)	(-7.11)	(-2.61)
Finance[c]	-16,119	-5,005	-3,433	-15,118	2,223	300.06
	(-3.34)	(-3.61)	(-2.26)	(-2.00)	(0.76)	(0.10)
Advertising[c]	-12,464	-6,019	1,213	-48,190	-5,754	-1,980
	(-0.95)	(-2.79)	(0.53)	(-2.88)	(-1.35)	(-0.43)
Law[c]	-6,803	-4,679	-2,202	19,247	14,415	1,647
	(-0.65)	(-2.50)	(-1.05)	(1.66)	(3.78)	(0.41)
Comp/Software[c]	9,030	-2,927	-332.95	-3,910	56.395	2,349
	(1.45)	(-1.83)	(-0.20)	(-0.44)	(0.02)	(0.68)
Eng/Mgmt[c]	-9,521	-7,069	-5,530	-18,185	3,849	-2,129
	(-2.12)	(-5.48)	(-4.06)	(-2.76)	(1.45)	(-0.74)
Manufacturing[c]	-15,047	-2,029	-1,676	-68,136	-15,106	-17,172
	(-4.24)	(-1.63)	(-1.28)	(-10.58)	(-5.55)	(-6.16)
Other[c]	-26,303	-85.607	1,265	-56,263	-3,394	-3,789
	(-14.25)	(-0.07)	(1.02)	(-20.61)	(-1.40)	(-1.59)
Small Building (Emp < 50)	-7,308	-	-	-43,722	-	-
	(-3.75)	-	-	(-14.39)	-	-
Sigma		1.0698e+10			1.5577e+10	
		(46.03)			(40.03)	
City Fixed Effects		5			5	
Observations		105,925			105,977	
Censored		73,133			99,056	
Uncensored		32,792			6,921	

[a] Sample includes buildings with 10 or more workers in New York County (Manhattan), Cook County (Chicago), Los Angeles MSA, San Francisco County, Washington DC County. Building level sales per worker is calculated omitting own-industry anchor establishments within the target building. T-ratios based on robust standard errors.

[b] Omitted category is retail share of employment.

[c] An establishment is coded as an anchor (1 if yes, 0 if no) if all three of the following conditions hold: (i) it is the largest establishment in the building; (ii) it includes at least 20% of building employment; (iii) it has 10 or more workers.

Appendix: Frequency Counts of Establishments by SIC 2 Industries

Table A-1: Frequency Counts of Establishments by Industry Groupings and SIC 2 Subcategories (775,638 establishments across all industries with total employment of 7,381,376)[a]

Panel A: Retail (10.09% of total)

SIC 2	Freq.	Prcnt.	Cum.
52	2,706	3.46	3.46
53	2,232	2.85	6.31
54	15,383	19.66	25.97
56	17,556	22.44	48.41
57	8,974	11.47	59.88
59	31,390	40.12	100
Total	**78,241**	**100**	

Panel B: Finance (5.87% of total)

SIC 2	Freq.	Prcnt.	Cum.
62	8,833	19.40	19.40
63	1,585	3.48	22.88
6512 -6519	15,423	33.87	56.74
67	19,701	43.26	100
Total	**45,542**	**100**	

Panel C: Advertising (0.60% of total)

SIC 2	Freq.	Prcnt.	Cum.
7311 - 7313 & 7319	4,641	100	100
Total	**4,641**	**100**	

Panel D: Law (3.96% of total)

SIC 2	Freq.	Prcnt.	Cum.
81	30,702	100	100
Total	**30,702**	**100**	

Panel E: Software/Data Processing (2.03% of total)

SIC 2	Freq.	Prcnt.	Cum.
7370 - 7379	15,774	100	100
Total	**15,774**	**100**	

Panel F: Management/Engineering (7.93% of total)

SIC 2	Freq.	Prcnt.	Cum.
87	61,493	100	100
Total	**61,493**	**100**	

Panel G: Manufacturing (4.44% of total)

SIC 2	Freq.	Prcnt.	Cum.
20	1,721	4.99	4.99
21	28	0.08	5.07
22	1,373	3.98	9.06
23	4,225	12.26	21.32
24	849	2.46	23.78
25	804	2.33	26.11
26	499	1.45	27.56
27	7,814	22.67	50.23
28	1,645	4.77	55.00
29	106	0.31	55.31
30	581	1.69	57.00
31	568	1.65	58.65
32	600	1.74	60.39
33	451	1.31	61.70
34	1,889	5.48	67.18
35	2,427	7.04	74.22
36	2,015	5.85	80.06
37	588	1.71	81.77
38	1,308	3.80	85.57
39	4,975	14.43	100
Total	**34,466**	**100**	

Continued on next page

Table A-1: Frequency Counts of Establishments by Industry Groupings and SIC 2 Subcategories[a]

Panel H: Other Industries (65.08% of total)				**Panel H continued: Other Industries**			
SIC 2	**Freq.**	**Prcnt.**	**Cum.**	**SIC 2**	**Freq.**	**Prcnt.**	**Cum.**
1	752	0.15	0.15	60	4,145	0.82	31.18
2	381	0.08	0.22	61	5,167	1.02	32.21
7	2,695	0.53	0.76	64	7,911	1.57	33.77
8	42	0.01	0.77	65 (if not in Panel B)	26,480	5.25	39.02
9	51	0.01	0.78	70	3,096	0.61	39.63
10	53	0.01	0.79	72	25,194	4.99	44.62
12	11	0.00	0.79	73 (if not in Panels C & E)	79,907	15.83	60.45
13	239	0.05	0.84	75	9,767	1.93	62.39
14	84	0.02	0.85	76	7,407	1.47	63.86
15	11,694	2.32	3.17	78	16,209	3.21	67.07
16	669	0.13	3.30	79	17,625	3.49	70.56
17	14,911	2.95	6.26	80	63,189	12.52	83.08
40	74	0.01	6.27	82	9,584	1.90	84.98
41	3,363	0.67	6.94	83	13,455	2.67	87.64
42	5,096	1.01	7.95	84	1,561	0.31	87.95
43	256	0.05	8.00	86	25,661	5.08	93.03
44	436	0.09	8.08	89	9,039	1.79	94.82
45	797	0.16	8.24	91	509	0.10	94.92
46	6	0.00	8.24	92	683	0.14	95.06
47	10,338	2.05	10.29	93	75	0.01	95.08
48	11,186	2.22	12.51	94	300	0.06	95.13
49	1,513	0.30	12.81	95	371	0.07	95.21
50	23,591	4.67	17.48	96	258	0.05	95.26
51	24,259	4.81	22.29	97	613	0.12	95.38
55	4,814	0.95	23.24	99	23,318	4.62	100.00
58	35,944	7.12	30.36	**Total**	**504,779**	**100**	

*Sample includes establishments used in Table 5.

ONLINE APPENDICES

Appendix A: Coding Buildings to Street and Blockface Locations

1. Overview

This appendix describes the data sources for the shapefiles used to code buildings to their respective street and blockface locations. Also discussed are the steps followed to prepare the shapefiles for these purposes. This is followed by an overview of how Stata code was used to merge the shapefile and Dun and Bradstreet data, and to code each building's location.

In all instances, ArcMap (ArcGIS) was used to clean the shapefiles although other GIS software could be used. Separate shapefiles were obtained for each of the five cities used in our analysis, each of which was cleaned interactively in a manner that allows for map-specific details and structure. Section 2 below provides the source URL for the shapefiles for each of the five cities. Section 3 describes procedures used to clean the shapefiles and to assign buildings to street block and blockface locations. The cleaned shapefiles and all related programs can be shared. The Dun and Bradstreet establishment data cannot be shared because of terms of the Syracuse University site license.

2. Data sources for the shapefiles

Shapefiles used in the analysis were downloaded from publicly available sources over the web. The data sources are below.

<u>New York City (five boroughs)</u>

Overall Data Source: NYC Planimetric Database

URL for the Planimetric database:
https://github.com/CityOfNewYork/nyc-planimetrics/blob/master/Capture_Rules.md

URL for the street map shapefile:
https://data.cityofnewyork.us/City-Government/NYC-Street-Centerline-CSCL-/exjm-f27b

Opening the shapefile produces the following files:

street_centerlines.sbn	street_centerlines.sbx
street_centerlines.shp	street_centerlines.shp.xml
street_centerlines.shx	street_centerlines.cpg
street_centerlines.dbf	street_centerlines.prj

URL for the blockface shapefile:

https://data.cityofnewyork.us/dataset/Sidewalk-Centerline/a9xv-vek9

Opening the shapefile produces the following files:

Blockface.shx	blockface_xy.txt
MidPoint_Blockface.cpg	MidPoint_Blockface.dbf
MidPoint_Blockface.prj	MidPoint_Blockface.shp
MidPoint_Blockface.shx	Blockface.cpg
Blockface.dbf	Blockface.prj
Blockface.sbn	Blockface.sbx
Blockface.shp	

Washington DC

Overall Data Source: Open Data City of Washington, DC

URL for the Open Data portal for Washington DC:

https://opendata.dc.gov/datasets/roadway-subblock/explore?location=38.946243%2C-77.065798%2C15.92

URL for the street map shapefile:

https://opendata.dc.gov/datasets/roadway-subblock/explore?location=38.946243%2C-77.065798%2C15.92

Opening the shapefile produces the following files:

Street_centerline_segments.prj	Street_centerline_segments.sbn
Street_centerline_segments.sbx	Street_centerline_segments.shp
Street_centerline_segments.shx	Street_centerline_segments.xml
Street_centerline_segments.cpg	Street_centerline_segments.dbf

URL for the blockface shapefile:
https://opendata.dc.gov/datasets/47945b50c4f245b58850e81d297e90b9_164/explore?location=38.900078%2C-77.035013%2C16.78

Opening the shapefile produces the following files:

Blockface_xy.txt.xml	MidPoint_Blockface.cpg"
MidPoint_Blockface.dbf	MidPoint_Blockface.prj
MidPoint_Blockface.shp	MidPoint_Blockface.shx
MidPoint_Blockface.xml	Blockface.cpg
Blockface.dbf	Blockface.prj
Blockface.sbn	Blockface.sbx
Blockface.shp	Blockface.shx
Blockface.xml	Blockface_xy.txt

San Francisco

Overall Data Source: San Francisco Municipal Transportation Agency

URL for San Francisco municipal transportation agency data:

https://datasf.org/

URL for the street map shapefile:

https://data.sfgov.org/Geographic-Locations-and-Boundaries/Streets-Active-and-Retired/3psu-pn9h

Opening the shapefile produces the following files:

street_centerlines.prj	street_centerlines.sbn
street_centerlines.sbx	street_centerlines.shp
street_centerlines.shp.xml	street_centerlines.shx
street_centerlines.cpg	street_centerlines.dbf

URL for the blockface shapefile:

https://data.sfgov.org/Transportation/Blockfaces/pep9-66vw

Opening the shapefile produces the following files:

Blockface.sbx	Blockface.shp
Blockface.shx	Blockface.xml
Blockface_xy.csv	Blockface_xy.txt.xml
MidPoint_Blockface.cpg	MidPoint_Blockface.dbf
MidPoint_Blockface.prj	MidPoint_Blockface.shp
MidPoint_Blockface.shx	MidPoint_Blockface.xml
Blockface.cpg	Blockface.dbf
Blockface.prj	Blockface.sbn

Chicago (Cook County):

Overall Data Source: Cook County GIS

URL for Cook County GIS data:

https://hub-cookcountyil.opendata.arcgis.com/datasets/4569d77e6d004c0ea5fada54640189cf_5/explore?location=41.812536%2C-87.874200%2C9.75

Opening the shapefile produces the following files:

Streets.sbn	Streets.sbx
Streets.shp	Streets.shp.xml
Streets.shx	Streets.cpg
Streets.dbf	Streets.prj

Los Angeles

Overall Data Source: City of Los Angeles Geo Hub

URL for the City of Los Angeles Geo Hub:
https://geohub.lacity.org/datasets/6b7e5c319b5543fcb35b8507c3b7e2bf_34/explore?location=34.023750%2C-118.411050%2C9.84

Opening the shapefile produces the following files:

Street_segments\Streets.prj	Street_segments\Streets.shp
Street_segments\Streets.shp.xml	Street_segments\Streets.shx
Street_segments\Streets.cpg	Street_segments\Streets.dbf

3. Preparing the shapefiles and assigning buildings to street and blockface locations

For each of the five cities above, the shapefiles provide detailed information on the line segments that make up a city's street network. Figure B-1 illustrates. Panel A highlights a street block segment in blue that is defined as the street line between two intersections. For comparison, Panel B highlights all of the street segments (in blue) that belong to a given street, running several blocks in length. Panel C highlights an individual blockface on a single city block, while Panel D highlights all four blockfaces that define the perimeter of the block. We coded buildings to their respective street blocks and blockfaces as defined in Panels A and C. This is possible because each street and blockface line segment is treated as a separate record in the shapefiles, corresponding to a separate row in the table attributes of the file, with latitude/longitude coordinates and street name in columns. In instances where a blockface shapefile is not available (Chicago and Los Angeles), street address information from Dun and Bradstreet is used to assign buildings to their respective blockface in a manner described below.

To clean and organize the shapefile data, we first assigned a unique identifying number to each street block and blockface line segment. This was done interactively in a single step for each map using ArcGIS. For most line segment records, the shapefiles include data fields in tabular form for the latitude and longitude of the end points of the segment, and in some instances the latitude/longitude coordinates of the line segment's midpoint. For records for which that information was missing, ArcGIS was used to fill in the missing coordinate information, and once again interactively in a single step for each map. Following these steps, all line segments in the cleaned shapefiles files contain complete latitude/longitude coordinate information as part of their records. These files were then saved.

The remaining steps used to assign buildings to street blocks and blockfaces were programmed in Stata. From Dun and Bradstreet, the set of building addresses reported in the establishment records was first determined. Those addresses were then merged with the cleaned shapefiles, customizing the code for each city to allow for city-specific differences in the structure of the GIS files. The Dun and Bradstreet establishment address records include street name and number as well as address latitude and longitude coordinates.

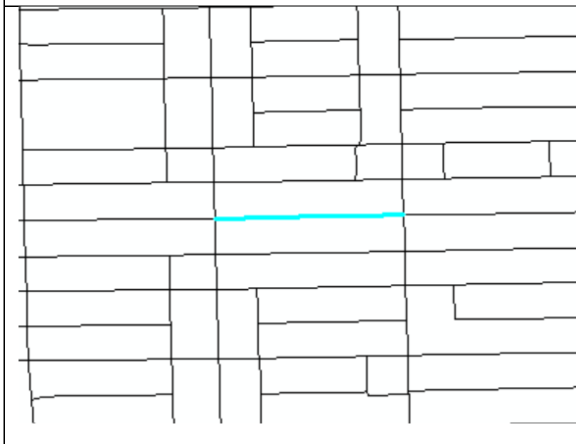
The precision of the location information is high. For the shapefiles, latitude/longitude precision is within one foot. For the Dun and Bradstreet data, latitude/longitude coordinates for a building address are accurate to within 3 to 30 feet.

For New York City, Washington DC, and San Francisco, blockface shapefiles were available. For these cities, we used latitude/longitude coordinates in both the shapefiles and Dun and Bradstreet to match building addresses to the closest street block and blockface segment. For Chicago and Los Angeles, blockface maps are not available. For those two cities, we first used latitude/longitude coordinates in Dun and Bradstreet and the street map shapefiles to match addresses to the closest street segment (city block). We then used the street address number from Dun and Bradstreet to code whether a given address was on the even or the odd side of a street and assigned the blockface location on the city block using that information.

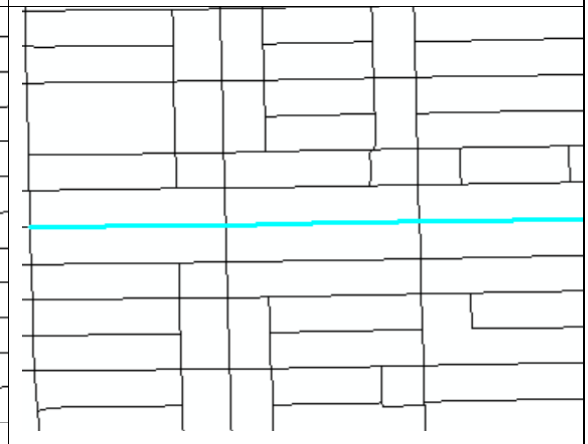
Once each Dun and Bradstreet address was coded to a street block and blockface, it was straightforward to compute other geographic features used in our analysis. This includes the level and composition of employment at different locations relative to a target building, including activity within 0.1 miles of the building, across the street, elsewhere on the blockface, and in the building itself.

Figure A-1: Example of street and blockface line segments coded in shapefiles

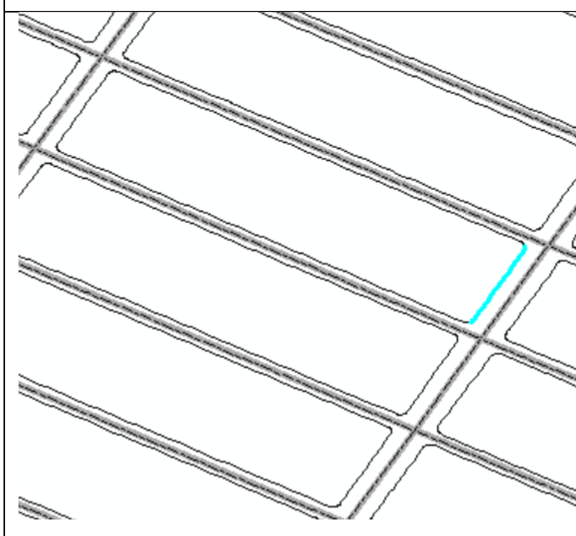
Panel A: Street segment



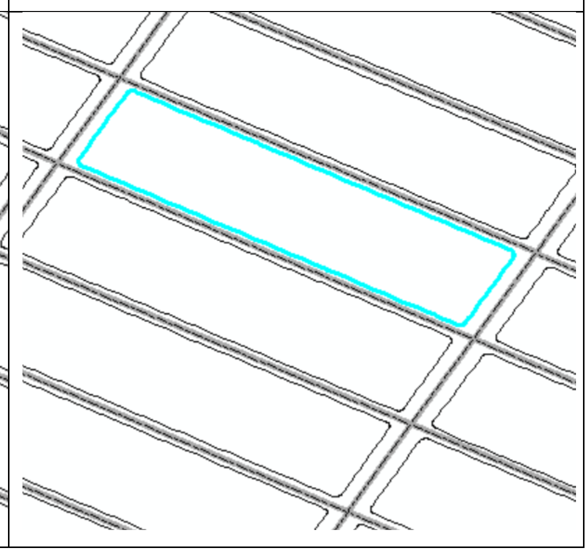
Panel B: All street segments for a given street



Panel C: Individual blockface segment



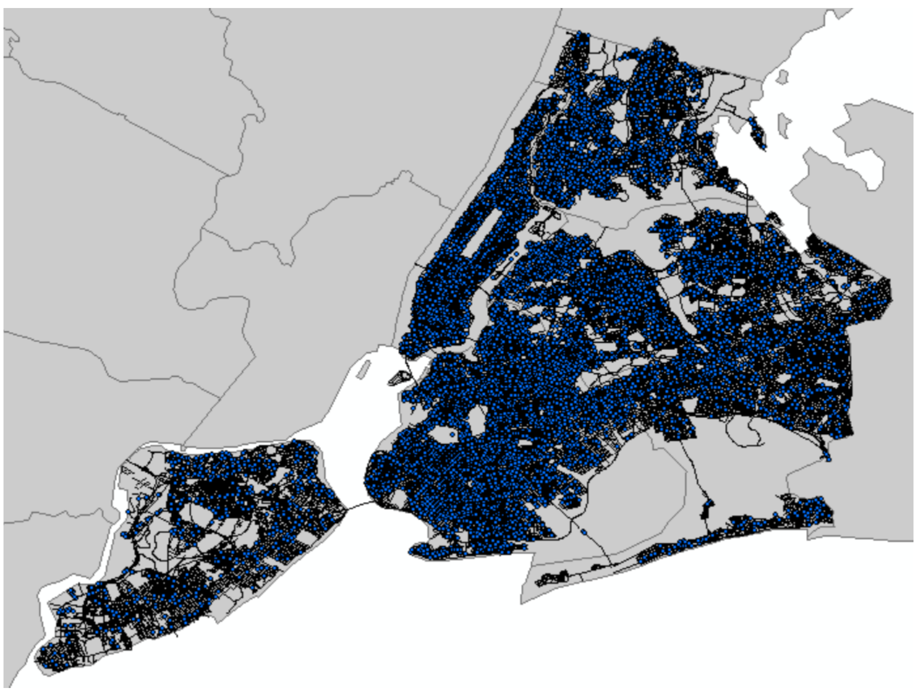
Panel D: All blockface segments in a block



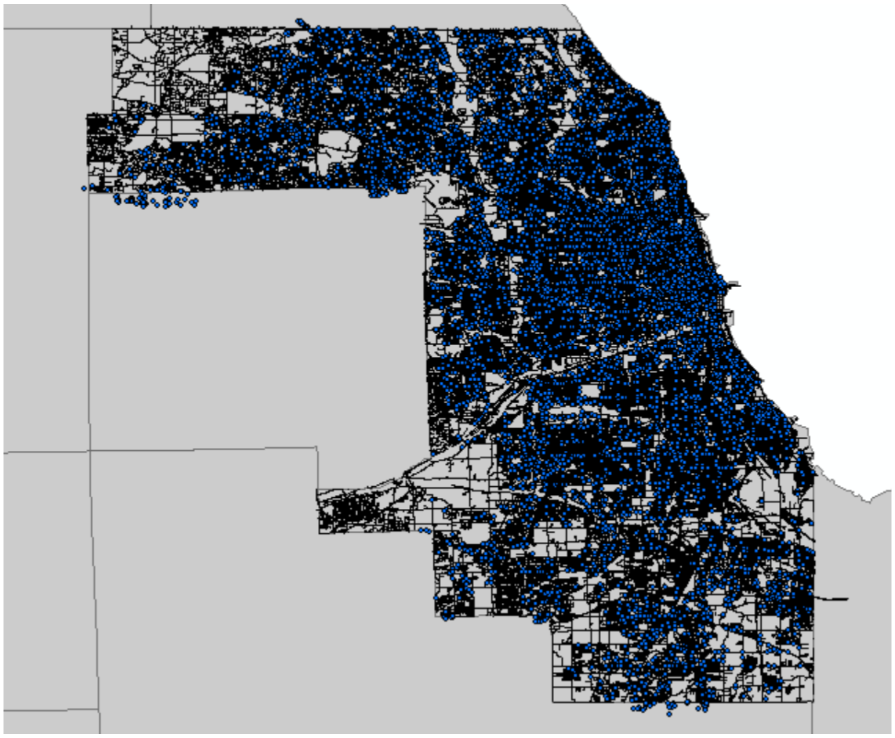
Appendix B: Geography of Sample Street Map Shape Files

The figures below highlight the geographic areas used in our analysis for New York City, Chicago (Cook County), Los Angeles, San Francisco and Washington DC. In each case, the figures below adhere to the following convention. Buildings are represented by blue dots. Streets are coded as black lines. Counties are grey-shaded areas.

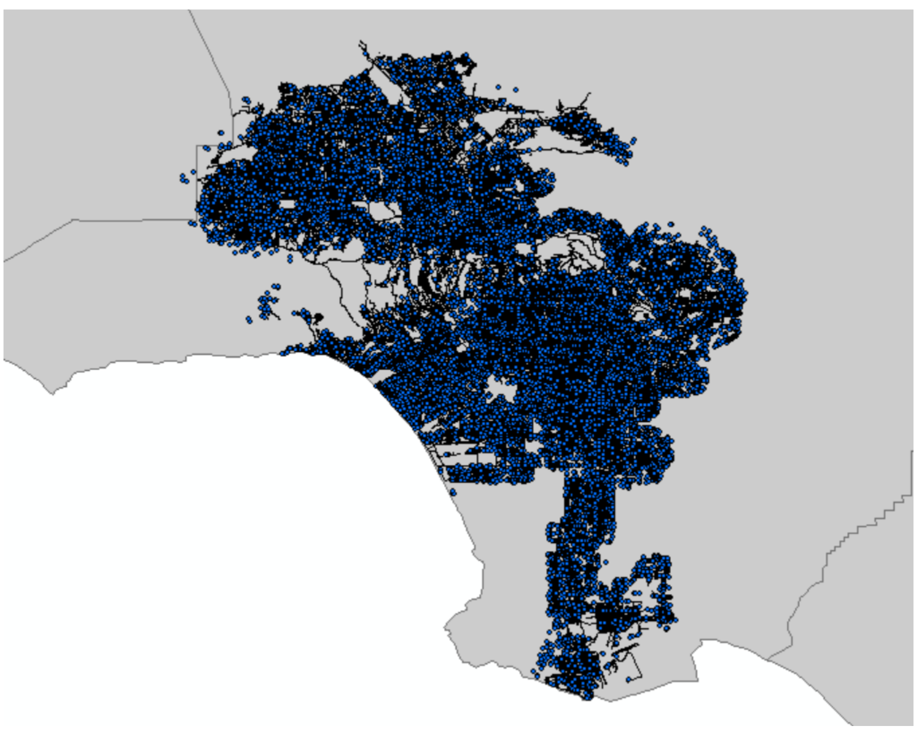
Panel A: New York City Building Sample



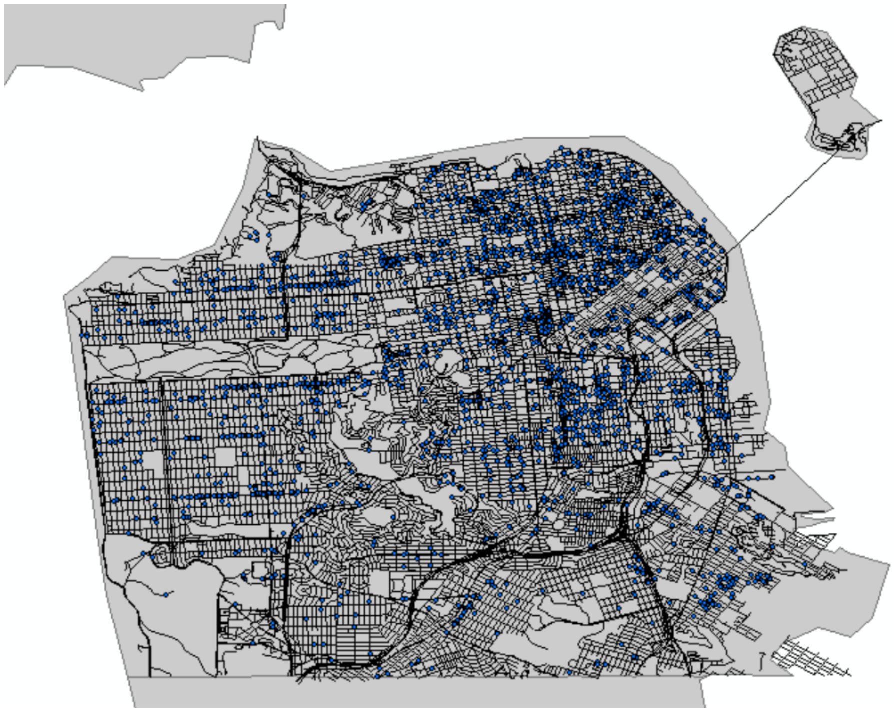
Panel B: Cook County Building Sample



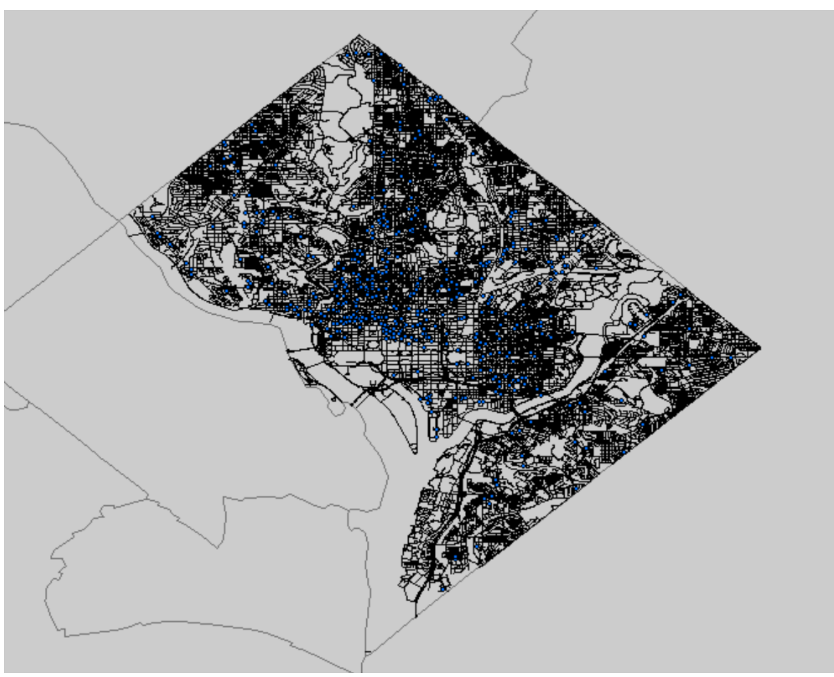
Panel C: Los Angeles Building Sample



Panel D: San Francisco Building Sample



Panel E: Washington DC Building Sample



Appendix C: Zipcode Fixed Effect Tobit Models

Table 5 Tobit Model Employment Share Within Target Building (not including own-industry anchor)[a]

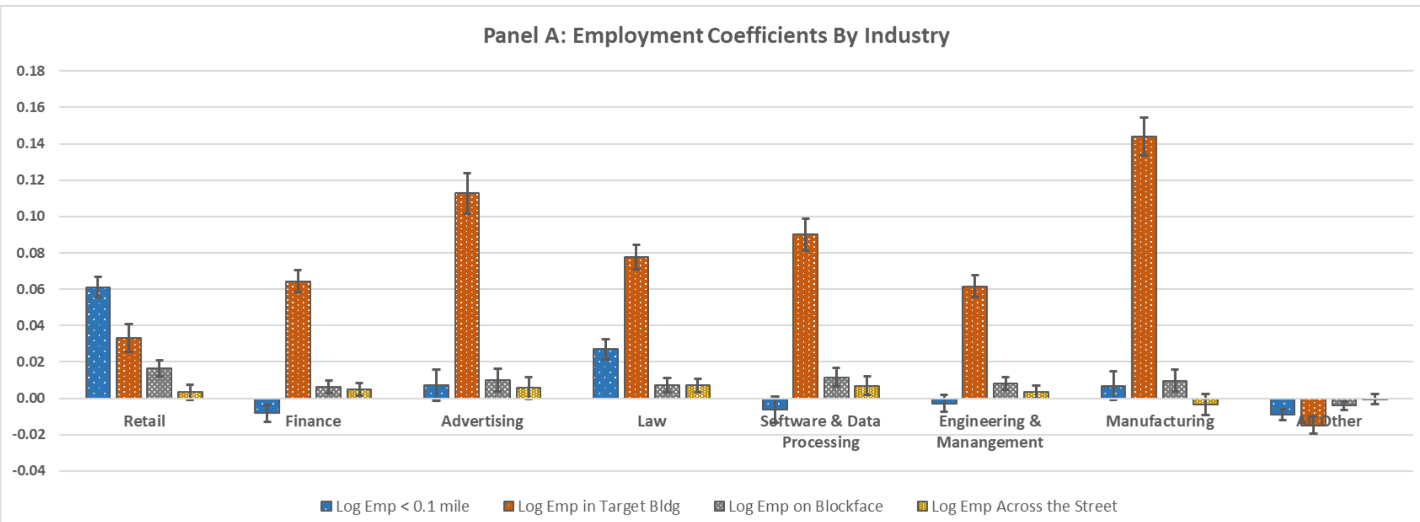
Group 1: Employment and Employment Share Coefficients

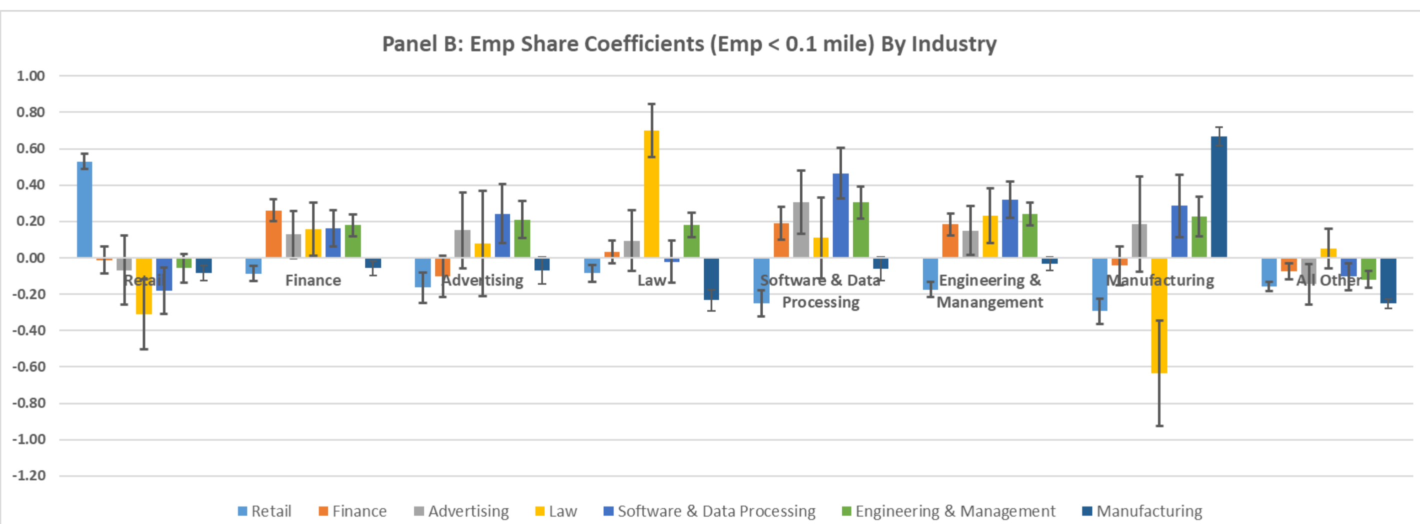
Regression Model Summary Statistics and Notes[a]

	Retail	Finance	Advertising	Law	Software & Data Proc	Eng & Mgmt	Manf	All Other
Zipcode FE	664	664	664	664	664	664	664	664
Observations	105,982	105,982	105,982	105,982	105,982	105,982	105,982	105,982
Censored	69,615	90,399	102,958	98,945	98,154	84,014	87,834	9,996
Uncensored	36,367	15,583	3,024	7,037	7,828	21,968	18,148	95,986
Pseudo R-sq	0.0703	0.135	0.277	0.230	0.185	0.104	0.136	0.0499

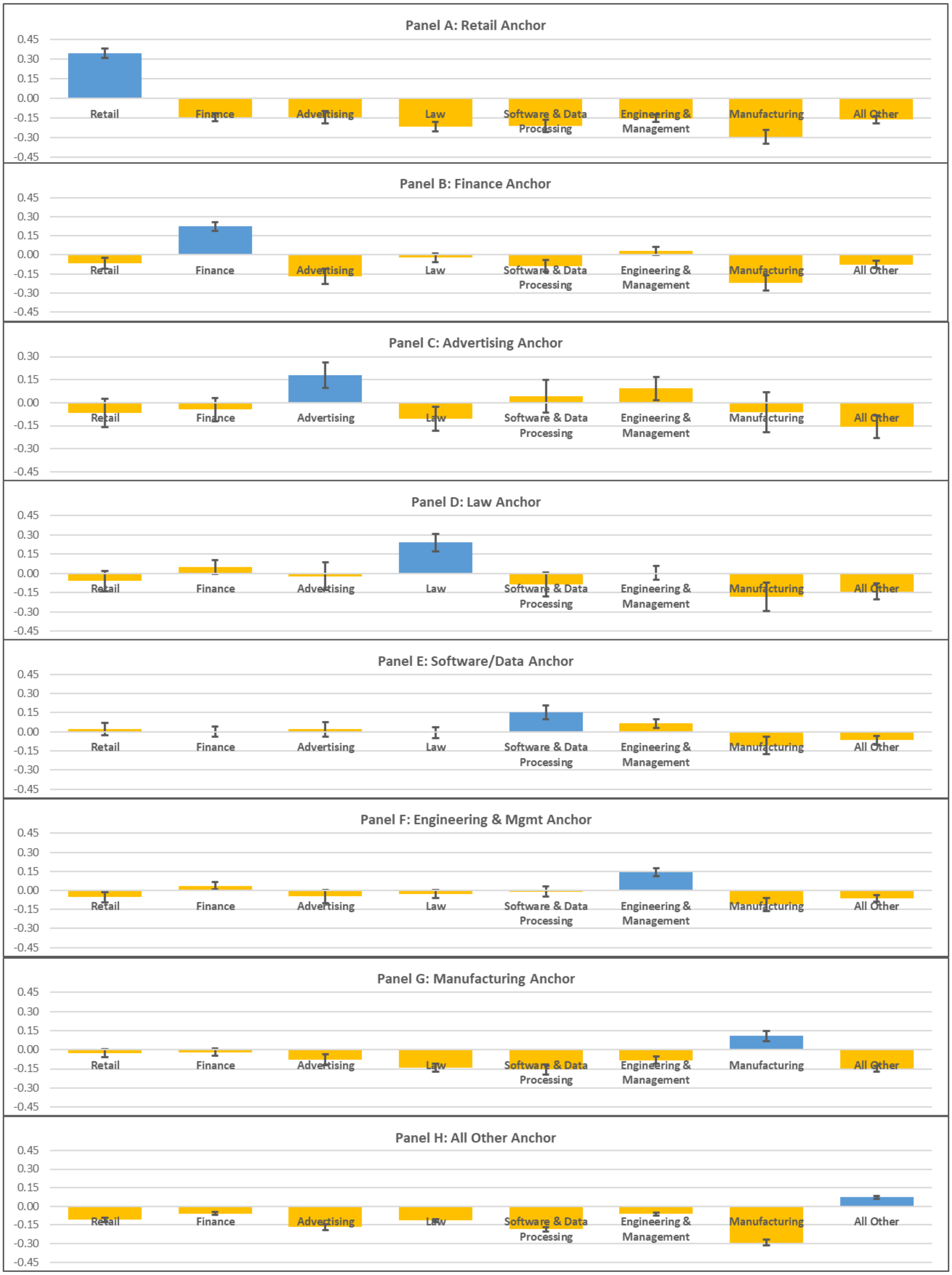
[a] Sample includes buildings with 10 or more workers in the following counties: New York (Manhattan), Cook (Chicago) (Chicago), Los Angeles, San Francisco, Washington DC. Confidence bands based on robust standard errors.

[b] An establishment is coded as an anchor (1 if yes, 0 if no) if all three of the following conditions hold: (i) it is the largest establishment in the building; (ii) it includes at least 20% of building employment; (iii) it has 10 or more workers.

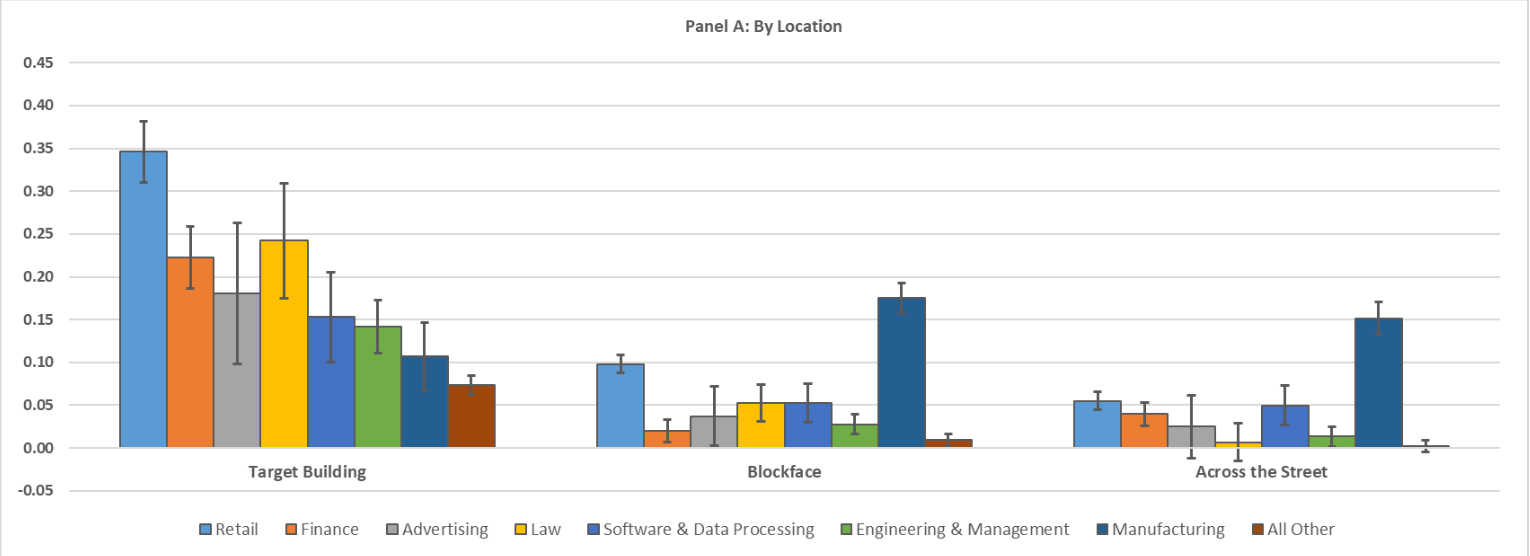


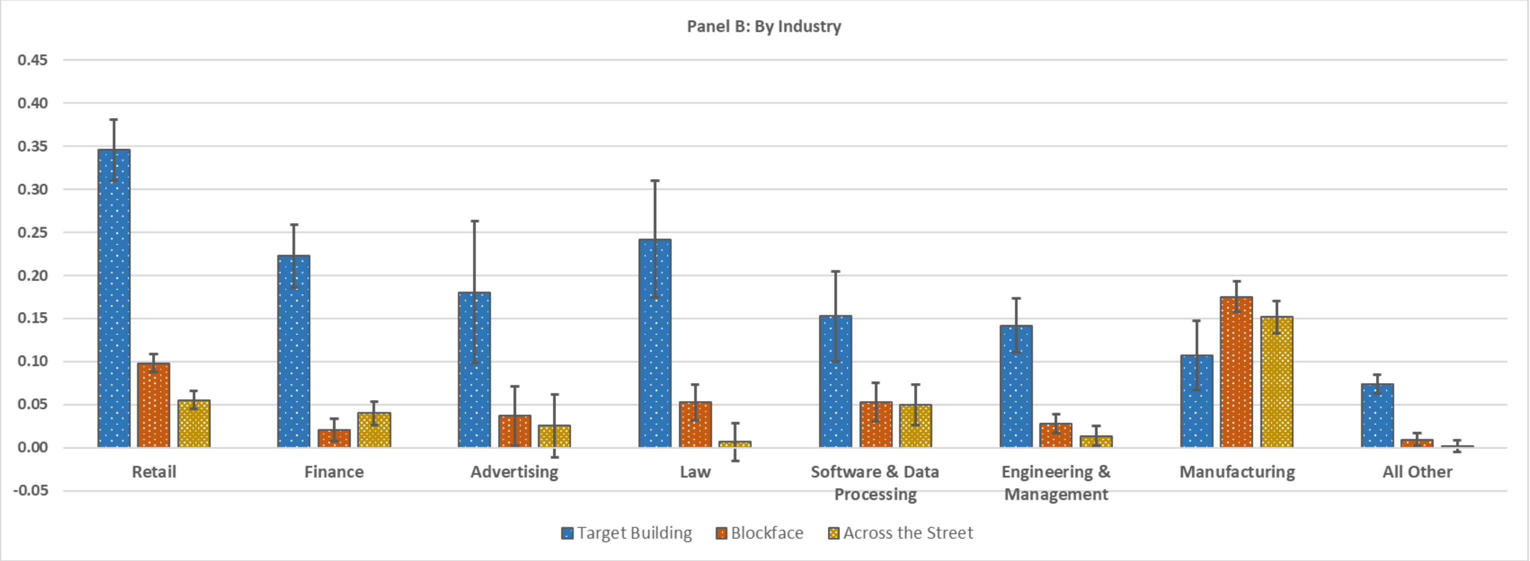


Group 2: Target Building Anchor Coefficients



Group 3: Own-Industry Anchor Coefficients Organized by Location and Industry





Appendix D: Alternate Definitions of Anchor Establishments

Two alternate versions of Table 5 are reported below, each using a different definition of an anchor. In both cases, and as in the text, anchors are only defined for buildings with 50 or more workers and must also be the largest establishment in their building. In the text, we further require that anchors account for 20% or more of employment in the building. That threshold is set to 15% in the first set of alternate tables below and 25% in the second set.

To distinguish the two sets of tables, the table labels below are augmented with 15% Anchor or 25% Anchor depending on which definition of an anchor establishment is adopted.

.

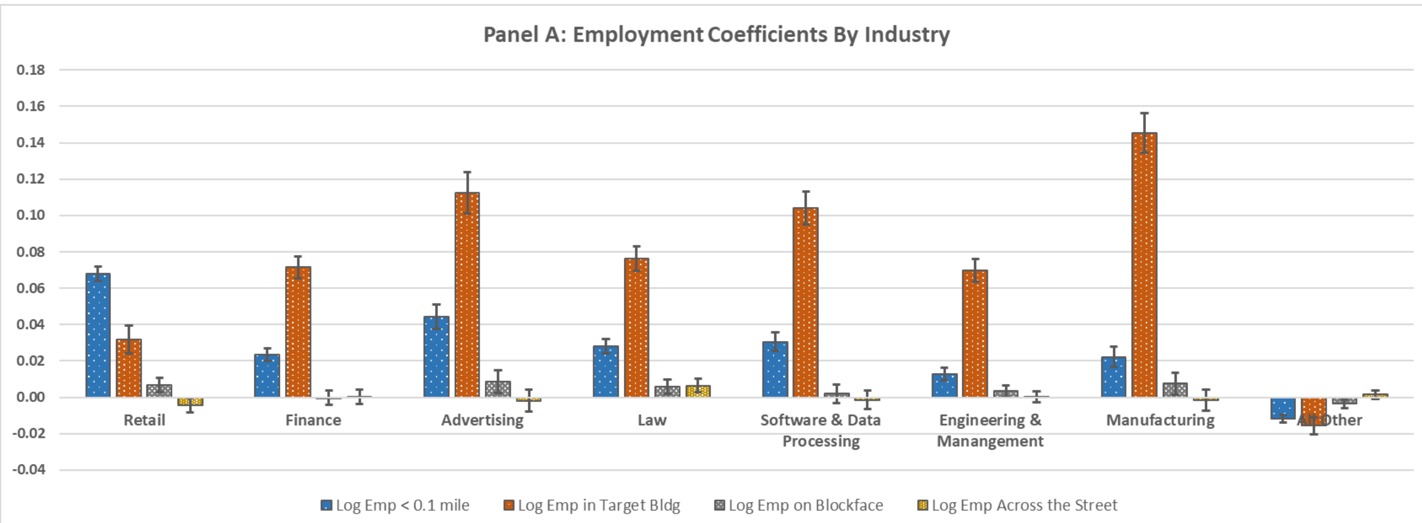
Table 5 - 15% Anchor: Tobit Model Employment Share Within Target Building (not including own-industry anchor)[a]

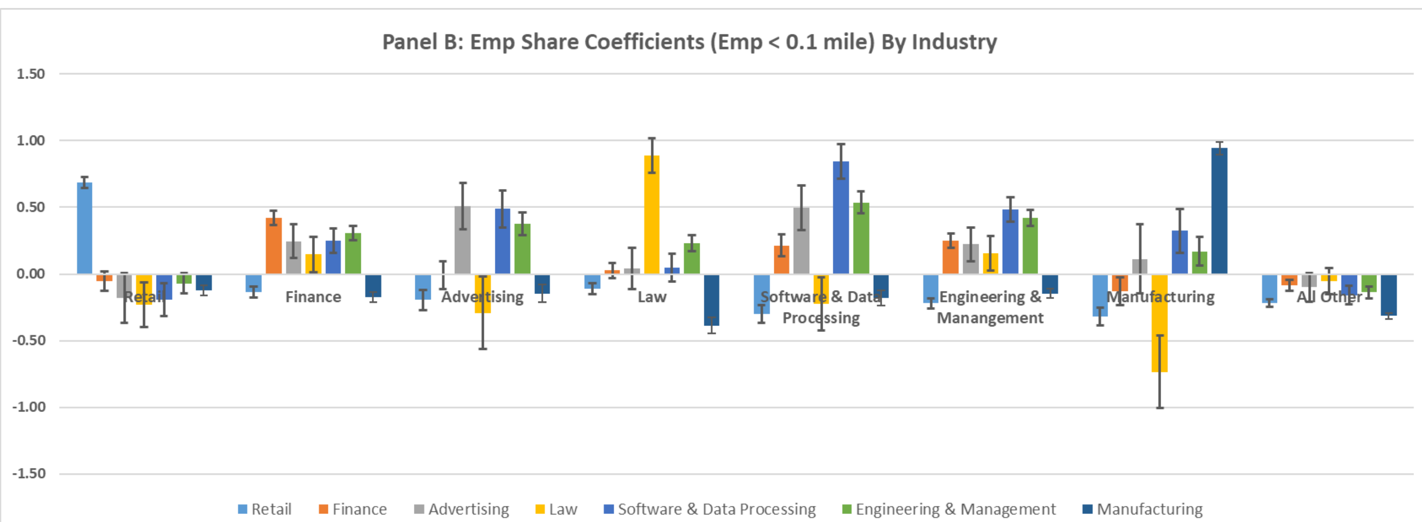
Regression Model Summary Statistics and Notes[a]

	Retail	Finance	Advertising	Law	Software & Data Proc	Eng & Mgmt	Manf	All Other
Anchor Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	5	5	5	5	5	5	5	5
Observations	105,982	105,982	105,982	105,982	105,982	105,982	105,982	105,982
Censored	69,625	90,406	102,969	98,945	98,168	84,021	87,846	9,996
Uncensored	36,357	15,576	3,013	7,037	7,814	21,961	18,136	95,986
Pseudo R-sq	0.0562	0.103	0.221	0.185	0.137	0.0784	0.111	0.0352

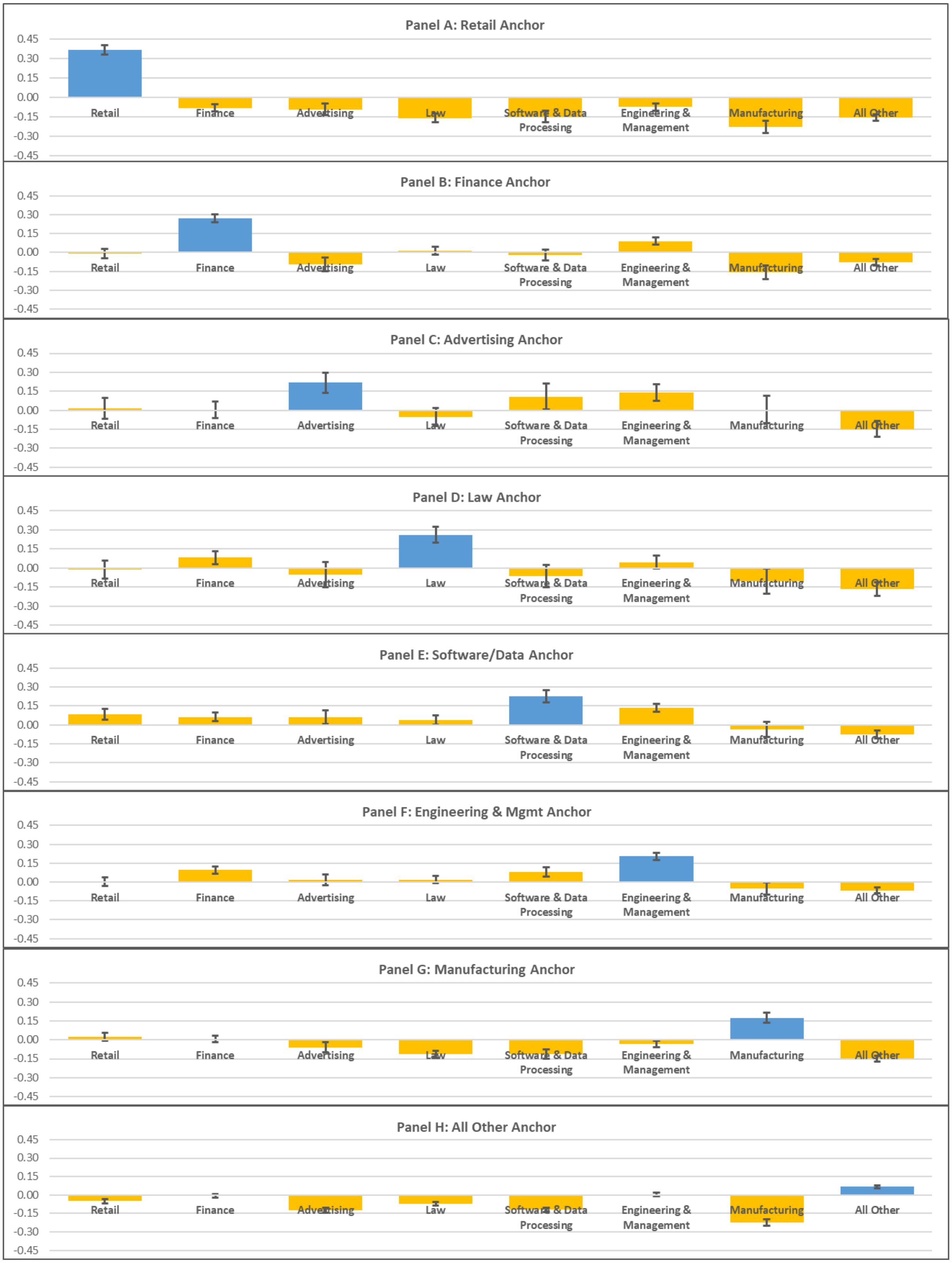
[a] Sample includes buildings with 10 or more workers in the following counties: New York (Manhattan), Cook (Chicago) (Chicago), Los Angeles, San Francisco, Washington DC. Confidence bands based on robust standard errors.
[b] An establishment is coded as an anchor (1 if yes, 0 if no) if all three of the following conditions hold: (i) it is the largest establishment in the building; (ii) it includes at least 15% of building employment; (iii) it has 10 or more workers.

Group 1: Employment and Employment Share Coefficients

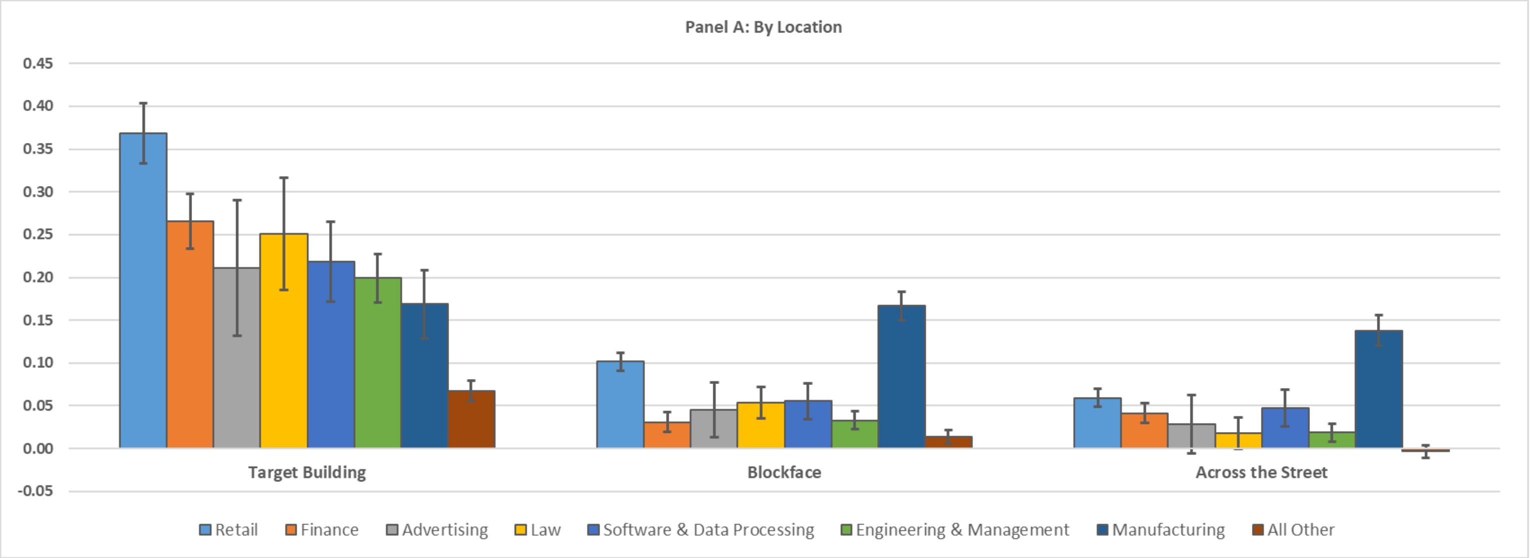




Group 2: Target Building Anchor Coefficients



Group 3: Own-Industry Anchor Coefficients Organized by Location and Industry



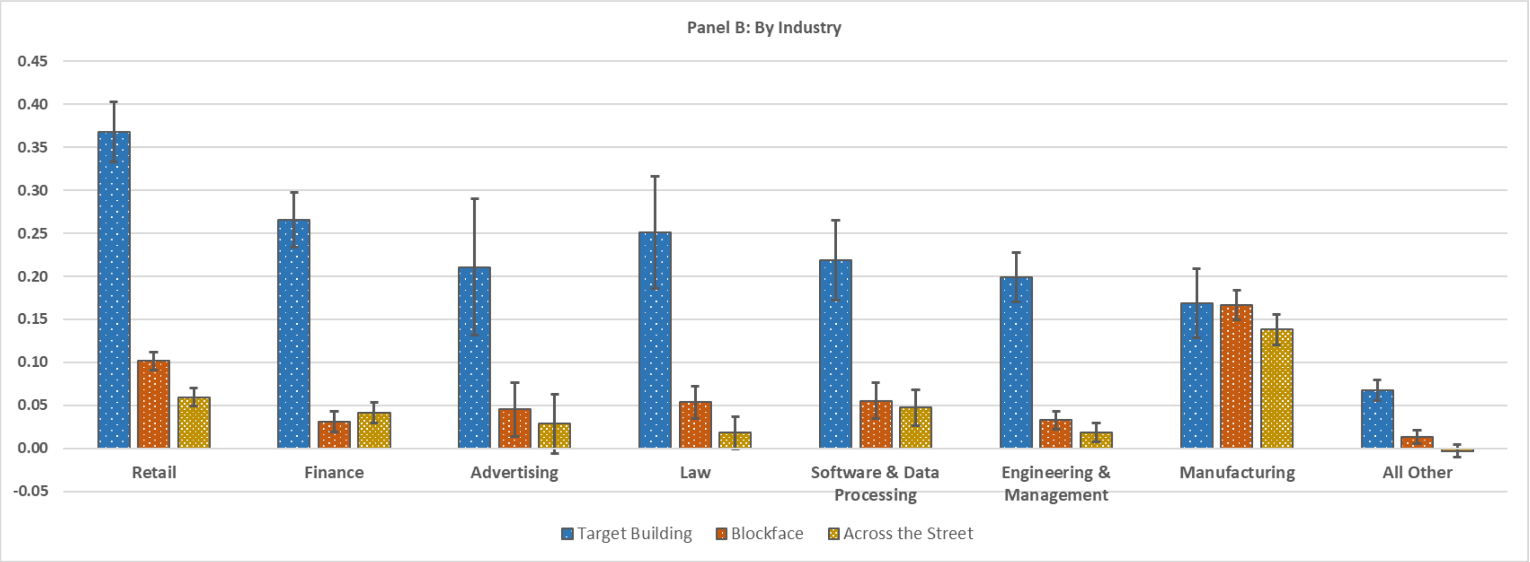


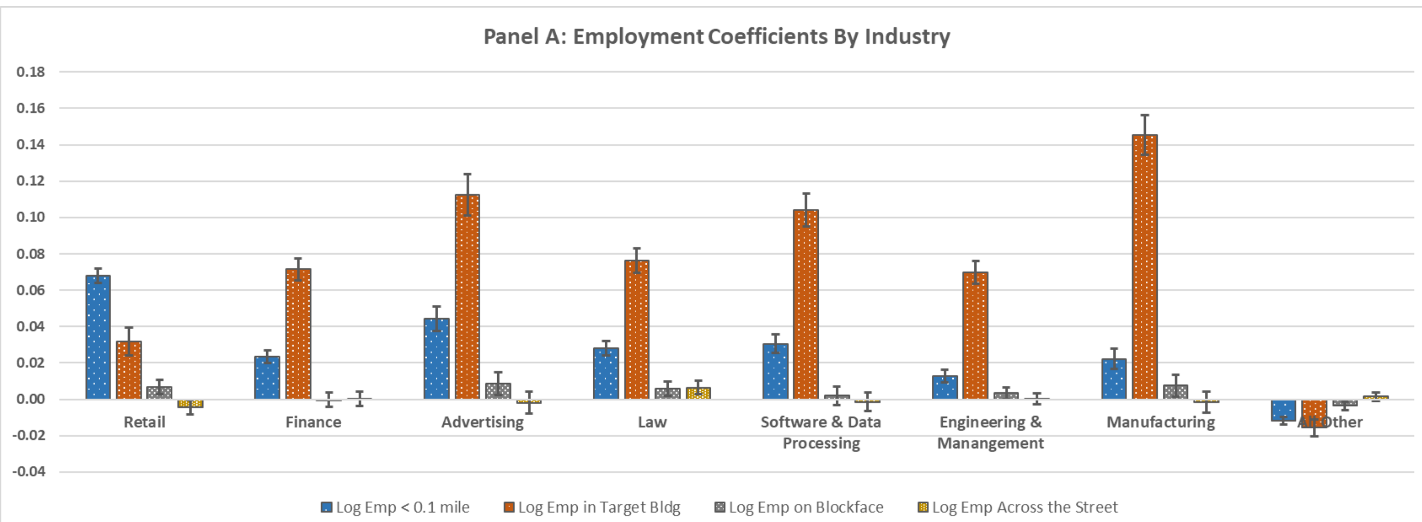
Table 5 - 25% Anchor: Tobit Model Employment Share Within Target Building (not including own-industry anchor)[a]

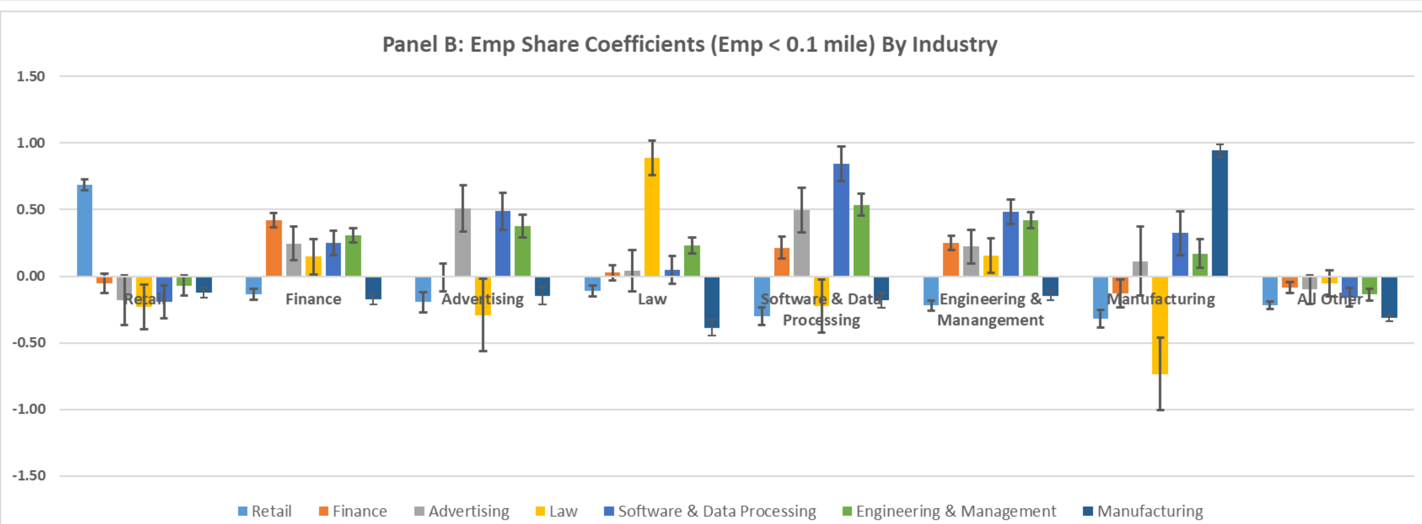
Regression Model Summary Statistics and Notes[a]

	Retail	Finance	Advertising	Law	Software & Data Proc	Eng & Mgmt	Manf	All Other
Anchor Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	5	5	5	5	5	5	5	5
Observations	105,982	105,982	105,982	105,982	105,982	105,982	105,982	105,982
Censored	69625	90406	102969	98945	98168	84021	87846	9996
Uncensored	36357	15576	3013	7037	7814	21961	18136	95986
Pseudo R-sq	0.0562	0.103	0.221	0.185	0.137	0.0784	0.111	0.0352

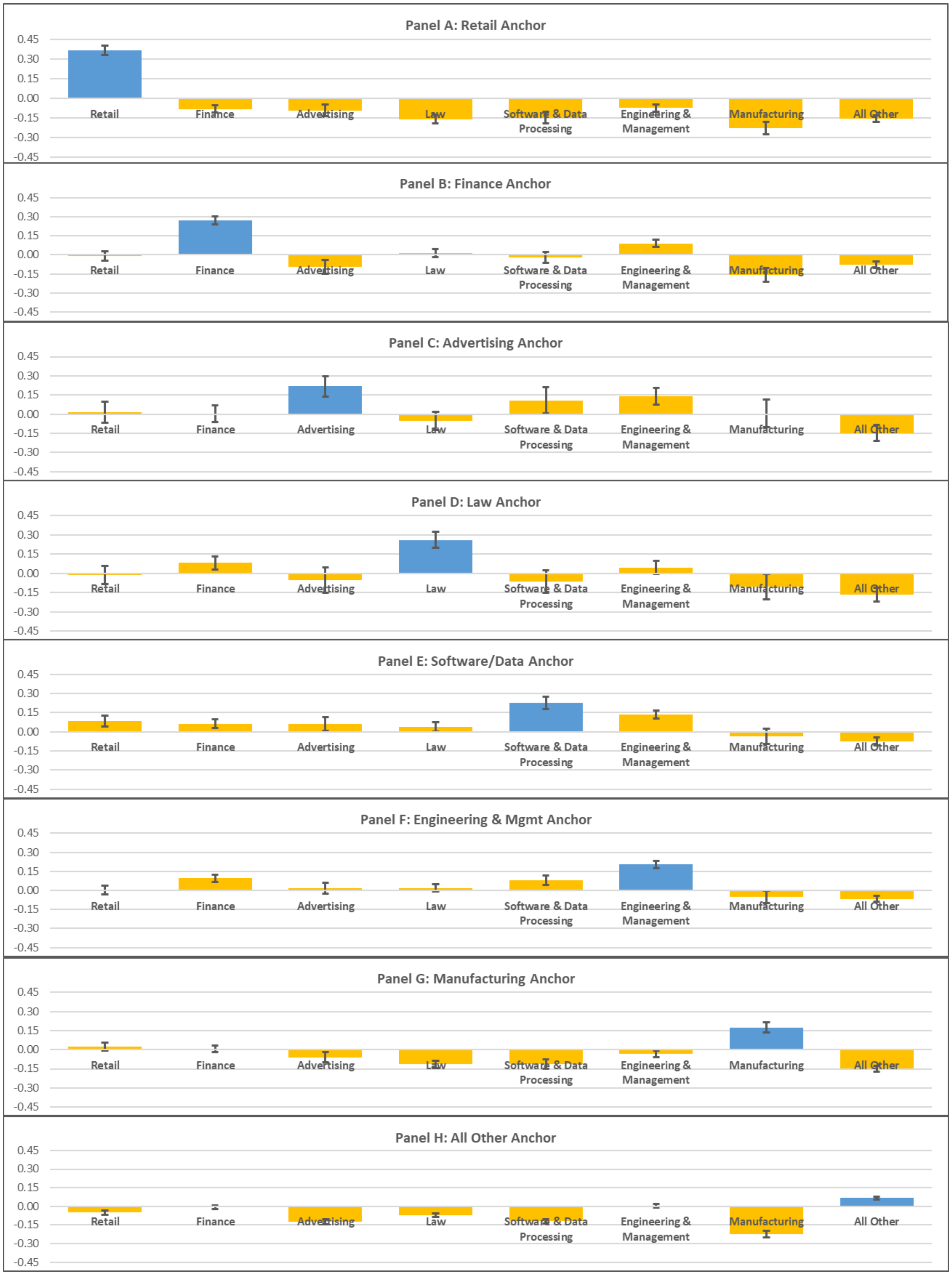
[a] Sample includes buildings with 10 or more workers in the following counties: New York (Manhattan), Cook (Chicago) (Chicago), Los Angeles, San Francisco, Washington DC. Confidence bands based on robust standard errors.
[b] An establishment is coded as an anchor (1 if yes, 0 if no) if all three of the following conditions hold: (i) it is the largest establishment in the building; (ii) it includes at least 25% of building employment; (iii) it has 10 or more workers.

Group 1: Employment and Employment Share Coefficients





Group 2: Target Building Anchor Coefficients



Group 3: Own-Industry Anchor Coefficients Organized by Location and Industry