Eyes on the Street, Spatial Concentration of Retail Activity and Crime

Stuart S. Rosenthal
Maxwell Advisory Board Professor of Economics
Department of Economics and Center for Policy Research
Syracuse University Syracuse, New York, 13244-1020
Email: ssrosent@syr.edu

Joaquin A. Urrego
Department of Economics and Center for Policy Research
Syracuse University Syracuse, New York, 13244-1020
Email: jaurrego@syr.edu

December 29, 2021

* We thank Giulia Faggio, William Horrace, Amy Schwartz, and participants in the Urban Economic Association fall, 2021 meeting for helpful comments. We also thank Safegraph for access to cellphone data including Mike Sussman, Mitch McGillis, and Jonathan Wolf. All errors are our own.

Abstract

We show that if spatial concentration of retail activity amplifies the effect of “eyes on the street”, this should lower neighborhood crime rates and reduce investment in anti-crime measures, with benefits capitalized into higher retail rent. Point-specific data for New York City supports the model predictions while comparisons between nighttime versus daytime activity, pre-pandemic versus COVID-19 lockdown, and different measures of spatial concentration shed light on mechanisms. Increasing block-level retail spatial concentration by one standard deviation reduces property crime and police stops by 8.5% and 11%, respectively, and causes retail rent to increase by at least 7.8%.

JEL Codes: R00, R33, K00.
Key Words: Eyes-on-the-street, Retail Spatial Concentration, Crime, Police, Rent Capitalization

1. Introduction

Longstanding arguments suggest that "eyes on the street" deter crime by increasing the likelihood that criminals will be caught (Jacobs, 1961; Browning and Jackson, 2013; Chang and Jacobson, 2017; Carr and Doleac, 2018; McMillen et al., 2019). Drawing on this idea, we argue that concentrating retail establishments at the street level should amplify the effect of eyes on the street by crowding shoppers into smaller areas and by making it possible to observe multiple store fronts at once (referred to below as crowding and visibility effects, respectively). Related arguments suggest that this should reduce public and private investment in anti-crime measures while lowering local crime rates, with benefits capitalized into higher retail rent.[1] These and other predictions are confirmed using point-specific data for New York City.[2]

The potential for crime deterrent effects of spatial concentration of retail activity to yield substantial savings is large. A 2018 National Retail Federation (NRF) survey of U.S. retailers found that respondents lost an average of 1.38% of sales to theft of merchandise and services, similar to a 1.3% loss rate reported for retailers in Europe in the 2018-2020 World Bank Enterprise survey.[3] These same surveys indicate that U.S. retailers allocated 0.74% of sales to private security measures while European retailers spent roughly 0.8% of sales. Benchmarked against 2018 retail sales in NYC (roughly $100 billion), the NRF estimates suggest that NYC retailers lost roughly $1.38 billion to theft in 2018 while also spending $740 million on security.[4] Adding to these costs, in 2018 the New York City Police Department (NYPD) directed $1.60 billion of its budget for police patrols (New York City Council, 2018), a portion of which

---

[1]Property crime refers to petit larceny, grand larceny, burglary, theft of services and fraud, all of which are classified as non-violent. Auto theft and robbery are classified separately, where the latter is considered a violent crime.

[2] For related work on commercial activity and crime, see Hakim and Shachmurove (1996), Greenbaun and Tita (2004), Bowes (2007), Stucky and Ottensmann (2009), Browning et al. (2010), Rosenthal and Ross (2010), Lee and Alshalan (2012), Weterings (2014), Groff and Lockwood (2014), Hipp (2016), and Tillyer and Walter (2019).

[3] The 2018 NRF survey is at https://cdn.nrf.com/sites/default/files/2018-10/NRF-NRSS-Industry-Research-Survey-2018.pdf. Details of the 2018-2020 World Bank Enterprise survey are at https://www.enterprisesurveys.org. We based our 1.3% measure above on pooled loses from Austria, Belgium, Croatia, Czech Republic, Denmark, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Slovenia, and Sweden.

[4] The U.S. Census reports that NYC sales in 2012 were $92.265 billion which is roughly $100,000 billion in 2018 dollars (https://www.census.gov/quickfacts/fact/table/newyorkcitynewyork/RTN130212).

would have been intended to protect against property crime.[5] In comparison, estimates later in the paper suggest that a one standard deviation increase in block-level concentration of retail activity would reduce property crime by 8.5%. That represents a total savings among NYC retailers of roughly $117 million.

To frame our analysis, we develop a conceptual model based on the idea that criminals trade off potential return and costs.[6] We also draw on extensive literature in which eyes-on-the-street is treated as a local public good (Jacobs, 1961; Browning and Jackson, 2013; Chang and Jacobson, 2017; McMillen et al., 2019). Central to our arguments is the idea that spatial concentration of retail activity increases the productivity of police patrols and private sector investment in protection measures by making it easier to see and apprehend criminals.

Our model assumes that retail concentration amplifies the effect of public and private protection against crime. Comparative statics then indicate that spatial concentration of retail outlets should reduce equilibrium levels of crime while lowering investment in public and private protection. In the empirical work that follows, we use multiple strategies to help ensure that these relationships are identified. This includes using geographically granular data, extensive controls, and various forms of differencing, details of which are described shortly. A threat to identification is that even after conditioning on these and other features, retailers may concentrate in high crime areas so as to gain better protection. Should that occur, our estimates will understate the crime deterrent effect of spatial concentration.

The second part of our model focuses on rent capitalization. Starting from the retailer's profit function, we show with competitive markets, equilibrium factor input costs per dollar of sales (based on space rented and labor) are equal to industry markup over wholesale cost, where markup allows for the cost of inventory lost to crime. Neighborhood spatial concentration reduces inventory lost to crime, savings from which are capitalized into higher local rent. A possible threat to identification in this part of the model is that spatial concentration also creates shopping externalities that likely increase sales for a

[5] Details of the NYPD 2018 budget are at http://council.nyc.gov/budget/wp-content/uploads/sites/54/2017/03/056-NYPD-exec-1.pdf .
[6] This is in the spirit of Becker (1968). See Freedman and Owens (2016) for recent related empirical evidence.

given amount of space and labor, profits from which will also be capitalized into higher rent. However, because crime deterrence affects cost while shopping externalities affect sales, here too we are able to identify a reliable lower bound estimate of the rent capitalization effect of crime deterrence.[7]

We use geocoded point-specific data from New York City (NYC) to test the model predictions. In all cases, we divide NYC into 0.2 by 0.2 mile grid cells (approximately a 4 to 7 minute walk), each of which is treated as a separate neighborhood. For the crime and police patrol models, activity is analyzed at the neighborhood level. For the capitalization models, rent is analyzed at the establishment level using the location of each establishment (within roughly 3 feet). In both cases, most controls are at the neighborhood level having aggregated up from point-specific data. The small size of our neighborhood units helps to reduce potential for unobserved factors. The small neighborhood design is also consistent with evidence that the effect of crime in urban areas is highly localized.[8]

The New York Police Department (NYPD) provides the coordinates of all reported crimes and police stops since 2006. Police stops refer to pedestrian stops made by the NYPD under the Stop-Question-Frisk policy (SQF).[9] Extensive neighborhood attributes are then matched to these data. This includes information on neighborhood buildings from the NYC Department of City Planning. Also included are numerous measures obtained from the NYC Fire Department, Department of IT and Telecommunication, Department of Public Health, and the Department of Parks and Recreation. Cellphone data on foot traffic are obtained from SafeGraph, while extensive information on neighborhood

---

[7] For previous work on the effect of shopping externalities on sales, see Gould et al. (2005), Pashigian and Gould (1998), Koster et al. (2014), Johansen and Nilssen (2016), Clapp et al. (2019) and Koster et al. (2019). For evidence that crime and other local attributes affect commercial property values and/or rent, see Sivitanidou (1995), Lens and Meltzer (2016), and Rosenthal, Strange and Urrego (2021).

[8] In related work, Ellen et al (2013) report that crime in New York City increases on city blocks where a mortgage default has recently occurred, likely because of deleterious effects from undermaintained and/or vacant properties. Linden and Rockoff (2008) find that the presence of a registered sex offender in Mecklenburg County, North Carolina has a negative effect on residential property values within 0.1 miles. Pope (2008) obtains similar results for Hillsborough County, Florida. In all three studies, estimated effects attenuate rapidly with distance.

[9] The SQF policy was widely criticized up to roughly 2012 as contributing to discriminatory police behavior against minorities, prompting a sharp shift in policy implementation. We use police stop data from 2016-2018 to mitigate concerns about these issues. Details are provided later in the paper.

establishments is obtained from Dun & Bradstreet. For the rent models we match establishment-level commercial leases from CompStak Inc. with establishment-level information from Dun and Bradstreet.

Results indicate that increasing neighborhood spatial concentration of retail activity from the 25th percentile to the 75th percentile neighborhood reduces property crime by 9.4% and police stops by 12.1%. These findings are obtained from negative binomial regressions for counts of property crime and police stops.[10] Included in these models is an extensive set of neighborhood and building specific controls. Most important, this includes the level and composition of employment in the neighborhood as we expect property crime to increase with the level of nearby activity and retail inventory. We also control for spatial concentration of non-retail industries within the defined neighborhoods. Cell phone data is used to further control for the overall level of foot traffic in the neighborhood as this allows for both business and non-business related pedestrian activity. Other controls include the presence of trees (as a proxy for amenities), share of residential units, building age and assessed value, and many other attributes.

Additional sample designs help to shed light on crowding and visibility as underlying mechanisms. We compare crime rates at night to those during the day and also crime rates throughout 2018 to the first two weeks of the NYC COVID-19 lockdown (March 22nd – April 5th, 2020).[11] Crowds are greatly diminished at night and during the lockdown for reasons unrelated to crime, and those shifts should weaken the effect of eyes on the street. For these comparisons we split property crime into different sub-categories (petit larceny, grand larceny plus burglary, theft of services plus fraud) and also consider effects on robbery and auto theft. Evidence suggests that both crowding and enhanced visibility help to explain why spatial concentration of retail activity deters crime.

In a separate set of models, we measure spatial concentration in three ways, based on spatial patterns of employment, the location of store fronts, and sales. We argue that the first measure is

[10] Negative binomial count models help to address overdispersion in our dependent variables that arise from zero values in many grid squares. Similar results were obtained using OLS.

[11] Later in April 2020 up to 20% of the NYPD police force was out sick with COVID-19. This would have reduced the ability of police to patrol (https://www.cnn.com/2020/04/07/us/nypd-coronavirus-out-sick/index.html). Focusing on the first weeks of the lockdown avoids this issue.

especially effective at capturing crowding effects, the second targets visibility, and the third measure is more of a placebo check having conditioned on the first two. Once again, results suggest that crowding and visibility both enhance crime deterrent effects from retail spatial concentration.

Estimates from our rent models also confirm that crime deterrence is capitalized into higher local rent. For the average neighborhood, a one standard deviation increase in retail spatial concentration is associated with a 7.8% increase in expenditures on space and labor per dollar of sales. A corresponding estimate for wholesale establishments is smaller and serves as a further robustness check. The absence of shoppers from warehouse facilities reduces the threat of shoplifting and also allows for aggressive protection measures that would discourage retail shoppers. Both effects should reduce the crime deterrent effect of retail spatial concentration on wholesale establishment rent, which is what we find.

To establish these and related results, the next section presents our conceptual model. Section 3 describes the data and summary statistics. Section 4 presents the results, and Section 5 concludes.

2. Theoretical Framework

This section develops a conceptual model that motivates the empirical work to follow. The first part considers the effect of spatial concentration of retail activity on equilibrium patterns of property crime and police protection. The second part examines rent capitalization.

2.1 Equilibrium protection and the level of crime

We begin by assuming that all stores in a neighborhood j have identical valued inventory, I_j, some of which is stolen while the rest is sold,

$$I_j = I^{stolen}(P_j) + I^{sold}(P_j) . \tag{2.1}$$

In (2.1), stolen inventory declines with the quality of protection services enjoyed by each store in the community, P_j, where P_j is generated based on a Cobb-Douglas production technology,

$$P_j = G_j Pu_j^{\alpha} Pr_j^{\beta} \quad \text{with } \alpha + \beta \leq 1 . \tag{2.2}$$

In (2.2), Pu_j and Pr_j denote public and private expenditures on security, respectively, where Pu_j is treated as a non-congestible local public good. Pu_j includes neighborhood police patrols, something we examine in the empirical work to follow, while Pr_j includes privately funded measures specific to each store, as with security alarms, door locks, surveillance cameras, etc. Also included as an argument in (2.2) is the spatial concentration of retail activity in neighborhood j, denoted by G_j. This enters as a Hicks neutral shift factor that amplifies the productivity of public and private investment in security. Reasons for why spatial concentration may enhance the productivity of anticrime measures are as described in the Introduction and are motivated by previous literature on "eyes on the street" (e.g. Chang and Jacobson, 2017; McMillen et al., 2019; Gonzales and Komisarow, 2020).[12]

To solve for the efficient mix of public and private investment in protection, suppose next that there are N_j retail outlets in neighborhood j, each of which pays an equal tax share to finance the cost of public protection. Establishments incur the following expenditures to support private and public security,

$$exp_j = Pr_j + \frac{1}{N_j} Pu_j \ , \tag{2.3}$$

where $1/N_j$ is the price of public protection, and the price of private protection is normalized to 1.

In the simplest setting, local government acts as a social planner and chooses Pu_j and Pr_j to minimize protection costs for each store while providing an efficient level of protection, P_j^*.[13] From (2.2) and (2.3), the Lagrangian is given by,

$$\mathcal{L} = Pr_j + \frac{1}{N_j} Pu_j + \eta \left(P_j - G_j Pu_j^{\alpha} Pr_j^{\beta} \right) . \tag{2.4}$$

Taking first-order conditions and rearranging, the efficient levels of public and private investment for a given level of P_j are,

---

[12] Spatial concentration could also facilitate collaboration with neighbors on protection measures as seems likely to occur in business improvement districts, BIDs (see, for example, Hoyt, 2005; Brooks, 2008 and Faggio, 2021). We consider this point in a robustness check later in the paper. BID presence does not affect our core results.

[13] Our results remain the same if local government seeks to maximize private sector profit by choosing the optimal level of Pu_j for a given Pr_j, and similarly, that the private sector chooses an optimal Pr_j for a given Pu_j. Expressions (2.5a) and (2.5b) then describe the corresponding reaction functions having substituted in P_j from (2.2). This yields the same combination of Pu_j and Pr_j as in (2.6) and also P_j^* as in (2.9) below.

$$Pu_j^* = \left[\frac{P_j}{G_j}\right]^{\frac{1}{\alpha+\beta}} N_j^{\frac{\beta}{\alpha+\beta}} \left(\frac{\alpha}{\beta}\right)^{\frac{\beta}{\alpha+\beta}} \tag{2.5a}$$

$$Pr_j^* = \left[\frac{P_j}{G_j}\right]^{\frac{1}{\alpha+\beta}} N_j^{-\frac{\alpha}{\alpha+\beta}} \left(\frac{\alpha}{\beta}\right)^{-\frac{\alpha}{\alpha+\beta}} \tag{2.5b}$$

Dividing, for each store the efficient ratio of public to private investment in protection is,[14]

$$\frac{Pu_j}{Pr_j} = N_j \frac{\alpha}{\beta} \quad . \tag{2.6}$$

while the cost of providing P_j protection to each store, $C_{P,j}$, is obtained by substituting (2.5a) and (2.5b) into (2.3). Grouping the production parameters into a positive constant κ,[15] and solving,

$$C_{P,j} = \kappa \left[\frac{P_j}{G_j}\right]^{\frac{1}{\alpha+\beta}} \left[\frac{1}{N_j}\right]^{\frac{\alpha}{\alpha+\beta}} \quad . \tag{2.7}$$

To determine the efficient level of protection, we next characterize the deterrent effect of protection and the related amount of inventory that is stolen from each store. This is expressed in (2.8) as,

$$I_j^{sold} = I_j(1 - P_j^{-\lambda}), \text{ where } \lambda > 0 \text{ and } P_j > 1 \; . \tag{2.8}$$

In this expression, deterrence increases with λ as criminal skill at evading protection measures diminishes, or equivalently, as penalties that discourage criminal activity become increasingly severe.

Subtracting (2.7) from (2.8) yields a measure of the net gain to each store from its investment in private and public protection against property crime. Taking first order conditions and rearranging, each store in the community receives a level of protection given by,[16]

$$P_j^* = \lambda^{\frac{\alpha+\beta}{1+\lambda(\alpha+\beta)}} I_j^{\frac{\alpha+\beta}{1+\lambda(\alpha+\beta)}} G_j^{\frac{1}{1+\lambda(\alpha+\beta)}} N_j^{\frac{\alpha}{1+\lambda(\alpha+\beta)}} \left(\frac{\alpha+\beta}{\kappa}\right)^{\frac{\alpha+\beta}{1+\lambda(\alpha+\beta)}} \; . \tag{2.9}$$

Inventory lost to crime is then obtained by substituting (2.9) into (2.8) and using the identity in (2.1),

[14] In (2.6), the equilibrium ratio of public to private protection does not depend on G_j because G_j is specified as a Hicks neutral shifter above. N increases the equilibrium ratio of *Pu* to *Pr* because it lowers the price of public protection relative to private protection measures.

[15] $\kappa = \left(\frac{\alpha}{\beta}\right)^{-\frac{\alpha}{\alpha+\beta}} + \left(\frac{\alpha}{\beta}\right)^{\frac{\beta}{\alpha+\beta}}$

[16] The first order condition requires that $\lambda I_j P_j^{-(\lambda+1)} - \frac{\kappa}{\alpha+\beta} G_j^{-\frac{1}{\alpha+\beta}} N_j^{-\frac{\alpha}{\alpha+\beta}} P_j^{\frac{1}{\alpha+\beta}-1} = 0$.

$$I_j^{stolen} = \left(\frac{1}{\lambda}\right)^{\frac{\lambda(\alpha+\beta)}{1+\lambda(\alpha+\beta)}} I_j^{\frac{1}{1+\lambda(\alpha+\beta)}} \left(\frac{1}{G_j}\right)^{\frac{\lambda}{1+\lambda(\alpha+\beta)}} N_j^{-\frac{\lambda\alpha}{1+\lambda(\alpha+\beta)}} \left(\frac{\alpha+\beta}{\kappa}\right)^{-\frac{\lambda(\alpha+\beta)}{1+\lambda(\alpha+\beta)}} \tag{2.10}$$

In (2.10), as λ shrinks to zero with rising criminal skill, I_j^{stolen} converges to I_j indicating that all inventory is stolen regardless of P (as is evident in (2.8)). At the opposite extreme, as λ becomes large, criminal skill goes to zero, deterrence increases, and inventory stolen goes to zero. In this instance, investment in protection converges to its lower bound of 1 in (2.9). Also evident from (2.10), higher G_j reduces the amount of inventory stolen. This is because spatial concentration among retailers lowers the cost of protection in (2.7) while increasing equilibrium protection in (2.9). Expression (2.10) also makes clear that crime increases with the level of retail activity, as measured by inventory. This is consistent with previous literature (Lee and Alshalan, 2005; Tillyer and Walter, 2019).

2.2 Retail rent capitalization

We next consider the effect of inventory lost to crime on retail rent. Suppose initially that all retail establishments are identical. This assumption is relaxed later in the section and in the empirical work. We also defer discussion of shopping externalities to the end of this section where we show that shopping externalities cause our model to underestimate capitalization effects of crime deterrence.

2.2.1 Homogenous establishments

Each retail establishment sells q units of merchandise at p. Retailers hire labor L at a wage w, rent space S at a rent per square foot r, and purchase inventory from wholesalers at a per unit cost c. Retail product price, wage, and inventory cost (p, w, and c, respectively) are determined at the metropolitan level and do not vary across neighborhoods. The share of inventory lost to crime shrinks with the neighborhood spatial concentration of retail activity, G_j. Output q is produced using labor and space, the productivity of which are assumed to be amplified by shopping externalities that also increase with G_j.

Collecting terms, profit for an establishment in neighborhood j is given by,

$$\pi = pq(L_j, S_j; G_j) - wL_j - r(G_j)S_j - C(c, G_j)q(L_j, S_j; G_j) \ , \tag{2.11}$$

In (2.11), observe that $C(c, G_j)$ is the cost of inventory for each unit sold where C increases with c and is inversely related to G_j. Because G_j enhances both crime deterrence and shopping externalities, higher G_j should be capitalized into higher local rent, r_j.

We define the cost of inventory $C(c, G_j)$ as,

$$C(c, G_j) = c + cP^{-\lambda}(G_j) \ , \tag{2.12}$$

where c is the wholesale price per unit of inventory as above, P is the level of protection as defined in (2.10), and $P^{-\lambda}(G_j)$ is the share of inventory lost to crime as defined in (2.8). Substituting (2.12) into (2.11) and setting $\pi = 0$ with competitive markets, expenditure on space per dollar sold can be written as,

$$\frac{r(G_j)S_j}{pq(G_j)} = \theta - \frac{wL_j}{pq(G_j)} - \gamma(G_j) \ , \tag{2.13}$$

where $\theta = \frac{p-c}{p}$ and $\gamma(G_j) = \frac{cP^{-\lambda}(G_j)}{p}$.

In (2.13), θ is the percentage markup of retail to wholesale price and is common across establishments in the metropolitan area. The term $\gamma(G_j) = \frac{cP^{-\lambda}(G_j)}{p}$ is a neighborhood-specific markup that allows for the share of inventory lost to crime. The remaining term, $\frac{wL_j}{pq(G_j)}$, is labor cost per dollar sold. In our data we observe $\frac{L_j}{pq(G_j)}$ which is included as a control in some of the regressions. The coefficient on that term provides an estimate of w. In an alternate specification, we shift $\frac{wL_j}{pq(G_j)}$ to the left side of the equation and use earnings data for New York City from the U.S. Bureau of Labor Statistics (BLS) to measure labor cost per dollar of sales. Results from the two specifications are quite similar. Bearing this in mind, we rewrite (2.13) as,

$$\frac{r(G_j)S_j + wL_j}{pq(G_j)} = \theta - \gamma(G_j) \ , \tag{2.14}$$

where the dependent variable measures non-inventory costs per dollar of sales.

2.2.2 Heterogeneous establishments

In this section we highlight three sources of heterogeneity that affect the dependent variable in (2.14). One is that companies belong to different industries, k = 1, … K, each of which may have its own markup, θ_k. In the estimation to follow, we allow for this by including industry SIC 2-digit fixed effects. A second source of heterogeneity are neighborhood attributes apart from G_j that may also affect productivity. These terms are represented by z_j and could include spatial concentration of non-retail economic activity, neighborhood level proxies for potential demand, and more. A third source is establishment-level skill that reduces S and L for a given q, denoted as θ_i, for i = 1, … I. Adding these terms and suppressing the establishment-specific i subscripts on S and L to simplify, (2.14) becomes,

$$\frac{r(G_j)S_j+wL_j}{pq(G_j)} = \theta_k - \gamma\left(G_j\right) + bz_j + \theta_i \quad . \tag{2.15a}$$

Our primary goal with (2.15a) is to measure the effect of G_j on $\gamma\left(G_j\right)$. We do this in three ways. In the first approach, we estimate (2.15a) using neighborhood fixed effects to measure $\gamma\left(G_j\right)$. These are then regressed on G_j to summarize the average relationship between $\gamma\left(G_j\right)$ and G_j. A second, more general approach is to estimate (2.15a) using Robinson's (1988) partial linear model in which $\gamma\left(G_j\right)$ is estimated nonparametrically while specifying a parametric structure for the other model terms. Results from both approaches indicate that $\gamma\left(G_j\right)$ is approximately linear in G. Partly for that reason and to simplify presentation, in a third approach we impose a linear approximation on $\gamma\left(G_j\right)$. Taking a first order Taylor expansion of $\gamma\left(G_j\right)$ around $\bar{G}$, the sample mean of G, (2.15a) becomes,

$$\frac{r(G_j)S_j+wL_j}{pq(G_j)} = \bar{\theta}_k - \gamma'(\bar{G})G_j + bz_j + \theta_i \quad , \tag{2.15b}$$

where $\bar{\theta}_k = \theta_k - \gamma(\bar{G}) + \gamma'(\bar{G})\bar{G}$. Notice that if $\gamma(\bar{G})$ is linear then $\gamma(\bar{G}) = \gamma'(\bar{G})\bar{G}$ and $\bar{\theta}_k = \theta_k$ so that $\bar{\theta}_k$ equals industry markup. Also, and of primary interest, with further manipulation, the coefficient on G_j

in (2.15b) can be written as $\gamma'(\bar{G}) = -aP_j^{-\lambda}\bar{G}^{-1}$, where a is a positive constant.[17] This measures the marginal effect of G on the cost of inventory lost to crime per dollar sold evaluated at $\bar{G}$.

In (2.15a) and (2.15b), note that neighborhood attributes that enhance productivity and/or reduce crime deterrent costs are expected to increase non-inventory costs per dollar sold while a business owner's skill does the reverse. This is because local productivity advantages from $\gamma(G_j)$ and bz_j should be capitalized into higher local rent. Entrepreneur skill, as captured by θ_i, should instead increase profit. In the empirical work to follow, we proxy for θ_i using establishment age as it is well established that older companies tend to be more productive. As anticipated, non-inventory costs per dollar sold decline with establishment age but increase with neighborhood level retail spatial concentration.

2.2.3 Lower bound on rent capitalization from crime deterrence

As suggested earlier, the model above likely yields a lower bound on the degree to which crime deterrence arising from retail spatial concentration is capitalized into higher neighborhood rent. To see why, recall that shopping externalities also increase with spatial concentration of retail activity (e.g. Gould et al., 2005; Koster et al., 2014; and Koster et al., 2019). For a given level of q, shopping externalities likely increase productivity by reducing the need for advertising and by improving the ability of store managers to anticipate flows of shoppers, enabling them to use space and labor more efficiently. These advantages will also be capitalized into higher local rent.[18]

---

[17] Recall from (2.13) that $\gamma(G_j) = \frac{cP^{-\lambda}(G_j)}{p}$. Differentiating with respect to G,

$$\gamma'(G_j) = -\frac{c}{p}\lambda P_j^{-(1+\lambda)}\frac{\partial P_j}{\partial G_j} = -aP_j^{-\lambda}G^{-1} < 0 \quad , \tag{N.1}$$

where $a = \left(\frac{c}{p}\right)\frac{\lambda}{1+\lambda(\alpha+\beta)} > 0$ and from (2.9), $\frac{\partial P_j}{\partial G_j} = \frac{1}{1+\lambda(\alpha+\beta)}P_j G^{-1}$. Taking a first order Taylor expansion of $\gamma(G_j)$ around $\bar{G}$ and substituting (N.1) into (2.15a) gives $\gamma'(\bar{G}) = -aP_j^{-\lambda}\bar{G}^{-1}$.

[18] Shopping externalities also of course have potential to increase q for a given retailer, requiring purchase of additional S, L and inventory. By itself, however, higher q would not affect the firm's profit margin or the dependent variable in (2.15b) if production is roughly constant returns to scale as seems likely for in-person retail establishments. Instead, it is the productivity advantages associated with shopping externalities that enhance net profit. See Koster et al. (2019) for related discussion.

Suppose now that crime deterrent effects are absent so that spatial concentration only affects profit through shopping externalities. Also, hold constant the level of space and labor used in production. Taking the derivative of the dependent variable in (2.15b) with respect to G_j and manipulating, productivity advantages from shopping externalities cause input costs per dollar sold to shrink if the following condition holds:

$$\frac{r\prime(G_j)S_j}{r(G_j)S_j+wL_j} < \frac{q\prime(G_j)}{q(G_j)} \tag{2.16}$$

where $r'(G_j)$ and $q'(G_j)$ are derivatives with respect to G_j.

Note now that $wL_j > 0$ so that $\frac{r\prime(G_j)S_j}{r(G_j)S_j+wL_j} < \frac{r\prime(G_j)}{r(G_j)}$. Also, $\frac{r\prime(G_j)}{r(G_j)}$ and $\frac{q\prime(G_j)}{q(G_j)}$ are approximately equal to $\%\Delta r$ and $\%\Delta q$, respectively. A sufficient condition for (2.16) to hold is that productivity gains from shopping externalities have a smaller percentage effect on r than on q, which is extremely likely. An even more general condition is that (2.16) will hold provided that $\%\Delta r$ is not substantially *larger* than $\%\Delta q$ (since $\frac{r\prime(G_j)S_j}{r(G_j)S_j+wL_j} < \%\Delta r$). This condition is undoubtedly met. Shopping externalities, therefore, shrink the dependent variable in our capitalization expressions while crime deterrence has the opposite effect. For this reason, our model will understate the rent capitalization effect of crime deterrence.[19]

3. Data, Neighborhoods, and Summary Statistics

Multiple data sources were used to estimate the models above, all of which focus on New York City for roughly 2018. For the crime and police stop models, both dependent variables and all of the controls vary at the neighborhood level. For the capitalization models, the dependent variable is at the establishment level and controls vary at that level or at the neighborhood level. All of the data, including the dependent variables and controls, are initially obtained as point-specific measures with street address and/or latitude and longitude. Information is then aggregated up to create neighborhood level measures.

[19] Note also that if capitalization of G into higher rent prompts retailers to substitute L for S, related cost savings will further cause our model to understate capitalization effects from crime deterrence.

Partly for that reason, below we first describe how neighborhoods are measured. This is followed by further description of the data and summary measures.

3.1 Measuring neighborhoods and spatial concentration

3.1.1 Defining neighborhoods

For all of our models, we divide NYC into 0.2 by 0.2 mile grid squares. This corresponds to roughly 2 Manhattan blocks traveling east-west and three blocks traveling north-south, or roughly six city blocks. Grid square boundaries are independent of political and other administrative boundaries. Each grid square is treated as a separate neighborhood. It is worth emphasizing that the grid squares are small enough to be relatively homogenous. This helps to control for unobserved local attributes. At the same time, the grid squares are large enough to allow for within-grid square variation in the degree of spatial concentration of economic activity and other measures.

3.1.2 Measuring spatial concentration within grid squares

We use the Getis-Ord statistic to calculate spatial concentration in a given grid square (Getis and Ord, 1992; Ord and Getis, 1995). This statistic is widely used for Hot-Spot analysis, especially for policing strategies that target hot-stop crime areas. Adopting the notation from Section 2, we refer to the spatial concentration measure as *G*. An appealing feature of the Getis-Ord measure is that it allows for point-specific location with a distance weighting function based on continuous distance. To simplify exposition, a given target establishment is always indexed by *i* while all other establishments in our NYC sample are indexed by *e* = 1, … *n*. The Getis and Ord expression for G_i is given by,

$$G_i = \frac{\sum_{e=1}^{n} \omega_{ie} x_e - \bar{X} \sum_{e=1}^{n} \omega_{ie}}{\sqrt{\frac{\sum_{e=1}^{n} x_e^2}{n} - \bar{X}^2} \sqrt{\frac{\left[n \sum_{e=1}^{n} \omega_{ie}^2 - \left(\sum_{e=1}^{n} \omega_{ie}\right)^2\right]}{n-1}}} \quad . \tag{3.1}$$

In this expression, x_e is employment at establishment e and $\bar{X}$ is the average size of an establishment throughout our NYC sample.

A key feature when implementing (3.1) is to specify a function for ω_{ie}, the weight placed on employment as distance, d_{ie}, increases from establishment *i*. We adopt the following weight function:

$$\omega_{ie}(d_{ie}) = \begin{cases} 1, & \text{if} \quad d_{ie} \leq 250 \\ 1/(d_{ie} - 250)^{0.7}, & \text{if} \quad 250 < d_{ie} \leq 1{,}000 \\ 0, & \text{if} \quad d_{ie} > 1{,}000 \end{cases} \tag{3.2}$$

This function sets ω_{ie} to 1 for all establishments within 250 feet of *i*, roughly equal to one block going north-south in Manhattan. For distances between 250 to 1,000 feet from *i*, ω_{ie} is assumed to decline with distance at rate $1/(d_{ie} - 250)^{0.7}$, where the exponent 0.7 was chosen to set ω_{ie} to roughly 1% at 1,000 feet. Beyond 1,000 feet, ω_{ie} is set to zero. Measured in this fashion, the weight function will often apply positive weight to employment beyond the border of a grid square.[20] It is also worth noting that specified as above, G_i is measured separately for each establishment and varies within a given neighborhood.

Two additional measurement issues are important to note. First, *G* is calculated separately for retail and non-retail industries (including service, manufacturing and finance). This allows us to evaluate whether spatial concentration of retail activity has a different effect on our dependent variables (crime, police stops and rent) relative to spatial concentration in other industries. This also means that our sample of neighborhoods are restricted to those in which all industries just noted have some presence.

Second, for the crime and police stop models, the unit of analysis is the neighborhood. For that reason, we use a neighborhood level measure of spatial concentration in those models. We also normalize that measure across neighborhoods to simplify interpretation. The resulting measure $\tilde{G}_j$, is formed as,

$$\tilde{G}_j = \frac{1}{sd(G_j)}\left(\bar{G}_j - \frac{1}{m}\sum_{j=1}^{m} \bar{G}_j\right) \quad . \tag{3.3a}$$

In this expression, $\bar{G}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} G_i$ is the average level of spatial concentration in neighborhood *j* (with n_j establishments) while *m* is the number of grid squares in the estimating sample. Measured as above, $\tilde{G}_j$ is positive if the average level of spatial concentration in neighborhood *j* is high relative to the typical

[20] Results were similar for alternate thresholds from 250 and 1,000 feet. Results were also robust for other rates of decay for the inverse distance portion of the ω_{ie} function, including exponents equal to 0.5, 1, and 2.

neighborhood. Also, a 1 unit increase in $\tilde{G}_j$ represents a 1 standard deviation increase in spatial concentration across grid squares.

For the rent models, the unit of analysis is the individual establishment. In those models, we still normalize spatial concentration in a fashion analogous to above while allowing the concentration measure to vary across establishments within individual neighborhoods. Specifically, we form,

$$\tilde{G}_i = \frac{1}{sd(G_i)}\left(\bar{G}_i - \frac{1}{I}\sum_{i=1}^{I}\bar{G}_i\right) \quad . \tag{3.3b}$$

where I is the number of establishments throughout the entire estimating sample.

3.1.3 Alternate measures of $\tilde{G}$

As described above, $\tilde{G}$ is a proxy for crowding and visibility that may enhance the crime deterrent effects of eyes on the street (hereafter we drop the *j* and *i* subscripts for convenience). In the estimation to follow, our primary measure of $\tilde{G}$ is based on the spatial distribution of employment and is denoted as $\tilde{G}_{Emp}$. In some models we also add a measure of $\tilde{G}$ based on the spatial concentration of store fronts, $\tilde{G}_{Stores}$. This measure targets visibility and helps to parse out its separate role from crowding. We further experiment with a third measure of $\tilde{G}$ based on sales, denoted as $\tilde{G}_{Sales}$. Controlling for $\tilde{G}_{Emp}$ and $\tilde{G}_{Stores}$, we have little reason to expect that $\tilde{G}_{Sales}$ would enhance the effect of eyes on the street. For this reason, $\tilde{G}_{Sales}$ acts as a placebo and a check on our model specification.

We use the same sample to measure all three types of $\tilde{G}$ to ensure consistency across measures. That sample is restricted to single site firms which account for 95 percent of establishments in New York City and are typically much smaller than those belonging to multi-site firms.[21] This ensures that when $\tilde{G}$ is large that is not because of the presence of a single large establishment regardless of whether $\tilde{G}$ is based on employment, store fronts or sales.[22] Also, establishment-level sales are more reliably measured for

[21] Among retail establishments, those that belong to multi-site firms have 20, 75, and 222 workers at the 75th, 95th and 99th size percentiles, respectively. For single site companies the corresponding values are 4, 12, and 35.
[22] We also omit a small number of establishments with sales per worker above \$500,000 when measuring $\tilde{G}$.

single-site firms for which sales are assigned to only one establishment. We also estimated all of our models having measured $\tilde{G}$ using all establishments no larger than 15 workers, including single-site and multi-site companies (15 workers is roughly equal to the 95^{th} size percentile among single-site retailers). Estimates were close to those reported later in the paper. In a further robustness check, in Appendix B we define neighborhoods using 3 by 3 configurations of grid squares and measure spatial concentration as the sum of squared employment shares across the 9 squares. Results were again similar to those in Table 2.

3.2 Data

3.2.1 Property crime

The NYPD provides data on all criminal complaints reported since 2006. In most instances, we use only 2018. For each crime, the data includes the date, time, precise location, and type of crime of the incident. Property crime includes petit larceny, grand larceny, burglary, theft of services and fraud. We aggregate these crimes together for our core models but estimate separate models in other instances, in addition to models for robbery and auto theft which also entail theft of property.

We drop all crimes that took place on a bus, subway or at a subway station as the location where the event was initiated may not be accurately recorded. We also drop all complaints that refer to attempted crimes where the perpetrator left the scene before fully committing the offense (less than 2% of all events), and crimes that extend beyond one day, as with kidnapping and/or hostage situations. This leaves 101,896 property crimes in the analysis to follow.

3.2.2 Police stops

Police stop data were obtained from the New York Police Department (NYPD) and are reported as part of the Stop-Question-Frisk (SQF) policy. We pool police stops from 2016-2018 as this helps to ensure a large enough number of stops to obtain reliable estimates.[23]

[23] Restricting police stop data to 2018 did not affect the results but increased the share of grid squares with zero stops from 50% to 75%. Using Zero-Inflated Negative Binomial models also did not affect the results.

We recognize that SQF has been controversial because of concerns about racial profiling and bias. Although the SQF policy was initiated in NYC prior to 2000, it became widely used during the Bloomberg administration (2002-2013) when stops grew from 97,296 in 2002 to over 500,000 in 2012 (Evans et al., 2014). By 2012, numerous newspaper articles had reported on allegations of excessive force being used when stopping African Americans and many court cases had followed. In response, by 2016 SQF stops were substantially reduced and better targeted at potential criminal activity as evidenced by sharply higher arrest rates upon making a stop (Urrego, 2021). In 2016, 2017 and 2018 the number of SQF stops were just 12,053, 11,204 and 11,008, respectively.[24]

SQF stops include only those of pedestrians. The SQF data report the date, time of day and location of each stop. Also reported is whether an arrest was made and if so for what type of crime. We drop any stops prompted by a 911 call or which were associated with an ongoing investigation. Instead, we retain only stops in which police were acting proactively. This reduces the number of stops in the analysis, from 34,265 (for the 2016-2018 period) to 7,277.

3.2.1 Employment, sales, and industry

Dun & Bradstreet provides information on more than one million establishments in the New York City area. For each establishment, we observe employment, sales, industry code, address, and latitude-longitude coordinates. The data were downloaded from the Syracuse University library (which has a site license) between October 2018 and February 2019 and are current as of that time.

3.2.2 Commercial rent

Commercial lease data were obtained from CompStak and are used to estimate the rent

[24] Studies based on early years of SQF found that African Americans and Hispanics were stopped at a higher rate than their white counterparts, even after controlling for neighborhood racial composition and criminal activity by race (Gelman et al., 2007; Ridgeway, 2007; Hanink, 2013; Evans et al., 2014; Ferrandino, 2018;). However, MacDonald and Braga (2019) show that racial patterns associated with SQF appear to have decreased in more recent years. Also worth noting, other studies have found that SQF stops reduce crime (Weisburd et al., 2014; Wooditch and Weisburd, 2016; MacDonald et al., 2016; Rosenfeld and Fornango, 2017; Ferrandino, 2018).

capitalization models. For each lease, CompStak reports effective rent per square foot of space leased, location of the lease (including street address and latitude and longitude), and tenant name. Our lease sample includes over 60,000 leases in NYC that were executed up to December, 2019. We match these data at the establishment level with the D&B data using information on tenant name, street address, and latitude/longitude coordinates. In total, we are able to reliably match almost 50% of the CompStak leases. Of the leases that were matched, 4,000 are classified as retail establishments in D&B with a primary SIC code 52-59. Of these, more than half include missing information on sales, employment or space leased, measures needed to estimate the models in expressions (2.14), (2.15a) and (2.15b). This leaves us with roughly 1,600 observations for the retail rent capitalization portion of the analysis. An additional roughly 550 matched observations are used to estimate analogous models for wholesale establishments.

3.2.3 Additional neighborhood attributes

Numerous measures were obtained from various New York City government agencies and coded up to the neighborhood level. This includes data obtained from the New York City Department of City Planning and the Department of Finance MapPLUTO 18v2 map. This map provides detailed information on the attributes of each tax lot in NYC. This includes information about the building situated on the lot, zoning, tax assessments, and many other lot specific characteristics. Additional neighborhood level data were obtained from the NYC Department of Health and Mental Hygiene, Fire Department, Department of IT and Telecommunications, The NYC Community Air Survey, the NYC Open Data portal, and the 2015 Tree Census conducted by the NYC Department of Parks and Recreation.

In our more robust models, we also control for visits to Points of Interest (POI) using cellphone data obtained from SafeGraph. SafeGraph defines over 110,000 POI in New York City and measures visits to each POI using cellphone GPS information combined with information on building footprints and other relevant information (e.g., store open hours). In cleaning these data, we first calculate the average number of monthly visitors to each individual POI during 2018. For each grid square, we then average monthly visits across POI within a grid square. That measure is included in our models to further control for grid square level foot traffic. It is worth noting that POI visits include activity associated with business and non-business related activity. Along with neighborhood level employment, these measures provide considerable information on the level of activity in a neighborhood.[25]

All of the models in the estimation to follow always include neighborhood-level measures of the share of residential units, total number of trees in the grid square, whether the grid square overlaps multiple police precincts, average age of buildings, average assessed value of buildings, and average sales per worker for single-site establishments (including companies in all industries). In some robustness checks, up to 25 additional neighborhood attributes are added to the crime and police stop models.

A complete list of neighborhood controls and their sources is provided in Appendix A.

3.3 Summary statistics

In total we define 3,506 grid squares across the NYC area. Table 1 provides the summary statistics of the estimating sample. In 2018, a grid square experienced an average of 29 property crimes, 63% of which were petit larcenies, with the total number of property crimes equal to 101,896. The number of police stops used in the analysis is smaller, just 7,277 as described above. Note also that roughly half of grid squares experience no police stops whereas the number of grid squares that reported zero property crime is below 2%. Because the crime and police stop data are count measures, and to allow for zeros, we estimate both the police stop and crime models using a negative binomial specification. This model is well suited to sample distributions such as ours for which variance of the outcome measures exceed their means.[26]

In Table 1, notice that the service industry accounts for the highest share of employment among the industries highlighted (46%), followed by retail (19%), finance (7%), and manufacturing (5%).

[25] We also calculated spatial *G* measures for POI. Correlation between that measure and spatial concentration of retail employment was just 8%. Including spatial concentration of POI visits also had no effect on the coefficients on the other model estimates and was dropped from the regressions to simplify specification and discussion.
[26] We also estimated OLS regressions for both police stops and property crime, setting the dependent variables to log(X+1) with X suitably defined in each instance. Results were similar.

Observe also that of the four industries highlighted, retail employment is the most spatially concentrated based on both the median and 75th percentile values across the sample of neighborhoods.

Also of note in Table 1, observe that the distribution of the ratio of retailer cost of space per dollar of sales often exceeds one and especially so for non-inventory costs per dollar sold (which includes labor costs). This is to be expected given the heavy reliance of small businesses on financing. It also likely reflects the high tendency of businesses to fail in their first few years, presumably because they are not able to generate sufficient revenue to cover costs. Based on data from the US Bureau of Labor Statistics, roughly 20% of small businesses fail in their first year, 50% by their fifth year, and 70% in their first ten years. We anticipate therefore that the cost/sales ratio will typically exceed one as in Table 1.

4. Results

4.1 Property crime and police stops

4.1.1 Core estimates

Table 2 reports estimates of the effect of employment-based spatial concentration of retail activity on property crime and police stops, $\tilde{G}_{Emp}$. As noted earlier, these estimates are obtained from negative binomial count models that address zeros in the data. Marginal effects evaluated at the mean of the full set of control measures are reported.[27] Recall that $\tilde{G}_{Emp}$ is normalized to have mean zero and standard deviation of 1 so that a one-unit change in $\tilde{G}_{Emp}$ equals one standard deviation across neighborhoods. Columns 1-4 present estimates of property crime while columns 5-8 repeat the estimation with police stops as the dependent variable.

For both the crime and police stop models, controls increase moving from left to right across the table. Columns 1 and 5 control for aggregate employment in the grid square, retail share of grid square employment, and spatial concentration of retail employment. Columns 2 and 6 add in controls for the

[27] Marginal effects from a negative binomial regression can be calculated using the expression, $\exp(\beta_x \Delta x)$, where Δx represents the change in a control variable and β_x is its corresponding coefficient. This expression gives the change in counts. β_x also approximates a semi-elasticity because a one unit change in x represents a change in the log count of the dependent variable equal to β_x.

distribution of grid square employment across industries, including retail, finance insurance and real estate (FIRE), manufacturing and service. Also included in these models are measures of $\tilde{G}_{Emp}$ for the non-retail industries just noted. Columns 3 and 7 add in controls for additional neighborhood level features. These include log of grid square sales per worker (based on single site establishments), the log number of trees in the neighborhood, average age of the buildings, log of the average building assessed value, whether the grid square overlaps with more than one police precinct, and the share of residential building units from among all buildings in the grid square. In columns 4 and 8 we include average monthly visits to POI in the grid square to further control for the level of neighborhood activity.

In Table 2, notice first that the coefficients on the primary measures of interest as reported in columns 1 and 5 are quite robust. This includes total neighborhood employment, retail share of employment and retail spatial concentration. The coefficients on these measures are little changed moving from left to the right to the most fully specified models in columns 4 and 8. Also worth noting is that business establishments may shy away from high crime areas and/or may spatially concentrate in such locations to gain better protection. To the extent that occurs, the estimates in Table 2 will understate the effect of employment levels and spatial concentration on crime and police stops.

Focus next on columns 4 and 8 and consider the level of activity in a grid square. Doubling grid square employment increases property crime and police stops by 55% and 35%, respectively. The corresponding effects from a doubling of visits to POI are 52% for property crime and 89% for police stops. These estimates confirm that property crime and police stops increase with the overall level of foot traffic in a neighborhood (as captured by POI visits) and especially so with an increase in business activity (as reflected in total employment).

Also noteworthy in Table 2, the coefficients on retail share of employment in both the crime and police stop models are large, positive, highly significant, and much larger in magnitude than for the other industries (finance, manufacturing and service): a 1 percentage point increase in retail share of employment is associated with a 2.3% increase in property crime and a 2% increase in the number of police stops. These patterns confirm that retail activity, with its lucrative inventory, has an especially large effect on property crime and police activity.

Observe next that in column 4, a one unit increase in $\tilde{G}_{Ret,Emp}$ – equal to a one standard deviation increase in spatial concentration of retail employment across neighborhoods – is associated with an 8.5% decline in crime. In contrast, coefficients on spatial concentration for the other industries (finance, manufacturing, and service) are much smaller and mostly not significant. This pattern suggests that spatial concentration of retail activity has a particularly important effect on eyes on the street, reducing the cost of protection and causing equilibrium levels of crime to decrease.

The same pattern is present in column 8 for the police stop model with the exception that spatial concentration of service employment also has a similar negative effect as for the retail sector. Increasing $\tilde{G}_{Ret,Emp}$ by 1 standard deviation decreases police stops by 11%. This reinforces the view that because spatial concentration reduces crime it also reduces investment in protection measures. We recognize that a critique of this interpretation could be that police may limit pedestrian stops when their own behavior is more readily observed (see Owens, 2019, 2020, for related discussion). That concern is mitigated, however, by evidence in column 8 that total neighborhood employment and foot traffic to POI both have strong positive effects on police stops. Also, non-retail employment always has a near zero effect on stops as does spatial concentration for FIRE and manufacturing. These patterns would not be anticipated if police shy away from making stops when others are present.

4.1.2 Additional controls

Table 3 adds in up to 25 additional neighborhood level controls to columns 4 and 8 of Table 2 and serves as a further check on robustness. The full list of additional measures is provided in Appendix A. These include controls for establishment attributes (age and risk profile of neighborhood establishments), zoning (historic district, FAR restrictions), distance to important sites (e.g. subway stations, public parks), other neighborhood features (e.g. presence of rats, problems with light shine), and tax exemptions on buildings. Each row in Table 3 adds in a different subgroup of controls to columns 4 and 8 of Table 2. The bottom row of Table 3 includes all 25 additional controls while the top row repeats the estimates from columns 4 and 8 of Table 2 for reference. Results are reported for just the coefficients on retail share of employment and retail spatial concentration of employment.[28]

Reading down the rows of Table 3 it is apparent that estimates are largely similar when different groups of additional controls are added to the crime and police stop models. This suggests that our core estimates are robust.

4.1.3 Mechanisms: crowding and visibility

Table 4 presents estimates of the crime model for different sample periods and crimes (we revisit the police stop models in Table 5). The models in Table 4 help to shed light on crowding and visibility as mechanisms that may account for why spatial concentration of retail activity deters crime. Recall that crowding may make it more difficult for criminals to steal merchandise without being observed although we note that this need not be the case if criminals are able to hide among a crowd. Visibility also makes it more difficult for criminals to go unobserved if police patrols, other people, and/or security cameras have better lines of sight on potential targets of property crime.

Each row in Table 4 corresponds to a different crime. To conserve space, only coefficients on retail share of employment and spatial concentration of retail are reported. Estimates for property crime are in the top row with different sub-categories of property crime immediately below. This includes Grand Larceny plus Burglary, Petit Larceny, and Theft of Services plus Fraud. Grand Larceny and Petit Larceny differ based on the value of merchandise stolen and could occur with or without breaking into a

[28] We also explored adding controls for the portion of a grid square that belongs to a Business Improvement District (BID) as BIDs may pool resources and invest in local security measures (see Faggio, 2021, for related work on the relationship between BIDs and crime). To do so, we experimented with different combinations of dummy variables for the share of a grid square that is within the boundaries of a BID, including 0, more than 0 and less than 25%, 25% to 50%, etc. In some specifications there was evidence of a positive relationship between BID presence and neighborhood crime. That could arise if BIDs form in high crime areas in part to help protect against crime (see Faggio, 2021, for similar results), in which case BID presence would be endogenous. We also found that BID presence did not have any discernible effect on police stops, and most important, including BID measures had no effect on the core estimates including those for spatial concentration. For these reasons, we chose not to include controls for BID presence in the crime and police stop models.

store. Break-ins are a defining feature of burglary. Theft of services often occurs when patrons leave a restaurant or hotel without paying for services provided. Fraud occurs when an individual uses someone else's credit card, forges signatures on a check, etc. Appearing below these crimes are Robbery and Auto Theft. Robberies occur when a victim is physically threatened regardless of whether a weapon is displayed. This is considered always to be the case if a victim is present during a break-in at a retail establishment. It should also be noted that roughly half of robberies reported in the NYC data occur on the street and not within a building. We did not attempt to decompose robberies by place of occurrence given the relatively small number of robberies in the data (see Table 1).

The columns in Table 4 refer to different time periods. Columns (1) and (2) include all hours of the day and periods. Columns (3) and (4) pertain to crime during daytime hours, while columns (5) and (6) refer to crime at night. Columns (7) and (8) focus on crime during the first two weeks of the COVID-19 lockdown in New York City (for all hours of the day), March 22 to April 5 in 2020.[29] Crowding is sharply reduced at night and especially during the lockdown when individuals were ordered to remain at home.[30] Visibility is reduced but not eliminated at night because of extensive street lighting but would have been fully viable during the lockdown.

Consider first the differences between daytime and nighttime patterns for crime. For property crime, at night the coefficient on retail share of employment is reduced by roughly 25% but remains large and highly significant. This suggests that inventory continues to attract criminal activity though thieves will need to break into stores that are closed, adding burglary to their crime. A different pattern is present for retail concentration. The coefficient on that measure shrinks by roughly 95% at night and is no longer significant. This suggests that deterrent effects from crowding that arise from spatial concentration of retail activity are greatly reduced at night. These patterns are present for all categories of property crime.

[29] During these two weeks, stores were largely closed for in-person visits but the NYPD was still largely fully staffed. As noted in the Introduction, not long after, the high rate of COVID-19 cases left the NYPD short staffed which may have reduced patrol activity.
[30] SafeGraph cell phone data indicate that foot traffic in New York City fell by 60% in April, 2020 relative to mid-January of that year.

For robbery, estimates are similar to property crime during the day but weaken only slightly at night: both retail employment and spatial concentration continue to have strong effects of the anticipated signs (positive and negative, respectively). Because crowding is absent at night, this suggests that visibility remains a viable mechanism in helping to prevent robbery, possibly because roughly half of robberies occur on the street as noted above.

Auto Theft exhibits a similar pattern as property crime during the day but the coefficients on retail employment and spatial concentration are noticeably smaller. At night, however, the coefficient on spatial concentration increases in magnitude, is negative, and is strongly significant. If the nighttime concentration of parked cars is higher closer to retail establishments (which includes bars and restaurants), then a similar explanation as for robberies may apply: Nighttime police patrols may observe more cars at once where vehicle density is higher, and this helps to deter auto theft.

Columns (7) and (8) of Table 4 provide analogous estimates for the first two weeks of the spring 2020 COVID-19 lockdown in New York City. The dominant patterns are the same as for nighttime crime in column 3. For property crime, spatial concentration has much less effect relative to 2018 (in column 2) but spatial concentration has a similar deterrent effect on robberies as for the pre-pandemic period.

Bearing in mind that crowding is sharply reduced at night and was almost eliminated during the lockdown, the patterns in Table 4 suggest that crowding and visibility have different effects on the different types of crimes considered in these tables. For Petit Larceny and Theft of Services/Fraud, the effect of retail spatial concentration is especially small and not significant both at night and during the lockdown. This along with other patterns suggests that crowding is an important deterrent of these crimes. For robbery, grand larceny (during the pandemic lockdown), and auto theft, the patterns suggest that visibility also acts as a deterrent.

Table 5 presents additional models that use a more targeted approach to uncover evidence of mechanisms, in this instance for both crime and police stops. We control for the three different types of $\tilde{G}$ discussed earlier, including measures based on employment ($\tilde{G}_{Emp}$), storefronts ($\tilde{G}_{Stores}$), and sales

($\tilde{G}_{Sales}$). The employment-based measure targets crowding while the storefront measure proxies the potential to observe multiple storefronts from a single location, a feature of visibility. Controlling for these measures and the other model controls, spatial concentration of sales is not especially related to eyes on the street as a deterrent of crime. For that reason, $\tilde{G}_{Sales}$ is included as a placebo check.

Consider now Panel A of Table 5, which reports the correlation across neighborhoods between the three measures of spatial concentration for the retail industry. As would be anticipated, correlation between $\tilde{G}_{Ret,Emp}$ and $\tilde{G}_{Ret,Sales}$ is relatively high, 54%. Correlation between $\tilde{G}_{Ret,Emp}$ and $\tilde{G}_{Ret,Stores}$, however, is just 14%, nearly the same as for correlation between $\tilde{G}_{Ret,Sales}$ and $\tilde{G}_{Ret,Stores}$. These summary measures confirm that $\tilde{G}_{Ret,Emp}$, $\tilde{G}_{Ret,Stores}$ and $\tilde{G}_{Ret,Sales}$ contain different information.

Estimates of the crime and police stop models are presented in Panel B of Table 5. Models are estimated for the same mix of time periods as in Table 4. The exception is for the first two weeks of the COVID-19 lockdown for which there were not enough police stops to estimate the model. Observe that each row corresponds to a different time-period regression with corresponding coefficients arrayed across columns. These include retail share of employment and the three different measures of retail spatial concentration, $\tilde{G}_{Ret,Emp}$, $\tilde{G}_{Ret,Stores}$ and $\tilde{G}_{Ret,Sales}$. Coefficients on the other model controls are suppressed. As before, the crime models are reported in columns 1-4 and the police stop models in columns 5-8.

Focusing on police stops first, in Table 5 notice that $\tilde{G}_{Ret,Emp}$ reduces stops both during the day and at night, but $\tilde{G}_{Ret,Stores}$ has no effect in either period. This confirms that police stops are less prevalent in crowded areas but are apparently not so sensitive to lines of sight to storefronts.

The patterns for property crime reinforce those for police stops and also display differences. In this case, crowding as proxied by $\tilde{G}_{Ret,Emp}$, clearly deters property crime during the day and in the lockdown but as in Table 4, the effect of $\tilde{G}_{Ret,Emp}$ is greatly reduced at night when crowds are mostly absent. Along with estimates from the police stop models, this pattern appears to confirm that crowding enhances the crime deterrent effects of eyes on the street, reducing the need and for police patrols.

Different from the police stop models, notice also that visibility, as proxied by $\tilde{G}_{Ret,Stores}$, deters property crime during the day, at night and during the COVID-19 lockdown. The corresponding coefficients are highly stable and significant, ranging between -0.12 and -0.14 across sample periods. This pattern suggests that visibility and the potential for police, other individuals, and/or security cameras to observe multiple storefronts at once also enhances eyes on the street and helps to deter property crime.

A last point to note in Table 5 is that for both property crime and police stops, the coefficients on $\tilde{G}_{Ret,Sales}$ are always small and not significant. This was largely anticipated.

4.2 Rent capitalization

Our final models present estimates of retail rent capitalization associated with the crime deterrent effects of neighborhood spatial concentration of retail activity. For reasons described earlier, these estimates are likely lower bounds. Throughout this section, we also include only the employment-based measure of spatial concentration. We do this primarily to simplify presentation and also because $\tilde{G}_{Emp}$ is the only one of the three measures of spatial concentration that affect police patrols in Table 5.

We begin by estimating expressions (2.14) and (2.15a) from Section 2. Figure 1 displays scatter plots of the estimated fixed effects against neighborhood spatial concentration of retail employment on the horizontal axis. Panel A presents estimates without any controls for other factors (as in expression (2.14)) while Panel B controls for other neighborhood and establishment attributes that may affect the dependent variable (as in (2.15a)). In both panels, the scatter plots clearly increase with spatial concentration of retail employment. This confirms that benefits from crime deterrence are capitalized into higher rent.

Figure 2 displays an alternate set of estimates of $\gamma(\tilde{G}_{Ret,Emp})$. In this case, we estimate (2.14) and (2.15a) using Robinson's (1988) two-step partial linear model drawing on the semipar routine in Stata developed by Verardi and Debarsy (2012). This estimates the $\gamma(\tilde{G}_{Ret,Emp})$ function non-parametrically with optimal smoothing. As before, Panel A does not allow for other factors while Panel B controls for additional neighborhood and establishment-level attributes. In both panels the gamma function is clearly increasing with retail spatial concentration and the confidence bands are narrow relative to the overall pattern. It is also evident that $\gamma(\tilde{G}_{Ret,Emp})$ is approximately linear in $\tilde{G}_{Ret,Emp}$. This last observation supports our remaining empirical exercise.

Table 6 reports estimates of expression (2.15b) in which we impose a linear approximation on the relationship between $\gamma(\tilde{G}_{Ret,Emp})$ and $\tilde{G}_{Ret,Emp}$. Panel A uses rent per square foot per dollar sold as the dependent variable and includes labor per dollar sold as a control measure (as in expression (2.13)). Panel B shifts labor costs based on census data to the left side of the equation and uses non-inventory costs per dollar sold as the dependent variable as in Figures 1 and 2. Additional estimates in Panels C and D are based on 538 wholesale establishments and serve as a robustness check. In these panels, rent, labor costs and sales are all specific to the wholesalers in the sample, but spatial concentration is still measured based on retail employment. As noted earlier, wholesalers are less prone to shoplifting and may also be more able to adopt aggressive anti-crime measures that would discourage retail shoppers. For these reasons, wholesale rent should be less sensitive to crime deterrent effects from retail spatial concentration.

In each panel, five models are presented with increasing numbers of controls for neighborhood and establishment-level attributes. We control for spatial concentration of other industries in column 2. In column 3 we add in the neighborhood level control for monthly visits to POI. Column 4 accounts for police precinct fixed effects, and column 5 adds in fixed effects for establishment age. A quick review across columns confirms that retail spatial concentration is associated with higher retail and wholesale rent. The magnitude of the spatial concentration effects also tends to shrink with further controls.

Focus now on column 5 for Panels A and C. The coefficient on employment per dollar of sales in Panel A for retail establishments is roughly \$38,000. The analogous estimate in Panel C for wholesale establishments is \$70,000. In comparison, for 2019, the office of the New York State comptroller reports that average retail earnings per worker in Manhattan were \$59,400 and for all of NYC \$46,600. For wholesale industry workers the US Bureau of Labor Statistics reports average 2019 earnings in NYC of $88,000.[31] Our estimates are close to these values.

Consider now Panel B where we measure labor costs per dollar sold directly by combining the BLS measure of average earnings among NYC retail workers with Dun and Bradstreet data on employment and sales. That term is moved into the dependent variable as noted above and as in expression (2.15b). The estimated effect of spatial concentration of retail employment in Panel B is very similar to the corresponding estimate in Panel A. A one standard deviation increase in spatial concentration of retail increases non-inventory costs per dollar sold by roughly 32 cents, an increase of 7.8% relative to the mean value for non-inventory costs per dollar sold across the sample.

Panels C and D focus on wholesale establishments and provide a robustness check. Notice that a one standard deviation increase in $\tilde{G}_{Ret,Emp}$ for neighborhood retail activity has a significant effect on wholesale non-inventory costs per dollar of sales. The point estimates are 13.2 cents in Panel C and 20.9 cents in Panel D. These estimates suggest that wholesalers also benefit from enhanced eyes on the street associated with retail spatial concentration, but to a lesser degree than retail establishments. This is as expected.[32]

5. Conclusions

U.S. retailers lose over 2% of sales to property crime each year from a combination of theft and expenditures to protect against crime, a substantial amount relative to retail profit margins that average roughly 3%.[33] Local public authorities also devote considerable resources to patrolling retail districts. This paper considers the degree to which such costs can be reduced by spatial concentration of retail

[31] See https://www.bls.gov/cew/data.htm for BLS estimates and The Retail Sector in New York City: Recent Trends and the Impact of COVID-19 - December 2020 (state.ny.us) for discussion by the New York State comptroller.
[32] Other coefficients in the retail and wholesale models in Table 6 are in line with priors. This includes establishment age fixed effects. These display a strong monotonic pattern in which older companies have smaller non-inventory costs per dollar sold. This is consistent with priors that older companies that have survived a competitive weeding out process are more productive and enjoy lower cost-to-sale ratios than younger establishments.
[33] See discussion by the Small Business Resource Center at https://sbrc.employers.com/retail/whats-a-good-profit-margin-for-retailers/ and the associated report by Deloitte Inc. (2018).

activity that may enhance "eyes on the street", making it easier to apprehend and deter criminals. We begin with a conceptual model in which spatial concentration is assumed to amplify the effect of eyes on the street. Comparative statics predict that spatial concentration should reduce neighborhood property crime and police patrols, savings from which should be capitalized into higher local rent. The model is estimated using rich point-specific data for New York City.

Findings indicate that for a shift from the 25th to the 75th percentile neighborhood based on retail spatial concentration, property crime decreases by 9.4%, police stops are reduced by 12.1%, and retail rent increases by at least 9.6%. These estimates are robust to an extensive set of controls that describe attributes of small (0.2 by 0.2 square mile) neighborhoods across NYC. Our estimates are also large enough to be important for local policy makers and business establishments. This is especially so in that the primary threats to identification likely cause our models to underestimate the effect of spatial concentration on crime and rent capitalization.

Additional exercises shed light on mechanisms. The first is based on comparisons of crime and police stops in daytime versus nighttime and also pre-pandemic versus COVID-19 lockdown. Crowds are greatly diminished at night and during lockdown. Visibility is also reduced at night but not during the lockdown. Evidence based on these and other features of the models suggest that crowding and visibility both contribute to eyes-on-the-street. In an alternate exercise, we create separate measures of spatial concentration based on employment – which targets crowding – and proximity of storefronts – which targets the ability to observe multiple stores from a single location. Once again, evidence suggests that crowding and visibility both enhance eyes on the street.

Taken together, the various models and estimates in our paper suggests that spatial concentration of retail activity enhances eyes-on-the-street, and much more so than concentration of other industries. Our estimates are also large enough to be important. Local government and the private sector can reduce the cost of crime by encouraging retailers to concentrate spatially within their neighborhoods.

References

Anselin, L. and Smirnov, O. (1996). Efficient algorithms for constructing proper higher order spatial lag operators. *Journal of Regional Science*, 36(1), 67–89.

Arzaghi, M. and Henderson, J. (2008). Networking off Madison Avenue. *The Review of Economic Studies*, 75(4), 1011–1038.

Bartel, A.P. (1975). An Analysis of Firm Demand for Protection against Crime. *The Journal of Legal Studies*, 4(2), 443–478.

BLS -U.S. Bureau of Labor and Statistics-.(2018). The Quarterly Census of Employment and Wages. Accessed on April 27th, 2020. https://www.bls.gov/cew/data.htm .

Bowes, D. R. (2007). A Two-Stage Model of the Simultaneous Relationship Between Retail Development and Crime. *Economic Development Quarterly*, 21(1), 79–90.

Brooks, L. (2008). Volunteering to Be Taxed: Business Improvement Districts and the Extra-Governmental Provision of Public Safety. *Journal of Public Economics*, 92(1-2), 388-406.

Browning, C.R.; Byron, R.; Calder, C.A.; Krivo, L.J.; Kwan, M.P.; Lee, J.Y. and Peterson, R.D. (2010). Commercial Density, Residential Concentration, and Crime: Land Use Patterns and Violence in Neighborhood Context. *Journal of Research in Crime and Delinquency*, 47(3), 329–357.

Browning, C.R. and Jackson, A.L. (2013). The Social Ecology of Public Space: Active Streets and Violent Crime in Urban Neighborhoods. *Criminology*, 51, 1009–1043.

Carr, J.B and Doleac, J.L. (2018). Keep the Kids Inside? Juvenile Curfews and Urban Gun Violence. *The Review of Economics and Statistics*, 100(4), 609–618.

Chang, T.Y. and Jacobson, M. (2017). Going to pot? The impact of dispensary closures on crime. *Journal of Urban Economics*, 100, 120–136.

Clapp, J.M.; Ross, S.R. and Zhou, T. (2019). Retail Agglomeration and Competition Externalities: Evidence from Openings and Closings of Multiline Department Stores in the U.S. *Journal of Business & Economic Statistics*, 37(1), 81–96.

Clotfelter, C.T. (1977). Public Services, Private Substitutes, and the Demand for Protection against Crime. *The American Economic Review*, 67(5), 867–877.

Cook, P.J. (1986). The Demand and Supply of Criminal Opportunities. *Crime and Justice*, 7, 1-27.

Cook, P.J. and MacDonald, J. (2011). Public Safety Through Private Action: An Economic Assessment of BIDs. *The Economic Journal*, 121(552), 445-462.

Deloitte Inc. (2018). "Global Powers of Retailing 2018 Transformative change, reinvigorated commerce", https://www2.deloitte.com/content/dam/Deloitte/at/Documents/about-deloitte/global-powers-of-retailing-2018.pdf .

Ellen, I.G.; Lacoe, J. and Sharygin, C.A. (2013). Do foreclosures cause crime? *Journal of Urban Economics*, 74, 59–70.

Evans, D. N.; Maragh, C. L. and Porter, J. R. (2014). What do We Know About NYC's Stop and Frisk Program? A Spatial and Statistical Analysis. *Advances in Social Sciences Research Journal*, 1(2), 129–144.

Faggio, G. (2021). The impact of Business Improvement Districts on crime. Working Paper.

Ferrandino, J.A. (2018). The effectiveness and equity of NYPD stop and frisk policy, 2003–2014. *Journal of Crime and Justice*, 41(2), 119–135.

Freedman, M. and E. Owens (2016). Your Friends and Neighbors: Localized Economic Development, Inequality, and Criminal Behavior. *Review of Economics and Statistics*, 98(2), 233–253.

Gelman, A.; Fagan, J. and Kiss, A. (2007). An Analysis of the New York City Police Department's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias. *Journal of the American Statistical Association*, 102(479), 813–823,

Getis, A. and Ord, J.K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 2 (3), 189–206.

Gould, E.D.; Pashigian, B.P. and Prendergast, C.G. (2005). Contracts, Externalities, and Incentives in Shopping Malls. *The Review of Economics and Statistics*, 87(3), 411–422.

Greenbaum, R. T., and Tita, G. E. (2004). The Impact of Violence Surges on Neighbourhood Business Activity. *Urban Studies*, 41(13), 2495–2514.

Groff, E. R., and Lockwood, B. (2014). Criminogenic Facilities and Crime across Street Segments in Philadelphia: Uncovering Evidence about the Spatial Extent of Facility Influence. *Journal of Research in Crime and Delinquency*, 51(3), 277–314.

Hakim, S. and Shachmurove, Y. (1996). Spatial and Temporal Patterns of Commercial Burglaries: The Evidence Examined. *The American Journal of Economics and Sociology*, 55(4), 443–456.

Hanink, P. (2013). Don't Trust the Police: Stop Question Frisk, Compstat, and the High Cost of Statistical over-Reliance in the NYPD. *Journal of the Institute of Justice and International Studies*, 13, 99–114.

Hipp, J.R. (2016). General Theory of Spatial Crime Patterns. *Criminology*, 54(4), 653–679.

Hoyt, L. M. (2005). Do Business Improvement District Organizations Make a Difference?: Crime In and Around Commercial Areas in Philadelphia. *Journal of Planning Education and Research*, 25(2), 185–199.

Huq, A. Z. (2017). The consequences of disparate policing: Evaluating stop and frisk as modality of urban policing. *Minnesota Law Review*, 101(6), 2397–2480.

Jacobs, J. (1961). The Death and Life of Great American Cities. New York: Vintage Books.

Johansen, B.O. and Nilssen, T. (2016), The Economics of Retailing Formats. *Journal of Industrial Economics*, 64(1), 109–134.

Koster, H.R.A.; van Ommeren, J. and Rietveld, P. (2014), Agglomeration Economies and Productivity: A Structural Estimation Approach Using Commercial Rents. *Economica*, 81(321), 63–85.

Koster, H.R.A.; Pasidis, I. and van Ommeren, J. (2019). Shopping externalities and retail concentration: Evidence from dutch shopping streets. *Journal or Urban Economics*, 114, 103194.

Lee, M.R. and Alshalan, A. (2005). Geographic Variation in Property Crime Rates: A Test of Opportunity Theory. *Journal of Crime and Justice*, 28(2), 101–127.

Lens, M.C. and Meltzer, R. (2016), Is Crime Bad for Business? Crime and Commercial Property Values In New York City. *Journal of Regional Science*, 56(3), 442–470.

Linden, L. and Rockoff, J.E. (2008). Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. *American Economic Review*, 98(3), 1103–1127.

MacDonald, J. and Braga, A.A. (2019). Did Post-*Floyd et al.* Reforms Reduce Racial Disparities in NYPD Stop, Question, and Frisk Practices? An Exploratory Analysis Using External and Internal Benchmarks, *Justice Quarterly*, 36 (5), 954–983.

MacDonald, J.; Fagan, J. and Geller, A. (2016). The Effects of Local Police Surges on Crime and Arrests in New York City. *PLoS ONE*, 11 (6): e0157223.

McMillen, D.; Sarmiento-Barbieri, I. and Singh, R. (2019). Do more eyes on the street reduce Crime? Evidence from Chicago's safe passage program. *Journal of Urban Economics*, 110, 1–25.

NRF-National Retail Federation-. (2019). National Retail Security Survey. Accessed on March 29th, 2021. https://nrf.com/research/national-retail-security-survey-2019

NYCC-New York City Council (2017). Report to the Committee on Finance and the Committee on Public Safety on the Fiscal 2018 Executive Budget for New York Police Department. Accessed on March 29th, 2021. http://council.nyc.gov/budget/wp-content/uploads/sites/54/2017/03/056-NYPD-exec-1.pdf

Ord, J.K. and Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, 27 (4), 286–306.

Owens, E. (2020). The Economics of Policing. In D. Marcotte and K. Zimmermann (Eds.), *Economics of Risky Behavior*. Springer.

Owens, E. (2019). Policy Essay: An Economic Approach to De-Policing. *Criminology and Public Policy*, 18(1), 77–80.

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.

Pope, J. (2008). Fear of Crime and Housing Prices: Household Reactions to Sex Offender Registries. *Journal of Urban Economics*, 64, 601–614.

Ridgeway, G. (2007). Analysis of Racial Disparities in the New York Police Department's Stop, Question, and Frisk Practices. *RAND Corporation*, TR-534-NYCPF, 2007. As of February 24, 2021: https://www.rand.org/pubs/technical_reports/TR534.html

Robinson, P.M. (1988). Root-n-consistent semiparametric regression. Econometrica 56, 931–954.

Rosenfeld, R. and Fornango, R. (2017). The Relationship Between Crime and Stop, Question, and Frisk Rates in New York City Neighborhoods. *Justice Quarterly*, 34(6), 931–951.

Rosenthal, S.S. and Ross, A. (2010). Violent crime, entrepreneurship, and cities. *Journal of Urban Economics*, 67(1), 135–149.

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, S.S.; Strange, W.C. and Urrego, J.A. (2021). Are city centers losing their appeal? Commercial real estate, urban spatial structure, and COVID-19. *Journal of Urban Economics: Insights*, forthcoming.

Sivitanidou, R. (1995). Urban Spatial Variations in Office-Commercial Rents: The Role of Spatial Amenities and Commercial Zoning. *Journal of Urban Economics*, 38(1), 23–49.

Stucky, T.D. and Ottensmann, J.R. (2009). Land Use and Violent Crime. *Criminology*, 47(4), 1223–1264.

Tillyer, M. S., and Walter, R. J. (2019). Busy Businesses and Busy Contexts: The Distribution and Sources of Crime at Commercial Properties. *Journal of Research in Crime and Delinquency*, 56(6), 816–850.

Urrego, J.A. (2021). Perceptions of Police Bias and the Location of Minority-Owned Establishments. Working paper.

Verardi, V. and Debarsy, N. (2012). Robinson's square root of N consistent semiparametric regression estimator in Stata. *The Stata Journal*, 12(4), 726–735.

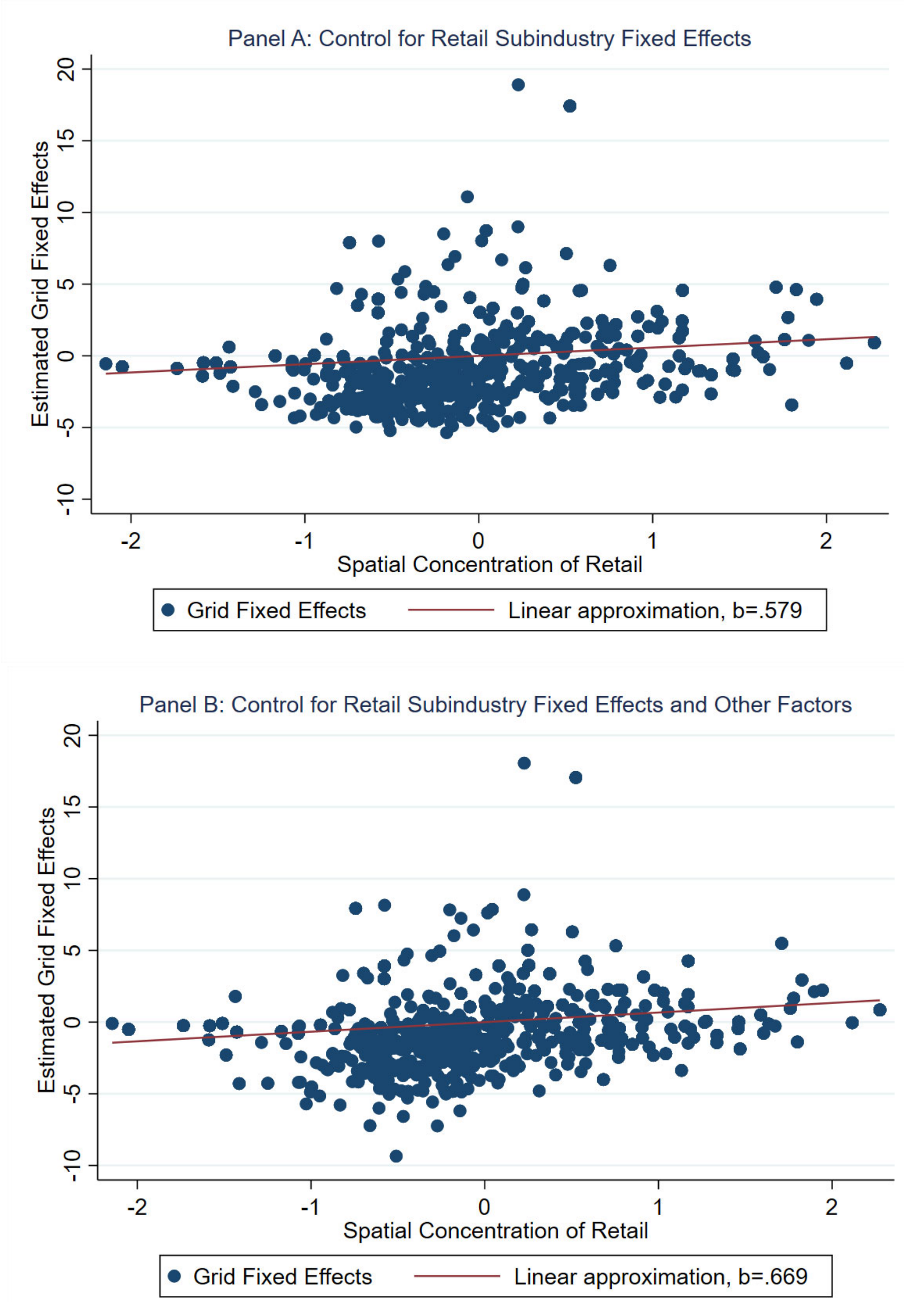
Weisburd, D.; Telep, C.W. and Lawton, B.A. (2014). Could Innovations in Policing have Contributed to the New York City Crime Drop even in a Period of Declining Police Strength? The Case of Stop, Question and Frisk as a Hot Spots Policing Strategy. *Justice Quarterly*, 31 (1), 129–153.

Weisburd, D.; Wooditch, A.; Weisburd, S. and Yang, S.M. (2016). Do Stop, Question, and Frisk Practices Deter Crime? Evidence at microunits of space and time. *Criminology & Public Policy*, 15 (1), 31–56.

Weterings, A. (2014). What Makes Firms Leave the Neighbourhood? *Urban Studies*, 51(8), 1613–1633.

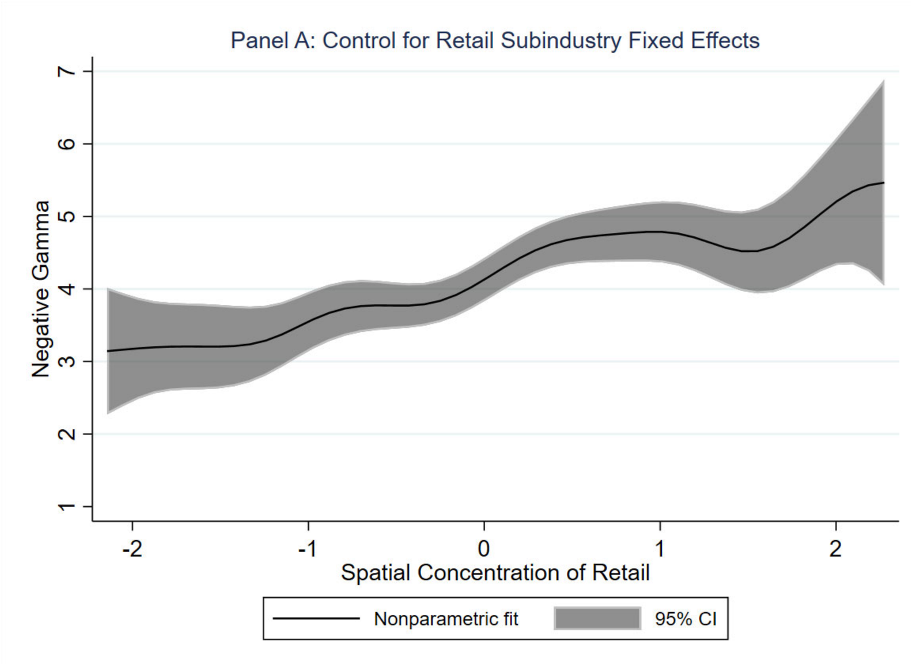
Wooditch, A. and Weisburd, D. (2016). Using Space–Time Analysis to Evaluate Criminal Justice Programs: An Application to Stop-Question-Frisk Practices. *Journal of Quantitative Criminology*, 32 (2), 191–213.

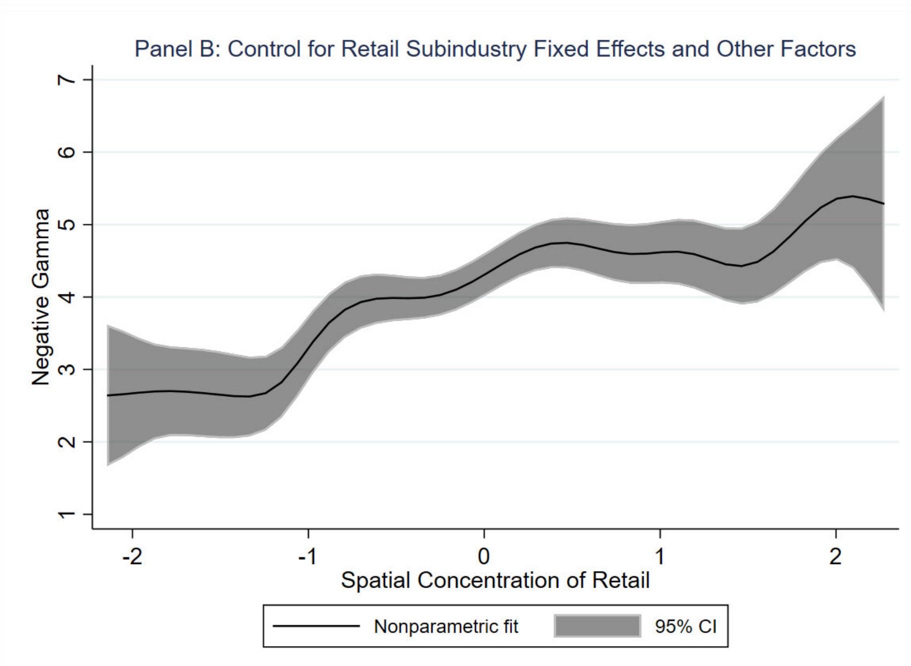
Figure 1: Neighborhood square fixed effect estimates of - $\gamma(\tilde{G})$ plotted against $\tilde{G}$[a]



[a] Panel A plots the estimated grid square fixed effects from expression (2.14). Panel B does the same but includes controls for other neighborhood and establishment attributes as in (2.15a) including spatial concentration of other industries, visits to POI, share of residential units, number of trees and dummies for different age categories of retail establishments.

Figure 2: Partial linear model estimates of the effect of $\tilde{G}$ on - $\gamma(\tilde{G})$[a]





[a] Panel A plots the estimated $-\gamma(\tilde{G})$ function from a partial linear model where non-inventory cost per dollar of sales depends, in addition to γ, on subindustry fixed effects. Panel B adds other controls linearly to the specification in Panel A. Those are spatial concentration of other industries, visits to POI, share of residential units, number of trees and fixed effect for different age categories of retail establishments.

Table 1: Summary Statistics[a]

	Obs.	Mean	Std. Dev.	p25	p50	p75
Panel A: At the grid square level[b]						
Property crime	3,506	29.06	54.78	6	14	31
Grand Larceny & Burglary	3,506	9.76	16.37	2	5	11
Petit Larceny	3,506	18.40	40.45	3	8	18
Theft of Services & Fraud	3,506	0.90	2.13	0	0	1
Robbery	3,506	2.57	3.48	0	1	4
Auto Theft	3,506	0.95	1.18	0	1	1
Police Stops	3,506	2.08	4.37	0	0	2
At Least 1 Police Stop (Stops > 0)	3,506	0.49	0.50	0	0	1
Share of Employment						
Retail	3,506	0.19	0.12	0.09	0.17	0.27
Finance	3,506	0.07	0.07	0.03	0.06	0.09
Manufacture	3,506	0.05	0.08	0.01	0.02	0.05
Services	3,506	0.46	0.18	0.34	0.46	0.58
Spatial Concentration						
$\tilde{G}_{Retail,Emp}$[d]	3,506	0.00	1.00	-0.63	-0.11	0.47
$\tilde{G}_{Finance,Emp}$[d]	3,506	0.00	1.00	-0.54	-0.21	0.26
$\tilde{G}_{Manf,Emp}$[d]	3,506	0.00	1.00	-0.67	-0.24	0.40
$\tilde{G}_{Service,Emp}$[d]	3,506	0.00	1.00	-0.52	-0.19	0.23
Number of Trees	3,506	119.25	55.53	80	118	157
Average Age Buildings	3,506	80.26	17.04	70.48	82.44	91.83
Average Building Assessment	3,506	2,116,857	15,500,000	75,979	177,548	665,259
Overlaps Multiple Police Precinct	3,506	0.24	0.43	0	0	0
Share of Residential Units	3,506	0.87	0.20	0.87	0.94	0.97
Neighborhood Sales per Worker	3,506	58,934	60,118	40,379	51,627	65,304
Average Monthly Visitors POI	3,506	179.91	118.88	104.92	148.85	215.43

	Obs.	Mean	Std. Dev.	p25	p50	p75
Panel B: At the establishment level[c]						
Spatial Concentration of Retail[d]	1,596	0.01	0.98	-0.57	0.02	0.65
Retailer Rent per sq foot (psf) per month	1,596	178.02	211.27	61.38	110.98	219.75
Retailer Employment	1,596	7.21	11.49	2	3	8
Retailer Sales	1,596	471,918	1,438,367	89,625	150,000	289,590
Retailer Cost of space / Sales	1,596	2.95	4.03	0.55	1.38	3.59
Retailer Non-inventory cost / Sales[c]	1,596	4.15	4.29	1.45	2.73	5.27
Wholesaler Cost of space / Sales	538	0.66	1.06	0.06	0.21	0.74
Wholesaler Non-inventory cost / Sales[c]	538	1.44	1.41	0.43	0.97	1.89

[a] Crime and police stops data are from the New York Police Department. Crimes lasting more than one day and all crimes that take place in a transportation system (e.g. on the subway) are dropped. Police stops prompted by 911 calls and those for ongoing investigations are dropped. Employment and sales are from Dun & Bradstreet while rent and space leased are from CompStak.
[b] The unit of analysis is a grid cell of 0.2 square miles across the New York City area.
[c] Panel B includes only establishments matched in the CompStak and Dun and Bradstreet files.
[d] Spatial concentration is calculated as described for $\tilde{G}$ in the text.
[e] Non-inventory cost refers to the sum of cost of space and labor cost. Estimated annual earnings from BLS at the NAICS 6 digits code are used to calculate labor cost.

Table 2: Negative Binomial Marginal Effects for Property Crime and Police Stops in 0.2 Mile Grid Squares[a]

	Property crime[b]				Police Stops[c]			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employment								
Log total employment	0.700***	0.752***	0.522***	0.554***	0.394***	0.463***	0.304***	0.349***
	(0.017)	(0.019)	(0.024)	(0.022)	(0.037)	(0.041)	(0.053)	(0.056)
Share of Employment: Retail	2.651***	2.813***	2.706***	2.307***	2.411***	2.640***	2.761***	2.017***
	(0.135)	(0.169)	(0.168)	(0.165)	(0.279)	(0.373)	(0.348)	(0.359)
Share of Employment: Finance	-	-0.166	-0.577**	-0.451**	-	-0.574	-0.650	-0.752
	-	(0.265)	(0.235)	(0.217)	-	(0.626)	(0.585)	(0.608)
Share of Employment: Manufacture	-	-0.265	-0.281	-0.226	-	0.293	0.183	0.027
	-	(0.278)	(0.261)	(0.247)	-	(0.757)	(0.551)	(0.536)
Share of Employment: Services	-	0.739***	-0.063	-0.025	-	0.941***	0.109	0.055
	-	(0.127)	(0.139)	(0.129)	-	(0.293)	(0.299)	(0.318)
Log monthly visits to POI	-	-	-	0.516***	-	-	-	0.894***
	-	-	-	(0.030)	-	-	-	(0.069)
Spatial concentration[d]								
Retail ($\tilde{G}_{Ret,Emp}$)	-0.061***	-0.050***	-0.102***	-0.085***	-0.087**	-0.062	-0.133***	-0.110***
	(0.018)	(0.018)	(0.016)	(0.015)	(0.040)	(0.040)	(0.039)	(0.036)
Finance ($\tilde{G}_{Fin,Emp}$)	-	0.024	0.021	0.005	-	0.081**	0.060	0.042
	-	(0.019)	(0.017)	(0.016)	-	(0.041)	(0.040)	(0.040)
Manufacture ($\tilde{G}_{Manf,Emp}$)	-	-0.033	0.025	0.007	-	-0.037	0.007	-0.020
	-	(0.020)	(0.019)	(0.018)	-	(0.045)	(0.041)	(0.041)
Services ($\tilde{G}_{Serv,Emp}$)	-	-0.120***	-0.008	-0.029	-	-0.239***	-0.108**	-0.137***
	-	(0.020)	(0.021)	(0.019)	-	(0.042)	(0.046)	(0.046)
Other Neighborhood Controls[e]	No	No	6	6	No	No	6	6
Overdispersion Poisson parameter	0.58	0.55	0.47	0.42	2.59	2.51	2.26	2.03
Observations	3,506	3,506	3,506	3,506	3,506	3,506	3,506	3,506

[a] Marginal effects based on the data means are reported. Significance is denoted as: * p<0.1, ** p<0.05, *** p<0.01. Robust standard errors in parentheses.
[b] Property crime includes 2018 incidents of grand larceny, burglary, petit larceny, theft of services, and fraud.
[c] Police stops include discretionary stops pooled from 2016-2018.
[d] Spatial concentration is calculated as described for $\tilde{G}$ in expression (3.3a).
[e] Neighborhood controls include log number of trees, average building age, log of average building assessment, whether the grid square overlaps with more than one police precinct, share of residential building units, and log of neighborhood sales per worker.

Table 3: Adding Additional Neighborhood, Building and Establishment Controls[a,b]

	Property Crime[d]		Police Stops[e]	
	% Retail Employment	$\tilde{G}_{Ret,Emp}$	% Retail Employment	$\tilde{G}_{Ret,Emp}$
	(1)	(2)	(3)	(4)
Root specification (columns 4 & 8 of Table 2)[c]	2.307***	-0.085***	2.017***	-0.110***
	(0.165)	(0.015)	(0.359)	(0.036)
Establishments Characteristics (2 controls)	2.286***	-0.067***	1.887***	-0.084**
	(0.164)	(0.015)	(0.362)	(0.037)
Zoning Restrictions (8 controls)	2.319***	-0.085***	1.203***	-0.107***
	(0.170)	(0.015)	(0.379)	(0.033)
Distance to Landmarks (4 controls)	2.379***	-0.082***	2.145***	-0.110***
	(0.165)	(0.016)	(0.355)	(0.037)
Amenities (9 controls)	2.198***	-0.103***	1.479***	-0.096***
	(0.162)	(0.016)	(0.346)	(0.036)
Irregular Shape & Tax-Exempt Bldgs. (2 controls)	2.323***	-0.083***	1.926***	-0.089**
	(0.162)	(0.015)	(0.353)	(0.037)
All of the above (25 controls)	2.080***	-0.076***	0.762**	-0.070**
	(0.167)	(0.015)	(0.358)	(0.031)

[a] See Appendix A for a description of the additional 26 variables included in this table.
[b]Marginal effects based on the data means are reported. Significance is denoted as: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Robust standard errors in parenthesis.
[c]The root specifications for property crime and police stops are the same as for columns 4 and 8 of Table 2. For each row, additional controls are included as indicated and are measured at the 0.2 mile neighborhood grid level.
[d] Property crime includes 2018 incidents of grand larceny, burglary, petit larceny, theft of services, and fraud.
[e] Police stops include discretionary stops pooled from 2016-2018.

Table 4: Number of Crimes for Alternate Time Periods by Type of Crime[a,b]

	All Hours and Periods		**Daytime**[c]		**Nighttime**[c]		**COVID-19 Lockdown**[d]	
	% Retail Employment	$\tilde{G}_{Ret,Emp}$	% Retail Employment	$\tilde{G}_{Ret,Emp}$	% Retail Employment	$\tilde{G}_{Ret,Emp}$	% Retail Employment	$\tilde{G}_{Ret,Emp}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Property crime	2.307***	-0.085***	2.650***	-0.104***	1.685***	-0.004	2.647***	-0.034
	(0.165)	(0.015)	(0.182)	(0.018)	(0.198)	(0.018)	(0.339)	(0.032)
Grand Larceny & Burglary	1.792***	-0.053***	2.034***	-0.072***	1.766***	0.003	2.493***	-0.070*
	(0.159)	(0.014)	(0.179)	(0.016)	(0.223)	(0.021)	(0.424)	(0.040)
Petit Larceny	2.650***	-0.104***	3.050***	-0.121***	1.733***	-0.008	2.810***	-0.009
	(0.189)	(0.018)	(0.208)	(0.021)	(0.238)	(0.022)	(0.418)	(0.040)
Theft of Services & Fraud	1.819***	-0.087***	2.204***	-0.125***	1.298***	-0.007	2.641*	0.001
	(0.319)	(0.030)	(0.404)	(0.042)	(0.459)	(0.039)	(1.474)	(0.279)
Robbery	2.330***	-0.144***	2.217***	-0.175***	2.265***	-0.142***	3.352***	-0.177***
	(0.220)	(0.021)	(0.254)	(0.027)	(0.265)	(0.028)	(0.669)	(0.058)
Auto Theft	0.833***	-0.057**	0.586*	-0.042	1.152***	-0.086***	1.018	0.008
	(0.212)	(0.023)	(0.333)	(0.034)	(0.320)	(0.033)	(0.854)	(0.087)

[a] Marginal effects based on the data means are reported. Significance is denoted as: $p<0.1$, ** $p<0.05$, *** $p<0.01$. Robust standard errors in parenthesis. Spatial concentration is calculated as described for $\tilde{G}$ in expression (3.3a).
[b] Every pair of coefficients, share of retail employment and spatial concentration of retail, is obtained from the preferred specification in Table 2 column (4). Note that the coefficients for property crime during all hours are identical to those in Table 2 column (4). The sample for all models is 3,507 grid cells (0.2 square miles) in the NYC area. Crime data refers to incidents in 2018.
[c] Daytime hours include crimes between 10 am and 6 pm. Nighttime hours include crimes between 10 pm and 5 am.
[d] COVID-19 lockdown refers to the first two weeks of the NYC lockdown, March 22nd to April 5th of 2020.

Table 5: Alternate Measures of Spatial Concentration[a]

Panel A: Correlation between alternate measures of spatial concentration

	$\tilde{G}_{Ret,Emp}$	$\tilde{G}_{Ret,Stores}$	$\tilde{G}_{Ret,Sales}$
$\tilde{G}_{Ret,Emp}$	1.0	-	-
$\tilde{G}_{Ret,Stores}$	0.15	1.0	-
$\tilde{G}_{Ret,Sales}$	0.54	0.14	1.0

Panel B: Crime and police stop regressions

	Property Crime[b]				**Police Stops[c]**			
	% Retail Employment	$\tilde{G}_{Ret,Emp}$ [d]	$\tilde{G}_{Ret,Stores}$ [d]	$\tilde{G}_{Ret,Sales}$ [d]	% Retail Employment	$\tilde{G}_{Ret,Emp}$ [d]	$\tilde{G}_{Ret,Stores}$ [d]	$\tilde{G}_{Ret,Sales}$ [d]
Regression Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All Hours/Periods	2.646***	-0.084***	-0.122***	0.002	1.955***	-0.104**	-0.046	0.035
	(0.178)	(0.016)	(0.021)	(0.016)	(0.390)	(0.041)	(0.051)	(0.037)
Daytime[e]	2.974***	-0.101***	-0.120***	-0.001	1.671***	-0.112**	-0.085	0.043
	(0.198)	(0.019)	(0.025)	(0.018)	(0.458)	(0.052)	(0.061)	(0.045)
Nighttime[e]	2.109***	-0.021	-0.141***	0.028	1.727***	-0.102**	0.040	0.015
	(0.210)	(0.020)	(0.024)	(0.020)	(0.436)	(0.047)	(0.059)	(0.045)
COVID-19 Lockdown[f]	3.096***	-0.078**	-0.142***	0.052	-	-	-	-
	(0.355)	(0.036)	(0.048)	(0.036)	-	-	-	-

[a] Marginal effects based on the data means are reported. Significance is denoted as: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Robust standard errors in parenthesis. The sample for all models is 3,506 grid cells (0.2 square miles) in the NYC area.

[b] Property crime includes 2018 incidents of grand larceny, burglary, petit larceny, theft of services, and fraud.

[c] Police stops include discretionary stops pooled from 2016-2018. During COVID-19 lockdown police stops were likely motivated by other confounding factors such as compliance of the stay at home executive orders, for that reason results for police stops during lockdown are not estimated.

[d] Spatial concentration of retail employment, $\tilde{G}_{\text{Ret,Emp}}$, is calculated as described for $\tilde{G}$ in expression (3.3). Similar structure is applied to the spatial concentration of retail sales, $\tilde{G}_{Ret,Sales}$, where establishment's sales is used instead of employment in (3.3a). Spatial concentration of storefronts, $\tilde{G}_{\text{Ret,Stores}}=\sum_e \omega_{ie}(d_{ie})\,/n_i(d_{ie})$, where n_i is the number of nearby establishments within a given distance, d_{ie}, and ω_{ie} is defined as in (3.2).

[e] Daytime hours include crimes between 10 am and 6 pm. Nighttime hours include crimes between 10 pm and 5 am.

[f] COVID-19 lockdown refers to the first two weeks of the NYC lockdown, which refers to March 22nd to April 5th of 2020.

Table 6: Cost of Inventory Lost to Crime Per Dollar Sold[a]

	(1)	(2)	(3)	(4)	(5)
Panel A: Retailer cost of space/$ sold					
Number of Workers/Sales	71,574***	70,893***	71,129***	71,828***	37,983***
	(6,196)	(6,116)	(6.116)	(6,432)	(7,180)
Spatial Concentration of Retail: $\tilde{G}_{Ret,Emp}$	0.373***	0.413***	0.436***	0.384***	0.353***
	(0.106)	(0.116)	(0.115)	(0.123)	(0.120)
Observations	1,596	1,596	1,596	1,596	1,596
R2	0.47	0.48	0.48	0.50	0.52
Panel B: Retailer non-inventory cost/$ sold					
Spatial Concentration of Retail: $\tilde{G}_{Ret,Emp}$	0.457***	0.517***	0.533***	0.377***	0.324***
	(0.119)	(0.132)	(0.132)	(0.141)	(0.125)
Observations	1,596	1,596	1,596	1,596	1,596
R2	0.538	0.544	0.544	0.567	0.63
Panel C: Wholesaler cost of space/$ sold					
Number of Workers/Sales	84,814***	82,089***	81,671***	81,802***	70,405***
	(8,355)	(8,707)	(8,864)	(9,410)	(9,734)
Spatial Concentration of Retail: $\tilde{G}_{Ret,Emp}$	0.168***	0.185***	0.177***	0.162***	0.132**
	(0.053)	(0.056)	(0.054)	(0.058)	(0.058)
Observations	538	538	538	538	538
R2	0.471	0.49	0.49	0.529	0.549
Panel D: Wholesaler non-inventory cost/$ sold					
Spatial Concentration of Retail $\tilde{G}_{Ret,Emp}$	0.247***	0.288***	0.253***	0.244***	0.209**
	(0.073)	(0.078)	(0.074)	(0.080)	(0.083)
Observations	538	538	538	538	538
R2	0.579	0.59	0.593	0.626	0.685
Neigh Controls (Table 2) and Corner location	Yes	Yes	Yes	Yes	Yes
SIC 2 Fixed Effects[b]	Yes	Yes	Yes	Yes	Yes
Lease Execution Year Fixed Effects[b]	Yes	Yes	Yes	Yes	Yes
Spatial Concentration other industries	-	Yes	Yes	Yes	Yes
POI Visits	-	-	Yes	Yes	Yes
Police Precinct Fixed Effects[c]	-	-	-	Yes	Yes
Establishment Age Categories[d]	-	-	-	-	Yes

[a] Significance is denoted as: * p<0.1, ** p<0.05, *** p<0.01. Robust standard errors are in parenthesis. Rent data comes from CompStak and employment and sales from Dun & Bradstreet. Establishments located above the 25th floor are dropped as are those in the top 1% of the distribution of sales, employment, space leased, and employment divided by sales. Non-inventory cost is the sum of expenditures on space and labor. Establishments with non-inventory cost in the top 5% are dropped.
[b] Models in Panels A and B include 24 lease transaction year fixed effects and 8 SIC2 fixed effects. Panels C and D include 19 lease transaction year fixed effects and 2 SIC2 fixed effects.
[c] In column (4), 69 police precinct fixed effects are present for the retail regressions and 34 for the wholesale regressions.
[d] Column (5) includes fixed effects for establishment age categories: less than 2 years, 2 to 5 years, 6 to 10 years, 10 to 25 years, and more than 25 years in business.

Appendix A: Data Sources and Variable Construction

A.1 Data sources and access

A.1.1 Proprietary data

Our two primary data sources are proprietary. This includes establishment level Dun & Bradstreet data obtained from the Syracuse University library. Syracuse University has a site license with Merge Intellect that makes these data available to members of Syracuse University. Many other universities have similar licensing arrangements. CompStak provides information on commercial leases. These data were obtained by purchasing an individual user license from CompStak Inc.[34] The licensing arrangements for both data sources do not permit us to share these data. We are able, however, to share our programs that clean the data and run our regressions. A detailed list of all of the data sources used in the paper is below.

A.1.2 Data Sources

1995 Street Tree Census. Department of Parks and Recreation, NYC Open Data. Downloaded on December 21, 2019. https://data.cityofnewyork.us/Environment/1995-Street-Tree-Census/kyad-zm4j

2005 Street Tree Census. Department of Parks and Recreation, NYC Open Data. Downloaded on December 21, 2019. https://data.cityofnewyork.us/Environment/2005-Street-Tree-Census/29bw-z7pj

2015 Street Tree Census. Department of Parks and Recreation, NYC Open Data. Downloaded on December 16, 2019. https://data.cityofnewyork.us/Environment/2015-Street-Tree-Census-Tree-Data/uvpi-gqnh

311 Service Requests from 2010 to Present. Department of Information Technology and Telecommunications (DITT), NYC Open Data. Downloaded on March 8, 2020. https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9

Commercial Leases – Compstak. Downloaded on August 1, 2021. https://compstak.com/.

Establishments Directory – Dun & Bradstreet. Downloaded on October 13, 2018 and updated on February 27, 2019. https://www.mergentintellect.com/index.php/search/index.

In-Service Alarm Box Locations. Fire Department, NYC Open Data. Downloaded on December 16, 2019. https://data.cityofnewyork.us/Public-Safety/In-Service-Alarm-Box-Locations/v57i-gtxb

[34] The CompStak data are populated by leasing agents who provide CompStak information on leases they have executed in exchange for being able to draw other leases from the CompStak database that may be helping in guiding a new client.

National Registry of Historic Places. NY State Parks, Recreation and Historic Preservation. Downloaded on March 8, 2020. https://www.nps.gov/subjects/nationalregister/index.htm

New York City Community Air Survey (NYCCAS). Downloaded on March 9, 2020. https://www1.nyc.gov/site/doh/data/data-sets/air-quality-nyc-community-air-survey.page

New York Police Department (NYPD) Complaint Data History, NYC Open Data. Downloaded on September 13, 2020. https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i.

New York Police Department (NYPD) Stop, Question and Frisk Data. Downloaded on October 7, 2020. https://www1.nyc.gov/site/nypd/stats/reports-analysis/stopfrisk.page.

NYC Condom Availability Program - HIV condom distribution locations. Department of Health and Mental Hygiene, NYC Open Data. Downloaded December 20, 2019. https://data.cityofnewyork.us/Health/NYC-Condom-Availability-Program-HIV-condom-distrib/4kpn-sezh

Open Space – Parks. Department of Parks and Recreation, NYC Open Data. Downloaded on March 6, 2020. https://data.cityofnewyork.us/Recreation/Open-Space-Parks-/g84h-jbjm

PLUTO and MapPLUTO version 18v2. Department of City Planning. Downloaded on January 1, 2020. https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-pluto-mappluto.page

Rodent inspections. Department of Health and Mental Hygiene (DOHMH), NYC Open Data. Downloaded on March 8, 2020. https://data.cityofnewyork.us/Health/Rodent-Inspection/p937-wjvj

SafeGraph Cellphone Data. Downloaded on March 2, 2020. https://www.safegraph.com/.

Subway Entrances. Metropolitan Transportation Authority (MTA), NYC Open Data. Downloaded on March 8, 2020. https://data.cityofnewyork.us/Transportation/Subway-Entrances/drex-xx56

A.2 Matching D&B and CompStak establishment level data

For the rent analysis in Table 6 we used establishment level matched records from D&B and CompStak. Our match routine took advantage of street addresses and establishment names which are reported in both files and utilize two similarity indexes. One uses n-gram with three characters. This divides a sentence into sequences of three characters and calculates how many of those three-character words match.[35] The second index calculates how many changes have to be made on one name to make it identical to the other, normalized by the difference in the two names' length. We define a "perfect" match

[35] We use the Stata program matchit to apply this algorithm and apply log weights to the three-character words based on frequency which minimizes false positive matches when encountering words like Inc or Cor.

between a lease in CompStak and an establishment in D&B if both observations belong to the same building and the similarity score between the two is the highest among all potential establishment matches in at least one of the indexes. “Good” matches are defined when the similarity scores in both indexes are the highest based on establishment name but we cannot definitively confirm the records correspond to the same building (based on street address). “Good” matches represent 23% of our estimating sample in Table 6. We also estimated Table 6 restricting the sample to “perfect” matches; results were robust.

A.3 Data sources for each measure in the paper

Table A1 below provides a complete listing of each variable and its source used in different parts of the paper.

Table A1: Data Sources by Variable

Neighborhood Controls	Variable	Definition	Source
Baseline Controls	Distribution and level of employment	Total employment and employment in each selected industry (SIC)	Dun & Bradstreet
	Spatial Concentration Measures (*G*)	Spatial *G* for selected industries based on employment and sales	Dun & Bradstreet
	Monthly visits to POI	Average monthly visits to each POI in a grid square based on cellphone data	SafeGraph
	Number of Trees	Total number of trees on the street based on 2015 Street Tree Census	Department of Parks and Recreation
	Average Building Age	Average age of buildings across the grid square.	MapPLUTO
	Average Building Assessed Value	Average assessed value of the building in the grid square	MapPLUTO
	Overlapping Police Precincts	Grid square overlaps multiple police precincts	MapPLUTO
	Share Residential Units	Share of all units in the grid square that are residential	MapPLUTO
	Neighborhood Sales per Worker	Total sales of single-site establishments in the grid square over total employment	MapPLUTO
Establishment Characteristics	Neighborhood Average Market Risk	Marketing Pre-screen Ranking: predicts the likelihood of a company to pay bills on-time. Ranges from 1 to 5, being 1 indicates most likely to pay	Dun & Bradstreet
	Neighborhood Avg. Establishment Age	Establishment Age = 2019 – Founding Year	Dun & Bradstreet
Zoning Restrictions	Share Special District	Share of lots in the grid located in special purpose districts.	MapPLUTO
	Share Commercial allowed in Residential	Share of lots in the grid that are allow for commercial overlay within a residential zoning district	MapPLUTO
	Share Multiple Zoning	Share of lots in the grid that are between multiple zoning features.	MapPLUTO
	Average Density Allowed by Residential Zoning	For lots in a residential district they are assigned a code from R1-1 to R10H, where the higher the number immediately after R the higher the density or intensity of land use permitted. We calculate the average of that number across the grid.	MapPLUTO
	Average Density Allowed by Commercial Zoning	For lots in a commercial district they are assigned a code from C1-6 to C8-4, where the higher the number immediately after C the higher the density or intensity of land use permitted. We calculate the average of that number across the grid.	MapPLUTO
	Share buildings with height restriction	Share of lots in the grid that are in a limited height district	MapPLUTO
	Average residential FAR	Maximum allowable residential floor area ratio across the grid	MapPLUTO
	Average commercial FAR	Maximum allowable commercial floor area ratio across the grid	MapPLUTO
Distance to Landmarks	Distance Central Park	Distance between the grid centroid and Central Park	Department of Parks and Recreation
	Distance Nearest Park	Distance from grid centroid to nearest park	Department of Parks and Recreation
	Distance Nearest Subway	Distance from grid centroid to nearest subway entrance (0 if entrance inside grid)	Metropolitan Transportation Authority
	# Subway Entrances	Number of subway entrances inside the grid square	Metropolitan Transportation Authority

Table A1: Data Sources by Variable (continued)

Neighborhood Controls	**Variable**	**Definition**	**Source**
Amenities	Average PM 2.5	Annual average fine particulate matter < 2.5 microns (2018), 300 mt resolution	Community Air Survey Air Pollution
	Ln(Reported rat problems)	Total 2018 rodent inspections that resulted in active rat signs.	Department of Health and Mental Hygiene
	Ln(Failed rodent inspections)	Total 2018 rodent inspections that did not pass the inspection.	Department of Health and Mental Hygiene
	Ln(Complaints about traffic lights)	311 Complaints (requests) related to traffic signal condition	Department of Information Technology and Telecommunications (DITT)
	Ln(Complaints about street lights)	311 Complaints (requests) related to street light condition	DITT
	Newly planted trees 2005-2015	Difference between 2005 and 2015 Tree Census	Department of Parks and Recreation
	Historic Places and Landmarks	Historic places registered before 2018 to the National Register of Historic Places	NY State Parks, Recreation and Historic Preservation
	Public alarm boxes on the street	Fire alarm boxes in the grid: includes Emergency Reporting System (ERS) and Box Alarm Reporting System (BARS)	Fire Department
	Active Sites: HIV testing and condom distribution locations	Active venues distributing free safer sex products under the NYC Condom Availability Program – HIV.	Department of Health and Mental Hygiene
Building with Irregular shape and Tax exemptions	Share of buildings with irregular shape	Share of lots in the grid that have an irregular shape	MapPLUTO
	Share of buildings that are tax exempt	Share of lots in the grid that have at least 20% of their assessment value exempt of property tax.	MapPLUTO

Appendix B: Alternate Definition of Neighborhoods and Spatial Concentration

As a robustness check, we also define neighborhoods and spatial concentration using a very different design than in the main body of the paper. Specifically, working with the same 0.2 by 0.2 mile grid squares, we define a neighborhood as 3x3 configuration of grid squares. The square in the center is designated as the target block grid square that identifies the neighborhood. Each grid square serves as a target for one neighborhood and for that reason, neighborhoods overlap as we move across NYC.

All neighborhood level variables in Table 2 are measured for each neighborhood based on this new definition of a neighborhood. Spatial concentration of employment within a 9-block neighborhood is measured as the sum of squared employment shares across the nine member grid squares as below:

$$G_i = \sum_{j=1}^{9} \left(\frac{E_{ij}}{\sum_{j=1}^{9} E_{ij}} \right)^2 \quad . \tag{B.1}$$

In this expression, if all employment is concentrated in a single grid square, G_i takes on a maximum value of 1, whereas if employment is spread equally across all 9 grid squares, G_i converges to its minimum value of 1/9.

Estimates based on this alternate definition of a neighborhood and of spatial concentration are in Table B1 and mirror the structure of Table 2. The important thing to note is that results are quite similar to those in Table 2.[36] This includes that higher spatial concentration of retail activity decreases the incidence of property crime and also decreases the number of police stops. As a further robustness check, we used the same design as above but restricted our estimating sample to non-overlapping neighborhoods. This did not change the results just described.

[36] Sample size for the 9-grid square neighborhoods is larger than for the individual 0.2 square mile neighborhoods used in the main body of the paper. This is because a number of the 0.2 square mile neighborhoods have zero employment in at least one of the industries highlighted in our control measures, retail, services, finance, and manufacturing. Because it is not possible to measure spatial concentration for missing industries, neighborhoods for which all four industries are not present were dropped from the sample. This occurred less frequently for the larger neighborhoods considered here in Appendix B.

Table B1: Negative Binomial Estimation for Property Crime and Police Stops at the 9-Grid Community Level [a,b]

	Property Crime[c]				Police Stops[d]			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Community level control								
Log total employment	0.802***	0.838***	0.715***	0.665***	0.526***	0.561***	0.487***	0.378***
	(0.010)	(0.012)	(0.016)	(0.014)	(0.027)	(0.030)	(0.039)	(0.036)
Share of Employment: Retail	3.696***	4.430***	3.519***	2.879***	3.564***	4.094***	3.025***	1.643***
	(0.132)	(0.181)	(0.173)	(0.163)	(0.306)	(0.406)	(0.405)	(0.388)
Share of Employment: Finance	-	-1.909***	-1.536***	-1.002***	-	-3.828***	-3.344***	-3.230***
	-	(0.243)	(0.217)	(0.204)	-	(0.982)	(0.707)	(0.556)
Share of Employment: Manufacture	-	-0.430	0.485	0.412	-	-1.916***	-0.176	-0.650
	-	(0.406)	(0.379)	(0.321)	-	(0.605)	(0.642)	(0.637)
Share of Employment: Services	-	1.105***	0.311**	0.363***	-	1.031***	0.006	-0.101
	-	(0.124)	(0.126)	(0.114)	-	(0.245)	(0.290)	(0.269)
Log monthly visits to POI	-	-	-	0.611***	-	-	-	1.284***
	-	-	-	(0.023)	-	-	-	(0.055)
Spatial Concentration $\tilde{G}$[e]								
Retail	-1.542***	-1.399***	-1.501***	-1.058***	-2.947***	-3.116***	-2.730***	-2.235***
	(0.132)	(0.140)	(0.131)	(0.125)	(0.383)	(0.401)	(0.418)	(0.322)
Finance	-	0.274**	-0.070	-0.179	-	1.037***	0.823***	0.693***
	-	(0.137)	(0.121)	(0.112)	-	(0.315)	(0.281)	(0.234)
Manufacture	-	0.244***	0.175**	0.036	-	-0.031	-0.108	-0.357***
	-	(0.073)	(0.070)	(0.062)	-	(0.145)	(0.149)	(0.132)
Services	-	-0.618***	-0.531***	-0.647***	-	-0.810**	-0.703**	-1.120***
	-	(0.157)	(0.152)	(0.145)	-	(0.329)	(0.301)	(0.248)
Other 9-Grids Controls	No	No	6	6	No	No	6	6
Overdispersion Poisson parameter	0.341	0.314	0.265	0.217	1.462	1.402	1.309	1.115
Observations	5,122	5,122	5,122	5,122	5,122	5,122	5,122	5,122

[a] Marginal effects based on the data means are reported. Significance is denoted as: * p<0.1, ** p<0.05, *** p<0.01. Robust standard errors in parenthesis.
[b] The analysis unit is 0.6 miles overlapping grid squares, defined as the usual target 0.2 mi grid plus its eight neighboring cells. Each 0.6 mi unit has nine 0.2 mi grids squares.
[c] Property crimes include grand larceny, burglary, petit larceny, theft of services, and fraud (reference year 2018). Model specifications follow those in Table 2.
[d] Only self-initiated police stops between 2016-2018 are included. Stops initiated by 911 calls and ongoing investigations are removed.
[e] Spatial concentration is calculated as described in expression (B.1).