Agglomeration and Green Commercial Development within Cities

Evan Perry*

March 1, 2022

Abstract

Buildings account for nearly 30% of all US greenhouse gas emissions. Energy-efficient, green buildings have the potential to reduce these emissions – a vital component of a comprehensive approach to climate policy. This research explores the distribution of green buildings within cities by developing a model that connects firms' agglomeration economies and their decision to occupy green real estate. Using data from the two leading green building certification programs in the US, I estimate a reduced-form of the model and find that its main predictions hold empirically. The results imply that green building incentives targeted at specific neighborhoods may be more effective than broader incentive programs.

1 Introduction & Motivation

Climate change is a grim reality. In its sixth assessment report, the UN's Intergovernmental Panel on Climate Change finds "it is unequivocal that human influence has warmed the atmosphere, ocean and land," and further that "observed warming is driven by emissions from human activities" (IPCC, 2021). Under a high emissions, do-nothing scenario, estimates suggest climate change induced deaths will rise to 73 deaths annually in a population of 100,000 by 2100 (Climate Impact Lab, 2020). This means that without significant reductions in emissions, we can expect climate

*Coe College, eaperry19@coe.edu. This research benefited from the invaluable comments of the students in the Spellman Fellows Program and members of the Coe College Economics faculty: Ryan Baranowski, Rick Eichhorn, and Chelsea Lensing. I am especially indebted to my faculty research advisor, Drew Westberg. Generous financial support from the Spellman Fund made this research possible.

change moralities will have a magnitude similar to other leading causes of death worldwide by the end of the century, like infectious disease (74 deaths per 100,000) and all cancers (125 deaths per 100,000). Mitigating climate change damages by reducing greenhouse gas emissions is essential.

In the US, buildings are a significant part of this problem. Buildings like our homes, offices, and schools account for nearly 30% of all greenhouse gas emissions in the US (Resources for the Future, 2021). Some of these emissions occur indirectly through electricity consumption (18% of all emissions) while other emissions are released directly at the source (12% of all emissions). These direct emissions often come from heating systems that burn natural gas. While investing in clean, renewable electricity can reduce indirect building emissions, decarbonization of the US electric grid will not slow direct building emissions. Consequently, reducing direct building emissions requires action specific to the structures we live and work in.

One strategy to reduce both direct and indirect emissions is to construct and remodel green buildings. For present purposes, we define a green building as any building with an energy-efficiency or sustainability certification that shows it has smaller carbon footprint than otherwise comparable buildings. Although the exact criteria for green building certification vary between programs, all programs emphasize that the buildings they certify have lower energy costs and emissions. Historically, energy-efficiency has been among the most politically feasible approaches to climate policy. Not only can energy-efficiency improvements reduce greenhouse gas emissions, but they allow consumers to save money and promote energy independence. As a result, energy-efficiency improvements attract broad political support. Current policy has recognized the necessity of green buildings and the potential benefit of public policy to support this need. The November 2021 infrastructure spending plan contains billions of dollars to support green infrastructure projects. This includes additional funding for the Weatherization Assistance Program, which makes energy-efficient retrofits to homes and buildings, mostly in impoverished areas.

Unfortunately, energy-efficiency improvements are not without their own issues. One of the primary concerns with subsidies of energy-efficient goods is heterogeneity in how consumers and firms respond to energy-efficient technologies (Allcott and Greenstone, 2012). There is evidence that some consumers and firms may face barriers to adopting energy-efficient technology, possibly due to imperfect information or other investment inefficiencies (Gerarden, Newell, and Stavins, 2017). Blanket subsides cannot differentiate between those consumers and firms that would adopt an energy-efficient technology with and without the extra incentive. If Firm A is hypersensitive to environmental policy it might over-invest in energy-efficient technologies. More concerning, if Firm B is unresponsive to energy-efficient upgrades, it might fail to adopt an energy-efficient technology even when doing so is profitable. An untargeted approach induces Firm A to continue purchasing new energy-efficient technologies more so than Firm B, where emissions reductions would be achievable at a lower cost. Further, subsidies may not be effective if there are other misaligned incentives. Targeted subsides and improvements have a smaller risk of being welfare decreasing than untargeted programs, but in practice it is often difficult to find suitable mechanisms that policymakers can use to accurately target incentives towards consumers and firms.

In the context of energy-efficient, green buildings, location may provide a useful tool for targeting these incentives. If firms that would not otherwise "go green" choose to locate in the same neighborhoods, then local policymakers might prevent welfare losses by targeting incentives to specific neighborhoods.

Although there are a number of reasons similarly green-minded firms may locate in the same neighborhoods, this paper focuses on agglomeration economies as a possible explanation. Specifically, I analyze whether or not agglomeration economies can help explain why some neighborhoods have so many more green commercial buildings than others. Section two reviews related literature on the spatial distribution of green buildings within cities and identifies a connection in the literature between agglomeration economies and the propensity of firms to occupy green buildings. Section three builds a spatial equilibrium model to explore the implications of this connection on the location of green buildings within cities and the location of ecologically responsive firms. In sections four and five, I pull data to estimate a reduced-form of the earlier model, and section six concludes.

2 Literature Review

Much of the economic research on green buildings has focused on premiums. Energy-efficient homes and offices should cost less to operate, giving prospective property owners an incentive to pay more upfront for energy-efficient buildings. Walls, Gerarden, Palmer, and Bak (2017) find evidence that energy efficient features within homes translate into higher selling prices with premiums over similar homes that are roughly equivalent to their energy cost savings. Both Eichholtz, Kok, and Quigley (2010) and Wiley, Benefield, and Johnson (2010) find evidence that green-certified office buildings earn rental premiums. In these papers, the researchers match green buildings to commercial real estate data and run hedonic regressions to determine the green building premium for office buildings. Eichholtz, Kok, and Quigley (2010) further note that these premiums as a percent of total rent tend to be somewhat smaller in "prime locations." Likely this is because land costs are larger in prime locations, making the green premium smaller relative to the total rent.

Other evidence suggests that there may be more to the relationship between green buildings and prime locations. Braun and Bienert (2015) find evidence that green buildings appear most frequently in prime locations, but note that over this pattern appears to be less apparent as green building adoption becomes more frequent in many cities. They map the location of green buildings in the 103 most populous metropolitan statistical areas. Then they assume a monocentric city with rents that are monotonically decreasing as the distance from the city center increases, as in the bid-rent curves developed by Alonso (1964). While green buildings are not exclusively found in prime locations – defined by proximity to the city center – green buildings are more concentrated in these locations. This finding matches evidence in Kaza, Lester, and Rodriguez (2013), which uses nearest neighbor methods to find that green buildings often develop in clusters.

Kaza, Lester, and Rodriguez (2013) and Braun and Bienert (2015) both suggest similar reasons green commercial development may be more common in downtown areas, dense with workers. Some of these conjectures rely entirely on developers' incentives. For instance, the additional expense of constructing an energy efficient building on an expensive piece of land may seem negligible to some developers. More convincing is the explanation that the location of green buildings is tied to firm sorting. Fitting with the sorting theory, Eichholtz, Kok, and Quigley (2016) find evidence that firms in service-based industries disproportionately choose to lease green buildings. These firms in the service industry also tend to locate in dense, downtown neighborhoods.

While these stylized facts and their suggested explanations are intriguing, the literature has stopped short of producing and testing a formalized model of the spatial nature of green commercial development. To capture the firm sorting and selection between green buildings and locations, such a model will need to address (1) why firms choose to locate in dense, urban environments, and (2) why firms choose to adopt eco-friendly technologies and practices. I do so using agglomeration economies and ecological responsiveness respectively.

Agglomeration economies provide an economic explanation for the clustering behavior of firms in densely populated areas. With high land costs and congestion in these environments, there must be a compensating factor for firms. Glaeser (2010) explains this through agglomeration economies:

> Agglomeration economies are the benefits that come when firms and people locate near one another together in cities and industrial clusters. These benefits all ultimately come from transport cost savings: the only real difference between a nearby firm and one across the continent is that it is easier to connect with a neighbor. Of course, transportation costs must be interpreted broadly, and they include the difficulties in exchanging goods, people, and ideas.

Transportation of goods is easier now than at any other point in human history, yet firms and people both continue to crowd together into cities. Modern agglomeration economies rely heavily on explanations including benefits from the pooling together of specialized labor forces and knowledge spillovers, where new ideas and innovations act as proximity-based positive externalities.

Firms are heterogeneous in their benefits from dense environments – that is, firms have heterogeneous agglomeration economies. A tech start-up may benefit more from the ability to connect with its specialized labor force and other potential clients than a cereal manufacturer. Gaubert (2018) incorporates the heterogeneity of firms' agglomeration economies and their innate ability into a model of city formation. Firms sort into different cities based on agglomeration economies and ability. Firms in sectors with high agglomeration economies sort into denser cities where they can benefit more than lower agglomeration firms. Gaubert develops and estimates this structural model using French firm-level data. The empirical results confirm the prediction that firms with high sectoral agglomeration economies sort into denser cities.

Firms are also heterogeneous in their ability to benefit from occupying green real estate. In studying why some firms adopt environmentally friendly technologies and practices, Bansal and Roth (2000) popularized the concept of ecological responsiveness. Differences in accrued benefits from environmentalism make some firms more receptive to go green than others. Eichholtz, Kok, and Quigley (2016) examine the ecological responsiveness of several industries in the context of green buildings. Their results suggest that a primary factor that drives firms in an industry to occupy green buildings is the need for a highly skilled workforce. Nyborg and Zhang (2013) find evidence that firms with high environmental-social-governance (ESG) ratings can negotiate and pay lower wages than similar firms with lower ESG ratings. This may explain why firms and industries dependent on highly skilled labor would choose to occupy greener buildings.

This research connects these related, but currently disparate, literatures. With a common driver of agglomeration economies and ecological responsiveness, I demonstrate the clustering behavior of firms, and consequently the clustering of green commercial real estate in worker-dense neighborhoods.

3 The Model

3.1 Model Environment & Overview

Here I create a model that focuses on two interrelated decisions that firms face: (i) what neighborhood to locate in and (ii) whether to lease green or brown real estate in that neighborhood. To create the needed variation in firms' decisions, firms will be heterogeneous in their exposure to agglomeration economies and the production benefit they receive from leasing a green building.

Assume a city comprised of many neighborhoods. Each neighborhood has an identical quantity of land available for commercial development, $\bar{\ell}$. The worker and housing aspects of neighborhood formation are outside the scope of the model, so let each neighborhood have a fixed and unique number of workers, N. Ex ante, neighborhoods differ only in N, so the number of workers is enough to characterize a neighborhood. The set of all neighborhoods is the city. We proceed by constructing the model's two sets of actors: firms and developers.

3.2 Firms

3.2.1 Firm Heterogeneity

Assume all firms produce a homogeneous, composite good. Firms can trade this good freely to consumers outside their neighborhood and city. Prices for the good are set by national markets and firms take this price as exogenous. We normalize prices so the price of the good is one. Firms differ in two dimensions: (i) their benefits from agglomeration economies, and (ii) their benefits from occupying a green building.

First, consider the benefits firms receive from agglomeration economies. In this model specification, agglomeration is the productive benefit that a firm receives from the proximity of workers to one another. We model this by allowing the firm's agglomerative benefit to be a function of the local density of workers $\frac{N}{\ell}$. Not all firms benefit from worker density identically, so let firm i's unique sectoral agglomeration parameter be α_i. Then denote the agglomeration productivity multiplier of firm i in neighborhood N as $\psi(\alpha_i, \frac{N}{\ell})$. Let $\frac{\partial\psi}{\partial\alpha} > 0$, so firms with a higher sectoral agglomeration parameter α receive a greater agglomerative benefit than other firms in the same neighborhood. Similarly, let $\frac{\partial\psi}{\partial(N/\ell)} > 0$, so identical firms receive greater agglomerative benefits when they locate in denser neighborhoods. I leave the functional form of ψ unspecified for now.

Second, consider the benefits firms receive from occupying a green building. The building a firm inhabits has design d. This design can either be energy efficient and green, g, or inefficient and brown, b (i.e. $d \in \{g, b\}$). In general, firms with energy-efficient real estate are more productive, but the magnitude of this productivity differential varies from firm to firm. Let θ_j be a parameter that represents the magnitude of firm j's benefits from occupying a green building. We can think of this as the *ecological responsiveness* parameter. Firms with a high θ will benefit more from "going green" than firms with a low θ. Denote the green productivity multiplier of firm type j with design d as $\lambda(d, \theta_j)$. We assume that a firm is always more productive when it leases energy-efficient real estate, such that $\lambda(g, \theta_j) > \lambda(b, \theta_j) = 1$ for any firm j. Here, we set$\lambda(b, \theta_j) = 1$ to require that firms can only receive a green productivity boost when they actually occupy a green building. This green benefit is increasing in the firm's ecological responsiveness, $\frac{\partial\lambda}{\partial\theta} > 0$. This formulation of ecological responsiveness is analogous to the formulation of agglomeration economies.

I do not specify what determines either α_i nor θ_j, and thus leave out any explicit link between firms' agglomeration economies and their green benefits. Let $f(\theta_j \mid \alpha_i)$ be the probability density distribution of ecological responsiveness types θ_j, given the agglomeration type α_i. Denote the corresponding cumulative distribution function $F(\,\cdot \mid \alpha_i)$. Based on the suggested link between agglomeration economies and ecological responsiveness (Braun and Bienert, 2015; Eichholtz, Kok, and Quigley, 2016), I invoke the following assumption: for $\alpha_1 < \alpha_2$, $F(\,\cdot \mid \alpha_2)$ first-order stochastically dominates $F(\,\cdot\,|\,\alpha_1)$.[1] This assumption means that firms with higher agglomeration economies will have lower probabilities of having low ecological responsiveness types and higher probabilities of having high ecological responsiveness types.

3.2.2 The Firm's Problem

Firm heterogeneity enters the firm's problem through its production function. Both agglomeration economies $\psi(\alpha_i, \frac{N}{\ell})$ and green benefits $\lambda(g, \theta_j)$ enter into a type ij firm's production function as scalars. Then a type ij firm leasing a building with design d in neighborhood N will produce output $y(L,R)$ according to the Cobb-Douglas production function,

$$y(L,R) = A\ \psi(\alpha_i, N)\ \lambda(d, \theta_j)\ L^{\beta} R^{\gamma} \bar{K}^{1-\beta-\gamma}. \tag{1}$$

Firms have three inputs: labor L in number of workers, commercial real estate R in square feet, and capital K. We are primarily concerned with the adjustment of wages and commercial real estate prices across neighborhoods, so we will take capital to be fixed at a common quantity for all firms. Apart from the agglomeration and green benefit scalars, there is also a city-specific constant A. This represents the raw productivity of firms and cities, and acts as a catchall for other sources of productivity. Then $\psi(\alpha_i, \frac{N}{\ell})$ represents productivity improvements from agglomeration, $\lambda(d, \theta_j)$ represents potential productivity improvements from green building adoption, and A represents all other productive environmental improvements. For ease of notation, let $\mathcal{A} = A\ \psi(\alpha_i, N)\ \lambda(d, \theta_j)$, so the production function can be rewritten as

$$y(L,R) = \mathcal{A}\ L^{\beta} R^{\gamma} \bar{K}^{1-\beta-\gamma}. \tag{2}$$

The firm's objective is to choose its inputs, the design of its real estate, and a neighborhood to maximize its profits. Firms face the market-determined wage W, price of green real estate per square foot p_g, and price of brown real estate per square foot p_b. Then the firm's problem is

[1] A distribution F first-order stochastically dominates a distribution G if and only if $F(x) \leq G(x)$ for all x in the support of the distributions.

summarized as,

$$\max_{L,R,d,N}\{\pi(L,R,d,N)\} = \max_{L,R,d,N}\left\{\mathcal{A}\, L^{\beta} R^{\gamma} \bar{K}^{1-\beta-\gamma} - WL - p_d R - k_{ij}\right\} \tag{3}$$

Given its type, the market wage, and the market price for green and brown commercial real estate, firms choose their inputs of labor L and commercial real estate R, the design of this real estate d, and the neighborhood to locate in N.

Firms also pay a fixed entrance cost that varies by type, k_{ij}. We assume that k_{ij} is increasing in α_i, so firms that experience higher agglomeration economies pay a higher price to enter the market. Suppose we did not make this assumption. Then high agglomeration firms would experience their agglomeration productivity gain without any additional cost. This would imply that the firm type with the highest α_i could outbid all other firm types in every neighborhood and we would end up with a single type of firm. The assumption of increasing entrance costs in α_i is crucial for allowing several firm types to exists, and thus crucial to creating firm sorting. This assumption also offers interesting interpretations of what high α_i firms look like. Through k_{ij} we see that not only do high α_i firms experience more agglomeration through processes such as knowledge spillovers, they also experience more market pressure at entry. Industries with these firms should be difficult to enter.

If we take the firm's choice of green or brown real estate and neighborhood as fixed for now, then the firm's problem yields the two first-order conditions:

$$W = \beta \mathcal{A} L^{\beta-1} R^{\gamma} \bar{K}^{1-\beta-\gamma} \tag{4}$$

$$p_d = \gamma \mathcal{A} L^{\beta} R^{\gamma-1} \bar{K}^{1-\beta-\gamma}. \tag{5}$$

The first of these first-order conditions says that the market wage must be equal to the marginal value of labor. Similarly, the second of these first-order conditions says the leasing price of commercial real estate with design d is equal to the marginal value of this real estate. We can solve this system of equations to obtain the optimal quantity of labor L^* and real estate of the given design R^*.

$$L^* = \left[\beta^{1-\gamma}\gamma^{\gamma}\mathcal{A}W^{\gamma-1}p_d^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} \tag{6}$$

$$R^* = \left[\beta^{\beta}\gamma^{1-\beta}\mathcal{A}W^{-\beta}p_d^{\beta-1}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} \tag{7}$$

Given the market price of real estate with design d per square foot, equation (6) is the firm's demand curve for labor. Analogously, given the market wage, equation (7) is the firm's demand curve for real estate with design d. If we substitute (6) and (7) in the firm's profit function, we can rewrite the firm's profits as a function of W and p_d,

$$\pi(W, p_d : d, N) = \Phi\left[\mathcal{A}W^{-\beta}p_d^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} - k_{ij} \tag{8}$$

where Φ is a constant term made up of β and γ.[2]

3.3 Developers

To complete the commercial real estate market, we need a developer agent to construct green and brown buildings for firms. The developer agent purchases land from the residual claimant and chooses the height of green and brown buildings on this land. Assume that developers can enter and exit the market freely, but must bid on all commercial land in a neighborhood. This implies that each neighborhood will have a single developer who earns zero-profits. The price of land will adjust for any neighborhood such that if any developer could earn a positive profit, another developer could freely enter, submit a marginally higher bid on the land, and earn a positive profit as well. Thus in equilibrium, the price of land in each neighborhood will guarantee that the developer will earn zero-profits. The zero-profit condition of the developer is also its spatial equilibrium condition.

There are two building designs the developer can choose between: an energy-efficient, green design, g, and a standard, brown design, b. The developer produces commercial real estate with a design d using height h_d and land ℓ_d, so that the production function for commercial real estate with design d is $R_d = h_d\ell_d$.[3] The developer faces a material cost of construction $c_d h_d^{\delta}\ell_d$, where

[2] A full derivation of the firm's problem is given in Appendix A.

[3] This production function and the corresponding cost function follow closely from Glaeser (2008).

c_d is a design-specific, exogenous variable reflecting the construction costs per square foot of a single-story building. In this formulation, we assume $\delta > 1$ such that developers face additional costs or "frictions" when they try to build up. Let $c_g > c_b$ so that green buildings have higher material costs than their otherwise identical brown versions.

Although a single developer agent constructs all commercial buildings in a neighborhood, it can build both green buildings and brown buildings in the same neighborhood. The goal of the developer is to maximize the sum of its profits from green and brown development subject to its land use constraint, $\ell_g + \ell_b = \bar{\ell}$. It chooses the height and land for green and brown buildings. Then the developer's problem is,

$$\max_{h_g, h_b, \ell_g, \ell_b} p_g h_g \ell_g - c_g h_g^\delta \ell_g + p_b h_b \ell_b - c_b h_b^\delta \ell_b - p_\ell \bar{\ell} \quad \text{s.t.} \quad \bar{\ell} = \ell_g + \ell_b \tag{9}$$

where p_g and p_b are the price per area unit of green and brown commercial real estate respectively. This constrained optimization yields the following three first-order conditions:

$$h_g = \left(\frac{p_g}{\delta c_g}\right)^{\frac{1}{\delta-1}} \tag{10}$$

$$h_b = \left(\frac{p_b}{\delta c_b}\right)^{\frac{1}{\delta-1}} \tag{11}$$

$$p_g h_g - p_b h_b = c_g h_g^\delta - c_b h_b^\delta. \tag{12}$$

Conditions (10) and (11) describe the optimal height of green and brown buildings respectively given their equilibrium prices in the commercial real estate market. We can consider these functions to represent the supply of building height in the city.

The developer's profit maximization problem does not determine the optimal allocation of land between green and brown construction. Its third first-order condition illustrates this. In (12) we see that on a per unit of land basis, the additional revenues a green building earns must equal the additional cost incurred. Then for the same quantity of land, profits from green buildings must be equal to the profits from brown buildings. Absent of the preferences of firms, developers are indifferent between all divisions of land between green and brown construction. To find the developer's optimal division will require us to incorporate firms' adoption decisions.

Green buildings earn a proportional premium over brown buildings per area unit. Taking the optimal height for green and brown buildings in (10) and (11) and substituting these into (12) yields this premium.

$$\frac{p_g}{p_b} = \left(\frac{c_g}{c_b}\right)^{\frac{1}{\delta}} \tag{13}$$

Green buildings are proportionally more expensive per square foot than brown buildings as $c_g > c_b$. This proportion is determined by the ratio of the input costs and the upward friction the developer faces when building taller buildings, δ. The premium is unaffected by the relative scarcity of green buildings in a neighborhood. Once again, because the developer is indifferent between building a green and brown building on any given piece of land, then the premium reflects differences in input costs only.

We assume that the developer earns zero-profits. This is accomplished through the price of land: if any developer earned a positive economic profit, then another builder could profitably enter by bidding slightly higher on the land. Thus the price of land will adjust so that no developer can earn a positive profit. Then the price of land is such that,

$$p_\ell = \left(p_g h_g - c_g h_g^\delta\right)\left(\frac{\ell_g}{\bar{\ell}}\right) + \left(p_b h_b - c_b h_b^\delta\right)\left(1 - \frac{\ell_g}{\bar{\ell}}\right). \tag{14}$$

This is necessary for our assumptions of free-entry of developers, as well as spatial equilibrium. When all developers earn zero-profits, then no developers can gain by deviating. Equation (14) says that the price of land will be equal to a weighted average of profits without land costs from green and brown development. The weights on green and brown profits are the proportion of land in the neighborhood allocated for the given design. Using (12) we can rewrite (14) as

$$p_\ell = p_g h_g - c_g h_g^\delta = p_b h_b - c_b h_b^\delta. \tag{15}$$

This further shows that profits from green development and profits from brown development will both be zero. The price of land is equal to the value the develop can create on the land, either from green or brown development.

3.4 The Adoption Decision

Firms choose the design of commercial real estate that maximizes their profits. With just two designs, green and brown, then a type ij firm will choose to buy green real estate if and only if[4]

$$\pi(W, p_g : g, N) > \pi(W, p_b : b, N). \tag{16}$$

Making use of equation (8), this simplifies to

$$\lambda(g, \theta_j) > \left(\frac{p_g}{p_b}\right)^{\gamma} \tag{17}$$

This says that for a firm to purchase green real estate, the productive benefit of a green building $\lambda(g, \theta_j)$ must outweigh the green premium, scaled by γ.

The previous condition must hold for the firm to purchase green real estate, but what must hold for the developer to build green real estate? Recall that the developer does not have a particular preference for building green or brown buildings; commercial real estate prices are set such that the additional cost of building a green building exactly offsets the premium it earns. We found this premium in equation (13). Then if we substitute in (13) into (17), the adoption condition now represents both firms and the developer and is made of only exogenous variables:

$$\lambda(g, \theta_j) > \left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta}} \tag{18}$$

With our assumption that $\lambda(d, \theta_j)$ is continuous and unbounded in θ_j and that $\lambda(g, \theta_j)$ is strictly increasing in θ_j, the adoption condition (equation (18)) implies the existence of a threshold type θ^*. That is, there must exist a θ^* for which

$$\lambda(g, \theta^*) = \left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta}}. \tag{19}$$

Then for any two firms with θ_1 and θ_2 such that $\theta_1 < \theta^* < \theta_2$, firm 1 will choose brown real estate and firm 2 will choose green real estate. Given the existence of threshold types, I choose

[4] Here we assume that profits from a green building must be strictly greater than profits from a brown building for firms to adopt.

to simplify the model so firms can have just two possible ecological responsiveness types. To condense notation, denote these with the benefits λ_g and λ_b, where the λ_g is the green benefit to an ecologically responsive firm that leases a green building in equilibrium and $\lambda_b = 1$ is the green benefit to an ecologically unresponsive firm that leases a brown building in equilibrium

3.5 Equilibrium Characterization

Before solving for an equilibrium, we need to first consider what an equilibrium looks like in this model. Any equilibrium must solve both the firm's problem and the developer's problem, which implies the following conditions to be satisfied in equilibrium:

(i) Given the market wage and price of green and brown real estate, firms maximize their profits

(ii) Given market demand for green and brown real estate, developers maximize their profits

(iii) All developers earn zero profits

(iv) All firms earn zero profits

(v) In each neighborhood, the labor market clears

(vi) In each neighborhood, the commercial real estate market clears

Like any conventional economic model, equilibrium requires that firms and developers choose the optimal combination of inputs in order to maximize their profits. For the firm, this will mean finding the profit-maximizing combination of labor and commercial real estate, and deciding whether that commercial real estate will be green or brown. For the developer, this will mean finding the profit-maximizing height and land for both green and brown construction. Conditions (i) and (ii) say this, and specify that firms will pay the market-determined wage and price for their chosen design of real estate, and developers will charge a set premium on green real estate based on their input costs.

Conditions (iii) and (iv) are the spatial equilibrium conditions. Spatial equilibrium requires that agents have no incentive to locate to another neighborhood. If firms could earn a higher profit in another neighborhood than the neighborhood they currently occupy, then the free movement of firms implies that we would not be in spatial equilibrium as the agents are not maximizing over the possible choices of location.

Equilibrium also requires that the labor and commercial real estate markets clear. There is a fixed supply of workers in each neighborhood N, and the market wage will adjust such that each worker is employed and firms can only hire workers in the neighborhood they locate in. Additionally, both commercial real estate markets need to clear, so that the supply of both green and brown real estate meets demand. It is worth noting that the product market does not need to clear, as firms can freely export their production. Likewise, consumers can freely import the composite good.

3.6 Equilibrium

For simplicity, assume the agglomeration type is homogeneous within a neighborhood after sorting occurs. We proceed by solving for the neighborhood N where firms with this homogeneous agglomeration type will choose to locate. Along with this, we solve for the equilibrium market outcomes, including the amount of green real estate.

Let M be the number of firms in the neighborhood. For the single agglomeration type under consideration α, let μ be the probability that firms with this agglomeration type have a green benefit λ_g and choose to occupy a green building. Note that μ is increasing in α, as higher agglomeration type firms will be more likely to go green. With identical agglomeration types within the neighborhood and two green benefit types, we have just two types of firms within the neighborhood. In solving in the model, let the agglomeration benefits take the functional form adapted from Gaubert (2018)

$$\psi\left(\alpha, \frac{N}{\bar{\ell}}\right) = \left(\frac{N}{\bar{\ell}}\right)^{\alpha}. \tag{20}$$

The conditions and equations derived earlier ensure that firms and developers maximize their profits. Additionally, the price of land adjusts such that developers earn zero profits. This is implicit in the derivation of the premium on green buildings. These satisfy conditions (i), (ii), and (iii) in the equilibrium characterization.

We proceed with the three remaining conditions, but save the details of the derivation for Appendix A. There are four equations we need to solve to yield a solution for our model: the firm's zero-profit condition, the labor market clearing condition, and the commercial real estate market clearing condition (both supply and demand). These make up a system of four equations in four unknowns, N, W, p_g, and R_g.

The system of equations has a unique solution. This unique solution offers two main insights. First, the equilibrium neighborhood size N is increasing in α. That is, firms with higher agglomerative benefits will choose denser neighborhoods. This is consistent with related models of heterogeneous agglomeration type sorting (Gaubert, 2018). Second, the proportion of green real commercial real estate is higher in denser neighborhoods. This point follows immediately from the first. High agglomeration firms are more likely to have a high ecological responsiveness, so the areas with the densest populations and highest agglomeration type firms, will also have the most green commercial real estate as a proportion of all commercial real estate.

We can still find this second relationship of interest without looking at the specifics of this equilibrium. Consider market demand for green and brown real estate, where there are μM green firms and $(1-\mu)M$ brown firms:

$$R_g = \mu M \left[\beta^\beta \gamma^{1-\beta} A \left(\frac{N}{\bar{\ell}} \right)^\alpha \lambda_g W^\beta p_g^{\beta-1} \right]^{\frac{1}{1-\beta-\gamma}} \bar{K} \tag{21}$$

$$R_b = (1-\mu) M \left[\beta^\beta \gamma^{1-\beta} A \left(\frac{N}{\bar{\ell}} \right)^\alpha \lambda_b W^\beta p_b^{\beta-1} \right]^{\frac{1}{1-\beta-\gamma}} \bar{K}. \tag{22}$$

We are interested in how the ratio of green to brown real estate will vary from neighborhood to neighborhood. This ratio is

$$\frac{R_g}{R_b} = \left(\frac{\mu}{1-\mu} \right) \left(\frac{\lambda_g}{\lambda_b} \right)^{\frac{1}{1-\beta-\gamma}} \left(\frac{p_g}{p_b} \right)^{\frac{\beta-1}{1-\beta-\gamma}}. \tag{23}$$

Now, equation (13) found the equilibrium relationship between p_g and p_b, so substituting in the green building premium we have

$$\frac{R_g}{R_b} = \left(\frac{\mu}{1-\mu} \right) \left(\frac{\lambda_g}{\lambda_b} \right)^{\frac{1}{1-\beta-\gamma}} \left(\frac{c_g}{c_b} \right)^{\frac{\beta-1}{\delta(1-\beta-\gamma)}}. \tag{24}$$

This equations might appear to show that the ratio of green to brown real estate is fixed. Recall though that μ is implicitly an increasing function of the degree of agglomeration economies α. Then in denser neighborhoods, where the agglomeration effect is strongest and high agglomeration firms locate, the ratio of green real estate to brown real estate is higher because this ratio is increasing in μ. Not only is there more green real estate in denser neighborhoods, but green real estate makes up a larger share of all commercial real estate in these neighborhoods.

4 Data

I do not have firm-level data and cannot calibrate or structurally estimate the model. Instead, I use a reduced-form estimation of the model.

4.1 Green Building Data

Green building data come from the Energy Star program and the Leadership in Energy and Environmental Design (LEED) program. These two programs are the primary green building certification programs in the US. Although there are other green building certification programs, researchers have used the Energy Star and LEED data in the past as a sufficiently complete registry of US green buildings (Eichholtz, Kok, and Quigley, 2010; Braun and Bienert, 2015).

The Energy Star program, ran by the US Department of Energy (DOE) and the US Environmental Protection Agency (EPA), evaluates, certifies, and labels durables that meet certain energy efficiency standards. The agencies eventually began applying the Energy Star label to residential and commercial buildings, extending the label to homes in 1995 and to office buildings in 1998. For a commercial building to earn the Energy Star certification, it must rank in the top 25% of comparable buildings for energy efficiency at the time of certification. Data on these buildings are available through the Registry of Energy Star Certified Buildings and Plants. This publicly-available registry contains over 30,000 commercial buildings from across the US and lists their addresses, their size, and a description of their use (e.g. school, grocery store, office).

Figure 1 displays the frequency of different buildings within the Energy Star registry. Although there are over 2 million Energy Star certified houses in the US, these houses are not available in the public database. Over half of all Energy Star certified buildings in the data are either schools or offices.

The US Green Building Council (USGBC) runs the LEED program. While the Energy Star certification focuses solely on energy efficiency, the LEED certification is more inclusive of other aspects of sustainable building, such as water efficiency and eco-friendly building materials. During construction or remodeling, trained green building raters award points to buildings based on the program's certification criteria. Buildings that meet a point threshold earn LEED certification. The USGBC maintains a publicly-available database of the buildings that earn the LEED certi-

Figure 1: Energy Star Building Classifications

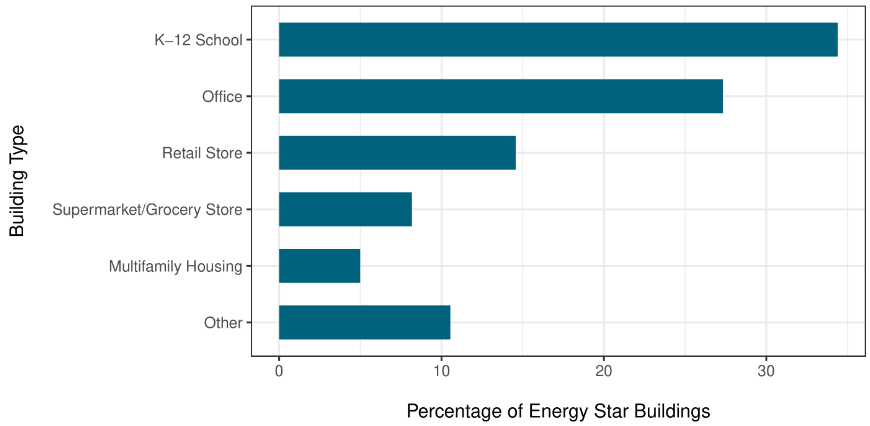
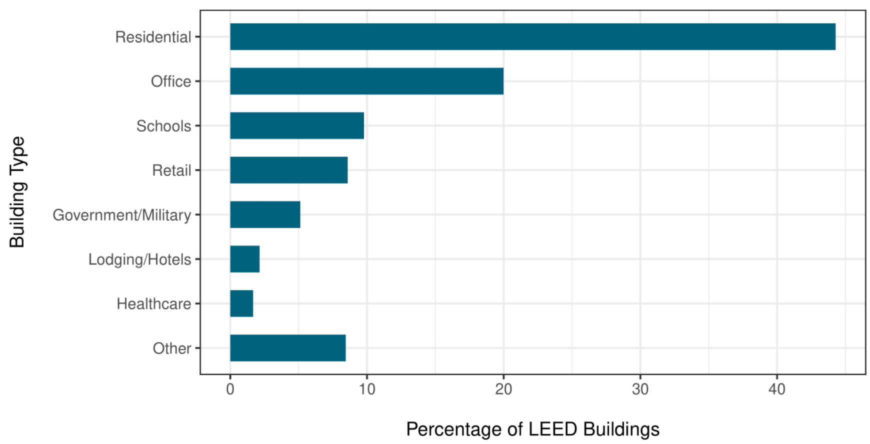


Figure 2: LEED Building Classifications



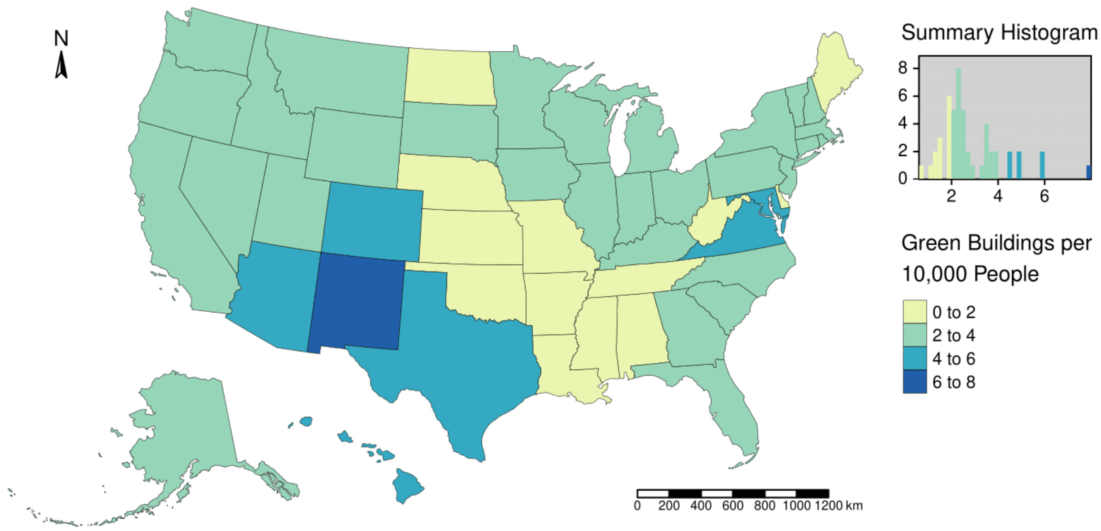
fication. The fields in this database are similar to those in the Registry of Energy Star Certified Buildings and Plants. The LEED registry contains over 100,000 buildings with US addresses.

The Energy Star certification is more popular with residential construction, and the LEED certification is more popular with commercial construction. However, the LEED registry includes certified single-family homes, where the Energy Star registry does not. Despite its commercial popularity, the LEED registry contains far more residential buildings than the Energy Star database. Offices are relatively more common than schools in the LEED registry than in the Energy Star registry. Retail buildings are again frequently found in this green building database.

I clean and merge the two datasets. From there I use a geocoding service to match each address to a latitude and longitude.[5] Using census tract shapefiles and the coordinates of the buildings,

[5]Specifically, I use the HERE Developer geocoding service through the `tidygeocoder` package in R. This popular

Figure 3: Green Buildings per Capita by State



I match each green building to a census tract.[6] Then I aggregate the green building data at the census tract level, forming counts for the number of commercial green buildings and total commercial green real estate area (in square feet).

Figure 3 maps green buildings per capita by state. There are a few places with a noteworthy concentration. First, green buildings appear to be more common in particularly warm climates, like in the southwestern US and Hawaii. These places have high cooling costs year round, so green homes and offices in these areas likely have higher energy savings than elsewhere. Hawaii also has the highest average electricity costs of any US state. Second, green buildings are somewhat more common in Maryland and Virginia. The Energy Independence and Security Act of 2007 requires federal agencies to lease Energy Star certified real estate, or other known energy efficient space (e.g. LEED certified). The large number of federal agencies and federal contractors in these states explains the relative frequency of green buildings in these states.

geocoding platform freely processes batches large enough for this dataset, and is more accurate than other free geocoding services, like the Census'. It also returns information about the strength of the match. I filter the data so it includes only buildings that are matched to at least the street. This avoids problems where the geocoder only recognizes the city, county, or state of a building's address.

[6]The US Census Bureau splits each county into census tracts. These units of geography typically have between 1,200 and 8,000 residents.

4.2 Census Tract Data

Additional census tract data come from the 2019 American Community Survey.[7] This includes demographic information about the residential population (race, ethnicity, income, age) as well as data on the housing stock in each census tract. Although these variables are not the focus of the analysis, they have entirely plausible relationships with the location and quantity of green commercial real estate.

The model focuses on the distribution of green commercial real estate within cities, so we also need to match census tracts to cities. Cities are often an imprecise unit of geography. Instead, I use urbanized areas and will use the term city and urbanized area interchangeably. The Census defines an urbanized area as a core area with a population of at least 50,000 and additional requirements based on population density. In 2010, there were 486 urbanized areas in the US, and these areas contained 71.2% of the US population. I pull the shapefiles from all urbanized areas in the US from the Census and match census tracts to their appropriate urbanized area (if the census tract is in one).[8]

The previous data measures green commercial real estate and neighborhood characteristics, but does not contain any relevant information on agglomeration economies. Although I can retrieve population-density for each census tract from the ACS, agglomeration economies are driven by dense worker-populations, not residential-populations. Instead, I use cellphone data from SafeGraph to proxy the number of people in each census tract during the workweek. SafeGraph collects data on the GPS pings of over 20 million anonymous devices. The company can then trace these GPS pings back to places of interests or common units of geography. The data I have come aggregated at the census block group level. Along with location data, SafeGraph records the timing of GPS pings. They can then count the number of cellphones in their database that ping within a census block group over a specified time frame, like during the workday. SafeGraph uses this to form what I will refer to as the worker population of a census block group.[9]

[7] I access this data through the `tidycensus` package in `R` and the Census API.

[8] The Census defines urbanized areas in census block groups, a smaller level of geography than census tracts. As a result, there are many census tracts that are only partially within an urbanized area. I will proceed by considering only those census tracts that are fully contained within an urbanized area.

[9] This is called the `work_hours_device_home_areas` in the raw data from SafeGraph.

Table 1: Summary Statistics

	Full Sample			≥ 1 Green Comm. Building			No Green Comm. Buildings		
	Mean	Median	Stn Dev	Mean	Median	Stn Dev	Mean	Median	Stn Dev
Green Commercial Real Estate									
Green Comm. Buildings	1	0	3	3	1	5	0	0	0
Green Comm. Real Estate (ft^2)	131,266	0	893,299	422,467	96,786	1,563,763	0	0	0
Cellphone Data									
Worker Population	3,042	2,130	3,135	4,813	3,662	4,358	2,243	1,734	1,911
Worker Density (workers/mi^2)	6,029	3,069	13,089	8,385	3,965	19,461	4,967	2,723	8,615
Residents									
Residential Population	4,248	4,040	1,931	4,546	4,334	2,069	4,113	3,904	1,850
Median Household Income	67,913	60,136	36,238	72,400	64,495	37,616	65,894	58,054	35,416
Median Housing Value	323,127	240,100	277,600	364,609	278,200	298,319	304,676	222,950	265,805
Median Age of Residents	38	37	7	37	37	8	38	37	7
Median Year Housing Built	1970	1969	32	1973	1974	47	1968	1967	20
Proportion White	0.619	0.689	0.265	0.654	0.710	0.231	0.603	0.674	0.277
Proportion Black	0.191	0.078	0.254	0.156	0.071	0.207	0.207	0.082	0.271
No. of Observations		35,853			11,140			24,713	

The worker population of a census block group is the number of devices that are located in the census block group during conventional working hours (7:30 am to 5:30 pm, Monday through Friday) more than any other census block group over the most recent 45 day period. I aggregate this further to the census tract level. All cellphone data are taken from February 2019.

Table 1 contains summary statistics for the dataset. Observations in the sample are neighborhoods (census tracts). The distribution of green commercial buildings and real estate (in square feet) has a heavy right skew. Over half of all neighborhoods in the sample do not have a commercial green building. The considerable variation in the quantity of green commercial real estate makes the central task of explaining this variation challenging. The distributions of the worker populations and the worker densities are also heavily right skewed. While the cellphone data have more modest variations than the green building data, they still have large variations.

Table 1 also presents summary statistics for two subsets of the data: census tracts with at least some green commercial real estate and census tracts with no green commercial real estate. Among those census tracts with green commercial real estate, there is still large variation in the amount of green commercial real estate, both when looking at the number of buildings and the square footage of these buildings.

There are notable differences between neighborhoods with and without green commercial buildings in both the cellphone data and the residential data. The mean and median worker populations in neighborhoods with green commercial buildings are more than the double the values in neighborhoods without green commercial buildings. Although this certainly fits with the prediction that worker-dense neighborhoods with higher agglomeration economies have more green commercial real estate, it may also just reflect that these neighborhoods have more commercial real estate. The residential population of two types of neighborhoods is comparable, with green neighborhoods having slightly more residents than non-green neighborhoods. Green neighborhoods tend to have wealthier residents, more expensive housing, and more white residents. These differences appear to be relatively small. The distribution of the median ages of residents is comparable between the two subsets of data. Overall, the clearest differences between green and non-green neighborhoods are the worker population and worker density. At a minimum, this is suggestive of the model's relationship.

5 Estimation

As mentioned earlier, I do not have data to estimate the model precisely. This would require data on individual firms that would allow me to calibrate the magnitude of agglomeration economies and ecological responsiveness that these firms face. Instead, I focus on verifying the key prediction that worker-dense neighborhoods will contain more green commercial real estate, relative to the total stock of commercial real estate in the neighborhood.

Unfortunately, I do not have data on the total stock of commercial real estate at the neighborhood level. To evaluate the model's prediction, I make the assumption that the total quantity of commercial real estate in a neighborhood is a function of the number of workers in that neighborhood. Presumably, firms need more space when there are more workers, so this function should be increasing. While this assumption may generally be true and is convenient for estimation purposes, certainly this assumption would not always hold in practice. The relationships between the number of workers and the total square footage of commercial real estate will also vary by use. Even if there are more workers in an office building than a warehouse, the warehouse may be the larger commercial space. This assumption limits the implications of the estimation.

Let equation (25) be the population regression function for neighborhood i in city k where R_g is green commercial real estate (square feet), N is the number of workers, $\mathbf{X}$ is a vector of neighborhood characteristics, N_{ik}/ℓ_{ik} is the worker density of neighborhood ik, and c_k is a dummy variable for city k.

$$\log(R_g)_{ik} = \alpha + \beta \log(N_{ik}) + \gamma \mathbf{X}_{ik} + \sum_{k=1}^{K} \left[\delta_k \log\left(\frac{N_{ik}}{\ell_{ik}}\right) c_k \right] + \varepsilon_{ik} \tag{25}$$

The inclusion of the $\log(N_{ik})$ term reflects the assumption that the total stock of commercial real estate is a function of the number of workers in a neighborhood. Controlling for the total number of workers amounts to controlling for the stock of commercial real estate with this assumption. By taking $\log(N_{ik})$, I assume that this function takes the form $R = N^{\beta}$, where R is the total commercial real estate stock. Appendix B discusses and estimates an alternative specification where this function is exponential. The exponential functional form is less compelling and does not hold as well empirically.

Although the original model did not include any other neighborhood characteristics, I include these based on the summary statistics in table 1. In reality there are many factors that might contribute to a firm's decision to go green in certain neighborhoods. This vector of covariates attempts to control for these in estimation.

The primary coefficients of interest are the δ_k's. The fourth term in (25) represents the city-level effects of worker density on green commercial real estate. The density of workers and residents varies from city to city along with the strength of agglomeration economies. Along with policy differences, this means neighborhoods with identical characteristics and worker-densities might vary in their green commercial real estate. For this reason, the population regression function allows δ to vary from city to city. Then if worker-dense neighborhoods do have more green commercial real estate relative to the total stock of commercial real estate, then δ_k should be positive for each city k.

There may be some concern that because the δ_k's are the only coefficients that vary from city to city, that the effect may represent differences between cities more than differences between the worker-density of neighborhoods within cities. To accommodate this concern, consider the alternative population regression function in equation (26).

$$\log(R_g)_{ik} = \alpha + \beta \log(N_{ik}) + \gamma \mathbf{X}_{ik} + \sum_{k=1}^{K} \delta_k c_k + \varepsilon_{ik} \tag{26}$$

This specification omits worker-density entirely and instead includes city-level constants.

One challenge in estimating these regressions on the sample is the right-skew of green commercial real estate. The population regression and sample regression present the relationship between the quantity of green commercial real estate and worker-density in log-log form which will help correct for the right skew. However, most census tracts contain no green commercial real estate. This creates a problem as the log of green commercial real estate, $\log(R_g)$, is often undefined. Given the model's derivation and its use of the logarithmic adjustment, I decide to proceed by using logarithms rather than making other transformations to handle to right skew with zeros. There are two conventional approaches for estimating a log-log model with zeros in the dependent variable: (1) adding a positive constant to R_g so that $R_g > 0$ for all census tracts,[10] or (2)

[10] Bellego, Benatia, and Pape (2021) finds that 48% of papers in the American Economic Review published between

excluding census tracts where $R_g = 0$. I consider both approaches. Adding a positive constant has the advantage of keeping a much larger sample, but the interpretation of estimates is tumultuous. Omitting neighborhoods without any green commercial real estate comes with a much stronger interpretation, but costs over half of the sample size.

Table 2 takes the first approach by estimating the sample regression function,

$$\log(\widehat{R_g + 1})_{ik} = \widehat{\alpha} + \widehat{\beta}\log(N_{ik}) + \widehat{\gamma}\mathbf{X}_{ik} + \sum_{k=1}^{K}\left[\widehat{\delta_k}\log\left(\frac{N_{ik}}{\ell_{ik}}\right)c_k\right] + \widehat{\varepsilon_{ik}} \tag{27}$$

in specifications (1), (3), (4), and (6), and estimating the sample regression function,

$$\log(\widehat{R_g + 1})_{ik} = \widehat{\alpha} + \widehat{\beta}\log(N_{ik}) + \widehat{\gamma}\mathbf{X}_{ik} + \sum_{k=1}^{K}\widehat{\delta_k}c_k + \widehat{\varepsilon_{ik}} \tag{28}$$

in specifications (2) and (5). This dataset contains both neighborhoods with and without green commercial real estate by adding the constant 1 to R_g for each observation. To preserve the quality of the $\widehat{\delta_k}$, we should ensure there are enough observations (neighborhoods) in a city. I consider two thresholds, 20 census tracts and 40 census tracts, for this. In specifications (1), (2), and (3), a census tract is only included in the sample if there are 20 census tracts in sample located in the same city. In specifications (4), (5), and (6), a census tract is only included in sample if there are 40 census tracts in the sample located in the same city. This restriction prevents the possibility of that $\widehat{\delta_k}$ is estimated on just a couple neighborhoods. Specifications (1) and (4) contain no neighborhood covariates, (i.e. the vector $\mathbf{X}$ is empty).

Overall, the coefficient estimates in table 2 are almost all significant. The variation in commercial green real estate is sizeable, and the model does a modest job at explaining this variation. The inclusion of the neighborhood characteristics noticeably improves the model's explanatory power. Within the vector $\mathbf{X}$ I include a dummy variable indicating whether or not the median year housing was built was in 1998 or after. The Energy Star program first began certifying commercial buildings in 1998. Housing in these neighborhoods was mostly developed after the introduction of the program. If commercial development occurred roughly at the same time as residential development, then this may help explain why these, recently developed neighborhoods have more

2016 and 2020 take this approach. Other popular, non-logarithmic, fixes include Poisson and inverse hyperbolic sine transformations.

Table 2: Neighborhoods with and without Green Real Estate

	Log [Green Real Estate (ft^2) +1]					
	$\geq$ 20 Tracts per City			$\geq$ 40 Tracts per City		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Worker Population	3.221**	3.416**	3.138**	3.225**	3.409**	3.138**
	(0.038)	(0.038)	(0.045)	(0.039)	(0.039)	(0.046)
Log Median Resident Age		−3.021**	−2.501**		−3.166**	−2.634**
		(0.194)	(0.198)		(0.201)	(0.205)
Log Income per Capita		2.149**	2.221**		2.243**	2.312**
		(0.112)	(0.113)		(0.115)	(0.115)
Log Median Housing Value		0.200*	0.151		0.178*	0.124
		(0.087)	(0.087)		(0.089)	(0.089)
Proportion Black		1.337**	1.146**		1.317**	1.120**
		(0.146)	(0.147)		(0.149)	(0.149)
Median Housing Built After 1998?		0.648**	0.751**		0.679**	0.778**
		(0.135)	(0.134)		(0.138)	(0.138)
City-Level Constants		✓			✓	
Log Worker Density – City	✓		✓	✓		✓
Observations	33,691	28,100	28,100	31,714	26,376	26,376
R^2	0.262	0.292	0.296	0.259	0.291	0.295
Adjusted R^2	0.258	0.287	0.291	0.256	0.288	0.292

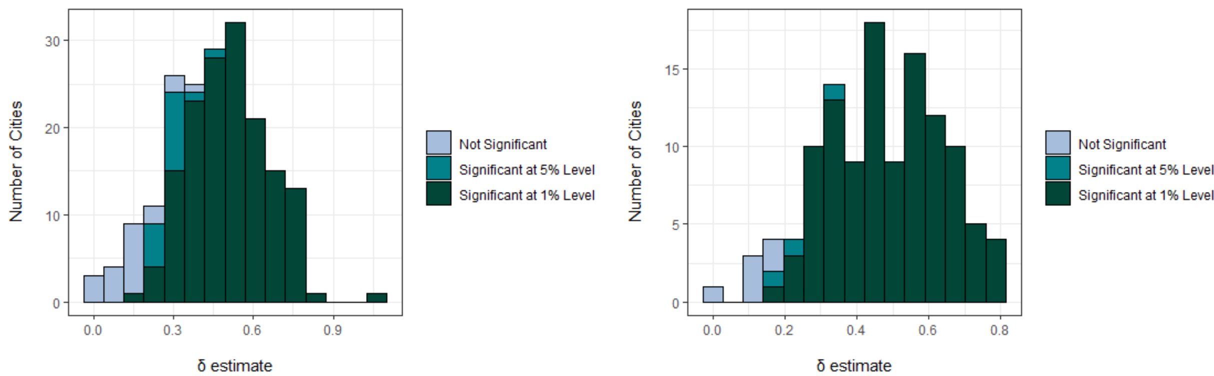
Notes: **Significant at the 1% level; *Significant at the 5% level

green development. This effect is both statistically and economically significant: neighborhoods where the median build year for housing is 1998 or after have 64.8% to 77.8% more green commercial real estate than other neighborhoods. This interpretation comes with the necessary caveat that it applies only to neighborhoods with at least some green commercial development.

Specifications (2) and (5) use city level constants instead of incorporating worker density. These two models explain just negligibly less variation in green commercial real estate than the alternative specifications, (3) and (6), that use worker density.

There are 190 cities represented in the sample with the restriction of at least 20 tracts per city, and 119 cities represented in the sample restriction of at least 40 tracts per city. To validate the prediction that worker-dense neighborhoods have more green commercial real estate, we

Figure 4: Distribution of $\widehat{\delta}$ in Table 2



(a) Specification (3)

(b) Specification (6)

check that the $\widehat{\delta}$ for each of these cities is significant and positive. Figure 4 verifies this hypothesis. Under both of the specifications in table 2 that include both neighborhood characteristics and worker-density, the distribution of the $\widehat{\delta}$ are predominantly significant, and conditional on their significance, they are always positive. The center of the significant estimates under both specifications is near 0.5. This indicates that under the assumption each neighborhood has at an additional square foot of green commercial real estate, a 1% increase in worker density is associated with around 0.5% increase in green commercial space.

The choice of adding 1 in to R_g is somewhat arbitrary, although common practice. Bellego, Benatia, and Pape (2021) indicates though that the choice of what constant to add can affect estimates and that the smallest computationally possible constant is not necessarily the best. I explore this in table 7, estimating the same model but changing the dependent variable by adding different constants. The choice of constant does affect coefficient estimates. Changing the constant does appear to preserve the sign of the estimates though, indicating that the hypothesized relationship holds with the true data generating process.

I take the alternative approach to the log(0) problem by estimating the sample regression functions that correspond with equations (25) and (26) on the subset of neighborhoods with at least one green commercial building. These regression results appear in table 3 and contain the specifications analogous to the specifications in table 2. The interpretation of these results is less ambiguous than the results in 2, though still limited.

The regression results in 3 differ from the results in 2 in a few key ways. First, when the sample contains just those neighborhoods with green commercial real estate, the models have less explanatory power in each case. Removing the predictability of the neighborhoods without green commercial real estate exposes the model to variation that is not explained as well. Second, coefficient estimates are almost always smaller in magnitude in table 3 than in table 2. Because this approach restricts the possible values of the dependent variable, this is unsurprising. Lastly, the samples I estimate each specification over are less than a third of the size of the corresponding samples in table 3.

In spite of these differences, the two sets of regression results bear many of the same conclusions. With few exceptions, the same terms have significant coefficient estimates despite the reduced sample size. The signs of coefficient estimates are almost always preserved as well. These coefficients have cleaner interpretations. Using specification (3), a 1% increase in the worker population is associated with 0.941% increase in green commercial real estate. Note that if in fact the total stock of commercial real estate is a function of the worker population and that the *density* of workers determines the proportion of commercial real estate that is green, the population coefficient would be 1 exactly, nearly the estimate. That is, a 1% increase in the worker population fits with a 1% increase in the stock of commercial real estate. We do not predict that the proportion of green real estate will change with the worker population alone, and expect then that all else equal, a 1% increase in all commercial real estate will correspond with a 1% increase in green commercial real estate. These estimates fit well with this prediction.

The coefficient estimates on the dummy variable for median housing build year of 1998 or after are still statistically and economically significant with this subset of the original sample. A neighborhood with green commercial real estate where the median year housing was built is after 1998 is expected to have 42.3% more green commercial real estate than an identical neighborhood where the median year housing was built is before 1998. Clearly green commercial development is more prevalent in neighborhoods that have experienced more recent development.

The neighborhoods in this restricted sample come from fewer cities. There are 89 cities that have at least 20 neighborhoods with green commercial buildings, and 53 cities that have at least 40 neighborhoods with green commercial buildings. Even with these restrictions, figure 5 indicates similar results to those in figure 4. For a neighborhood in the average city, a 1% increase in worker density is associated with a 0.15% increase in the stock of green commercial real estate.

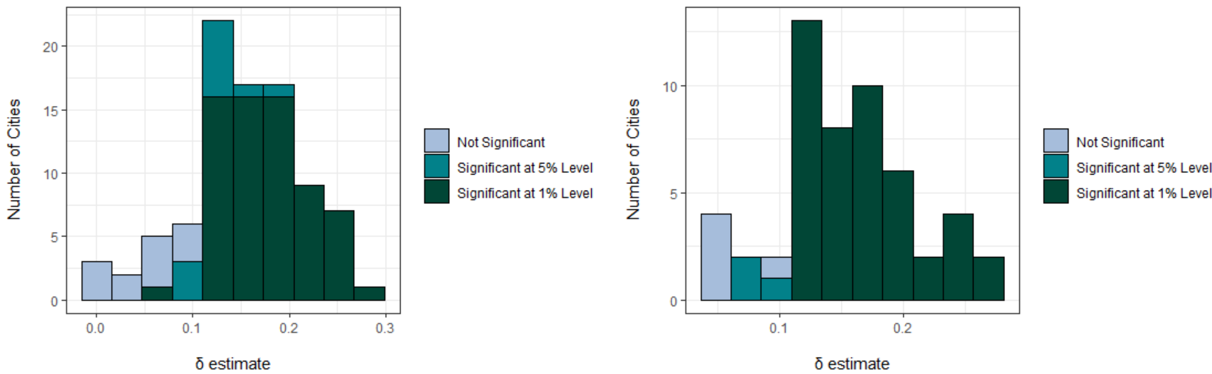
Table 3: Neighborhoods with Green Real Estate

	Log Green Real Estate					
	≥ 20 Tracts per City			≥ 40 Tracts per City		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Worker Population	0.961**	1.037**	0.941**	0.972**	1.040**	0.947**
	(0.028)	(0.030)	(0.033)	(0.030)	(0.032)	(0.035)
Log Median Resident Age		−1.792**	−1.500**		−1.844**	−1.552**
		(0.129)	(0.136)		(0.138)	(0.145)
Log Income per Capita		0.953**	0.898**		0.991**	0.935**
		(0.073)	(0.073)		(0.077)	(0.077)
Log Median Housing Value		−0.120*	−0.129*		−0.110	−0.120
		(0.060)	(0.060)		(0.064)	(0.064)
Proportion Black		0.738**	0.613**		0.739**	0.611**
		(0.122)	(0.123)		(0.129)	(0.130)
Housing Built After 1998?		0.392**	0.423**		0.401**	0.430**
		(0.077)	(0.077)		(0.080)	(0.080)
City-Level Constants		✓			✓	
Log Worker Density – City	✓		✓	✓		✓
Observations	9,495	8,064	8,064	8,550	7,271	7,271
R^2	0.191	0.225	0.229	0.186	0.223	0.227
Adjusted R^2	0.183	0.215	0.220	0.180	0.217	0.221

Notes: **Significant at the 1% level; *Significant at the 5% level

In both regression tables, there is a strong difference in the stock of green commercial real estate between neighborhoods where most housing was built 1998 or after and neighborhoods where most housing was built before 1998. Given this, I also estimate the regression models on the set of neighborhoods whose median year housing was built is 1998 or after. This attempts to control for the presence of any land development in the neighborhood. There are few urban neighborhoods that had such widespread residential development since 1998. Even rarer are cities with many of these neighborhoods. Regression results with specifications analogous to those in tables 2 and 3 are in tables 4 and 5 in Appendix B. The sample sizes are small, and estimates struggle with statistical significance.

Figure 5: Distribution of $\widehat{\delta}$ in Table 3



(a) Specification (3)

(b) Specification (6)

6 Conclusion

This paper has developed a theoretical model of green commercial development within cities and validated one of its major predictions. It contributes to the existing literature by formalizing intuition into an economic model. This model makes predictions that future empirical work could build on and attempt to verify. The analysis provides evidence that worker-dense neighborhoods do have more green commercial real estate relative to their total commercial real estate, although the difference is modest.

The empirical work in this paper expands on the empirical work of other papers that evaluate the spatial distribution of green commercial buildings within cities. Braun and Bienert (2015) relies on the strong assumption of monocentric cities to argue that green buildings are concentrated in downtown, dense locations. Like Braun and Bienert (2015) and Kaza, Lester, and Rodriguez (2013), the lack of data on the total stock of commercial real estate limits the empirical analysis in this paper. The assumption made to adjust for this is strong, but the estimates in table 3 are suggestive that this assumption may hold to some degree empirically. Future empirical work might attempt to refine the dataset by considering specific types of green commercial buildings to reduce the strength of this assumption. Additional work that looks at other evidence of agglomeration economies within industries and matches this to the clustering of green development may prove useful as well.

Together, the model and empirical analysis suggest that ecologically responsive firms tend to cluster together in certain neighborhoods. This has implications for place-based policies. Because firms sort into neighborhoods , the location of a firm sends a rough signal to policymakers about the firm's propensity to adopt green buildings, and possibly green practices more generally. Empirical models like those in this paper could be used to identify places where green development is unlikely to happen and where environmentally-unresponsive firms might locate. This provides a useful tool to policymakers looking to target green building incentives. For economists, it provides a theoretical foundation for such policies.

References

Allcott, H., and M. Greenstone (2012): "Is there an energy efficiency gap?," *Journal of Economic Perspectives*, 26(1), 3–28.

Alonso, W. (1964): *Location and land use*. Harvard University Press.

Bansal, P., and K. Roth (2000): "Why companies go green: A model of ecological responsiveness," *Academy of Management Journal*, 43(4), 717–736.

Bellego, C., D. Benatia, and L.-D. Pape (2021): "Dealing with logs and zeros in regression models," *CREST-Série des Documents de Travail*, (2019-13).

Braun, T., and S. Bienert (2015): "Is green (still) a matter of prime? Stylized facts about the location of commercial green buildings," *Journal of Sustainable Real Estate*, 7(1), 160–182.

Climate Impact Lab (2020): "Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits," Discussion paper, TA Carleton, AM Jina, T Micheal, M Greenstone, T Houser, SM Hsiang, A Hultgren, RE Kopp, E Robert, KE McCusker, and IB Nath.

Eichholtz, P., N. Kok, and J. M. Quigley (2010): "Doing well by doing good? Green office buildings," *American Economic Review*, 100(5), 2492–2509.

——— (2016): "Ecological responsiveness and corporate real estate," *Business & Society*, 55(3), 330–360.

Gaubert, C. (2018): "Firm sorting and agglomeration," *American Economic Review*, 108(11), 3117–53.

Gerarden, T. D., R. G. Newell, and R. N. Stavins (2017): "Assessing the energy-efficiency gap," *Journal of Economic Literature*, 55(4), 1486–1525.

Glaeser, E. L. (2008): *Cities, agglomeration, and spatial equilibrium*. Oxford University Press.

Glaeser, E. L. (2010): *Agglomeration economics*. University of Chicago Press.

IPCC (2021): "Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," Discussion paper, Masson Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.), Cambridge University Press. In Press.

Kaza, N., T. W. Lester, and D. A. Rodriguez (2013): "The spatio-temporal clustering of green buildings in the United States," *Urban Studies*, 50(16), 3262–3282.

Nyborg, K., and T. Zhang (2013): "Is corporate social responsibility associated with lower wages?," *Environmental and Resource Economics*, 55(1), 107–117.

Resources for the Future (2021): "US carbon emissions at a glance," *Federal Climate Policy Toolkit*.

Walls, M., T. Gerarden, K. Palmer, and X. F. Bak (2017): "Is energy efficiency capitalized into home prices? Evidence from three US cities," *Journal of Environmental Economics and Management*, 82, 104–124.

Wiley, J. A., J. D. Benefield, and K. H. Johnson (2010): "Green design and the market for commercial office space," *The Journal of Real Estate Finance and Economics*, 41(2), 228–243.

Appendix A

Firm's Derivation

Let the firm's profit function be given by

$$A\psi(\alpha_i, N)\lambda(d, \theta_j)L^{\beta}R^{\gamma}K^{1-\beta-\gamma} - WL - p_d R - k_i \tag{29}$$

Yields the first-order conditions:

$$\frac{\partial \pi}{\partial L} = \beta A\psi(\alpha_i, N)\lambda(d, \theta_j)L^{\beta-1}R^{\gamma}\bar{K}^{1-\beta-\gamma} - W = 0 \tag{30}$$

$$\frac{\partial \pi}{\partial R} = \gamma A\psi(\alpha_i, N)\lambda(d, \theta_j)L^{\beta}R^{\gamma-1}\bar{K}^{1-\beta-\gamma} - p_d = 0 \tag{31}$$

Because all terms in (31) are non-zero, then the two first-order conditions yield:

$$\frac{\beta}{\gamma}\frac{R}{L} = \frac{W}{p_d}$$

This allows us to solve for the firm's optimal combination of labor and commercial real estate in terms of each other.

$$R = \frac{\gamma}{\beta}\frac{W}{p_d}L \tag{32}$$

$$L = \frac{\beta}{\gamma}\frac{p_d}{W}R \tag{33}$$

Taking (32) and substituting it into (30),

$$\beta A\psi(\alpha_i, N)\lambda(d, \theta_j)L^{\beta-1}\left(\frac{\gamma}{\beta}\frac{W}{p_d}L\right)^{\gamma}\bar{K}^{1-\beta-\gamma} = W$$

When we simplify this, we have the firm's optimal quantity of labor in terms of wages and the price of commercial real estate with a given design d. That is, given the price of commercial real estate the firm pays, we have the firm's individual labor demand curve.

$$L^{*} = \left[\beta^{1-\gamma}\gamma^{\gamma}A\psi(\alpha_i, N)\lambda(d, \theta_j)W^{\gamma-1}p_d^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} \tag{34}$$

We can apply a similar process to solve for the firm's optimal quantity of real estate. If we substitute (33) into (31), we obtain

$$\gamma A\psi(\alpha_i, N)\lambda(d,\theta_j)\left(\frac{\beta}{\gamma}\frac{p_d}{W}R\right)^{\beta} R^{\gamma-1}\bar{K}^{1-\beta-\gamma} = p_d.$$

This simplifies to the firms demand curve for commercial real estate given the neighborhood wage,

$$R = \left[\beta^{\beta}\gamma^{1-\beta}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{\beta-1}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K}. \tag{35}$$

With the equilibrium values of L and R, we can substitute these back into the firm's original profit function.

$$\begin{aligned}
\pi(L^*, R^*) &= A\psi(\alpha_i, N)\lambda(d,\theta_j)(L^*)^{\beta}(R^*)^{\gamma}K^{1-\beta-\gamma} - WL^* - p_d R^* - k_i \\
&= A\psi(\alpha_i, N)\lambda(d,\theta_j)\left[\beta^{1-\gamma}\gamma^{\gamma}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{\gamma-1}p_d^{-\gamma}\right]^{\frac{\beta}{1-\beta-\gamma}}\bar{K}^{\beta} \\
&\quad \left[\beta^{\beta}\gamma^{1-\beta}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{\beta-1}\right]^{\frac{\gamma}{1-\beta-\gamma}}\bar{K}^{\gamma} \\
&\quad - W\left[\beta^{1-\gamma}\gamma^{\gamma}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{\gamma-1}p_d^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} \\
&\quad - p_d\left[\beta^{\beta}\gamma^{1-\beta}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{\beta-1}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} - k_i \\
&= \left[\beta^{\beta}\gamma^{\gamma}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} - \left[\beta^{1-\gamma}\gamma^{\gamma}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} \\
&\quad - \left[\beta^{\beta}\gamma^{1-\beta}A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} - k_i \\
&= \left(\left(\beta^{\beta}\gamma^{\gamma}\right)^{\frac{1}{1-\beta-\gamma}} - \left(\beta^{1-\gamma}\gamma^{\gamma}\right)^{\frac{1}{1-\beta-\gamma}} - \left(\beta^{\beta}\gamma^{1-\beta}\right)^{\frac{1}{1-\beta-\gamma}}\right)\left[A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{-\gamma}\right]\bar{K} - k_i \\
&= \Phi(\beta,\gamma)\left[A\psi(\alpha_i, N)\lambda(d,\theta_j)W^{-\beta}p_d^{-\gamma}\right]\bar{K} - k_i
\end{aligned} \tag{36}$$

where $\Phi(\beta,\gamma) = \left(\beta^{\beta}\gamma^{\gamma}\right)^{\frac{1}{1-\beta-\gamma}} - \left(\beta^{1-\gamma}\gamma^{\gamma}\right)^{\frac{1}{1-\beta-\gamma}} - \left(\beta^{\beta}\gamma^{1-\beta}\right)^{\frac{1}{1-\beta-\gamma}}$. This is the firm's profit function, now written as a function of the market wage W and the or commercial real estate per square foot with design d.

Zero Profit Conditions for Firms

There are two types of firms and both types must earn zero profits. Using the profit functions with profit-maximizing inputs in equation (8), yields two conditions.

$$k_1 = \Phi\left[A\left(\frac{N}{\overline{\ell}}\right)^{\alpha}\lambda_g W^{-\beta}p_g^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K} \tag{37}$$

$$k_2 = \Phi\left[A\left(\frac{N}{\overline{\ell}}\right)^{\alpha}\lambda_b W^{-\beta}p_b^{-\gamma}\right]^{\frac{1}{1-\beta-\gamma}}\bar{K}. \tag{38}$$

Our preferred form of these conditions will move our endogenous variables to one side of the equation. Simplifying and rearranging these two conditions are equivalent to,

$$N^{-\alpha}W^{\beta}p_g^{\gamma} = \Phi^{1-\beta-\gamma}\left[A\lambda_g\beta^{\beta}\gamma^{\gamma}\overline{\ell}^{-\alpha}\right]\left(\frac{\bar{K}}{k_1}\right)^{1-\beta-\gamma} \tag{39}$$

$$N^{-\alpha}W^{\beta}p_g^{\gamma} = \Phi^{1-\beta-\gamma}\left[A\lambda_b\beta^{\beta}\gamma^{\gamma}\overline{\ell}^{-\alpha}\right]\left(\frac{\bar{K}}{k_2}\right)^{1-\beta-\gamma}\left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta}} \tag{40}$$

Here we have eliminated p_b by using the premium in equation (13). We do not need both equations – with the premium these are equivalent. Notice that in order for these to be equivalent, the ratio of fixed costs must adjust. We have fixed the number of firms, so fixed costs will adjust here to accommodate this. Then

$$C_1 = \Phi^{1-\beta-\gamma}\left[A\lambda_g\beta^{\beta}\gamma^{\gamma}\overline{\ell}^{-\alpha}\right]\left(\frac{\bar{K}}{k_1}\right)^{1-\beta-\gamma}. \tag{41}$$

Labor Market Clearing

Proceeding, we require that the the labor market clears. That is, the sum of labor demand from green and brown firms in the neighborhood must equal the number of workers in the neighborhood. There are μM green firms and $(1-\mu)M$ brown firms, each of which has the labor demand function in equation (6). Let $\eta = 1-\beta-\gamma$. Then the labor market clearing condition is:

$$
\begin{aligned}
N &= \mu M\left[\beta^{1-\gamma}\gamma^{\gamma}A\left(\frac{N}{\bar{\ell}}\right)^{\alpha}\lambda_g W^{\gamma-1}p_g^{-\gamma}\right]^{\frac{1}{\eta}}\bar{K} \\
&\quad + (1-\mu)M\left[\beta^{1-\gamma}\gamma^{\gamma}A\left(\frac{N}{\bar{\ell}}\right)^{\alpha}\lambda_b W^{\gamma-1}p_b^{-\gamma}\right]^{\frac{1}{\eta}}\bar{K} \\
&= M\left[\beta^{1-\gamma}\gamma^{\gamma}A\left(\frac{N}{\bar{\ell}}\right)^{\alpha}W^{\gamma-1}\right]^{\frac{1}{\eta}}\bar{K}\left[\mu\left(\lambda_g p_g^{-\gamma}\right)^{\frac{1}{\eta}} + (1-\mu)\left(\lambda_b p_b^{-\gamma}\right)^{\frac{1}{\eta}}\right] \\
&= M\left[\beta^{1-\gamma}\gamma^{\gamma}A\left(\frac{N}{\bar{\ell}}\right)^{\alpha}W^{\gamma-1}p_g^{-\gamma}\right]^{\frac{1}{\eta}}\bar{K}\left[\mu\left(\lambda_g\right)^{\frac{1}{\eta}} + (1-\mu)\left(\lambda_b\right)^{\frac{1}{\eta}}\left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta\eta}}\right]
\end{aligned}
$$

which makes use of the premium derivation. We want the endogenous variables on one side of the equation.

$$
N^{\frac{\eta-\alpha}{\eta}}W^{\frac{1-\gamma}{\eta}}p_g^{\frac{\gamma}{\eta}} = M\left[\beta^{1-\gamma}\gamma^{\gamma}A\bar{\ell}^{-\alpha}\right]^{\frac{1}{\eta}}\bar{K}\left[\mu\left(\lambda_g\right)^{\frac{1}{\eta}} + (1-\mu)\left(\lambda_b\right)^{\frac{1}{\eta}}\left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta\eta}}\right] \tag{42}
$$

$$
N^{\eta-\alpha}W^{1-\gamma}p_g^{\gamma} = \left[\beta^{1-\gamma}\gamma^{\gamma}A\bar{\ell}^{-\alpha}\right](M\bar{K})^{\eta}\left[\mu\left(\lambda_g\right)^{\frac{1}{\eta}} + (1-\mu)\left(\lambda_b\right)^{\frac{1}{\eta}}\left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta\eta}}\right]^{\eta}. \tag{43}
$$

From this, we can let

$$
C_2 = \left[\beta^{1-\gamma}\gamma^{\gamma}A\bar{\ell}^{-\alpha}\right](M\bar{K})^{\eta}\left[\mu\left(\lambda_g\right)^{\frac{1}{\eta}} + (1-\mu)\left(\lambda_b\right)^{\frac{1}{\eta}}\left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta\eta}}\right]^{\eta}. \tag{44}
$$

Real Estate Market Clearing: Demand

Unlike the labor market, the real estate market is segmented between green and brown real estate. We will have two demand functions and two supply functions. However, we can reduce this to just a single supply and a single demand function. Market demand for green and brown real estate is given by

$$
R_g = \mu M\left[\beta^{\beta}\gamma^{1-\beta}A\left(\frac{N}{\bar{\ell}}\right)^{\alpha}\lambda_g W^{\beta}p_g^{\beta-1}\right]^{\frac{1}{\eta}}\bar{K} \tag{45}
$$

$$
R_b = (1-\mu)M\left[\beta^{\beta}\gamma^{1-\beta}A\left(\frac{N}{\bar{\ell}}\right)^{\alpha}\lambda_b W^{\beta}p_b^{\beta-1}\right]^{\frac{1}{\eta}}\bar{K} \tag{46}
$$

$$
\tag{47}
$$

To start, we can rearrange green real estate demand to get it into our standard form.

$$R_g = \mu M \left[\beta^\beta \gamma^{1-\beta} A \left(\frac{N}{\bar{\ell}} \right)^\alpha \lambda_g W^\beta p_g^{\beta-1} \right]^{\frac{1}{\eta}} \bar{K} \tag{48}$$

$$N^{\frac{-\alpha}{\eta}} W^{\frac{\beta}{\eta}} p_g^{\frac{1-\beta}{\eta}} R_g = \mu M \left(\beta^\beta \gamma^{1-\beta} A \lambda_g \right)^{\frac{1}{\eta}} \bar{K} \tag{49}$$

$$N^{-\alpha} W^\beta p_g^{1-\beta} R_g^\eta = \left(\beta^\beta \gamma^{1-\beta} A \lambda_g \right) (\mu M \bar{K})^\eta. \tag{50}$$

This is the form in the main text, where

$$C_3 = \left(\beta^\beta \gamma^{1-\beta} A \lambda_g \right) (\mu M \bar{K})^\eta. \tag{51}$$

It turns out that in equilibrium, R_b will be proportion to R_g. Consider the ratio between green and brown real estate on the firm's side:

$$\frac{R_g}{R_b} = \left(\frac{\mu}{1-\mu} \right) \left(\frac{\lambda_g}{\lambda_b} \right)^{\frac{1}{\eta}} \left(\frac{p_g}{p_b} \right)^{\frac{\beta-1}{\eta}} \tag{52}$$

$$= \left(\frac{\mu}{1-\mu} \right) \left(\frac{\lambda_g}{\lambda_b} \right)^{\frac{1}{\eta}} \left(\frac{c_g}{c_b} \right)^{\frac{\beta-1}{\delta\eta}} \tag{53}$$

making use of equation (13).

Real Estate Market Clearing – Supply

The two market supply functions for the segmented market for real estate are

$$R_g = h_g \ell_g \tag{54}$$

$$R_b = h_b \ell_b. \tag{55}$$

The developer is not terribly interested in how much green real estate its building relative to brown real estate. Using the first two of the developer's first-order conditions and taking the ratio of green to brown real estate yields,

$$\frac{R_g}{R_b} = \left(\frac{c_b}{c_g} \right)^{\frac{1}{\delta}} \frac{\ell_g}{\ell_b}. \tag{56}$$

The developer remains agnostic about how to divide up its land between green and brown development. To find this division of land will require market demand. Setting (53) equal to (56), we get

$$\left(\frac{\mu}{1-\mu}\right)\left(\frac{\lambda_g}{\lambda_b}\right)^{\frac{1}{\eta}}\left(\frac{c_g}{c_b}\right)^{\frac{\beta-1}{\delta\eta}} = \left(\frac{c_b}{c_g}\right)^{\frac{1}{\delta}}\frac{\ell_g}{\ell_b} \tag{57}$$

$$\frac{\ell_g}{\ell_b} = \left(\frac{\mu}{1-\mu}\right)\left(\frac{\lambda_g}{\lambda_b}\right)^{\frac{1}{\eta}}\left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta\eta}} \tag{58}$$

For ease, let $\rho = \left(\frac{\mu}{1-\mu}\right)\left(\frac{\lambda_g}{\lambda_b}\right)^{\frac{1}{\eta}}\left(\frac{c_g}{c_b}\right)^{\frac{\gamma}{\delta\eta}}$. Then because $\ell_b = \overline{\ell} - \ell_g$,

$$\ell_g = \left(\frac{\rho}{1+\rho}\right)\overline{\ell} \tag{59}$$

We can the plug this and the developer's first order condition into the supply curve for green real estate.

$$R_g = h_g \ell_g \tag{60}$$

$$R_g = \left(\frac{p_g}{\delta c_g}\right)^{\frac{1}{\delta-1}}\left(\frac{\rho}{1+\rho}\right)\overline{\ell} p^{\frac{-1}{\delta-1}} R_g = \left(\frac{1}{\delta c_g}\right)^{\frac{1}{\delta-1}}\left(\frac{\rho}{1+\rho}\right)\overline{\ell} \tag{61}$$

Then we let

$$C_4 = \left(\frac{1}{\delta c_g}\right)^{\frac{1}{\delta-1}}\left(\frac{\rho}{1+\rho}\right)\overline{\ell}. \tag{62}$$

Equilibrium

Using the results from the previous subsections, we write out our system of four equations in four unknowns:

$$N^{-\alpha}\, W^{\beta}\, p_g^{\gamma} = C_1 \tag{63}$$

$$N^{(1-\alpha-\beta-\gamma)}\, W^{1-\gamma}\, p_g^{\gamma} = C_2 \tag{64}$$

$$N^{-\alpha}\, W^{\beta}\, p_g^{1-\beta}\, R_g^{1-\beta-\gamma} = C_3 \tag{65}$$

$$p_g^{\frac{-1}{\delta-1}}\, R_g = C_4 \tag{66}$$

where C_1, C_2, C_3, and C_4 are constants made up of exogenous variables and parameters. To solve this system of equations, we take the logarithm of these equations and obtain the following system of linear equations.

$$-\alpha \log(N) + \beta \log(W) + \gamma \log(p_g) = \log(C_1) \tag{67}$$

$$(1-\alpha-\beta-\gamma)\log(N) + (1-\gamma)\log(W) + \gamma \log(p_g) = \log(C_2) \tag{68}$$

$$-\alpha \log(N) + \beta \log(W) + (1-\beta)\log(p_g) + (1-\beta-\gamma)\log(R_g) = \log(C_3) \tag{69}$$

$$\frac{-1}{\delta-1}\log(p_g) + \log(R_g) = \log(C_4) \tag{70}$$

This system of equations has full rank, and consequently a unique solution. We use Matlab to solve this system of equations and confirm existence of a unique solution.

Appendix B

Table 4: Recently Developed Neighborhoods

	Log [Green Real Estate (ft^2) +1]					
	≥ 20 Tracts per City			≥ 40 Tracts per City		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Worker Population	2.970** (0.225)	3.988** (0.202)	3.157** (0.229)	2.794** (0.250)	3.793** (0.224)	2.991** (0.255)
Log Median Housing Value		−6.788** (1.057)	−3.831** (1.126)		−6.089** (1.140)	−3.320** (1.224)
Log Median Resident Age		5.162** (0.587)	3.584** (0.623)		5.347** (0.651)	3.768** (0.699)
Log Income per Capita		−0.173 (0.544)	0.116 (0.533)		−0.112 (0.614)	0.278 (0.604)
Proportion Black		2.977* (1.160)	2.095 (1.140)		2.688* (1.224)	1.927 (1.209)
City-Level Constants		✓			✓	
Log Worker Density – City	✓		✓	✓		✓
Observations	901	868	868	724	701	701
R^2	0.452	0.487	0.508	0.426	0.466	0.486
Adjusted R^2	0.441	0.474	0.496	0.417	0.455	0.474
Notes:	**Significant at the 1% level; *Significant at the 5% level					

Table 4 displays regression results for the regression models identical to those in table 2. The sample here is just those census tracts where the median year housing was built is 1998 or after. While this regression contains a much higher R^2 than the other regressions, the sample is too small to obtain significant estimates in many cases. This is true of the $\widehat{\delta}$.

Table 5: Recently Developed Neighborhoods with Green Real Estate

	Log [Green Real Estate (ft^2) +1]					
	≥ 20 Tracts per City			≥ 40 Tracts per City		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Worker Population	0.696**	1.137**	0.857**	0.391	1.137**	0.606**
	(0.136)	(0.131)	(0.142)	(0.199)	(0.208)	(0.228)
Log Median Housing Value		−3.916**	−2.558**		−4.630**	−1.329
		(0.638)	(0.710)		(1.037)	(1.249)
Log Median Resident Age		2.225**	1.601**		1.726**	0.577
		(0.328)	(0.357)		(0.487)	(0.536)
Log Income per Capita		−0.509	−0.460		0.019	0.169
		(0.303)	(0.300)		(0.430)	(0.410)
Proportion Black		0.096	−0.235		−0.441	−0.420
		(0.706)	(0.693)		(0.891)	(0.843)
City-Level Constants		✓			✓	
Log Worker Density – City	✓		✓	✓		✓
Observations	438	419	419	183	177	177
R^2	0.352	0.385	0.410	0.364	0.325	0.392
Adjusted R^2	0.331	0.359	0.383	0.346	0.293	0.359

Notes: **Significant at the 1% level; *Significant at the 5% level

Table 5 displays regression results for the regression models identical to those in table 3. The sample here is just those census tracts where the median year housing was built is 1998 or after. While this regression contains a much higher R^2 than the other regressions, the sample is too small to obtain significant estimates in many cases. This is true of the $\widehat{\delta}$.

Table 6: Exponential Relationship between Commercial Real Estate and Workers

	Log [Green Real Estate +1]					
	≥ 20 Tracts per City			≥ 40 Tracts per City		
	(1)	(2)	(3)	(4)	(5)	(6)
Worker Population	0.070^{**}	0.080^{**}	0.069^{**}	0.070^{**}	0.079^{**}	0.069^{**}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log Median Housing Value		-3.110^{**}	-2.217^{**}		-3.270^{**}	-2.345^{**}
		(0.199)	(0.201)		(0.205)	(0.208)
Log Median Resident Age		2.001^{**}	2.174^{**}		2.105^{**}	2.271^{**}
		(0.115)	(0.115)		(0.117)	(0.117)
Log Income per Capita		0.247^{**}	0.126		0.221^{*}	0.092
		(0.089)	(0.088)		(0.091)	(0.090)
Proportion Black		0.638^{**}	0.438^{**}		0.606^{**}	0.392^{**}
		(0.149)	(0.148)		(0.151)	(0.150)
Housing Built After 1998?		0.304^{*}	0.544^{**}		0.322^{*}	0.558^{**}
		(0.138)	(0.137)		(0.142)	(0.141)
City-Level Constants		✓			✓	
Log Worker Density – City	✓		✓	✓		✓
Observations	33,691	28,100	28,100	31,714	26,376	26,376
R^2	0.234	0.262	0.276	0.229	0.260	0.273
Adjusted R^2	0.230	0.257	0.271	0.226	0.256	0.270
Notes:	**Significant at the 1% level; *Significant at the 5% level					

Table 6 displays regression results for the regression models identical to those in table 2. This assumes an exponential form of for the function relating worker population to commercial building stock. Hence, the only difference is these regressions do not log the worker population. The model does not hold as well as the model in table 2 and seems less plausible.

Table 7: Alternative Transformations Effect on Estimates

	Log [Green Real Estate $+a$]				
	$a = 0.001$	$a = 0.01$	$a = 0.1$	$a = 1$	$a = 10$
	(1)	(2)	(3)	(4)	(5)
Log Worker Population	4.923**	4.328**	3.733**	3.138**	2.543**
	(0.074)	(0.065)	(0.055)	(0.046)	(0.037)
Log Median Resident Age	−3.897**	−3.476**	−3.055**	−2.634**	−2.212**
	(0.328)	(0.287)	(0.246)	(0.205)	(0.165)
Log Income per Capita	3.464**	3.080**	2.696**	2.312**	1.928**
	(0.184)	(0.161)	(0.138)	(0.115)	(0.093)
Log Median Housing Value	0.280*	0.228	0.176	0.124	0.072
	(0.142)	(0.124)	(0.107)	(0.089)	(0.071)
Proportion Black	1.715**	1.517**	1.318**	1.120**	0.922**
	(0.239)	(0.209)	(0.179)	(0.149)	(0.120)
Housing Built After 1998?	1.005**	0.929**	0.854**	0.778**	0.703**
	(0.221)	(0.193)	(0.166)	(0.138)	(0.111)
Log Worker Density – City	✓	✓	✓	✓	✓
Observations	26,376	26,376	26,376	26,376	26,376
R^2	0.283	0.286	0.290	0.295	0.302
Adjusted R^2	0.280	0.283	0.287	0.292	0.298
Notes:	**Significant at the 1% level; *Significant at the 5% level				

Table 7 displays regression results for the regression models similar to those in specification (6) of table 2. The difference is that we consider variations on the constant added to the green commercial real estate before the taking the log. This shows that the estimates do vary with the constant added.