Bunching in Real-Estate Markets: Regulated Building Heights in New York City

by

Jan K. Brueckner
University of California, Irvine

David Leather
Chapman University

Miguel Zerecero
University of California, Irvine

January 2024, revised April 2004

Abstract

This paper presents a real-estate application of the bunching methodology widely used in other areas of applied microeonomics. The focus is on regulated building heights in New York City, where developers can exceed a parcel's regulated height by incurring additional costs. Using the bunching methodology, we estimate the magnitude of these extra costs, with the results showing a modest increase in the marginal cost of floor space beyond the regulated building height. We use these estimates to predict the additional floor space that would be created by complete removal of building-height regulation in NYC. While this last exercise is circumscribed by our focus on a limited number of zoning categories, the results suggest that New York could secure notably more housing through lighter height regulation.

Bunching in Real-Estate Markets: Regulated Building Heights in New York City

by

Jan K. Brueckner, David Leather, Miguel Zerecero[†]

1. Introduction

This paper presents a rare application in the real-estate context of the bunching methodology widely used in other areas of applied microeonomics. The application is building-height regulation in New York City, where several costly actions allow a developer to exceed the regulated height for his parcel. The paper aims to use the observed bunching patterns around regulated heights to estimate the marginal-cost premium for exceeding those heights, thus capturing the size of the cost-function kink faced by developers. Our approach reverses the usual application of the bunching methodology, under which the kink size is known and the goal is to estimate a behavioral parameter. Our behavioral parameter (the exponent in a housing production function) has been reliably estimated, and we use its value to identify the unknown size of the cost-function kink. We also use our estimates to predict the increase in floor space in our sample that would result from eliminating height regulation.

The bunching methodology introduced by Saez (2010), which we adopt in its original form, has been widely used in a variety of applications. In these applications, consumers or firms bunch at a kink point of some function that enters their optimization problem, and the extent of the bunching can be used to estimate the value of an unknown parameter of interest. Saez (2010), for example, uses the extent of bunching at an income-tax schedule's kink point, where the slope changes discontinuously, to estimate the elasticity of labor supply, and Chetty et al. (2011) carry out a similar exercise. As explained in the recent surveys by Kleven (2016) and Bertanha, Caetano, Jales and Seegert (2023), the method has also been applied in many other taxation contexts as well as in research on financial markets, health care, environmental regulation, education and energy demand.[1] In addition to surveying applications, Bertanha et

---

[1] Blomquist, Newey, Kumar and Liang (2021) point out an identification issue in the basic bunching method.

al. (2023) also discuss the most recent econometric advancements in bunching methodologies.

In most studies, the size of the kink faced by optimizing agents is known (the increase in a marginal tax rate, for example), while the behavioral parameter that influences the extent of bunching is the unknown quantity. Alternatively, however, it is possible to use Saez's method to estimate the unknown size of a kink faced by optimizers, provided that the relevant behavioral parameters have been specified. A fellow researcher fittingly described this approach as "reverse engineering."[2] The present paper carries out such an exercise, with a focus on bunching due to regulated building heights for residential housing, using data from New York City. The only other studies of real-estate bunching of which we are aware are Kopczuk and Munroe (2015) and Slemrod, Weber and Shan (2017), who study transfer taxes.[3] But our focus on building heights and their regulation follows a growing literature in urban economics.[4]

In New York and most other US cities, heights are regulated via limits on a building's floor-area ratio, or FAR, which equals the square feet of floor space in the building divided by the lot size. If the building fully covers the entire lot, the FAR is simply equal to the number of floors. FAR limits vary across parcels, being low in some locations and high in others, depending on the decisions of the local planning authority.

In New York, however, several costly options allow developers to build beyond $\overline{\text{FAR}}$, the parcel's FAR limit. For example, under the Privately Owned Public Spaces (POPS) program, a developer is granted a maximum FAR above $\overline{\text{FAR}}$ in return for devoting a portion of the lot to open space with public access.[5] Because some of the developer's land goes unused in the production of floor space, reliance on the POPS program raises the cost of that space. Alternatively, a developer can purchase "air rights" from a nearby existing building whose FAR value is below its limit. When this building is contiguous to the developer's lot, air rights

---

[2] We thank Mazhar Waseem for this appellation.

[3] Since transfer taxes only apply above a large sales-price threshold, the taxes generate a "notch," a discontinuous jump in the tax burden, rather than a continuous kink. Kleven and Waseem (2013) developed an estimation method for notches analogous to Saez's kink methodology. Like Kopczuk and Munroe (2015) and Slemrod et al. (2017), applications in the research areas mentioned above often involve notches rather than kinks. In an unpublished paper, Levy (2024) studies bunching at regulated housing-quality thresholds.

[4] See Bertaud and Brueckner (2005), Brueckner and Sridhar (2012), Brueckner and Singh (2020), Ahlfeldt and McMillen (2018), Ahlfeldt and Barr (2022), Barr and Jedwab (2023), among others.

[5] See this NY Deparment of City Planning website for more information: `https://www.nyc.gov/site/plan-ning/plans/pops/pops.page`. Evidently, the amount of additional FAR is negotiated with the city.

are acquired through a "zoning lot merger," which allows the unused FAR to be transferred to the new building, with compensation to the existing building's owner. Air rights from a noncontiguous property can be bought under a different mechanism. Weinberger (2023) offers a fuller explanation of these rules while also stating that the cost of air rights in NYC ranges between \$100 and \$300 per square foot of floor space. Another path to a higher FAR is available for parcels adjacent to subway stops, where expenditures on station improvements can raise the allowed FAR by 20%.[6] In addition, developers can get FAR bonuses though NYC's inclusionary housing program.[7]

Thus, developers can build above $\overline{\text{FAR}}$, but at a cost, which leads to a kink at $\overline{\text{FAR}}$ in the marginal cost of floor space. This conclusion follows because the extra cost depends on the extent to which $\overline{\text{FAR}}$ is exceeded, which is clearest in the case of air rights. Since we observe a parcel's $\overline{\text{FAR}}$, the location of the kink on the marginal-cost schedule is known. However, its size is unknown. The goal of the paper is to estimate the size of this cost kink using the bunching methodology. To do so, we must impose a value for the capital exponent in a Cobb-Douglas housing production function, which plays the role of the unknown behavioral parameter in our model. A recent and reliable exponent estimate is drawn from Combes, Duranton and Gobillon (2021), who estimate the housing production function using French data.[8]

Using data for 2017, we proceed by focusing only on NYC parcels containing residential buildings constructed since 2000, so that observed FAR values have been set relatively recently rather than decades earlier. We divide this collection of parcels into different $\overline{\text{FAR}}$ groups, of which five have enough observations (at least 1000) to be usable for our purposes. For example, one group has $\overline{\text{FAR}}= 2.0$, while another has $\overline{\text{FAR}}= 0.9$. Then, within each $\overline{\text{FAR}}$ group, we use

---

[6] See https://zr.planning.nyc.gov/article-xi/chapter-5/115-21.

[7] See https://www.nyc.gov/site/planning/zoning/districts-tools/inclusionary-housing.page.

[8] Garicano, Lelarge, and Van Reenen (2016) carry out an exercise somewhat similar to ours in their study of the French labor market. French firms are subject to additional regulations when their employment exceeds 50 workers. As the impact of these regulations on firm employment is hard to quantify, the paper estimates the tax equivalent of the regulations by exploiting the bunching pattern of firms at the threshold. Like us, they rely on external estimates of production parameters to generate their tax-function estimates. Rather than using Saez's (2010) methodology or other bunching estimating techniques that rely on a few parametric assumptions, the paper includes a normally distributed measurement error on observed employment. This approach allows estimation of the parameters of the model using maximum likelihood. An advantage of their method is that they can use the entire distribution of employment to identify the relevant parameters.

the bunching methodology to estimate the increase in the marginal cost of floor space above $\overline{\text{FAR}}$, expressed as a proportion of the cost below $\overline{\text{FAR}}$. We estimate the standard error of our marginal-cost penalty, and thus its 95% confidence interval, by the bootstrap method, using 10,000 draws with replacement.

Glaeser et al. (2005) and Brueckner and Singh (2020) study the stringency of land-use regulation in NYC and other cities, and the latter paper's methodology allows the free-market FAR to be estimated, showing that NYC's regulated FAR values lie well below it. This result appears to confirm the allegations of many observers (including Glaeser and his coauthors) that, despite its high density, New York is *not dense enough*, with more housing and thus taller buildings needed. Our results allow us to advance this debate by computing the extra floor space that would be generated by removing the FAR limit in any of our $\overline{\text{FAR}}$ groups.[9]

While this exercise is carried out in section 5 of the paper, section 2 shows how the bunching methodology can be applied to the building-height case. Section 3 describes our data sources, and section 4 presents the estimation results. Section 6 offers conclusions.

2. Saez's analytics adapted to building heights

Saez's (2010) bunching methodology can be adapted to a developer's choice of building height. This section explains the adaptation.

2.1. Housing production

To start, we depict the production of housing floor space, relying on the standard approach in urban models. Let Q denote output of floor space, which is produced with inputs of capital K (building materials) and land ℓ using the CRS production function $H(K, \ell)$. Floor space per acre of land, denoted F, is given by $F = H(K, \ell)/\ell = H(K/\ell, 1) = H(S, 1) \equiv f(S)$, where $S = K/\ell$ is capital per acre. The developer's profit per acre of land, exclusive of land cost, is then $pf(S) - S$, where p is the rental price per square foot floor space and the price of capital is normalized to 1. As usual in urban models, developers are price-takers, treating p as parametric. The first-order condition for choice of S is $pf'(S) = 1$ (note that $f'' < 0$).[10]

---

[9] Application of our methods to other cities might be interesting, but before doing so, researchers must learn whether any such city offers developers ways of exceeding a regulated FAR limit, as does NYC.

[10] While the function f is assumed to be smooth, a recent study by Eriksen and Orlando (2022) documents

Floor space per acre, F, is commonly known as FAR (the floor-area ratio), and the developer can be depicted as choosing it rather than S, which is more convenient for our purposes. Making a change of variable from S to F yields $S = f^{-1}(F)$. The developer's capital cost per acre can then be written $S = f^{-1}(F) \equiv C(F)$. Profit per acre is then $pF - C(F)$, and the first-order condition is $p = C'(F)$. The two approaches using S and F are, of course, equivalent. If H takes a Cobb-Douglas form with capital exponent $\rho < 1$, then $f(S)$ is proportional to S^ρ and $C(F)$ is proportional to $F^{1/\rho} \equiv F^\gamma$, where $\gamma = 1/\rho > 1$.

For notational simplicity, let $\overline{F}$ instead of $\overline{\text{FAR}}$ denote the regulated FAR. Then suppose that the marginal cost $C'(F)$ is larger above $\overline{F}$ than below it, as would be the case if air rights need to be purchased to set F above $\overline{F}$. To capture this behavior, let $C(F)$ include a multiplicative factor that equals α/γ below $\overline{F}$ and $(\alpha + \beta)/\gamma$ above it, as follows:

$$C(F) \;=\; \begin{cases} \alpha F^\gamma/\gamma & \text{if } F \leq \overline{F} \\ (\alpha + \beta)F^\gamma/\gamma \;-\; \beta \overline{F}^\gamma/\gamma & \text{if } F > \overline{F}. \end{cases} \tag{1}$$

Note that the $\beta \overline{F}^\gamma/\gamma$ term in the second line of (1) ensures that $C(F)$ is continuous at $\overline{F}$. Using (1), the previous first-order condition $p = C'(F)$ for F can be written as

$$p \;=\; \begin{cases} \alpha F^\lambda & \text{if } F \leq \overline{F} \\ (\alpha + \beta)F^\lambda & \text{if } F > \overline{F}, \quad \text{where } \lambda \;\equiv\; \gamma - 1 \;=\; 1/\rho - 1 \;>\; 0. \end{cases} \tag{2}$$

Solving (2) for F gives

$$F \;=\; \begin{cases} (p/\alpha)^{1/\lambda} & \text{if } p \leq \alpha \overline{F}^\lambda \equiv p^* \\ (p/(\alpha + \beta))^{1/\lambda} & \text{if } p > (\alpha + \beta)\overline{F}^\lambda \equiv p^{**}. \end{cases} \tag{3}$$

While (2) and (3) apply when the developer purchases air rights to set F above $\overline{F}$, the appendix shows that the same first-order conditions apply when the developer secures a higher FAR limit by creating public open space.

---

nonsmooth marginal costs at the fourth and eighth floors in tall buildings. Addressing such jumps would complicate the present analysis, but (as seen below) our sample buildings tend to have fewer than 4 floors, which is below the problematic range.

2.2. Bunching

Inspection of (2) and (3) shows that the optimal F equals $\overline{F}$ when $p = \alpha\overline{F}^{\lambda} \equiv p^{*}$ and also when $p = (\alpha + \beta)\overline{F}^{\lambda} \equiv p^{**}$. For p values in the range (p^{*}, p^{**}), the optimal F also equals $\overline{F}$, although neither of the tangency conditions in (2) and (3) are satisfied. In this case, for a range of p values, the highest profit is reached at the kink in the C function without a tangency occurring, leading to the bunching of optimal F values at $\overline{F}$. Figure 1 provides an illustration, with the two dotted lines in the figure, having slopes p^{*} and p^{**}, being tangent to the two separate portions of $C(F)$ at the kink. A line with an intermediate slope would touch the kink without a tangency.

We consider a large group of land parcels sharing a common $\overline{F}$, but the price p is assumed to differ across the parcels in a manner described by the density function $t(p)$. Price differences across parcels would be caused by differences in neighborhood characteristics, including access to subway stations and parks or other amenities.[11] Optimal values of F thus differ across parcels depending on the relevant p values. Since F is the choice variable, we need to derive the density of F over the relevant ranges.[12] With $p = \alpha F^{\lambda}$ holding below $\overline{F}$, the density of F in this range can be derived by using the change-of-variable formula on p's density $t(p)$. Doing so, the density of F for $F < \overline{F}$ is given by

$$t(\alpha F^{\lambda})\alpha\lambda F^{\lambda-1} \equiv h_0(F), \tag{4}$$

where h_0 denotes F's density in this range. Note that p as a function of F is substituted in t, with result multiplied by the derivative of this relationship. The resulting transformation changes the density's scale on the horizontal axis as well as its height.

Similarly, to find the density of F for $F \geq \overline{F}$, $p = (\alpha + \beta)\overline{F}^{\lambda}$ is substituted into $t(p)$,

---

[11] While large differences in p across neighborhoods would presumably lead the city to set different $\overline{F}$ values, contrary to the assumption of a common, fixed $\overline{F}$, smaller variations in p (captured by $t(p)$) are likely to occur with a given $\overline{F}$ zone. Note that the assumption of price-taking behavior by developers is viable even though p is neighborhood specific. The overall price level is parametric to developers, with deviations in p relative to this level serving to equalize utility across neighborhoods in the presence of amenity differences.

[12] Note that in our model, the price p is analogous to the unobserved ability n in Saez (2010), while F is analogous to Saez's before-tax income z, the worker's choice variable.

yielding

$$t((\alpha + \beta)F^\lambda)(\alpha + \beta)\lambda F^{\lambda - 1} \equiv h_1(F), \tag{5}$$

where h_1 denotes F's density in this range.

To relate all this information to the extent of bunching at $\overline{F}$, recall that developers facing p values in the interval $[p^*, p^{**}] = [\alpha \overline{F}^\lambda, (\alpha + \beta)\overline{F}^\lambda]$ bunch at $\overline{F}$. Intuitively, for a given λ, the range of p values leading to bunching, and thus the number of developers who bunch, is larger the greater is the ratio $(\alpha + \beta)/\alpha$ and thus the larger is the marginal-cost penalty for exceeding $\overline{F}$.

We can derive the size of this group of bunching developers using the density $h_0(F)$, which applies in the range below $\overline{F}$. To do so, suppose for a moment that the marginal-cost kink at $\overline{F}$ did not exist, with the density $h_0(F)$ applying for all F values. To use this density, note that the developer with $p = p^{**} = (\alpha + \beta)\overline{F}^\lambda$ facing the marginal-cost factor α would choose $F = [(\alpha + \beta)/\alpha]^{1/\lambda}\overline{F}$, as can be seen from the first line of (3). Therefore, the number of developers in the $[p^*, p^{**}]$ interval would be the number of developers choosing F between $\overline{F}$ and $F = ((\alpha + \beta)/\alpha)^{1/\lambda}\overline{F}$ in the absence of the marginal-cost kink. Let the last expression be written as $\overline{F} + \Delta\overline{F}$, where

$$\Delta\overline{F} = \left[\left(\frac{\alpha + \beta)}{\alpha} \right)^{1/\lambda} - 1 \right] \overline{F}. \tag{6}$$

Then, the number of bunchers B is equal to the integral of the density $h_0(F)$, which applies in the absence of the kink, between these two values:

$$B = \int_{\overline{F}}^{\overline{F} + \Delta\overline{F}} h_0(z)dz. \tag{7}$$

This integral can be approximated by the area of a trapezoid with its corners on $h_0(F)$ at the limits of integration, yielding

$$B = \Delta\overline{F} \frac{h_0(\overline{F}) + h_0(\overline{F} + \Delta\overline{F})}{2}. \tag{8}$$

7

The problem, though, in operationalizing (8) is that we do not observe $h_0(\overline{F} + \Delta\overline{F})$, given that h_0 in the presence of the marginal-cost kink only applies up to $\overline{F}$ and not above it. But we can use the relationship between h_0 and h_1 implied by (4) and (5) to replace $h_0(\overline{F} + \Delta\overline{F})$ by expression involving the (observable) density h_1, circumventing this obstacle. As shown in the appendix, doing so yields

$$h_0(\overline{F} + \Delta\overline{F}) \;=\; \left(\frac{\alpha + \beta}{\alpha}\right)^{-1/\lambda} h_1(\overline{F}). \tag{9}$$

Substituting (9) into (8), while recalling the definition of $\Delta\overline{F}$ in (6), (7) can be written as[13]

$$
\begin{aligned}
B \;&=\; \Delta\overline{F}\,\frac{h_0(\overline{F}) \;+\; ((\alpha + \beta)/\alpha)^{-1/\lambda}\, h_1(\overline{F})}{2} \\
&=\; (\theta - 1)\overline{F}\,\frac{h_0(\overline{F}) \;+\; (1/\theta)\, h_1(\overline{F})}{2},
\end{aligned}
\tag{10}
$$

where

$$\theta \;\equiv\; \left(\frac{\alpha + \beta}{\alpha}\right)^{1/\lambda}. \tag{11}$$

Note that θ equals the ratio of the marginal-cost parameters just above and below $C(F)$'s kink, which is then raised to a power equal to the price elasticity of F, equal to $1/\lambda$ from above.

Since B, $h_0(\overline{F})$, and $h_1(\overline{F})$ can be measured in the data, (9) can be used to solve for θ. With a θ estimate, denoted $\widehat{\theta}$, in hand, the marginal-cost ratio $(\alpha + \beta)/\alpha$ can be estimated via

$$(\alpha \widehat{+\beta})/\alpha \;=\; \widehat{\theta}^{\,\lambda}, \tag{12}$$

using an independent value of λ.

As explained in Saez (2010), the elements in (10) can be generated by creating three FAR intervals around $\overline{F}$, defined by a width factor δ. One interval is centered at $\overline{F}$, consisting of

---

[13] Eq. (10) corresponds to eq. (5) in Saez (2010). In Saez's case, $\theta = [(1 - t_0)/(1 - t_1)]^e$, where t_0 and t_1 are the income-tax rates below and above the kink in the net-of-tax earnings schedule and e is the compensated elasticity of earnings with respect to 1 minus the tax rate.

observed FAR values satisfying $F \in [\overline{F} - \delta, \overline{F} + \delta]$. Two additional intervals lie just below and just above this interval, consisting of FAR values satisfying $F \in [\overline{F} - 2\delta, \overline{F} - \delta)$ and $F \in (\overline{F} + \delta, \overline{F} + 2\delta]$. These intervals yield an estimate of the extent of bunching, captured by the excess mass in the middle interval relative to the masses in the two outer intervals (mass being the number of F observations). In addition, the latter masses allow estimates of $h_0(\overline{F})$ and $h_1(\overline{F})$.

Specifically, let N denote the number of FAR observations in the central interval, H_- denote the number of observations in the lower outer interval, and H_+ denote the number of observations in the upper outer interval. Then the estimated magnitude of bunching equals $\widehat{B} = N - H_- - H_+$. The value of $h_0(\overline{F})$ in (10) is estimated by the average height of the density in the range $(\overline{F} - 2\delta, \overline{F} - \delta)$, or $h_- = H_-/\delta$. Similarly, the value of $h_1(\overline{F})$ is estimated by average height of the density in the range $(\overline{F} + \delta, \overline{F} + 2\delta)$, or $h_+ = H_+/\delta$. In both cases, the average density height equals the number of FAR observations in each interval divided by its width. This setup is illustrated in Figure 2, which pertains to parcels with $\overline{\text{FAR}} = 0.6$ (see below). Note that the distance between the dashed vertical lines equals δ. The areas H_- and H_+ are not explicitly labeled in the figure, but the density heights h_- and h_+ are labeled.

The estimated values, as seen in the figure, are substituted into (10), and the equation is then solved to yield $\widehat{\theta}$, the estimate of θ. Note that after multiplying through by θ, (10) becomes a quadratic equation in θ, which can be solved by the quadratic formula. Given $\widehat{\theta}$, we use the existing estimate of λ from Combes, Duranton and Gobillon (2021) to yield $\widehat{\theta}^\lambda$, the estimate of the marginal-cost ratio from (12). That paper provides a recent and reliable estimate of the parameters of the Cobb-Douglas housing production function based on French data.[14] Their estimate of the capital exponent for Paris, a large city comparable to NYC, is 0.54. Using the second line of (2), this exponent translates into a λ value of $(1/0.54) - 1 = 0.85$.

Bootstrapping, based on repeated sampling with replacement from the data set (using 500 draws), generates a mean $\widehat{\theta}$ value along with a standard error and confidence interval. In addition, our θ estimate, and hence the estimate of $(\alpha + \beta)/\alpha$, obviously depend on the value

---

[14] While the paper uses data on single-family houses rather than taller buildings to estimate the production function, the estimates appear to be the most reliable ones available.

of the interval parameter δ, and this dependence can be appraised via sensitivity analysis.

3. Data

The Primary Land Use Tax Lot Output (PLUTO) dataset (Release 17v1.1) provided by the New York City Department of City Planning is used as our analysis dataset. The sample year is 2017. PLUTO contains information on the physical dimensions of the tax lot and the building(s) that sit on the lot, the economic uses of the tax lot, the year when the building was built along with the years of the last major alterations, the zoning designation(s) that pertain to the lot, and the assessed value of the property and land, amongst other fields. The initial dataset has 859,223 observations. The geography of New York City is spread across the five boroughs: Brooklyn (32.29% of tax lots), Bronx (10.46%), Manhattan (5%), Queens (37.77%), and Staten Island (14.46%).

We apply several filters to the data to ensure a clean focus. First, we drop commercially zoned and industrially zoned buildings from our sample, focusing only on residential buildings. Next, we drop any observation that has more than one zoning designation, is zoned as a special purpose district, a historical district, a limited-height district, or where no zoning designation or year built is indicated. We also drop any lot that does not contain a single building. These filters exclude 428,142 observations, or 49.82% of the initial sample. We also filter out zoning designations for which we either do not have adequate information on height restrictions, or whose rules are governed by "height-factor" zoning regulations, which confound our analysis.[15]

We then generate values for FAR limits using the PLUTO's "Zoning Data Tables," keeping only FAR values that pertain to at least 1,000 tax lots.[16] This final filter results in five

---

[15] Height-factor zoning regulations confound our analysis, since the developer can increase the FAR limit by providing more open space, building a taller skinnier building, as under the POPS program. The problem is that FAR varies with open space according to a complicated formula, with no regular FAR limit stated. To further confound the analysis of these properties, in 1987 "The Quality Housing Program" (QHP) was initiated, and developers were to choose between height-factor zoning and the QHP, which imposed a number of design restrictions while allowing for wider buildings more in line with the historic character of the neighborhoods. These filters remove an additional 65,680 observations, or 7.64% of the initial sample.

[16] As explained in detail in Peng (2023), NYC instituted a land-use reform over the 2002-2013 period that raised regulated FAR values in many locations, with an average increase of 23% (from an average value of 1.6 to 2.34, although values for some parcels actually decreased). Since $\overline{FAR}$ values were changing over this period for many parcels, we assigned each parcel in our sample the $\overline{FAR}$ prevailing in its construction year (2000 or later). Because of data limitations, however, parcels with construction occurring in 2001-2002 were assigned the 2000 $\overline{FAR}$. Assigning all parcels their 2017 $\overline{FAR}$ values could have introduced measurement error if the

subsamples with FAR limits of 0.5, 0.6, 0.9, 1.25, and 2.0.[17] The final analysis dataset consists of 23,369 properties, or roughly 2.72% of the initial sample. Of these properties, 2,656 (11.37%) are in the Bronx, 4,580 (19.6%) are in Brooklyn, 9,247 (39.56%) are in Queens, and 6,885 (39.56%) are in Staten Island. Only a single parcel lies in Manhattan. Compared to the initial sample, our analysis sample is more heavily weighted to the boroughs of Staten Island, the Bronx, and Queens, and much less representative of Brooklyn and especially Manhattan.

4. Results

4.1. Summary statistics

As just mentioned, the data we use consist of parcels with five different $\overline{F}$ values, as shown in Table 1 along with the number of observations for each group and the zoning categories they contain. The table also shows the average number of floors in each group, as well as the average value of the floors/FAR ratio. These numbers show that the parcels in our sample do not contain the very tall buildings for which NYC is famous. The number of observations in those categories is just too small for the application of our method. In addition, recall that, if a building fully covers its lot, FAR is equal to the number of floors. With the average floors/FAR ratio exceeding 1, it follows that, on average, buildings in our sample do not cover their lots. For the FAR $= 1.25$ and 2.00 groups, the ratio value of approximately 2 indicates that lot coverage is around 50% for these groups, with the fraction lower in the groups with smaller FAR values (and higher ratios).

Figures 3-7 show FAR histograms for the 5 different groups, with the group's $\overline{\text{FAR}}$ value shown by the vertical line. It is important to note that, with more than 1000 observations in each group, these histograms do not capture all the detail in the FAR distributions. However, they are easier to read than more disaggregated histograms while adequately capturing the bunching patterns around the $\overline{\text{FAR}}$ values, which tend to be prominent. Note that the vertical line is sometimes just to the left rather to the right of the modal FAR, which could reflect FAR measurement error in the data. However, since values in a range around $\overline{\text{FAR}}$ are counted as

---

reform raised $\overline{\text{FAR}}$ after the construction year, making the recorded $\overline{\text{FAR}}$ exceed the one the developer faced.

[17] Note that the price density $t(p)$ from the theoretical analysis is likely to differ across these $\overline{\text{FAR}}$ groups, implicitly depending on $\overline{F}$.

bunched, this discrepancy has no effect on our computations. Figure 8 shows a map of the sample parcel locations, which must be read on the screen. As can be seen, parcels are lacking in Manhattan, which partly accounts for the absence, on average, of tall buildings. However, note that the $\overline{F} = 2.0$ group is closest to Manhattan, accounting for its taller buildings relative to other groups.

4.2. Main results

Table 2 shows the θ and $(\alpha + \beta)/\alpha$ estimates for the different FAR groups along with the assumed δ values used in their computation. The bootstrap confidence intervals for θ as well as the mean $\widehat{\theta}$ from the bootstrap are shown, along with the implied confidence intervals for the marginal-cost ratio $(\alpha + \beta)/\alpha$. It is important to note that the δ interval value for each $\overline{\text{FAR}}$ group is chosen by careful examination of a detailed histogram of the FAR distribution for the group, like those in Figures 3-7, so as to reasonably capture the bunching area. The sensitivity analysis presented below shows how the $\widehat{\theta}$ values are affected by the δ choices.

Consider first the $\overline{\text{FAR}} = 2.0$ group. The $\widehat{\theta}$ value equals 1.090, which yields a $\widehat{(\alpha + \beta)}/\alpha$ value of 1.077, indicating that the marginal cost of additional floor space is about 8% higher above $\overline{\text{FAR}}$ than below it, a moderate cost penalty. At 1.091, the average bootstrap $\widehat{\theta}$ is very close to the sample θ estimate, indicating little bias from the nonlinearity of the estimation, and the distribution of the bootstrap $\widehat{\theta}$ values around this mean is fairly symmetric, as shown in Figure 9. The 95% confidence interval for θ, also shown in the table, ranges between 1.060 and 1.129, while the implied confidence interval for the marginal-cost ratio $(\alpha + \beta)/\alpha$ ranges between 1.051 and 1.109.

Turning to the $\overline{\text{FAR}} = 1.25$ group, the $\widehat{\theta}$ value of 1.032, and the associated $\widehat{(\alpha + \beta)}/\alpha$ value of 1.027, indicate that the cost penalty above $\overline{\text{FAR}}$ is now smaller, at only about 3%. The mean bootstrap $\widehat{\theta}$ is again close to the sample estimate, and the $\widehat{\theta}$ distribution (not shown) is again quite symmetric. Confidence intervals are again shown.

Results for the $\overline{\text{FAR}} = 0.9$ and 0.6 groups are similar, with the $\widehat{\theta}$ estimates of 1.076 and 1.062, respectively. The associated $\widehat{(\alpha + \beta)}/\alpha$ values are 1.064 and 1.052, yielding cost premia of 5–6%. The average bootstrap $\widehat{\theta}$ values are again very close to the sample estimates, and the $\widehat{\theta}$ distributions are symmetric.

The $\overline{\text{FAR}} = 0.5$ group has a larger $\widehat{\theta}$ like that of the 2.0 group. It equals $\widehat{\theta} = 1.117$, yielding an $\widehat{(\alpha + \beta)}/\alpha$ estimate of 1.099 and a cost penalty of 10%. The other previous features of the bootstrap results remain present.

Recall that the extent of bunching (the number of developers with p values between p^* and p^{**}) depends on the magnitude of the marginal-cost penalty for exceeding $\overline{F}$, given by $(\alpha + \beta)/\alpha$. This relationship is evident in the F histograms in Figures 3-7. The estimated values of the penalty are largest (8–10%) in the $\overline{F} = 2.0$ and $\overline{F} = 0.5$ cases, and inspection of the figures shows that bunching appears to be most prominent in these cases, confirming the expected association. However, the smaller sizes of the other penalities (3-6%) and the absence of truly large penalities in any of the $\overline{\text{FAR}}$ groups may reflect the absence of truly dramatic bunching patterns, like those often seen in tax-related studies.

The results in Table 2 thus show that a marginal-cost penalty exists above $\overline{\text{FAR}}$ in each of the groups, roughly ranging between 3% and 10%. Moreover, in all groups, the $\widehat{\theta}$ confidence intervals never cover $\theta = 1.0$, which would indicate the absence of a penalty. Therefore, our findings confirm that, if they pay an additional cost, developers in NYC can build above the regulated FAR value for their parcel.

The connection between these results and those presented by Brueckner and Singh (2020) deserves discussion. That paper estimated the stringency of FAR regulation for five US cities, including NYC. The procedure was to regress the log of land value for vacant parcels in a particular city on the log of $\overline{\text{FAR}}$ for the parcel, with a high coefficient indicating a large gap between the free-market and regulated FARs and thus stringent regulation. By assuming a value for the Cobb-Douglas capital exponent, the ratio of free-market to regulated FAR can be computed, and with an exponent value of 0.6, the ratio equals 0.77 for NYC. With such a substantial gap and the relatively low cost of exceeding $\overline{\text{FAR}}$ that we have estimated, there would appear to be substantial impetus to make the expenditures needed to do so. Perhaps the large amount of mass above $\overline{\text{FAR}}$ in the histograms in Figures 3-7 is evidence of such expenditures being widely undertaken.

4.3. Sensitivity analysis

As discussed above, the $\widehat{\theta}$ estimate for an $\overline{\text{FAR}}$ group, and the implied estimate of the

marginal-cost ratio, depend on the assumed value of the interval parameter δ. Table 3 shows sensitivity analysis, with the first line within each group showing our assumed δ value and $\widehat{\theta}$ values from Table 2, and the second and third lines showing the $\widehat{\theta}$ estimates using smaller and larger δ values. As can be seen, the $\widehat{\theta}$ estimates vary somewhat with the value of δ. But the only striking change occurs in the $\overline{\text{FAR}} = 0.9$ group, where raising δ from the assumed value of 0.15 to 0.175 increases $\widehat{\theta}$ from 1.076 all the way to 1.232. Overall, the sensitivity analysis does not change the conclusion that building above $\overline{\text{FAR}}$ is costly for developers.

5. Gains in housing floor space from eliminating FAR regulation

As explained in the introduction, FAR regulation reduces the amount of housing that can be produced in NYC, a city that many observers view as under-supplying housing floor space. Our analysis allows us to compute the extra housing floor space that could be gained by eliminating the regulation.[18] Since the exercise assumes that existing buildings would be replaced after FAR deregulation, the computed impacts should be viewed as long-run effects. In addition, the exercise ignores any price effects from greater housing supply, as discussed further below.

5.1. Analysis based on the theoretical model

To understand how our procedure works, focus on the theoretical model of section 2, and consider first a developer who chooses some $F = F_{now} > \overline{F}$ under the current limitations ("now" denotes current). From (2), the price p faced by this developer, denoted $\widetilde{p}$, satisfies $\widetilde{p} = (\alpha + \beta)F_{now}^{\lambda}$. Inverting this relationship (as in (3)) to find F_{now} yields $F_{now} = (\widetilde{p}/(\alpha + \beta))^{1/\lambda} > \overline{F}$. With FAR regulation eliminated, the $\alpha + \beta$ factor becomes α, indicating a lower marginal cost in the FAR range above the old $\overline{F}$. The developer would then choose a new and larger F value satisfying $\widetilde{p} = \alpha F_{new}^{\lambda}$, or $F_{new} = (\widetilde{p}/\alpha)^{1/\lambda} > F_{now}$. The ratio between the new

---

[18] Peng (2023) examines the effects on NYC floor space of the land-use reform described in footnote 16 above. The paper estimates a dynamic equilibrium model where forward-looking developers face fixed costs to redevelop a parcel of land, which allows for a slow readjustment of floor space supply after the relaxation in FAR limits. After estimating the model, the paper performs a long-run counterfactual analysis and finds an increase in floor space by the year 2050 of around 0.7%. We find similar increases in floor space from a marginal increase in $\overline{\text{FAR}}$, as explained below. Leather (2023) carries out a related exercise, using a data set of which the current one is a subset.

and current FAR values is then given by

$$\frac{F_{new}}{F_{now}} \;=\; \frac{(\widetilde{p}/\alpha)^{1/\lambda}}{(\widetilde{p}/(\alpha+\beta))^{1/\lambda}} \;=\; \left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda} \;=\; \theta. \tag{13}$$

The chosen FAR thus rises by the factor θ with elimination of FAR regulation, so that $F_{new} = \theta F_{now}$. Using this relationship and recalling that F equals floor space per acre of land, the new supply of floor space from the developer's lot is given by

$$\ell \times F_{new} \;=\; \ell \theta F_{now}, \tag{14}$$

where ℓ is the lot size. Thus, new floor space on the parcel equals lot size times θ times the current FAR. By contrast, the current amount of floor space equals ℓF_{now}. The expressions ℓF_{now} and $\ell \theta F_{now}$ can be summed across all parcels with $F_{now} > \overline{F}$ to get total floor space above $\overline{F}$, both before and after the elimination of FAR regulation.[19]

Turning to bunching developers, recall that in the theoretical model, developers who bunch at $\overline{F}$ have p values in the range $[p^*, p^{**}] = [\alpha\overline{F}^\lambda, (\alpha+\beta)\overline{F}^\lambda]$. In the absence of FAR regulation, the developer facing $p^* = \alpha\overline{F}^\lambda$ would choose $F = \overline{F}$, using the first line of (3). Similarly, a developer facing $p^{**} = (\alpha + \beta)\overline{F}^\lambda$ would choose $F = ((\alpha + \beta)/\alpha)^{1/\lambda}\overline{F} = \theta\overline{F}$, again using the first line of (3). Approximating based on an average of these endpoint values, bunching developers on average would thus choose F equal to $(\overline{F} + \theta\overline{F})/2 = (1/2)(1 + \theta)\overline{F} > \overline{F}$.

Therefore, for current bunchers, the average floor space in the absence of FAR regulation would equal $\ell \times (1/2)(1 + \theta) \times \overline{F}$, assuming ℓ is the same for all bunchers. The current floor space for bunchers equals $\ell\overline{F}$. Both these expressions would be summed across bunchers to get the total floor space with and without FAR regulation. Finally, for parcels with $F_{now} < \overline{F}$, FAR deregulation yields no change in floor space.

5.2. Taking the theory to the data

In taking the previous theoretical formulas to the data, we must recognize that, unlike in the model, bunching is viewed as occurring over a range of F values instead of at the single

---

[19] Note that the floor-space gain depends on θ, which equals the marginal-cost ratio raised to a power, not directly on the ratio itself.

point $\overline{F}$. In the estimation, the bunching range has been set at $[\overline{F} - \delta, \overline{F} + \delta]$. As a result, we view developers who breach the FAR limit as those who set F above $\overline{F} + \delta$. Developers who bunch are those in the bunching interval, but to be consistent with the model, we treat these developers as choosing $F = \overline{F}$ rather than the values near $\overline{F}$ that they actually choose. With these amendments, both the total existing floor space and the floor space that would be produced in the absence of FAR regulation can be computed for each of the five $\overline{F}$ groups.

Letting i denote the parcel and $\mathbf{1}$ denote an indicator function, total current floor space for a given $\overline{F}$ group is given by[20]

$$SPACE_{now} \;=\; \sum_i \; \Big\{ \ell_i \, F_i \times \mathbf{1}[F_i < \overline{F} - \delta] \;\;+\;\; \ell_i \, \overline{F} \times \mathbf{1}[\overline{F} - \delta < F_i < \overline{F} + \delta]$$

$$+\;\; \ell_i \, F_i \times \mathbf{1}[F_i > \overline{F} + \delta] \Big\}. \tag{15}$$

Recall that, as explained above, a parcel's floor space in the bunching range is set equal to $\ell_i \overline{F}$ rather than $\ell_i F_i$.[21]

Total floor space without FAR regulation is given by

$$SPACE_{new} \;=\; \sum_i \; \Big\{ \ell_i \, F_i \times \mathbf{1}[F_i < \overline{F} - \delta] \;\;+\;\; 0.5(1 + \theta) \, \ell_i \, \overline{F} \times \mathbf{1}[\overline{F} - \delta < F_i < \overline{F} + \delta]$$

$$+\;\; \theta \, \ell_i \, F_i \times \mathbf{1}[F_i > \overline{F} + \delta] \Big\}. \tag{16}$$

Note that, following the discussion above, floor space in the bunching range is inflated by the factor $0.5(1 + \theta)$, while floor space above $\overline{F} + \delta$ is inflated by the factor θ. Floor space below $\overline{F} - \delta$ is unchanged.[22]

The results of computing (15) and (16) for each of the $\overline{FAR}$ groups are shown in Table 4. The largest percentage gains in floor space from the elimination of FAR regulation are in

---

[20] While lot size for bunchers was, for simplicity, assumed equal in the theoretical discussion above, (14) allows it to differ across parcels

[21] This choice will have little effect given that the average of F_i over the bunching range will be close to $\overline{F}$.

[22] The percentages of parcels contained in the bunching intervals $[\overline{F} - \delta, \overline{F} + \delta]$ equal 20, 21 and 24% for the 1.25, 2.0, and 0.6 $\overline{FAR}$ groups, respectively, but are larger for the remaining groups, equal to 31 and 38% for the 0.5 and 0.9 $\overline{FAR}$ groups, respectively.

the $\overline{\text{FAR}} = 2.0$ and 0.5 groups, where gains are 6.3% and 8.1% respectively. The gains for the other groups are smaller, in the 1-5% range. The results show that, even focusing on $\overline{\text{FAR}}$ groups containing buildings that are relatively short by NYC standards, the floor-space gains from the removal of FAR regulation can be appreciable.

The extra floor generated by FAR deregulation would put downward pressure on the prices p per square foot across all parcels. These price declines would in turn reduce the F values that developers would choose. As a result, the floor-space gains in Table 4 should be interpreted as upper bounds on the predicted gains. Moreover, since any floor-space gains would only be realized in the long run as buildings are replaced, such a distant horizon would provide more room for secular price changes that could also affect chosen FARs and thus changes in floor space.

5.3. Floor-space gains from a marginal or non-marginal increase in $\overline{F}$

The online appendix shows how to calculate the change in floor space for a marginal increase in $\overline{F}$ rather than for a complete elimination of FAR regulation, again relying on the theoretical model. As above, there are multiple groups to consider. The calculation remains complicated, but in the end, it leads to an intuitive conclusion: the rate of increase of total floor space when $\overline{F}$ increases by an infinitesimal amount ϵ is just equal to the total amount of space bunched at the original $\overline{F}$. F for each of these buildings increases by ϵ, so that they become bunched at the marginally higher $\overline{F}$. Dividing the FAR gain by ϵ to evaluate the derivative, the result is simply B, the original amount of bunching (multiplying by lot size gives the floor-space gain).

The online appendix also generates formulas for a non-marginal increase in $\overline{F}$, denoted ε, which raises the $\overline{\text{FAR}}$ limit to $\overline{F} + \varepsilon$. The analysis develops formulas for the case where some of the original bunchers continue to bunch at the higher $\overline{F}$, which requires that the old upper bound p^{**} of the bunching p-range (equal to $(\alpha + \beta)\overline{F}^\lambda$) is larger than the lower bound of the new bunching p-range under the higher $\overline{F}$, equal to $\alpha(\overline{F} + \varepsilon)^\lambda$. However, the effect of the $\overline{F}$ increase is most easily evaluated when these two quantities are equal (so that no old bunchers become new bunchers), which reduces to the requirement that $\overline{F} + \varepsilon$ equals $\theta\overline{F}$, using (11). In this case, the online appendix shows that the incremental floor space from the larger $\overline{F}$ is

given by the average lot size times

$$\frac{((\theta - 1)\overline{F})^2}{2} \frac{1 + \theta}{\theta} h_1(\overline{F}).$$ (17)

Replacing $h_1(\overline{F})$ in (17) by its sample equivalent h_+ and evaluating the expression for each of the $\overline{\text{FAR}}$ groups yields the results shown in Table A1 in the online appendix. The table shows the elasticity of predicted incremental floor space with respect to the non-marginal increase in $\overline{\text{FAR}}$. The incremental floor space is highly variable across the groups, leading to a range of elasticities, most of which are quite small (the largest elasticity is the 0.08 value for the $\overline{F} = 2.0$ group).[23] The reasons for this variability are that $(\theta - 1)^2$ is very small in the groups where θ is closest to 1.0 and that the h_+ values vary considerably (by as much as a factor of 60) across the groups. Nevertheless, the results in Table A1 show that computing the floor-space gain from a non-marginal increase in $\overline{F}$ is feasible.

6. Conclusion

This paper has presented a real-estate application of the bunching methodology widely used in other areas of applied microeonomics. Our results show that the marginal cost of floor space increases modestly when the developer exceeds a parcel's regulated building height. We use our estimates to predict the additional floor space that be created by a complete removal of building-height regulation in NYC. This last exercise is circumscribed by our focus on a limited number of zoning categories, but the results suggest that New York could secure notably more housing through lighter regulation.

Our method could be applied in other real-estate contexts. For example, minimum parking requirements (MPRs) impose costly restrictions on developers throughout the US, forcing provision of off-street parking to prevent congestion in street parking around a new development (see Cutter and Franco, 2012). If developers are able to go below the specified MPR for a parcel (thus providing fewer parking spaces) by incurring some additional costs, the bunching method could be used to estimate their magnitude, assuming that suitable data were available.

---

[23] The absolute increase in floor space in the 2.0 $\overline{\text{FAR}}$ group is 0.8% of the original value, a number very close to Peng's (2023) 0.7% percentage increase from the reform she studies, which was also marginal in nature.

The paper's method could also be applied to FAR regulation in other cities. Although analogs to the Privately Owned Public Spaces and air-rights purchase programs, which allow increments to FAR in NYC, may not be available elsewhere, other cities presumably provide alternative avenues for exceeding their height limits.

Finally, while this paper has for simplicity used the basic bunching methodology of Saez (2010), progress has been made since then in refining and strengthening the method. These new developments are surveyed by Bertanha, Caetano, Jales and Seegert (2023), and their application to our NYC data would be a useful extension of the present research.

Appendix

A1. The cost effect of public open space

Suppose that the developer devotes a fraction $1 - \phi$ of his lot to public open space in order to secure a more generous FAR limit, denoted $\overline{F}_{os} > \overline{F}$. Profit exclusive of land cost under the Cobb-Douglas assumption then equals $pK^\rho(\phi\ell)^{1-\rho} - K$, with land under the building equal to $\phi\ell$. Dividing $K^\rho(\phi\ell)^{1-\rho}$ by ℓ, FAR based on the entire lot area is thus given by $F = \phi^{1-\rho}S^\rho$, where S again equals K/ℓ. As a result, $S = (F/\phi^{1-\rho})^{1/\rho} \equiv C(F)$, and cost then equals $C(F) = \mu F^\phi$, where $\phi = 1/\rho$ and $\mu \equiv \phi^{(1-\rho)/\rho} > 1$.

Thus, cost exceeds F^ϕ, the $C(F)$ expression in the absence of open space, for all allowable values of F when open space is provided. This outcome contrasts with the air-rights case, where extra costs are incurred only above $\overline{F}$. Despite this difference, the first-order condition for choice of F will involve a multiplicative factor (1 vs. μ) that jumps to a higher value above $\overline{F}$, just as in (2). Crucially, if his chosen F is below $\overline{F}$, the developer will not provide open space, so that the lower cost function, and its smaller multiplicative factor of $1 < \mu$, is relevant for the first-order condition over this range.

However, since the shift to the open-space regime that occurs at $\overline{F}$ applies to the entire range of F values, it generates a discontinous increase in cost, in contrast to the continuity of the cost function under the purchase of air rights, as captured in (1). As a result, the open-space regime generates a "notch" in the developer's profit function along with a change in its slope, which would require a different empirical method than the one we use (Kleven and Waseem, 2013). However, since a simple purchase of air rights appears to be an easier (and presumably much more common) path to exceeding $\overline{F}$ than provision of open space, our use of the kink rather than the notch methodology appears to be appropriate.

A2. Derivation of equation (8)

Eq. (8) is derived as follows:

$$
\begin{aligned}
h_0(\overline{F} + \Delta\overline{F}) &= t[\alpha(\overline{F} + \Delta\overline{F})^\lambda]\,\alpha\lambda(\overline{F} + \Delta\overline{F})^{\lambda-1} \\[2mm]
&= t\left[\alpha\left(\left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda}\overline{F}\right)^\lambda\right]\alpha\lambda\left[\left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda}\overline{F}\right]^{\lambda-1} \\[2mm]
&= t((\alpha+\beta)\overline{F}^\lambda)\,\alpha\lambda\left(\frac{\alpha+\beta}{\alpha}\right)^{(\lambda-1)/\lambda}\overline{F}^{\lambda-1} \\[2mm]
&= \left(\frac{\alpha+\beta}{\alpha}\right)^{-1/\lambda}t(\alpha+\beta)\overline{F}^\lambda)(\alpha+\beta)\lambda\overline{F}^{\lambda-1} \\[2mm]
&= \left(\frac{\alpha+\beta}{\alpha}\right)^{-1/\lambda}h_1(\overline{F}). \hspace{5cm} (a1)
\end{aligned}
$$

The first line of $(a1)$ uses (3), the second line uses the definition of $\Delta\overline{F}$ in (5), the third line simplifies the second line, the fourth line further simplifies the third line, and the last line uses the definition of h_1 in (4).

Table 1: FAR groups

$\overline{FAR}$	Observations	Floors	Floors/FAR	Zoning categories
2.0	2831	3.57	1.89	R5D, R6B
1.25	3772	2.86	2.26	R5
0.9	5499	2.42	2.58	R4, R4-1, R4-A, R4-B
0.6	9898	2.11	3.34	R3-1, R3-2, R3-A, R3-X
0.5	1369	2.02	3.84	R1-1, R1-2, R1-2A, R2, R2-A

The *Floors* and *Floors/FAR* columns of this table show the mean values of these variables for each $\overline{FAR}$ level.

Table 2: Estimated θ values and confidence intervals

$\overline{FAR}$	Observations	δ	$\widehat{\theta}$	Confidence int.	avg. $\widehat{\theta}$	$\widehat{(\alpha+\beta)/\alpha}$	Confidence int.
2.0	2831	0.15	1.090	[1.060, 1.129]	1.091	1.077	[1.051, 1.109]
1.25	3772	0.10	1.032	[1.013, 1.054]	1.033	1.027	[1.011, 1.046]
0.9	5499	0.15	1.076	[1.048, 1.107]	1.076	1.064	[1.041, 1.090]
0.6	9898	0.04	1.062	[1.049, 1.075]	1.062	1.052	[1.042, 1.063]
0.5	1369	0.04	1.117	[1.073, 1.175]	1.119	1.099	[1.062, 1.147]

$\widehat{\theta}$ is the estimated value of $((\alpha+\beta)/\alpha)^{1/\lambda}$, generated by assuming the given value of the interval parameter δ. The 95% confidence interval for θ is based on a standard error generated through a 10,000-draw boostrap procedure with replacement, and the average $\widehat{\theta}$ generated by the bootstrap is also shown. The marginal-cost ratio $(\alpha+\beta)/\alpha$ is estimated by $\widehat{\theta}^{\lambda}$, using $\lambda = 0.85$, and its confidence interval (based on the θ interval) is also shown.

Table 3: Sensitivity analysis

$\overline{FAR}$	δ	$\widehat{\theta}$
2.0	0.15	1.090
	0.10	1.038
	0.20	1.111
1.25	0.10	1.032
	0.05	1.044
	0.15	1.017
0.9	0.15	1.076
	0.125	1.025
	0.175	1.232
0.6	0.04	1.062
	0.03	1.038
	0.05	1.108
0.5	0.04	1.117
	0.03	1.070
	0.05	1.134

This table shows how $\widehat{\theta}$ varies
as the interval value δ changes.

Table 4: Floor space gain without FAR regulation

$\overline{FAR}$	$SPACE_{now}$	$SPACE_{new}$	% gain
2.0	17,109,266	18,187,679	6.3%
1.25	16,138,644	16,337,226	1.2%
0.9	16,231,509	17,074,061	5.2%
0.6	18,317,031	19,041,412	4.0%
0.5	4,092,694	4,425,844	8.1%

$SPACE_{now}$ is current total floor space in parcels
with the given $\overline{FAR}$ value, while $SPACE_{new}$ is
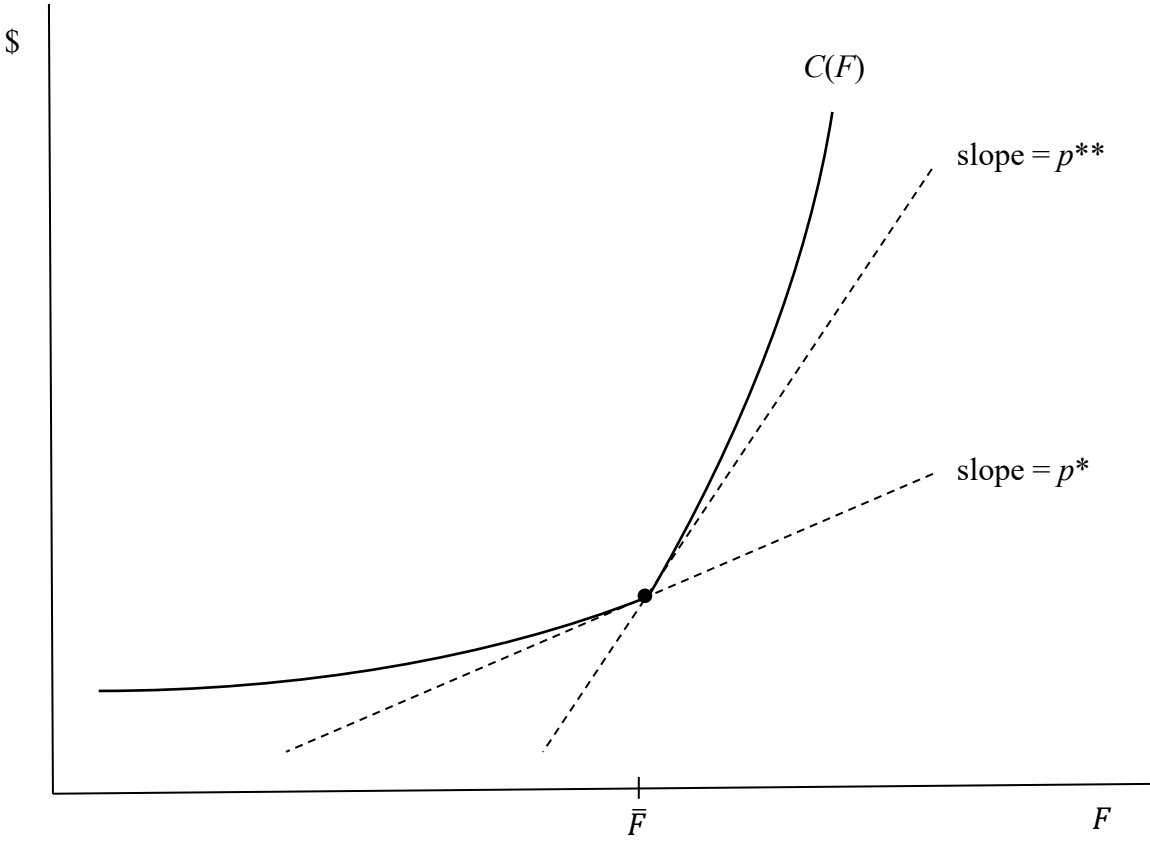predicted total floor space after removal of FAR
regulation.
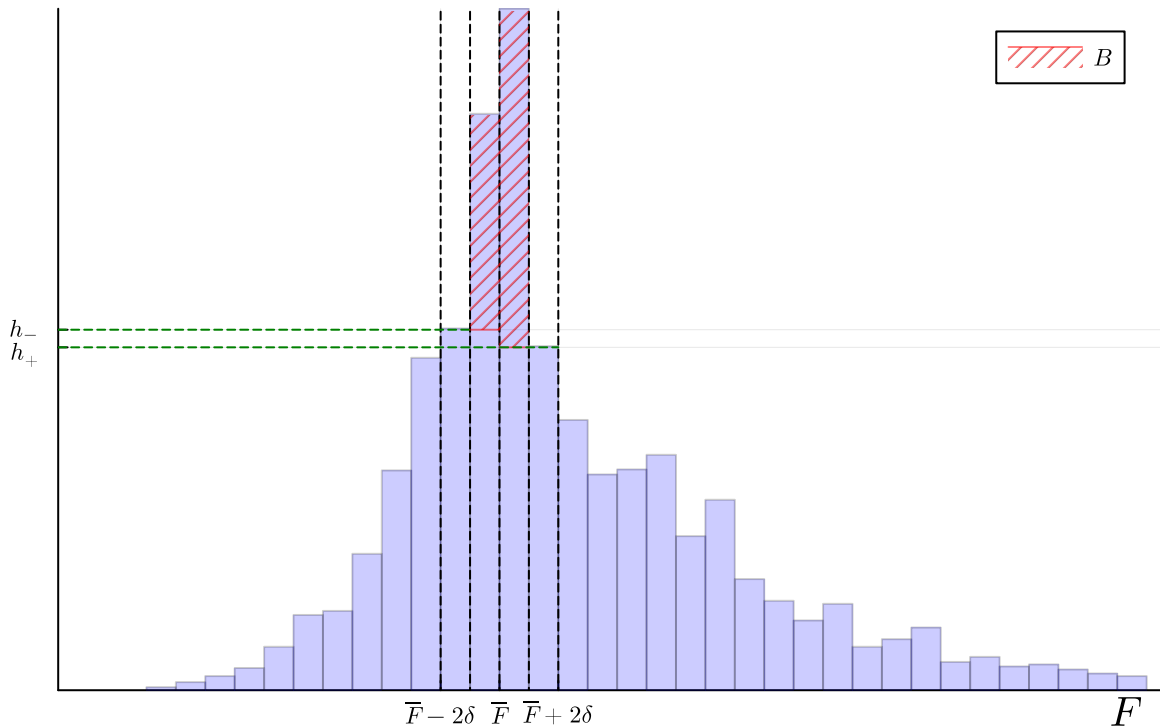
Figure 1: Kink in $C(F)$

Figure 2: Bunching Area

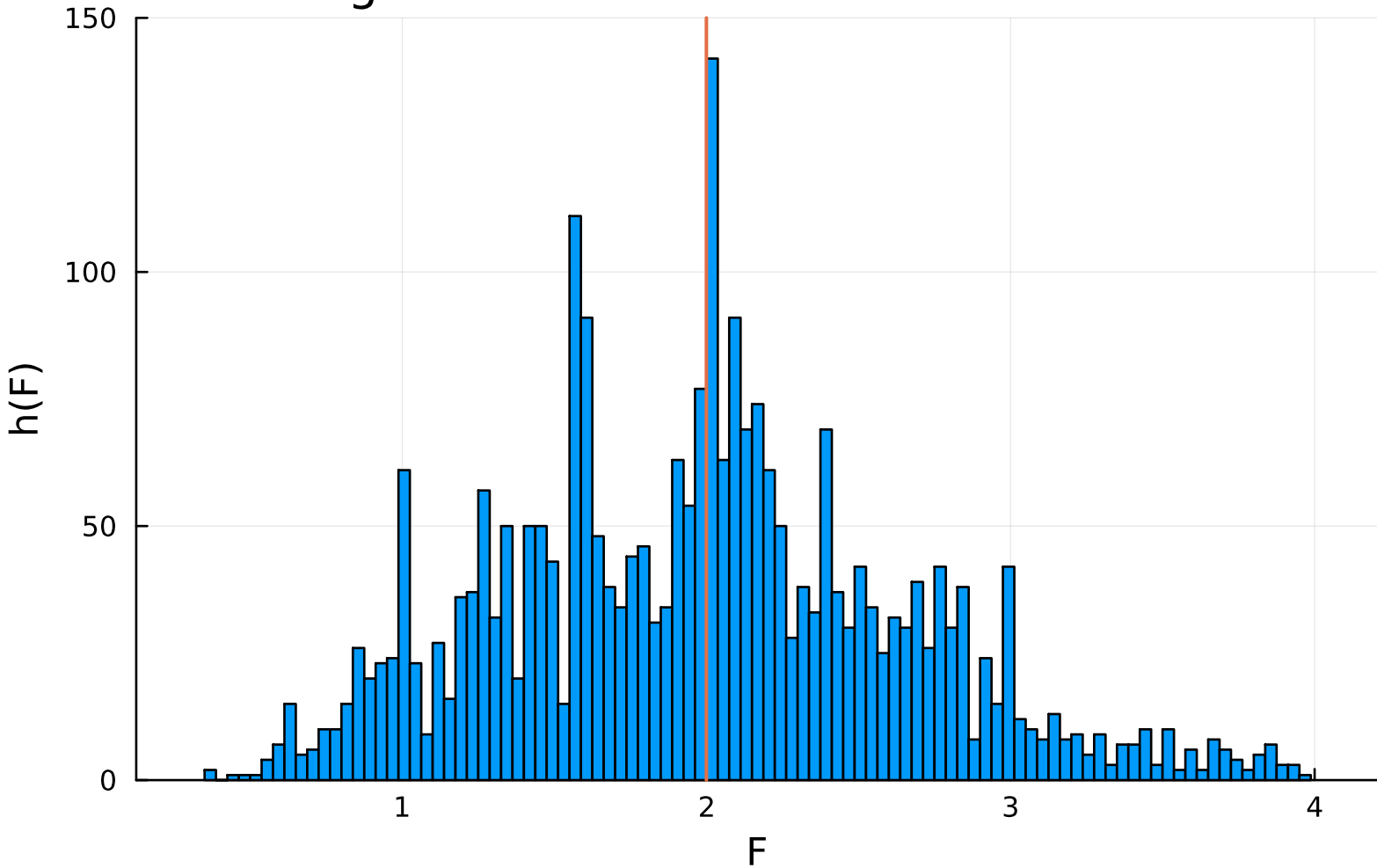Figure 3: FAR Distribution for $\overline{F} = 2.0$

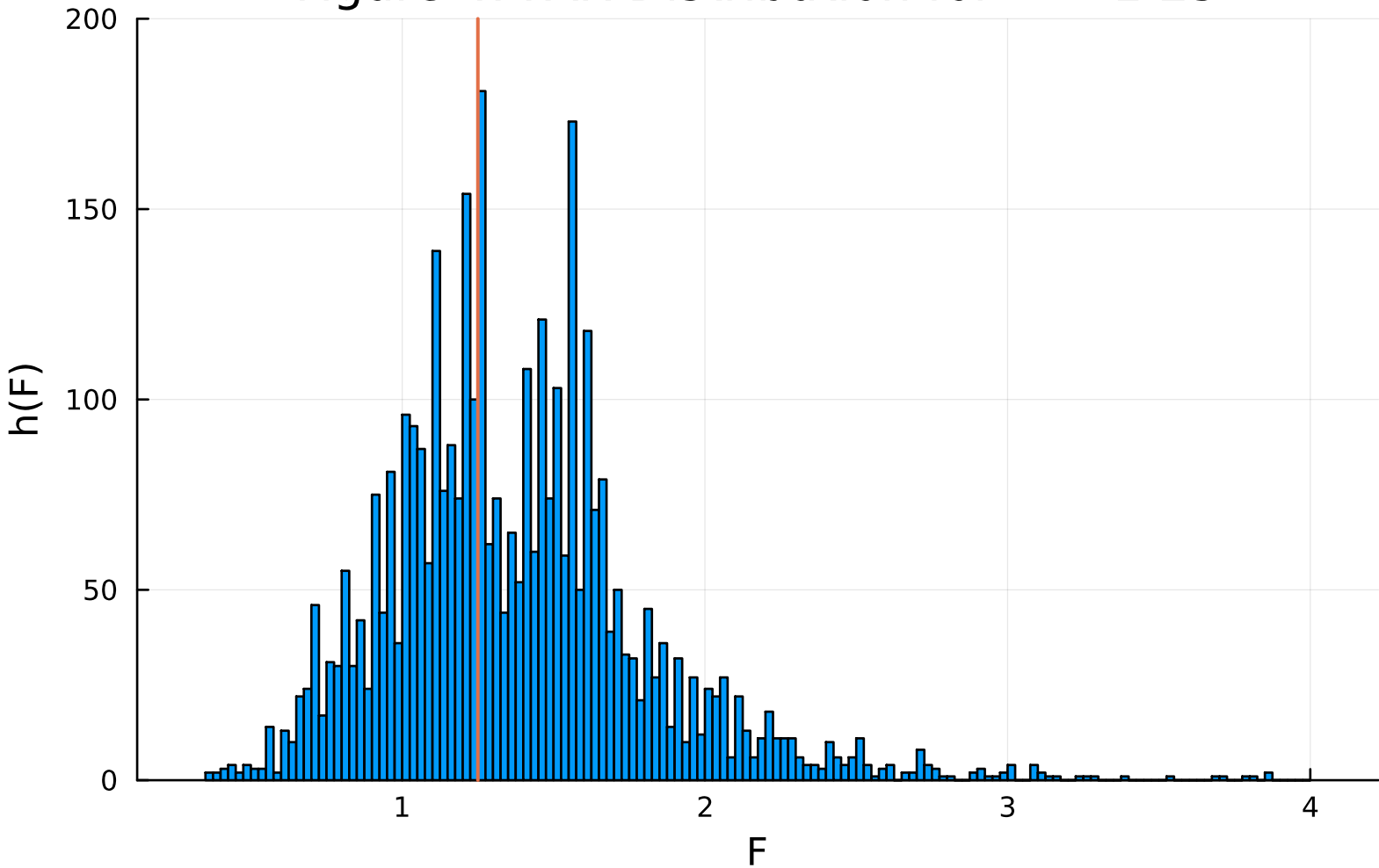Figure 4: FAR Distribution for $\overline{F} = 1.25$

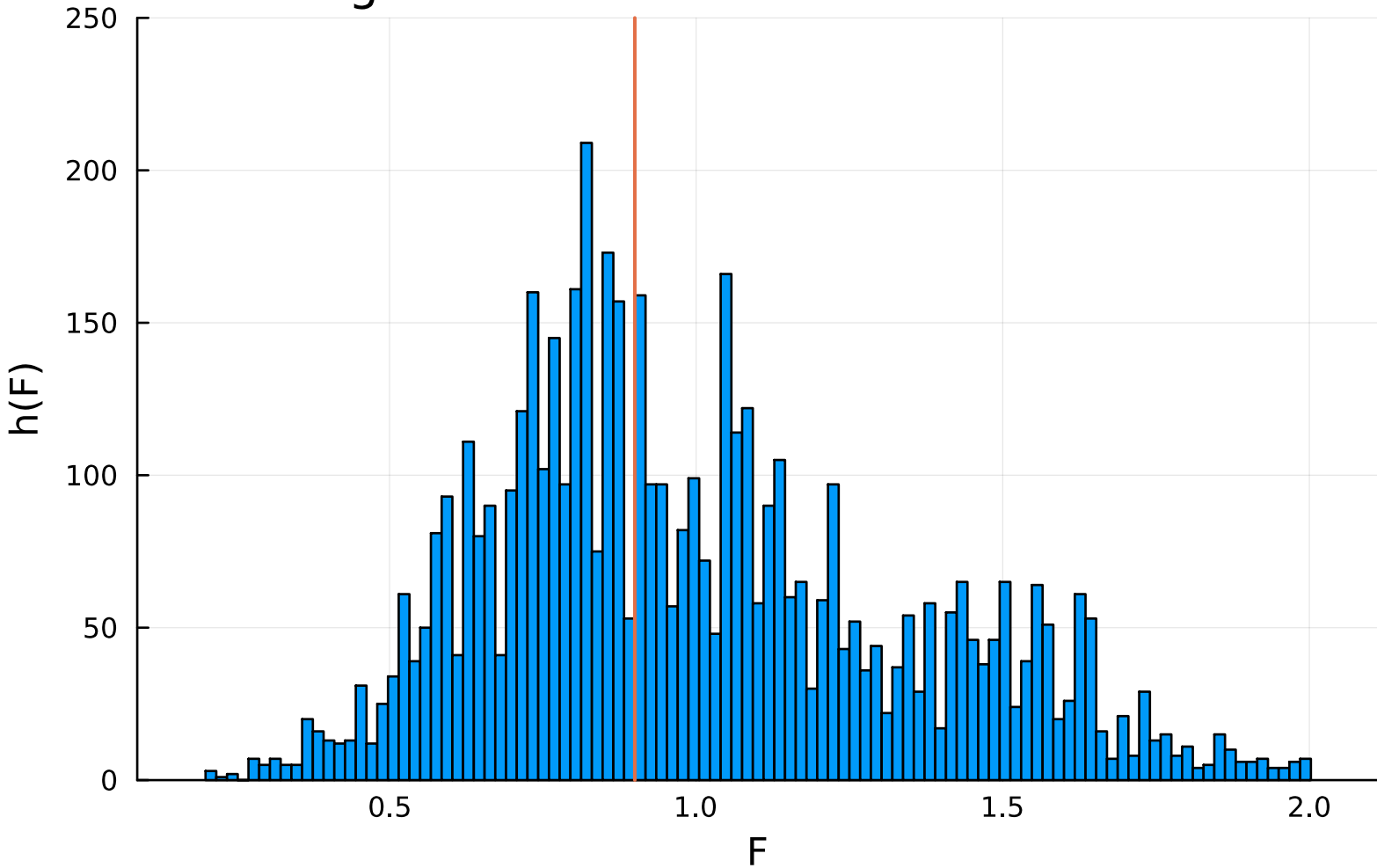Figure 5: FAR Distribution for $\overline{F} = 0.9$

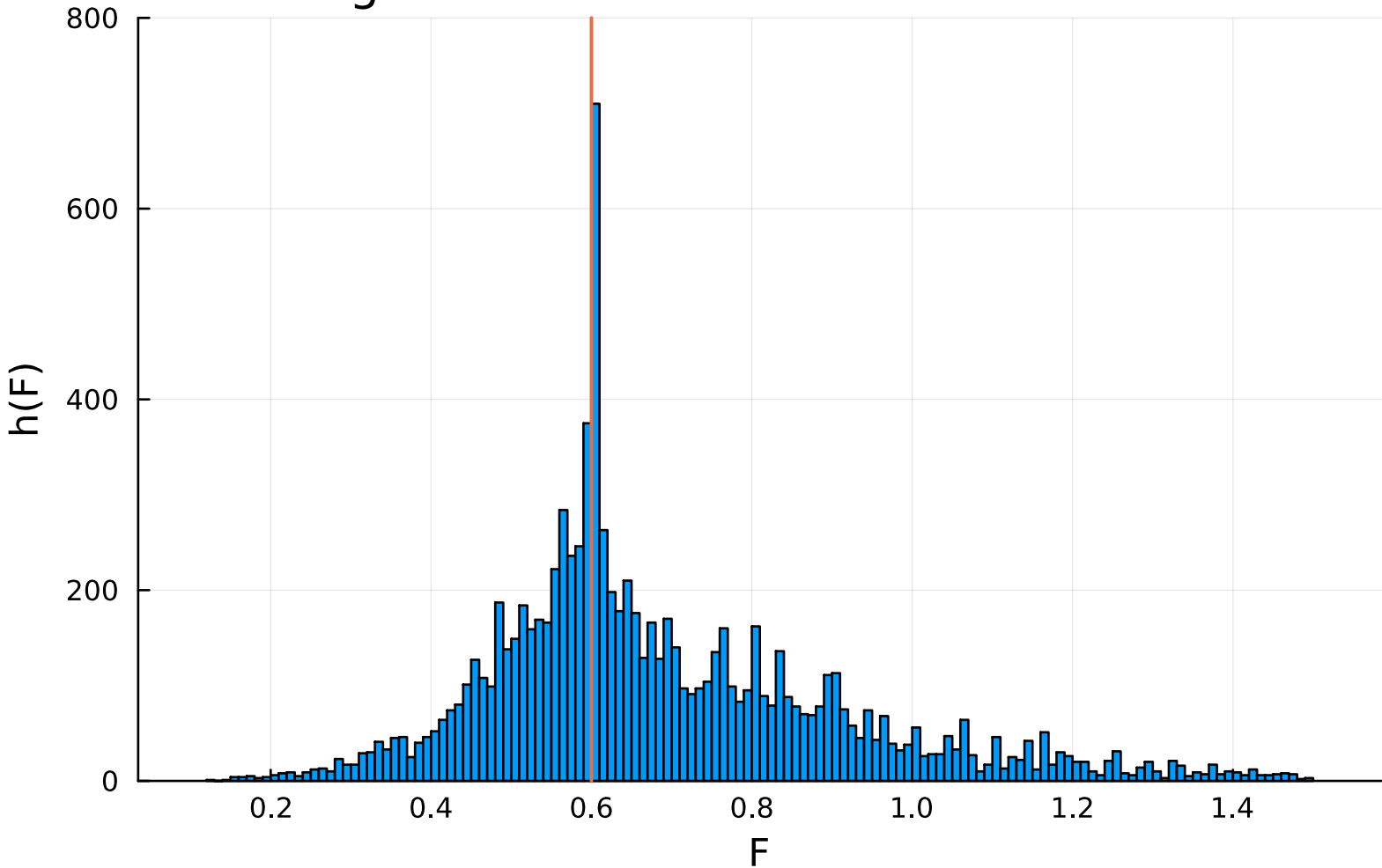Figure 6: FAR Distribution for $\overline{F} = 0.6$

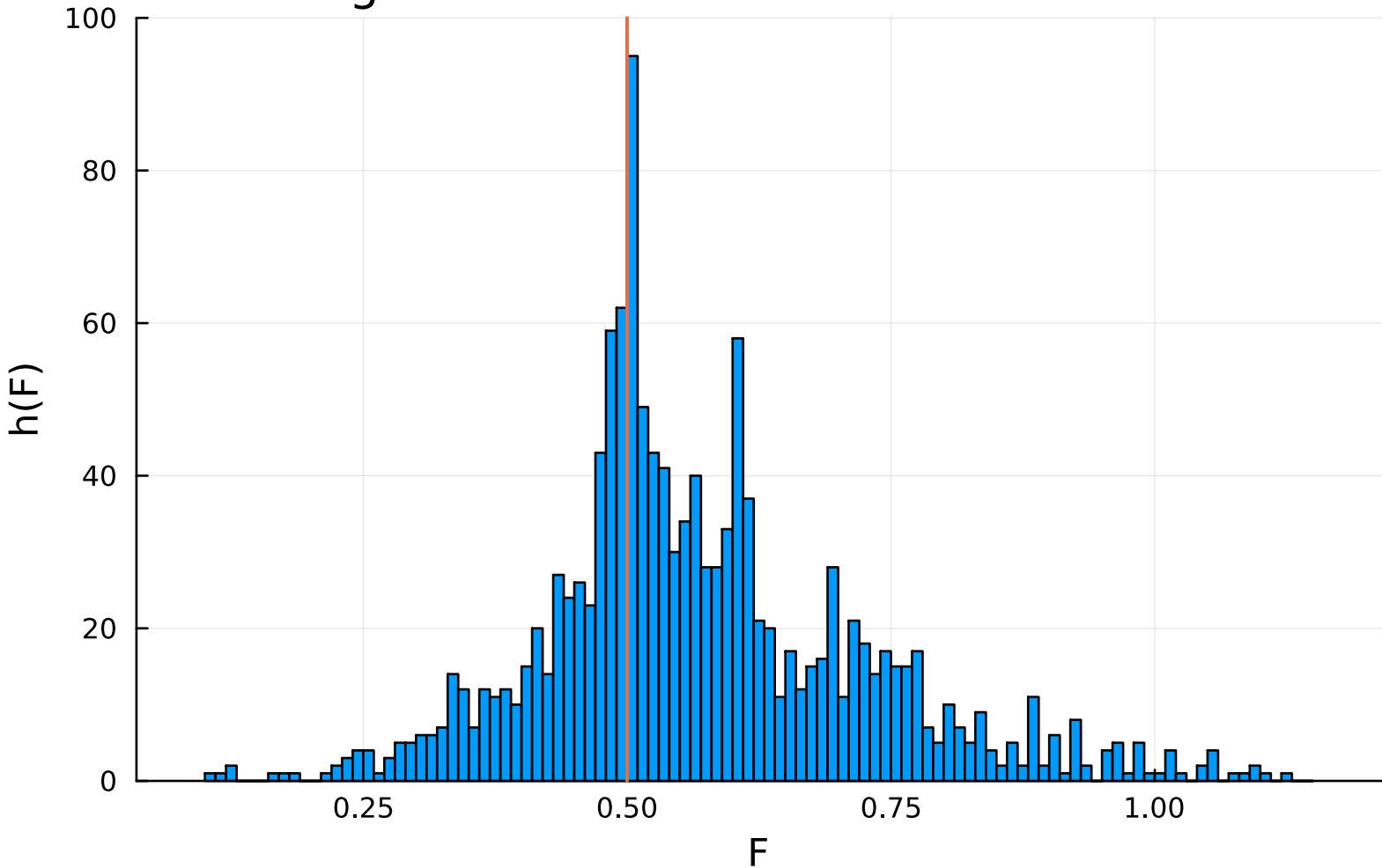Figure 7: FAR Distribution for $\overline{F} = 0.5$

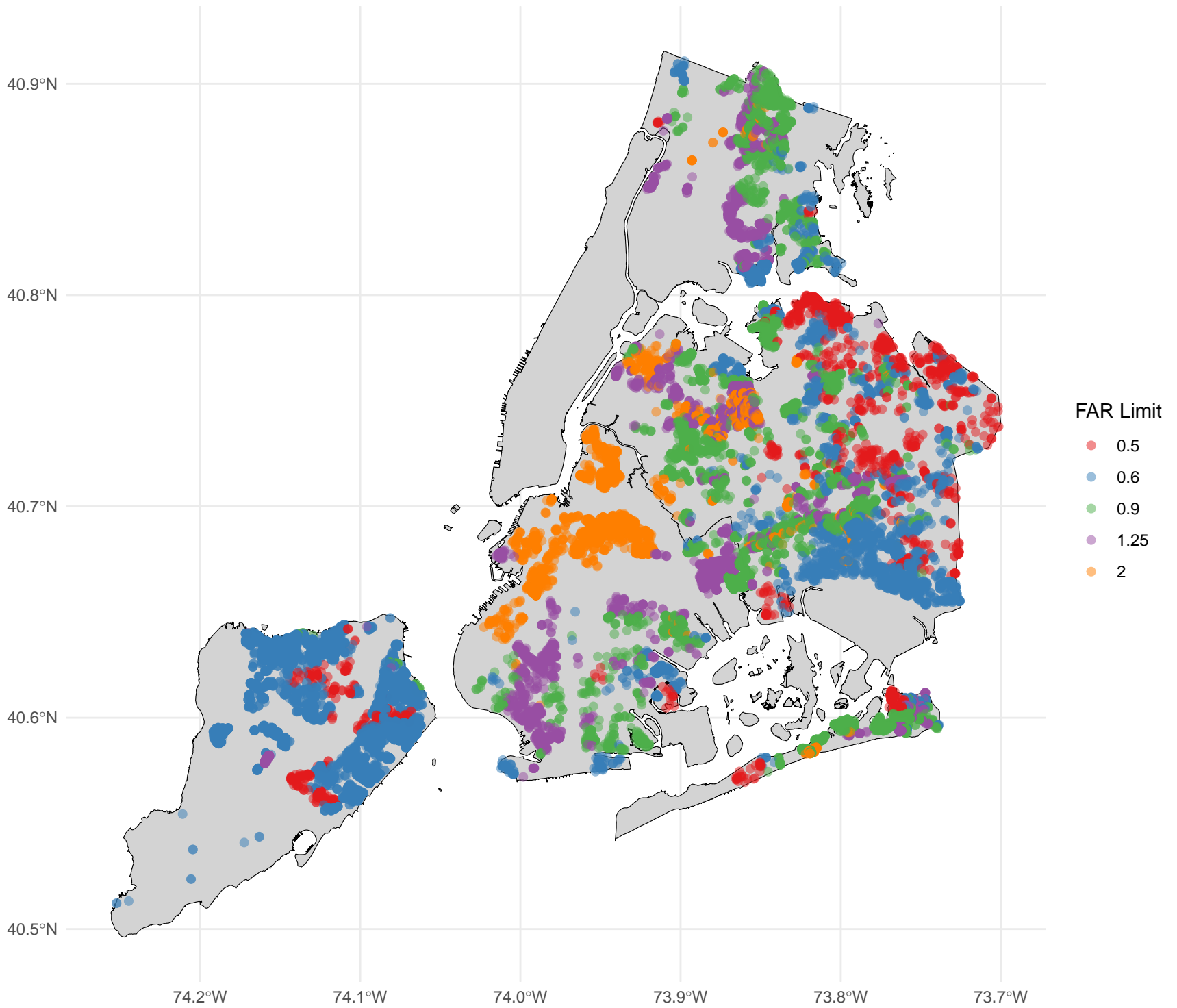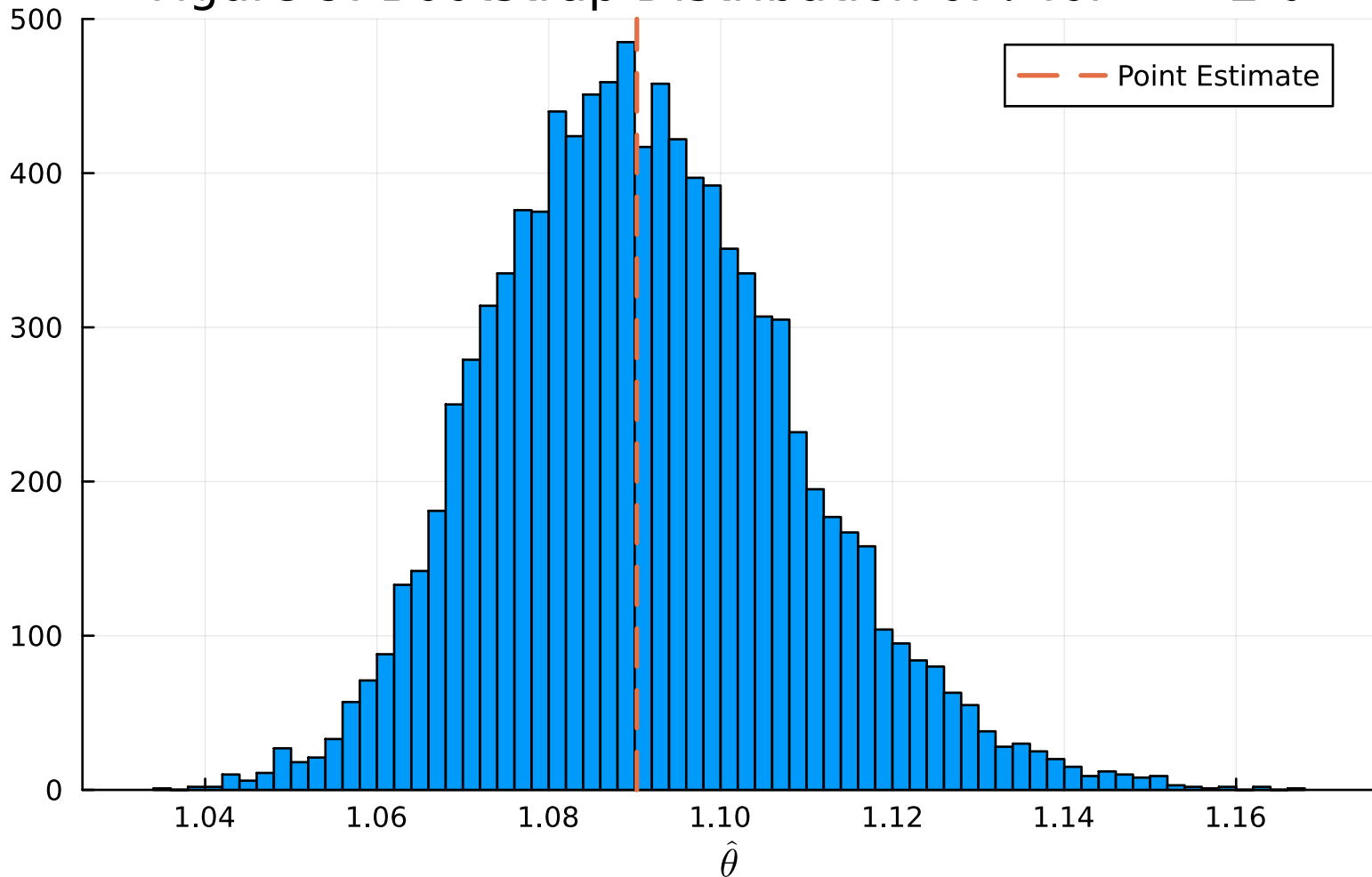Figure 8: Location of Sample Properties by FAR Limit

Figure 9: Bootstrap Distribution of $\hat{\theta}$ for $\overline{F} = 2.0$

References

AHLFELDT, G.M., BARR, J., 2022 The economics of skyscrapers: A synthesis? *Journal of Urban Economics* 129, article 103419.

AHLFELDT, G.M., MCMILLEN, D.P., 2018. Tall buildings and land values: Height and construction cost elasticities in Chicago, 1870-2010. *Review of Economics and Statistics* 100, 861-875.

BARR, J., JEDWAB, R.2023. Exciting, boring, and nonexistent skylines: Vertical building gaps in global perspective. *Real Estate Economics* 51, 1512-1546.

BERTANHA, M., CAETANO, C., JALES, H., SEEGERT, N., 2023. Bunching designs: A guide to practice. Unpublished paper

BERTAUD, A., BRUECKNER, J.K., 2005. Analyzing building height restrictions: Predicted impacts and welfare costs. *Regional Science and Urban Economics* 35, 109-125.

BLOMQUIST, S., NEWEY, W.K., KUMAR, A., LIANG, C.-Y., 2021. On bunching and identification of the taxable income elasticity. *Journal of Political Economy* 129, 2320-2342.

BRUECKNER, J.K., SRIDHAR, K.S., 2012. Measuring welfare gains from relaxation of land-use restrictions: The case of India's building-height limits. *Regional Science and Urban Economics* 42, 1061-1067.

BRUECKNER, J.K., SINGH, R., 2020. Stringency of land-use regulation: Building heights in US cities. *Journal of Urban Economics* 116, 102329.

CHETTY, R., FRIEDMAN, J.N., OLSEN, T., PISTAFERRI, L., 2011. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records. *Quarterly Journal of Economics* 126, 749-804.

CUTTER, W.B., FRANCO, S.F., 2012. Do parking requirements significantly increase the area dedicated to parking? A test of the effect of parking requirements values in Los Angeles County. *Transportation Research Part A* 46, 901-925.

DURANTON, G., COMBES, P.-P., GOBILLON, L., 2021. The production function for housing: Evidence from France. *Journal of Political Economy* 129, 2766-2816.

ERIKSEN, M.D., ORLANDO, A.W., 2022. Returns to scale in residential construction: The marginal impact of building height. *Real Estate Economics* 50, 534-564.

GARICANO, L., LELARGE, C., VAN REENEN, J., 2016. Firm size distortions and the productivity distribution: Evidence from France. *American Economic Review* 106, 3439-3479.

GLAESER, E.L., GYOURKO, J., SAKS, R., 2005. Why is Manhattan so expensive? Regulation and the rise in housing prices. *Journal of Law and Economics* 48, 331-369.

KLEVEN, H.J., WASEEM, M., 2013. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *Quarterly Journal of Economics* 128, 669-723.

KLEVEN, H.J., 2016. Bunching. *Annual Review of Economics* 8, 435-464.

KOPCZUK, W., MUNROE, D., 2015. Mansion tax: The effect of transfer taxes on the residential real estate market. *American Economic Journal: Economic Policy* 7, 214-257.

LEATHER, D., 2023. What's the use? Land use uncertainty, real estate prices, and the redevelopment option. Unpublished paper available at SSRN: `https://ssrn.com/abstract=4551548`.

LEVY, A., 2024. Regulating housing quality: Evidence from France. Unpublished paper.

PENG, X.E., 2023. The dynamics of urban development: Evidence from land use reform in New York. Unpublished paper.

SAEZ, E., 2010. Do taxpayers bunch at kinks? *American Economic Journal: Economic Policy* 2, 180-212.

SLEMROD, J., WEBER, C., SHAN, H., 2017. The behavioral response to housing transfer taxes: Evidence from a notched change in D.C. policy? *Journal of Urban Economics* 100, 137-153.

WEINBERGER, A., 2023. Air rights guide. On PropertyClub website. `https://propertyclub.nyc/article/air-rights-in-real-estate`.

Bunching in Real-Estate Markets: Regulated Building Heights in New York City

Jan K. Brueckner, David Leather, and Miguel Zerecero

We can ask how much floor space would increase if the FAR limit $\overline{F}$ were to increase by a small amount. First, recall that the optimal choice of FAR as a function of $\overline{F}$ is

$$\mathcal{F}(p, \overline{F}) = \begin{cases} (p/\alpha)^{1/\lambda} & \text{if } p < \alpha\overline{F}^{\lambda}, \\ \overline{F} & \text{if } p \in \left[\alpha\overline{F}^{\lambda}, \, (\alpha+\beta)\overline{F}^{\lambda}\right], \\ (p/(\alpha+\beta))^{1/\lambda} & \text{if } p > (\alpha+\beta)\overline{F}^{\lambda}. \end{cases}$$

We can compute what would be the change in the chosen FAR after $\overline{F}$ is increased by a small amount $\varepsilon \leq (\theta - 1)\overline{F}$. If the FAR change is smaller than this limit, at least some of the current bunchers also bunch under the new FAR limit. The increases in FAR after the change are given by

$$\mathcal{F}(p, \overline{F} + \varepsilon) - \mathcal{F}(p, \overline{F}) = \begin{cases} 0 & \text{if } p < \alpha\overline{F}^{\lambda}, \\ (p/\alpha)^{1/\lambda} - \overline{F} & \text{if } p \in \left[\alpha\overline{F}^{\lambda}, \, \alpha\left(\overline{F}+\varepsilon\right)^{\lambda}\right], \\ \varepsilon & \text{if } p \in \left(\alpha\left(\overline{F}+\varepsilon\right)^{\lambda}, \, (\alpha+\beta)\overline{F}^{\lambda}\right), \\ \overline{F} + \varepsilon - (p/(\alpha+\beta))^{1/\lambda} & \text{if } p \in \left[(\alpha+\beta)\overline{F}^{\lambda}, \, (\alpha+\beta)\left(\overline{F}+\varepsilon\right)^{\lambda}\right], \\ 0 & \text{if } p > (\alpha+\beta)\left(\overline{F}+\varepsilon\right)^{\lambda}. \end{cases}$$

The equation above spells out five relevant groups. For the first and last groups, the change does not affect their FAR choice. The second group corresponds to buildings that were bunched under the old $\overline{F}$ but not under the new one, $\overline{F} + \varepsilon$. The third group contains buildings that were bunched under the old restriction and bunch at the higher $\overline{F}$ under the new restriction. The fourth

group consists of building that were not bunched before, being above $\overline{F}$, but choose to bunch under the new restriction.

To obtain the total change in FAR we need to integrate the individual changes over the entire distribution of prices p, which has a density function $t(p)$. We can integrate separately for each of the relevant subsets of the support of p, starting with the new non-bunchers:

$$D_1(\varepsilon) \equiv \int_{\alpha \overline{F}^\lambda}^{\alpha(\overline{F}+\varepsilon)^\lambda} \left[(p/\alpha)^{1/\lambda} - \overline{F} \right] t(p)dp.$$

As we did previously, we do a change of variable. Defining $z = (p/\alpha)^{1/\lambda}$, we have that $dp = \alpha\lambda z^{\lambda-1}dz$. Recall that $t(\alpha z^\lambda)\alpha\lambda z^{\lambda-1} \equiv h_0(z)$. Then, we get

$$D_1(\varepsilon) = \int_{\overline{F}}^{\overline{F}+\varepsilon} \left(z - \overline{F} \right) h_0(z)dz.$$

Approximate this integral with the area of a trapezoid yields

$$D_1(\varepsilon) = \varepsilon \left(\frac{\overline{F}h_0(\overline{F}) + \left(\overline{F} + \varepsilon \right) h_0(\overline{F} + \varepsilon)}{2} - \overline{F}\frac{h_0(\overline{F}) + h_0(\overline{F} + \varepsilon)}{2} \right) = \frac{\varepsilon^2}{2}h_0(\overline{F} + \varepsilon).$$

For the always-bunchers we have:

$$D_2(\varepsilon) = \varepsilon \int_{\alpha(\overline{F}+\varepsilon)^\lambda}^{(\alpha+\beta)\overline{F}^\lambda} t(p)dp = \varepsilon \int_{\overline{F}+\varepsilon}^{\overline{F}+\Delta\overline{F}} h_0(z)dz = \varepsilon \left(\Delta\overline{F} - \varepsilon \right) \left(\frac{h_0(\overline{F} + \varepsilon) + h_0(\overline{F} + \Delta\overline{F})}{2} \right),$$

where the last equality uses the trapezoid approximation.

For the new-bunchers, doing the same change of variable, we have:

$$D_3(\varepsilon) = \int_{(\alpha+\beta)\overline{F}^\lambda}^{(\alpha+\beta)(\overline{F}+\varepsilon)^\lambda} \left[\overline{F} + \varepsilon - (p/(\alpha+\beta))^{1/\lambda} \right] t(p)dp = \int_{\overline{F}+\Delta\overline{F}}^{\overline{F}+\Delta\overline{F}+\theta\varepsilon} \left[\overline{F} + \varepsilon - \theta^{-1}z \right] h_0(z)dz.$$

Using again the trapezoid approximation, we get the following expression

$$D_3(\varepsilon) = \theta\varepsilon \left(\overline{F} + \varepsilon \right) \left(\frac{h_0(\overline{F} + \Delta\overline{F}) + h_0(\overline{F} + \Delta\overline{F} + \theta\varepsilon)}{2} \right)$$

$$- \frac{\theta\varepsilon}{\theta} \left(\frac{(\overline{F} + \Delta\overline{F})h_0(\overline{F} + \Delta\overline{F}) + (\overline{F} + \Delta\overline{F} + \theta\varepsilon)h_0(\overline{F} + \Delta\overline{F} + \theta\varepsilon)}{2} \right)$$

$$= \frac{\theta\varepsilon}{2} \left[h_0(\overline{F} + \Delta\overline{F}) \left(\overline{F} + \varepsilon - \left(\frac{\overline{F} + \Delta\overline{F}}{\theta} \right) \right) + h_0(\overline{F} + \Delta\overline{F} + \theta\varepsilon) \left(\overline{F} + \varepsilon - \left(\frac{\overline{F} + \Delta\overline{F} + \theta\varepsilon}{\theta} \right) \right) \right].$$

As $\overline{F} + \Delta\overline{F} = \theta\overline{F}$ we can simplify the expression above to

$$D_3(\varepsilon) = \frac{\theta\varepsilon^2}{2}h_0(\overline{F} + \Delta\overline{F}).$$

Denote the total change of FAR after increasing the limit by ε as $D(\varepsilon)$. It equals

$$D(\varepsilon) = D_1(\varepsilon) + D_2(\varepsilon) + D_3(\varepsilon)$$

$$= \frac{\varepsilon^2}{2} h_0(\overline{F} + \varepsilon) + \varepsilon \left(\Delta \overline{F} - \varepsilon \right) \left(\frac{h_0(\overline{F} + \varepsilon) + h_0(\overline{F} + \Delta \overline{F})}{2} \right) + \frac{\theta \varepsilon^2}{2} h_0(\overline{F} + \Delta \overline{F})$$

$$= \varepsilon \Delta \overline{F} \left(\frac{h_0(\overline{F} + \varepsilon) + h_0(\overline{F} + \Delta \overline{F})}{2} \right) + \frac{\varepsilon^2 \left(\theta - 1 \right)}{2} h_0(\overline{F} + \Delta \overline{F}).$$

Rate of change of FAR to a marginal increase in $\overline{F}$. From the expression above, we can calculate the derivative of total FAR with respect to the FAR limit $\overline{F}$. By definition, this derivative is equal to

$$\lim_{\varepsilon \to 0} \frac{D(\varepsilon)}{\varepsilon} = \Delta \overline{F} \left(\frac{h_0(\overline{F}) + h_0(\overline{F} + \Delta \overline{F})}{2} \right) + \lim_{\varepsilon \to 0} \left(\frac{\varepsilon \left(\theta - 1 \right)}{2} h_0(\overline{F} + \Delta \overline{F}) \right) = B.$$

Therefore, the derivative of total FAR with respect to the FAR limit is equal to the amount of bunching.

However, this derivative is only informative for marginal increments to the FAR limit $\overline{F}$, which might not be of importance in a potential practical implementation of a reform. Thus, we are also interested in how much total FAR would change when ε is not marginal.

Change in FAR from non-marginal changes in $\overline{F}$. We can rearrange $D(\varepsilon)$ to be a function of B by adding and subtracting some terms:

$$D(\varepsilon) = \varepsilon \Delta \overline{F} \left(\frac{h_0(\overline{F}) + h_0(\overline{F} + \Delta \overline{F})}{2} \right) + \frac{\varepsilon^2 \left(\theta - 1 \right)}{2} h_0(\overline{F} + \Delta \overline{F}) - \varepsilon \Delta \overline{F} \left(\frac{h_0(\overline{F}) - h_0(\overline{F} + \varepsilon)}{2} \right)$$

$$= \varepsilon B + \frac{\varepsilon^2 \left(\theta - 1 \right)}{2} h_0(\overline{F} + \Delta \overline{F}) - \varepsilon \Delta \overline{F} \left(\frac{h_0(\overline{F}) - h_0(\overline{F} + \varepsilon)}{2} \right).$$

This formula is useful, as we can then use the amount of bunching, which is observable, to calculate $D(\varepsilon)$. However, there are still obstacles. As explained in Section 2, we do not observe h_0 for values above $\overline{F}$. For $h_0(\overline{F} + \Delta \overline{F})$, equation (8) already gives us that this is equal to $\theta^{-1} h_1(\overline{F})$. However, now that we have an estimate for θ we can create a correspondence between h_0 and h_1 for *any* value F.

To see this point, recall the definitions for h_0 and h_1:

$$h_0(F) \equiv t\left(\alpha F^\lambda\right)\alpha\lambda F^{\lambda-1}, \quad h_1(F) \equiv t\left((\alpha+\beta)F^\lambda\right)(\alpha+\beta)\lambda F^{\lambda-1}.$$

Now note that

$$h_1\left((F+\varepsilon)\theta^{-1}\right) = t\left((\alpha+\beta)((F+\varepsilon)\theta^{-1})^\lambda\right)(\alpha+\beta)\lambda((F+\varepsilon)\theta^{-1})^{\lambda-1}$$

$$= t\left((\alpha+\beta)((F+\varepsilon))^\lambda\left(\frac{\alpha}{\alpha+\beta}\right)\right)(\alpha+\beta)\lambda(F+\varepsilon)^{\lambda-1}\left(\frac{\alpha}{\alpha+\beta}\right)^{\frac{\lambda-1}{\lambda}}$$

$$= t\left(\alpha((F+\varepsilon))^\lambda\right)\alpha\lambda(F+\varepsilon)^{\lambda-1}\left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda} = h_0\left(F+\varepsilon\right)\theta.$$

Therefore we have

$$h_0(\overline{F}+\varepsilon) = \theta^{-1}h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right).$$

Putting everything together, and noting that $\Delta\overline{F} = (\theta-1)\overline{F}$, we have an expression for $D(\varepsilon)$ as a function of objects that are observable:

$$D(\varepsilon) = \varepsilon B + \frac{\varepsilon^2}{2}\left(\frac{\theta-1}{\theta}\right)h_1(\overline{F}) - \frac{\varepsilon(\theta-1)\overline{F}}{2}\left[h_0(\overline{F}) - \theta^{-1}h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right)\right]$$

$$= \varepsilon B + \frac{\varepsilon}{2}\left(\frac{\theta-1}{\theta}\right)\varepsilon h_1(\overline{F}) - \frac{\varepsilon}{2}\left(\frac{\theta-1}{\theta}\right)\left[\theta\overline{F}h_0(\overline{F}) - \overline{F}h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right)\right]$$

$$= \varepsilon B + \frac{\varepsilon}{2}\left(\frac{\theta-1}{\theta}\right)\left[\varepsilon h_1(\overline{F}) + \overline{F}h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right) - \theta\overline{F}h_0(\overline{F})\right].$$

The only new object to estimate after estimating θ is $h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right)$. If one does not want to do so, one could compute a lower bound instead using previously computed objects:

$$D(\varepsilon) \geq \varepsilon B + \frac{\varepsilon}{2}\left(\frac{\theta-1}{\theta}\right)\left(\varepsilon h_1(\overline{F}) - \theta\overline{F}h_0(\overline{F})\right).$$

We have constrained ε to be smaller than $(\theta-1)\overline{F}$. We can then compute what is the change in FAR when we relax $\overline{F}$ up to the limit imposed by this restriction, i.e. $\varepsilon = (\theta-1)\overline{F}$. Substituting back the expression for B in the main text and after some simplifications, we get:

$$D((\theta-1)\overline{F}) = \frac{\left((\theta-1)\overline{F}\right)^2}{2}\frac{1+\theta}{\theta}h_1(\overline{F}).$$

With θ and $h_1(\overline{F})$, it is immediate to evaluate the expression above. In this case, there is no need to estimate new objects after estimating θ.

Change in floor space after a change in $\overline{F}$. To calculate the change in floor space after relaxing the FAR limit $\overline{F}$ for an individual parcel, we need to multiply the lot size of each parcel to the individual change in FAR. To get the total change in floor space we need to integrate over all of the parcels.

Let $d(p, \overline{F}, \varepsilon) \equiv \mathcal{F}(p, \overline{F} + \varepsilon) - \mathcal{F}(p, \overline{F})$ denote the change in FAR for a developer with rental price p, current FAR limit $\overline{F}$ and increase in the limit ε. Then the change in floor space for a developer facing price p and that has a lot size ℓ is simply

$$d_{(}p, \overline{F}, \varepsilon)\ell.$$

Let $g(\ell, p)$ denote the joint density of parcels with lot size ℓ and price p, with $\int \int g(\ell, p) \mathrm{d}\ell \, \mathrm{d}p = 1$. Also, let N be the total number of parcels. Then, the total change in floor space after increasing $\overline{F}$ by ε is

$$\Delta S \equiv N \underbrace{\int \int d(p, \overline{F}, \varepsilon)\ell \, g(\ell, p) \, \mathrm{d}\ell \, \mathrm{d}p}_{\text{Average change in floor space}} = N \int \int d(p, \overline{F}, \varepsilon)\ell \, g(\ell|p) \, g_P(p) \, \mathrm{d}\ell \, \mathrm{d}p,$$

where $g(\ell|p)$ is the conditional density of parcels with lot size ℓ conditional on price p, and $g_P(p)$ is the marginal density of parcels with price p.

As we do not observe prices, we cannot assign a lot size to each counterfactual change in FAR. However, if floor space and prices are independent, then

$$\Delta S = N \underbrace{\int d(p, \overline{F}, \varepsilon)g_P(p) \, \mathrm{d}p}_{\text{Average change in FAR}} \int \ell \, g(\ell) \, \mathrm{d}\ell = D(\varepsilon) \underbrace{\int \ell \, g(\ell) \, \mathrm{d}\ell}_{\text{Average Lot Area}},$$

where we use the fact that $g(\ell|p) = g(\ell)$ because of independence of ℓ and p, and $D(\varepsilon)$ is, as before, the total change in FAR. If the independence assumption does not hold, we can regard the expression above as an approximation.

Table A1: Floor space elasticity from non-marginal increase in $\overline{\text{FAR}}$

$\overline{FAR}$	Elasticity
2.0	0.082
1.25	0.0043
0.9	0.0019
0.6	0.00017
0.5	0.00093

The table gives elasticity of the predicted increment to floor space from a non-marginal increase in $\overline{\text{FAR}}$, as described in the text.