# 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

**Abstract**

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 pattern near a regulated height to estimate the marginal-cost penalty for exceeding that height, 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 (often the increment to a marginal tax rate) is known and the goal is to estimate a behavioral parameter (often a labor-supply elasticity). By contrast, 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 a 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.

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

*by*

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

## 1. Introduction

The bunching methodology introduced by Saez (2010) has been widely used in a variety of applications, usually to estimate the value of an unknown behavioral parameter. 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 et al. (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.

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.[1] The present paper carries out this type of exercise, with a focus on the real-estate market. In particular, we study bunching due to regulated building heights for residential housing, using data from New York City. The only other studies of

[†] We thank David Agrawal, Gilles Duranton, Teju Velayudhan, Mazhar Waseem, and especially Jinwon Kim for helpful comments and discussions. The usual disclaimer applies.

[1] After an hearing an explanation of our approach, Mazhar Waseem described it as "reverse engineering." Although they do not use Saez's (2010) methodology, Garicano, Lelarge and Van Reenen (2016) carry out an excercise 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.

real-estate bunching of which we are aware are Kopczuk and Munroe (2015) and Slemrod, Weber and Shan (2017), who study transfer taxes.[2] But our focus on building heights and their regulation follows a growing literature in urban economics.[3]

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.[4] 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 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.

Thus, these two options allow the developer to build above $\overline{\text{FAR}}$, but at a cost. In effect, they raise the marginal cost of floor space beyond $\overline{\text{FAR}}$, creating a kink in the cost function. Since a parcel's $\overline{\text{FAR}}$ is observed, the location of the kink is known. However, its size is

[2] 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.

[3] 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.

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

unknown. The goal of the paper is to estimate the size of this cost kink using Saez's (2010) 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 production function using French data.

We proceed by dividing residential parcels in NYC 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. Within each group, we drop buildings constructed prior to the year 2000, so that observed FAR values have been set relatively recently. Then, within each $\overline{\text{FAR}}$ group, we apply our adaptation of Saez's method 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 500 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.

While this exercise is carried out in section 5 of the paper, section 2 shows how Saez's 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

### *2.1. Housing production*

Emmanuel Saez's (2010) bunching methodology can be adapted to a developer's choice of building height. This section explains the adaptation, drawing in detail on Saez's analyis.

To start, we depict the production of housing floor space. Let $Q$ denote output of floor space, which is produced with inputs of capital $K$ (building materials) and land $\ell$ 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. The first-order condition for choice of $S$ is $pf'(S) = 1$ (note that $f'' < 0$).

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 $S^\rho$ 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 $\alpha/\gamma$ below $\overline{F}$ and $(\alpha+\beta)/\gamma$ above it, as follows:

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

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

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

It is easily seen from (2) that $1/\lambda$ is the elasticity of $F$ with respect to $p$, both below and above $\overline{F}$. While (2) applies 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.

*2.2. Bunching*

Inspection of (2) 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) is 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 $t(p)$. Therefore, optimal values of $F$ differ across parcels depending on the relevant $p$ values. Since $F$, not $p$, is observed, we need to derive the density of $F$ over the relevant ranges. Since $p = \alpha F^{\lambda}$ holds 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

$$t(\alpha F^{\lambda})\alpha\lambda F^{\lambda-1} \quad \equiv \quad h_0(F), \tag{3}$$

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 $t(p)$, yielding

$$t((\alpha+\beta)F^{\lambda})(\alpha+\beta)\lambda F^{\lambda-1} \quad \equiv \quad h_1(F), \tag{4}$$

where $h_1$ denotes $F$'s density in this range.

Let $h(F)$ give the overall density of $F$, which equals $h_0(F)$ below $\overline{F}$ and $h_1(F)$ above $\overline{F}$. In addition, let the limits of $h(F)$ as $F$ approaches $\overline{F}$ from below (above) be denoted $h(\overline{F})_- \equiv h_0(\overline{F})$ and $h(\overline{F})_+ \equiv h_1(\overline{F})$, respectively. From (3) and (4), it is clear that these limits are unequal, so that a density discontinuity exists at $\overline{F}$.

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 $\lambda$, 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 $\alpha$ would choose $F = [(\alpha+\beta)/\alpha]^{1/\lambda}\overline{F}$, as can be seen from setting $p = p^{**}$ in the first line of (2). 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

$$\Delta\overline{F} \quad = \quad \left[\left(\frac{\alpha+\beta)}{\alpha}\right)^{1/\lambda} \; - \; 1\right]\overline{F}. \tag{5}$$

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:

$$B \quad = \quad \int_{\overline{F}}^{\overline{F}+\Delta\overline{F}} h_0(z)dz. \tag{6}$$

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

$$B \quad = \quad \Delta\overline{F}\,\frac{h_0(\overline{F}) \;+\; h_0(\overline{F}+\Delta\overline{F})}{2}. \tag{7}$$

The problem, though, in operationalizing (7) 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 (3) and (4) to replace $h_0(\overline{F}+\Delta\overline{F})$ by expression involving the (observable) density $h_1$, circumventing this obstacle. Doing so yields

$$\begin{aligned}
h_0(\overline{F}+\Delta\overline{F}) &= t[\alpha(\overline{F}+\Delta\overline{F})^{\lambda}]\,\alpha\lambda(\overline{F}+\Delta\overline{F})^{\lambda-1} \\
&= t\left[\alpha\left(\left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda}\overline{F}\right)^{\lambda}\right]\alpha\lambda\left[\left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda}\overline{F}\right]^{\lambda-1} \\
&= t((\alpha+\beta)\overline{F}^{\lambda})\,\alpha\lambda\left(\frac{\alpha+\beta}{\alpha}\right)^{(\lambda-1)/\lambda}\overline{F}^{\lambda-1} \\
&= \left(\frac{\alpha+\beta}{\alpha}\right)^{-1/\lambda} t(\alpha+\beta)\overline{F}^{\lambda})(\alpha+\beta)\lambda\overline{F}^{\lambda-1} \\
&= \left(\frac{\alpha+\beta}{\alpha}\right)^{-1/\lambda} h_1(\overline{F}). \qquad (8)
\end{aligned}$$

The first line of (8) 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 implies the third line, and the last line uses the definition of $h_1$ in (4).

Substituting (8) along with $h_0(\overline{F}) = h(\overline{F})_-$ and $h_1(\overline{F}) = h(\overline{F})_+$ into (7), while recalling the definition of $\Delta\overline{F}$ in (5), (6) can be written as[5]

$$\begin{aligned}
B &= \Delta\overline{F}\,\frac{h(\overline{F})_- \;+\; ((\alpha+\beta)/\alpha)^{-1/\lambda}\,h(\overline{F})_+}{2} \\
&= (\theta-1)\overline{F}\,\frac{h(\overline{F})_- \;+\; (1/\theta)\,h(\overline{F})_+}{2}, \qquad (9)
\end{aligned}$$

where

$$\theta \;\equiv\; \left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda}. \qquad (10)$$

Note that $\theta$ 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.

[5] Eq. (9) 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.

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

$$\widehat{(\alpha+\beta)}/\alpha \quad = \quad \widehat{\theta}^{\,\lambda}, \tag{11}$$

using an independent value of $\lambda$.

As explained in Saez (2010), the elements in (9) can be generated by creating three FAR intervals around $\overline{F}$, defined by a width factor $\delta$. One interval is centered at $\overline{F}$, consisting of 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 $h(\overline{F})_-$ and $h(\overline{F})_+$.

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 estimated average height of the density $h_-$ below $\overline{F}$ equals $\widehat{h}_- = H_-/\delta$, and the estimated average height of the density $h_+$ above $\overline{F}$ equals $\widehat{h}_+ = H_+/\delta$. In each case, the average density height equals the number of FAR observations in each interval divided by its width. This discussion is illustrated below in a diagram for the $\overline{F} = 2.0$ case.

These estimated values are substituted into (9), and the equation is then solved to yield $\widehat{\theta}$, the estimate of $\theta$. Note that after multiplying through by $\theta$, (9) becomes a quadratic equation in $\theta$, which can be solved by the quadratic formula. Given $\widehat{\theta}$, we use the existing estimate of $\lambda$ from Combes, Duranton and Gobillon (2021) to yield $\widehat{\theta}^{\,\lambda}$, the estimate of the marginal-cost ratio from (11). That paper provides a recent and reliable estimate of the parameters of the Cobb-Douglas housing production function based on French data. Their estimate of the capital exponent is 0.65, which (using the second line of (2)) translates into a $\lambda$ value of $(1/0.65) - 1 = 0.54$.

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 $\theta$ estimate, and hence the estimate of $(\alpha+\beta)/\alpha$, obviously depend on the value of the interval parameter $\delta$, 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 859,225 observations. The geography of New York City is spread across the five boroughs: Brooklyn (32.23% of tax lots), Bronx (10.64%), Manhattan (5%), Queens (37.77%), and Staten Island (14.48%).

We apply several filters to the data to ensure a clean focus. 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 drop any tax lot that does not contain a single building, and we drop buildings in building class categories that are not of interest.[6] These filters exclude 403,176 observations, or roughly 46.92% 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.[7]

[6] These categories are "Garages or Gasoline Stations," "Hospitals and Health," "Theatres," "Loft Buildings," "Churches, Synagogues, etc.," "Asylums and Homes," "Places of Public Assembly (Indoor) and Cultural," "Condominium Buildings," "Outdoor Recreation Facilities," "Transportation Facilities," "Utility Bureau Properties," "Vacant Land," "Educational Structures," "Selected Government Installations," and "Misc" which contains uses such as tennis courts, court houses, public parking areas, the post office, and the United Nation buildings.

[7] 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

We also drop commercially and industrially zoned buildings from our sample, leaving only residential parcels. Most commercial zoning designations in NYC have a residential equivalent zone. This feature allows any commercially zoned building with a residential equivalent to be treated as a mixed-use building that is assigned an FAR limit equal to a combination of the FAR limits on commercial and residential space. This mixture is problematic as it creates many FAR limits, and since there are few commercially zoned buildings in the city, they do not provide a sample size large enough to run our analysis. This filter removes 10,060 observations, or 1.17% of the initial sample.

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. This final filter results in five subsamples with FAR limits of 0.5, 0.6, 0.9, 1.25, and 2.0. The final analysis dataset consists of 22,381 properties, or roughly 2.6% of the initial sample. Of these properties, 2,631 (11.75%) are in the Bronx, 4,089 (18.27%) are in Brooklyn, 8,881 (39.68%) are in Queens, and 6,779 (30.29%) 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 and the Bronx, slightly less representative of 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

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.

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 2-6 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 much of 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 striking. Figure 7 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 $\theta$ and $(\alpha + \beta)/\alpha$ estimates for the different FAR groups along with the assumed $\delta$ values used in their computation. The bootstrap confidence intervals for $\theta$ 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 $\delta$ interval value for each $\overline{\text{FAR}}$ group is chosen by careful examination of a detailed histogram of the FAR distribution for the group (more detailed than those in Figures 2-6), so as to reasonably capture the bunching area. The sensitivity analysis presented below shows how the $\widehat{\theta}$ values are affected by the $\delta$ choices.

Consider first the $\overline{\text{FAR}} = 2.0$ group. The $\widehat{\theta}$ value equals 1.119, which yields a $\widehat{(\alpha + \beta)}/\alpha$ value of 1.063, indicating that the marginal cost of additional floor space is about 6% higher above $\overline{\text{FAR}}$ than below it, a moderate cost penalty. At 1.121, the average bootstrap $\widehat{\theta}$ is very close to the sample $\theta$ estimate, and the distribution of the bootstrap $\widehat{\theta}$ values around this mean is fairly symmetric, as shown in Figure 8. The 95% confidence interval for $\theta$, also shown in the table, ranges between 1.08 and 1.16, while the implied confidence interval for the marginal-cost ratio $(\alpha + \beta)/\alpha$ ranges between 1.042 and 1.085.

Figure 9 illustrates the areas used to compute the extent of bunching, with the width of the cells equal to $\delta$. While the areas are illustrated using a kernel density that approximates the FAR distribution, the numerical computations use the actual distribution. As explained above, $B$ is equal to the area under the density between the second and fourth vertical lines (reading from the left) minus the sum of the areas between the first and second lines ($H_-$) and between the fourth and fifth vertical lines ($H_+$). This difference can be decomposed into the the area between lines 2 and 3 minus $H_-$ (the left red area) plus the area between lines 3 and 4 minus $H_+$ (the right red area).

Turning to the $\overline{\text{FAR}} = 1.25$ group, the $\widehat{\theta}$ value of 1.034, and the associated $(\widehat{\alpha + \beta})/\alpha$ value of 1.018, indicate that the cost penalty above $\overline{\text{FAR}}$ is now much smaller, at only about 2%. The mean bootstrap $\widehat{\theta}$ is again close to the sample estimate, and the $\widehat{\theta}$ distribution (not shown) is again quite symmetric.

Results for the $\overline{\text{FAR}} = 0.9$ and 0.6 groups are very similar, with the $\widehat{\theta}$ estimate now larger at about 1.06 and the implied cost premia just above 3%. The average bootstrap $\widehat{\theta}$ values are again 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 $\widehat{\theta} = 1.123$, yielding an implied a cost penalty of more than 6%. 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 2-6. The estimated values of the penalty are largest 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 relatively modest sizes of all the estimated penalities 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 2% and 6%. 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.

*4.3. Sensitivity analysis*

While we use a reliably estimated value of $\lambda$ equal to 0.54, based on a Cobb-Douglas capital exponent $\rho$ of 0.65, larger estimated values of $(\alpha+\beta)/\alpha$ emerge from larger $\lambda$ values. For example, if $\lambda = 1$, corresponding to $\rho = 0.5$ (see (2)), the estimate of the marginal-cost penalty is equal to $\widehat{\theta}$ itself, as seen in (11). In this case, the penalty ranges between 3% and 12% rather than between 2% and 6%, as can be seen from the $\widehat{\theta}$ column of Table 2. However, empirical support for such a low $\rho$ value appears to be lacking.

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 $\delta$. Table 3 shows sensitivity analysis, with the first line within each group showing our assumed $\delta$ value and $\widehat{\theta}$ values from Table 2, and the second and third lines showing the $\widehat{\theta}$ estimates using smaller and larger $\delta$ values. As can be seen, the $\widehat{\theta}$ estimates vary somewhat with the value of $\delta$. But the only striking change occurs in the $\overline{\text{FAR}} = 0.9$ group, where raising $\delta$ from the assumed value of 0.15 to 0.175 increases $\widehat{\theta}$ from 1.065 all the way to 1.237. 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.[8]

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 (in other words, "now"). The price $p$ faced by this developer, denoted $\widetilde{p}$, satisfies $\widetilde{p} = (\alpha+\beta)F_{now}^{\lambda}$, with $F_{now} = (\widetilde{p}/(\alpha+\beta))^{1/\lambda} > \overline{F}$. With FAR regulation eliminated, the $\alpha+\beta$ factor becomes $\alpha$, and this developer would instead choose a new $F$ value satisfying $\widetilde{p} = \alpha F_{new}^{\lambda}$, or $F_{new} =$

[8] Peng (2023) carries out a related exercise by using an estimated structural model to predict housing supply increases resulting from a reform that relaxed some FAR limits prior to our sample year.

$(\widetilde{p}/\alpha)^{1/\lambda}$. The ratio between the new and current FAR values is given by

$$\frac{F_{new}}{F_{now}} \quad = \quad \frac{(\widetilde{p}/\alpha)^{1/\lambda}}{(\widetilde{p}/(\alpha+\beta))^{1/\lambda}} \quad = \quad \left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda} \quad = \quad \theta \tag{12}$$

The chosen FAR thus rises by the factor $\theta$ with elimination of FAR regulation. Using (12), the new supply of floor space from the developer's lot is given by

$$F_{new} \times \ell \quad = \quad \ell\, \theta\, F_{now}, \tag{13}$$

where $\ell$ is the lot size (recall that $F$ equals floor space per acre). Thus, new floor space equals $\theta$ times the current FAR times lot size. Note that this discussion assumes that existing buildings that are constrained by FAR regulation would be replaced. By contrast, the current amount of floor space equals $\ell F_{now}$. These expressions can be summed across all parcels with $F > \overline{F}$ to get total floor space above $\overline{F}$, both before and after the elimination of FAR regulation. It is important to note that the floor-space gain depends on $\theta$, which equals the marginal-cost ratio raised to a power, not directly on the ratio itself.

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 $F = \overline{F}$, while a developer facing $p = (\alpha+\beta)\overline{F}^{\lambda}$ would choose $F = ((\alpha+\beta)/\alpha)^{1/\lambda}\overline{F} = \theta\overline{F}$. Approximating based on an average of these endpoint values, bunching developers on average would thus choose $F$ equal to $(\overline{F} + \theta\overline{F})/2 = (1/2)(1+\theta)\overline{F}$.

Therefore, for current bunchers, the average floor space in the absence of FAR regulation would equal $\ell \times [(1+\theta)/2] \times \overline{F}$, assuming $\ell$ 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.

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. As above, there are multiple groups to consider. The calculation is again 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 $\epsilon$ is just equal to the total amount of space bunched at the original $\overline{F}$. $F$ for each of these buildings increases by $\epsilon$, so that they become bunched at the marginally higher $\overline{F}$. Dividing the FAR gain by $\epsilon$ 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 $\overline{F}$.

In taking the 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 point $\overline{F}$. While this difference can be handled for the case of full removal of FAR regulation, it prevents clean calculations for the case of a non-marginal increase in $\overline{F}$. Hence, we present numbers only for the full-removal case.

In the previous computations, 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[9]

$$SPACE_{now} \ = \ \sum_i \{\ell_i\, F_i \times \mathbf{1}[\mathrm{F_i} < \overline{\mathrm{F}} - \delta] \ + \ \ell_\mathrm{i}\, \overline{\mathrm{F}} \times \mathbf{1}[\overline{\mathrm{F}} - \delta < \mathrm{F_i} < \overline{\mathrm{F}} + \delta] \ + \ \ell_\mathrm{i}\, \mathrm{F_i} \times \mathbf{1}[\mathrm{F_i} > \overline{\mathrm{F}} + \delta]\}. \tag{14}$$

Recall that, as explained above, a parcel's floor space in the bunching range is set equal to $\ell_i\overline{F}$ rather than $\ell_i F_i$.[10] Total floor space without FAR regulation is given by

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

[9] While lot size for bunchers was, for simplicity, assumed equal in the theoretical discussion above, (13) allows it to differ across parcels

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

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

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 $\theta$. Floor space below $\overline{F}-\delta$ is unchanged.[11]

The results of computing (14) and (15) for each of the $\overline{\mathrm{FAR}}$ groups are shown in Table 4. The largest percentage gains in floor space from the elimination of FAR regulation are in the $\overline{\mathrm{FAR}} = 2.0$ and 0.5 groups, where gains are 6.8% and 9.5% respectively. The gains for the other groups are smaller, in the 2-4% range. The results show that, even focusing on $\overline{\mathrm{FAR}}$ groups containing buildings that are relatively short by NYC standards, the floor-space gains from the removal of FAR regulation can be appreciable.

## 6. Conclusion

This paper has presented a rare application in the real-estate context of the bunching methodology widely used in other areas of applied microeconomics. The application is to regulated building heights in New York City, where several costly actions allow a developer to exceed the regulated height for his parcel. The goal of the paper is to use the observed bunching pattern near a regulated height to estimate the marginal-cost penalty for exceeding that height, 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 (often the increment to a marginal tax rate) is known and the goal is to estimate a behavioral parameter (often a labor-supply elasticity). By contrast, 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 a cost-function kink.

Our results show a modest increase in the marginal cost of floor space above a parcel's regulated building height. We use these estimates to gauge the extent of the increase in floor space that would result from the removal of building-height regulation in NYC. This exercise

[11] 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 $\overline{\mathrm{FAR}}$ groups, respectively, but are larger for the remaining groups, equal to 31 and 38% for the 0.5 and 0.9 $\overline{\mathrm{FAR}}$ groups, respectively.

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.

## Appendix

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 $\ell$, FAR based on the entire lot area is thus given by $F = \phi^{1-\rho}S^{\rho}$, where $S$ again equals $K/\ell$. 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 $F^{\phi}$, 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. $\mu$) 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.

**Table 1: FAR groups**

| $\overline{FAR}$ | *Observations* | *Floors* | *Floors/FAR* | *Zoning categories* |
|---|---|---|---|---|
| 2.0 | 2549 | 3.20 | 1.83 | R5D, R6B, M1-2, M1-4, M2-1, M2-3, M3-1, M3-2 |
| 1.25 | 3509 | 2.84 | 2.26 | R5 |
| 0.9 | 5295 | 2.39 | 2.59 | R4, R4-1, R4-A, R4-B |
| 0.6 | 9694 | 2.09 | 3.31 | R3-1, R3-2, R3-A, R3-X |
| 0.5 | 1334 | 2.02 | 3.84 | R1-1, R1-2, R1-2A, R2, R2-A |

**Table 2: Estimated $\theta$ values and confidence intervals**

| $\overline{FAR}$ | *Observations* | $\delta$ | $\widehat{\theta}$ | *Confidence int.* | *avg.* $\widehat{\theta}$ | $\widehat{(\alpha+\beta)}/\alpha$ | *Confidence int.* |
|---|---|---|---|---|---|---|---|
| 2.0 | 2549 | 0.15 | 1.119 | [1.080, 1.163] | 1.121 | 1.063 | [1.042, 1.085] |
| 1.25 | 3509 | 0.10 | 1.034 | [1.013, 1.058] | 1.035 | 1.018 | [1.007, 1.031] |
| 0.9 | 5295 | 0.15 | 1.065 | [1.036, 1.095] | 1.066 | 1.035 | [1.019, 1.050] |
| 0.6 | 9694 | 0.04 | 1.061 | [1.047, 1.075] | 1.061 | 1.032 | [1.025, 1.040] |
| 0.5 | 1334 | 0.04 | 1.123 | [1.073, 1.176] | 1.125 | 1.065 | [1.039, 1.091] |

$\widehat{\theta}$ is the estimated value of $((\alpha+\beta)/\alpha)^{1/\lambda}$, generated by assuming the given value of the interval parameter $\delta$. The 95% confidence interval for $\theta$ is based on a standard error generated through a 500-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 $\lambda = 0.54$, and its confidence interval (based on the $\theta$ interval) is also shown.

Table 3: Sensitivity analysis

| $\overline{FAR}$ | $\delta$ | $\widehat{\theta}$ |
| --- | --- | --- |
| 2.0 | 0.15 | 1.119 |
| | 0.10 | 1.075 |
| | 0.20 | 1.128 |
| 1.25 | 0.10 | 1.035 |
| | 0.05 | 1.041 |
| | 0.15 | 1.016 |
| 0.9 | 0.15 | 1.065 |
| | 0.125 | 1.030 |
| | 0.175 | 1.237 |
| 0.6 | 0.04 | 1.061 |
| | 0.03 | 1.037 |
| | 0.05 | 1.106 |
| 0.5 | 0.04 | 1.123 |
| | 0.03 | 1.076 |
| | 0.05 | 1.144 |

Table 4: Floor space gain without FAR regulation

| $\overline{FAR}$ | $SPACE_{now}$ | $SPACE_{new}$ | % gain |
| --- | --- | --- | --- |
| 2.0 | 16,477,906 | 17,602,509 | 6.8% |
| 1.25 | 12,194,765 | 12,466,904 | 2.2% |
| 0.9 | 14,611,072 | 15,233,868 | 4.3% |
| 0.6 | 17,718,632 | 18,404,479 | 3.9% |
| 0.5 | 3,997,145 | 4,377,899 | 9.5% |

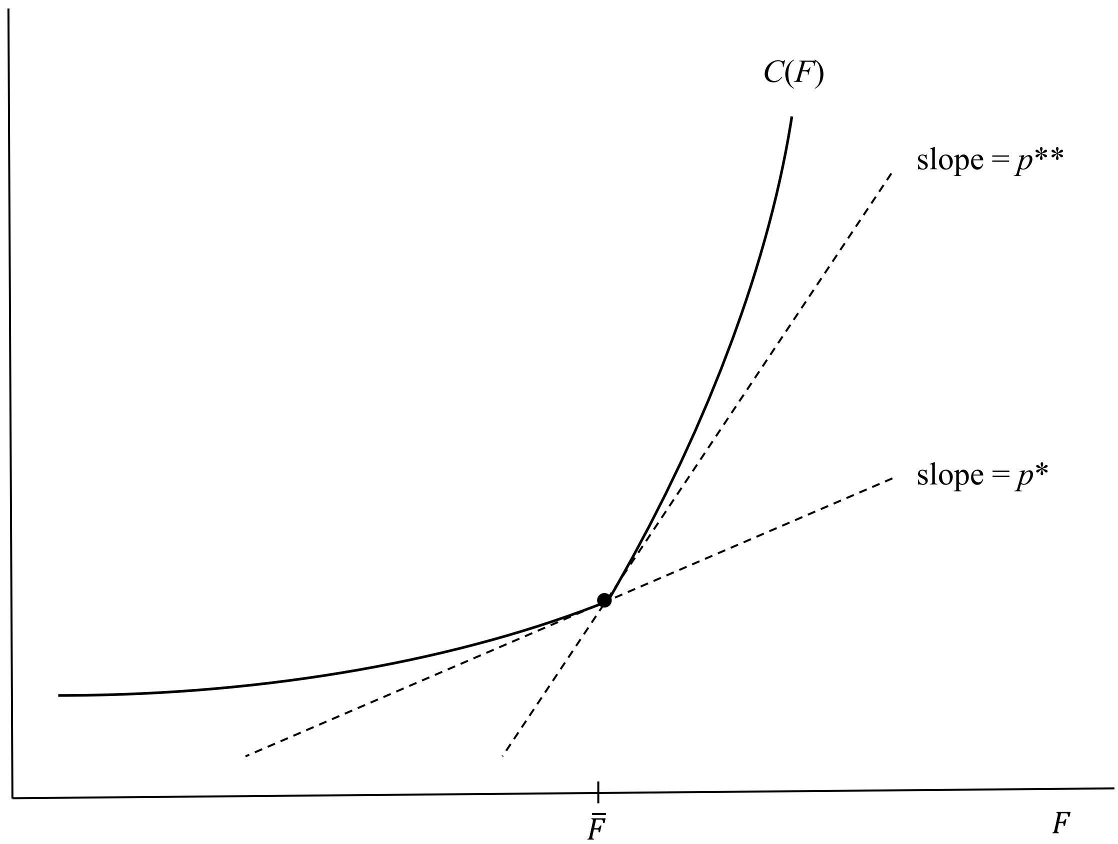
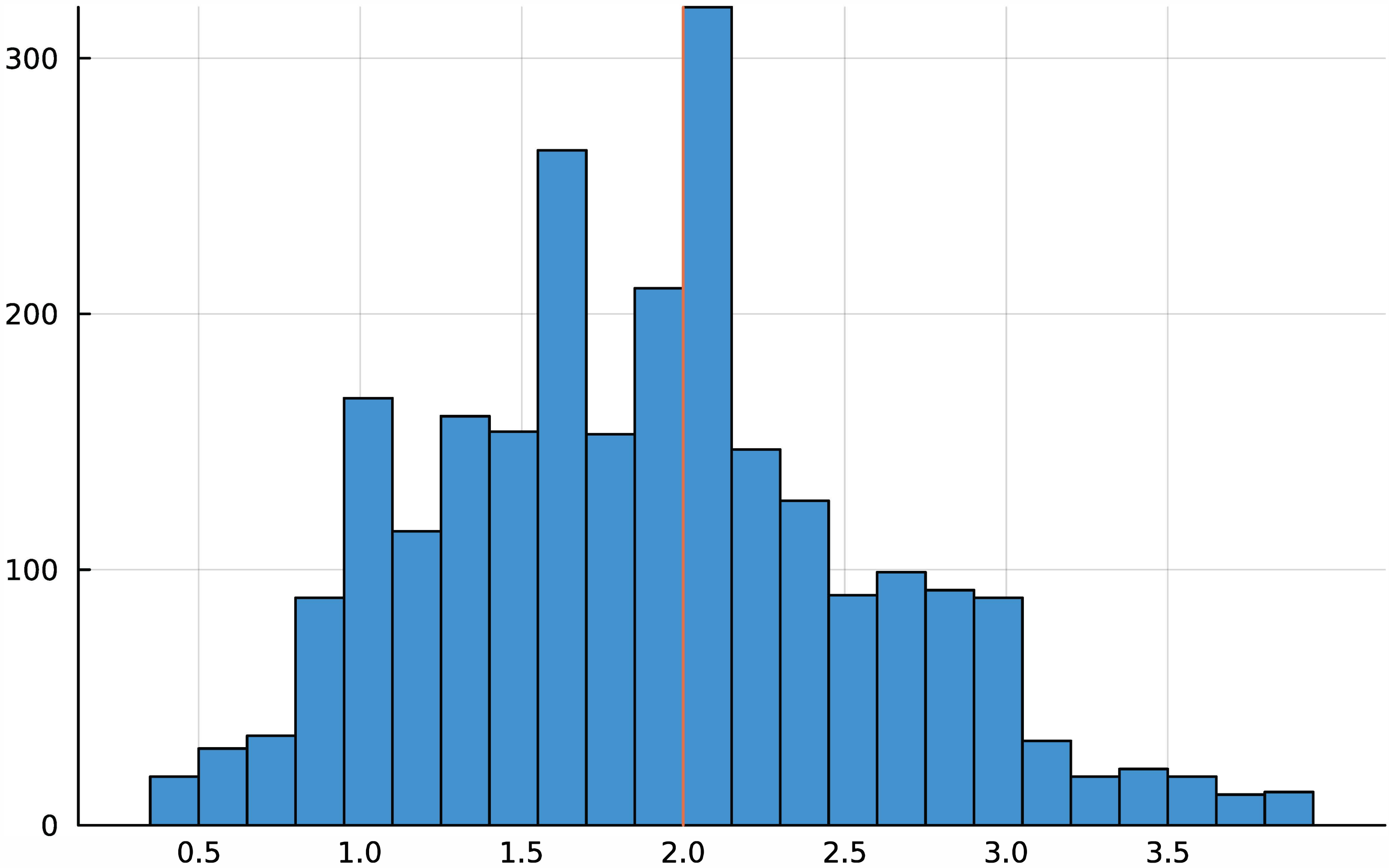



Figure 1: Kink in $C(F)$

Figure 2: FAR Distribution for $\bar{F}$ = 2.0



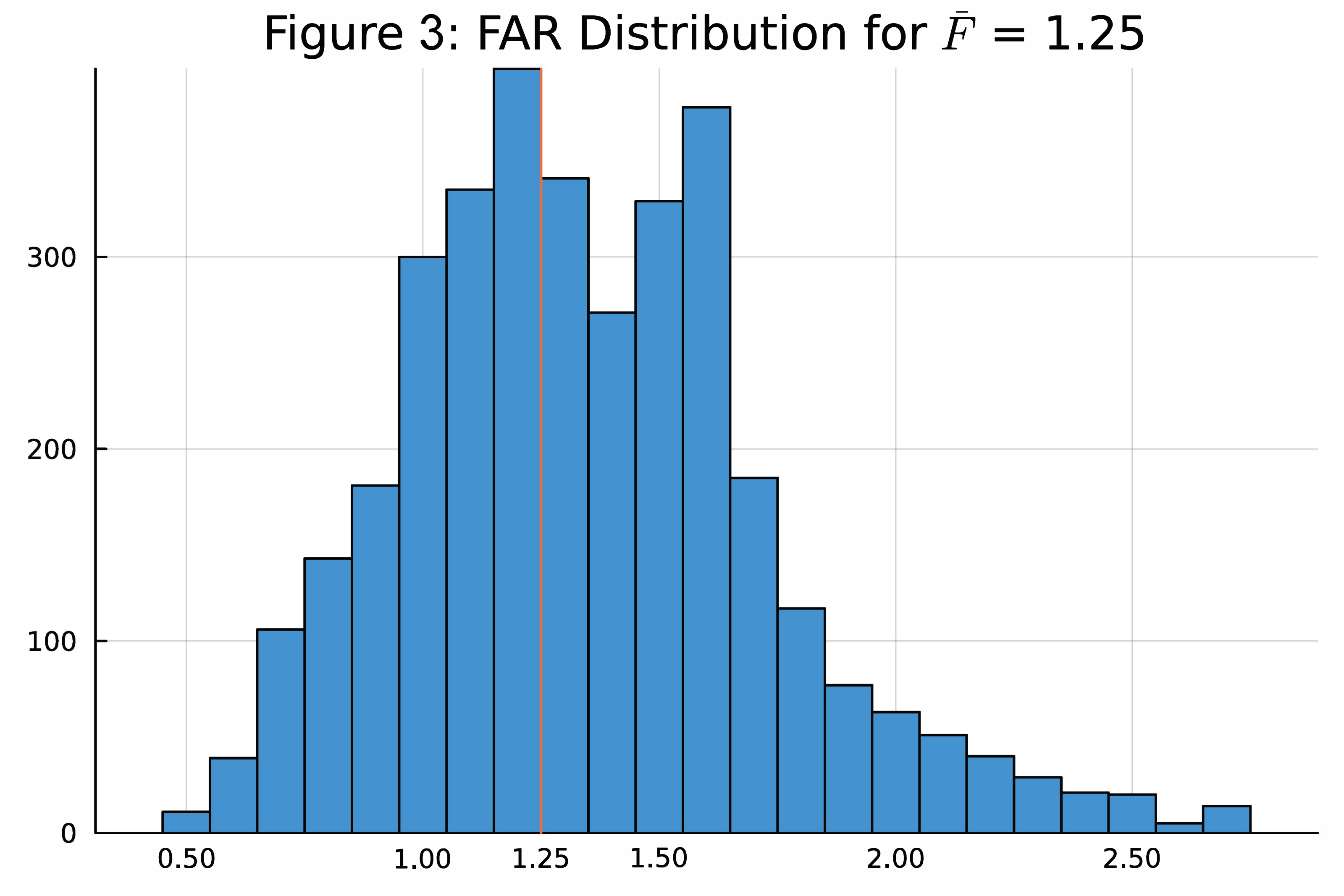


Figure 3: FAR Distribution for $\bar{F}$ = 1.25
300
200
100
0
0.50
1.00
1.25
1.50
2.00
2.50

Figure 4: FAR Distribution for $\bar{F}$ = 0.9

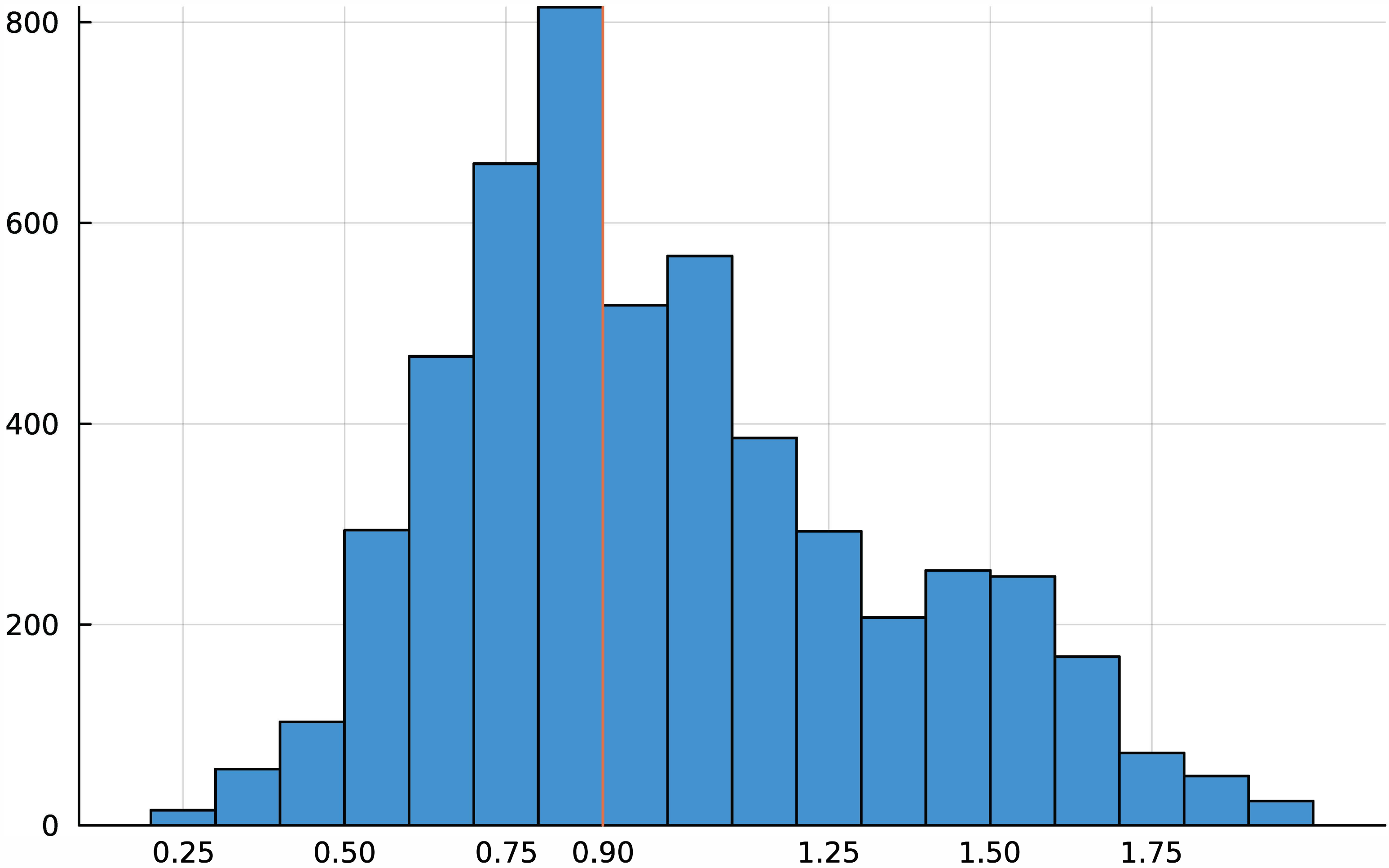


Figure 5: FAR Distribution for $\bar{F}$ = 0.6

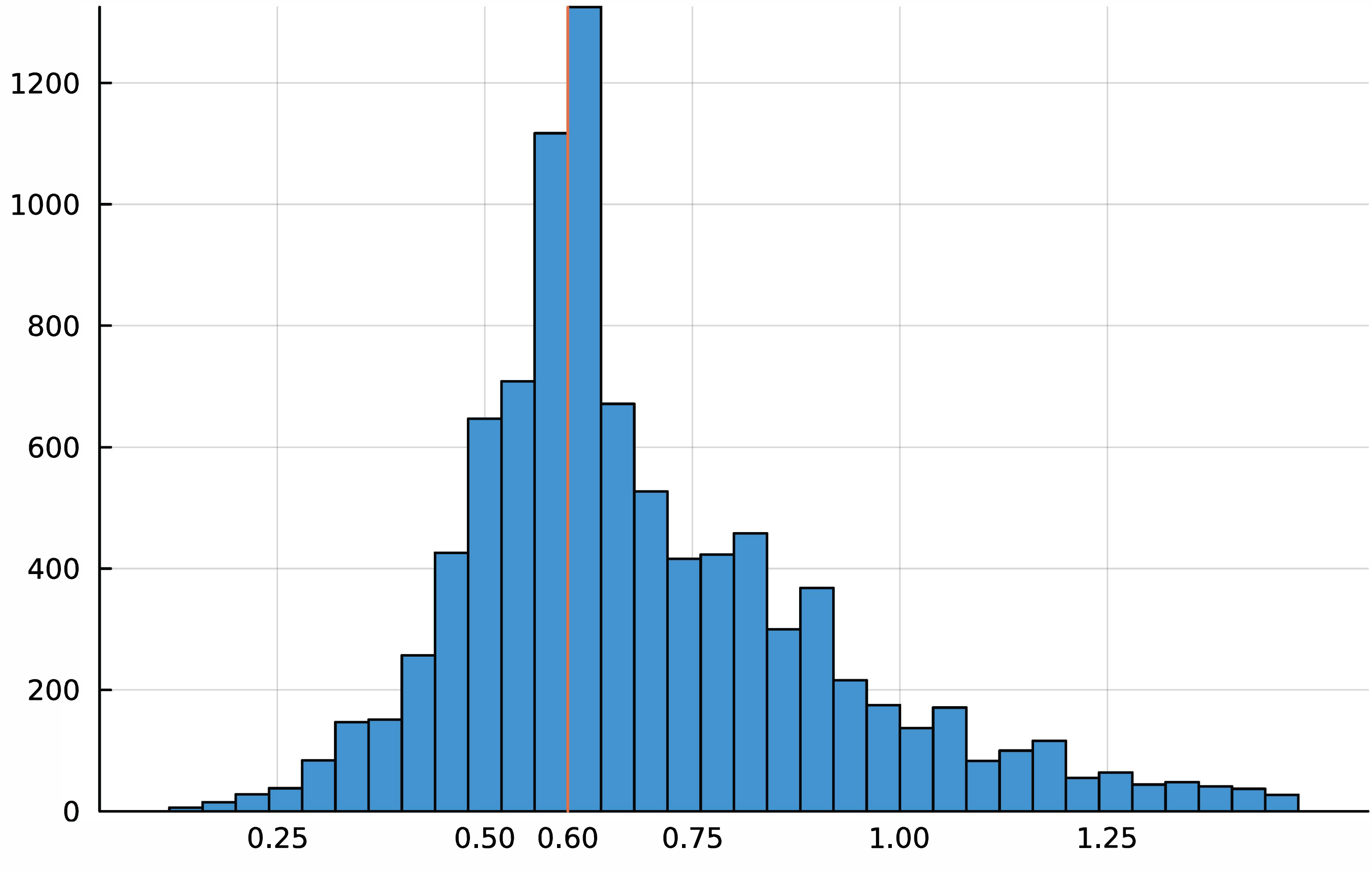


Figure 6: FAR Distribution for $\bar{F}$ = 0.5

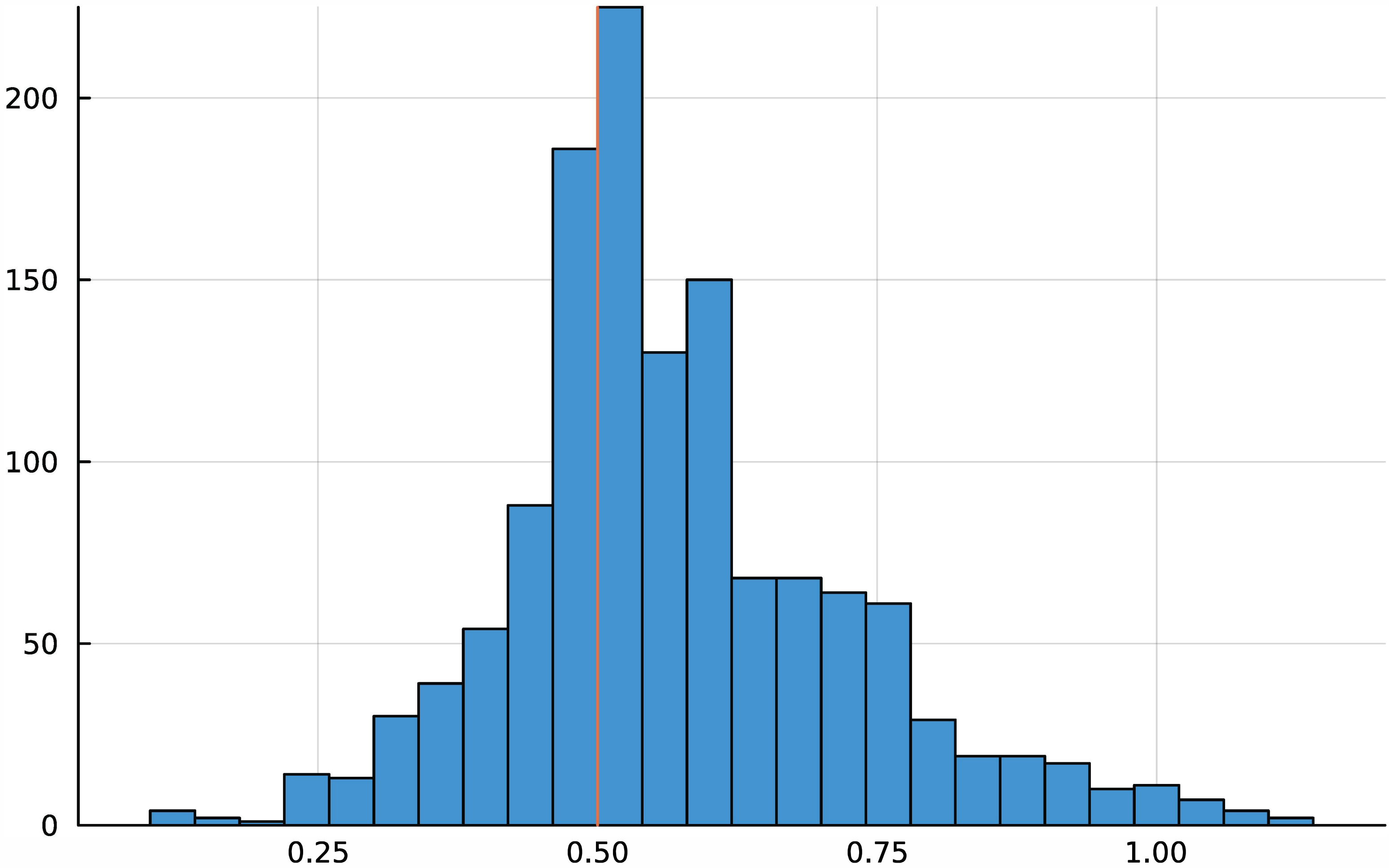
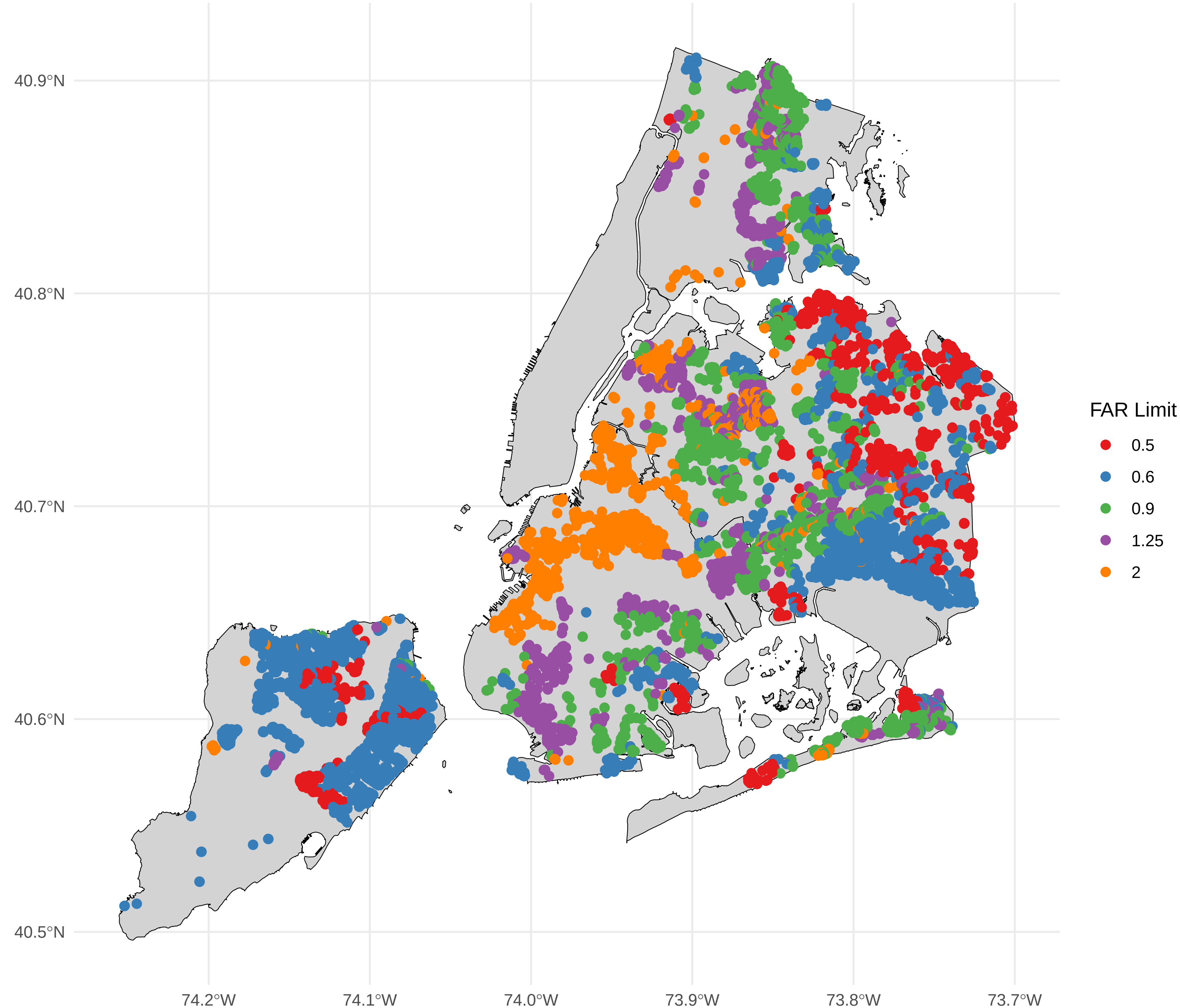


Figure 7: Location of Sample Properties by FAR Limit



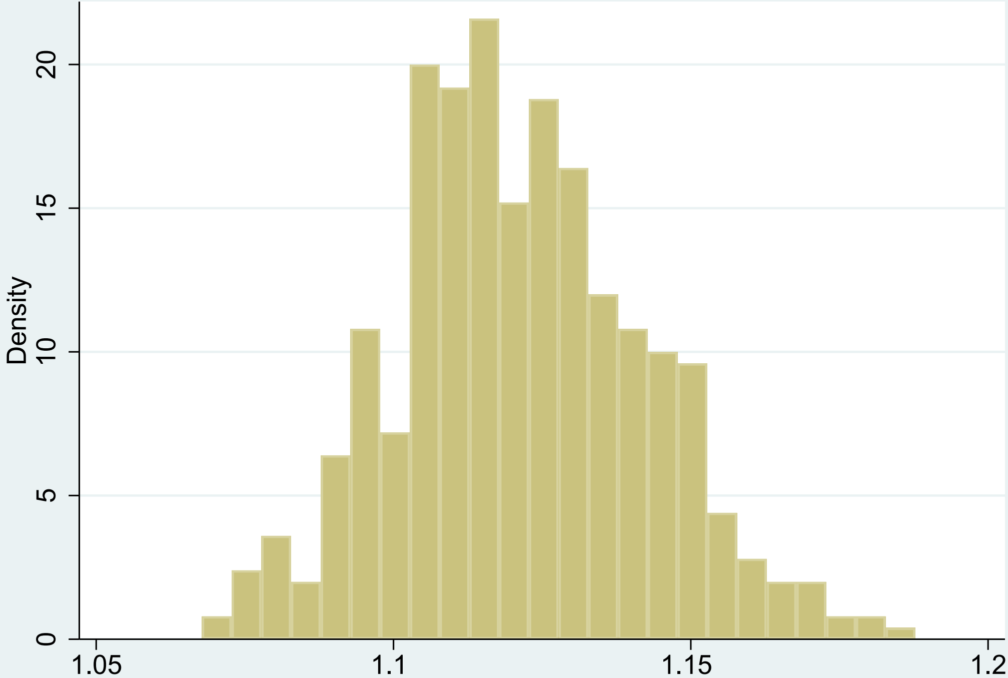
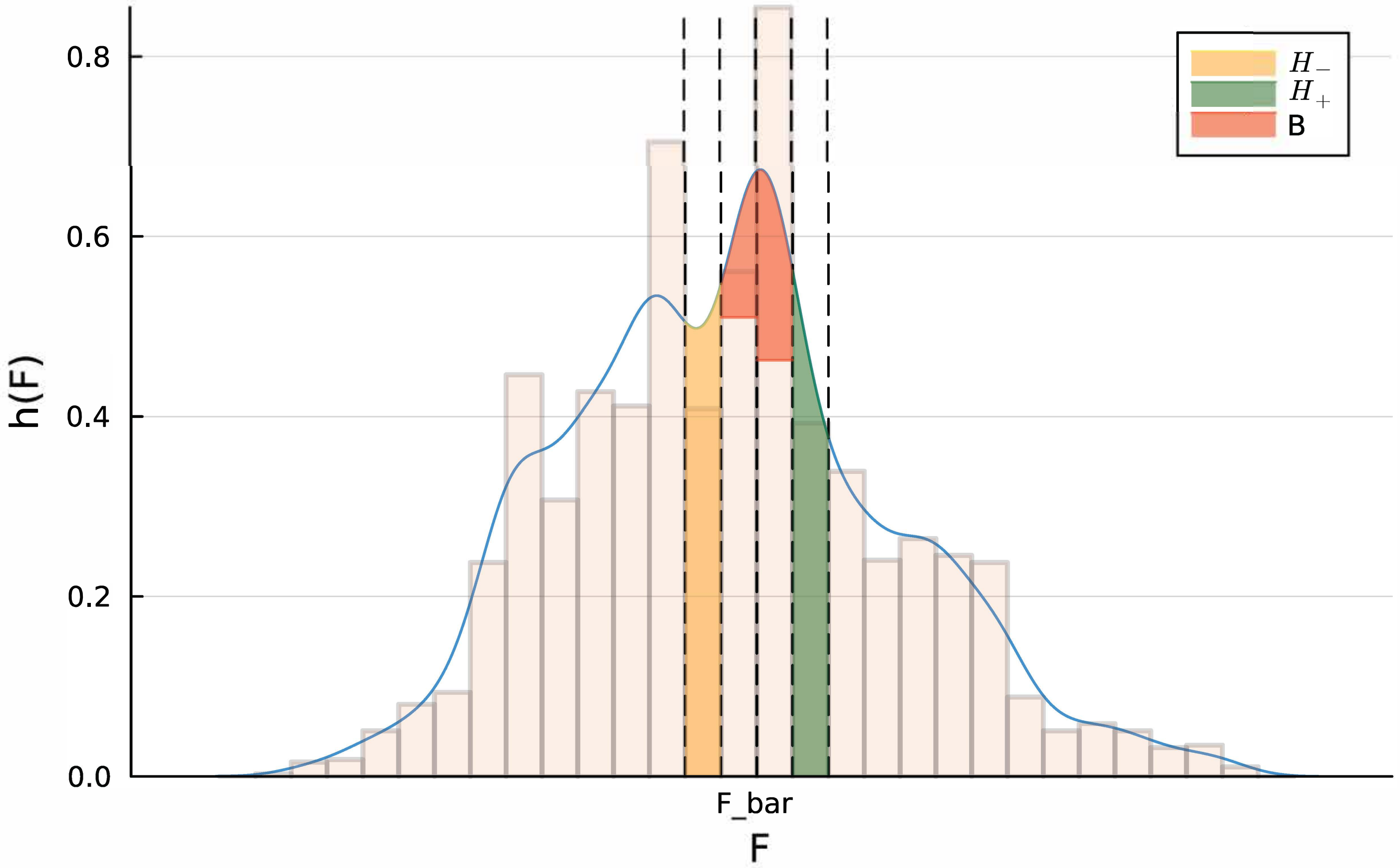


Figure 8: Distribution of boostrap theta estimates

# Figure 9: Bunching Areas



# 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.

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.

Duranton, G., Combes, P.-P., Gobillon, L., The production function for housing: Evidence from France. *Journal of Political Economy* 129, 2766-2816.

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.

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`.

Online Appendix

# 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 marginally. 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 $\varepsilon \leq (\theta-1)\,\overline{F}$. The increases 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:

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

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:

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

where the last equality uses the trapezoid approximation.

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

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

Using again the trapezoid approximation, we get the following expression

$$\begin{aligned}
D_3(\varepsilon) &= \theta\varepsilon\left(\overline{F}+\varepsilon\right)\left(\frac{h_0(\overline{F}+\Delta\overline{F}) + h_0(\overline{F}+\Delta\overline{F}+\theta\varepsilon)}{2}\right) \\
&\quad - \frac{\theta\varepsilon}{\theta}\left(\frac{(\overline{F}+\Delta\overline{F})h_0(\overline{F}+\Delta\overline{F}) + (\overline{F}+\Delta\overline{F}+\theta\varepsilon)h_0(\overline{F}+\Delta\overline{F}+\theta\varepsilon)}{2}\right) \\
&= \frac{\theta\varepsilon}{2}\left[h_0(\overline{F}+\Delta\overline{F})\left(\overline{F}+\varepsilon - \left(\frac{\overline{F}+\Delta\overline{F}}{\theta}\right)\right) + h_0(\overline{F}+\Delta\overline{F}+\theta\varepsilon)\left(\overline{F}+\varepsilon - \left(\frac{\overline{F}+\Delta\overline{F}+\theta\varepsilon}{\theta}\right)\right)\right].
\end{aligned}$$

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 $\varepsilon$ as $D(\varepsilon)$. It equals

$$
\begin{aligned}
D(\varepsilon) &= D_1(\varepsilon) + D_2(\varepsilon) + D_3(\varepsilon) \\
&= \frac{\varepsilon^2}{2} h_0(\overline{F}+\varepsilon) + \varepsilon\left(\Delta\overline{F} - \varepsilon\right)\left(\frac{h_0(\overline{F}+\varepsilon) + h_0(\overline{F}+\Delta\overline{F})}{2}\right) + \frac{\theta\varepsilon^2}{2} h_0(\overline{F}+\Delta\overline{F}) \\
&= \varepsilon\Delta\overline{F}\left(\frac{h_0(\overline{F}+\varepsilon) + h_0(\overline{F}+\Delta\overline{F})}{2}\right) + \frac{\varepsilon^2\left(\theta-1\right)}{2} h_0(\overline{F}+\Delta\overline{F}).
\end{aligned}
$$

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. The derivative of total floor space with respect the limit is $B$ times lot size.

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 $\varepsilon$ is not small.

We can rearrange $D(\varepsilon)$ to be a function of $B$ by adding and subtracting some terms:

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

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 $\theta$ 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

$$
\begin{aligned}
h_1\left((F+\varepsilon)\theta^{-1}\right) &= t\left((\alpha+\beta)((F+\varepsilon)\theta^{-1})^{\lambda}\right)(\alpha+\beta)\lambda((F+\varepsilon)\theta^{-1})^{\lambda-1} \\
&= t\left((\alpha+\beta)((F+\varepsilon))^{\lambda}\left(\frac{\alpha}{\alpha+\beta}\right)\right)(\alpha+\beta)\lambda(F+\varepsilon)^{\lambda-1}\left(\frac{\alpha}{\alpha+\beta}\right)^{\frac{\lambda-1}{\lambda}} \\
&= t\left(\alpha((F+\varepsilon))^{\lambda}\right)\alpha\lambda(F+\varepsilon)^{\lambda-1}\left(\frac{\alpha+\beta}{\alpha}\right)^{1/\lambda} = h_0\left(F+\varepsilon\right)\theta.
\end{aligned}
$$

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:

$$
\begin{aligned}
D(\varepsilon) &= \varepsilon B + \frac{\varepsilon^2}{2}\left(\frac{\theta-1}{\theta}\right)h_1(\overline{F}) - \frac{\varepsilon(\theta-1)\overline{F}}{2}\left(h_0(\overline{F}) - \theta^{-1}h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right)\right) \\
&= \varepsilon B + \frac{\varepsilon}{2}\left(\frac{\theta-1}{\theta}\right)\varepsilon h_1(\overline{F}) - \frac{\varepsilon}{2}\left(\frac{\theta-1}{\theta}\right)\left(\theta\overline{F}h_0(\overline{F}) - \overline{F}h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right)\right) \\
&= \varepsilon B + \frac{\varepsilon}{2}\left(\frac{\theta-1}{\theta}\right)\left(\varepsilon h_1(\overline{F}) + \overline{F}h_1\left((\overline{F}+\varepsilon)\theta^{-1}\right) - \theta\overline{F}h_0(\overline{F})\right).
\end{aligned}
$$

The only new object to estimate after estimating $\theta$ 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:

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