

Insper Instituto de Ensino e Pesquisa  
Ciências Econômicas

Gustavo Theil

## **Dez anos de plano diretor: Repercussões da regulação urbana na densidade populacional**

São Paulo - SP  
2024

Gustavo Theil

# **Dez anos de plano diretor: Repercussões da regulação urbana na densidade populacional**

Trabalho de Conclusão de Curso

Insper Instituto de Ensino e Pesquisa  
Ciências Econômicas

Orientador: Adriano Borges Costa

São Paulo - SP  
2024

---

Gustavo Theil

Dez anos de plano diretor: Repercussões da regulação urbana na densidade populacional/ Gustavo Theil. - São Paulo - SP: 2024

48 p.

: Insper Instituto de Ensino e Pesquisa  
Ciências Econômicas , 2024.

Orientador: Adriano Borges Costa

1. Economia Urbana 2. Política pública I. Gustavo Theil. II. Dez anos de plano diretor: Repercussões da regulação urbana na densidade populacional.

---

# Resumo

Resumo

Palavras-chave:

# Abstract

Abstract

Keywords:

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>7</b>
1.1	Motivação	7
1.2	Plano Diretor Estratégico (PDE)	8
1.3	O problema	11
<b>2</b>	<b>REVISÃO LITERÁRIA</b>	<b>13</b>
2.1	Importância da regulamentação	13
2.2	Evidências do PDE de 2014	15
<b>3</b>	<b>DADOS</b>	<b>16</b>
3.1	Padrões Construtivos	16
3.2	Regulação	17
3.3	Densidade populacional	17
3.4	Cruzamento dos dados	18
<b>4</b>	<b>METODOLOGIA E RESULTADOS</b>	<b>20</b>
4.1	<b>Do padrão construtivo à densidade populacional</b>	<b>20</b>
4.1.1	Abordagem de regressão linear	21
4.1.2	Abordagem de <i>machine learning</i>	22
4.1.3	Resultados principais	23
4.2	<b>Da regulação ao padrão construtivo</b>	<b>24</b>
4.2.1	Abordagem <i>diff-in-diff</i> geográfico (e <i>event study</i> )	25
4.2.2	Abordagem de RDD	27
4.2.3	Resultados principais	27
4.3	<b>Da regulação à densidade populacional</b>	<b>29</b>
4.3.1	Abordagem via diferenças em diferenças com efeito heterogêneo	32
4.3.2	Resultados principais	32
<b>5</b>	<b>CONCLUSÃO</b>	<b>34</b>
5.1	Resultados principais	34
5.2	Contribuições para o debate público	34
	<b>REFERÊNCIAS</b>	<b>35</b>

<b>APÊNDICES</b>	<b>36</b>
<b>APÊNDICE A – DISCUSSÕES SOBRE A VERTICALIZAÇÃO . . . . .</b>	<b>37</b>
<b>APÊNDICE B – DISCUSSÕES SOBRE A INFORMALIDADE . . . . .</b>	<b>39</b>
<b>APÊNDICE C – METODOLOGIA DE CRUZAMENTO GEOGRÁFICO DOS DADOS . . . . .</b>	<b>40</b>
<b>APÊNDICE D – OUTRAS FIGURAS E TABELAS . . . . .</b>	<b>41</b>

# 1 Introdução

## 1.1 Motivação

As cidades são repletas de agentes, cada um com diferentes interesses. Os cidadãos desejam maximizar a qualidade de vida, o que geralmente significa habitar residências espaçosas, em lugares tranquilos, repletos de natureza, mas que ainda tenham acesso às oportunidades de emprego, lazer, saúde, educação, entre outros. As imobiliárias e construtoras desejam construir prédios, casas e condomínios que maximizem seus lucros, escolhendo as localizações com maior demanda e menores custos. Os proprietários de terras e imóveis buscam comprar ativos baratos com a expectativa de que valorizem no futuro ou extraiam renda de aluguel.

Porém, a dinâmica das cidades é moldada por inúmeros conflitos de interesses, visto que muitos deles são antagônicos. É impossível, por exemplo, que todo mundo more em contato com a natureza, já que na medida em que cada família constrói sua casa, a natureza do vizinho diminui. Dessa forma, há duas maneiras de coexistir no espaço: através das relações de poder, nas quais, o agente mais poderoso tem seu interesse atendido (ou negociado) em detrimento de outro, ou um planejador central intermedia este conflito de forma a gerar o resultado mais favorável para todos.

A principal instituição responsável por intermediar os conflitos no que tange a esfera urbana de São Paulo, a cidade mais populosa das américas, é o Plano Diretor Estratégico (PDE). O PDE é uma lei municipal que deve ser revisado a cada 10 anos, com uma revisão intermediária depois de 5 anos<sup>1</sup>. O plano diretor de 2014 trouxe mudanças significativas em relação ao seu predecessor, sendo reconhecido internacionalmente pela sua abordagem de combate às desigualdades e incentivo ao uso do transporte público.

Enquanto o plano de 2002 descentralizou grande parte das decisões sobre zoneamento nas subprefeituras, o PDE de 2014 seguiu uma abordagem moderna, em linha com a literatura de urbanismo de *Transit Oriented Development* (TOD), em português, “desenvolvimento orientado ao transporte”. Antigamente, cada subprefeitura teria protagonismo no zoneamento de sua região, enquanto no novo plano, a regulação foi definida a nível municipal e centralizado. O sistema antigo apresentava grande dificuldade em lidar com os conflitos de interesse e grupos locais tinham a capacidade de conter o adensamento em áreas desejáveis, prejudicando o desenvolvimento sustentável da cidade.

A literatura de TOD, que surgiu no final da década de 80, mas popularizou-se apenas

---

<sup>1</sup> Estas datas historicamente não foram respeitadas à risca. O PDE de 2002 ficou em vigor 12 anos, até a vigência do plano de 2014 começar. A revisão do PDE de 2014 ocorreu em 2023. Isso geralmente decorre dos atrasos administrativos e debates prolongados

depois dos anos 2000, defende a mobilidade como um dos principais pilares do desenvolvimento das cidades (IBRAEVA et al., 2020). A proposta do TOD é aproximar as famílias às oportunidades, através do direcionamento do desenvolvimento urbano no entorno de infraestrutura de transporte coletivo, causando uma melhora na mobilidade e maior adensamento. Dessa forma, diminui-se a dependência do carro, que gera externalidades negativas, e incentiva-se o uso do transporte público de alta capacidade, cujas externalidades são mais positivas, ou menos negativas do que o transporte individual motorizado.

As ideias do TOD não apenas influenciam o PDE de 2014, como são uma parte estruturante. Entre os objetivos enunciados do plano diretor, destacam-se os quatro primeiros:

Art. 7º A Política de Desenvolvimento Urbano e o Plano Diretor Estratégico se orientam pelos seguintes objetivos estratégicos:

I - conter o processo de expansão horizontal da aglomeração urbana, contribuindo para preservar o cinturão verde metropolitano;

II - acomodar o crescimento urbano nas áreas subutilizadas dotadas de infraestrutura e no entorno da rede de transporte coletivo de alta e média capacidade

III - reduzir a necessidade de deslocamento, equilibrando a relação entre os locais de emprego e de moradia;

IV - expandir as redes de transporte coletivo de alta e média capacidade e os modos não motorizados, racionalizando o uso de automóvel;

Dito isso, a pergunta motivadora desta pesquisa é se 10 anos depois do início da vigência do plano diretor é possível identificar se ele está cumprindo com seus objetivos – em especial, o objetivo de adensar as áreas na proximidade da infraestrutura de transporte público.

## 1.2 Plano Diretor Estratégico (PDE)

Esta seção é dedicada a compreender quais ferramentas estão à disposição do plano diretor para que alcance seus objetivos. Todavia, o PDE apresenta diversos instrumentos que agem não apenas em prédios novos, como também atua na requalificação de lotes já construídos. O foco desta pesquisa está nos três principais instrumentos desenhados para a regulamentação de novas construções. Estes são o coeficiente de aproveitamento, a cota parte e o gabinete. Outros instrumentos, apesar de importantes, não têm como objetivo incentivar ou desincentivar a densidade populacional de cada área.

O PDE dividiu o mapa da cidade em macroáreas, macrozonas e zonas especiais, cada uma com objetivos específicos e respectivas restrições para cada instrumento. Além disso, criou um regime especial para lotes que estejam próximos à infraestrutura de transporte público de alta capacidade, região nomeada de Eixos de Estruturação da Transformação Urbana (EETUs).

## Eixos de Estruturação da Transformação Urbana (EETUs)

Os eixos são ativados pela proximidade ao transporte público de alta capacidade, no caso de São Paulo, isso pode se dar de duas maneiras. A primeira forma de ativação é através das estações de trens, metrôs ou monotrilhos, que criam uma área de influência em um raio de 400m ao redor do ponto de acesso à estação. A segunda forma é através de corredores de ônibus municipais e intermunicipais, que geram uma área de influência de 150m de distância para cada lado da via ao longo do corredor.

Nas zonas de eixos, além haver um regime especial para os três instrumentos que serão apresentados, há outras medidas que almejam incentivar o uso do transporte coletivo. Para que as pessoas utilizem menos o carro, por exemplo, criou-se restrições para o número de vagas de estacionamento, que antes eram incentivadas no PDE de 2002.

## Coeficiente de Aproveitamento (CA)

O Coeficiente de Aproveitamento (CA) é um instrumento utilizado mundialmente<sup>2</sup> em planos diretores para regular a **densidade construtiva**. O CA determina quantas vezes a área do lote pode ser construída. A legislação paulistana prescreve no direito de propriedade de todos os lotes da cidade um coeficiente de aproveitamento básico igual a 1, ou seja, é permitido construir uma vez a área do terreno. Se um proprietário desejar construir 4 andares em seu lote, com uma CA de 1, pode ocupar apenas 1/4 da área do terreno. Na ótica da legislação, o potencial de construir metros quadrados a mais de uma vez a área do terreno não está incluso no direito de propriedade, e, portanto pertence à sociedade, não ao proprietário do lote.

Em cada macroárea da cidade há restrições diferentes para o CA mínimo e máximo – o básico se mantém 1 em toda a cidade. Na Macrozona de Estruturação e Qualificação Urbana, o CA mínimo varia entre 0,3 e 0,5 a depender de qual macroárea se encontra, e o máximo é sempre 2. Na Macrozona de Proteção e Recuperação Ambiental, não há CA mínimo e o CA máximo varia entre 0,1 e 1, a depender do nível de proteção que foi designado à região. Nessas áreas de preservação outras regulações também podem atuar<sup>3</sup>.

Quando a região é ativada por transporte público de alta capacidade, há modificadores sobre o CA. Na Macrozona de Estruturação e Qualificação Urbana, caso haja região de eixo, o CA máximo aumenta para 4. Na Macrozona de Proteção e Recuperação Ambiental, contanto que a área esteja fora da região de proteção aos mananciais, o CA máximo torna-se 2. Dessa forma, o PDE permite o dobro de densidade construtiva no entorno da infraestrutura de transporte, demonstrando seu compromisso com o TOD.

<sup>2</sup> Em inglês, *Floor Area Ratio* (FAR) ou *Floor Area Ratio*

<sup>3</sup> Aplica-se a legislação estadual pertinente, especialmente as leis específicas das Bacias Billings e Guarapiranga

## Gabarito

O gabarito também é um instrumento comum de se regular em planos diretores para controlar a **verticalização** na cidade. O gabarito, dentre os instrumentos discutidos, é o mais visível ao olho das pessoas que passam na rua ou da vista aérea, quando se chega de avião no aeroporto de Guarulhos ou Congonhas. Na Macrozona de Estruturação e Qualificação Urbana, o gabarito máximo é de 28m ou, contabilizando-se em pavimentos, o térreo mais 8 andares. Nas áreas de Macrozona de Proteção e Recuperação Ambiental, o gabarito máximo é de 15m, ou térreo mais 4 pavimentos.

Como uma forma de incentivo à produção imobiliária nos eixos, na Macrozona de Estruturação e Qualificação Urbana não há limite de gabarito para as construções, que passam a ser limitadas apenas pelo CA. Os eixos em Macrozona de Proteção e Recuperação Ambiental apresentam gabarito máximo de 28m.

Um detalhe importante de se pontuar é que há uma relação de identidade entre CA, verticalização e taxa de ocupação. A taxa de ocupação é o percentual da área do terreno que recebe edificação. Por exemplo, em um terreno com CA igual a 4, caso sejam construídos 8 andares, a taxa de ocupação precisa ser 50%. Analogamente, se taxa de ocupação for definida em 100%, por exemplo, o número de pavimentos deve ser 4. Mais detalhes sobre esta discussão, bem como um comparativo entre a densidade construtiva e a verticalização encontram-se no Apêndice A.

## Cota Parte

A cota parte, por sua vez, não é um instrumento comum de se observar em planos diretores e pode ser considerada uma medida relativamente experimental – diferentemente do CA e gabarito que são instrumentos consolidados. Este instrumento é responsável por regulamentar a **densidade habitacional**. Sabendo-se a cota parte ( $Q$ ) e a área do terreno ( $A_t$ ), é possível identificar o número mínimo de unidades habitacionais ( $N_{min}$ ) que o lote deve apresentar, seguindo a Equação 1.1 a seguir.

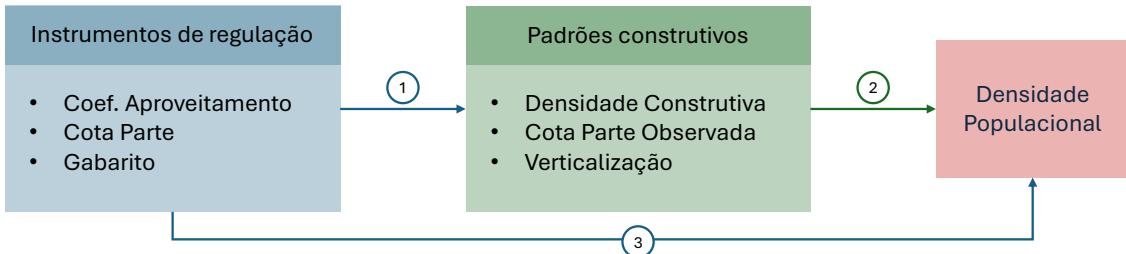
$$N_{min} = \frac{CA_{utilizado}}{CA_{max}} \cdot \frac{A_t}{Q} \quad (1.1)$$

Um lote, por exemplo, de  $1000m^2$  com cota parte de 20, caso utilize o CA máximo, precisa construir no mínimo 50 unidades. É interessante notar que o CA do lote não interfere no número mínimo de unidades habitacionais que devem ser produzidas, ao menos que o CA utilizado seja menor do que o máximo, e, neste caso, o número mínimo de unidades que devem ser providas é menor. Todavia, a cota parte não é um instrumento que funciona para a cidade toda. Na verdade, a única região em que a cota parte atua é a região de EETUs, nas quais a cota parte é de 20 no caso da Macrozona de Estruturação e Qualificação Urbana e de 40 em Macrozona de Proteção e Recuperação Ambiental.

## 1.3 O problema

Para avaliar se o PDE de fato conseguiu cumprir com seus objetivos, três perguntas precisam ser respondidas. O diagrama da Figura 1 contribui para compreender cada uma delas.

Figura 1 – Dos instrumentos de regulação até o resultado esperado



**A regulação foi capaz de influenciar os padrões construtivos na cidade da maneira esperada?**

Através dos instrumentos apresentados na Seção 1.2, o PDE pretendia que o desenvolvimento imobiliário se concentrasse na Macrozona de Estruturação e Qualificação Urbana, em especial nos EETUs, que receberam um novo conjunto de regras. Em grande parte, essas mudanças permitiram um adensamento construtivo maior, verticalização ilimitada e impuseram uma densidade habitacional mínima. Todavia, como mencionado na Seção 1.1, a cidade é composta de diversos interesses, sendo o mercado imobiliário um agente central nesse aspecto.

O mercado imobiliário, responsável por de fato construir novas habitacões, responde à regulação sempre buscando maximizar seu lucro. Como a regulação não criou incentivos e na verdade, apenas mudou permissões em cada região da cidade, não necessariamente o mercado imobiliário aceita o novo conjunto de regras como um convite para construir a tipologia que o PDE deseja ver nos EETUs. Dentro dos limites da legislação de CA entre 2 e 4, há muita liberdade para o mercado escolher o que construir. Ademais, sempre existe a opção de não construir nessas regiões, quando o benefício do maior potencial construtivo não é suficiente para compensar a taxação advinda da extração do CA básico.

**Os padrões construtivos regulamentados e incentivados são de fato capazes de gerar a densidade esperada?**

Em outras palavras, é possível afirmar que mais densidade construtiva, mais verticalização e mais densidade habitacional equivale a mais densidade populacional? Apesar de simples a pergunta, há certas nuances. Primeiramente, é importante identificar se os três componentes do padrão construtivo são importantes e qual apresenta maior relevância na determinação da densidade populacional.

Caso a densidade construtiva seja um determinante para a densidade populacional, o instrumento do CA é importante. De maneira análoga, caso a densidade habitacional ou a verticalização se destaquem, a cota parte e o gabarito também são importantes. Nesse sentido, responder esta pergunta ajuda a identificar quais instrumentos analisados na pergunta 1 são mais importantes.

## A regulação foi capaz de influenciar a densidade populacional na cidade?

O PDE propôs uma relação causal, ilustrada na Figura 1, que pressupõe que, se as duas primeiras condições forem atendidas, o aumento da densidade populacional será uma consequência direta. Ou seja, a terceira pergunta se tornaria desnecessária caso as duas primeiras se confirmassem.

Contudo, essa relação não é garantida, pois outros fatores também podem influenciar a densidade populacional, independentemente das regulamentações introduzidas. Assim, é essencial avaliar o efeito total da regulação sobre a densidade populacional, a fim de compreender o impacto completo e real das políticas implementadas.

## Direcionamento no artigo

No Capítulo 3 serão apresentados os dados utilizado nas análises. No Capítulo 2 é feita uma breve revisão literária. No Capítulo 4 será discutida a metodologia e resultados para cada uma das perguntas, na Seção 4.1, 4.2 e 4.3, respectivamente. No Capítulo 5 serão trazidas as reflexões finais, na Seção 5.1 o foco é nos principais resultados, e na 5.2 é feita uma discussão com implicação mais direta no debate público. Por fim, nos Apêndices há discussões de detalhes importantes, mas que tangenciam a narrativa do trabalho.

Os códigos utilizados para gerar os resultados apresentados estão disponíveis de maneira pública para verificação, replicação ou utilização para outros fins, contanto que este trabalho seja citado de maneira adequada<sup>4</sup>.

---

<sup>4</sup> O código pode ser acessado na página do GitHub [neste link](#).

## 2 Revisão Literária

### 2.1 Importância da regulamentação

Na literatura microeconômica há diversos modelos que tentam compreender as dinâmicas econômicas da cidade, analisando como as forças de mercado agem na ausência da regulação. Utilizando premissas formais e provas matemáticas, é possível resolver os modelos para algumas variáveis endógenas, como preço por metro quadrado na cidade, uso do solo, densidade populacional, entre outros. O que geralmente há de comum entre estes modelos, é que se o mercado receber um tratamento *laissez-faire*, se observará um maior adensamento no centro, onde há oferta de emprego. Quando há múltiplos centros, ou pontos de transporte público de alta capacidade, também há maior adensamento no entorno deles (PAPAGEORGIOU; PINES, 2012; FUJITA, 1989).

Isso acontece porque o adensamento traz diversos benefícios econômicos. A aglomeração tecnológica (*i*) aumenta a produtividade dos trabalhadores, na medida em que os empregos são mais concentrados e há “transbordamento” de conhecimento entre as firmas da região. Além disso, uma oferta maior e mais diversa de trabalho, causa maior competitividade e eficiência na escolha da pessoa certa para cada cargo. Aglomerações pecuniárias (*ii*) reduzem os custos das firmas, sem alterar sua produtividade. Com maior demanda por serviços como segurança, limpeza, contratação e advocacia, estes mercados se desenvolvem, tornam-se mais competitivos, eficientes e baratos. Inclusive, há serviços especializados de nicho, que podem estar disponíveis e acessíveis apenas em grandes centros urbanos. Aglomeração de varejo (*iii*) traz ganhos para os consumidores e comerciantes. Quando o comércio está aglomerado, o consumidor pode escolher entre mais opções e se desloca menos entre seus destinos caso queria comprar mais de um item. Dessa forma, os consumidores ganham e os comerciantes também, visto que com mais consumidores e maior fluxo, maiores as vendas. Por fim, o custo de transporte (*iv*) é um dos fatores que mais mudam quando há densidade. A redução do custo de transporte, que pode ser considerada uma economia de aglomeração pecuniária, acontece não apenas para os trabalhadores, que se deslocam menos às oportunidades de emprego, mas também às firmas que gastam menos transportando seus bens e serviços (BRUECKNER, 2011).

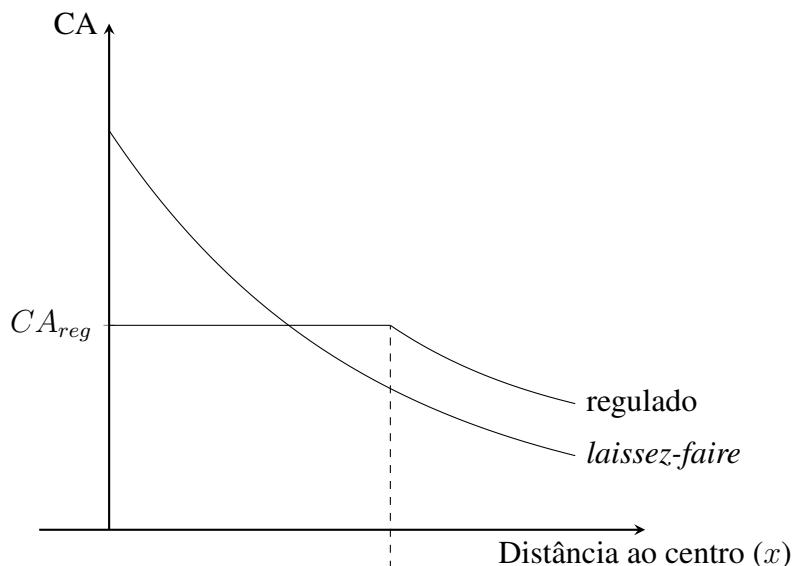
Entretanto, ao introduzir falhas de mercado nesses modelos, a organização da cidade na ausência de regulamentação pode estar associada cenários não desejáveis. Muitas vezes estas falhas de mercado encadeam uma expansão da mancha urbana, com implicações negativas no meio ambiente, além da perda dos benefícios da aglomeração. Um exemplo de falha de mercado está relacionada ao tráfego de carros. Este meio individual diminui os custos de transporte para longas distâncias, possibilitando que os cidadãos habitem residências mais distantes de seus

destinos. Todavia, além de causar uma expansão da mancha urbana, introduzir risco de acidentes e emitir poluentes, cada carro a mais que está na rua torna a viagem de todas as outras pessoas um pouco mais lenta. Apesar desses malefícios serem negligenciáveis nas escolhas individuais de utilizar ou não o carro, ou seja, um indivíduo racional não vai deixar de usar o carro, em uma cidade com milhões de habitantes, o custo individual negligenciado de cada carro se soma, de forma a ter um impacto significativo para a sociedade.

Nesse sentido, o governo poderia intervir, de forma a corrigir estas falhas de mercado, por exemplo, introduzindo um pedágio urbano. Dessa forma, seriam internalizadas as externalidades negativas dessa escolha individual de usar o carro. Com o dinheiro coletado desse imposto, poderiam ser compensados os danos causados, através de investimentos públicos. Entretanto, tal medida é difícil de ser implementada, levando os tomadores de decisão a tomar outro caminho no combate às externalidades. Regular os padrões construtivos, pode ser uma forma alternativa de lidar com o problema, impedindo que as externalidades aconteçam. Nessa frente, o CA é o instrumento mais bem estabelecido na literatura do urbanismo para regular o padrão construtivo em cidades modernas nos Estados Unidos, Canadá, Alemanha, França, etc.

Caso o CA máximo regulado seja maior do que o mercado demanda, o CA observado será inferior ao limite, ou seja, a regulação não impacta essa região diretamente. Nas regiões em que a demanda por habitação é maior do que o permitido, parte da demanda precisam ser remanejada territorialmente, para que o limite não seja ultrapassado. A ideia do PDE é limitar o adensamento nas regiões periféricas, direcionando o adensamento para as regiões centrais. Entretanto, caso o excesso de demanda esteja em uma área central, haverá um espraiamento para as periferias, causando o efeito contrário do estipulado. Na Figura 2 é possível observar um exemplo de CA máximo na cidade como um todo e a demanda por habitação excedente no centro, que com o limite máximo no CA, se dispersa em territórios mais distantes.

Figura 2 – Impacto da regulação no CA da cidade



Portanto, caso o planejador central identifique incorretamente a demanda de mercado em cada região da cidade, é possível que a demanda reprimida encadeie um processo de expansão da mancha urbana e todas as externalidades negativas discutidas anteriormente. Nesse sentido o CA, apesar de ser uma ferramenta poderosa no arsenal do planejador urbano, ela também é perigosa, visto que pode levar a efeitos adversos.

## 2.2 Evidências do PDE de 2014

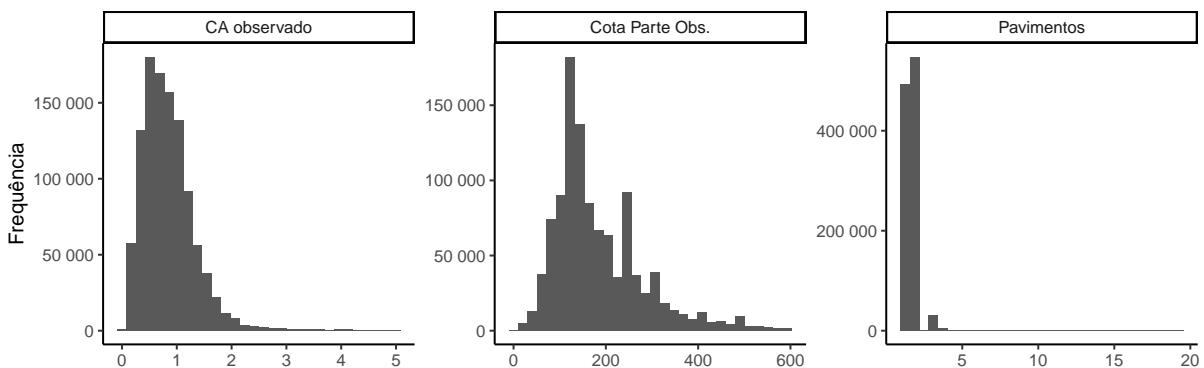
# 3 Dados

## 3.1 Padrões Construtivos

Em relação aos dados sobre os empreendimentos imobiliários, a base de dados escolhida foi do IPTU. Para os fins deste artigo, ela é a mais completa, visto que não representa um fluxo de novos imóveis construídos todos os anos como a base da Embraesp, mas apresenta um estoque imobiliário de tudo que já foi construído de moradia formal na cidade. Dito isso, estão cadastrados os 3.096.719 números únicos de contribuintes, que, segundo a definição da documentação dos dados no GeoSampa, “A cada imóvel urbano corresponderá um número de inscrição no Cadastro Imobiliário Fiscal, entendendo-se como imóvel: I - a área de terreno, construído ou não, definida em matrícula do competente Serviço de Registro de Imóveis ou em transcrições ainda vigente”. Os únicos dados que não constam nessa base são relativos a lotes irregulares ou não registrados.

Entre os dados disponíveis do IPTU, se destacam a área do terreno, a área construída, a área ocupada, o tipo de uso e o número de pavimentos do empreendimento. Logo, com estes dados é possível identificar os padrões construtivos na cidade, mantendo os cálculos os mais fiés possível à forma como são calculados pela regulação. A densidade construída, por exemplo, é definida pelo CA observado na construção – não é considerando nesta análise o CA regulado da região.

Figura 3 – Distribuição do padrão construtivo em cada lote residencial de São Paulo



De fato, ao observar os indicadores na Figura 3, é possível notar que o perfil habitacional de São Paulo é bastante horizontal. É surpreendente que 95% dos lotes residenciais em São Paulo apresentam 2 ou menos pavimentos. Ainda, a mediana do CA observado é 0,8, o que indica que em mais da metade dos lotes da cidade foi construído menos do que a área do terreno. Ainda, a cota parte observada mediana dos lotes da cidade é de  $155m^2$ , o que representa um

uso do terreno para poucas unidades habitacionais em média – a cota parte observada mediana é aproximadamente 8 vezes maior do que a cota parte máxima nos EETUs. Na Figura 14 do Apêndice D é possível observar a evolução dos padrões construtivos no tempo e na Figura 15 está a distribuição da área construída dedicada a cada tipo de uso e perfil de densidade.

## 3.2 Regulação

Os dados da regulação são disponibilizados também de maneira pública no site da prefeitura junto ao PDE, mas pode também ser acessado pelo GeoSampa. O principal dado utilizado foi das áreas de influência dos ativadores de eixo, para que cada lote pudesse ser classificado como dentro ou fora da região de EETU.

## 3.3 Densidade populacional

Os dados de densidade populacional, são produzidos pelo Censo demográfico do IBGE, que ocorre a cada 10 anos. O censo de 2020 se atrasou para 2022 por conta da pandemia e os dados começaram a ser divulgados em 2024. Até o momento, o que foi disponibilizado de maneira pública é a malha preliminar do censo. Considerando que o PDE entrou em vigor em 2014, foi utilizado o censo de 2010 como um período pré PDE e o censo de 2022 como período pós. Segundo o levantamento de 2022, a atual população do município de SP se encontra em 11.451.999 de habitantes, divididos em 4.996.529 de domicílios, dos quais apenas 4.316.336 estão ocupados. Na Figura 16 do Apêndice D é possível observar quais são as áreas mais densas da cidade.

A unidade de observação do censo mais próxima de micrdados que pode ser georreferenciada é o setor censitário<sup>1</sup>. Os limites dos setores são construídos de forma que o agente de coleta do censo consiga cobrir o setor inteiro, minimizando os erros de medição e otimizando o processo. A metodologia para a delimitação dos setores censitários leva em consideração diversos fatores, como elementos na paisagem que se constituam em barreiras naturais ou artificiais e dificultam o trabalho do agente, pontos de referência estáveis e de fácil identificação no terreno, limites das estruturas territoriais como bairros e quadras, entre outros (IBGE, 2024). Um fator determinante para o tamanho do setor censitário é o número de domicílios a serem entrevistados, que em áreas urbanizadas deve estar entre 250 e 400 para os mais densos e entre 150 e 250 para os menos densos.

---

<sup>1</sup> Para estar em conformidade com a Lei Geral de Proteção de Dados (LGPD), é importante proteger a anonimidade dos entrevistados, então deve haver algum tipo de agregação geográfica.

### 3.4 Cruzamento dos dados

Parte da dificuldade de discutir os temas urbanos com dados é a complexidade de trabalhar com eles. Como é possível observar na Tabela 1, cada dado apresenta um nível de observação e frequência diferente. Portanto, sempre é necessário que haja alguma agregação territorial ou temporal para que a análise seja possível. Como comentado anteriormente, a metodologia de processamento de dados utilizada neste artigo está disponível em um repositório público.

Tabela 1 – Informações sobre cada base de dados

Dado	Fonte	Frequência	Escala
Censo	Instituto Brasileiro de Geografia e Estatística (IBGE)	Decenal	Setor censitário
IPTU	Cadastro Territorial predial de Conservação e Limpeza mantido pelo Departamento de Cadastro (DECAD) da Secretaria Municipal da Fazenda da Prefeitura de São Paulo	Anual	Contribuinte
Lotes	Cadastro Imobiliário Fiscal da Secretaria da Fazenda	Anual	Lote
Regulação	Secretaria de Urbanismo e Licenciamento (SMUL)	2014	Perímetro de zona

Os dados do IPTU não são georreferenciados, então desacompanhados de outros dados não é possível cruzá-los com os do Censo. O que possibilita fazer essa junção é que o número único de contribuinte do IPTU, também é o código do Setor, Quadra e Lote (SQL) que o empreendimento se encontra. Nesse sentido, é possível decompor o SQL e cruzar com as bases de lotes, que contêm a geometria de cada um dos 1.677.980 lotes da cidade. O que explica a diferença entre o número de lotes e número de números dos contribuintes são os lotes que contém um condomínio, que pode haver diversos números únicos de contribuintes em apenas um lote. Nos dados de IPTU, há 1.314.353 contribuintes em lotes com apenas uma unidade habitacional e 33.129 condomínios, que contêm 1.782.366 unidades. Ao cruzar o SQL do IPTU com a base de lotes, 44.319 contribuintes não encontram um par na outra base, o que representa uma perda de 1,67% das unidades.

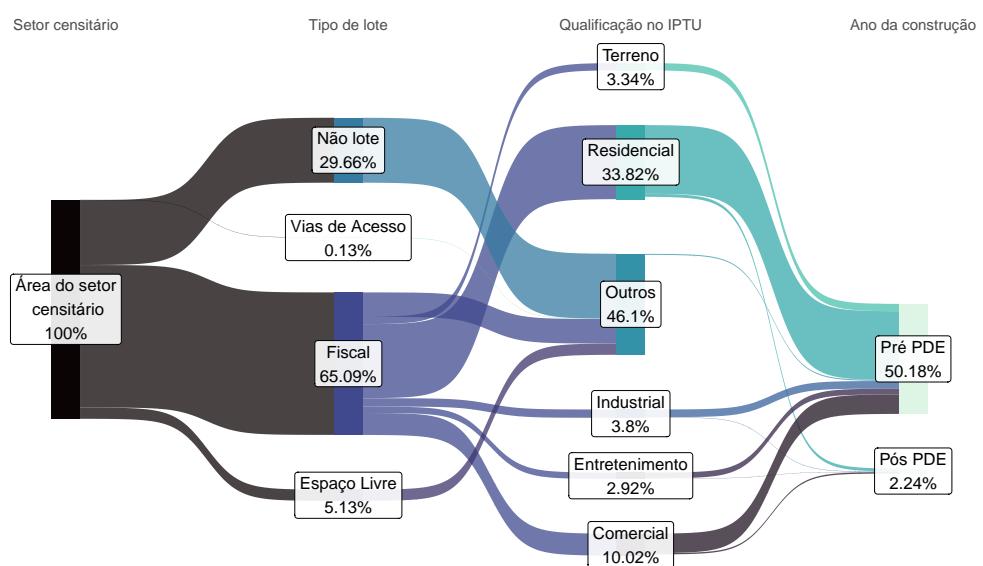
Os dados da regulação não são codificados com o SQL, mas podem ser cruzados geograficamente. O perímetro de zona é um limite parecido com a quadra, que pode ser recortada em alguns casos. Para determinar quais áreas são de EETUs, os perímetros de zona foram unidos, de forma a considerar todo o espaço interior e um contorno das regiões de EETU foi criada. O mapa com o resultado pode ser visto no Apêndice D, Figura 17.

Como o perímetro de zona é bastante aderente à estrutura das quadras, a classificação dos lotes em estar dentro ou fora da região de eixo é direta. Por outro lado, os setores censitários seguem uma estrutura completamente diferente. Há vezes que um setor censitário contém diversos lotes, mas é possível que um lote contenha também mais de um setor censitário. Para fins práticos, os dados do IPTU, censo e regulação foram cruzados geograficamente. Uma

discussão com maior detalhamento pode ser consultada no Apêndice C.

Os setores censitários são uma partição<sup>2</sup> do espaço da cidade de São Paulo, o que implica que a área do setor censitário é composta não apenas pelos lotes, mas também por trechos que contém ruas, parques, usos não residenciais, entre outros fatores. Na Figura 4 é possível identificar que apenas um terço da área do setor censitário é dedicado ao uso residencial. Além disso, no raio de 1km dos eixos, aproximadamente 4% dos empreendimentos residenciais foram construídos após 2014<sup>3</sup>, que foi o período em que o PDE entrou em vigor.

Figura 4 – Dedicação do espaço do setor censitário para cada uso em um raio de 1km dos eixos



<sup>2</sup> Uma partição é uma divisão que ocupa o espaço por completo, sem que haja intersecção entre os elementos

<sup>3</sup> Este valor está superestimado, visto que grande parte dos prédios que ficaram prontos entre 2014 a 2017 foram aprovados durante a vigência do PDE de 2002

# 4 Metodologia e Resultados

Como comentado na Seção 1.3, há três perguntas a ser respondidas, que podem ser visualizadas no diagrama da Figura 1. Nas seções a seguir, cada uma das perguntas será respondida através da análise quantitativa dos dados apresentados.

## 4.1 Do padrão construtivo à densidade populacional

O objetivo desta seção é responder: **Os padrões construtivos regulamentados e incentivados são de fato capazes de gerar a densidade esperada?** Para tanto, foram considerados os dados do Censo de 2022 e os do IPTU do mesmo ano. Para esta seção, a regulação de cada região da cidade não é importante, visto que o foco é em identificar qual a relação entre a densidade construtiva, densidade habitacional e pavimentos com a densidade populacional e esta relação não é modificada pelo PDE. A influência do PDE recai sobre estes três componentes do padrão construtivo, através da regulação do CA, cota parte e gabarito, respectivamente. Dessa forma, validar a importância desses padrões construtivos significa validar também os instrumentos de regulação.

Para inferir a relação entre as 3 variáveis e a densidade populacional, foram adotadas duas abordagens complementares. A primeira foi de utilizar uma regressão linear. A principal limitação dessa abordagem é que a forma funcional da regressão não está fundamentada economicamente, o que a torna sujeita a problemas de endogeneidade, e, portanto, enfraquece a validade da inferência. A abordagem alternativa, de árvores de regressão, transita do universo da econometria para o universo do *machine learning*. Apesar de ter uma interpretabilidade mais restrita, as árvores de regressão não precisam de uma forma funcional como a regressão linear, mas conseguem encontrar relações não lineares entre as variáveis do modelo.

O primeiro passo para implementar estas metodologias consiste em combinar as bases de dados em um nível de observação único. Para manter os dados na escala mais desagregada possível, os dados poderiam ser agregados ao nível do setor censitário. Porém, como foi discutido na Seção 3.3, o tamanho do setor censitário depende diretamente tanto da população que o habita, quanto de sua densidade, o que traria viés para a regressão<sup>1</sup>. Desse modo, o mapa da cidade foi recortado em um *raster* (quadriculado) com células de 800m de largura e os dados tanto do censo quanto do IPTU foram agregados nessa escala. Os detalhes mais técnicos podem ser encontrados no Apêndice C.

---

<sup>1</sup> Neste caso, a densidade populacional depende da área do setor censitário, que não pode ser omitida da regressão. Entretanto, se for utilizada como um regressor, apresentará uma relação de simultaneidade com a variável dependente, já que quanto mais densa uma região é, menor será sua área. Sendo assim, seria necessário fazer uma estimativa por equações simultâneas, trazendo complexidade para a identificação, já que não há variáveis exógenas suficientes para recuperar os parâmetros de interesse

Por fim, foram selecionadas apenas as células do raster que apresentavam um baixo grau de informalidade. O espectro de irregularidade foi calculado comparando o número de domicílios registrados no Censo com o total de unidades registradas no IPTU. Quando uma região se encontra no extremo esquerdo do espectro (próximo de 0%), isso indica que, embora haja domicílios registrados no Censo, a área não possui moradias formais. Valores próximos a 50% refletem equilíbrio entre os dados do Censo e do IPTU, indicando ausência de informalidade. Valores acima de 50%, que representam um número maior de unidades no IPTU do que no Censo, são raros. Para mais detalhes sobre a informalidade, consulte o Apêndice Apêndice B.

#### 4.1.1 Abordagem de regressão linear

Tabela 2 – Resultados da regressão, para espectro de irregularidade entre 45 e 55%

	y = População		y = Densidade	
	(A) Nível	(B) Log	(C) Nível	(D) Log
Intercepto	2121.210***	7.814***	6240.590***	8.583***
Unidades	1.476***	0.000***		
Área construída (m <sup>2</sup> )	0.000	0.000+		
Verticalização	-131.656***	-0.036**	24.817	-0.033
Densidade construtiva			705.142	0.199**
Densidade habitacional			0.187***	0.000*
Num.Obs.	332	332	332	332
R <sup>2</sup>	0.861	0.602	0.319	0.223
R <sup>2</sup> Adj.	0.860	0.598	0.313	0.216

+ p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Na Tabela 2 é possível observar os resultados de quatro regressões, todas com o mesmo intuito, porém com formas funcionais diferentes. Nas primeiras duas colunas, a população é regredida em unidades, metros quadrados de área construída e um indicador de verticalização. Interpretando a primeira coluna (A), é possível inferir que caso se mantenham constantes os outros componentes, uma unidade habitacional a mais está associada a 1,47 moradores a mais na região. O número de metros quadrados construídos não apresenta significância estatística. O coeficiente da verticalização indica que um aumento de 1 na verticalização<sup>2</sup> em média está associado a 130 pessoas a menos na célula do *raster*. Para ter um parâmetro de comparação, a média da população por célula é de 4.600 habitantes, com uma dispersão considerável. A distribuição da população dos *rasters* pode ser vista na Figura 18 do Apêndice D.

<sup>2</sup> Um aumento de 1 na verticalização implicaria que todas as construções na célula do *raster*, *ceteris paribus* tivessem um pavimento a mais

Na segunda coluna da tabela (B), os resultados são semelhantes, mas a área construída passa a ser estatisticamente significativo e positivo, mas negligenciável. O coeficiente associado à verticalização continua negativo, indicando que é esperado que um aumento de 1 na verticalização esteja associado a 3,6% menos pessoas habitando a região, caso todos os outros componentes continuem constantes. Nas colunas C e D da Tabela 2, apesar das variáveis dependente e explicativas serem referentes a densidade<sup>3</sup> e não valores absolutos, os resultados permanecem parecidos. A principal diferença é que a verticalização ao invés de ter uma relação inversa com a população, passa a ser estatisticamente irrelevante.

Apesar de contraintuitivo, o resultado apresenta uma interpretação razoável. O modelo econométrico (Coluna A) aponta que em cada unidade habitacional há em média 1,47 habitantes, mas caso estas mesmas unidades ganhem metragem de área construída é possível observar mais habitantes nestes domicílios, visto que se tornam mais espaçosos. O resultado associado à verticalização pode indicar que o perfil dos núcleos domiciliares que habitam imóveis mais verticalizados é de famílias menores. Outro resultado surpreendente é o intercepto. De maneira intuitiva, quando há zero pavimentos, nenhuma área construída e nenhuma unidade habitacional, a população deveria ser zero também, mas o modelo aponta para 2.121 pessoas. A explicação para isso são as moradias informais. Quando a regressão é feita para vários recortes do espectro de irregularidade, na medida em que o recorte se aproxima de selecionar apenas células equilibradas (espectro de irregularidade igual a 50%), o intercepto se torna estatisticamente igual a zero. Em outras palavras, quando há irregularidade, o número de unidades é subestimado e o efeito é capturado pelo intercepto. A Figura 19 apresenta uma análise de robustez com todos os coeficientes e seus respectivos intervalos de confiança para cada intervalo de espectro de irregularidade selecionado.

#### 4.1.2 Abordagem de *machine learning*

Na abordagem de *macinhe learning*, foi construído um modelo de *random forest* (WRIGHT; ZIEGLER, 2015). Uma das principais vantagens dessa técnica é que não exige a especificação de uma forma funcional, pois o próprio modelo seleciona as variáveis mais importantes e ajusta uma regressão não linear. Outro benefício da *random forest* é sua capacidade de avaliar a importância relativa das variáveis, algo que não é diretamente testável em uma regressão linear.

Para desenvolver o modelo de floresta de regressão, a base de dados foi dividida em conjuntos de treino e teste, prevenindo overfitting, ou seja, garantindo que o modelo aprenda padrões em vez de simplesmente memorizar os dados. Foram geradas 100 mil árvores com base nas células que apresentam um espectro de irregularidade entre 40% e 60%. Para avaliar a performance do modelo, foi replicada a regressão da Tabela 2 (Coluna C) com o método *random*

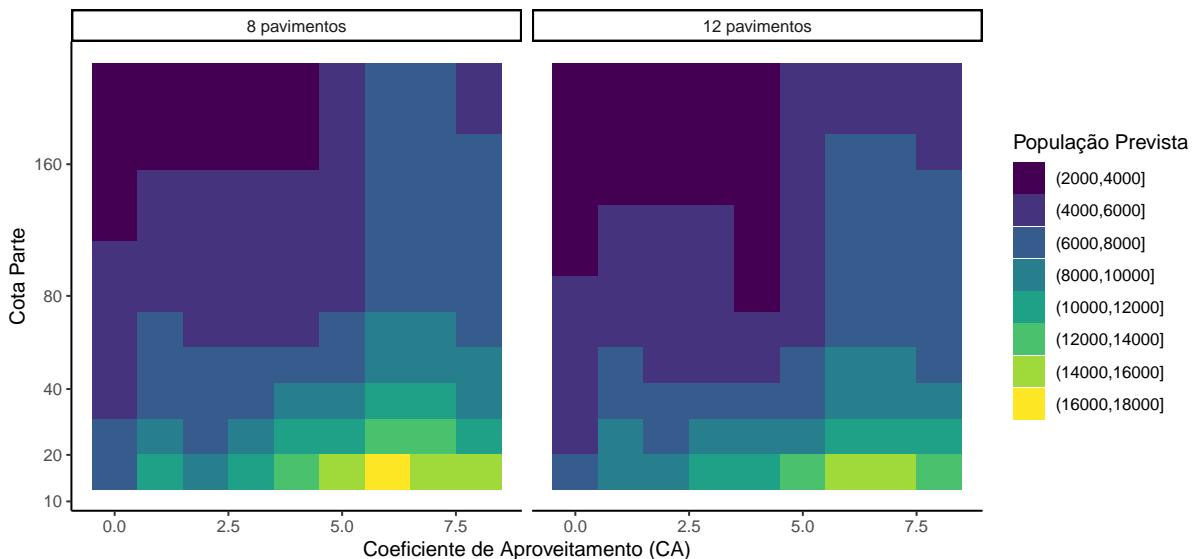
<sup>3</sup> A densidade construtiva é calculada igual o CA (área construída / área do terreno) e a densidade habitacional é o inverso da cota parte (número de unidades / área do terreno)

*forest*. Observou-se um  $R^2$  de 81,9% na base de teste, indicando um ganho significativo em poder explicativo. Esse ganho decorre exclusivamente da capacidade do modelo de capturar relações não lineares entre as variáveis, visto que os dados de entrada são os mesmos da regressão.

Para identificar quais variáveis são mais relevantes no modelo, foi aplicado um procedimento de permutação, no qual se introduz ruído nas variáveis explicativas para avaliar o impacto disso no erro do modelo (BREIMAN, 2001; NEMBRINI; KÖNIG; WRIGHT, 2018). Esse método permite quantificar a importância de cada variável para explicar a variável dependente. Como resultado, a *random forest* identificou a densidade habitacional como a mais importante, seguida pela densidade construtiva e por fim a verticalização. Estes resultados são condizentes com os achados apresentados na Tabela 2.

Com base na *random forest* construída, também é possível simular a população de uma determinada região, com base em seus padrões construtivos. Para fins ilustrativos, foi considerada uma célula do raster, que possui largura de 800m. Nessa célula hipotética, a depender do CA, cota parte e verticalização observados, pode-se estimar a população que a habitaria. Na Figura 5 é possível analisar os resultados. Em linha com que foi distorcido até então, uma cota parte baixa parece ser o fator mais importante para obter uma densidade alta, principalmente se aliada à um CA alto. Um número maior de pavimentos, como já comentado, na maioria das situações está associado à uma densidade mais baixa ou igual, caso a densidade construtiva e habitacional se mantenham as mesmas.

Figura 5 – Simulações da *random forest* para a população de uma célula hipotética de 800x800m



#### 4.1.3 Resultados principais

Apesar das diferenças metodológicas, ambas as abordagens convergem para um resultado comum: a densidade habitacional é o componente do padrão construtivo mais importante para

definir a densidade populacional. Esse achado é intuitivo, pois a quantidade de domicílios desempenha um papel central na determinação do número de habitantes. A densidade construtiva e o número de pavimentos, dados uma densidade habitacional e pavimentos constantes, podem ser vistos como modificadores do perfil desses domicílios.

Na medida que a densidade construtiva aumenta, as unidades habitacionais tendem a ser mais espacosas, o que acomoda mais moradores em média. No caso da verticalização, com a densidade construtiva e habitacional fixas, qualquer aumento nos pavimentos exige uma redução na taxa de ocupação do terreno (conforme discutido na Seção 4.2 e no Apêndice A), resultando em edifícios com maiores recuos. Os dados sugerem que o mesmo número de unidades habitacionais e de área construída, quando inseridos em empreendimentos mais verticais, apresentam menos moradores em média.

À luz dos resultados apresentados, a discussão sobre os instrumentos instituídos pelo PDE tomam um novo rumo. Os dados mostram que a verticalização, que foi incentivada nas zonas de eixos, não é uma aliada do adensamento, por vezes inclusive trabalhando no sentido contrário. O efeito da densidade construtiva foi em grande parte obstruído pelo potencial que a densidade habitacional tem de gerar densidade. Nesse sentido, se o PDE almeja definir quais regiões da cidade devem apresentar maior ou menor densidade, seria mais efetivo focar na cota parte e permitir um CA e gabarito suficientes para a construção das unidades habitacionais<sup>4</sup>.

## 4.2 Da regulação ao padrão construtivo

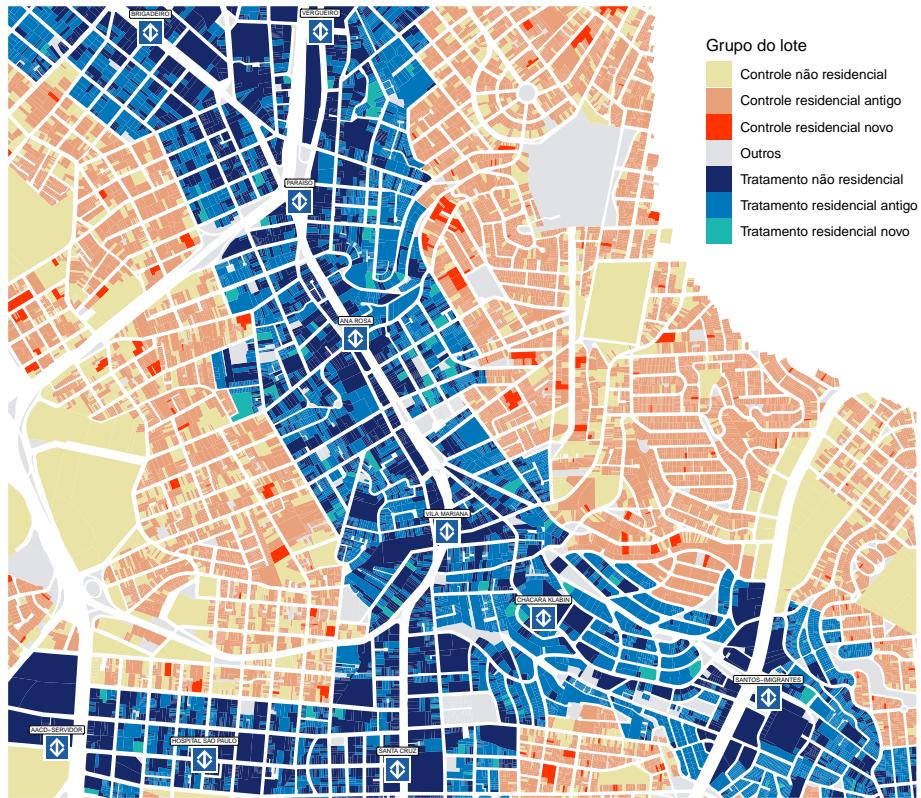
O objetivo desta seção é responder: **A regulação foi capaz de influenciar os padrões construtivos na cidade da maneira esperada?** Para tanto, foi escolhida também duas abordagens. Ambas são, em sua essência, metodologias de regressão em discontinuidade (RDD<sup>5</sup>). Essa escolha é ideal por conta da exogeneidade na qual foi definido quais lotes pertenceriam ou não às áreas de eixo perto de sua fronteira. Como comentado na Seção 1.2, para se tornar uma zona de EETU, o lote deve estar dentro da área de influência do transporte coletivo de alta capacidade. Nesse sentido, os lotes vizinhos aos classificados como dentro da zona de eixo são um controle perfeito, visto que a diferença entre um lote de tratamento e o outro de controle é apenas atravessar a rua. Outras regiões da cidade também tiveram suas regulações alteradas, mas o foco da identificação

Na Figura 6 é possível ver um exemplo de classificação dos lotes como controle e tratamento na Vila Mariana, região de eixos ativados por estações de metrô, em uma região de união de duas linhas, a verde e a azul. Um empreendimento é considerado novo caso ele tenha

<sup>4</sup> Uma cota parte de 20 significa que, com o CA básico, cada unidade habitacional apresentará 20 metros quadrados. Estas mesmas unidades com um CA de 4 podem apresentar em média  $80m^2$ . Todavia, é impossível atingir um CA 4 com menos de 4 pavimentos. Inclusive, quando há recuos obrigatórios, é necessário um número de pavimentos maior do que o CA para que seja possível construir todos os metros quadrados de área

<sup>5</sup> Do inglês, *Regression Discontinuity Design*

Figura 6 – Exemplo de separação dos grupos de controle e tratamento na Vila Mariana



sido terminado de construir depois de 2014, ano de início da vigência do PDE.

#### 4.2.1 Abordagem *diff-in-diff* geográfico (e *event study*)

O *diff-in-diff* (*DiD*) geográfico é uma categoria de regressão de diferenças em diferenças que leva em consideração a proximidade espacial das observações para a definição dos critérios de controle e tratamento. Neste caso, foram definidos como tratamento todos os lotes residenciais que estão dentro da distância de 50m entre o centroide e a borda do eixo e como controle, a mesma distância, mas para os lotes que estão fora da zona de EETUs. A hipótese de identificação é igual a de um DiD comum: os grupos dois grupos, na ausência do tratamento, precisam apresentar tendências paralelas.

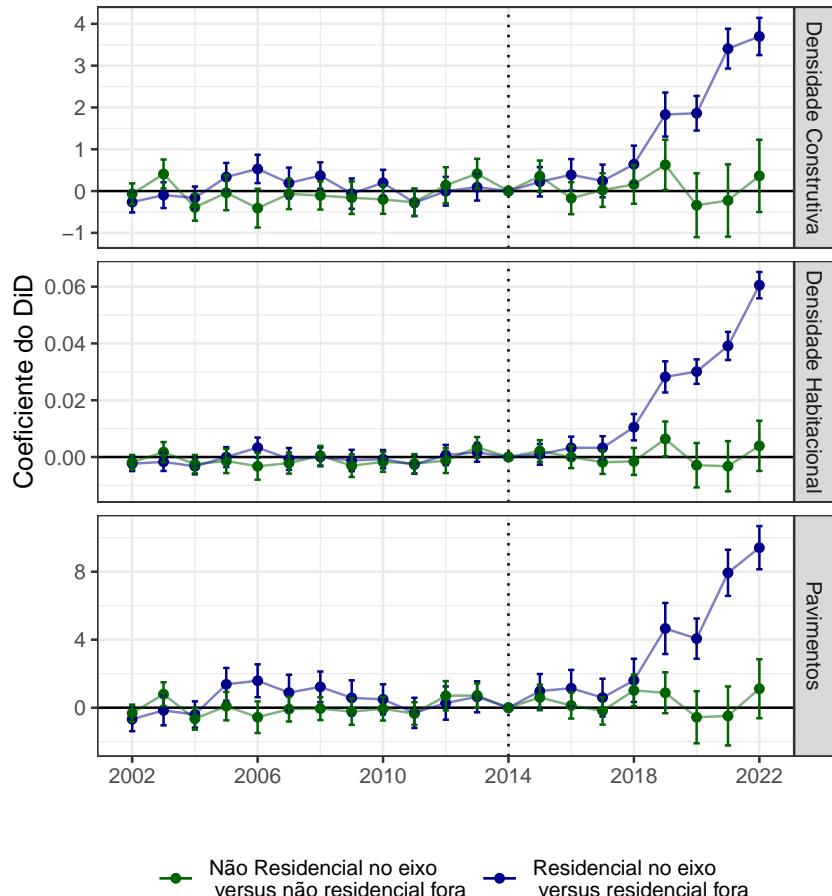
A Tabela ?? traz os resultados do DiD escolhendo cada componente do padrão construtivo como variável de interesse. Os dados apontam que a introdução dos eixos é responsável por aumentar em média a densidade construtiva em 0,9, a densidade habitacional em 0,011 domicílios por metro quadrado de terreno e 2,4 pavimentos. O eixo não altera a probabilidade do lote ser ou não residencial, como mostra a quarta coluna da regressão. Além disso, para este recorte de 50m, não há diferença estatisticamente significante entre os grupos de controle e tratamento antes do PDE. Os resultados são robustos às diferentes escolhas para o recorte de proximidade à borda. Na Figura 20 do Apêndice D é possível verificar o coeficiente associado ao efeito do tratamento

Tabela 3 – Resultados do DiD para unidades dentro de 50m da fronteira de eixo

	densidade_construtiva	densidade_habitacional	pavimentos	residencial
Intercepto	1.031***	0.008***	1.992***	0.780***
Tratamento	-0.001	0.000	0.027	-0.011***
Pós PDE	0.891***	0.009***	1.661***	-0.100***
Tratamento x Pós PDE	0.903***	0.011***	2.392***	0.001
Num.Obs.	59904	59904	59904	77625
R2	0.027	0.032	0.026	0.001
R2 Adj.	0.027	0.032	0.026	0.001

+ p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

para cada escolha de distância da borda.

Figura 7 – *Event study* para os padrões construtivos na fronteira dos eixos

Para fortalecer a análise e identificar heterogeneidade temporal nos resultados, foi conduzido também um estudo de evento. Os resultados estão apresentados na Figura 7. O *Event study* também aponta para um aumento da densidade construtiva, densidade habitacional e pavimentos depois do PDE, mas o efeito parece se intensificar após 2018. Isso pode ser explicado

pelo fato de que os empreendimentos entregues entre 2014 e 2017 em grande parte ainda seguiam a regulação do PDE de 2002. Além disso, foram observadas tendências paralelas no período pré PDE, indicando que de fato os grupos são comparáveis e o critério escolhido para separação dos grupos é exógeno. Por fim, um resultado surpreendente é de que os eixos parecem ter afetado apenas os empreendimentos residenciais.

#### 4.2.2 Abordagem de RDD

Considerando o sucesso da metodologia de assumir como parecidos os grupos que estão próximos da fronteira dos eixos, a estratégia do RDD faz bastante sentido<sup>6</sup>. A hipótese de identificação do RDD consiste em considerar que no entorno do ponto de corte, as observações imediatamente acima e abaixo do limiar são similares em todas as características observáveis e não observáveis, exceto pela variável de tratamento. Como discutido anteriormente, essa exogeneidade existe, visto que a diferença entre um lote de tratamento e o outro de controle é apenas atravessar a rua.

Entretanto, quando é implementado o RDD, não se observa efeito do tratamento – este resultado pode ser visto na Figura 8. O motivo pelo qual isso acontece, é que como são poucos os lotes que receberam edificação pós PDE (consultar Figura 4), mesmo que este tenha um efeito significativo, ele é imperceptível quando se considera todos os lotes. O resultado torna-se significativo quando se seleciona apenas os lotes que receberam edificação após 2014. Neste caso, é observado um aumento na densidade construtiva, densidade habitacional e pavimentos (Figura 9).

Na Tabela 4 é possível observar os resultados do RDD filtrado apenas para lotes com edificações novas e utilizando *kernel* triangular com largura de banda escolhida automaticamente. Os resultados são semelhantes aos encontrados via diferenças em diferenças e os apresentados na Figura 9, com exceção da probabilidade do lote ser residencial. Como teste placebo, foi replicado o mesmo procedimento, mas considerando o *cutoff* em -100 ao invés de zero. A largura de banda selecionada automaticamente foi menor que 100, o que implica que nesse procedimento houve uma comparação do grupo controle com o próprio grupo controle. O teste placebo não apresentou efeito para os três componentes do padrão construtivo, validando o teste e trazendo mais evidências para a exogeneidade do tratamento na fronteira. Todavia, o coeficiente associado ao lote ser residencial apresentou significância, invalidando a inferência sobre esta variável da Tabela 4.

#### 4.2.3 Resultados principais

Os dados mostram que construções nos eixos após o início da vigência do PDE apresentam maior densidade construtiva, densidade habitacional e pavimentos. As evidências apontam

---

<sup>6</sup> Inclusive, o RDD se torna igual à estimação por diferença, quando o grau do polinômio é igual a zero e a *bandwidth* é a mesma.

Figura 8 – RDD para os padrões construtivos no entorno dos eixos

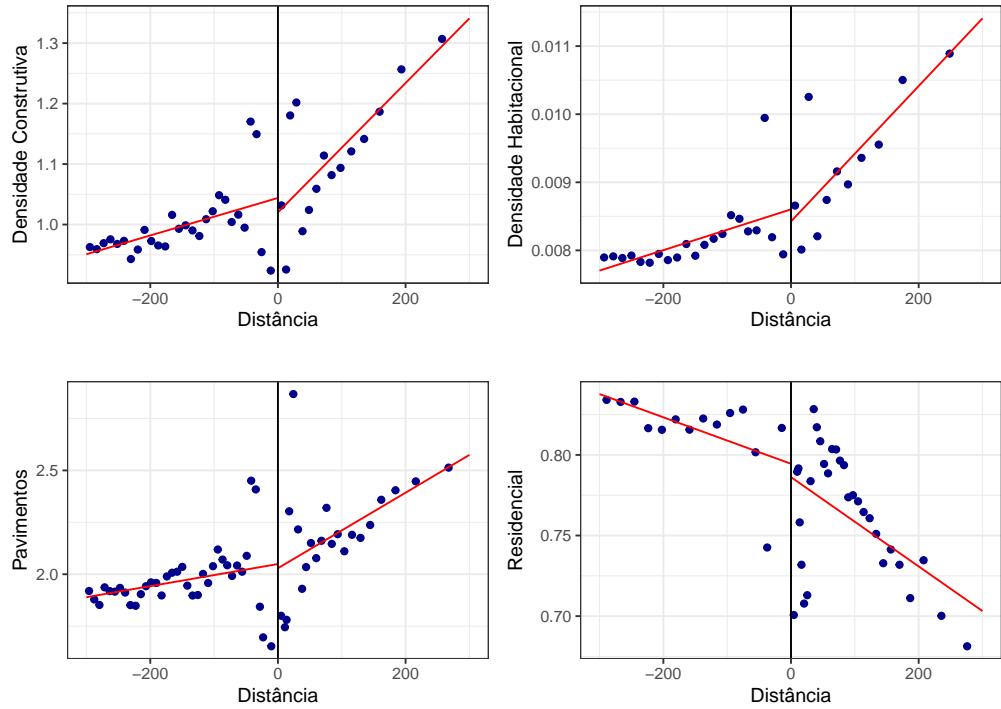
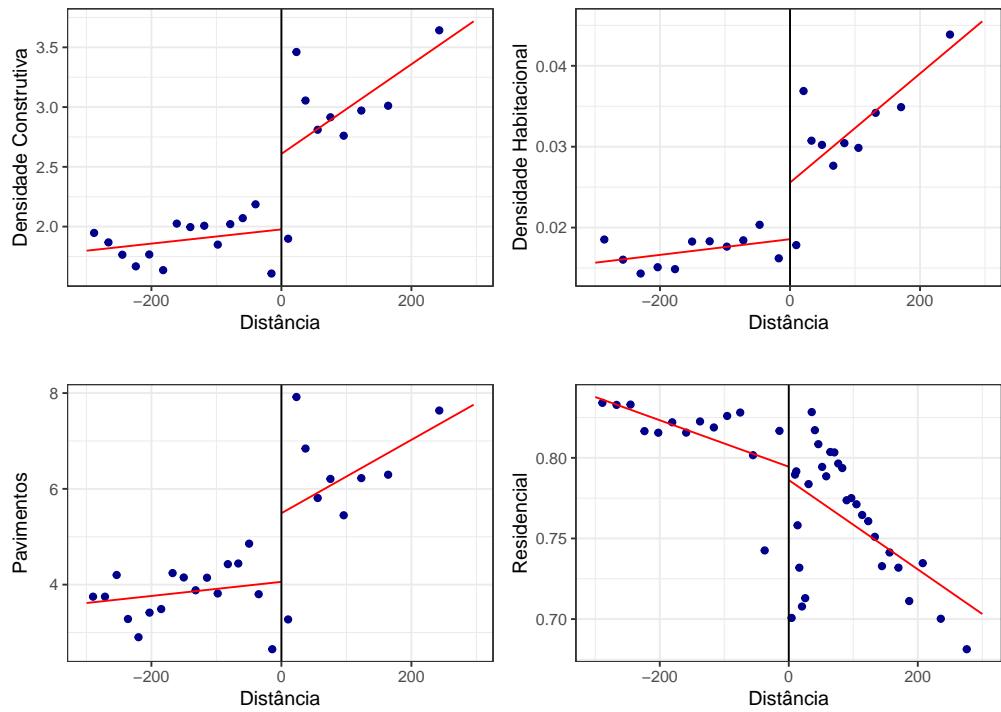


Figura 9 – RDD para os padrões construtivos no entorno dos eixos, filtrando por lotes com edificações recentes



também para uma intensificação do efeito do PDE começando a partir de 2018 e alcançando seu ápice nos dados mais recentes, em 2022. Além disso, as mudanças do PDE parecem ter afetado apenas os empreendimentos residenciais, enquanto nos outros usos não se observa uma diferença

Tabela 4 – Resultados do RDD apenas com lotes edificados pós PDE

	Densidade Construtiva	Densidade Habitacional	Pavimentos	Residencial
Conventional	0.707**	0.010***	1.955**	-0.261***
Bias-Corrected	0.638**	0.010**	1.821**	-0.290***
Robust	0.638*	0.010*	1.821*	-0.290***
Observações.à.esquerda	622	799	624	884
Observações.à.direita	684	841	689	1002
Largura.de.Banda	61.2	76.3	61.7	59.5

+ p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Tabela 5 – Teste placebo do RDD com o cutoff deslocado 100 metros para fora do eixo

	Densidade Construtiva	Densidade Habitacional	Pavimentos	Residencial
Conventional	0.031	0.000	0.211	0.128**
Bias-Corrected	0.096	0.001	0.407	0.147***
Robust	0.096	0.001	0.407	0.147**
Observações.à.esquerda	863	921	890	1022
Observações.à.direita	990	1070	1027	1202
Largura.de.Banda	76.9	82.7	79	68.1

+ p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

significativa entre a região de eixo e fora dos eixos.

Entretanto, um resultado que se destaca também é que a mudança causada pelo PDE nos últimos 10 anos não apresentou intensidade e/ou velocidade suficientes para se observar uma divisão territorial clara de eixo e não eixo nos padrões construtivos. Para observar os resultados, é necessário olhar exclusivamente para os novos empreendimentos, que, como representado na Figura 4, representam menos de 5% dos empreendimentos.

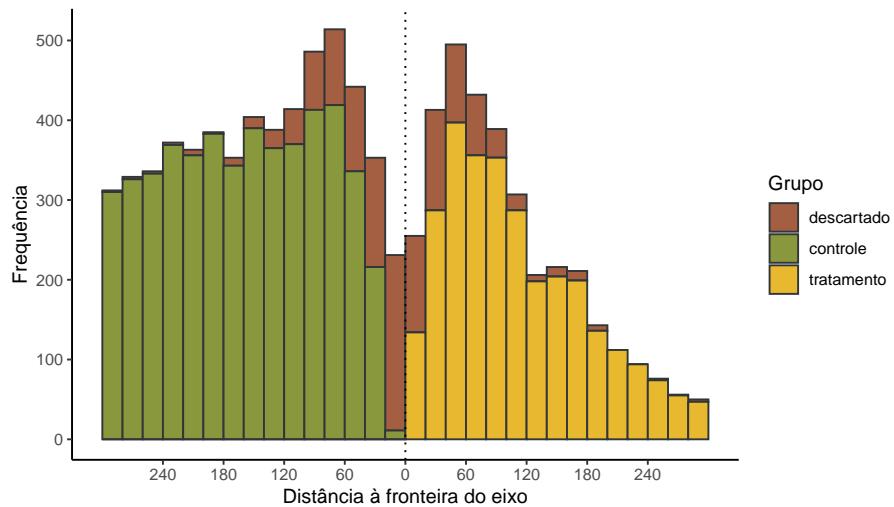
### 4.3 Da regulação à densidade populacional

O objetivo desta seção é responder: **A regulação foi capaz de influenciar a densidade populacional na cidade?** Como comentado na Seção 4.1, a densidade habitacional é o componente mais importante para determinar a densidade populacional, e segundo os resultados apresentados na Seção 4.2 os eixos causaram um aumento da densidade habitacional. Dessa forma, a hipótese que mais faz sentido é que a densidade populacional nos eixos aumentou. Para testar esta hipótese foi realizado um procedimento parecido com o apresentado na Seção 4.2, mas trocando a variável resposta por densidade populacional.

A principal diferença é que, no caso da densidade, a unidade de observação (setor censitário) é geograficamente muito maior do que os lotes. Assim, é comum que um setor censitário próximo à fronteira ocupe tanto áreas do grupo de controle quanto do grupo de tratamento, o que exige sua exclusão da análise. No entanto, a ocorrência desse fenômeno é tão frequente que, nas proximidades imediatas da fronteira, há um número insuficiente de unidades para realizar a análise usando RDD. Esse fenômeno é ilustrado na Figura 10.

Além disso, há um viés de sobrevivência, no qual apenas setores censitários pequenos, com menor probabilidade de estarem em ambos os lados da fronteira, permanecem próximos à borda após aplicar o critério de exclusão<sup>7</sup>. O problema com esses setores é que, como discutido na Seção 3.3, quanto menor o setor censitário, maior tende a ser sua densidade, o que faz com que a densidade média na fronteira aumente de forma atípica tanto no grupo de controle quanto no de tratamento. Esses fatores inviabilizam a análise via RDD.

Figura 10 – Número de observações por grupo em função da distância à fronteira do eixo

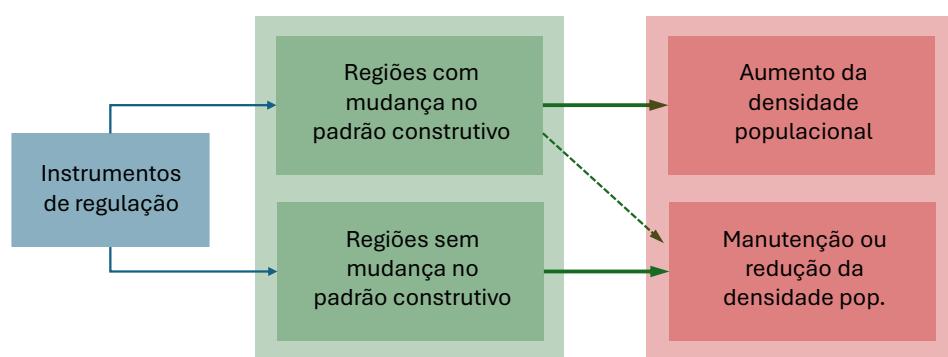


Porém, antes de fazer as estimativas, é importante definir com muita clareza o que é o tratamento. Como avaliado na Seção 4.2, mesmo que os padrões construtivos dos novos empreendimentos tenham mudado, apenas um número seletivo de lotes foram recentemente edificados. Nesse sentido, em grande parte das regiões de tratamento não será observado o efeito do tratamento

Nesse sentido, se o tratamento for o recebimento uma nova regulação, os setores censitários que não contém lotes com edificações recentes apresentarão, em teoria, efeito zero. Como estes setores censitários são a grande maioria, é possível que o efeito de receber prédios novos com a regulação nova seja ofuscado – esta relação pode ser observada na Figura 11. Portanto, é necessário identificar efeitos heterogêneos do tratamento, com relação à quão recentes são as edificações no setor censitário.

<sup>7</sup> No caso do setor censitário, é considerado de tratamento quando seu centroide está na região de eixo e mais de 95% de sua área está nessa área. O controle é definido pela presença do centroide fora do eixo e menos de 5% de área de eixo. Qualquer setor censitário que não cumpre com algum dos critérios é excluído

Figura 11 – Identificação do efeito do tratamento (Figura 1 modificada)



### 4.3.1 Abordagem via diferenças em diferenças com efeito heterogêneo

Apesar da impossibilidade de conduzir a análise via RDD, o sucesso da abordagem por *diff-in-diff* nas seções anteriores é um forte indicativo de que para a densidade populacional essa abordagem também é adequada. Na Equação 4.1 está representada a fórmula da regressão, na qual  $Y$  representa a densidade populacional,  $P$  o período pós tratamento,  $T$  o grupo de tratamento e  $Q$  a proporção da área do setor censitário que recebeu novas edificações durante o período de análise<sup>8</sup>.

No modelo, o  $\beta_1$  representa a diferença de densidade entre 2022 e 2010. O  $\beta_2$  captura a diferença de densidade entre o grupo de tratamento e de controle na ausência do tratamento. O  $\beta_3$  é o efeito que novas edificações apresentam na densidade. A primeira interação, cujo coeficiente é o  $\beta_4$ , captura o efeito do tratamento caso nenhuma nova edificação seja construída<sup>9</sup>, como demonstrado na Equação 4.3. O  $\beta_5$  qualifica as diferença entre o efeito de qualificações em 2022 e 2010 na densidade populacional. Analogamente, o  $\beta_6$  faz o mesmo, mas entre o grupo controle e tratamento na ausência do tratamento. Por fim, o  $\beta_7$  é o parâmetro de interesse, que captura o efeito do tratamento, a depender da proporção da área do setor censitário que recebeu novas edificações, como demonstrado na Equação 4.5.

$$\mathbb{E}(Y|P, T, Q) = \beta_0 + \beta_1 P + \beta_2 T + \beta_3 Q + \beta_4 PT + \beta_5 PQ + \beta_6 TQ + \beta_7 PQT \quad (4.1)$$

$$\frac{\partial E(Y|T, P, Q = 0)}{\partial T} = \beta_2 + \beta_4 P \quad (4.2)$$

$$\mathbb{E}(Y|T = 1, P = 1, Q = 0) - \mathbb{E}(Y|T = 1, P = 0, Q = 0) = \beta_4 \quad (4.3)$$

$$\frac{\partial E(Y|T, P, Q)}{\partial T} = \beta_2 + \beta_4 P + \beta_6 Q + \beta_7 PQ \quad (4.4)$$

$$[\mathbb{E}(Y|T = 1, P = 1, Q = q) - \mathbb{E}(Y|T = 1, P = 0, Q = q)] = \beta_4 + \beta_7 q \quad \frac{\partial E(Y|T, P, Q)}{\partial Q} = \beta_3 + \beta_5 P + \beta_7 T \quad (4.5)$$

$$\frac{\partial E(Y|T, P, Q)}{\partial Q} = \beta_3 + \beta_5 P + \beta_6 T + \beta_7 TP \quad (4.6)$$

Os resultados das estimativas podem ser consultados na Tabela ??

### 4.3.2 Resultados principais

<sup>8</sup> Para os dados do censo de 2022, o período em que as edificações são consideradas novas as é entre 2014 e 2022. Para o censo de 2010, o período foi considerado 2002 até 2010.

<sup>9</sup> Este coeficiente, em teoria, deve ser zero, visto que com nenhuma edificação nova, ninguém pode ser afetado pelo tratamento

Tabela 6 – Resultado diff-in-diff do Censo, para setores censitários em um raio de 100m dos eixos

	(A)	(B)	(B)	(B)
Intercepto	0.057***	0.055***	533.359***	192.072***
T (Grupo de tratamento)	-0.004	-0.006	-22.741+	8.449+
P (Período pós PDE)	-0.016***	-0.015***	-193.710***	-27.108***
Q (Percentual residencial novo)		0.054*	0.029	-2.989
T x P	0.000	0.002	-6.418	-6.807
T x Q		0.016	-84.899	-39.625
P x Q		0.015	-46.080	48.249
T x P x Q		-0.036	187.922	57.080
Num.Obs.	3225	3225	3225	3225
R2	0.013	0.028	0.171	0.032
R2 Adj.	0.013	0.026	0.169	0.030

+ p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Tabela 7 – Estimativas de habitantes de um setor censitário médio para cada cenário, segundo resultados apresentados em ?? (A)

Cenário	Coeficientes	Tratamento	Controle
Em 2010, em regiões que <b>não receberão</b> prédios novos em 2022	$\beta_0 + \beta_2 t$	50.364	50.364
Em 2010, em regiões que <b>receberão</b> prédios novos em 2022	$\beta_0 + \beta_2 t + 0,6\beta_3 + 0,6\beta_6 t$	81.986	50.364
Em 2022, em regiões que <b>não receberam</b> prédios novos	$\beta_0 + \beta_1 + \beta_2 t + \beta_4 t$	35.502	50.364
Em 2022, em regiões que <b>receberam</b> prédios novos em 2022	$\beta_0 + \beta_1 + \beta_2 t + 0,6\beta_3 +$ $\beta_4 t + 0,6\beta_5 + 0,6\beta_6 t +$ $0,6\beta_7 t$	65.769	50.364

## 5 Conclusão

5.1 Resultados principais

5.2 Contribuições para o debate público

# Referências

- BREIMAN, L. Random forests. **Machine learning**, Springer, v. 45, p. 5–32, 2001.
- BRUECKNER, J. K. **Lectures on urban economics**. [S.l.]: MIT press, 2011.
- FUJITA, M. **Urban economic theory**. [S.l.]: Cambridge university press, 1989.
- IBGE. **Censo Demográfico 2022: Malha de Setores Censitários Preliminares**. Rio de Janeiro, Brasil: Instituto Brasileiro de Geografia e Estatística, 2024. Acesso em: 2024-06-23. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv102072.pdf>>.
- IBRAEVA, A. et al. Transit-oriented development: A review of research achievements and challenges. **Transportation Research Part A: Policy and Practice**, Elsevier Ltd, v. 132, p. 110–130, 2 2020. ISSN 09658564.
- Município de São Paulo. **Lei nº 10.235, de 16 de dezembro de 1986**. 1986. Dispõe sobre a forma de apuração do valor venal de imóveis, para efeito de lançamento dos Impostos sobre a Propriedade Predial e Territorial Urbana, concede descontos sobre os valores venais dos imóveis sujeitos à incidência desses impostos, no exercício de 1987, e dá outras providências.
- NEMBRINI, S.; KÖNIG, I. R.; WRIGHT, M. N. The revival of the gini importance? **Bioinformatics**, Oxford University Press, v. 34, p. 3711–3718, 11 2018. ISSN 14602059.
- PAPAGEORGIOU, Y. Y.; PINES, D. **An essay on urban economic theory**. [S.l.]: Springer Science & Business Media, 2012. v. 1.
- WRIGHT, M. N.; ZIEGLER, A. ranger: A fast implementation of random forests for high dimensional data in c++ and r. **arXiv preprint arXiv:1508.04409**, 2015.

# Apêndices

# APÊNDICE A – Discussões sobre a verticalização

Assumindo prédios como paralelepípedos, ou seja, todos os andares apresentam a mesma metragem, é possível traçar relações geométricas nas quais a área ocupada é a superfície que representa a base do paralelepípedo, o número de pavimentos sua altura e a área construída seu volume (Figura 12). Nesse sentido, as relações apresentadas na Equação A.1, que apresenta diferentes formas de calcular a verticalização, ficam mais claras. Na equação, AC representa a área construída; AO, a área ocupada; AT, a área do terreno; CA o coeficiente de aproveitamento e tx\_ocupacao, o percentual da área do terreno que é ocupada. Usando o procedimento apresentado, em regiões que possuem o mesmo CA, quanto menor for a taxa de ocupação do terreno, maior será a verticalização. Analogamente, para dois lotes com a mesma ocupação de terreno, quanto maior for o CA, maior é a verticalização.

Figura 12 – Representação do prédio como um paralelepípedo

- Coeficiente de Aproveitamento

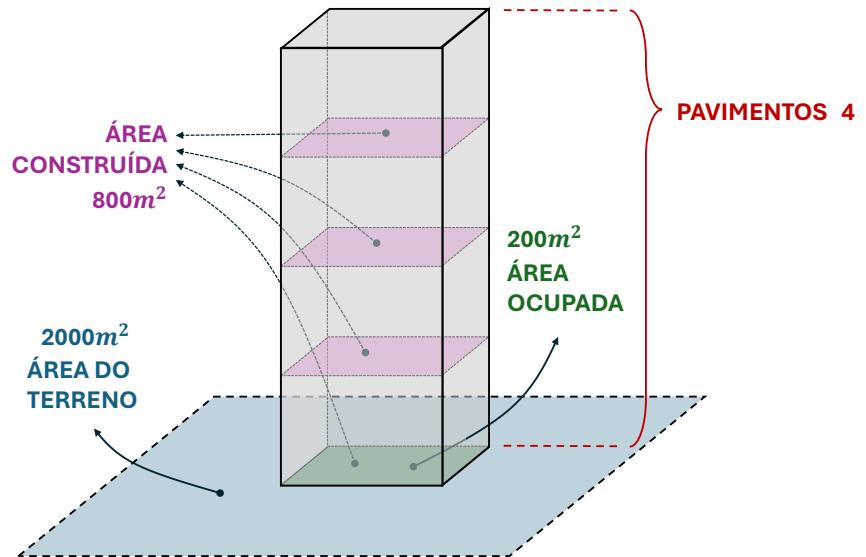
$$CA = \frac{AC}{AT} = \frac{800}{2000} = 0,4$$

- Taxa de ocupação

$$tx_{ocup} = \frac{AO}{AT} = \frac{200}{2000} = 10\%$$

- Verticalização

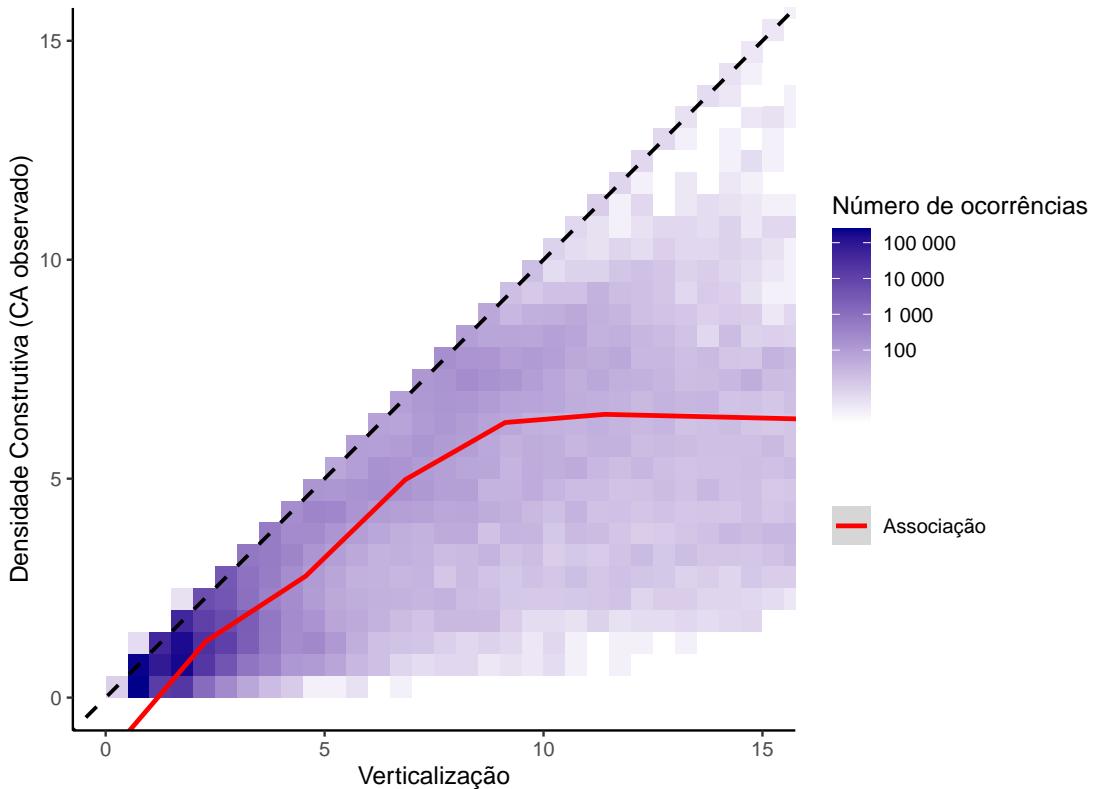
$$vert = \frac{CA}{tx_{ocup}} = \frac{0,4}{0,1} = 4$$



$$\text{Pavimentos} = \frac{AC}{AO} = \frac{AC}{AT \cdot \underbrace{\frac{AO}{AT}}_{tx_{ocup}}} = \frac{AC}{AT} \div \frac{AO}{AT} = \frac{CA}{tx_{ocup}} \quad (\text{A.1})$$

Dessa forma, o número de pavimentos nunca será maior do que o CA, dado que a taxa

Figura 13 – Relação entre densidade construtiva e verticalização



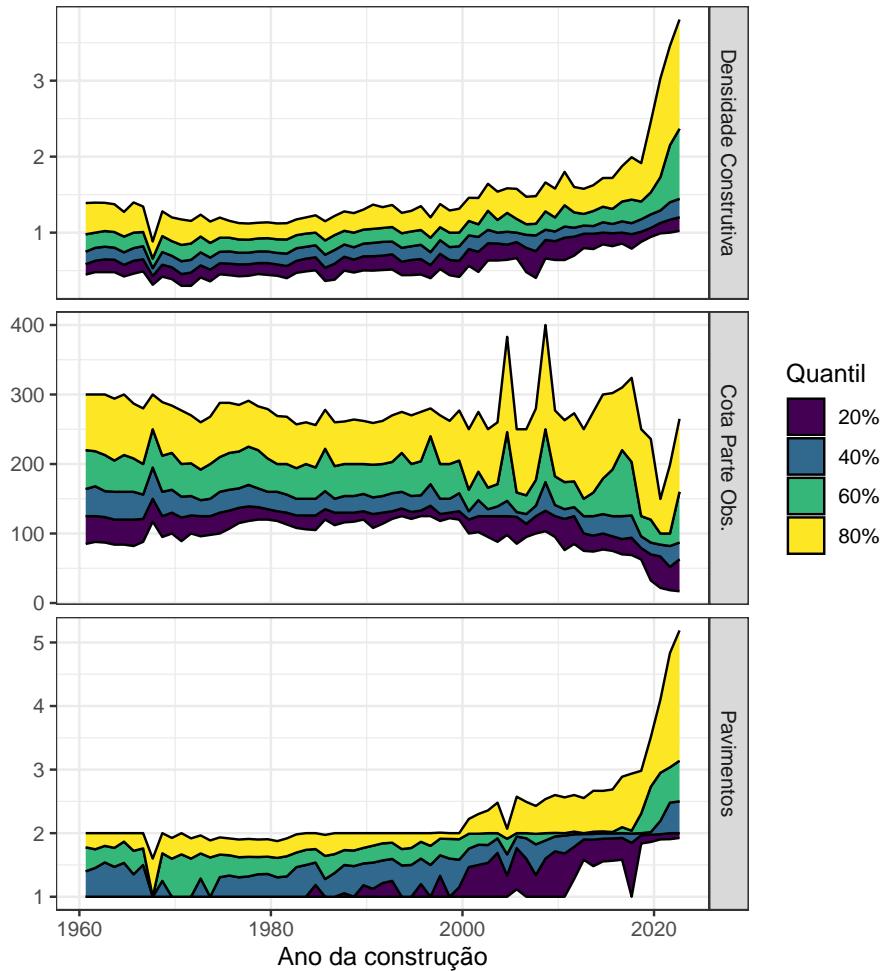
de ocupação sempre é um número que varia entre zero e um. Na Figura 13 é possível observar a relação entre as duas variáveis. Para lotes muito pequenos, as duas medidas geralmente são iguais, visto que geralmente se ocupa 100% da área do terreno. Na medida em que a verticalização aumenta, é observável que o CA não acompanha este crescimento, o que indica que há uma queda na taxa de ocupação.

## APÊNDICE B – Discussões sobre a informalidade

## APÊNDICE C – Metodologia de cruzamento geográfico dos dados

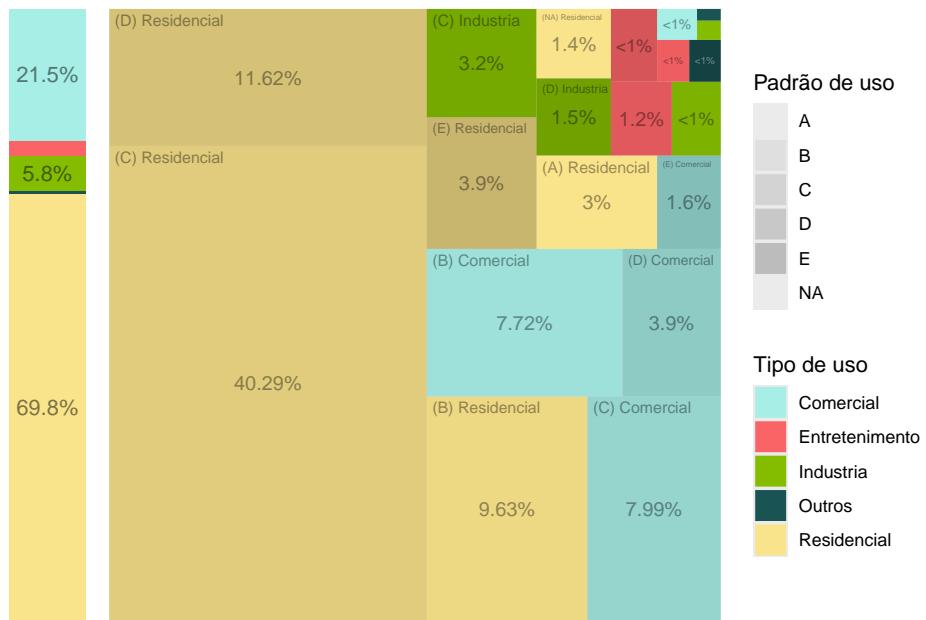
## APÊNDICE D – Outras figuras e tabelas

Figura 14 – Variação no tempo dos padrões construtivos residenciais observados em São Paulo



Ao analisar os padrões construtivos ao longo do tempo na Figura 14, fica evidente a trajetória de adensamento. Na figura, cada linha representa um quantil dos lotes do IPTU. Ao ordenar as construções no ano 2020 de maneira crescente de CA observado, o empreendimento na posição 80% dessa fila apresenta um CA observado de 3,8. Este mesmo procedimento nos anos 2000 retornaria um empreendimento com CA observado igual a 1,46. Analogamente, a cota parte observada que em 2020 apresenta o quantil 20% de meros 21,7 metros quadrados, no ano 2000 apresentaria 100 metros quadrados. Isso evidencia que nos últimos anos São Paulo passou por um forte processo de adensamento.

Figura 15 – Área construída no município de São Paulo por tipo e padrão de uso



Na Figura 15, é possível observar a distribuição da área construída por tipo e padrão de uso. Em São Paulo, 70% da área construída é residencial, sendo que o padrão residencial mais comum é o "C". O padrão é uma divisão feita no cálculo do IPTU ([Município de São Paulo, 1986](#)), para criar descontos do imposto para imóveis com características desejáveis do ponto de vista do planejamento urbano, como maior adensamento, ou para imóveis de pessoas com menor poder aquisitivo. Medidas como metragem, pé direito, vaga de garagem, número de pavimentos, elementos arquitetônicos, materiais de construção, etc., são fatores levados em consideração para determinar o padrão. No caso do "C", ele está no meio da escala de densidade, sendo "A" o mais denso, e "E" o menos. O segundo uso mais comum é o comercial, seguido do industrial e entretenimento. Na categoria de entretenimento, se encontram templos religiosos, clubes, estádios esportivos, cinemas, aeroporto, museu, zoológico, entre outros.

Figura 16 – Densidade populacional em São Paulo por setor censitário (Censo 2022)

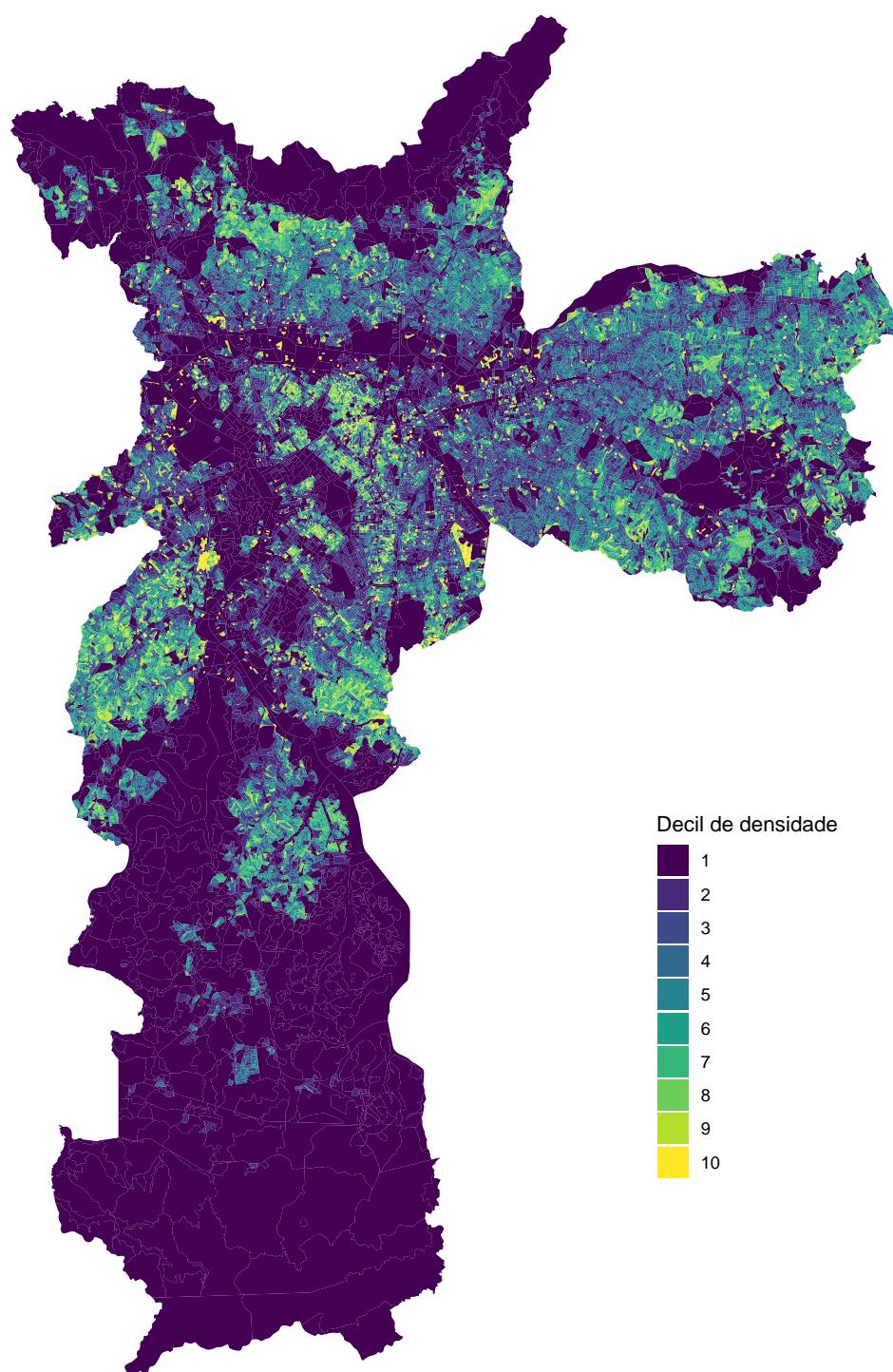


Figura 17 – Regiões de eixos (pintadas em azul)

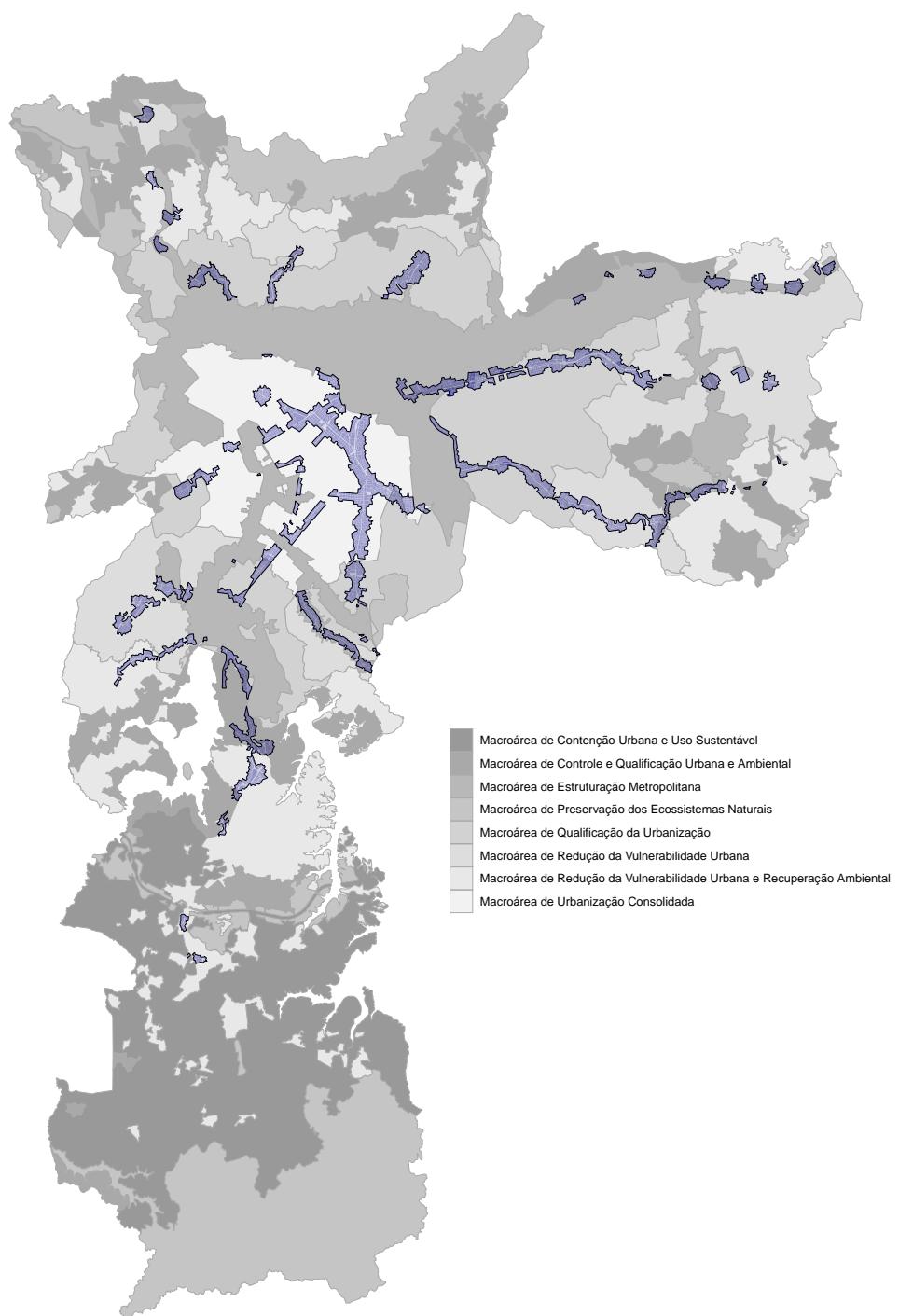


Figura 18 – Distribuição da quantidade de moradores em cada célula do raster

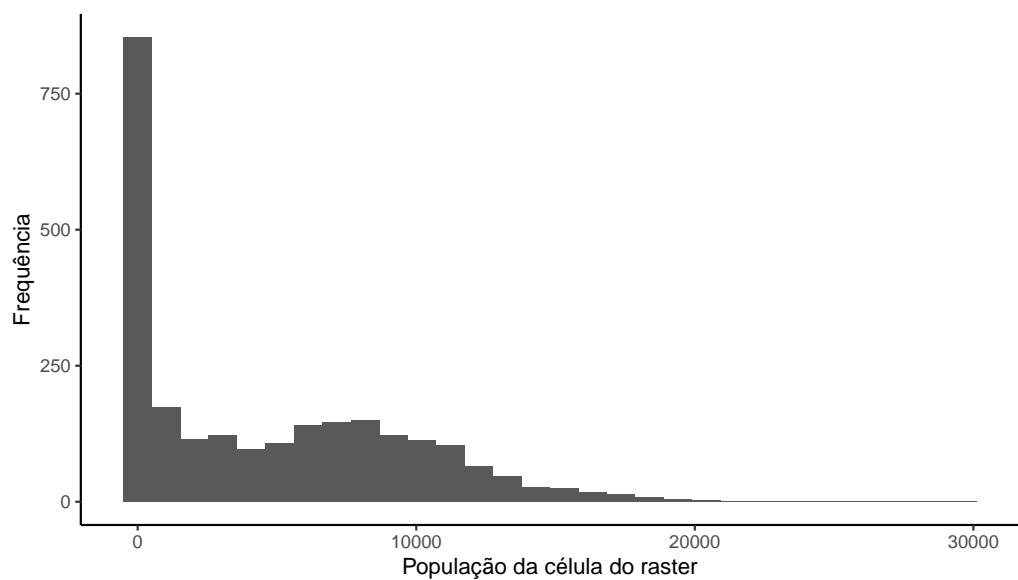


Figura 19 – Robustez dos resultados da regressão para diferentes espectros de irregularidade

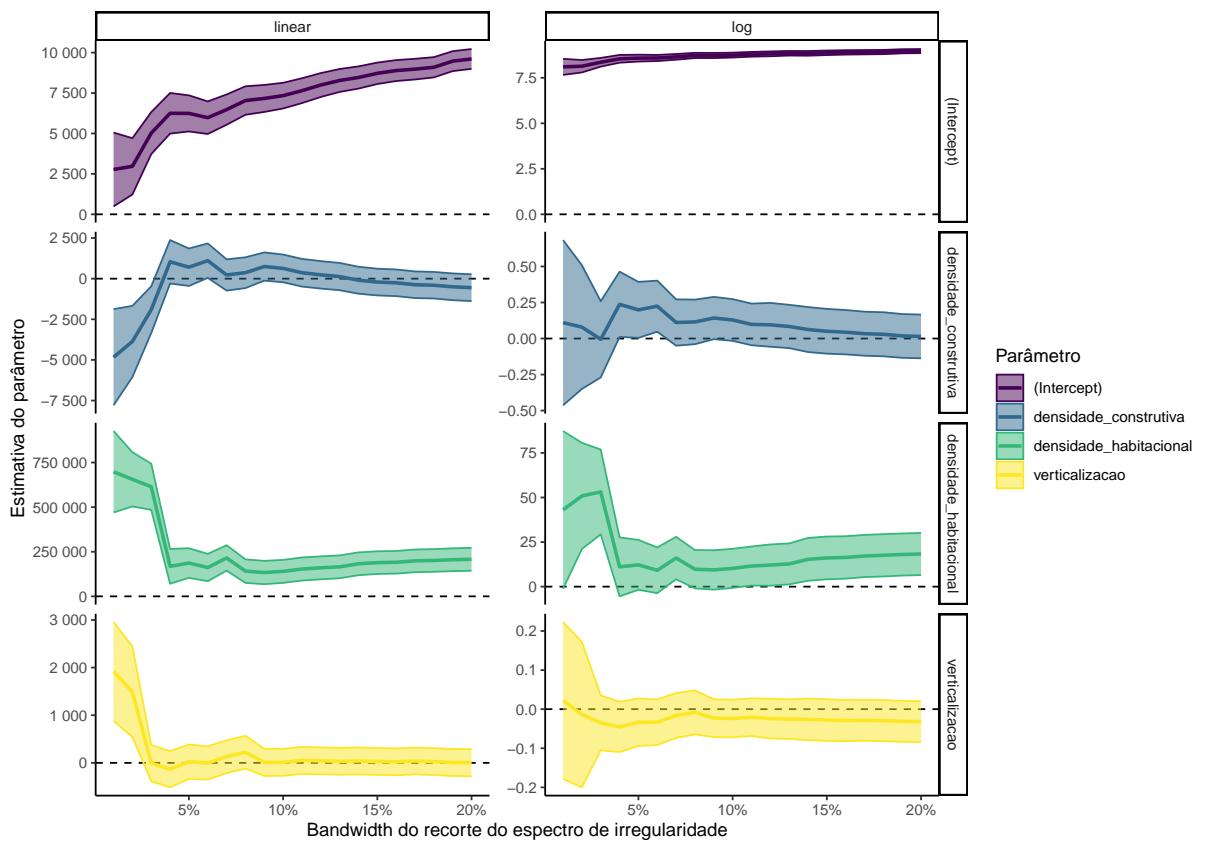


Figura 20 – Robustez dos resultados do DiD para diferentes recortes de proximidade

