

**TEXTO PARA DISCUSSÃO**

**2849**

**INADEQUAÇÕES HABITACIONAIS E  
INFORMALIDADE NO MERCADO DE  
HABITAÇÃO BRASILEIRO: EVIDÊNCIAS  
EMPÍRICAS A PARTIR DO CENSO 2010**

**VANESSA GAPRIOTTI NADALIN**



**INADEQUAÇÕES HABITACIONAIS E  
INFORMALIDADE NO MERCADO DE  
HABITAÇÃO BRASILEIRO: EVIDÊNCIAS  
EMPÍRICAS A PARTIR DO CENSO 2010<sup>1</sup>**

**VANESSA GAPRIOTTI NADALIN<sup>2</sup>**

---

1. A autora agradece aos pareceristas Bruno Cruz e Cleandro Krause pelos comentários que muito contribuíram para a melhora do texto. Erros remanescentes são de responsabilidade exclusiva da autora.

2. Técnica de planejamento e pesquisa na Diretoria de Estudos e Políticas Regionais, Urbanas e Ambientais do Instituto de Pesquisa Econômica Aplicada (Dirur/Ipea). *E-mail*: <vanessa.nadalin@ipea.gov.br>.

**Governo Federal**

**Ministério do Planejamento e Orçamento**

**Ministra** Simone Nassar Tebet

**ipea** Instituto de Pesquisa  
Econômica Aplicada

Fundação pública vinculada ao Ministério do Planejamento e Orçamento, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiros – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

**Presidenta (substituta)**

LUCIANA MENDES SANTOS SERVO

**Diretor de Desenvolvimento Institucional (substituto)**

SÉRGIO VINÍCIUS MARQUES DO VAL CÔRTEZ

**Diretor de Estudos e Políticas do Estado,  
das Instituições e da Democracia (substituto)**

BERNARDO ABREU DE MEDEIROS

**Diretor de Estudos e Políticas  
Macroeconômicas (substituto)**

FRANCISCO EDUARDO DE LUNA ALMEIDA SANTOS

**Diretor de Estudos e Políticas Regionais,  
Urbanas e Ambientais (substituto)**

BOLÍVAR PÊGO FILHO

**Diretor de Estudos e Políticas Setoriais, de Inovação,  
Regulação e Infraestrutura (substituto)**

EDISON BENEDITO DA SILVA FILHO

**Diretora de Estudos e Políticas Sociais (substituta)**

ANA LUIZA MACHADO DE CODES

**Diretor de Estudos Internacionais (substituto)**

FERNANDO JOSÉ DA SILVA PAIVA RIBEIRO

**Coordenador-Geral de Imprensa e Comunicação Social**

JOÃO CLÁUDIO GARCIA RODRIGUES LIMA

**Ouvidoria:** <http://www.ipea.gov.br/ouvidoria>

**URL:** <http://www.ipea.gov.br>

## Texto para Discussão

Publicação seriada que divulga resultados de estudos e pesquisas em desenvolvimento pelo Ipea com o objetivo de fomentar o debate e oferecer subsídios à formulação e avaliação de políticas públicas.

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2023

Texto para discussão / Instituto de Pesquisa Econômica Aplicada.- Brasília : Rio de Janeiro : Ipea , 1990-

ISSN 1415-4765

1. Brasil. 2. Aspectos Econômicos. 3. Aspectos Sociais.  
I. Instituto de Pesquisa Econômica Aplicada.

CDD 330.908

As publicações do Ipea estão disponíveis para *download* gratuito nos formatos PDF (todas) e EPUB (livros e periódicos).  
Acesse: <http://www.ipea.gov.br/portal/publicacoes>

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério do Planejamento e Orçamento.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

JEL: R31; O18.

DOI: <http://dx.doi.org/10.38116/td2849>

# SUMÁRIO

SINOPSE

ABSTRACT

1 INTRODUÇÃO .....	6
2 CONJUNTO DAS CIDADES COM AGSN .....	8
3 OS MERCADOS HABITACIONAIS DAS MAIORES CIDADES.....	11
4 MAPAS DOS 70 ARRANJOS POPULACIONAIS E 106 MUNICÍPIOS ISOLADOS .....	18
5 OS MERCADOS HABITACIONAIS DOS MUNICÍPIOS .....	22
6 <i>CLUSTERS</i> DE SETORES CENSITÁRIOS NO RIO DE JANEIRO E EM SÃO PAULO: OS AGSNS SÃO UM <i>CLUSTER</i> SEPARADO? .....	32
7 COMENTÁRIOS FINAIS.....	47
REFERÊNCIAS .....	48
APÊNDICE A .....	51
APÊNDICE B .....	52
APÊNDICE C .....	53
APÊNDICE D .....	54



## SINOPSE

Este trabalho busca caracterizar o mercado habitacional formal e informal brasileiro e possíveis segmentações nessas duas primeiras divisões. O principal objetivo é explicitar as semelhanças e as diferenças, e as heterogeneidades em cada grupo. Levando em consideração a inadequação habitacional, há uma parte do mercado formal que é semelhante ao mercado informal. Assim como há uma parte do mercado informal que é semelhante ao mercado formal, e, portanto, livre de inadequações habitacionais mais graves. Dessa maneira, o trabalho auxilia a responder esta pergunta essencial: o quanto a informalidade influencia a disponibilização de moradia digna. Foi utilizado como definição de bairro informal o conceito de aglomerado subnormal (AGSN), do Instituto Brasileiro de Geografia e Estatística (IBGE) para o Censo Demográfico 2010. Este estudo inova ao caracterizar os dois mercados a partir de informações do questionário da amostra do Censo 2010, cuja identificação de pertencimento a um AGSN foi feita na sala do sigilo no IBGE. Duas informações importantes disponibilizadas são as que se referem ao valor do aluguel e ao tempo de deslocamento casa-trabalho. São feitas análises descritivas utilizando-se, entre outros, os indicadores de autocorrelação espacial local (Lisa) e os gráficos de dispersão e regressão (*binscatter regressions*). Uma análise de *clusters* para setores censitários nas metrópoles do Rio de Janeiro e de São Paulo indicou que o segmento de mercado de *baixíssima renda* foi encontrado no *cluster* que concentrou as inadequações habitacionais mais graves, chamado de *sem cidade*. Tanto em São Paulo quanto no Rio de Janeiro, a grande maioria dos moradores de *sem cidade* não estão em setores classificados como AGSN.

**Palavras-chave:** Brasil; favelas; aglomerados subnormais; assentamentos informais; Censo Demográfico; mercado de habitação.

## ABSTRACT

This work seeks to describe the Brazilian formal and informal housing markets and the possible segmentations within these first two divisions. The main objective is to make explicit the similarities and differences, and the heterogeneities within each group. Considering inadequate housing features, there is a part of the formal market that is similar to the informal market. Just as there is a part of the informal market that is similar to the formal market, and therefore is free from more serious housing inadequacies. In this way, the work helps to answer an essential question: how much does informality influence the availability of adequate housing? The IBGE's concept of informal settlements, *aglomerados subnormais* (AGSN), was used as a proxy for informal settlements. This work innovates by describing the two markets based on information from the 2010 Census sample questionnaire, whose identification of belonging to a subnormal agglomeration was made inside the IBGE's confidential data center. Two important pieces of information available are those referring to the rent value and time spent commuting from home to work. Descriptive analyzes are performed using, among others, local spatial autocorrelation indicators (Lisa), scatter plots and binscatter regressions. A cluster analysis for census enumeration districts in the metropolises of Rio de Janeiro and São Paulo indicated that the *very low-income* market segment was found in the cluster that concentrated the most serious housing inadequacies, called *no city*. In both São Paulo and Rio de Janeiro, the vast majority of residents of *no city* are not in sectors classified, by IBGE, as AGSN.

**Keywords:** Brazil; slums; informal settlements; Demographic Census; housing market.

## 1 INTRODUÇÃO

Como todos os mercados de bens, o mercado de habitação apresenta segmentações. Mas quais são os segmentos do mercado de habitação no Brasil? Tradicionalmente, os segmentos do mercado de uma cidade estão baseados em preços semelhantes em um mesmo *bairro*, ou porção territorial da cidade (Goodman e Thibodeau, 1998). Ao mesmo tempo, muito se fala sobre a cidade dual, a cidade formal e informal. O Fórum Urbano Mundial, organizado pela Organização das Nações Unidas (ONU) no Rio de Janeiro em 2010, por exemplo, tinha no título *Bridging the urban divide*, propondo uma ponte para essa divisão urbana. Nesse sentido, este trabalho busca caracterizar o mercado formal e informal brasileiro e possíveis segmentações nessas duas primeiras divisões.

O principal objetivo é explicitar as semelhanças e as diferenças, e as heterogeneidades em cada grupo. Levando em consideração a inadequação habitacional, há uma parte do mercado formal que é semelhante ao mercado informal. Assim como há uma parte do mercado informal que é semelhante ao mercado formal, e, portanto, livre de inadequações habitacionais mais graves. Portanto, este estudo auxilia a responder esta pergunta essencial: o quanto a informalidade influencia a disponibilização de moradia digna?

Os preços dos imóveis são resultado do equilíbrio do mercado para esse bem heterogêneo: da interação entre oferta e demanda. Refletem a valorização dos seus principais componentes, quais sejam: estrutura física (número de dormitórios, se é apartamento, por exemplo); amenidades urbanas (áreas verdes e paisagismo do bairro, proximidade a equipamentos de lazer) e inserção urbana (acessibilidade a oportunidades de emprego, equipamentos de saúde, educacionais) – Rosen (1974). Em termos das diferenças entre mercados, entre diferentes cidades, outras características agregadas explicam a variação dos preços, como os salários dos trabalhadores e as amenidades ambientais das cidades (estar na beira do mar, por exemplo) – Roback (1982).

A princípio, o mercado habitacional informal apresentaria componentes que se valorizariam de forma semelhante. O surgimento, o crescimento e as principais características das favelas no Brasil estão em sintonia com a ocupação de áreas não disponíveis para o mercado formal, ou de menor preço, mas ainda próximas aos polos econômicos das cidades. Ao mesmo tempo, as favelas também se identificam simplesmente com ocupações irregulares, que por serem irregulares são opções de moradia mais baratas, na periferia (Brueckner, Mation e Nadalin, 2019; Cavalcanti, Da Mata e Santos, 2019; Smolka e Biderman, 2011). Ou seja, em alguma medida, as favelas brasileiras são submercados de baixa renda que competem com o mercado formal de baixa renda. Há evidências de que o processo de valorização de localidades mais centrais cria condições para que bairros informais bem localizados sejam formalizados, fazendo com que a boa inserção urbana seja uma característica cada vez mais presente no mercado formal. Henderson, Regan e Venables (2021), por exemplo, verificam empiricamente e modelam teoricamente esse processo para Nairóbi,

no Quênia. Portanto, bairros informais mais centrais tenderiam a ser mais diferentes de seus vizinhos formais que os bairros informais periféricos. A partir do momento que os bairros informais centrais se formalizam, e se gentrificam, essas diferenças desapareceriam.

Foi utilizado como definição de bairro informal o conceito de aglomerado subnormal (AGSN), do Instituto Brasileiro de Geografia e Estatística (IBGE) para o Censo Demográfico 2010 (IBGE, 2010a). A definição de AGSN do IBGE é adequada, porém inacurada para muitos aspectos da informalidade. Assim, tomamos essa classificação como uma medida *proxy* da informalidade.<sup>1</sup> Este trabalho inova ao caracterizar os dois mercados a partir de informações do questionário da amostra do Censo 2010, cuja identificação de pertencimento a um AGSN foi feita na sala do sigilo do IBGE. Duas informações importantes disponibilizadas são as que se referem ao aluguel e ao tempo de deslocamento. Além disso, também é inovadora a análise conjunta da proximidade a empregos formais geolocalizados, a partir das informações dos empregos formais da Relação Anual de Informações Sociais (Rais), da declividade média do setor censitário e da distância dos setores censitários ao CBD de cada cidade.

Utilizamos como amostra do estudo 510 municípios brasileiros que ou apresentavam AGSN, ou pertenciam a um arranjo populacional (IBGE, 2016) em que algum dos municípios possuía AGSN. Segundo o Censo 2010, o total de população desses 510 municípios correspondia a 56,2% da população do Brasil em 2010, sendo a população em AGSN 11,4% dessa população (IBGE, 2010b).<sup>2</sup> São feitas análises descritivas utilizando-se, entre outros, os indicadores de autocorrelação espacial local (Lisa), os gráficos de dispersão e regressão (*binscatter regressions*) e uma análise de *clusters* para setores censitários nas metrópoles do Rio de Janeiro e de São Paulo. Também analisamos os agrupamentos urbanos que congregam esses 510 municípios em 70 aglomerados e 106 municípios isolados. Estes agrupamentos são chamados de arranjos populacionais. Os arranjos populacionais são “agrupamentos de dois ou mais municípios com forte integração populacional, devido aos movimentos pendulares para trabalho ou estudo, ou à contiguidade entre manchas urbanas” (IBGE, 2016).

Exploramos as diferenciações dos mercados em cinco níveis geográficos, quais sejam: i) o conjunto das cidades com AGSN (média dos 510 municípios); ii) os mercados habitacionais das maiores cidades (os treze maiores arranjos populacionais); iii) os 106 municípios isolados e os 70 arranjos populacionais (representados em mapas do Brasil); iv) as diferenças entre os 510 municípios (visualizando a

---

1. A pesquisa Núcleos Urbanos Informais, capitaneada pelo Ipea, em parceria com diversas instituições, e encomendada pelo Ministério do Desenvolvimento Regional (MDR), investiga os limites da definição de AGSN do IBGE para captar o fenômeno da informalidade habitacional urbana.

2. População do Brasil: 178.360.003; população dos 510 municípios: 100.229.069; população em AGSN: 11.413.993.

diferença entre municípios-sede, do entorno e isolados); e v) os setores censitários em São Paulo e no Rio de Janeiro (na análise de *clusters* de setores censitários). Desse modo, a estrutura deste trabalho segue os níveis geográficos, do mais abrangente para o menos abrangente.

## 2 CONJUNTO DAS CIDADES COM AGSN

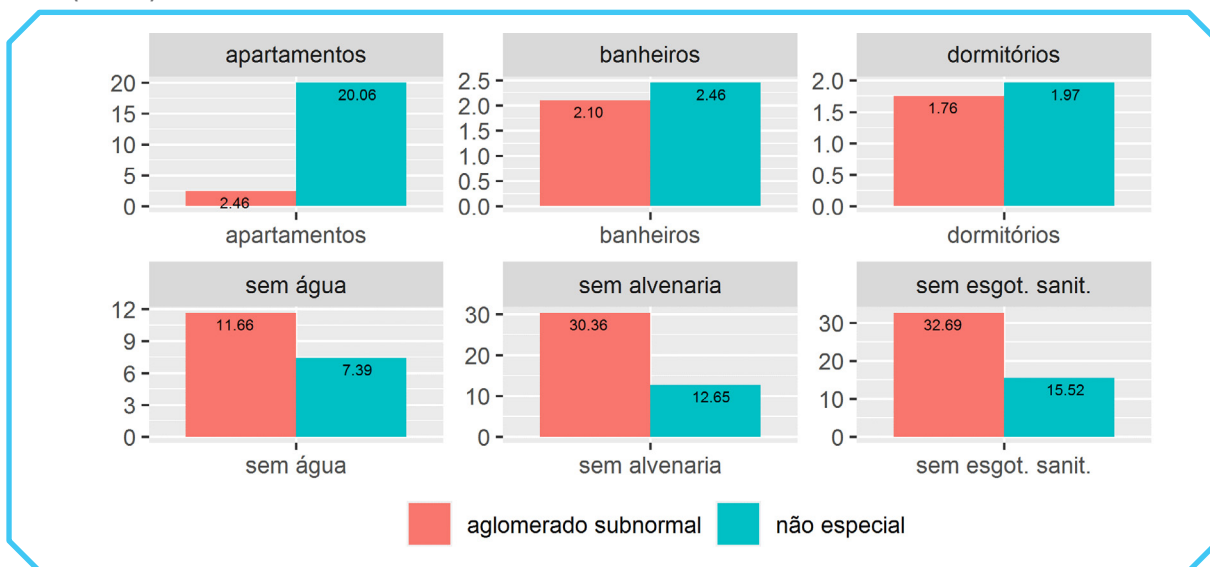
Esta seção traz as estatísticas descritivas provenientes das informações de todos domicílios da amostra nos 510 municípios analisados: 2.089.661 amostrados, desses, 196.150 em AGSN. Um bom “resumo” das diferenças nas características dessa divisão.

O gráfico 1 mostra as médias das características das estruturas físicas dos domicílios. Domicílios em AGSN, na grande maioria, não são apartamentos, têm menos banheiros, menos dormitórios, menos paredes de alvenaria, menos abastecimento de água e esgotamento adequados – definições segundo o Objetivo de Desenvolvimento Sustentável 11 (ODS 11).<sup>3</sup>

### GRÁFICO 1

#### Características dos domicílios: médias

(Em %)



Fonte: Censo 2010. Disponível em: <<https://bit.ly/3kPWyuS>>. Acesso em: 30 jan. 2023.

Obs.: 1. Ver definições de água e esgotamento inadequados em: <<https://bit.ly/3G6D8dm>>.

2. Unidades: proporção de domicílios com as características, à exceção do número médio de dormitórios e banheiros.

3. Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

3. Disponível em: <<https://bit.ly/3G6D8dm>>.

**TEXTO para DISCUSSÃO**

Já o gráfico 2 mostra a inserção urbana dos AGSNs, condizente com as descrições anteriores. São regiões da cidade que ocupam mais os terrenos de alta declividade, mas também os de menor declividade. Com mais frequência, os AGSNs se localizam mais periféricamente, estão próximos de menos empregos formais e levam a mais tempo no deslocamento de casa para o trabalho. A mediana da distância ao centro de negócios (central bussiness district – CBD) é um pouco maior para os AGSNs, mas esta distância varia menos que os não especiais.

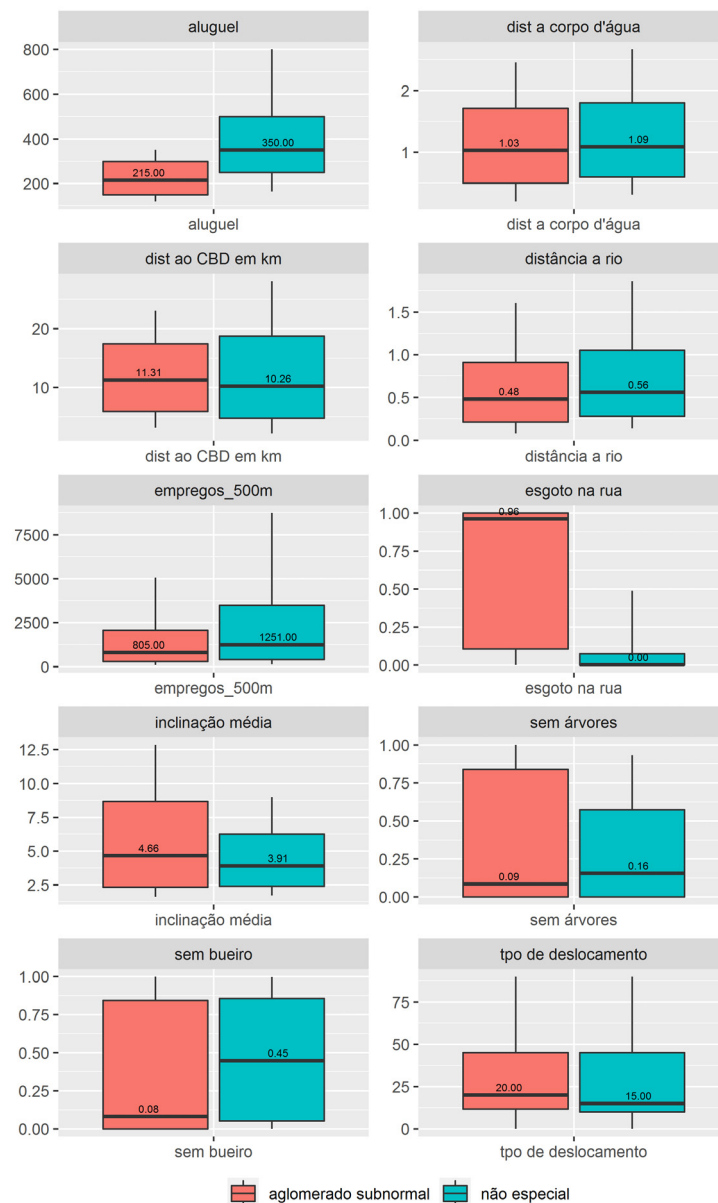
A declividade média do terreno é maior, com intervalo interquartil (IQR)<sup>4</sup> maior. Já o tempo de deslocamento é maior, com IQR menor e a proximidade a empregos formais menor, com IQR menor. A mediana do aluguel é bem menor, IQR com pouca sobreposição ao formal (150 a 300 *versus* 250 a 500), sendo a primeira evidência de que o mercado de habitação informal seria um mercado separado do mercado formal. Das variáveis do entorno, a que mais chama atenção é a presença de esgoto na rua: a mediana é 96% contra 0,4%, sendo que o percentil 25 dos AGSNs é maior (10,7%) que o percentil 75 dos domicílios em áreas não especiais (7,5%). O que evidencia uma condição urbana muito pior para os AGSNs.

---

4. O intervalo interquartil mede a diferença entre o terceiro quartil e o primeiro quartil; que também são os limites da “caixa” do gráfico *boxplot*. Os quartis são os valores dos limites de intervalos ao ordenar a variável em ordem crescente e dividir a amostra em quatro partes iguais.

## GRÁFICO 2

Boxplot da inserção urbana, do aluguel e das características do entorno: total da amostra



Fontes: Censo 2010. Para empregos, Rais (Disponível em: <<https://bit.ly/3wN0gs2>>). Para informações geográficas, projeto OpenStreet Map (disponível em: <<https://bit.ly/3JvkxtA>>). Para a topografia do terreno, sensor de satélite STRM (disponível em: <<https://bit.ly/3HwqSme>>). Acessos em: 30 jan. 2023.

Obs.: 1. Aluguéis em R\$ de 2010; distâncias em quilômetros; esgoto na rua, sem árvores e sem bueiro: proporção de domicílios com as características; inclinação média em porcentagem; tempo de deslocamento em minutos; corpos d'água são acumulações de água, como mares, lagos e reservatórios.

2. Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).



### 3 OS MERCADOS HABITACIONAIS DAS MAIORES CIDADES

Este trabalho explora a ideia de que o mercado de habitação é segmentado, no caso do Brasil entre mercado formal e informal. As segmentações mais comuns são em bairros, regiões da cidade que compartilham de preços semelhantes. Mesmo assim, os diferentes segmentos de mercado ou submercados de uma mesma metrópole também podem ser visualizados na sua totalidade, e considerados como um mercado habitacional único, específico a cada cidade. Isto acontece porque as diferentes localidades em uma mesma aglomeração urbana pertencem a um mesmo mercado de trabalho, onde potencialmente é possível trabalhar e morar em qualquer par de localizações dentro dessa região.

Portanto, vale a pena visualizar as principais características dos mercados habitacionais das principais cidades do país. Dadas as diferentes trajetórias de desenvolvimento institucional e econômico, os diferentes suportes geográficos e as diferentes localizações relativas há grandes desigualdades entre esses mercados habitacionais.

Mostraremos essas características para os treze maiores arranjos populacionais do IBGE. Nos parece que estes agrupamentos de municípios propostos pelo IBGE são boas delimitações das maiores regiões metropolitanas (RMs) brasileiras, e boas aproximações do que seriam os grandes mercados de trabalho e de habitação correspondentes a essas grandes *cidades*, ou RMs. Nesta seção chamaremos de *cidades* esses arranjos populacionais. As RMs brasileiras oficiais são delimitadas por lei estadual e não seguem critérios homogêneos para a inclusão de municípios. Para as grandes metrópoles, não há diferenças muito significativas, pois as diferenças entre os conjuntos de municípios considerados acabam correspondendo a municípios muito pequenos.

#### 3.1 Inserção urbana

Começamos com as diferenças na estrutura urbana, visíveis a partir da distância dos domicílios amostrados ao CBD e da quantidade de empregos formais existentes em um *buffer* de 500 m a partir dos limites do setor censitário no qual o domicílio está inserido. As diferenças entre os domicílios em setores não especiais e domicílios em AGSN exemplificam as desigualdades na inserção urbana desses domicílios. Os gráficos mostram, em formato de *boxplot*, os valores dos percentis 10, 25, 50, 75 e 90.<sup>5</sup> Em primeiro lugar, sobressaem-se as cidades com maiores valores:

---

5. Os percentis se referem aos valores que uma variável toma quando ordenamos em uma fila os indivíduos de uma amostra do menor valor ao maior valor da esquerda (menor) para a direita (maior), por exemplo. O décimo percentil, por exemplo, corresponde ao valor do indivíduo que está na posição da fila em que 10% dos indivíduos estão à sua esquerda, apresentando valores menores que o dele e 90% dos indivíduos estão à direita, estes apresentando valores maiores que o dele. O percentil 50 é a mediana da distribuição da variável.

Brasília e Rio de Janeiro, cidades que ocupam uma extensão muito ampla de território. Brasília é muito espalhada por conformação do tecido urbano fragmentado e descontínuo; o Rio de Janeiro é muito espalhado devido à geografia, barreiras físicas à ocupação urbana.

Evidentemente que quanto maior a população da cidade, maior a distância ao CBD, pois há a necessidade de usar mais terreno para construir mais casas para abrigar mais população. Mas mesmo cidades de porte semelhante – por exemplo, Brasília e Fortaleza, São Paulo e Rio de Janeiro ou Curitiba e Salvador – apresentam diferenças amplas nas medianas e nos demais percentis da distribuição das distâncias ao centro dos domicílios amostrados pelo censo, indicando diferenças na estrutura interna urbana, como diferentes valores de densidade interna e configurações intraurbanas de polos de concentração de população.

Outra característica muito diferenciada entre as cidades é a inserção urbana dos AGSNs. Em geral, os AGSNs se distribuem em distâncias ao centro menores que os domicílios em setores não especiais. A exceção claríssima é Manaus, onde uma minoria dos domicílios em AGSN fica próximo ao CBD; Belém também apresenta AGSN mais longínquos. Por sua vez, no Rio de Janeiro, em Porto Alegre e em Fortaleza, há uma boa porção de domicílios em AGSN localizados mais centralmente que os domicílios não especiais. Já em São Paulo e em Brasília, os AGSNs estão mais distantes do centro. Note que na seção 2 deste texto, para o conjunto dos domicílios amostrados pelo censo, a mediana da distância dos AGSNs era superior à dos domicílios não especiais. Aqui a informação segregada por cidade traz um retrato mais detalhado, ilustrando a grande heterogeneidade que o dado agregado estava ocultando.

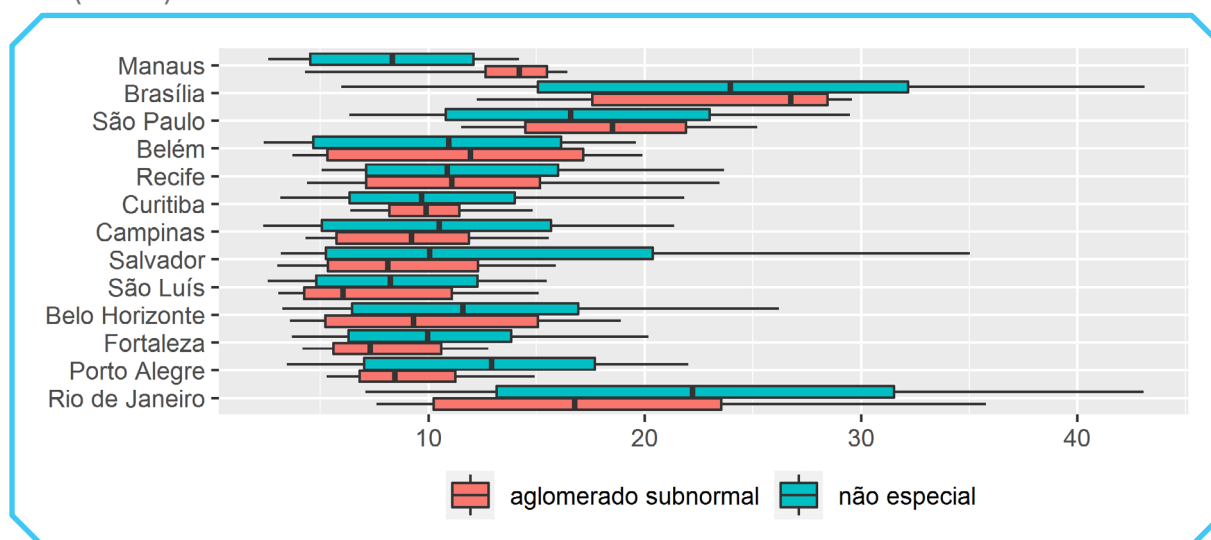
Esse padrão da proximidade ao CBD também se repete na distribuição dos totais de empregos formais próximos aos domicílios. No entanto, esse indicador também é sensível a o quanto os empregos estão concentrados, e quantos subcentros de empregos existem na metrópole, trazendo Belém como uma cidade com boa proximidade de empregos e domicílios, por exemplo. Em Manaus, a maior parte dos domicílios em AGSN fica próxima a menos empregos que as áreas não especiais, assim como em Belém. Já Brasília, que não apresentava AGSNs discrepantes no quesito distância ao centro, aqui apresenta domicílios em AGSN com nítida desvantagem aos não especiais no quesito proximidade a empregos. São Paulo, que no gráfico 3, de distância ao centro, mostrou distâncias bem menores que o Rio de Janeiro, por exemplo, aqui apresenta AGSN com bem menos empregos formais próximos.

## TEXTO para DISCUSSÃO

### GRÁFICO 3

#### Distância ao centro de negócios: AGSN e setores não especiais

(Em km)



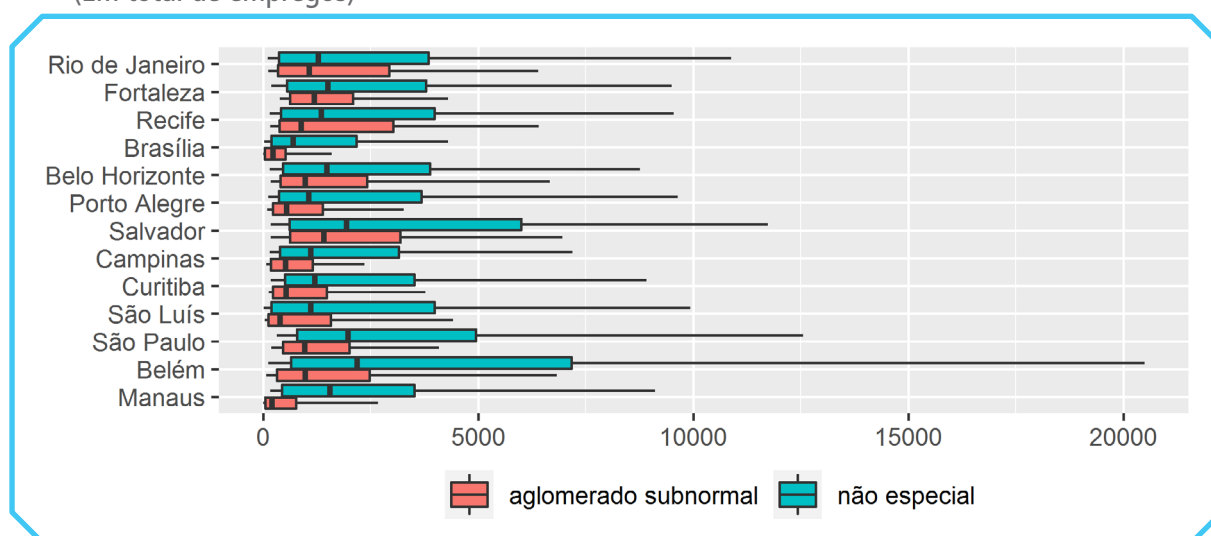
Fonte: Censo 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

### GRÁFICO 4

#### Proximidade e empregos dentro de *buffer* de 500 m: AGSNs e setores não especiais

(Em total de empregos)



Fonte: Censo 2010.

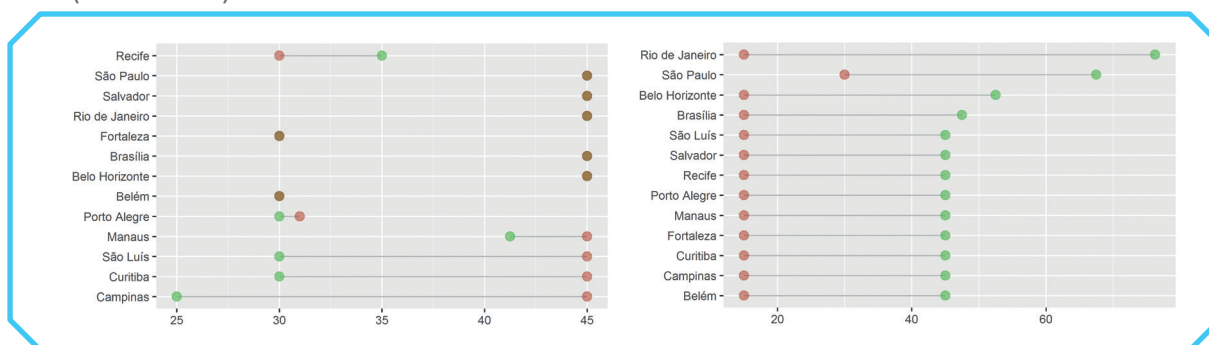
Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Esse quadro tanto da estrutura intraurbana diferenciada entre as cidades quanto da inserção urbana desigual entre domicílios não especiais e subnormais também se reflete no tempo de deslocamento casa-trabalho, uma das variáveis medidas no questionário da amostra do Censo 2010. A mediana do tempo de deslocamento é mostrada no gráfico 5, à esquerda, deixando claro como as desvantagens dos AGSNs em Manaus, descritas anteriormente, traduzem-se em maior tempo de deslocamento, por exemplo. As demais cidades com medianas subnormais superiores (Porto Alegre, São Luís, Campinas e Curitiba) também apresentam desvantagens em inserção urbana coerentes com esse tempo de deslocamento. Já o painel da direita do gráfico 5 apresenta a comparação entre a *melhor* porção dos AGSNs (percentil 25, menor tempo de deslocamento) e a *pior* porção dos não especiais (percentil 75, maior tempo de deslocamento). Nota-se que os melhores AGSNs sempre gastam menos tempo se deslocando (no mínimo 25 minutos a menos), sendo a diferença maior para o Rio de Janeiro. Isto nos mostra o quanto existe de heterogeneidade interna aos AGSNs. Mesmo em Manaus, há AGSNs muito bem localizados, muito melhores que os não especiais, apesar de serem poucos, como vimos anteriormente. Em Recife, a vantagem de inserção urbana dos AGSNs é mais homogênea, pois a própria mediana do tempo de deslocamento dos AGSNs é inferior à dos domicílios não especiais.

## GRÁFICO 5

**Medianas do tempo de deslocamento de AGSN e não especiais (à esquerda). Tempo de deslocamento: percentil 75 não especial e percentil 25 AGSN (à direita)**

(Em minutos)



Fonte: Censo 2010.

Obs.: 1. Quando há sobreposição dos pontos verdes (não especial) com os pontos vermelhos (AGSN), visualiza-se o ponto na cor marrom.

2. Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

### 3.2 Aluguel e amenidades urbanas

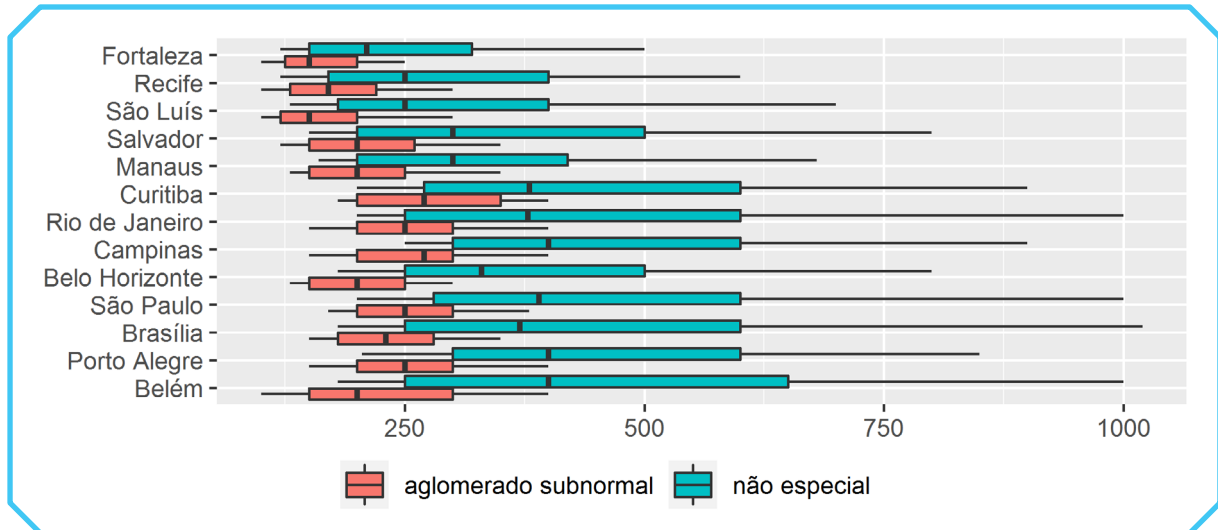
O questionário da amostra do Censo 2010 perguntou o valor do aluguel para imóveis alugados. O gráfico 6 traz essa preciosa informação para os mercados formal e informal nas treze cidades aqui analisadas. O mercado não especial apresenta uma amplitude maior de valores e valores de percentil 90 muito mais amplos que os valores de percentil 75. Indicando uma desigual distribuição de renda e de riquezas já muito conhecida e retratada por diversas fontes de dados. Os valores do mercado informal são menores, assim como sua amplitude (o intervalo interquartil) é menor. Os valores mais baixos do mercado informal não são alcançados pelo mercado formal.

Isso indica que há uma faixa de preços na qual há domicílios alugados tanto no mercado formal quanto no informal. Em São Paulo parece haver uma menor proporção de domicílios nessa faixa que em Curitiba, por exemplo. Em Belo Horizonte e Porto Alegre menos ainda. Para essas faixas de preços, ambos os mercados competem pelo mesmo consumidor, sendo interessante investigar as demais características desses mercados para compreender o que há de vantajoso na opção pelo mercado informal. A análise da inserção urbana, feita anteriormente, indica que há uma porção dos melhores domicílios subnormais que possui melhor inserção urbana que os piores domicílios em setores não especiais, nessas cidades, por exemplo. O caso do Rio de Janeiro indicou boas inserções urbanas dos AGSNs e, de fato, da mediana para cima o mercado informal carioca parece estar sobreposto ao mercado formal. Para ilustrar o quanto as amenidades de vizinhanças acompanham essa variação do aluguel, o gráfico 7 traz informações relativas ao entorno dos domicílios: a proporção de domicílios em ruas com calçadas no setor censitário do domicílio, coletadas pelos recenseadores do Censo 2010.

A diferença entre as medianas dos AGSNs e dos setores não especiais nos mostra como em Recife, Salvador e Curitiba há maior proporção de AGSN com calçadas que proporção de domicílios com calçadas em setores não especiais. Nas demais cidades, o entorno dos domicílios subnormais é pior que os setores não especiais, o já que era de se esperar. A comparação da *melhor* porção dos AGSNs (percentil 25) com a *pior* porção dos não especiais (percentil 75) mostra que, em todas as cidades, os melhores domicílios subnormais estão em setores censitários com menos domicílios sem calçadas que os piores não especiais. No Rio de Janeiro e em Curitiba, locais onde vimos uma proporção grande de domicílios em faixas comuns de preços ao mercado formal e informal, a vantagem dos melhores domicílios subnormais é maior, algo que pode explicar a escolha por morar em AGSN em detrimento de um setor não especial de mesmo preço.

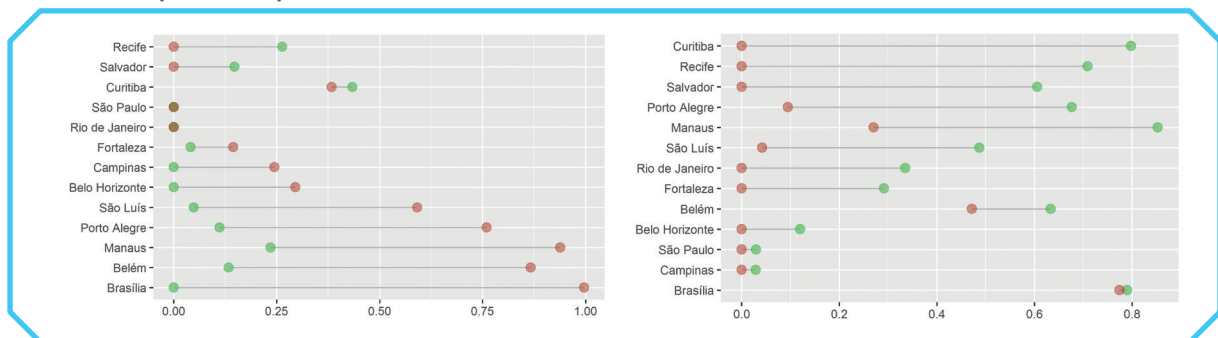
**GRÁFICO 6****Distribuição de valores de aluguéis: AGSNs e setores não especiais**

(Em R\$ de 2010)



Fonte: Censo 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

**GRÁFICO 7****Domicílios em ruas sem calçadas: medianas (à esquerda) e P75 não especial e P25 AGSN (à direita)**

Fonte: Censo 2010.

Obs.: 1. Não especial em verde, AGSN em vermelho.

2. Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

O gráfico 8 nos mostra outra característica do entorno dos domicílios: a presença de esgoto a céu aberto. Esta característica, à diferença da existência de calçada, é mais generalizadamente presente nos domicílios subnormais. Mesmo assim, para Belém, Fortaleza, São Luís, Manaus, Curitiba e Belo Horizonte, os *melhores* AGSNs são melhores que os *piores* não especiais.

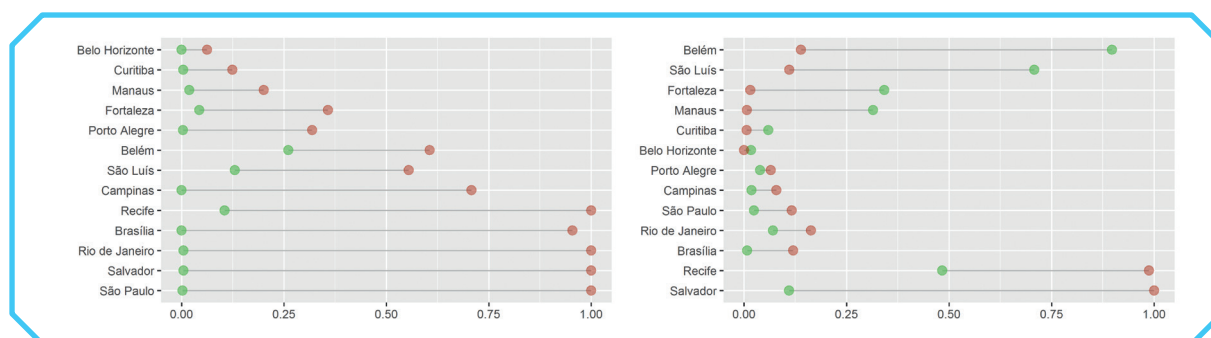


## TEXTO para DISCUSSÃO

Esses gráficos sustentam um retrato da desigualdade tanto entre o mercado formal e informal quanto em cada um desses mercados. O mercado informal possui características piores e aluguel inferior. A desigualdade de inserção urbana, de aluguel e de qualidade da vizinhança no mercado informal chama atenção, uma vez que comprova a heterogeneidade interna aos AGSNs e justifica que haja alguns bairros classificados como AGSN que são “concorrentes” do mercado formal, inclusive com amenidades urbanas melhores que estes.

### GRÁFICO 8

Esgoto a céu aberto: medianas (à esquerda) e P75 não especial e P25 AGSN (à direita)



Fonte: Censo 2010.

Obs.: 1. Não especial em verde, AGSN em vermelho.

2. Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Fica registrado também a heterogeneidade entre as treze *cidades*, ou arranjos populacionais, analisadas. Há combinações diferentes de desigualdades de inserção urbana/valor do aluguel e amenidades urbanas. Por exemplo, em Manaus, há clara desvantagem na inserção urbana dos AGSNs e das amenidades urbanas, mas não tão grande diferença no valor do aluguel. Já em Belo Horizonte, há grande diferença no aluguel e pouca diferença na inserção urbana e nas amenidades urbanas. O caso do Rio de Janeiro ilustra bem a presença de muitos domicílios em AGSN bem localizados; enquanto São Paulo tem AGSNs não bem localizados e com amenidades urbanas piores que os setores não especiais.

#### 4 MAPAS DOS 70 ARRANJOS POPULACIONAIS E 106 MUNICÍPIOS ISOLADOS

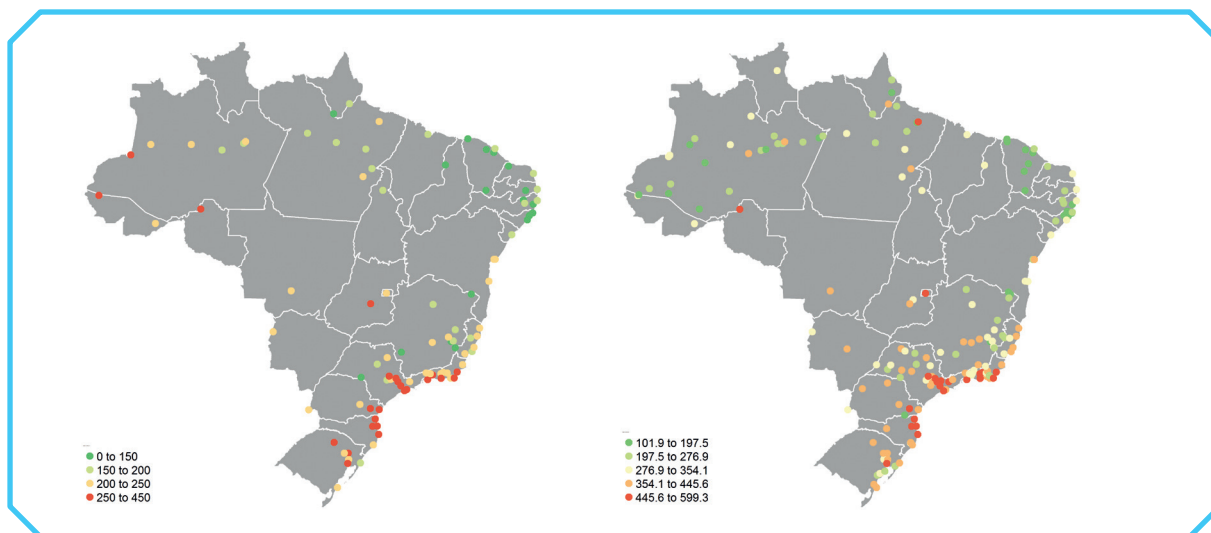
Aproveitando as informações dos 176 núcleos urbanos, sendo 70 arranjos populacionais e 106 municípios isolados com presença de AGSNs, visualizaremos as principais características no mapa do Brasil, buscando por padrões regionais de variação dessas características. Começaremos pelo mapa de aluguel e de renda do chefe do domicílio, na figura 1. Os mapas da distribuição dos valores dos aluguéis formais e informais sugerem aluguéis mais altos no Sul e no Sudeste e menores no Norte e no Nordeste. Note como as faixas de cores dos valores médios dos aluguéis informais correspondem a valores menores que as faixas de valores das mesmas cores do mapa de aluguéis médios formais, esses intervalos são calculados utilizando o método de Jenks. O mapa formal, inclusive, possui uma faixa de cores de R\$ 445,00 a R\$ 599,00 de aluguel médio que nem existe para o mercado informal.

No entanto, buscamos corroborar esse padrão espacial estatisticamente, para além da inspeção visual, calculando o indicador de autocorrelação espacial local (Lisa) – Anselin (1995). Esse indicador, além de quantificar o quanto observações próximas possuem valores próximos, indica se essa associação espacial é estatisticamente significativa, pois compara essa distribuição dos valores no espaço com possíveis distribuições *aleatórias* dos mesmos valores no espaço.

A análise do Lisa indicou a existência de regiões que concentram valores altos de aluguéis formais no Sul e no Sudeste e valores baixos de aluguéis no Nordeste, mas não encontramos padrão espacial no aluguel informal. A figura 2 traz esses resultados. Talvez isso aconteça porque não há muita variância nos valores dos aluguéis informais, uma vez que a porção mais humilde da população não é tão diferenciada entre as regiões brasileiras quanto a porção mais abastada. Há ricos muito mais ricos nas regiões mais ricas do país, enquanto os pobres não são tão diferenciados assim. A relação empírica entre os rendimentos das famílias, as amenidades urbanas e os valores dos aluguéis está documentada na literatura. Cravo, Resende e Cruz (2019) encontram o padrão regional compatível ao da figura 2 ao investigar essa relação.

**FIGURA 1**

Mapa coroplético da distribuição dos valores dos aluguéis AGSN (à esquerda) e não especiais (à direita): 70 arranjos populacionais e 106 municípios isolados

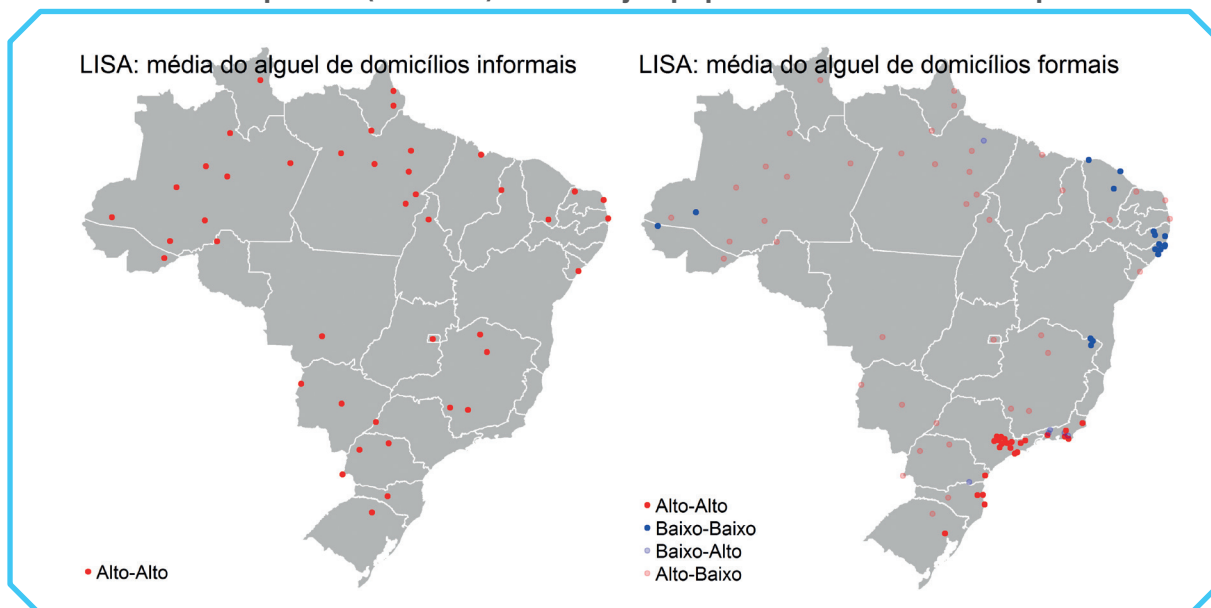


Fonte: Censo 2010.

Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

**FIGURA 2**

Mapa dos *clusters* Lisa dos valores dos aluguéis médios em AGSN (à esquerda) e setores não especiais (à direita): 70 arranjos populacionais e 106 municípios isolados



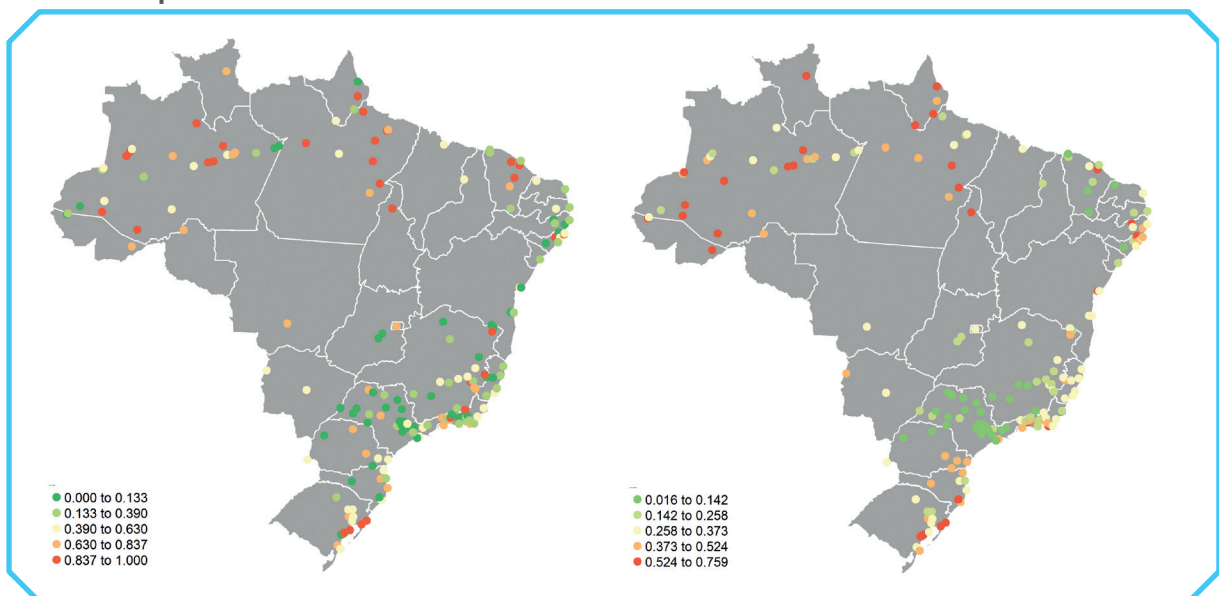
Fonte: Censo 2010.

Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Já para a distribuição das diferenças nas médias das amenidades urbanas entre os 176 núcleos urbanos analisados, o padrão espacial Sul/Sudeste com boas amenidades e Norte/Nordeste com piores amenidades é encontrado tanto para o mercado formal como para o mercado informal e confirmado para ambos os mercados pela análise do indicador Lisa. A figura 3 mostra os dois mapas da distribuição dos valores das médias de domicílios sem calçadas, e indica exatamente esse padrão de valores piorando com menores latitudes.

### FIGURA 3

Mapa coroplético da distribuição das proporções de domicílios em ruas sem calçada em AGSN (à esquerda) e não especiais (à direita): 70 arranjos populacionais e 106 municípios isolados



Fonte: Censo 2010.

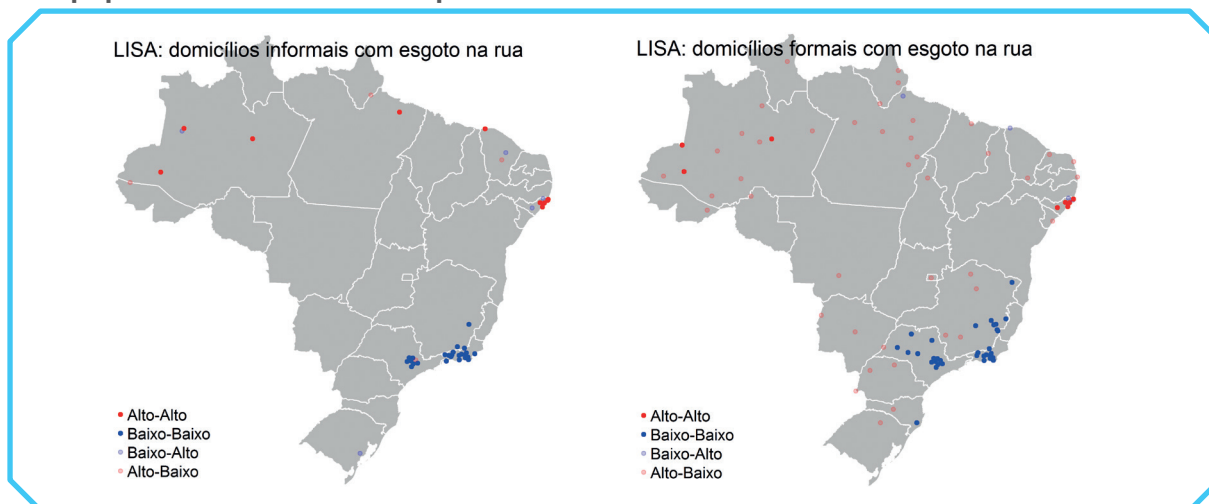
Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

A figura 4 mostra esses *clusters* para domicílios com esgoto a céu aberto, apresentando *clusters* de valores baixos (com melhor amenidade urbana) nos estados de São Paulo, Rio de Janeiro e Minas Gerais, e concentração de valores altos no Norte e no Nordeste.

Além disso, a figura 5 mostra um padrão espacial semelhante para *clusters* da variável de disponibilidade de esgotamento sanitário adequado, com a diferença que, para os domicílios formais, há *clusters* de valores baixos (melhor amenidade) no estado do Rio Grande do Sul. Note-se a ausência de associação espacial significativa para os domicílios em AGSN em toda a região Sul.

**FIGURA 4**

Mapa dos *clusters* Lisa das proporções médias de domicílios em ruas com esgoto a céu aberto em AGSN (à esquerda) e setores não especiais (à direita): 70 arranjos populacionais e 106 municípios isolados

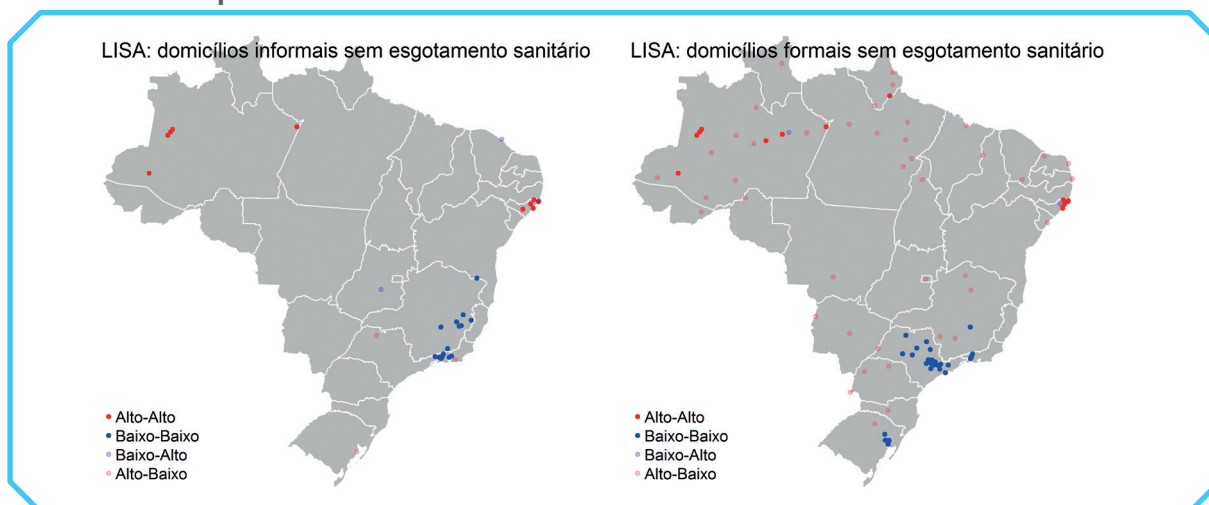


Fonte: Censo 2010.

Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

**FIGURA 5**

Mapa dos *clusters* Lisa das médias de domicílios sem esgotamento sanitário em AGSN (à esquerda) e setores não especiais (à direita): 70 arranjos populacionais e 106 municípios isolados



Fonte: Censo 2010.

Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

## 5 OS MERCADOS HABITACIONAIS DOS MUNICÍPIOS

Esta seção explora as informações para cada um dos 510 municípios da nossa amostra: os 106 municípios isolados, os 70 municípios-sede e os 334 municípios no “entorno” dos municípios-sede dos arranjos populacionais, conforme definição feita em IBGE (2010b). Analisaremos as diferenças na variação das características dos mercados habitacionais entre esses municípios. Apresentaremos os resultados da análise de regressão em gráficos de dispersão por grupos de intervalos (*binscatter regressions*).

A tabela 1 ilustra a distribuição dos domicílios particulares permanentes nas categorias de municípios da nossa amostra. Apesar de apenas 70 municípios serem sede, esses concentram 47% dos domicílios de todos os 510 municípios. Por seu turno, notamos como os municípios isolados são pequenos, pois os 106 municípios concentram somente 11% dos domicílios. A proporção de domicílios em AGSN varia de 7% a 11%. Os municípios-sede apresentam a maior proporção de AGSN, como era de se esperar, pois nesses municípios, no mercado formal, os preços são mais elevados devido à concentração de famílias de maior renda, assim como a concentração de melhores amenidades urbanas e oportunidades de emprego.

Assim como notado na diferença da média do conjunto dos municípios e dos treze maiores arranjos populacionais (seções 2 e 3 deste trabalho), as características das estruturas físicas dos domicílios variam bastante entre AGSN e domicílios em setores não especiais; e aqui notamos como variam também bastante entre os mercados dos municípios analisados. O gráfico 9 nos mostra essas distribuições a partir de estimativas de curvas de densidade de Kernel, que nada mais são que uma suavização do tradicional histograma.

**TABELA 1**

**Distribuição dos domicílios particulares permanentes nas categorias de municípios da amostra**

	Não especial		AGSN		Total	
	(A)	A/C (%)	(B)	B/C (%)	(C)	Total (%)
Entorno	21.645.540	93	1.654.420	7	<b>23.299.960</b>	<b>42</b>
Isolado	5.601.519	92	509.782	8	<b>6.111.301</b>	<b>11</b>
Sede	2.863.495	89	2.825.481	11	<b>25.688.976</b>	<b>47</b>
510 municípios	50.110.554	91	4.989.683	9	<b>55.100.237</b>	<b>100</b>

Fonte: Censo 2010.



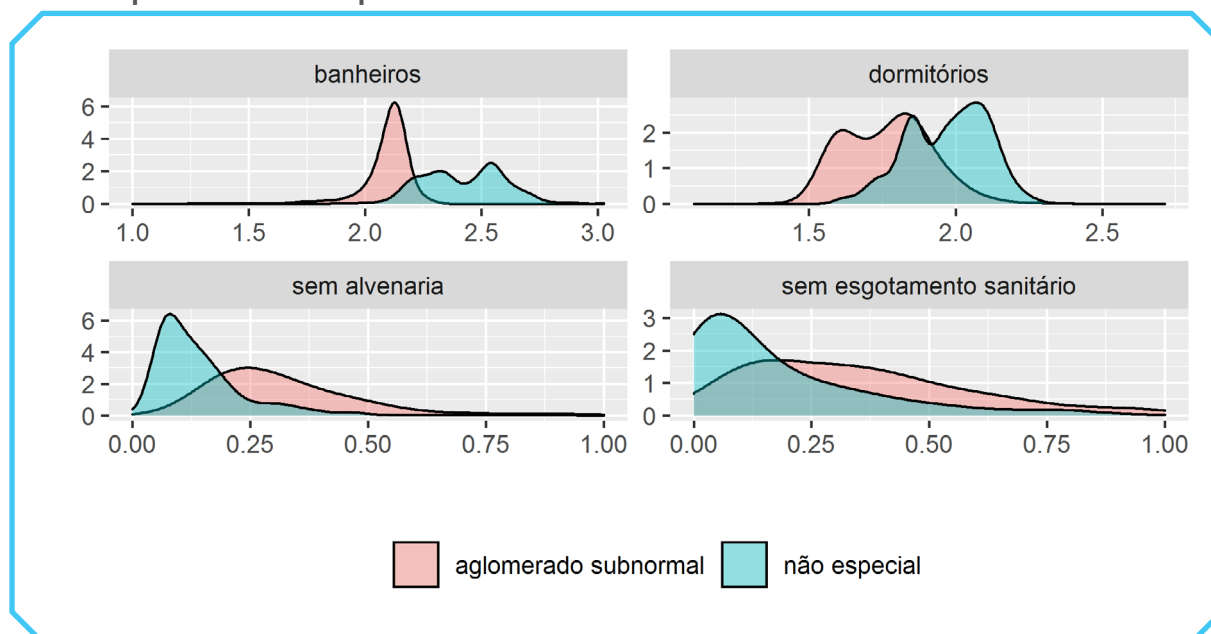
## TEXTO para DISCUSSÃO

Nota-se que, entre os municípios, os mercados formais apresentam maior qualidade e menor variabilidade na ausência de alvenaria e de esgotamento sanitário. Essas características, que indicam um nível de qualidade bastante baixo, são bem pouco frequentes nos mercados formais, enquanto há mercados informais melhores e piores nesses quesitos. Já a distribuição do número médio de dormitórios parece ter variabilidade semelhante, sendo, em geral, 0,25 dormitórios por domicílio maior nos domicílios não especiais. A distribuição de número médio de banheiros dos AGSNs fica bem concentrada ao redor de 2,1 banheiros, enquanto as médias dos domicílios não especiais se concentram entre 2,2 e 2,7 banheiros por domicílio.

A qualidade dos domicílios em termos de número de dormitórios e banheiros varia mais entre os domicílios em setores não especiais, o que é condizente com a ideia de que o mercado formal abrange tanto os segmentos de menor renda, quanto de maior renda da população.

### GRÁFICO 9

**Densidade de Kernel da distribuição de médias dos municípios de número de banheiros, dormitórios, domicílios sem alvenaria e sem esgotamento sanitário adequados nos municípios da amostra**



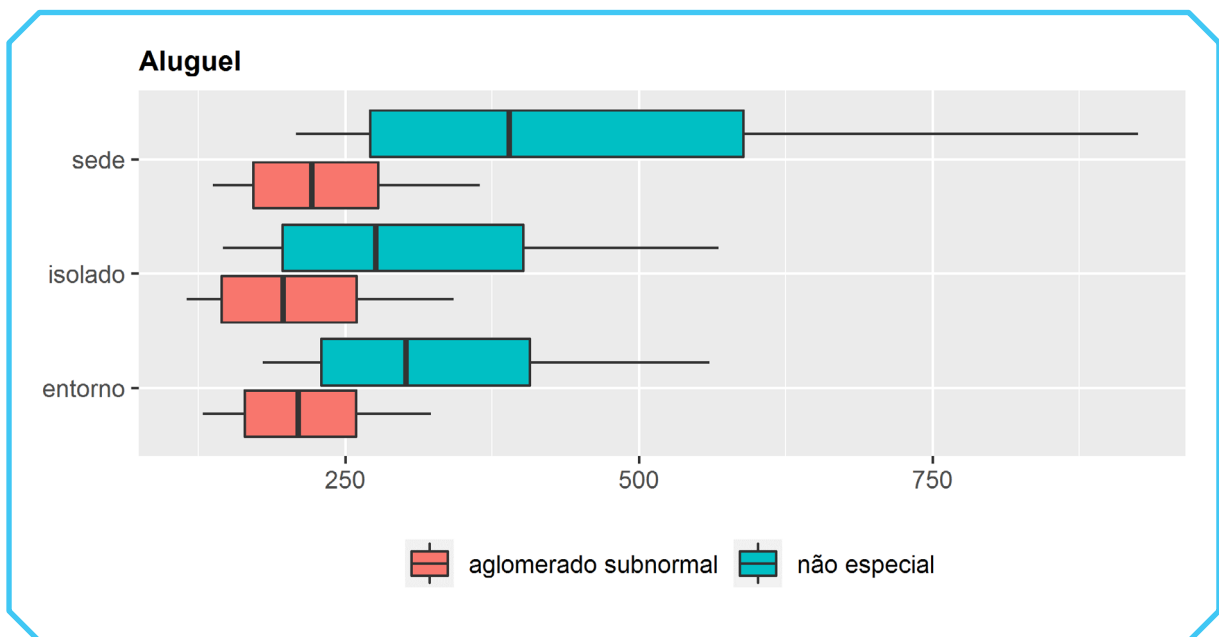
Fonte: Censo 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Já o gráfico 10 introduz uma diferenciação importante entre os mercados dos municípios-sede, do entorno e isolados, a média dos aluguéis. Nota-se como os aluguéis dos municípios-sede são, em geral, bem mais frequentes que os dos demais tipos de municípios, e concentram os valores mais altos. É de se esperar, pois esses municípios são os que apresentam melhores amenidades urbanas, melhor qualidade urbanística e mais oportunidades de emprego, educação e cultura. Nota-se, também, como os valores dos aluguéis de municípios do entorno e dos municípios isolados não são muito pronunciadamente diferentes.

### GRÁFICO 10

**Boxplot da distribuição dos valores de aluguéis nos municípios da amostra: sede, entorno e isolados**



Fonte: Censo 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Com relação aos aluguéis dos AGSNs, notamos como a diferença de valores entre não especial e AGSNs é muito mais ampla para os municípios-sede que para os demais municípios. E como a distribuição dos valores desses aluguéis em AGSN não parece se diferenciar muito entre municípios-sede, isolados e do entorno (gráfico 10).

TABELA 2

## Regressão de comparação de aluguéis dos municípios isolados

	Variável dependente			
	Mediana não especial	Mediana AGSN	Média não especial	Média AGSN
Municípios-sede	105,307*** (74,616, 135,998)	38,980*** (17,325, 60,636)	142,770*** (103,797, 181,743)	45,355*** (23,953, 66,757)
Municípios do entorno	23,313 (-3,889, 50,515)	24,877** (5,716, 44,038)	25,835 (-8,708, 60,377)	23,019** (4,082, 41,956)
Constante	251,889*** (229,519, 274,259)	175,333*** (159,549, 191,117)	292,438*** (264,032, 320,844)	190,137*** (174,538, 205,736)
Observações	190	191	190	191
R <sup>2</sup>	0,169	0,046	0,195	0,061
Estatística F	18,965*** (df = 2; 187)	4,483** (df = 2; 188)	22,580*** (df = 2; 187)	6,082*** (df = 2; 188)

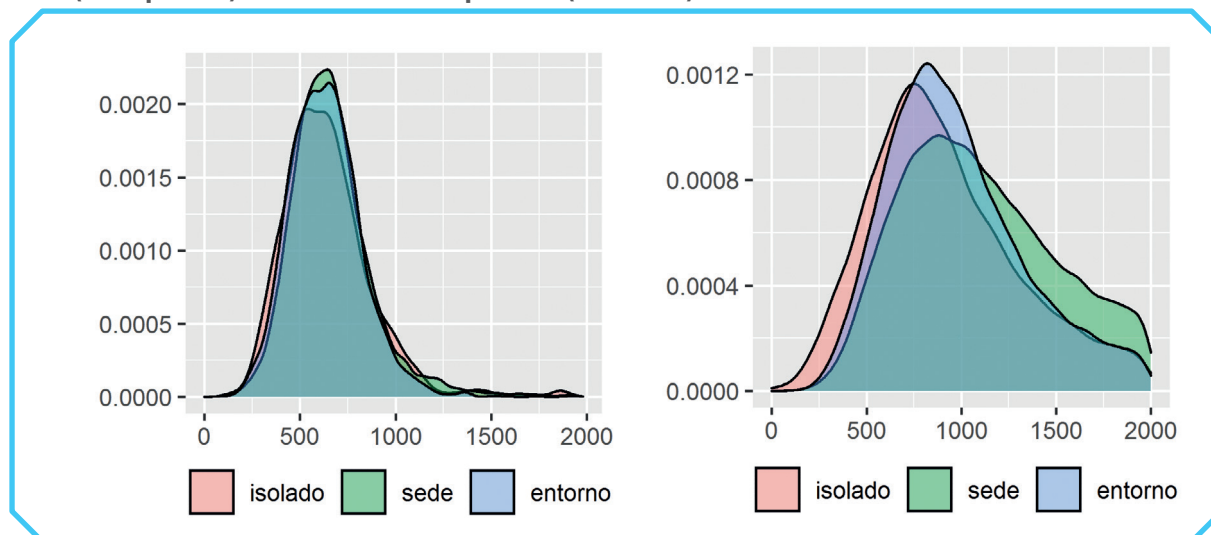
Fonte: Censo 2010.

Obs.: \*p < 0,1; \*\*p < 0,05; \*\*\*p < 0,01.

A tabela 2 mostra o resultado de regressões lineares simples de comparação das medianas e das médias dos aluguéis formais e informais. Esses valores são regredidos contra as variáveis binárias *dummies* indicando a condição de município-sede ou do entorno, tomando como base os municípios isolados. De fato, as diferenças entre municípios isolados e do entorno no mercado formal não são estatisticamente significantes, mas são significantes entre municípios isolados e municípios-sede, e em um montante grande, de R\$ 105,00 e R\$ 142,00. Já no mercado informal, todas as diferenças são estatisticamente diferentes, porém a significância da diferença é muito fraca, e os montantes das diferenças muito pequenas, entre R\$ 23,00 e R\$ 45,00.

**GRÁFICO 11**

Densidade de Kernel da renda média dos chefes dos domicílios dos setores censitários dos municípios da amostra, por categoria de município: setores em AGSN (à esquerda) e setores não especiais (à direita)



Fonte: Censo 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Uma das variáveis que está mais envolvida na formação dos aluguéis é a renda das famílias (Roback, 1982). Seriam esses rendimentos também mais homogêneos nos AGSN? Dispomos de informação dos rendimentos médios dos chefes dos domicílios por setor censitário. O gráfico 11 mostra a distribuição dessa renda nos 510 municípios de análise, densidade de Kernel, ponderada por número de domicílios em cada setor censitário e separada por setores que pertencem aos municípios-sede, do entorno e isolados. O gráfico é bem claro em ilustrar como também essa variável é mais homogênea nos setores subnormais, e como nos setores não especiais os chefes dos domicílios dos municípios-sede apresentam, em geral, renda superior à dos municípios do entorno e isolados.

Esse resultado já havia sido encontrado na análise dos padrões regionais dos aluguéis dos 70 arranjos populacionais e dos 106 municípios isolados. Ou seja, há uma homogeneidade regional no valor do aluguel em AGSN, assim como há uma homogeneidade em termos de categoria do município: se sede, isolado ou entorno.

Em seguida, passamos a comparar a distribuição somente dos municípios-sede e do entorno com relação às variáveis sensíveis às diferenças na estrutura urbana das cidades. Nem é necessário falar da distância ao centro das cidades, pois por definição esta é menor nos municípios-sede. Mas, com relação à posição relativa dos domicílios e dos empregos formais presentes no núcleo urbano, como era de se esperar, os domicílios do município-sede estão próximos a mais empregos que os do entorno.

## TEXTO para DISCUSSÃO

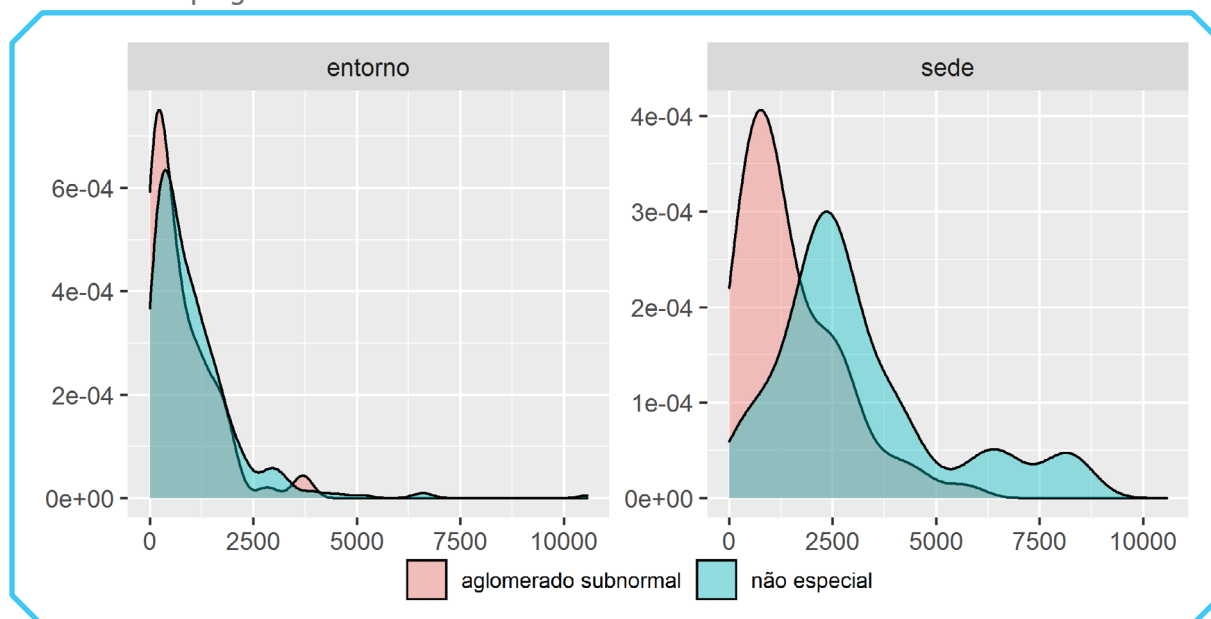
Outra característica importante, que nos mostra o gráfico 12A (empregos dentro de um *buffer* de 500 m), é que nos municípios do entorno praticamente não há diferenças entre AGSN e setores não especiais na quantidade de empregos próximos aos domicílios. Já nos municípios-sede, há menos variabilidade da quantidade de empregos próximos e menor quantidade para os domicílios em AGSN. Nas sedes, os domicílios em AGSN, em geral, estão próximos a 1 mil empregos menos que nos domicílios em setores não especiais.

Como a inserção urbana dos domicílios e sua posição relativa aos empregos impacta diretamente o tempo de deslocamento casa-trabalho, o gráfico 12B (tempo de deslocamento) também mostra uma situação semelhante à do gráfico 12A: o tempo de deslocamento é maior, em geral, no entorno, e é pior para AGSNs. No entanto, essa informação mostra uma diferença entre AGSN e não especiais nos municípios do entorno que na proximidade dos empregos não havia. Parece que há poucos empregos disponíveis no entorno, então, para o *buffer* utilizado de 500 m, há poucos empregos próximos tanto de AGSN como de não especiais. A diferença do tempo de deslocamento, no entanto, reforça que os AGSNs ocupam bairros mais distantes, ou com menos acesso a infraestrutura de mobilidade urbana.

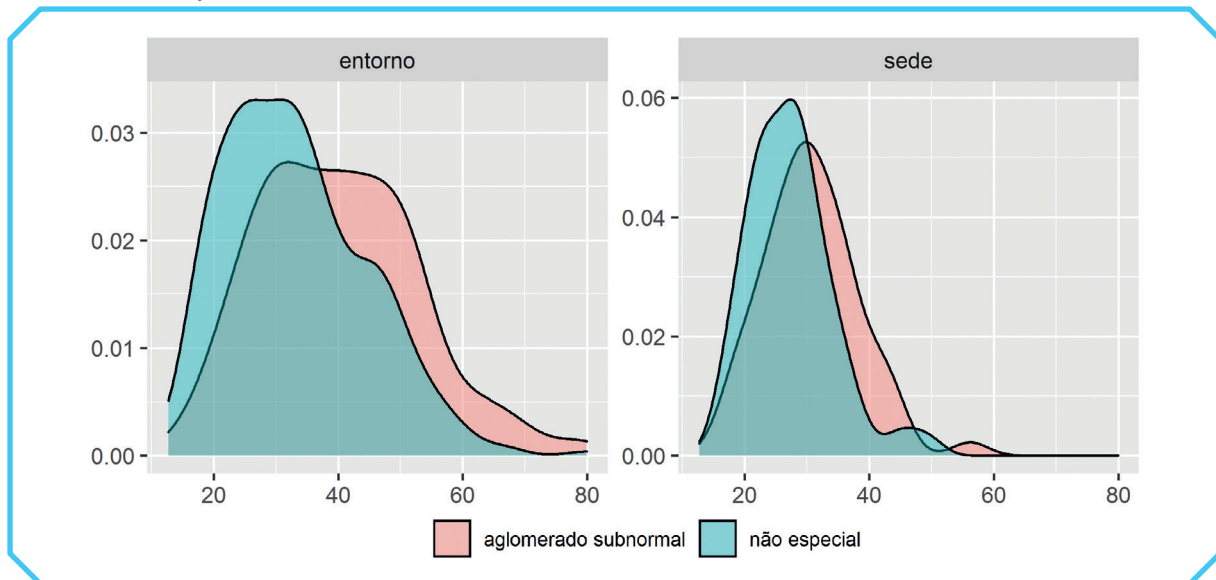
### GRÁFICO 12

#### Proximidade a empregos e tempo de deslocamento: municípios-sede e do entorno da amostra

12A – Empregos dentro de *buffer* de 500 m do setor censitário



12B – Tempo de deslocamento



Fonte: Censo 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

### 5.1 Binned scatterplot

Nesta subseção apresentaremos os resultados da análise quantitativa que indica uma correlação positiva entre o valor médio do aluguel formal e a diferença de médias do aluguel formal e informal em 190 municípios da nossa análise, para os quais há AGSN e há informação suficiente de aluguel informal. O gráfico 4 traz uma dispersão simples, ilustrando a correlação para os municípios-sede, do entorno e isolados.

Esse achado é relevante, pois indica algum mecanismo de ligação entre ambos mercados que deve ser melhor investigado. O gráfico 13 apresenta em amarelo as médias ressaltadas para grupos de intervalos e uma reta que corresponde a uma regressão linear simples da diferença de médias dos aluguéis formal e informal na média do aluguel formal. Esta regressão simples pode ser entendida como uma regressão de preços hedônicos, não tanto como em Rosen (1974), mas sim como em Roback (1982), pois cada unidade da regressão é um município, um mercado (ou submercado) distinto, e não estão presentes na análise as características da moradia que compõem os modelos hedônicos. No entanto, essa correlação simples não teria interpretação causal devido à explicação já conhecida do viés de variável omitida, como lembrado por Greenstone (2017). No caso, uma correlação entre as características não mensuradas dos municípios e a diferença dos aluguéis, por exemplo, poderia estar “mascarando” o efeito das características do mercado informal no formal.

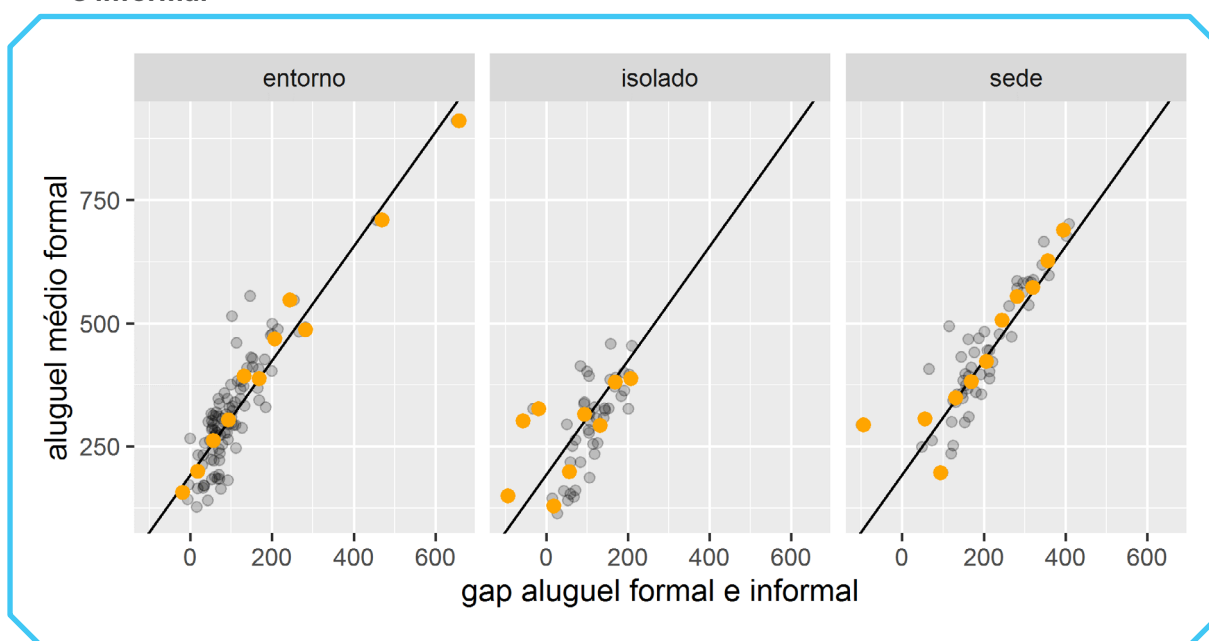


## TEXTO para DISCUSSÃO

Os principais determinantes das médias dos aluguéis formais que estão de fora da regressão simples do gráfico 13, e que poderiam estar correlacionados com o *gap* do mercado formal e informal, são três: as características das estruturas físicas formais; as amenidades ambientais dos bairros formais; e as diferenças entre os municípios de renda e atividade econômica – os salários do modelo de Roback (1982). Mas seria necessário aprofundar o conhecimento das possíveis relações entre os mercados para determinar como seriam as correlações entre esses determinantes do mercado formal e mercado informal.

### GRÁFICO 13

Dispersão do aluguel médio formal e diferença de médias do mercado formal e informal



Fonte: Censo 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Os mercados formal e informal são mercados substitutos em alguma medida, mas de que maneira? Há modelos econômicos teóricos que nos indicam que ambos mercados disputam os mesmos tipos de terra e que o mercado informal diminui a oferta de terra a ser ocupada pelo mercado formal, em um contexto de oferta fixa de terra (Brueckner e Selod, 2009) ou oferta flexível de terra (Selod e Tobin, 2018). Smolka e Biderman (2011) descrevem para a América Latina que esse seria o caso, pois, seria mais rentável para o loteador não seguir os padrões urbanísticos formais. Mesmo assim, há ocupação de áreas que não seriam ocupáveis pelo mercado formal, dando alguma “barreira de entrada” ao mercado formal, por exemplo, as áreas com riscos de desastres ocupadas por favelas. Diferentes custos de formalização (Cavalcanti, Da Mata e Santos, 2019),

diferentes riscos de remoção, diferentes possibilidades de acesso a serviços urbanos de saneamento básico e segurança pública e diferentes possibilidades de inserção urbana influenciam a relação entre os mercados formal e informal, certamente.

A técnica utilizada para investigar a robustez da correlação do gráfico 14 é a da dispersão em grupos de intervalos (*binned scatterplots*); uma técnica de visualização da correlação entre duas variáveis que aprimora o gráfico 14. Os grupos de intervalos (*bins*) não são tão essenciais no nosso caso, uma vez que não dispomos de uma quantidade de observações tão grande que faça a nuvem de pontos se tornar ilegível. O que é mais essencial, a técnica permite a inserção de variáveis de controle.

Além disso, a técnica permite analisar a função de esperança condicional de uma variável com relação à outra, avaliando se a relação é de fato linear, como pressuposto em uma regressão linear (Cattaneo *et al.* 2021). Além de ser uma técnica de suavização (não paramétrica), ela permite o cálculo de bandas de confiança, intervalos de confiança para uma função. Trata-se de uma função, pois, a suavização permite flexibilidade para que a relação entre as duas variáveis não seja linear, e, sim, seja uma “função”. O pacote utilizado é o *binsreg* no *software R*.

Infelizmente, não há estudos ou modelos que tratem as relações de causalidades anteriormente descritas, envolvendo as duas variáveis de interesse, de maneira a construir um grafo acíclico direcionado de causalidade (Pearl, 1995) que nos desse embasamento teórico para contornar assertivamente o problema de viés de variável omitida, indicando quais variáveis de controle são “bons controles”, ou seja, amenizam o viés de variáveis omitidas mas não trazem viés de variável confundidora, por exemplo (Pearl, 1995).

No entanto, tendo em vista a discussão anteriormente feita e a literatura citada, sabemos que as diferenças na qualidade das unidades habitacionais e as diferenças de renda explicam os preços/aluguéis formais. Podemos argumentar que a qualidade das unidades formais médias parece estar descolada do funcionamento do mercado informal, visto que somente uma parte do mercado formal é substituída pelo mercado informal, dada a grande desigualdade de renda brasileira. Já o padrão regional de desigualdade de renda não parece explicar as características do mercado informal, um resultado da nossa análise com o indicador Lisa feita na seção 4 deste trabalho. Isto nos encoraja a incluir variáveis dicotômicas, indicando pertencimento às grandes regiões como controle no *binscatter*, para levar também em consideração o papel do desenvolvimento econômico, ou as diferenças em salários, como no modelo de Roback (1982). O controle pela condição de municípios-sede ou do entorno também capta a diferenciação de renda intrametropolitana.

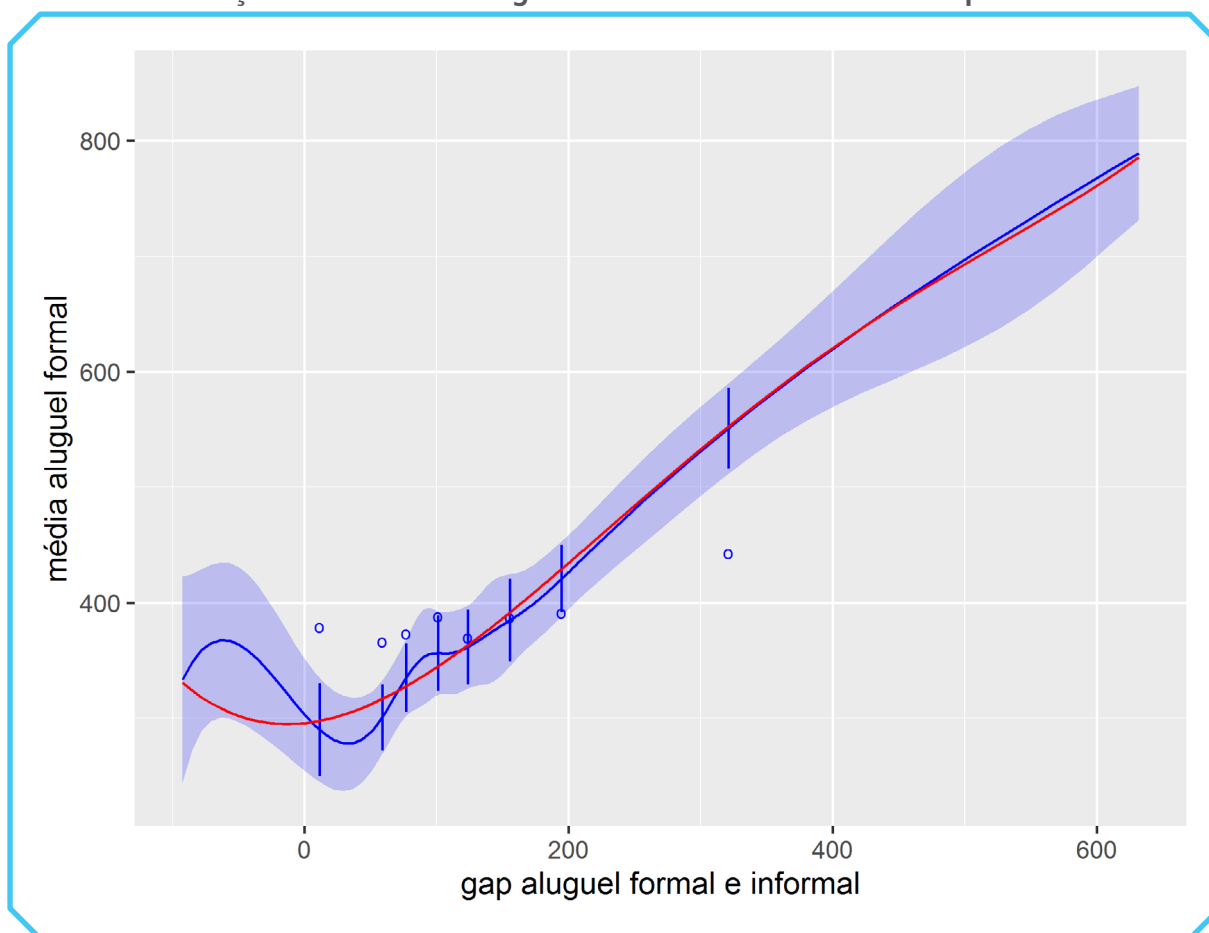
O gráfico 14 traz o resultado. No nosso caso, apesar da flexibilidade em assumir qualquer forma funcional, a relação entre as duas variáveis – a média do aluguel formal e a diferença das médias

## TEXTO para DISCUSSÃO

do aluguel formal e informal – parece de fato linear e aparece mesmo após a inserção das variáveis de controle. Esta “função” de correlação foi calculada levando em consideração como controles: a média do número de banheiros dos domicílios formais; a qual grande região o município pertencia; e se se encontrava na sede, no entorno ou se era município isolado. A regressão linear simples que corresponde a essa correlação linear é mostrada na tabela 3, indicando que uma diferença de R\$ 100,00 entre o aluguel formal e informal está associada a um aluguel médio formal R\$ 79,00 maior.

### GRÁFICO 14

Dispersão em grupos de intervalos (*binned scatterplots*) da média do aluguel formal e da diferença das médias do aluguel formal e informal nos municípios da amostra



Fonte: Censo 2010.

Obs.: 1. Variáveis de controle: média do número de banheiros dos domicílios formais, a qual grande região pertencia e se o município se encontrava na sede, no entorno ou se era município isolado. A linha vermelha corresponde a um modelo global de regressão polinomial cúbica de grau 3. A linha azul corresponde à regressão polinomial nos grupos de intervalos (*bins*). Já os pontos correspondem ao gráfico de dispersão em *bins* tradicional. A mancha azul é a banda de confiança.

2. Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

**TABELA 3****Resultados da regressão linear da média do aluguel formal, com covariáveis**

	Variável dependente
	Média do aluguel formal
Diferença das médias do aluguel formal e informal	0.79*** (0.68, 0.89)
Número médio de banheiros dos domicílios formais especiais	162.18*** (101.46, 222.91)
Nordeste	-103.01*** (-125.88, -80.15)
Sudeste	-21.40 (-43.91, 1.11)
Centro-Oeste	5.07 (-21.18, 31.33)
Sul	-20.73 (-62.63, 21.17)
Municípios-sede <sup>1</sup>	31.66*** (12.71, 50.61)
Municípios do entorno <sup>1</sup>	29.95*** (14.88, 45.01)
Constante	-123.43 (-246.80, -0.06)
Observações	190
R <sup>2</sup>	0.87
R <sup>2</sup> Ajustado	0.86
Estatística F	148.99*** (df = 8; 181)

Fonte: Censo 2010.

Nota:<sup>1</sup> Em comparação aos municípios isolados.

Obs.: \* p &lt; 0.1; \*\* p &lt; 0.05; \*\*\* p &lt; 0.01.

## 6 CLUSTERS DE SETORES CENSITÁRIOS NO RIO DE JANEIRO E EM SÃO PAULO: OS AGSNS SÃO UM CLUSTER SEPARADO?

Esta seção tem como objetivo investigar quais seriam os segmentos do mercado habitacional nas duas maiores metrópoles brasileiras, Rio de Janeiro e São Paulo, agrupando os setores censitários de acordo com suas características. Seriam segmentos de *alta renda*, de *média renda* e de *baixa renda*? Tomando os AGSNS como uma mensuração da cidade "informal", seriam os AGSNS um segmento de mercado distinto do segmento de *baixa renda* ou *baixíssima renda* formal?

O mercado de habitação pode ser segmentado com base em preços semelhantes em um mesmo bairro, ou porção da cidade (Goodman e Thibodeau, 1998). Vamos partir do princípio oposto: checar como pequenas porções do território (os setores censitários) compartilham

## TEXTO para DISCUSSÃO

características semelhantes e observar se esses grupos de setores censitários semelhantes estão aglomerados em territórios suficientemente contíguos. O princípio é semelhante, porém o ponto de vista é o inverso, pois os preços das habitações “resumem” as características tanto dos imóveis quanto da vizinhança.

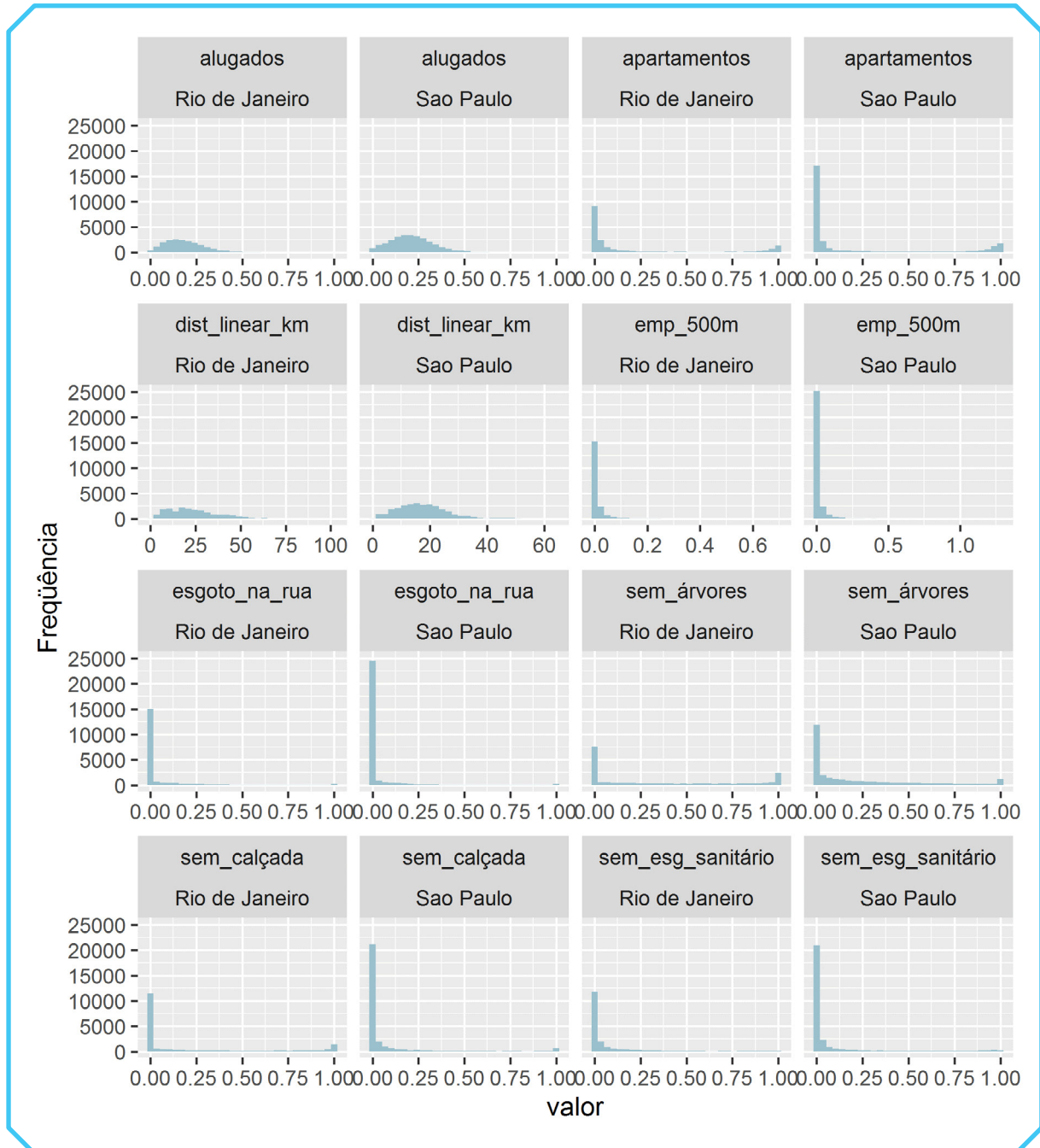
Uma técnica de análise exploratória de dados, os algoritmos de clusterização de aprendizado de máquina, é a ferramenta que propicia fazer uma partição de um conjunto de dados de maneira a obter casos mais semelhantes entre si dentro de um mesmo grupo e grupos suficientemente distintos entre si. À diferença dos demais algoritmos de aprendizado de máquina, na clusterização não se sabe de antemão a quais grupos cada observação pertence, pois o importante é “descobrir” quais são os grupos (Irizarry, 2019).

No nosso caso, as observações serão os setores censitários do Censo 2010, e os “previsores”, variáveis explicativas, ou características de cada grupo, serão as variáveis provenientes do questionário do universo do Censo 2010, enriquecidas com as variáveis que descrevem o entorno dos domicílios (IBGE, 2010b) e três variáveis externas ao Censo 2010, a distância do setor censitário ao ponto central da cidade: Sé, em São Paulo, e Praça Mauá, no Rio de Janeiro; a declividade do setor censitário e a quantidade de empregos formais dentro de um *buffer* de 500 m de cada setor censitário.<sup>6</sup>

Os valores observados passam por uma normalização para que variáveis de diferentes escalas possam ser comparadas. Esses valores são utilizados para calcular a distância euclidiana entre as observações, nossa medida de semelhança. Faremos uma breve análise de componentes principais para utilizar as características que “mais variam” entre os setores censitários. Com relação aos diferentes métodos, que utilizam diferentes passos para aglutinar as observações em grupos, testaremos duas técnicas de clusterização: k-means e *cluster* hierárquico (Irizarry, 2019).

Por fim, vale a pena visualizar o formato das distribuições dos previsores utilizados na separação dos *clusters*. Quase todas as variáveis são proporções de domicílios com determinada característica com relação ao total de domicílios de cada setor censitário. No gráfico 15, somente a distância ao centro não é uma porcentagem. Nota-se muita polarização em muitas das variáveis, ou pouca variação. A polarização parece ser maior em São Paulo que no Rio de Janeiro.

6. As informações de empregos formais geocodificados estão descritas em Pereira *et al.* (2020), a distância ao centro foi calculada pela autora utilizando os centroides dos setores censitários como referência, calculando a distância euclidiana.

**GRÁFICO 15****Histograma das variáveis descritoras selecionadas**

Fonte: IBGE (2010b).

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

## 6.1 Análise de componentes principais

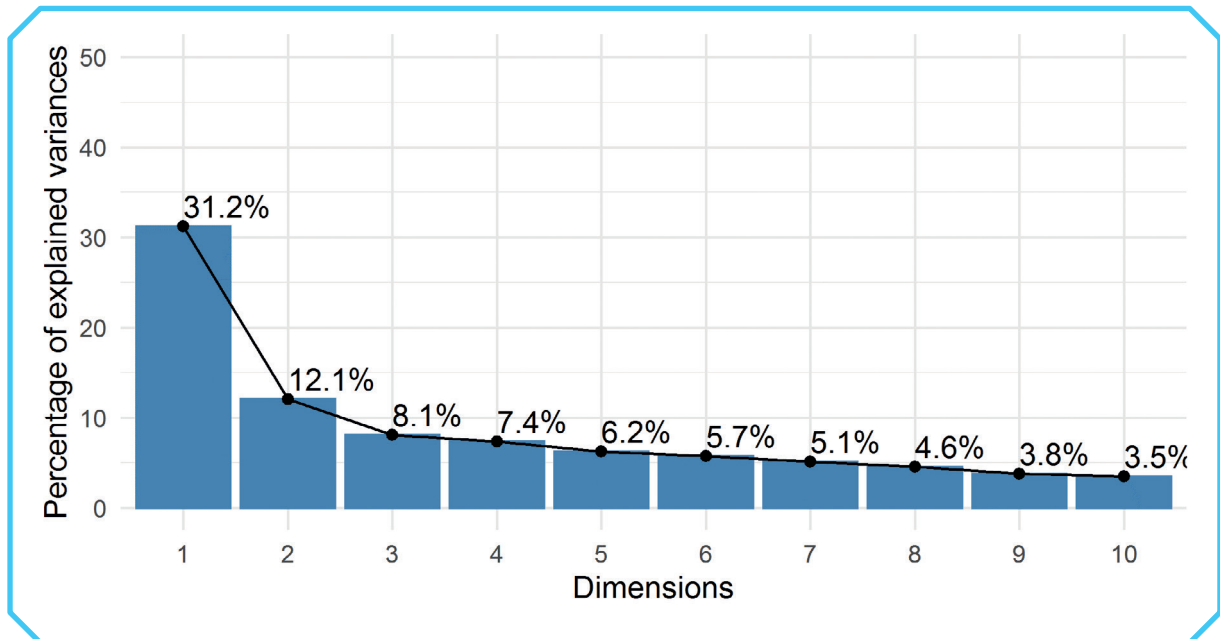
Partimos de uma seleção de dezesseis variáveis que descrevem os setores censitários: sete do entorno dos domicílios; seis do questionário do universo; e três externas (distância ao centro, declividade e proximidade a empregos formais). Algumas dessas variáveis contêm informações redundantes entre si, ou seja, acrescentá-las na análise não confere mais poder explicativo, somente adiciona mais ruído. Portanto, vale a pena investigar essas redundâncias entre os “previsores”, para conseguir melhor performance na análise de *cluster*. Quanto mais ruído entre os previsores, maior a tendência de existir um grupo que concentra quase todas as observações.

Na linguagem do *software R*, utilizamos os pacotes *factoextra* e *FactoMineR* para fazer as análises aqui descritas. Os dois painéis do gráfico 16 mostram a porcentagem da variância explicada de cada componente principal encontrado. Os componentes principais são combinações das dezesseis variáveis. Em São Paulo, os cinco primeiros componentes explicam 58,7% da variância; enquanto, no Rio de Janeiro, os cinco primeiros componentes principais explicam 65% da variância.

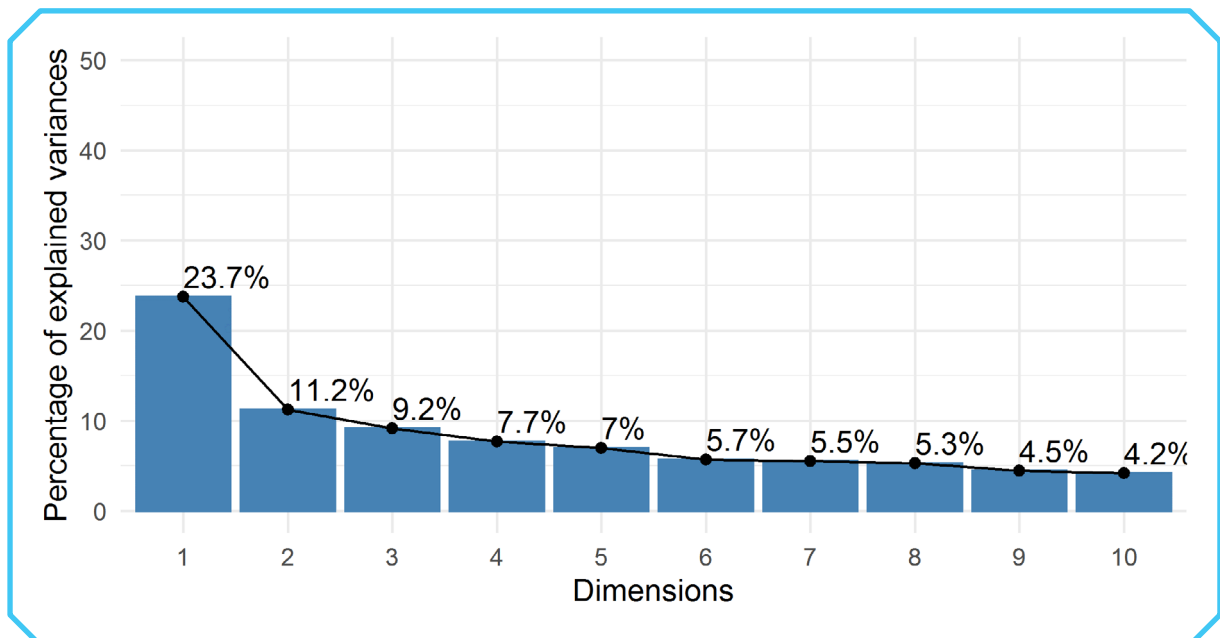


**GRÁFICO 16***Scree plot: variâncias explicadas a cada componente principal adicionado*

16A – Rio de Janeiro



16B – São Paulo



Fonte: IBGE (2010b).

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

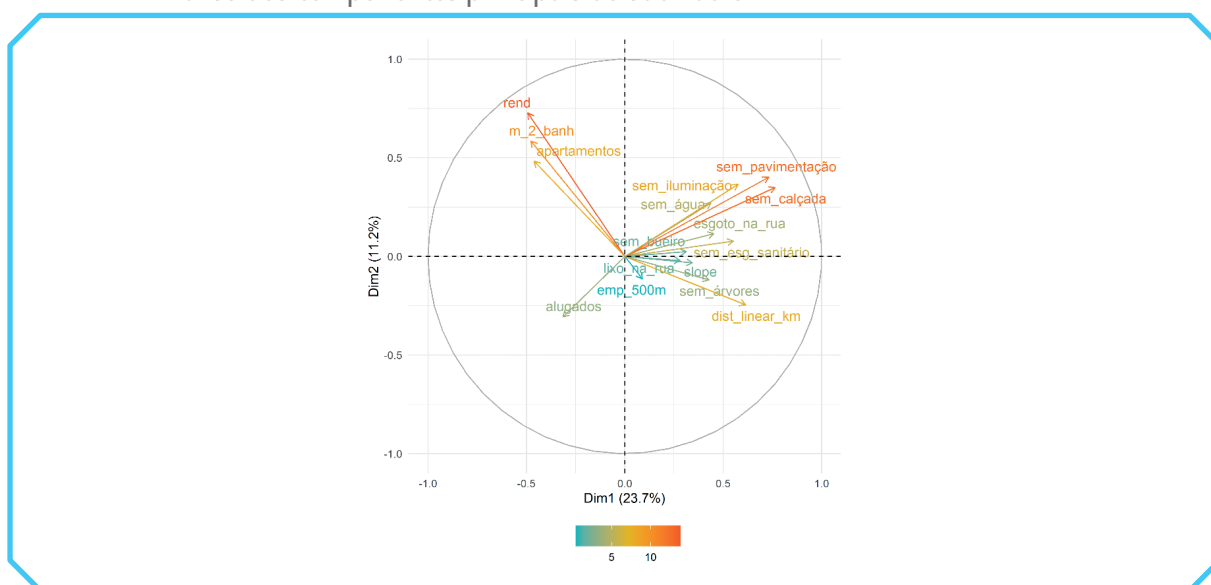
## TEXTO para DISCUSSÃO

Ao analisar a contribuição de cada uma das dezesseis variáveis para os dois primeiros componentes principais, é possível visualizar quais variáveis dão a mesma contribuição para estes componentes (flechas muito próximas entre si) sendo, portanto, redundantes, e quais contribuem de forma diferenciada. O gráfico 17 traz essas visualizações para São Paulo e Rio de Janeiro.

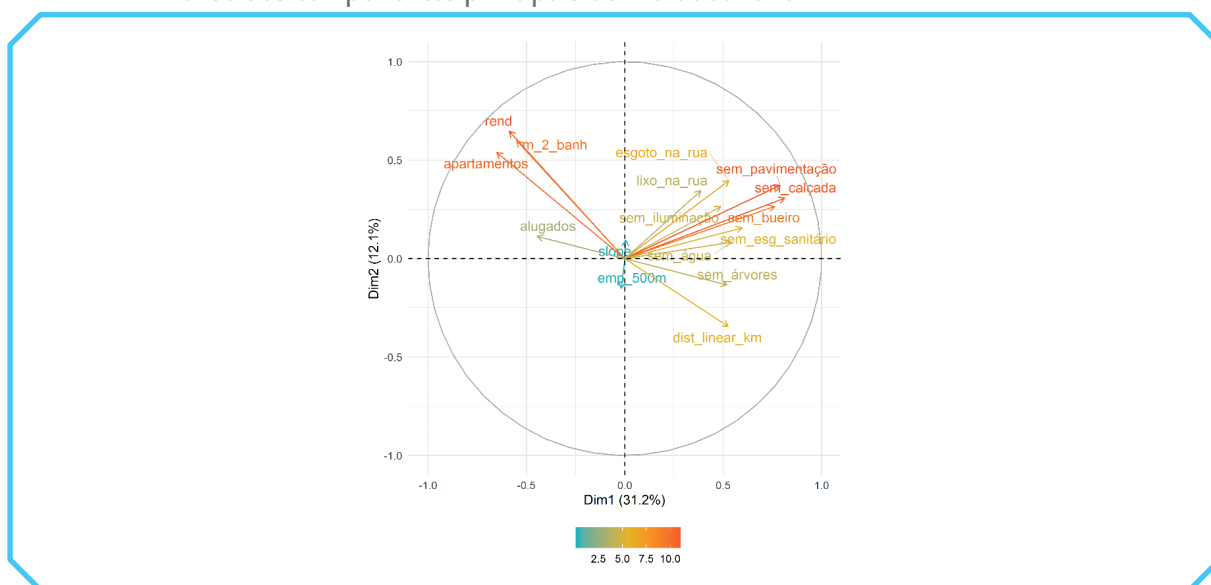
### GRÁFICO 17

#### Variáveis que contribuem para os dois primeiros componentes principais

##### 17A – Análise dos componentes principais de São Paulo



##### 17B – Análise dos componentes principais do Rio de Janeiro



Fonte: IBGE (2010b).

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Em congruência com essas análises descritivas, e com as principais variáveis que contribuem para os primeiros componentes principais, escolhemos como “previsores” para serem utilizados na análise de *cluster* os conjuntos de variáveis a seguir descritas.

- 1) São Paulo: *emp\_500m*; *dist\_linear\_km*; *sem árvores*; *lixo na rua*; *sem calçada*; *mais de dois banheiros*; e *alugados*.
- 2) Rio de Janeiro: *apartamentos*; *emp\_500m*; *dist\_linear\_km*; *sem árvores*; *sem calçada*; e *esgoto\_na\_ua*.

## 6.2 Resultado da análise de *clusters* e descrição dos *clusters* encontrados

Se as informações dos setores censitários fossem totalmente aleatórias, não seria possível identificar *clusters*, ou subgrupos, com características distintas. Então, para investigar se de fato há uma estrutura que diferencia os setores censitários, foram feitos testes que tentam descobrir qual seria o número ótimo de *clusters* presente nos dados, o teste do cotovelo (*elbow*) e da silhueta (*silhouette*). Eles se diferenciam entre si na maneira de medir o quanto os *clusters* são homogêneos internamente e heterogêneos entre os grupos.<sup>7</sup> Estas análises foram feitas utilizando as variáveis escolhidas na seção 6.1. Para São Paulo e Rio de Janeiro, valida-se a existência de sete *clusters*, os gráficos são mostrados nos apêndices. Mesmo assim, escolhemos separar São Paulo e Rio de Janeiro em seis *clusters*, já que não há perda significativa de poder explicativo.

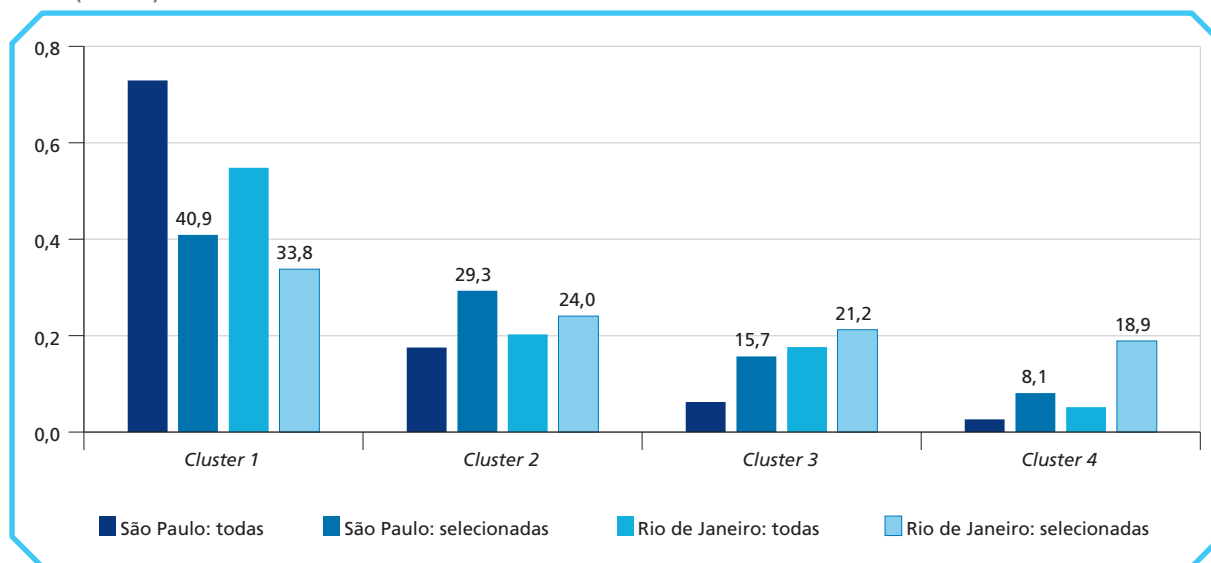
A redução da dimensionalidade de fato serviu para tirar ruído das variáveis “previsoras”. Utilizando o método *k-means* para formar os *clusters* com todas as dezesseis variáveis, o gráfico 18 mostra a proporção de moradores nos quatro maiores *clusters* formados. Nota-se como os moradores se concentram demasiadamente em um *cluster* só no caso de São Paulo (73%). Aplicando o algoritmo com as variáveis escolhidas há mais diferenciação das observações, com 41% dos moradores no *cluster* maior. No Rio de Janeiro, o problema não seria tão grave (de 55% para 34%).

---

7. O teste de cotovelo investiga a porcentagem da variância explicada como uma função do número de *clusters* utilizados. A partir de certo ponto, adicionar mais um *cluster* vai adicionar bem menos porcentagem de variância explicada. Este é o ponto do “cotovelo” do gráfico. Já o teste da silhueta procura o número de *clusters* que maximiza o valor médio da silhueta. Esse valor corresponde a o quanto cada observação é similar ao seu *cluster*, e distante dos demais *clusters*.

**GRÁFICO 18**

Proporção de moradores em cada *cluster* utilizando todas as dezesseis variáveis disponíveis ou as variáveis selecionadas via análise de componentes principais (Em %)



Fonte: IBGE (2010b).

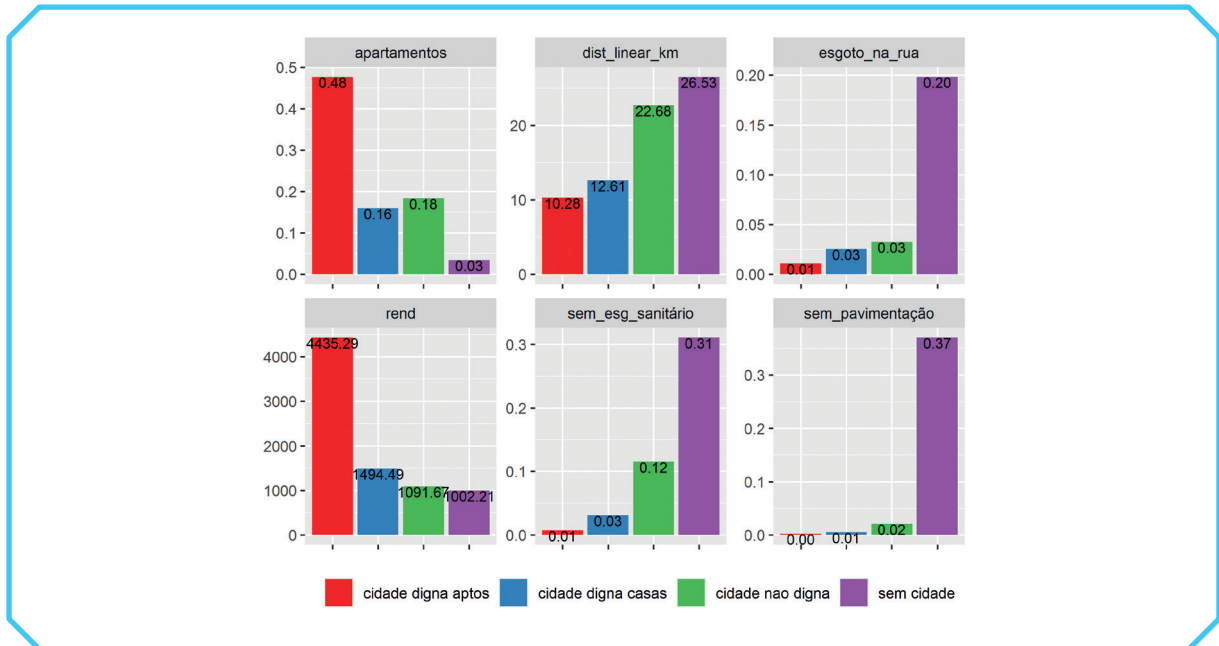
O gráfico 18 também nos indica que as proporções finais dos *clusters* são semelhantes entre as duas cidades. Em seguida, mostramos as diferenças entre as médias ponderadas pelo número de domicílios das variáveis "previsoras" para os diferentes *clusters*. As combinações de características nos levaram a escolher os nomes para os *clusters*. No caso de São Paulo, os dois menores *clusters* formados foram desconsiderados. No Rio de Janeiro, dois *clusters* de características muito próximas foram aglutinados e um desconsiderado, o de menor tamanho. Escolhemos os seguintes nomes: cidade digna apartamentos; cidade digna casas; cidade não digna; e sem cidade.

As médias ponderadas por total de domicílios dos previsores varia bastante entre esses *clusters*. De fato, esses nomes foram escolhidos por serem condizentes com as principais características de cada *cluster*. O gráfico 19 mostra essas diferenças para seis das variáveis consideradas, as demais são apresentadas no apêndice B.

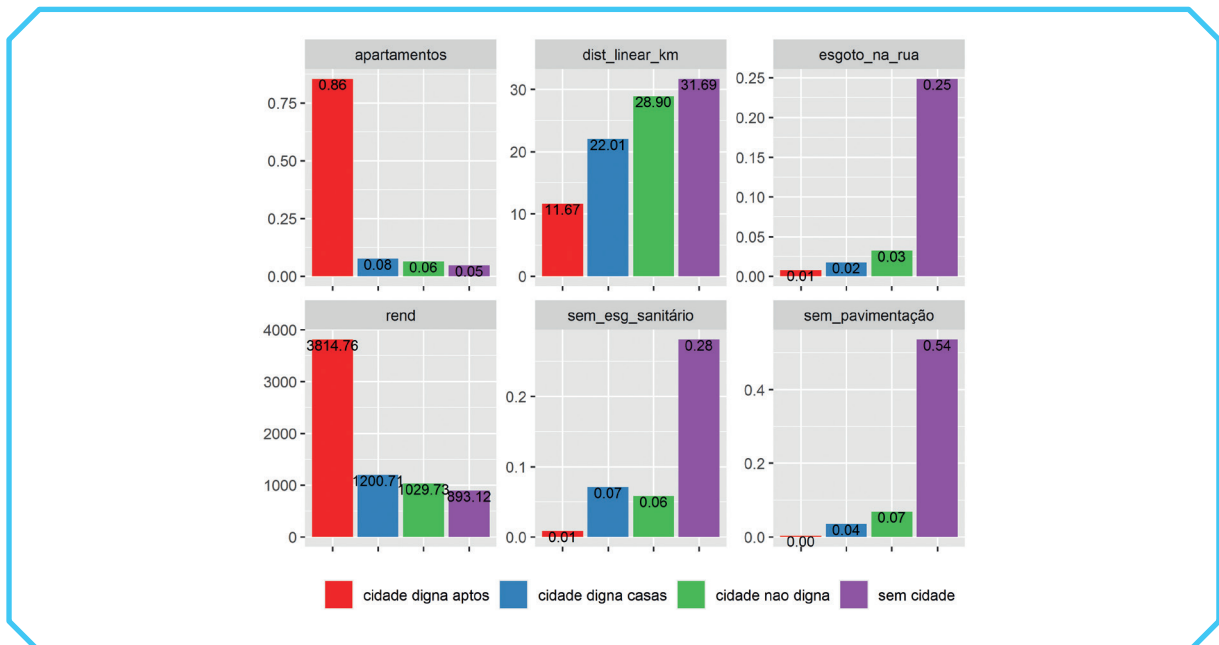
**GRÁFICO 19**Variáveis selecionadas: médias ponderadas por número de domicílios por *clusters*<sup>1</sup>

(Em %)

19A – São Paulo



19B – Rio de Janeiro



Fonte: IBGE (2010b).

Nota: <sup>1</sup> Distância em quilômetro; renda em R\$ de 2010.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Por exemplo, o *cluster cidade digna apartamentos* é composto, em sua maioria, por apartamentos (48% dos domicílios em São Paulo e 86% no Rio de Janeiro), apresenta renda muito maior que os demais *clusters*, e todos indicadores de infraestrutura urbana e serviços urbanos adequados. Em São Paulo, ele corresponde a 15,7% do total de moradores; no Rio de Janeiro, a 19%. Este *cluster* seria uma delimitação apropriada da cidade dos *ricos*, ou formal.

Já o *cluster cidade digna casas*, não é composto por apartamentos (16% em São Paulo e 8% no Rio de Janeiro). É o maior *cluster* no Rio de Janeiro (33% dos moradores); mas o segundo maior em São Paulo (29% dos moradores). Apresenta a segunda maior renda média do chefe do domicílio e indicadores de infraestrutura (pavimentação, calçada e iluminação) muito próximos ao *cluster* da cidade mais rica. Já os indicadores de serviços, como esgotamento, abastecimento de água, lixo na rua e esgoto na rua, podem ser parecidos com a cidade rica ou piores.

O nome *cidade não digna* implica um limite arbitrário de inadequação que este *cluster* teria ultrapassado. Lembrando que os valores limites das características de cada *cluster* são resultado do algoritmo de agrupamento. A *cidade digna de casas* apresenta indicadores de inadequação às vezes próximos à *cidade não digna*, mas este *cluster* é sempre pior que aquele. Certas características, como a falta de calçadas, pavimentação e iluminação públicas, são consideravelmente piores no Rio de Janeiro. A falta de esgotamento sanitário e abastecimento de água adequados são consideravelmente piores em São Paulo.

Por sua vez, o *cluster sem cidade* apresenta algumas características muito destoantes das demais, claramente concentrando inadequações. A falta de esgotamento sanitário e a falta de abastecimento de água em São Paulo são notórias, maiores que no Rio de Janeiro. A magnitude da inadequação do entorno do domicílio é maior no Rio de Janeiro: 78% de domicílios sem calçadas, 25% com esgoto na rua, 60% sem árvores; e em São Paulo: 73%, 20% e 55%, respectivamente. Em São Paulo, 8,8% da população está nesse *cluster*; enquanto, esta proporção é de 24% no Rio de Janeiro.

### 6.3 Localização intraurbana dos *clusters*

Nesta subseção apresentamos a localização dos *clusters*. Os mapas, nas figuras 6 e 7, mostram os centroides dos setores censitários diferenciados por cores que distinguem a que *cluster* pertencem. Como dito anteriormente, uma definição de submercado de habitação fala explicitamente em uma subdivisão do território de uma cidade, em que os preços sejam suficientemente semelhantes. No nosso caso, desconsideramos a estrutura espacial da cidade, e as relações de vizinhança física

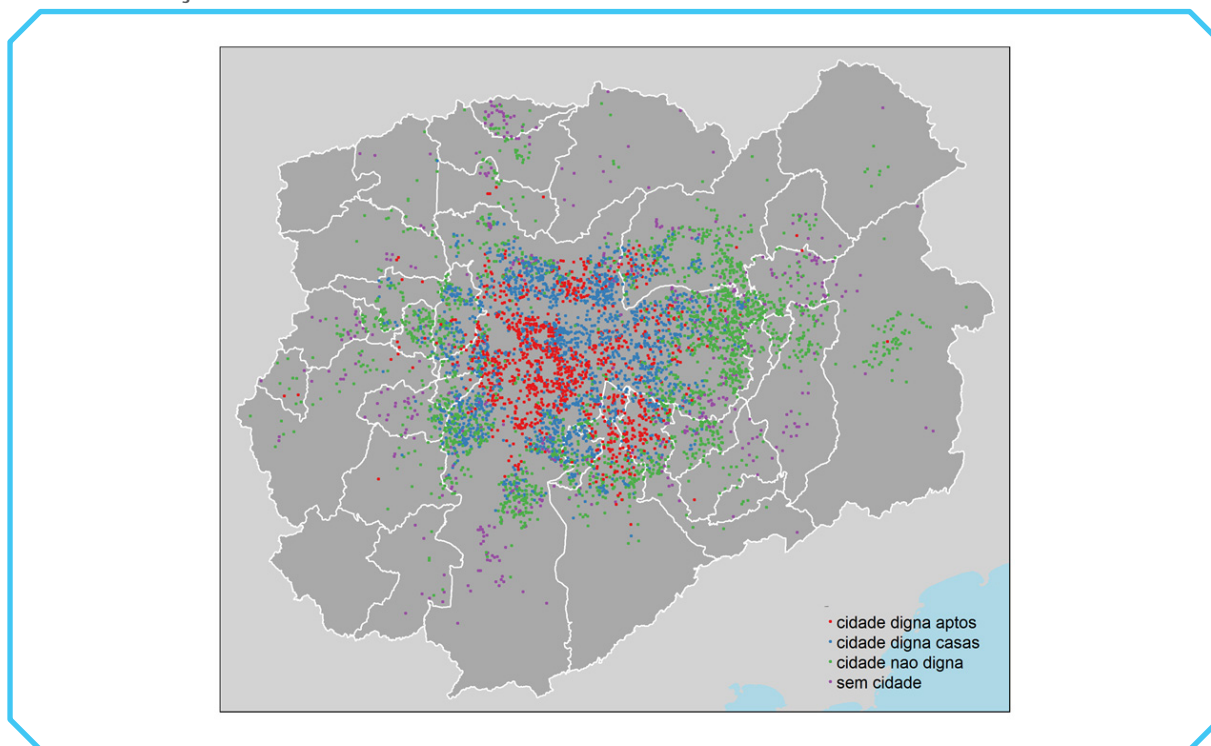
entre os setores censitários. As variáveis de distância ao centro e declividade média do terreno são as únicas que descrevem de alguma maneira as propriedades geográficas dos setores.

Mesmo assim, os *clusters* formados apresentam uma estrutura espacial visualmente identificável, conforme mostram as figuras 6 e 7, assim como a localização separada de cada *cluster* nos apêndices. Como era de se esperar, os setores censitários do *cluster cidade digna apartamentos* se concentram nas zonas centrais do Rio de Janeiro e de São Paulo, correspondendo aos bairros mais valorizados e de melhor infraestrutura e serviços urbanos. Já *cidade digna casas* ocupa os bairros centrais próximos aos da *cidade digna apartamentos*, mas de bairros mais identificados com a classe média, como a zona leste em São Paulo e Tijuca, Vila Isabel e Meier no Rio de Janeiro. A *cidade não digna* e *sem cidade* estão claramente localizados periféricamente, sendo que esta última tende a ser mais periférica que a anterior.

Os limites não são perfeitamente definidos e as regiões de maior concentração de um *cluster* muitas vezes se sobrepõem às regiões de concentração de outro *cluster*.

## FIGURA 6

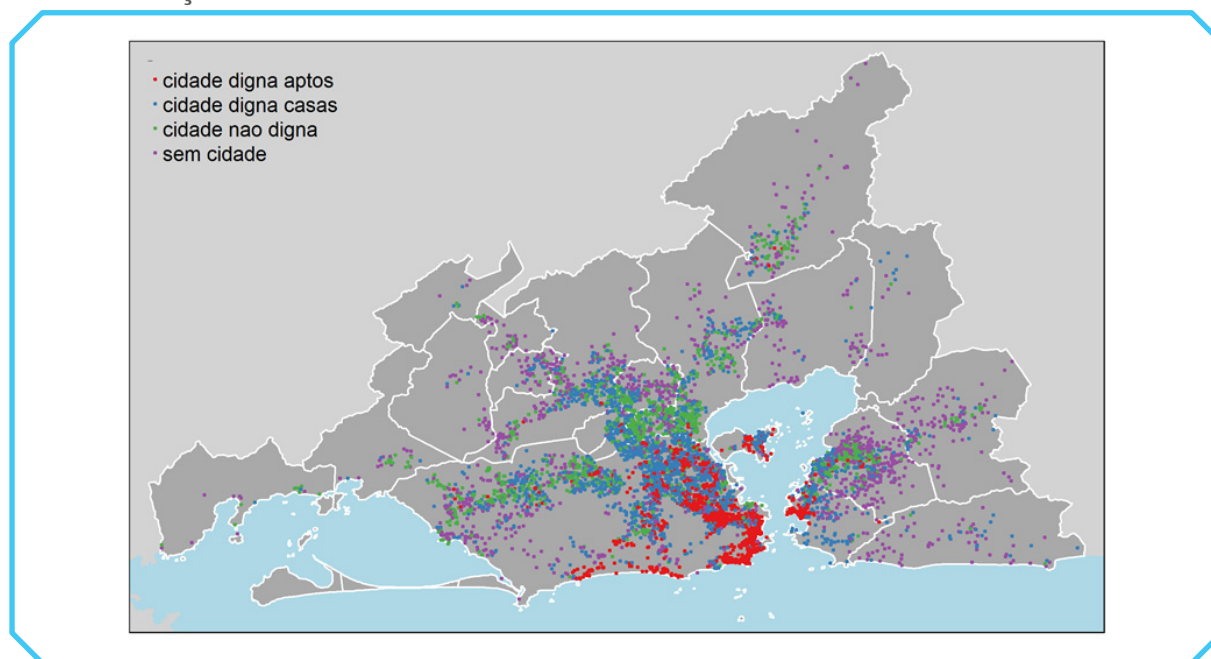
### Localização intraurbana dos *clusters* – São Paulo



Fonte: IBGE (2010b).

Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).



**FIGURA 7****Localização intraurbana dos *clusters* – Rio de Janeiro**

Fonte: IBGE (2010b.)

Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Assim, a partir dessa classificação seria impossível criar uma partição do território, gerando uma divisão da cidade em submercados, como na definição mais tradicional de submercados de habitação. Inclusive, no caso do Rio de Janeiro, o *cluster sem cidade* está muito mais entremeado nas regiões centrais em que os demais *clusters* se concentram que em São Paulo. Em São Paulo, *sem cidade* está bem claramente localizado nas zonas mais periféricas possíveis. A distância ao centro média é, de fato, maior no *cluster sem cidade* em São Paulo. No Rio e Janeiro, essa média é maior, porém a variância muito ampla,<sup>8</sup> havendo muitos setores do tipo *sem cidade* muito próximos aos do *cluster cidade digna apartamentos*.

8. O desvio-padrão da distância para os *sem cidade* no Rio de Janeiro é de 13,6 km, enquanto em São Paulo é de 9,1 km.

## 6.4 Correspondência com os setores subnormais

Já que os setores do *cluster sem cidade* são muito piores em termos de características urbanas e de renda, seria esse *cluster* o correspondente à cidade informal, ou às favelas? Na nossa classificação disponível, aos setores subnormais? No Rio de Janeiro, os *sem cidade* estão até localizados mais centralmente, visivelmente nos morros dos bairros nobres da cidade, onde ficam as favelas.

A tabela 4 mostra o quanto os moradores dos setores censitários classificados como AGSN se distribuíram nos *clusters* criados. No Rio de Janeiro, os moradores em AGSN se concentraram mais no *cluster cidade digna casas*; enquanto em São Paulo se concentraram mais em *cidade não digna*. Pouquíssimos moradores de AGSN estão em setores do *cluster cidade digna apartamentos*. Mas, apesar da concentração naqueles *clusters*, os moradores de AGSNs não correspondem à maioria dos totais de moradores em nenhum *cluster*.

**TABELA 4**

**Distribuição de moradores e moradores em AGSN nos *clusters* encontrados**

		A	B	A/B	$A/\sum A$	$B/\sum B$
		Total de moradores em AGSN	Total de moradores	Moradores em AGSN (%)	Total de moradores em AGSN (%)	Total de moradores (%)
Rio de Janeiro	Cidade digna aptos	46.283	2.272.588	2	3	19
	Cidade digna casas	1.117.262	4.060.509	28	64	34
	Cidade não digna	196.658	2.555.299	8	11	21
	Sem cidade	365.619	2.889.327	13	21	24
São Paulo	Cidade digna aptos	11.535	3.059.159	0	1	16
	Cidade digna casas	202.563	5.709.026	4	9	29
	Cidade não digna	1.476.500	7.972.023	19	68	41
	Sem cidade	355.058	1.585.994	22	16	8

Fonte: IBGE (2010b).

No caso do *cluster sem cidade*, tanto em São Paulo como no Rio de Janeiro, a grande maioria dos moradores de *sem cidade* não estão em setores classificados como AGSN. As grandes inadequações urbanas identificadas nesse *cluster* demonstram como é ele que concentra as necessidades habitacionais das duas metrópoles estudadas. Devido à definição de AGSN incluir algumas inadequações habitacionais – iluminação elétrica, coleta de lixo, água e esgotamento sanitário –, esperaríamos que ambas classificações fossem mais próximas. Cabe notar que, na definição de AGSN do IBGE, há dois aspectos que não são mensurados pelos questionários do censo, quais sejam: a urbanização fora do padrão e a ocupação de áreas restritas, por exemplo, faixas de domínio de rodovias e áreas contaminadas. Essas informações são coletadas antes da aplicação dos questionários, quando da classificação dos setores censitários em subnormais ou não.

Esse resultado ressalta, mais uma vez, a heterogeneidade entre os AGSNs. Alguns estudos empíricos já evidenciaram essa heterogeneidade (Brueckner, Mation e Nadalin, 2019; Mation, Nadalin e Krause, 2014). Há AGSN mais bem servidos, e menos bem servidos. Porém, essa classificação não corresponde à concentração de inadequações urbanas. Talvez a questão-chave para interpretar esse resultado seja que, nessas metrópoles analisadas, o fator que de fato distingue os AGSNs seja a inadequação fundiária; a urbanização fora do padrão e a ocupação de áreas restritas são informações não disponíveis e não utilizadas na análise aqui apresentada.

Encontramos evidências de que tanto o mercado formal quanto o mercado informal (aqui visto como os AGSNs) são segmentados, e que há regiões das metrópoles analisadas que se diferenciam bastante em termos de inadequações habitacionais e urbanas. Essas diferenciações podem não ser captadas pelos AGSNs.

**QUADRO 1**  
Principais achados

Variável	Fatos estilizados encontrados	Literatura com achado semelhante
Distância ao centro	Heterogeneidade entre cidades: poucas em que a maioria dos AGSNs são mais próximos.	Catalá e Carmo (2021); Nadalin e Mation (2018).
	Para o conjunto do Brasil, a mediana geral dos AGSNs é maior.	Inédito.
	São Paulo e Brasília: AGSNs mais distantes. Rio de Janeiro e Belo Horizonte: AGSNs mais próximos do centro.	Brueckner, Mation e Nadalin (2019); Catalá e Carmo (2021); Nadalin e Mation (2018).
Total de empregos em <i>buffer</i> de 500 m do setor censitário	AGSNs estão próximos a menos empregos formais.	Linke <i>et al.</i> (2016); Nadalin (2020); Nakamura e Avner (2018).
	No Rio de Janeiro e em Recife, os AGSNs são próximos a mais empregos que em São Paulo e em Manaus.	Inédito.
	A diferença é maior para municípios-sede de arranjos populacionais que para municípios do entorno.	Inédito.
Tempo de deslocamento	Medianas de AGSNs e não especiais coincidem para muitas cidades. Em geral, AGSNs apresentam tempos maiores.	Celhay e Gil (2020).
	Os tempos dos <i>melhores</i> AGSNs, percentil 25, são sempre bem menores que os dos <i>piores</i> não especiais, percentil 75.	Inédito.
	Há diferença nos tempos entre AGSNs e não especiais, tanto nos municípios-sede dos arranjos populacionais quanto nos municípios do entorno.	Inédito.
Valores de aluguel	Aluguéis de setores não especiais são maiores no Sul e no Sudeste.	Cravo, Resende e Cruz (2019).
	Não há padrão regional nos aluguéis em AGSN.	Inédito.
	Há associação linear entre aluguel médio formal e diferença de aluguéis médios de setores não especiais e de AGSNs.	Inédito.
	A diferença entre aluguéis médios de setores não especiais e de AGSNs é maior para municípios-sede de arranjos populacionais.	Inédito.
Análise de <i>cluster</i>	O <i>cluster</i> que concentra as inadequações urbanas não é aquele em que se concentram os moradores de AGSNs.	Catalá e Carmo (2021).

Elaboração da autora.

## 7 COMENTÁRIOS FINAIS

Este texto buscou caracterizar dois segmentos do mercado habitacional brasileiro, o formal e informal, conforme a definição dos AGSNs do IBGE. As principais semelhanças e diferenças foram explicitadas por meio da análise de cinco níveis geográficos. Encontramos desigualdades tanto regionais quanto intraurbanas, assim como na hierarquia das cidades. O que ficou mais claro é que os AGSNs são muito heterogêneos, e que não são, necessariamente, sinônimo de inadequações habitacionais. Passamos a resumir os principais resultados.

Com relação à inserção urbana, distância ao centro, proximidade a empregos e tempo de deslocamento, verificamos que as cidades se diferenciam bastante de acordo com diferenças na estrutura intraurbana. Mesmo assim, em geral, os AGSNs possuem distâncias ao centro menores, menos proximidade a empregos e mais tempo de deslocamento. No entanto, os melhores domicílios subnormais estão sempre melhor inseridos na cidade que os piores domicílios em setores não especiais.

Já com relação aos aluguéis, a distribuição dos valores em AGSN são sempre em valores menores que do mercado formal, mas há superposições de valores de ambos mercados. De acordo com a cidade, há maiores porções de ambos mercados superpostas. Em geral, as características da vizinhança e de estrutura física dos AGSNs são piores, justificando os alugueis menores, sendo que a ausência de revestimento em alvenaria e a presença de esgoto a céu aberto são as características que mais os distinguem. Nas cidades em que há mais domicílios em ambos mercados com valores superpostos, ou semelhantes, os domicílios em AGSN apresentam melhores características, relativamente às características dos domicílios em setores não especiais.

Há um padrão regional nos aluguéis e características da estrutura física e de vizinhança em que, em geral, o Norte e o Nordeste são piores que as demais regiões. Mesmo assim, encontramos que esse padrão espacial não é estatisticamente significativo para os aluguéis informais, somente para os formais. Isto nos indicaria que as condições de pobreza, e pouca capacidade de pagamento, não se diferenciam tanto entre as regiões do país quanto as condições de riqueza. Os maiores aluguéis de cada região apresentam variabilidade muito maior que os aluguéis informais, em geral. Este padrão de desigualdade maior para o mercado formal que para o informal também é encontrado para as diferenças intraurbanas, ao considerar a diferenciação dos municípios-sede, do entorno e isolados. A desigualdade intraurbana de inserção urbana também é mais pronunciada nos municípios-sede, onde os AGSNs são menos parecidos aos setores não especiais.

A primeira análise quantitativa apresentada no estudo encontrou uma correlação entre a média de aluguéis formais e a diferença das médias dos aluguéis formais e informais para 190

municípios para os quais haviam dados disponíveis, levando em consideração como variáveis de controle a região, se era município-sede e a média de banheiros dos domicílios formais. Essa correlação robusta é linear e documenta o quanto ambos mercados são substitutos.

Já a segunda análise quantitativa, utilizando algoritmos de clusterização de aprendizado de máquina, agrupou os setores censitários das metrópoles do Rio de Janeiro e de São Paulo de acordo com as suas características. Os grupos resultantes estão localizados em porções do território bastante definidas, o grupo de melhor qualidade correspondendo aos bairros sabidamente mais valorizados, por exemplo. Isto corroborou com a definição de segmentos de mercados que são porções do território com preços homogêneos.

No entanto, os AGSNs não têm a maioria dos moradores em nenhum dos *clusters*, nem sequer no *cluster* que claramente concentra mais inadequações habitacionais. Os AGSNs não são um *cluster* separado. Mesmo assim, encontram-se com maior frequência nos *clusters* de menor qualidade e quase não estão presentes no *cluster* de melhor qualidade. Esse resultado corrobora a visão de que os AGSNs são muito heterogêneos, e de que, para diminuir as inadequações habitacionais, as políticas públicas não devem restringir o foco, tratando apenas da questão da informalidade, ou dos AGSNs.

Este trabalho teve cunho principalmente descritivo, mas colabora por trazer informações muito ricas e inéditas, que contribuem para as reflexões necessárias à definição das políticas habitacionais e de desenvolvimento urbano no nosso país. Como futuros desenvolvimentos, ressaltamos a necessidade de resumir os achados teóricos das modelagens já disponíveis na literatura sobre o fenômeno em um grafo causal acíclico direcionado, de maneira a esclarecer os canais causais envolvidos. Assim, políticas mais assertivas e sustentáveis podem ser desenhadas.

## REFERÊNCIAS

ANSELIN, L. Local indicators of spatial association – LISA. **Geographical Analysis**, 27, n. 2, p. 93-115, Apr. 1995.

BRUECKNER, J. K.; MATION, L.; NADALIN, V. G. Slums in Brazil: where are they located, who lives in them, and do they 'squeeze' the formal housing market? **Journal of Housing Economics**, v. 44, p. 48-60, June 2019.

BRUECKNER, J. K.; SELOD, H. A theory of urban squatting and land-tenure formalization in developing countries. **American Economic Journal: Economic Policy**, v. 1, n. 1, p. 28-51, Feb. 2009.

CATALÁ, L. S.; CARMO, R. L. D. O conceito de aglomerado subnormal do IBGE e a precariedade dos serviços básicos de infraestrutura urbana. **Revista Brasileira de Estudos de População**, v. 38, p 1-24, 2021.

CATTANEO, M. *et al.* **On binscatter**. New York: Federal Reserve Bank of New York, Aug. 2021. (Staff Reports, n. 881).

CAVALCANTI, T.; DA MATA, D.; SANTOS, M. On the determinants of slum formation. **The Economic Journal**, v. 129, n. 621, p. 1971-1991, 2019.

CELHAY, P. A.; GIL, D. The function and credibility of urban slums: evidence on informal settlements and affordable housing in Chile. **Cities**, v. 99, p. 102-605, Apr. 2020.

CRAVO, T. A.; RESENDE, G. M.; CRUZ, B. D. O. Mensurando as disparidades regionais no Brasil: salários, custo de vida e amenidades locais. **Revista Brasileira de Economia**, v. 73, n. 1, p. 1-28, jan.-mar. 2019.

GOODMAN, A. C.; THIBODEAU, T. G. Housing market segmentation. **Journal of Housing Economics**, 7, n. 2, p. 121-143, June 1998.

GREENSTONE, M. The continuing impact of Sherwin Rosen's "Hedonic prices and implicit markets: product differentiation in pure competition". **Journal of Political Economy**, 125, n. 6, p. 1891-1902, Dec. 2017.

HENDERSON, J. V.; REGAN, T.; VENABLES, A. J. Building the city: from slums to a modern metropolis. **The Review of Economic Studies**, v. 88, n. 3, p. 1157-1192, May 2021.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Demográfico 2010: aglomerados subnormais – primeiros resultados**. Rio de Janeiro: IBGE, 2010a.

\_\_\_\_\_. **Censo demográfico: resultados do universo**. Rio de Janeiro: IBGE, 2010b.

\_\_\_\_\_. **Arranjos populacionais e concentrações urbanas do Brasil**. 2. ed. Rio de Janeiro: IBGE, 2016.

IRIZARRY, R. A. (Ed.). **Introduction to data science: data analysis and prediction algorithms with R**. Boca Raton: CRC Press, 2019.

LINKE, C. *et al.* **Inserção urbana de habitação de interesse social: um olhar sobre mobilidade cotidiana e uso do solo**. Rio de Janeiro: Ipea, fev. 2016. (Texto para Discussão, n. 2176).

MATION, L.; NADALIN, V.; KRAUSE, C. **Favelização no Brasil entre 2000 e 2010: resultados de uma classificação comparável**. Brasília: Ipea, out. 2014. (Texto para Discussão, n. 2009).



NADALIN, V. Slums, accessibility to opportunities and economic outcomes. An empirical analysis for São Paulo, Brazil. *In*: ENCONTRO NACIONAL DE ECONOMIA, 48., 2020, [s.l.]. **Proceedings...** [s.l.]: Anpec, 2020.

NADALIN, V. G.; MATION, L. F. **Localização intraurbana das favelas brasileiras**: o papel dos fatores geográficos. Brasília: Ipea, jun. 2018. (Texto para Discussão, n. 2390).

NAKAMURA, S.; AVNER, P. **Spatial distributions of job accessibility, housing rents, and poverty in Nairobi, Kenya**. Washington: The World Bank, Nov. 2018. (Policy Research Working Paper, n. 8654).

PEARL, J. Causal diagrams for empirical research. **Biometrika**, v. 82, n. 4, p. 669-688, Dec. 1995.

PEREIRA, R. H. M. *et al.* **Desigualdades socioespaciais de acesso a oportunidades nas cidades brasileiras**: 2019. Brasília: Ipea, jan. 2020. (Texto para Discussão, n. 2535).

ROBACK, J. Wages, rents, and the quality of life. **Journal of Political Economy**, v. 90, n. 6, p. 1257-1278, Dec. 1982.

ROSEN, S. Hedonic prices and implicit markets: product differentiation in pure competition. **Journal of Political Economy**, v. 82, n. 1, p. 34-55, Jan.-Feb. 1974.

SELOD, H.; TOBIN, L. **The informal city**. Washington: World Bank, June 2018. (World Bank Policy Research Working Paper, n. 8482).

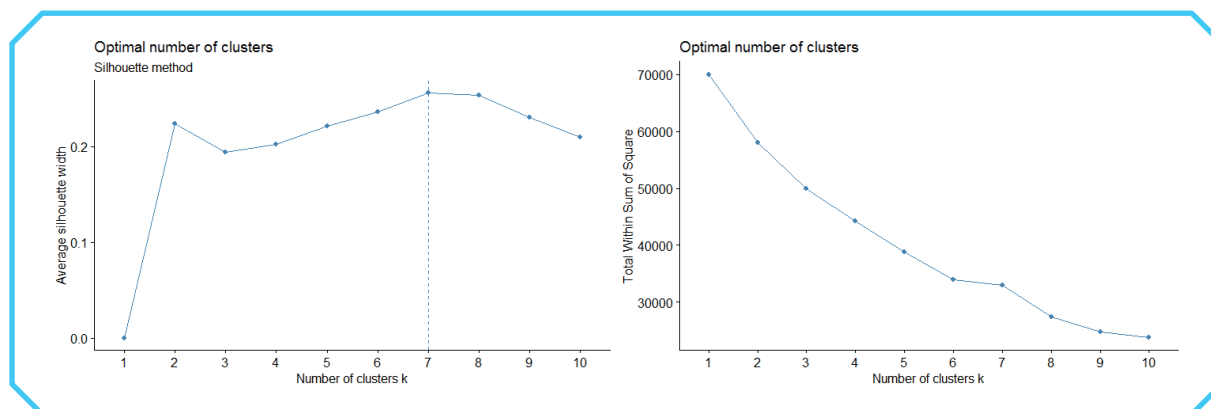
SMOLKA, M. O.; BIDERMAN, C. Housing informality: an economist's perspective on urban planning. *In*: BROOKS, N.; DONAGHY, K.; KNAAP, G.-J. (Ed.). **The Oxford handbook of urban economics and planning**. Oxford: Oxford University Press, 2011. p. 814-833.

## APÊNDICE A

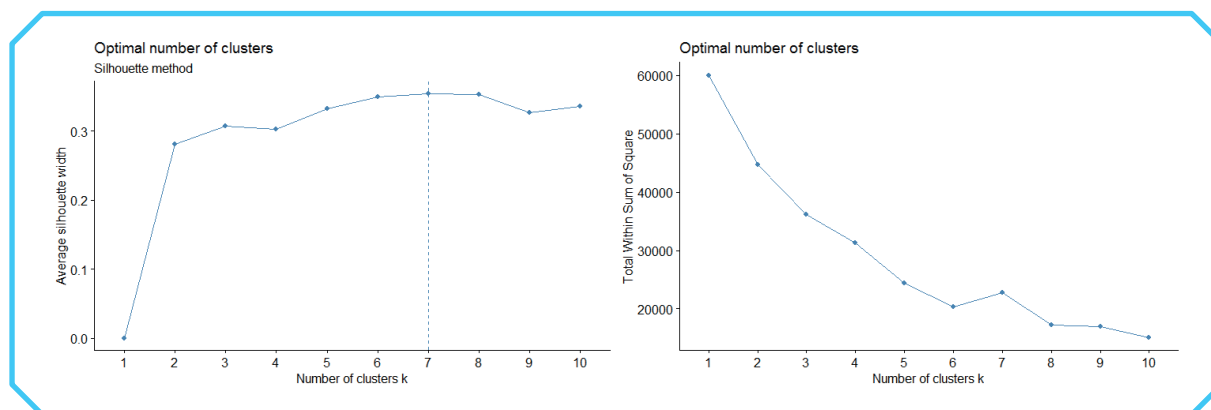
## GRÁFICO A.1

Validação ou ajuste dos números de *clusters*<sup>1</sup>

## A.1A – São Paulo



## A.1B – Rio de Janeiro



Elaboração da autora.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Nota: <sup>1</sup> Dezesesseis variáveis previsoras, amostra de 10 mil setores censitários.

## APÊNDICE B

## GRÁFICO B.1

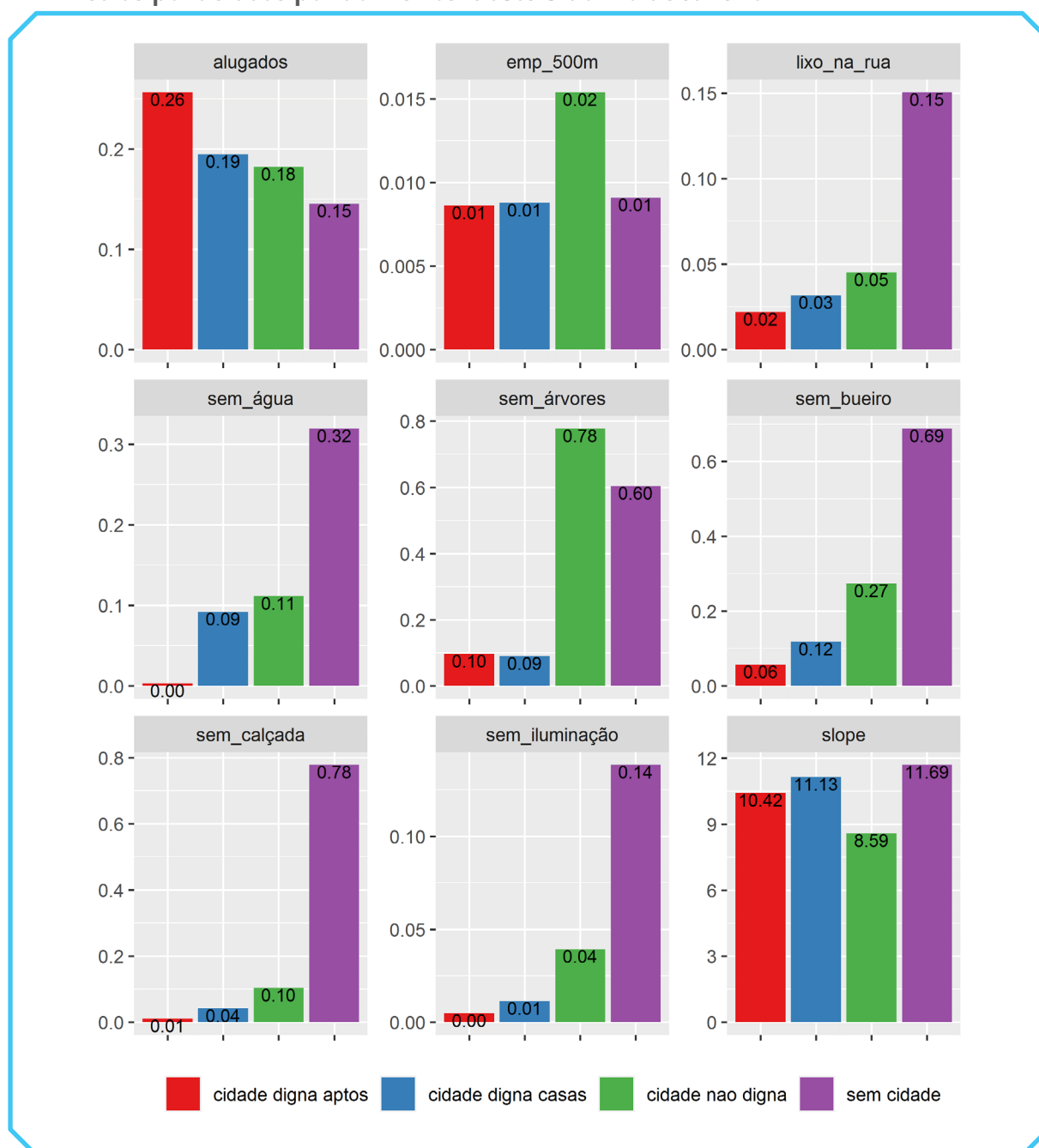
Médias ponderadas por domicílios: *clusters* de São Paulo

Elaboração da autora.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

## APÊNDICE C

## GRÁFICO C.1

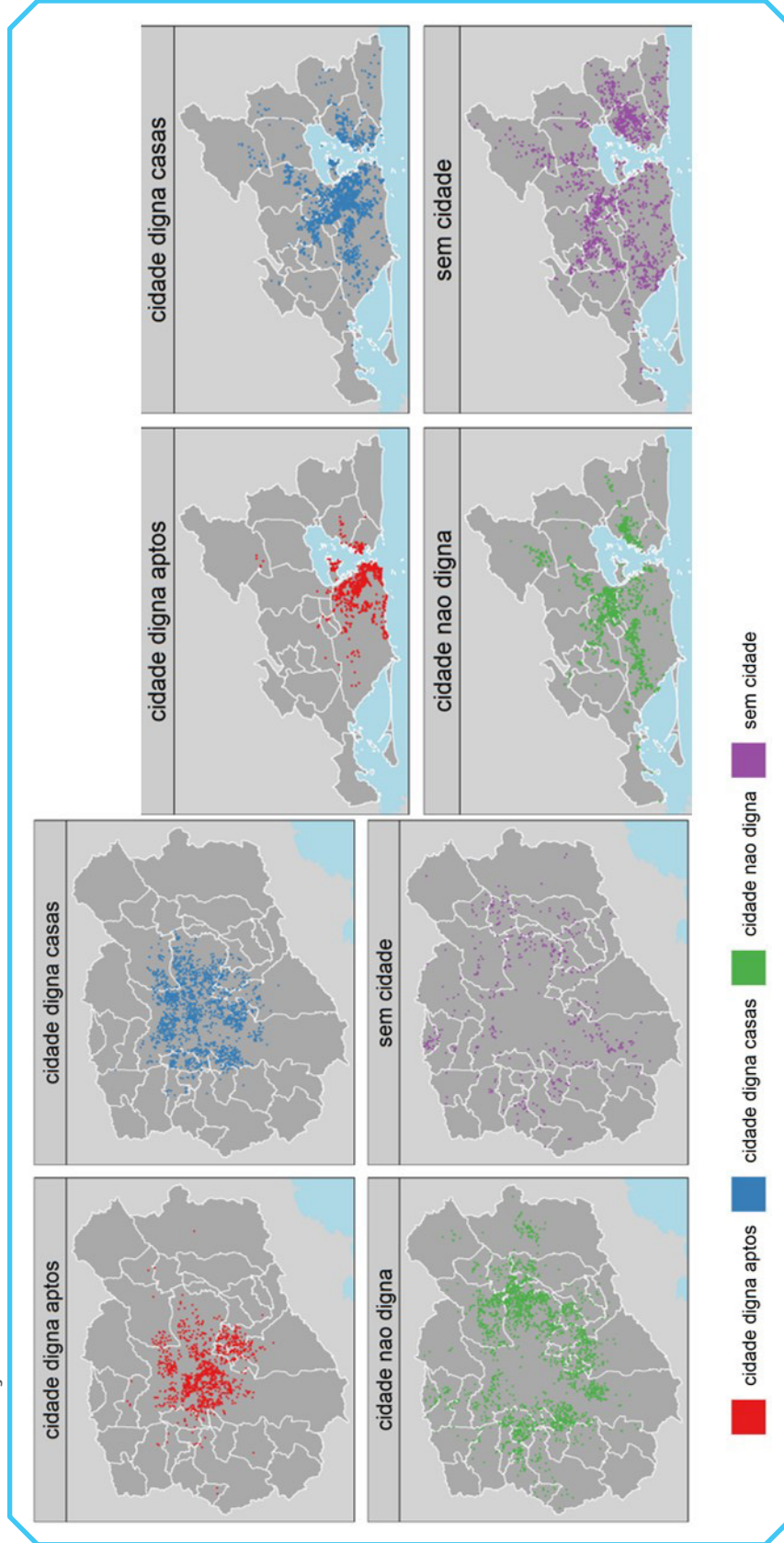
Médias ponderadas por domicílios: *clusters* do Rio de Janeiro

Elaboração da autora.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

## APÊNDICE D

### GRÁFICO D.1 Localização intraurbana dos *clusters*



Elaboração da autora.

Obs.: Gráfico cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

# Ipea – Instituto de Pesquisa Econômica Aplicada

## EDITORIAL

### Coordenação

Aeromilson Trajano de Mesquita

### Assistentes da Coordenação

Rafael Augusto Ferreira Cardoso

Samuel Elias de Souza

### Supervisão

Camilla de Miranda Mariath Gomes

Everson da Silva Moura

### Revisão

Alice Souza Lopes

Amanda Ramos Marques

Ana Clara Escórcio Xavier

Barbara de Castro

Clícia Silveira Rodrigues

Olavo Mesquita de Carvalho

Regina Marta de Aguiar

Reginaldo da Silva Domingos

Brena Rolim Peixoto da Silva (estagiária)

Nayane Santos Rodrigues (estagiária)

### Editoração

Anderson Silva Reis

Cristiano Ferreira de Araújo

Danielle de Oliveira Ayres

Danilo Leite de Macedo Tavares

Leonardo Hideki Higa

### Capa

Aline Cristine Torres da Silva Martins

### Projeto Gráfico

Aline Cristine Torres da Silva Martins

*The manuscripts in languages other than Portuguese published herein have not been proofread.*

### Ipea – Brasília

Setor de Edifícios Públicos Sul 702/902, Bloco C

Centro Empresarial Brasília 50, Torre B

CEP: 70390-025, Asa Sul, Brasília-DF

## **Missão do Ipea**

Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.



**ipea** Instituto de Pesquisa  
Econômica Aplicada

MINISTÉRIO DO  
PLANEJAMENTO  
E ORÇAMENTO

