



Universidade Federal de Pernambuco
Programa de Pós-Graduação em Economia – PIMES

Angelo Fernando Gonzaga dos Santos

Determinantes da Expansão Urbana no Brasil

Recife, 2020

Angelo Fernando Gonzaga dos Santos

Determinantes da Expansão Urbana no Brasil

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciências Econômicas (área de concentração: Teoria Econômica), como parte dos requisitos necessários para a obtenção do Título de Mestre em Economia.

Orientador: Raul da Mota Silveira Neto

Coorientador: Raydonal Ospina Martínez

Recife

2020

Catálogo na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

S237d

Santos, Angelo Fernando Gonzaga dos
Determinantes da expansão urbana no Brasil / Angelo Fernando
Gonzaga dos Santos. - 2020.
54 folhas: il. 30 cm.

Orientador: Prof. Dr. Angelo Fernando Gonzaga Silveira Neto e
Coorientador Prof. Dr. Raydonal Ospina Martinez.

Dissertação (Mestrado em Economia) – Universidade Federal de
Pernambuco, CCSA, 2020.

Inclui referências e apêndices.

1. Expansão urbana. 2. Modelo AMM. 3. Uso do solo. I. Silveira Neto,
Raul da Mota (Orientador). II. Ospina Martinez, Raydonal (Coorientador).
III. Título.

336 CDD (22. ed.)

UFPE (CSA 2020 – 075)

ANGELO FERNANDO GONZAGA DOS SANTOS

DETERMINANTES DA EXPANSÃO URBANA NO BRASIL

Dissertação ou Tese apresentada ao Programa de Pós-Graduação em 29 de maio de 2020 da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Economia.

Aprovado em: 29/05/2020.

BANCA EXAMINADORA

Profº. Dr. Raul da Mota Silveira Neto (Orientador)
Universidade Federal de Pernambuco

Profº. Dr. Raydonal Ospina Martínez (Examinador Interno)
Universidade Federal de Pernambuco

Profº. Dr. Ana Maria Bonomi Barufi (Examinador Externo)
Universidade Federal de São Paulo

Resumo

O presente trabalho tem como objetivo estudar os condicionantes da área urbana no Brasil nos anos de 1991, 2000 e 2010. Para isso, foi utilizado o modelo monocêntrico básico e construído uma base de painel inédita para o caso brasileiro. A motivação do presente trabalho é a falta de estudos sobre o crescimento da área urbana no Brasil e a disponibilidade de novos dados de satélite, land cover. Primeiro, este trabalho apresenta a literatura econômica sobre o espraiamento das cidades. Em segundo lugar, são apresentados a estratégia empírica e os dados utilizados para construir a base de dados. Em seguida, os resultados para as AMCs e Regis brasileiras são expostos, assim como as respectivas elasticidades. E por fim, são feitas considerações finais. As evidências do trabalho são comparadas com as de trabalhos feitos para o caso americano. Evidências inéditas sobre o papel da moradia informal no crescimento da área urbana são apresentadas. Os resultados encontrados para o caso brasileiro sugerem que os condicionantes da área urbana propostos pela teoria urbana tem as relações esperadas e estão de acordo com os trabalhos anteriores. Além disso, a moradia subnormal é positivamente correlacionada com o espraiamento das cidades. As elasticidades encontradas sugerem que as particularidades do Brasil podem afetar na magnitude das forças sobre a área urbana.

Palavras-chaves: Expansão Urbana. modelo AMM. Land Cover. Uso do solo.

Abstract

This work aims to study the conditioning forces of the urban area in Brazil for the years 1991, 2000, and 2010. The standard monocentric model and a unique panel database for Brazil are used to achieve this purpose. The motivation of the present work is the lack of evidence concerning the urban area sprawl in Brazil and the new satellite data available about the land cover in Brazil. First, this work will present the literature review about the economic theory of urban sprawling. Second, this work presents the data and methodology used to achieve the objectives of the paper. After that, the results for the Brazilian AMCs and Regis samples are presented with the estimated elasticities. In the end, a final discussion is shown. The results suggest that the conditioning forces shown in the theory were found for the Brazilian case and are consistent with other works. Results about the role of informal housing in urban sprawl are shown. Also, the smaller elasticities found for Brazil suggest that the particularities of the country can restrict the magnitudes of the forces in the urban sprawl.

Keywords: Urban expansion. AMM model. Land Cover. Land use.

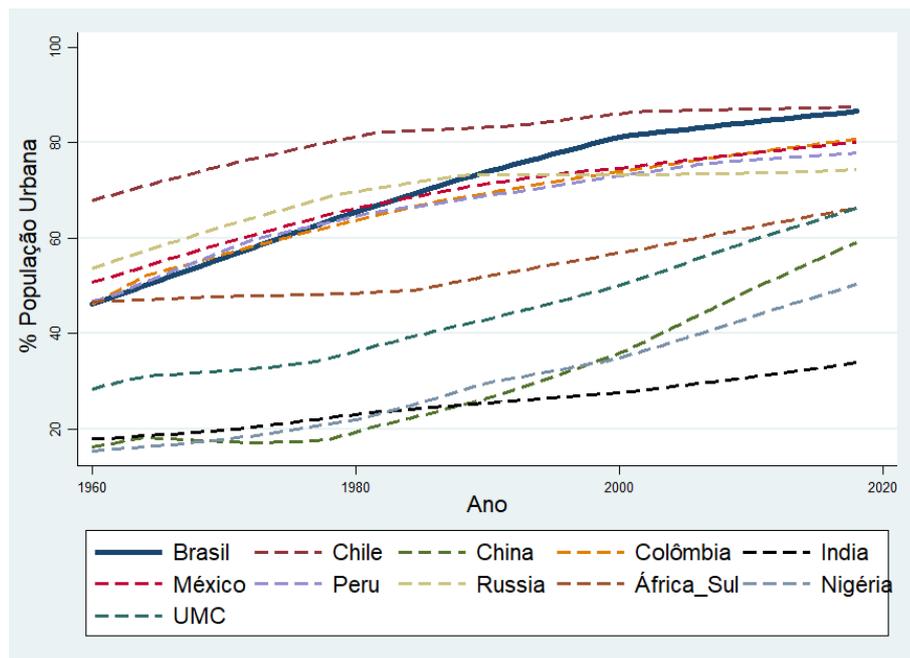
Sumário

1	INTRODUÇÃO	8
2	REVISÃO DE LITERATURA	14
2.1	Teoria Econômica e Área Urbana	14
3	ESTRATÉGIA EMPÍRICA E DADOS	21
3.1	MQO e Painel	21
3.2	Dados	23
3.2.1	Dados de imagem de satélite	23
3.2.2	Condicionantes da área urbana	26
3.2.3	Estatística descritivas	28
4	RESULTADOS	31
4.1	AMCs	31
4.2	Regis	35
4.3	Elasticidades	39
5	CONSIDERAÇÕES FINAIS	44
	REFERÊNCIAS	45
	APÊNDICE A - TESTES DE ROBUSTEZ	48

1 Introdução

Segundo os dados do Censo Demográfico (IBGE) em 2010, o Brasil apresentara 85% de sua população localizada no meio urbano. Isto se deve ao aumento da produtividade agrícola impulsionada pelo intenso processo de industrialização. Tal processo foi um dos fatores que desencadeou migrações do campo para cidade, criando assim novos centros urbanos e expandindo os já existentes (Martins e Silveira Neto, 2012). A figura 1 mostra a evolução do percentual da população total residente em áreas urbanas entre os anos de 1960 e 2018 em países com renda média alta, similar ao Brasil de acordo com definição do Banco Mundial. Também foi adicionado a média da população deste grupo de renda, identificado pela sigla UMC. Ao comparar o Brasil com os demais países, observa-se que o percentual da população urbana no país é maior em relação a maioria dos países incluídos no gráfico, com exceção do Chile. Embora existam trabalhos como os de Silva et al. (2017) e Da Mata (2007) que estudam o crescimento urbano do Brasil através do crescimento populacional das cidades, não foi encontrado muitos estudos, como o de Vila Nova (2014), que estudam a evolução da área urbana das cidades. Os novos dados de *land cover* disponíveis para o território brasileiro se somam aos fatores anteriormente mencionados como motivadores para a condução do presente trabalho no estudo da evolução da área urbana das cidades no Brasil.

Figura 1 – Evolução da População Urbana medida como % da População total. Fonte: Elaboração própria



Para a compreensão das forças que afetam a dinâmica do uso do espaço nas cidades,

economistas têm utilizado modelos teóricos para gerar evidências consistentes acerca deste processo. A compreensão destas forças é de fundamental importância para o planejamento público das cidades, dado que a extensão das cidades está ligada ao bem estar dos residentes, especialmente por questões relacionadas aos custos de aglomeração e de espraiamento. Alguns desses custos são o aumento do *commuting*, a poluição e a necessidade de melhor infraestrutura urbana para atender toda a extensão das cidades. Portanto, entender como melhor prever a dinâmica de crescimento das cidades é de fundamental importância. Além disso, o entendimento de como os agentes enfrentam o *trade-off* entre acessibilidade e consumo de espaço pode auxiliar no planejamento de políticas mais assertivas para o desenvolvimento da área urbana das cidades.

O modelo mais aceito pelos economistas para explicar este fenômeno é o proposto por Alonso-Mills-Muth (Alonso, 1964; Mills, 1967; Muth, 1969). Os autores consideram o *trade-off* entre acessibilidade e consumo de espaço, enfrentado pelas famílias residentes da cidade. Neste modelo, quatro forças explicam a distribuição das pessoas e famílias no espaço urbano. Estas forças atuam a favor ou contra a expansão da área das cidades a depender de como afetam a capacidade de escolha dos agentes em trocar espaço e acessibilidade. O primeiro condicionante desta dinâmica é a população. O número de pessoas residentes em uma cidade afeta a escolha locacional das famílias através do aumento da disputa por espaço. Ou seja, quanto mais pessoas residirem na cidade, a tendência de espraiamento da mesma tende se verificar. Esse movimento acontece pela necessidade do agente consumir espaço, que torna-se mais escasso e caro com o crescimento populacional e condiciona os agentes a consumirem mais espaço nas franjas da cidade, causando espraiamento. A segunda força é a renda, que possibilita que o agente troque acessibilidade por espaço mais facilmente, afetando assim sua escolha locacional e, por consequência, a composição da cidade. Terceiro, o preço da terra agrícola: quanto maior a produtividade do solo não-urbano, maior o custo de oportunidade de transformá-lo em uso urbano. Este fator é determinante para a definição do tamanho das cidades pois afeta diretamente o uso do entorno dos aglomerados urbanos. Por exemplo, se o rendimento da terra for elevado o suficiente para compensar seu uso como área urbana, a cidade tende a ter seus limites restringidos. Isto se deve ao fato de que o entorno será vastamente utilizado para atividades agrícolas ou agropecuárias, restando assim pouco espaço para uso urbano. Por outro lado, se a terra não possuir alto rendimento, a tendência é de espraiamento da cidade. Neste caso, utilizar o solo para fins urbanos ou como consumo de espaço torna-se mais rentável para os agentes. Por fim, a tecnologia de transporte também pode afetar a composição das cidades. Esse componente afeta os custos de transporte e por consequência a escolha entre espaço e acessibilidade dos agentes. Por exemplo, o advento do carro tornou possível acessar as franjas da cidade mais facilmente. Portanto, um agente poderia morar mais longe e acessar o seu centro de emprego no mesmo tempo ou até mais rapidamente. Reduções de custo de transporte como essa são fundamentais para a resolução do *trade-off* entre acessibilidade

e espaço decidido pelas famílias (Fujita 1989; Fujita e Thisse 1996).

Inicialmente o modelo AMM foi testado para o caso norte-americano, mostrando-se eficiente para explicar o processo de expansão das cidades. Os resultados mostram a existência de uma correlação positiva entre população e renda com o crescimento das cidades. A renda agrícola foi identificada com força limitante da expansão das cidades, tendo relação negativa com a área urbana. Além disso, Brueckner e Fansler (1983), McGrath (2005), Spivey (2008) e Paulsen (2012), mesmo utilizando diferentes especificações econométricas e bancos de dados distintos, encontram evidências consistentes com o modelo tradicional AMM, solidificando a tese de que esses fatores são de fato relevantes para explicar a dinâmica do uso do espaço urbano. O modelo também foi testado em cidades europeias. Garcia-López et al. (2018) encontram resultados similares aos observados para o caso norte-americano. A renda e população seriam forças pró expansão urbana, mas o rendimento agrícola teria o efeito contrário, como previsto pelo modelo teórico.

Quanto a países em desenvolvimento existem alguns estudos que buscam validar a hipótese do modelo nestes casos. Em uma pesquisa realizada sobre a China, Deng et al (2008) encontraram resultados muito semelhantes ao observado para as cidades norte-americanas em relação aos condicionantes da área urbana. Utilizando dados em painel com três anos, os autores verificaram se as forças citadas no modelo clássico teriam correlação com a dinâmica do crescimento urbano. Mais uma vez, as evidências indicam que a população e a renda estariam positivamente correlacionados com o espraiamento da cidade, efeito oposto ao rendimento agrícola.

Para o caso brasileiro, Vila Nova (2014) utilizam o mesmo modelo para verificar se os resultados são semelhantes aos casos internacionais. Os autores apresentam evidências de que renda e população e são associados as expansão do uso do solo urbano, ou seja, são positivamente correlacionadas com o espraiamento das cidades. Já a renda agrícola é negativamente correlacionada com o crescimento da área urbana.

A disponibilidade de novos tipos de dados de satélite motivaram pesquisas que os utilizam para testar a robustez do modelo e as evidências empíricas encontradas anteriormente. Alguns dos dados utilizados são os de luzes noturnas. Trabalhos utilizam esse tipo de dado com o objetivo de explicar a dinâmica econômica local e utilizar esta informação como *proxy* da atividade econômica (Henderson et al., 2012; Hodler e Raschky, 2014; Harari, 2016; Henderson et al., 2016; Storeygard, 2016). Outro tipo de informação advinda das imagens de satélite é o dado de *land cover*, que buscam caracterizar o uso e cobertura do solo terrestre. Na literatura de economia urbana tais informações foram exploradas em alguns trabalhos (Paulsen, 2012; Vila Nova, 2014; Deng et al. 2008) e têm sido consideradas como uma alternativa para solucionar problemas referentes aos dados utilizados anteriormente. Os estudos de Paulsen (2012) e Vila Nova (2014) utilizam dados de satélite como variável explicativa de suas regressões. Deng et al. (2008) também utiliza

informações extraídas de imagem de satélite em suas regressões para testar o modelo monocêntrico. O principal argumento para o uso deste novo tipo de dado é a possível incapacidade da densidade urbana de captar de forma eficiente o uso urbano do solo em alguns locais. Por exemplo, pode existir uma determinada região com um uso urbano intensivo, como fábricas, ou centros industriais, porém com baixa densidade. Os dados de *land cover* são derivados de imagens de satélites e extraídos através de classificação de imagem. Esse dado tem como objetivo classificar o uso do solo, como por exemplo área urbana ou agricultura. Com a identificação de áreas urbanas independentemente da densidade de pessoas em determinada área, tendo por base a mudança da natureza do solo, pode-se captar locais com baixa densidade mas com uso intensivo urbano classificado corretamente. Este tipo de dado tem sido discutido como uma boa alternativa para a densidade demográfica na classificação do uso urbano.

Essas novas variáveis são utilizadas por alguns estudos para verificar a robustez de evidências empíricas passadas de modelos. Paulsen (2012) mostra que considerar tais variáveis não altera os resultados encontrados considerando a densidade demográfica, o que dá robustez às evidências empíricas anteriores e corrobora com a consistência de trabalhos clássicos que utilizam a densidade. Além disso, considerando o elevado custo de acesso a dados de satélites, devido ao fato de em geral de serem privados ou o alto poder computacional exigido para os categorizar, talvez não valha a pena utilizá-los. Tendo em vista tal discussão, o presente estudo também tem como objetivo utilizar dados de *land cover* para o caso brasileiro para avaliar sua viabilidade. Além das informações de satélite, foram consideradas variáveis referentes à presença de aglomerados subnormais nas cidades brasileiras. A teoria urbana tem discutido o papel da moradia informal sobre o espraiamento da cidades. Espera-se que devido à característica mais horizontal destas habitações, impedindo construções mais verticais, a ocupação do espaço seria menos adensada, fazendo com que os limites da cidade sejam expandidos para acomodar mais residentes. Para o caso brasileiro, essa discussão é bastante importante devido à forte presença dos aglomerados subnormais em cidades do país.

O principal objetivo deste trabalho é adicionar evidências acerca da aplicabilidade do modelo monocêntrico e sobre os condicionantes do crescimento urbano para as cidades brasileiras. Outros objetivos secundários são utilizar dados de satélite - *land cover* - na avaliação do modelo monocêntrico e explorar o papel dos aglomerados subnormais sobre o crescimento das cidades brasileiras. Embora existam evidências e estudos bem consolidados sobre a dinâmica das cidades americanas (Brueckner e Fansler, 1983; McGrath, 2005; Spivey, 2008; e Paulsen, 2012), evidências sobre a mesma dinâmica no Brasil ainda são escassas. Além disso, o uso de bases de dados inéditas utilizadas para o Brasil é outra adição a literatura. Primeiro, foi desenvolvido um painel com três anos, a construção deste painel é importante pois possibilita controlar para fatores não observáveis dos AMCs. Segundo, o uso de dados de *land cover*, que vem sendo utilizado em modelos de expansão

urbana (Deng et al., 2008; Paulsen, 2012; Vila Nova, 2014). Neste trabalho foi utilizado a base do MapBiomas como fonte de dados de *land cover*. Metodologicamente, foram estimados painel e MQO com o objetivo de replicar trabalhos feitos para outras cidades do mundo (Paulsen, 2012; Spivey, 2008; McGrath, 2005; Brueckener e Fasler, 1983).

Com relação à unidade de análise, foram consideradas duas agregações para os municípios. Isso permite comparar os territórios intertemporalmente de maneira mais adequado e considerar efeitos pendulares. A primeira agregação que utilizada tem como base o trabalho de Reis et al. (2008). No Brasil, houve a criação de novos municípios ao longo do tempo, o que impossibilita a comparação intertemporal destas unidades de observação. Os pesquisadores propõem a união de alguns municípios em Áreas Mínimas Comparáveis, as AMCs, para criar unidades de estudo comparáveis ao longo do tempo, antes da desagregação. A segunda agregação a ser considerada é baseada nas Regiões Geográficas Imediatas, Regis, criadas pelo IBGE. O critério de agregação destes blocos regiões considera a interdependência entre municípios na provisão de trabalho, serviços públicos, consumo de bens, moradia, entre outros. Devido a essa integração, podem existir efeitos associados a movimentos pendulares, ou seja, existir influência de determinado município sobre a dinâmica urbana de outro.

Os resultados de análise indicam que o caso Brasileiro segue o proposto pela teoria urbana. População e renda são forças pró crescimento da área urbana das cidades, enquanto a renda agrícola tende a retrair este movimento. Este resultado foi encontrado para os modelos MQO e painel utilizados. Um resultado inédito e importante para o caso brasileiro diz respeito à população subnormal. Foi encontrada uma relação positiva entre a variável de moradia informal e a expansão da área urbana, indicando existir força pró espraiamento ocasionado pela presença de aglomerados subnormais. Este resultado foi encontrado utilizando os modelos MQO e painel, para as duas amostras, AMCs, e Regis. Quanto às elasticidades de renda e população, foram encontrados resultados importantes. Para as elasticidades encontradas utilizando MQO, as magnitudes encontradas são semelhantes às vistas no caso americano. Porém ao utilizar as estimações em painel, observa-se que as elasticidades de renda e população são inferiores quando comparadas às encontradas por Paulsen (2012). Este resultado sugere que fatores particulares do Brasil, como desigualdade elevada e precária infraestrutura urbana podem reduzir a demanda por espaço advinda de aumentos na renda e população. Ainda sobre estes resultados, observam-se elasticidades maiores para Regis menores, o que pode indicar que fatores particulares das estruturas de Regis maiores podem reduzir a demanda por espaço. Por exemplo, o tempo de *commuting* e desigualdade mais elevados em Regis com mais de 500 mil habitantes podem contribuir com a diminuição da magnitude das elasticidades.

O restante deste trabalho está organizado da seguinte forma. A segunda seção, traz a revisão de literatura que fundamenta a análise a ser realizada. Na seção três serão

apresentadas a estratégia empírica e dados utilizados. Após isto, serão discutidos os resultados do trabalho e, por fim, a última seção traz as considerações finais.

2 Revisão de Literatura

Esta seção explora a literatura relativa ao modelo monocêntrico e aos condicionantes da expansão das cidades. Além disso, são discutidas as aplicações do modelo teórico em dados de países diferentes e com metodologia distintas. Por fim, são discutidos estudos que utilizam dados de satélite com o objetivo de testar novas variáveis, acrescenta robustez a trabalhos anteriores e para explicar outros fatos econômicos.

2.1 Teoria Econômica e Área Urbana

Economistas defendem que a expansão urbana é resultado do processo coordenado de mercado que aloca a terra para uso agrícola ou urbano a depender dos custos e ganhos associados a cada tipo de uso do espaço (Bruckner e Fasler 1983). Com o intuito de comprovar esta tese, aplicações de modelos teóricos buscaram verificar se as hipóteses dos modelos econômicos são válidas para explicar a processo de expansão das cidades. Esses trabalhos têm como objetivo investigar os fatores associados à dinâmica de crescimento das áreas urbanas. O modelo mais utilizado e consolidado para explicar os condicionantes da urbanização é o modelo Alonso-Mills-Muth, ou AMM (Alonso, 1964; Mills, 1967; Muth 1969). Este modelo apresenta o arcabouço teórico mais utilizado para explicar as forças que afetariam a alocação espacial de agentes nas cidades. O modelo considera o *trade-off* entre acessibilidade e consumo de espaço resolvido pelas famílias ao escolherem aonde irão se morar. Segundo as hipóteses do modelo, esta escolha locacional seria afetada por quatro forças: população, renda, custos de transporte e produtividade agrícola. A população na medida em que cresce, acirra a competição por espaço, e dada uma fronteira r de uma cidade, as pessoas não poderiam ter o mesmo nível de utilidade com a mais pessoas residindo no mesmo espaço. Por isso, o aluguel e a fronteira da cidade crescem, para que seja possível acomodar mais pessoas. A renda, por sua vez, afeta o *tradeoff* entre acessibilidade e uso de espaço, pois tem efeito sobre os preços e os níveis de consumo de espaço. O uso do espaço aumenta na cidade e torna os aluguéis mais caros, fazendo com que a fronteira r da cidade se expanda. Com relação aos custos de transporte, a redução destes custos afetam o *tradeoff* entre espaço e acessibilidade, dado que se pode acessar locais mais distantes em menos tempo. Portanto, morar nas franjas da cidade torna-se menos desvantajoso. Ou seja, dado que o custo de morar no subúrbio é reduzido, as pessoas se dispõem a morar mais longe do centro e consumir mais espaço. Por fim, o aluguel agrícola define a competitividade da terra agrícola em relação aos aluguéis pagos pelo setor urbano. Com o aumento da produtividade agrícola, o custo de transformar o uso da terra agrícola em urbano aumenta, fazendo com que este entorno seja destinado

ao uso produtivo. Entornos mais produtivos tendem a restringir o desenvolvimento da fronteira das cidades, gerando cidades mais concentradas. Por outro lado, terras agrícolas de baixa produtividade que tem como alternativa produtiva o uso urbano, fazendo com que as cidades sejam mais espalhadas.

Para compreender melhor como estes condicionantes afetam a dinâmica de crescimento das cidades, o modelo teórico será exposto mais detalhadamente. Segundo Fujita (1989), as hipóteses básicas do modelo são três. A primeira, é a existência de um CDB, onde o emprego da cidade se concentra. A segunda, é a existência de custos monetários quer o agente enfrenta para acessar o trabalho e sua moradia. Terceiro, os agentes possuem um renda y e consomem dois bens, moradia e outros bens. Spivey (2008) também pontua as hipóteses do modelo AMM. Uma delas é que os consumidores têm o mesmo nível de renda no CDB e que possuem as mesmas preferências a cerca do consumo de espaço, consumindo apenas moradia e outro bem numérico. O autor também afirma existir um custo de transporte monetário enfrentado pelos agentes para que possam se deslocar na cidade.

Os agentes residentes na cidade, maximizam sua utilidade sujeito à sua restrição orçamentária para que consumam moradia e outro bens (no modelo, o numérico). Formalmente, a $U(z, s)$ é maximizada sujeito à restrição orçamentária definida por $z + p(x)q + tx$, onde z é o bem numérico, x a distância ao CBD, $p(x)$ o preço da moradia que é função da distância, q é consumo de moradia e t o custo de transporte enfrentado pelo agente. Formalmente será:

$$\max_{z,q,x} U(z, s) \quad s.t. \quad I = z + p(x)q + tx \quad (2.1)$$

A partir deste modelo e assumindo que ambos os bens são normais e que possuem efeito renda positivo, pode-se obter as derivadas que explicam como os condicionantes descritos acima, população, renda, custos de transporte e valor da terra, afetam os limites da cidade. As derivadas obtidas são as seguintes:

$$\frac{d\bar{x}}{dN} > 0, \frac{d\bar{x}}{dI} > 0, \frac{d\bar{x}}{dr_a} < 0, \frac{d\bar{x}}{dt} < 0, \quad (2.2)$$

Estas derivadas formalizam como os condicionantes afetam os limites da cidade, \bar{x} . Fujita (1989) e Wheaton (1974) explicam teoricamente estas derivadas. A população, N , tem efeito sobre a competição de espaço dos agentes na cidade. Com o aumento da população, a disputa de espaço tende a se acirrar devido ao aumento de demanda pela terra, ocasionando crescimento do preço da moradia. Na fronteira da cidade, isto faz com que o limite da cidade seja expandido, reduzindo a utilidade das famílias. O aumento da renda, I , faz com que as pessoas possam trocar acessibilidade e espaço mais facilmente, aumentando a demanda por espaço dos indivíduos na periferia. Considerando a demanda da cidade

cresce com a renda das famílias, esta maior demanda será suprida com maior consumo de terra nos limites da cidade, aumentando a área urbana. Já a produtividade da terra agrícola, r_a , tende a aumentar o custo de oportunidade das famílias em transformar área produtiva em área urbana. Se o entorno agrícola torna-se mais produtivo, o aluguel deste entorno sofrerá aumento, o que torna o consumo de espaço por parte das famílias mais caro. Por isso, as famílias reduzem seu consumo de espaço e os limites da cidade tendem a sofrer retração, reduzindo a área urbana da cidade. Os custos de transporte, t , afetam a escolha locacional das famílias ao determinar o custo de acesso ao CDB. Com a redução dos custos de transporte, o consumo de espaço nas franjas da cidade é menor, o que faz famílias consumirem mais espaço nos limites da cidade, expandindo a área urbana.

Brueckner e Fansler (1983) desenvolveram um dos primeiros trabalhos a utilizar o modelo AMM com o objetivo de explicar o crescimento das cidades. Neste trabalho os autores apresentam o background do modelo Alonso-Mills-Muth e chegam a duas conclusões acerca da dinâmica do crescimento das cidades. Primeiro, as evidências sugerem que a alocação de espaço entre uso urbano e rural seria determinada por um coordenado processo de mercado que considera os custos e ganhos de cada tipo de uso do solo. Diferentemente do que era sugerido por outros pesquisadores fora da teoria econômica, as cidades não cresceriam tendo por base um uso indiscriminado do solo agrícola. Segundo, as evidências mostram que o framework básico proposto pela teoria econômica urbana de fato pontua forças capazes de explicar a expansão das cidades, sendo elas: população, renda, renda agrícola e custos de transporte. Para testar o poder preditivo do modelo AMM, os autores utilizam dados de áreas urbanas definidas pelo Censo Norte-Americano de 1970. A base é constituída de 40 áreas urbanizadas com seus respectivos dados de área total em metros quadrados, população, proxy de custo de transporte, renda e rendimento agrícola. A renda agrícola é representada pela mediana do valor da terra agrícola por acre de cada *county* contendo área urbanizada presente na base. Para renda, é utilizada a média das rendas familiares. Para a variável de custo de transporte, foram utilizadas duas proxies. Uma considera o percentual de pessoas com um veículo ou mais sob a hipótese de que quanto mais carros usados, maior o custo de commuting. A segunda usa o percentual de pessoas usando o sistema público de transporte. Para estimar a equação, Brueckner e Fansler (1983) utilizam um estimador de máxima verossimilhança e encontram resultados que estão de acordo com as hipóteses apresentadas no modelo teórico. Os coeficientes da população e renda são positivos e significantes, mas a terra agrícola tem coeficiente negativo e significativo, como predito pela teoria. Isso indica que a área urbana é função crescente da população e da renda, porém decrescente com o valor da terra. Os custos de transportes não tiveram significância encontrada. Este trabalho trouxe evidências iniciais da validade da teoria proposta pelos economistas urbanos.

Posteriormente, devido à disponibilidade de mais dados e com mais períodos de tempo, alguns economistas utilizaram o modelo monocêntrico na tentativa de testar sua

validade em diferentes cidades e métodos de estimação. McGrath (2005) realiza uma reavaliação dos resultados obtidos por Brueckner e Fansler (1983) utilizando uma base de dados mais extensa com informações de regiões metropolitanas norte-americanas. O autor ainda destaca que após o trabalho feito por Brueckner e Fansler (1983), muitos estudos foram feitos com o objetivo de gerar mais evidências empíricas baseadas no modelo AMM e discutir o fenômeno da expansão da área urbana (Brueckner, 2000; Lopez e Hynes, 2003). Esses estudos trouxeram mais solidez à validade das hipóteses do modelo econômico para explicar o advento do crescimento dos aglomerados urbanos. O objetivo do trabalho de McGrath (2005) era testar a robustez do modelo clássico em explicar a dinâmica do espraiamento da área urbana das cidades americanas para regiões metropolitanas com um período de tempo maior. Para testar o modelo, McGrath (2005) utiliza uma base de dados com as 33 maiores regiões metropolitanas do EUA de 1950 até 1990 e suas respectivas medidas de população, renda per capita, custos de transporte e valor da terra agrícola para cada década e cidade. As áreas urbanas são definidas através de um corte de densidade igual ou maior que 1000 pessoas por metro quadrado. Metodologicamente são duas etapas. A primeira utiliza um cross section simples que regride o log da distância das franjas até o centro no log da população, renda, valor da terra agrícola e proxies de custo de transporte. Os resultados mostram os mesmo encontrados por Brueckner e Fansler (1983). População e renda com coeficiente positivos e significantes e valor da terra negativamente correlacionado com o crescimento da área urbana. Os custos de transporte são não significantes estatisticamente. Na segunda, o autor adiciona uma variável de tendência temporal na regressão. Os coeficientes das variáveis explicativas tem mesmo sentido encontrado no primeiro exercício, com exceção dos custos de transportes. O *commuting* passa a ter sinal negativo e estatisticamente significativo, como previsto pela teoria. Quanto à variável temporal, esta aumenta o poder preditivo do modelo, que explica aproximadamente 90% da variação da distância da franja da cidade até o centro. Além disso, a variável temporal é positivamente correlacionada com a expansão das regiões metropolitanas, sugerindo uma tendência sistêmica de expansão da área urbana com o tempo. O autor conclui o trabalho afirmando que os fatores apresentados pela teoria econômica urbana como condicionantes do uso do espaço das cidade são os principais para explicar o tamanho dos aglomerados urbanos. Este resultado corrobora a tese de Brueckner e Fansler (1983), que afirma existir um sistema de mercado bem coordenado que determina o uso do solo como urbano ou agrícola. Ainda, o trabalho comprova a robustez do modelo clássico para diferentes regiões e diversos períodos no tempo.

O trabalho de Spivey (2008) também tem como um dos seus objetivos avaliar a robustez do modelo AMM para bases mais recentes e maiores. O argumento do autor é de que a natureza das cidades havia mudado em relação ao período testado por Brueckner e Fansler (1983). Por exemplo, algumas cidades iniciaram um processo de policentrismo e de mudança nos padrões de commuting. Segundo Spivey (2008), testar a validade

da teoria econômica em explicar cidades mais modernas traria substancial solidez com relação ao poder preditivo do modelo AMM. Neste trabalho, o autor utiliza dados do Censo americano de 2000, e avalia todas as áreas urbanas do EUA, totalizando 452 áreas urbanizadas, bastante superior a bases utilizadas anteriormente. Os dados referentes à população, renda, custos de transporte e valor da terra agrícola são retirados do Censos. Spivey (2008) ainda propõe variáveis alternativas para mensurar o commuting das cidades com o objetivo de melhorar a forma de medir este condicionante. Metodologicamente, um equação Box-Cox e um estimador de máxima verossimilhança são estimados, relacionando área urbanizada em metros quadradas com população, valor médio da terra, mediana da renda familiar e a medida de custos de transporte. Os subcentros estimados por McMillen e Smith (2003) foram adicionados às áreas urbanas para considerar o policentrismo das cidades. O objetivo é observar se o poder de explicação do modelo se mantém ao adicionar esses subcentros, caracterizando uma mudança na natureza das cidades. Os resultados mostram que independentemente da especificação utilizada o modelo continuar a explicar bem as variações de área urbana observadas. Ainda, o modelo continua a explicar bem esse desenvolvimento do uso do solo urbano mesmo considerando toda a transformação existente na dinâmica das cidades modernas em relação as testadas por Bruckener e Fansler (1983). Os coeficientes das variáveis continuam com os sinais esperados e encontrados em estudos mais antigos. Estas evidências mostram o quão robusto é o modelo monocêntrico AMM para explicar a expansão das áreas urbanizadas.

O artigo de Paulsen (2012) utiliza o modelo monocêntrico para estudar as regiões metropolitanas americanas. Um dos objetivos do trabalho é verificar a robustez do modelo AMM, ao comparar com trabalhos anteriores (Bruckener e Fansler, 1983; McGrath, 2005; Spivey, 2008). O autor utiliza um painel de 3 anos com informações de mais de 300 áreas metropolitanas dos EUA. Uma grande contribuição deste trabalho é o uso desta base em painel para testar as hipóteses do modelo clássico. Outra contribuição do trabalho é o uso de dados de satélite como alternativa para medir a extensão das regiões urbanas. O argumento para o uso desta alternativa é de que a densidade populacional - critério mais utilizado para definir a área urbana das cidades - não captaria precisamente a área urbana. Esta imprecisão pode acontecer em locais com baixa densidade populacional, mas com uso urbano intensivo do solo. Por exemplo, locais no interior do país com alta concentração de atividade industrial possuem um intensivo uso urbano do solo, porém a densidade populacional tende a ser mais baixa. Ao utilizar dados de *land cover*, Paulsen (2012) avalia os ganhos desses dados em relação ao uso de faixas de densidade populacional para mensurar a área urbana. Paulsen (2012) também verifica se existem quebras estruturais no modelo condicionado a duas características das cidades. As características são a natureza policêntrica - mais de um subcentro de emprego - e o tamanho - utilizando o limite de 500 mil habitantes - das cidades.

Paulsen (2012) utiliza como variáveis explicativas os condicionantes da área urbana

indicados na teoria urbana, são eles: a renda, a população e o valor da terra agrícola. Metodologicamente, o autor utiliza regressões lineares simples, MQO e o modelo em painel com efeitos fixos de ano e de região metropolitana para que considere a influência de variáveis não observáveis sobre as estimações. Para ambos os modelos, as variáveis dependentes utilizadas foram a área medida através da imagem de satélite e a medida utilizando o corte da densidade populacional. Os resultados encontrados mostram que os coeficientes são consistentes com a teoria econômica independente da forma de mensurar a área urbana. A população e a renda tem relações positivas e significantes estatisticamente com o a área urbana das regiões metropolitanas. O valor da terra agrícola, tem coeficientes negativos nas estimações de MQO e painel. Os resultados de Paulsen (2012) mostram que, independente da forma de medir a área das cidades, os resultados são consistentes. Isto indica que os ganhos do uso de dados de satélite para o caso americano não são consideráveis quando comparados com a densidade populacional. Com relação as quebras estruturais do modelo condicional ao tamanho e policentrismo das cidades, as evidências mostram que existe quebra condicional as características das cidades. Indicando que a diferença de coeficientes existente para regiões metropolitanas com características diferentes não são consequência das amostras e sim destas características. Os resultados mostram que a elasticidades da renda e população são menores para cidades maiores. Com relação ao policentrismo das cidades, a elasticidade da renda é maior para cidades policêntricas, enquanto a elasticidade da população é menor. Em conclusão, o trabalho mostra que embora o modelo AMM seja simples, continua a explicar bem o desenvolvimento da área urbana e que os condicionantes têm magnitudes diferentes a depender da estrutura da cidade.

Dada a crescente disponibilidade de dados trabalhos tem elaborado extensões do modelo AMM considerando outras forças para explicar a dinâmica de crescimento das cidades. Um exemplo destas expansões é o estudo do papel da amenidades na escolha locacional das famílias (Garcia et al, 2018). Outro fator considerado é o crescimento do setor de serviço e da indústria (Deng et al. 2008). Estes resultados tem demonstrado a relevância de considerar mais uma força atuando na escolha locacional das famílias, pró concentração ou pró suburbanização. O sentido da força tem relação com o ganho de utilidade do agente em relação à distância às amenidades. Garcia et al. (2018) mostraram que, no caso das cidades europeias, as amenidades históricas possuem um efeito pró concentração, dado que grande maioria delas estão situadas no centro, atraindo agentes. Inicialmente as investigações feitas sobre a dinâmica do crescimento das cidade foram feitas para o caso das cidades americanas (Paulsen, 2012; Spivey, 2008; Brueckner e Fansler, 1983; McGrath, 2005), porém, mais recentemente extensões do modelo ou o modelo tradicional também foram utilizados para estudar cidades europeias, chinesas ou latinas (Deng et al., 2008; Garcia et al., 2018; Vila Nova, 2014).

Os resultados encontrados para as cidades americanas mostram que o modelo AMM

explica bem a dinâmica do uso do espaço urbano, mesmo sendo um modelo relativamente simples. A população, a renda, custos de transporte e o preço agrícola estão correlacionadas com a dinâmica do crescimento da área urbana. Além disso, os trabalhos apresentados mostram que o modelo continua a explicar bem mesmo com diferentes tipos de metodologias e bases de dados. A boa performance do modelo, fez com que economistas começassem a aplicá-lo em outros países e continentes como Europa, China e Brasil.

3 Estratégia Empírica e Dados

Nesta seção serão apresentadas a estratégia empíricas e os dados utilizados no trabalho. Na primeira seção são detalhados os modelos utilizados e as estratégias utilizadas para melhor estimar os resultados. Na segunda parte, são apresentados os dados referentes aos condicionantes do crescimento urbana e os dados de satélite e suas aplicações.

3.1 MQO e Painel

A metodologia utilizada propõe avanços na aplicação do modelo monocêntrico para o caso brasileiro. Os mesmos procedimentos propostos por Paulsen (2012) serão seguidos, que consiste em utilizar MQO e Painel para estimar os coeficientes das forças estudadas e medir elasticidades para cada uma das relações estatisticamente significantes. Primeiramente, será realizada uma série de exercícios utilizando regressões lineares, MQO, para cada ano censitário. Em segundo lugar, utiliza-se os modelos em painel para controlar por fatores não observáveis, AMCs e Regis utilizando efeitos fixos de tempo e município.

A endogeneidade potencialmente existente entre as variáveis independentes e dependentes apresenta duas origens. O primeiro é o Viés de variável omitida, que ocorre quando não foi incluído alguma variável que estaria correlacionada com a variável dependente ou alguma variável explicativa do modelos. Neste trabalho, por exemplo, algum fator que explica a expansão urbana não estaria sendo considerado no nosso modelo. Como consequência, os coeficientes do modelo estão superestimados ou subestimados, ou seja, estaria atribuindo a determinada variável do modelo uma correlação maior ou menor do que a existente. Ainda mais, pode estar afirmando existir alguma correlação entre as variáveis que na verdade não existe. O segundo problema é a causalidade reversa, onde variáveis independentes afetam a variável dependente e vice versa. No presente trabalho, a expansão urbana pode estar causando crescimento da renda, e a renda, por sua vez, causando a expansão da área urbana. Considerando este problema, o sentido e a magnitude das forças encontradas podem ser diferentes.

Com o objetivo de suavizar os problemas de endogeneidade citados acima, são utilizadas variáveis de controle e o modelo em painel. Os controles tem como objetivo solucionar parte dos problemas de viés de variável omitida. Ao incluir controles, consideram-se possíveis efeitos advindos deles. Os controle utilizados são as macro regiões e o percentual da população residente em aglomerados subnormais. Ao se utilizar controles de macro regiões são considerados os efeitos da dinâmica de cada macro região sobre a área urbana. Por exemplo, devido à estrutura produtiva de muitas cidades do Centro - Oeste depender fortemente da atividade agrícola, espera-se que as cidades desta regiões tendam a ter

menor desenvolvimento da área urbana. Por parte da moradia informal, o efeito encontrado é o contrário. Devido à característica horizontal destas moradias, o mesmo espaço será ocupado por menos pessoas considerando a existência da possibilidade de construções mais verticais no mesmo espaço. Devido a esta horizontalidade dos aglomerados subnormais, espera-se que locais com maior concentração deste tipo de habitação tenha sua área urbana expandida. O modelo em painel, por sua vez, possibilita considerar fatores não observáveis que podem influenciar as estimações. Para considerar esse fatores foram utilizados efeitos fixos de tempo, AMC e Regis. Através do efeito fixo de tempo podem ser captados efeitos advindos de características não observáveis relacionadas com os anos censitários estudados. Por exemplo, a área urbana pode ter sido afetada por determinado ciclo econômico positivo que ocorreu em algum ano de nosso painel. O segundo efeito fixo é o de município ou Regis. Neste caso, o objetivo é considerar o papel de fatores não observáveis do município e das Regis. Através disso consideram-se fatores fixos relacionados às unidades de observações que afetam a área urbana e lidam com parte da endogeneidade existente nas estimações.

Com relação às unidades de observação, foram feitas duas agregações diferentes para os municípios. A primeira agregação foi feita em AMCs, tendo como base a proposta feita pelos pesquisadores do IPEA no estudo de Reis et al. (2008). No Brasil trouxe a criação de novos municípios ao longo do tempo, o que muda a composição territorial destas unidades de observação. Devido a isto, existe a necessidade de criar-se unidades de observações comparáveis a composição territorial existente antes da desagregação de alguns municípios, possibilitando a comparação intertemporal destas unidades de observação. A segunda agregação foi feita tendo como base as Regiões Geográficas Imediatas (Regis) definidas no estudo do IBGE (2017). Estas regiões são definidas pelo IBGE baseado na proximidade dos centros urbanos que satisfazem as necessidades imediatas das populações. Algumas dessas necessidades são compras de bens de consumo duráveis e não duráveis, busca de trabalho, procura por serviços de saúde e de educação, e a prestação de serviços públicos, tais como postos de atendimento do INSS, do Ministério do Trabalho e Emprego e do Judiciário, entre outros. Considerando a interação entre os municípios incluídos na mesma Regis, efeitos sobre o espraiamento da Área Urbana de determinados municípios pode ser advindo de forças vindas de outra região. Considere, por exemplo, um município próximo a um determinado centro de emprego, este mesmo município pode ter sua área urbana aumentada devido ao aumento de trabalhadores no centro de emprego que moram neste município. A agregação pode considerar esses movimentos pendulares entre diversas unidades federativas.

Como citado anteriormente, o principal modelo é o modelo em painel. Formalmente o modelo terá a seguinte forma:

$$AreaUrbana_{it} = \beta_0 + \beta_1 Pop_{it} + \beta_2 Renda_{it} + \beta_3 Agrícola_{it} + \beta_4 \%PopSub_{it} + \theta_t + \gamma_i + \epsilon_{it} \quad (3.1)$$

A variável dependente é a Área Urbana medida em hectares para cada ano t e município ou Regis i da base. Na regressão principal utilizam-se dados de *land cover* como variável independente. As variáveis explicativas seguem a literatura do modelo monocêntrico clássico (Brueckner e Fansler, 1983; McGrath(2005); Spievy (2008); e Paulsen 2012). Baseado nisto, foram incluídos a população urbana, renda domiciliar anual média, rendimento agrícola e o percentual da população residente em aglomerados subnormais dos municípios ou Regis i em cada período de tempo t . Por fim, θ_t é o efeito fixo de cada ano de nosso painel e γ_i o efeito fixo de cada município. Não foram considerados como regressores os custos marginais devido a dificuldade de se obter informações para os diferentes anos censitários.

3.2 Dados

3.2.1 Dados de imagem de satélite

A revolução tecnológica trouxe consigo grande progresso na obtenção de dados para análise econômica. Associado ao uso de métodos de inteligência artificial, como *machine learning* e deep learning, as imagens de satélites passaram a ser insumo de novas informações. Duas informações começaram a ser bastante utilizadas na ciência econômica, são elas: cobertura do solo (Paulsen, 2012; Vila Nova, 2014) e luzes noturnas (Henderson et al., 2012; Hodler e Raschky, 2014; Harari, 2016; Storeygard, 2016; Henderson et al., 2016). Através da informação de cobertura do solo, é possível saber a área de florestas, áreas urbanas ou corpos d'água, como rios. Usando estes dados, várias aplicações vem sendo feitas, como o monitoramento de desmatamentos, saúde dos rios ou espraiamento de cidades. A outra informação citada, luzes noturnas, diz respeito à intensidade de luz captada em determinada área durante a noite. Com esses dados, economistas iniciaram o uso de dados de satélite na análise econômica. A maioria das aplicações utilizando estes dados tentaram mensurar a atividade econômica de algumas áreas através da intensidade de luz, supondo existir uma correlação entre luz e atividade econômica (Henderson et al. 2012; Jean et al., 2016; Keola et al., 2015). Embora essa correlação seja a mais explorada até o momento, a capacidade de extrair novas informações de imagens tiradas do espaço tem sido expandida e estudada.

Donaldson e Storeygard (2016) pontuam três fatores que fundamentam a importância dos dados de satélite para os economistas. Primeiro, o acesso a dados dificilmente coletados através de outros meios. Por exemplo, alguns países africanos apresentam grande dificuldade para executar surveys ou coletar dados de governo, devido a guerras, difícil acesso ou fatores burocráticos. Outro exemplo são as extensões de florestas ou áreas de agricultura, ambos muito custosos de mensurar. Por meio de imagens de satélite, é possível extrair informações sobre lugares a um custo muito baixo, o que possibilita o estudo de

locais ainda não explorados. Segundo, a resolução disponível é muito maior em relação a outras bases disponíveis. Atualmente, existem satélites que podem captar imagens bastante precisas: Engstrom et al. (2017) extraíram informações que exigem muita precisão como o número de carros, tamanho de rodovias e material dos telhados das construções. Terceiro, a vasta cobertura geográfica destes dados. A possibilidade de captar a mesma informação para qualquer região do planeta é algo ainda raro para muitas bases de dados usadas em Economia. Ao utilizar imagens remotas, essa possibilidade torna-se viável e real. Ainda é possível pensar em mais uma vantagem. Esses dados podem ser utilizados como alternativas para testar a robustez de resultados obtidos anteriormente através de bases comumente usadas. Por exemplo, na literatura de Economia urbana alguns autores tem utilizado cobertura da área urbana como alternativa aos dados obtidos através da densidade populacional.

Através do uso dessas informações, muitos estudos tem sido feitos nas mais diversas áreas do ciência econômica. Donaldson e Sterygart (2016) mostram várias aplicações destes dados em pesquisas aplicadas como Burgess et al. (2012) que estuda desmatamento utilizando modelos de competição de Cournot; e Jayachandran (2009) que estuda o impacto da poluição do ar resultante do desmatamento das florestas na mortalidade de crianças e fetos na Indonésia. No caso da ciência econômica, os dados de cobertura do solo e luzes noturnas são os mais utilizados.

Quanto aos dados de coberturas de solo, atualmente, existe um padrão (Delince et al. 2017) que determina a classificação de uso em cinco categorias: Floresta, Formação natural não florestal, Agropecuária, Área não vegetada, Corpo d'água e não observados. Cada uma destas categorias possui subclassificações que discriminam de forma mais detalhada a natureza de cada cobertura. Por exemplo, a Área não vegetada pode ser infraestrutura urbana, mineração ou outras áreas não florestais. Estas subclassificações permitem uma melhor desagregação do tipo de uso do solo da região.

O processo comumente utilizado é o método de aprendizado de máquina (*machine learning*). Através deste processo, as bandas espectrais ou cores emitidas por cada tipo de uso do solo serão associadas a uma das classificações citadas anteriormente. Esta categorização é possível devido ao aprendizado de padrões de outras imagens de satélite utilizadas para treinar o algoritmo de previsão.

O artigo do Paulsen (2012) avalia os ganhos associados ao uso deste tipo de dado para a literatura. Dito de outra forma, o autor procura testar se os resultados encontrados diferem daqueles obtidos utilizando o método clássico de densidade populacional. A motivação dessa comparação é fundamentada na crítica feita ao uso da densidade populacional como indicador de áreas urbanas. Este indicador poderia ser viesado, já que existe o risco de classificar locais urbanizados como rurais simplesmente pela baixa densidade. Por exemplo, regiões com concentração de indústrias dificilmente terão alta densidade, porém

o uso urbano do solo é extensivo. Os resultados do trabalho mostram que o tipo de dado utilizado, densidade ou cobertura de solo, não mudaram os resultados de forma substancial. Para ambos, a expansão da área urbana é positivamente correlacionada com a população e renda, e negativamente relacionada com a renda agrícola.

Para o caso brasileiro, Vila Nova (2014) utilizaram dados de cobertura do solo. Os estudos apontam que os dados de satélite tem resultado similar ao encontrado ao utilizar a densidade populacional como critério de área urbana. Assim como em Paulsen (2012), a população e renda são positivamente correlacionadas com a expansão da área urbana.

Outro tipo de dado utilizado na literatura econômica são as luzes noturnas. Da mesma forma que os dados de cobertura de solo, a intensidade da luz emitida à noite é captada através das bandas espectrais e transformada em uma variável de intensidade da luz atribuída às regiões. A hipótese é que esta intensidade estaria relacionada ao nível de atividade econômica de cada região. Devido a esta suposição, este tipo de informação começou a ser utilizado como proxy para a atividade econômica local (Delince et al. 2017). Henderson et al. (2012) testaram a hipótese de correlação entre luminosidade e atividade econômica para países africanos utilizando dados agregados de PIB para o nível de país. Embora os resultados difiram dos dados oficiais, os autores propõem a abordagem de análise mais desagregadas, ou seja, estudando algumas regiões do país ou continente. Baseado nisto, uma nova literatura se desdobrou e estudou unidades menos agregadas.

Hodler e Raschky (2014) estudaram a preferência de regional de líderes eleitos. Os autores utilizam luzes noturnas em unidades administrativas subnacionais para identificar se as regiões dos líderes eleitos sofrem maior alteração em relação às demais áreas. O resultado mostra que os locais relacionados aos candidatos vencedores demonstram maior crescimento de luminosidade após sua eleição. Harari (2016) utiliza os dados de intensidade de luz a nível de cidade para estudar a geometria urbana das cidades indianas. Storeygard (2016) também utiliza dados a nível de cidade para medir a atividade econômica em seu estudo sobre custos de transporte na África. Por fim, Henderson et al. (2016) utiliza dados de luzes noturnas em grid regionais feitos pelos autores. Esse exemplos mostram o uso mais comum deste tipo de informação, como proxy da atividade econômica das regiões de estudo.

Os dados de *land cover* utilizados neste trabalho foram extraídos da plataforma disponibilizada pelo MapsBiomass. O MapBiomass¹ é uma plataforma criada com a colaboração de órgãos governamentais e privados especializados em processamento de imagens de satélite. O projeto foi criado com o objetivo de gerar informações através de satélites sobre a superfície terrestre brasileira. A plataforma fornece dados anuais sobre uso e cobertura do solo brasileiro. Os dados têm abrangência nacional e podem ser analisados por estado, município e tipo de habitat. O satélite utilizado é o Landsat, mesmo considerando para a

¹ <https://mapbiomas.org/>

criação dos dados usados por Paulsen (2012). A classificação dos dados é feita através de um processo de *Machine Learning*. Neste processo, o algoritmo classifica imagens brutas de satélite em 6 categorias: Floresta, formação natural não-florestal, região agrícola ou agropecuária, Área não vegetada, água e não observada. O processo consiste na identificação das bandas espectrais observadas em fotos de satélites e de suas respectivas coberturas e usos de solo. Por exemplo, a vegetação florestal reflete determinado tipo de espectro de cor, desta forma, o robô pode aprender a associar banda e usos e coberturas do solo. Usando o método citado, o algoritmo utiliza um banco de dados de referência com o qual compara as bandas espectrais e as respectivas classificações de solo. Após criar medidas referentes a bandas de cada pixel, o algoritmo categoriza cada pixel em uma das 6 categorias citadas anteriormente. Após avaliar todos os pixels, um mapa de *land cover* é produzido com todos os pixels categorizados pelo método de previsão de *Machine Learning*.

3.2.2 Condicionantes da área urbana

Para adicionar dados dos condicionantes da área urbana a base de dados foram utilizadas quatro bases. Esta base contém informações de renda, população, rendimento agrícola por hectare. As bases utilizadas são: O atlas do desenvolvimento humano da PNUD, Pesquisa Agrícola municipal (PAM), Pesquisa da pecuária municipal (PPM), e os Censos do IBGE. O painel possui 3 anos 1991, 2000, 2010, com 4265 AMCs do Brasil, 357 Regis com 100 mil habitantes ou mais e 45 Regis para o corte de 500 mil habitantes ou mais. Abaixo são descritas de forma detalhada as fontes de dados.

Uma importante informação referente à base de dados diz respeito as AMCs. No Brasil, alguns municípios foram criados em anos distintos, mudando a composição territorial dos municípios em todo o país. Segundo o estudo de Reis et al. (2008), pesquisadores do Instituto de Pesquisa Econômica Aplicada (IPEA), o Brasil tinha 3.951 municípios em 1970 e devido à criação de novos este número subiu para 5.507 em 2000. Esta nova composição territorial dificulta a comparação de dados intertemporais de tais unidades de estudo. Com o objetivo de resolver este problema, o mesmo estudo propôs a agregação dos municípios em AMCs utilizada no presente trabalho. Através desta agregação foram criadas áreas comparáveis às composições territoriais dos municípios antes da desagregação, devido a isto pode-se estudar municípios intertemporalmente. Na base existem municípios criados após o ano de 1991, sendo necessário o uso do método de agregação por AMCs para os anos 2000 e 2010 tendo como base a Reis et al. (2008). Para isso, foi agregado por AMC valores de renda, população, produção agrícola e área urbana dos municípios que pertencem a mesma.

Outra agregação utilizou como base a divisão em Regis, para esse procedimento utilizou-se a agregação feita pelo IBGE. As Regis têm como objetivo unir municípios interdependentes em aspectos econômicos, sociais e urbanos. Considerando este elo entre

as unidades federativas, os efeitos de municípios em outros podem ser considerados.

O Programa das Nações Unidas para o Desenvolvimento (PNUD) tem dados dos Censos demográficos realizados pelo IBGE. Estas informações podem ser encontrados no Atlas do Desenvolvimento Humano disponibilizado pelo órgão em seu site ². Além de disponibilizar os dados dos censos, a base compatibilizou os dados de anos diferentes para facilitar a obtenção dessas informações. Através da PNUD, foi coletado dados de população total e população urbana para os anos censitários de 1991, 2000 e 2010.

A pesquisa da Produção Agrícola Municipal (PAM) é realizada pelo IBGE anualmente com o objetivo de monitorar a produção agrícola dos municípios. Como citado anteriormente, o rendimento agrícola é considerado no modelo clássico como um dos fatores que afetam a expansão dos limites das cidades. Os dados da PAM foram utilizados com o objetivo de criar uma proxy do rendimento da terra e assim testar a validade da hipótese levantada no modelo AMM para as cidades brasileiras. Através da PAM podem ser obtidos os valores da produção recebidos por cada município em determinada atividade agrícola, área plantada, área destinada ao plantio, área colhida, quantidade produzida e rendimento médio. Algumas das atividades pesquisadas anualmente são: coco, café, abacaxi e cana-de-açúcar. Para construir a variável de rendimento agrícola por hectare foi calculada a razão entre o valor da produção e área plantada ou destinada a colheita. Esta variável de rendimento agrícola criada através da PAM foi intitulada de `pa_pam` nos dados. No caso do valor da produção, foi realizada a conversão dos valores de 1991, em cruzeiros, para reais. Após todos os valores estarem em reais foi utilizado o IPR, índice de preços ao produtor produzido pela FGV, para atualizar todos os valores da produção agrícola para agosto de 2010. Utilizou-se o método de números índices considerando agosto como o mês de referência, atualizando os valores de 1991 e 2000 para 2010.

A Pesquisa da Pecuária Municipal (PPM) é realizada pelo IBGE anualmente e tem como objetivo monitorar a produção agropecuária e atinge todo o território nacional, com informações para o Brasil, Regiões Geográficas, Unidades da Federação, Mesorregiões Geográficas, Microrregiões Geográficas e Municípios. Através dessa base de dados é possível obter a produção e rendimento sobre efetivo dos rebanhos, ovinos tosquiados, vacas ordenhadas, produtos de origem animal e da aquicultura. Essas informações ainda são desagregadas para tipos de animal como bovino, equino e ovinos, para o caso da aquicultura existe a distinção por espécie como Tilápia, Pirarucu ou Camarão. Com as informações da PPM é possível obter a produção total de cada município advindo da atividade agropecuária. Para construir a outra variável de rendimento da terra foi utilizada a soma da produção agrícola - retiradas da PAM - com a da produção agropecuária - retirada da PPM - e dividir pela área de agricultura e pastagem retirada da base de cobertura de solo. Esta variável será intitulada `pt_lc` e usada em testes de robustez. Para os

² <http://atlasbrasil.org.br/2013/pt/consulta/>

valores de produção de 1991, foi realizado o mesmo procedimento utilizado para a Pesquisa da agricultura. Foi convertido os valores em cruzeiros para reais e depois atualizado para o valor de agosto de 2010 através do índice de preços ao produtor (IPR). Para atualizar os valores foi utilizado o método de números índices assim como feito para a PAM.

Os Censos do IBGE disponibilizam informações de renda domiciliar anual e aglomerados subnormais para os municípios. Para a construção da renda domiciliar anual de cada município, foi feita a agregação para AMCs e Regis e posteriormente, calculada a média da renda anual através do Censos de domicílios do IBGE. Para realizar o cálculo do percentual de pessoas residentes em aglomerados subnormais, foram utilizados os Censos de domicílios e pessoas. Primeiramente, foram identificados os domicílios presentes em aglomerados subnormais. Após isto, as pessoas que estavam residindo em moradia informais foram identificadas através do merge entre o censo de pessoas e de domicílios utilizando o id do domicílios. Por fim, calculou-se a razão entre a população em aglomerados subnormais e a população total.

3.2.3 Estatística descritivas

Nesta seção serão discutidos os resultados descritivos da base de dados. Serão apresentadas duas bases de dados. A primeira diz respeito à base de AMCs, com 4265 unidades de observações nos anos de 1991, 2000 e 2010. A segunda base de dados que vai discutir diz respeito as Regis acima de 100 mil habitantes, com 357 Regis nos três anos censitários. A tabela abaixo apresenta o número de observações, média, desvio padrão, valor mínimo e valor máximo para as amostras. As variáveis apresentadas são área urbana, população, renda domiciliar anual, percentual da população residente em aglomerados subnormais e rendimento agrícola em duas especificações, *pa_pam* e *pt_lc*. Esses dados possibilitam compreender de forma mais técnica as características das regiões estudadas e de sua distribuição.

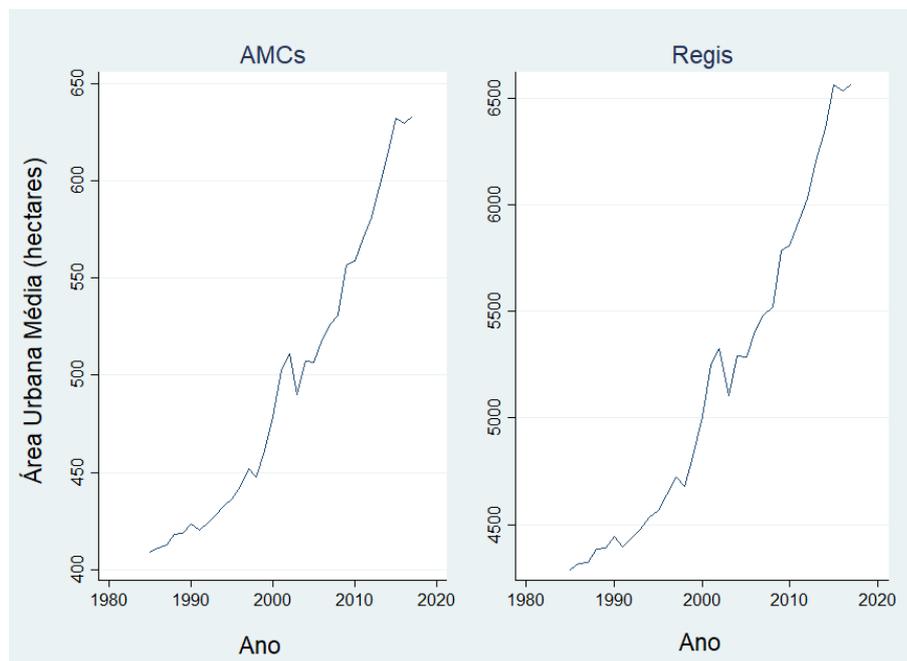
A Tabela 1 representa as estatísticas descritivas da amostra de AMCs e das Regis com população acima de 100 mil habitantes. É possível observar que os rendimento agrícola tem alta dispersão em ambas as especificações, *pt_lc* e *pa_pam*. Isto comprova a variedade de estruturas produtivas que existem no Brasil a respeito do valor da terra. O Brasil tem grande produtividade agrícola em algumas regiões, como a região do Centro - Oeste. Por outro lado, regiões como as Capitais e grandes regiões urbanas como as regiões metropolitanas, tendem a possuir pouca ou nenhuma estrutura produtiva agrícola. Outro fator de destaque é a distribuição de pessoas residentes em aglomerados subnormais. Os dados mostram que grande parte dos AMCs tem baixíssima ou nenhuma presença de aglomerados subnormais segundo os dados do IBGE. Este fato é esperado considerando que a moradia informal é mais presente em grandes centros, parcela relativamente pequena em relação a todos os AMCs o Brasil. Uma última característica importante é a concentração

Tabela 1 – Estatística Descritiva – Amostra de AMCs

	AMCs														
	1991					2010									
	Obs	Média	Desvio Padrão	Min	Máx	Obs	Média	Desvio Padrão	Min	Máx					
Área Urbana (ha)	4167	418,33	1928,35	0,248	81516,8	4198	472,88	2004,47	0,16873	83932,62	4210	551,02	2135,53	0,442815	84213,12
População Urbana	4265	25626,5	187252	254	9419100	4265	31835,02	203145,2	362	9816138	4265	37077,02	228239,1	427	11152344
Rendimento Agrícola - pa_pam (R\$/ha)	4237	7,53	138,91	0	7641,63	4228	4,75	19,89	0	856,75	4202	4,75	12,39	0	482,89
Rendimento Agrícola - pt_lc (R\$/ha)	4254	9,42	152,48	0	7715,56	4245	5,96	24,26	0	884,25	4227	5,86	16,57005	0	593,15
Renda domiciliar (R\$ de 2010)	4265	2302,26	1367,42	429,81	12953,52	4265	4132,74	2357,72	837,45	21298,67	4265	5873,70	2847,54	1439,96	24330,23
% População subnormal	4265	0,74	3,29	0	85,88	4265	0,28	2,03	0	55,15	4265	0,61	3,39	0	61,19
Regis com 100 mil habitantes ou mais															
	1991					2010									
	Obs	Média	Desvio Padrão	Min	Máx	Obs	Média	Desvio Padrão	Min	Máx	Obs	Média	Desvio Padrão	Min	Máx
Área Urbana (ha)	357	4540,47	11292,89	33,31	160139,2	357	5167,35	11974	254,59	167996,8	357	6013,90	12922,88	274,06	175378,2
População Urbana	357	291703,2	1000841	23910	15112086	357	361256,8	1144824	35090	17119104	357	419904	1292570	38623	19458888
Rendimento Agrícola - pa_pam (R\$/ha)	357	2,36	3,76	0,06	39,01	357	3,26	5,15	0,06	57,79	357	3,77	5,03	0,01	49,44
Rendimento Agrícola - pt_lc (R\$/ha)	357	2,91	4,43	0,12	53,62	357	3,88	5,62	0,12	66,46	357	4,45	5,29	0,10	51,45
Renda domiciliar (R\$ de 2010)	357	2287,52	1174,51	696,16	5751,22	357	4057,56	2048,30	1179,67	9425,76	357	5704,87	2479,52	1742	12488,97
% População subnormal	357	1,38	3,27	0	25,13	357	0,67	2,32	0	25,01	357	2,51	5,78	0	53,40

Fonte: elaboração própria.

Figura 2 – Evolução da Área Urbana média para AMCs e Regis. Fonte: Elaboração própria.



da distribuição da renda domiciliar em faixa de renda baixa nas duas amostras.

A figura 2 mostra a evolução da área urbana média para as AMCs e Regis com 100 mil habitantes ou mais. Para ambas as amostras pode ser vista a tendência de crescimento da área média. A variação percentual entre as amostras é semelhante. Entre 1991 e 2010, a área urbana cresceu 32,96 % nas AMCs e 32,30% nas Regis. Entre 1991 e 2000, o crescimento nas AMCs foi de 13,86%, enquanto que nas Regis foi de 14%. Por fim, entre 2000 e 2010, a área das cidades cresceu 16,77% nas AMCs e 16,04% nas Regis. Este aumento da área das cidades é consistente com as elevações da renda e população observadas na tabela 1.

4 Resultados

Nesta seção serão discutidos os resultados obtidos através da aplicação da estratégia empírica. Esta seção terá três partes. Na primeira, serão discutidos os resultados obtidos através das regressões de painel e OLS para a amostra de AMCs. Na segunda seção, serão apresentados os resultados para a amostra de Regis com mais de 100 mil habitantes, seguindo o mesmo procedimento feito para os AMCs. Por fim, são apresentados os resultados das elasticidades estimadas utilizando o modelo de MQO e painel para AMCs e Regis. Para as estimações de OLS e Painel, a estratégia de Paulsen (2012) foi seguida utilizando o *land cover* da área urbana como variável dependente, e a população, renda domiciliar, rendimento agrícola e o peso da população vivendo em setor subnormal como variáveis explicativas.

4.1 AMCs

As variável dependente utilizada nestas especificações é a cobertura de solo da área urbana em hectares medida através do processamento de imagens de satélites feito pelo projeto MapBiomass. Como variáveis independentes foram utilizadas a população urbana dos AMCs retiradas dos Censos, a renda anual domiciliar média retirada dos Censos e o rendimento agrícola construído com os dados da PAM. O rendimento agrícola foi construído através da razão entre o rendimento advindos da atividade agrícola recebidos pelos AMCs e a área plantada ou destinada ao plantio, encontrados na Pesquisa da Produção Agrícola Municipal realizada pelo IBGE.

Os resultados obtidos através do método de MQO para os anos de 1991, 2000 e 2010 podem ser vistos na Tabela 3. Nesta tabela existem três colunas para cada ano censitário. A primeira coluna contém os resultados do modelo sem controles, apenas com população, renda e rendimento agrícola explicando a área urbana. A segunda coluna apresenta estimativas do modelo com controles de macro regiões. Por fim, a terceira coluna adiciona a porcentagem de pessoas residindo em aglomerados subnormais no modelo.

Os resultados mostram que a população tem relação positiva e estatisticamente significativa com o crescimento da área urbana para todos os anos estudados. A magnitude dos coeficientes da população indicam a correlação entre 1 residente adicional nas AMCs e o crescimento da área urbana. Para o ano de 1991, um habitante adicional está relacionado com 0,00919 hectares, ou, 91,9 m^2 adicionais de área urbana. Para o ano de 2000, o coeficiente indica aumento de 90 m^2 na área urbana. Por fim, em 2010 a correlação indica 99,9 m^2 adicionais de área urbana.

A renda domiciliar é positivamente correlacionada com a expansão da área urbana das cidade em todos os anos. Além disso, as relações são estatisticamente significantes. Sobre a magnitude desta força no modelo com controles, foi encontrado o seguinte: Para o ano de 1991, 1 real a mais na renda domiciliar das famílias está correlacionado ao aumento de 0,186 hectares na área urbana, ou, 1860 m^2 . Para o ano de 2000 a relação é de 0,108 hectares, ou, 1080 m^2 . No ano de 2010, o valor é de 0,0914 hectares, 914 m^2 adicionais.

Com relação ao rendimento da terra agrícola observa-se relação negativa nos anos de 2000 e 2010. Porém, apenas nas especificações com controles no ano 2010 foram encontradas relações estatisticamente significantes. O sentido das forças é o esperado pela teoria urbana, indicando que a produtividade da terra tende a reduzir a expansão urbana. Para o ano de 1991, todas as especificações indicam relação positiva, porém insignificante. A magnitude do coeficiente do rendimento agrícola no ano de 2010 sugere que R\$ 1 adicional na produtividade da terra por hectare está relacionada com a redução de 1,117 hectares da área urbana.

Sobre o percentual de pessoas residentes em aglomerados subnormais, os resultados mostram coeficientes positivos em todos os anos, indicando que o maior número de pessoas residindo em moradias informais está correlacionado com cidades mais espalhadas. Embora todos os sinais tenham o sinal esperado, apenas nos anos de 1991 e 2000 foi encontrado significância estatística. A magnitude dos coeficientes indicam que 1% a mais de residentes em aglomerados subnormais estaria correlacionado com 25,61 ha adicionais na área urbana no ano de 1991. Já para o ano de 2000 o aumento da área urbana é de 41,40 ha. O R^2 dos modelos de MQO para a amostra de AMCs e AMCs varia entre 0,860 e 0,923 ao longo dos anos, indicando que o modelo explica aproximadamente 85% das variações dos dados.

Em relação à literatura de economia urbana, os resultados obtidos para a população e a renda das famílias corroboram as evidências encontradas em outros países e cidades. Como já exposto anteriormente, a teoria mostra que o aumento dessas forças resultaria na expansão da área das cidades. Sobre o rendimento agrícola, os resultados obtidos para os anos de 1991 e 2000 sem controle, são contrários ao que a teoria propõe. Como visto anteriormente, o modelo AMM, sugere que o rendimento da terra agrícola afetaria os limites da cidade ao restringir o crescimento da cidade de forma negativamente correlacionada. Dito de outra forma, maior produtividade agrícola restringiria a expansão da área urbana. Apesar deste anos apresentar sinais não esperados para a especificação sem dummies de macro regiões, as demais estimações para os anos censitários mostram sinais esperados. Ao utilizar o controle de macro regiões, o rendimento agrícola passa a ser negativamente correlacionado com o crescimento da área urbana, resultado que está de acordo com as evidências e teoria vistos na literatura econômica. Sobre o percentual da população em serores subnormais, foi encontrado o resultado esperado. Uma possível explicação para o crescimento da área urbana ocasionado pela moradia informal é a característica

Tabela 2 – OLS - Amostra de AMCs utilizando pa_pam - Variável dependente é a Área Urbana em hectares

Variáveis	(1) 1991	(2) 1991	(3) 1991	(4) 2000	(5) 2000	(6) 2000	(7) 2010	(8) 2010	(9) 2010
Renda Domiciliar	0,148*** (0,0150)	0,195*** (0,0229)	0,186*** (0,0212)	0,0807*** (0,00802)	0,112*** (0,0134)	0,108*** (0,0125)	0,0678*** (0,00820)	0,0927*** (0,0162)	0,0914*** (0,0153)
População	0,00931*** (0,000542)	0,00925*** (0,000526)	0,00919*** (0,000499)	0,00916*** (0,000489)	0,00910*** (0,000472)	0,00900*** (0,000449)	0,0102*** (0,00106)	0,0101*** (0,00105)	0,00999*** (0,00107)
Rendimento Agrícola (pa_pam)	0,0415 (0,0306)	0,0239 (0,0214)	0,0184 (0,0139)	0,343 (0,297)	-0,322 (0,389)	-0,313 (0,396)	-0,426 (0,403)	-1,201** (0,575)	-1,117** (0,533)
% Pop subnormal			25,61** (10,02)			41,40** (18,04)			12,71 (13,91)
Constant	-170,5*** (26,55)	-248,4*** (61,18)	-233,5*** (58,31)	-155,5*** (25,76)	-258,6*** (65,04)	-239,1*** (61,67)	-213,8*** (28,88)	-289,8*** (91,42)	-280,4*** (85,89)
Observations	4,140	4,140	4,140	4,162	4,162	4,162	4,149	4,149	4,149
R-squared	0,910	0,914	0,916	0,918	0,921	0,923	0,860	0,864	0,864
Macro	NO	YES	YES	NO	YES	YES	NO	YES	YES

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

Tabela 3 – Painel - Amostra de AMCs utilizando pa_pam - Variável dependente é a Área Urbana em hectares

Variáveis	(1)	(2)	(3)	(4)
Renda Domiciliar	0,0146*** (0,00193)	0,0178*** (0,00253)	0,0170*** (0,00303)	0,0169*** (0,00297)
População	0,00882*** (0,000416)	0,00669*** (0,00129)	0,00669*** (0,00130)	0,00668*** (0,00131)
Rendimento Agrícola (pa_pam)	-0,0114*** (0,00409)	-0,0129** (0,00488)	-0,0135*** (0,00460)	-0,0134*** (0,00449)
% Pop subnormal				1,477 (1,848)
Constant	139,4*** (27,65)	184,7*** (29,71)	190,8*** (29,88)	190,2*** (29,54)
Observations	12,451	12,451	12,451	12,451
R-squared	0,515	0,519	0,521	0,521
Number of Codigo	4,197	4,197	4,197	4,197
EF AMCs	NÃO	SIM	SIM	SIM
EF Ano	NÃO	NÃO	SIM	SIM

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

predominantemente horizontal destas habitações. Esta característica horizontal faz com que menos pessoas ocupem o mesmo espaço, comparado a construções mais verticais como prédios. Por isso, a cidade tende a ter seus limites expandidos para acomodar mais residente. Considerando a forte presença deste tipo de habitação no Brasil, é de fundamental importância entender o papel das mesmas na dinâmica das cidades brasileiras.

A seguir, serão apresentados os resultados obtidos com o método de painel. Esta abordagem econométrica possibilita o uso de efeitos fixos para considerar a influência de não-observáveis fixas no tempo relacionadas aos anos censitários e AMCs de nossa base. Ao considerar estes efeitos fixos no tempo, parte do viés ocasionado por fatores não observáveis dos AMCs e tempo em nossas estimações será resolvido.

A variável dependente continua sendo a área urbana calculada pelo MapBiomias em hectares para os AMCs do Brasil. As variáveis explicativas também são as mesmas utilizadas na abordagem MQO, a população urbana, renda domiciliar anual média e o rendimento agrícola. Os resultados são apresentados na Tabela 3, a seguir. Na primeira coluna da tabela foi utilizado um painel sem efeitos fixos, na segunda coluna foram adicionados efeitos fixos de AMCs, na terceira e quarta foi incluído efeitos fixos de tempo e AMCs. Para considerar a possível correlação entre nossas observações, os AMCs, foram utilizados clusters de Estado para estimar os erros de nossa estimação.

Os resultados do painel corroboram com as evidências encontradas anteriormente através do MQO. A população e a renda são positivamente correlacionados de maneira estatisticamente significativa em todas as especificações. Ou seja, com a expansão da área urbana dos AMCs. O rendimento agrícola tem sinal negativo e estatisticamente significativo, indicando que a produtividade da terra restringe o crescimento da cidade. A influência da moradia informal é positiva, indicando o mesmo sentido encontrado na MQO, porém, essa relação não é significativa. Esses resultados estão de acordo com o encontrado em estudos feitos em outros países, como o de Paulsen (2012). Outro fator importante a ser comentado é a diferença entre os coeficientes da MQO e painel. Utilizando o modelo em painel encontramos coeficientes menores em relação aos modelos de MQO, este fato também é observado nos resultados obtidos por Paulsen (2012). Isto ocorre devido aos efeitos fixos, que consideram a influência de fatores fixos no tempo sobre a nossa variável dependente. O R^2 do modelo em painel varia entre 0,515 e 0,521, indicando que este modelo explica aproximadamente 52% das variações dos dados.

A magnitude dos coeficientes, como mencionado anteriormente, são menores. Para a renda domiciliar, os resultados indicam que, tudo o mais constante, R\$ 1 adicional na renda anual domicilia está correlacionado com 0,0169 ha adicionais na área da cidade, ou 169 m^2 . Para a população, um residente a mais está relacionado a 0,0066 ha adicionais ou 66 m^2 . Por fim, o coeficiente do rendimento agrícola indica que R\$ 1 a mais na produtividade da terra por hectare está relacionado a redução de 0,0134 ha, ou, 134 m^2 na área da cidade.

4.2 Regis

Nesta seção, serão apresentados os resultados para a base de AMCs agregada em Regis. O motivo pelo qual serão utilizados esses aglomerados é a existência do movimento pendular entre os AMCs. Ou seja, alguns moradores de diferentes cidades podem participar de atividades econômicas e sociais em outras cidades. Por exemplo, trabalhadores de determinado centro de emprego podem residir em AMCs vizinhos, o que faz com que as dinâmicas urbanas dessas cidades sejam correlacionadas. Essa correlação intra regional faz com que as agregação em AMCs seja menos ajustada ao nosso modelo teórico. A aglomeração em Regis foi definida pelo IBGE considerando principalmente a interação entre os AMCs na suas estruturas urbanas e econômicas. Algumas destas iterações são a provisão de bens de consumo, acesso a trabalho, acesso a serviços públicos, entre outros. Ao considerar unidades mais agregadas foi considerada a interação econômica e social existentes entre os AMCs, tem-se melhor ajuste as observações ao modelo teórico. Esta abordagem é semelhante a proposta por McGrath (2005) que utiliza as regiões metropolitanas norte-americanas. Vale destacar que não foram utilizadas as Regiões metropolitanas como unidade de observação devido ao pequeno número das mesmas. Segundo o IBGE, atualmente existem 74 regiões metropolitanas em todo o Brasil. Porém,

em anos censitários anteriores o número é ainda menor, são 9 RMs em 1991, 22 em 2000 e apenas 34 em 2010.

Para as Regis, foram utilizados cortes de tamanho para avaliar a validade da teoria condicionado ao tamanho das regiões. O modelo principal utiliza a amostra de Regis com 100 mil habitantes ou mais, mas também foram testadas as amostra entre 100 mil e 500 mil habitantes, e por fim, 500 mil habitantes ou mais. Aqui foi replicada a estratégia seguida por Paulsen (2012), que utiliza de diferentes cortes de população e avalia os coeficientes das forças que estão sendo estudados, população, renda e rendimento agrícola. Para o autor, cidades maiores apresentam mais subcentros do emprego, o que pode gerar resultados diferentes referentes aos condicionantes da dinâmica da área urbana. No artigo, Paulsen (2012) utiliza o limite de 500 mil habitantes para definir regiões metropolitanas grandes e pequenas: acima de 500 mil são definidas como grandes, abaixo são definidas como pequenas. O objetivo desses cortes é testar se existem mudanças estruturais na dinâmica de crescimento da área urbana condicional à mudança na estrutura das cidades. Através dos resultados pode-se avaliar a diferença entre a magnitude das forças e os sentidos encontrados para diferentes tamanhos de cidades. Com relação aos erros estimados no modelo de painel, utilizou-se a estratégia de clusterizações por macro regiões para considerar correlações entre as Regis presente na mesma região.

Para verificar se existe quebra estrutural nas Regis com 500 mil habitantes ou mais, foi executado o teste de Chow, assim como feito por Paulsen (2012). Neste teste, foi realizado um pooling dos dados em painel para a amostra de Regis acima de 100 mil habitantes. As variáveis explicativas, população, renda domiciliar, renda agrícola e percentual da população subnormal foram iteradas com a dummy criada para Regis com 500 mil habitantes ou mais. Após realizar a regressão, o teste F foi utilizado para verificar se todos os coeficientes são conjuntamente diferentes de zero, indicando quebra estrutural. O resultado do teste é significativo a 1%, indicando que existe quebra estrutural no modelo para a Regis grandes.

Seguindo com a apresentação do resultados, primeiramente serão expostos os resultados obtidos através da MQO. Os resultados são apresentados na Tabela 4, a seguir. Assim como no caso dos AMCs são apresentadas três colunas para cada ano censitário. A primeira coluna consiste do modelo básico sem controles, apenas com renda, população e rendimento agrícola explicando área urbana. A segunda coluna introduz os controles de macro regiões e na terceira, foi incluído a variável de aglomerados subnormais, que será explicada posteriormente.

Como pode ser visto na Tabela 4, para cada ano censitário, os coeficientes da população e renda domiciliar têm sinal esperado quando estimado o modelo utilizando as Regis. A população e renda têm sinais positivos e significantes em todas as especificações e anos utilizados em nossa pesquisa. As magnitudes dos coeficientes são semelhantes ao

Tabela 4 – OLS - Regis com 100 mil habitantes ou mais utilizando pa_pam - Variável dependente é a Área Urbana em hectares

Variáveis	(1) 1991	(2) 1991	(3) 1991	(4) 2000	(5) 2000	(6) 2000	(7) 2010	(8) 2010	(9) 2010
Renda Domiciliar	0,837*** (0,113)	0,836*** (0,178)	0,800*** (0,169)	0,571*** (0,0701)	0,628*** (0,127)	0,616*** (0,125)	0,615*** (0,0729)	0,682*** (0,154)	0,655*** (0,150)
População	0,0107*** (0,000291)	0,0108*** (0,000297)	0,0106*** (0,000222)	0,00999*** (0,000296)	0,0100*** (0,000300)	0,00981*** (0,000252)	0,00946*** (0,000463)	0,00948*** (0,000468)	0,00940*** (0,000460)
Rendimento Agrícola (pa_pam)	107,6 (106,3)	106,6 (110,2)	61,84 (88,42)	5,793 (13,67)	-16,47 (11,60)	-42,05** (21,29)	14,34 (22,58)	-10,27 (20,89)	-15,88 (19,86)
% Pop subnormal			19,192* (10,011)			20,564* (12,048)			5,976 (6,292)
Constant	-760,8*** (258,9)	258,7 (672,9)	211,5 (667,0)	-776,8*** (245,9)	-129,5 (784,2)	32,17 (775,8)	-1,524*** (346,5)	-689,4 (1,178)	-509,8 (1,159)
Observations	357	357	357	357	357	357	357	357	357
R-squared	0,954	0,958	0,961	0,958	0,961	0,962	0,948	0,951	0,951
Macro	NO	YES	YES	NO	YES	YES	NO	YES	YES

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

longo dos anos e em diferentes especificações. Como já dito anteriormente, este é o sentido esperado tendo em vista as evidências observadas na literatura. Com relação ao rendimento agrícola, espera-se sinal negativo para esta variável, indicando sentido contrário a expansão da área urbana. No caso das Regis, os resultados pouco diferem qualitativamente dos vistos para AMCs. Primeiramente, nos anos de 2000 e 2010 foi encontrado sinal negativo para o valor da terra após a introdução das dummies de macro regiões, assim como visto nas amostra de AMCs. A única diferença em relação à amostra de AMCs se dá no ano onde existe significância estatística para o rendimento agrícola. Para a amostra de Regis, a relação é significativa em 2000, para as AMCs, a relação é significativa em 2010. Para o ano de 1991, em ambas as amostras não foi encontrado sinal negativo entre a produtividade da terra e a área urbana. Para a variável mensurando a moradia informal, existe correlação positiva, indicando que a presença dessa tipo de habitação tende a espriar a área urbana das cidades. Como dito para os AMCs, a característica horizontal desse tipo de moradia faz com que menos pessoas possam ocupar o mesmo espaço. Com isso, os limites da cidade devem expandir para acomodar mais pessoas.

A seguir serão apresentados os resultados para a amostra de Regis utilizando a abordagem de painel. Na tabela 5, os modelos utilizados são similares ao caso das AMCs. Na primeira coluna de cada especificação o painel foi feito sem efeitos fixos, na segunda coluna foi adicionado efeitos fixos de AMCs, na terceira foram adicionados efeitos fixos de tempo e AMCs, por fim, na quarta coluna foi adicionado o percentual da população em setores subnormais.

Sobre os resultados obtidos, primeiramente serão discutidos os resultados para as Regis com 100 mil habitantes ou mais, após isto, as Regis entre 100 mil e 500 mil habitantes, e por fim, com 500 mil habitantes ou mais.

Os resultados obtidos para através do painel utilizando 100 mil habitantes com corte segue os resultados esperados. A população e a renda domiciliar são positivamente correlacionados com a expansão da área urbana, sendo essas relações estatisticamente significantes. O rendimento agrícola, por sua vez, tem relação negativa mas não significativa.

O corte entre 100 e 500 mil habitantes tem sinais positivos e significantes para a população e renda domiciliar. Para a renda agrícola não foram encontrados sinais significantes, o sinal da relação é positivo, contrário ao esperado pela literatura. Com relação a população subnormal, o sinal encontrado também não é o esperado, negativo, porém a relação não é significativa.

Ao utilizar a amostra mais restrita, com o corte de 500 mil habitantes, a população e renda continuam com o mesmo sinal positivo e significativo estatisticamente. O rendimento agrícola, mesmo que continue com sinal negativo, perde significância estatística em relação ao primeiro corte. O efeito do percentual da população vivendo em setores subnormais é positivo, mas a relação é insignificante. O R^2 dos modelos em painel é menor que os vistos

para o modelo MQO, variando entre 0,633 e 0,775, indicando que o painel explica mais de 63% das variações dos dados utilizados.

4.3 Elasticidades

Para que possam ser comparados os resultados do presente trabalho com outros feitos anteriormente e visando a melhor interpretação das forças em estudo, foram calculadas as elasticidades de cada variável explicativa. Foi utilizado as médias amostrais de cada variável para cada ano censitário. Este cálculo foi feito para todas as variáveis com relações significantes em nosso painel e MQO. A elasticidade nos informa o quanto a área urbana varia percentualmente dado a variação de 1% na variável explicativa. Por exemplo, para o ano de 1991 a variação de 1% na renda domiciliar anual média está correlacionado com um aumento de %1,02 na área urbana.

As elasticidades foram calculadas para todos os coeficientes estatisticamente significantes das estimações MQO. Na tabela 6 são apresentadas as elasticidades medidas para os coeficientes que tem relação estatisticamente significativa. Seguindo Paulsen (2012), foram incluídos informações com elasticidades encontradas em outros trabalhos para que possam ser comparados com os resultados encontrados no caso brasileiro. Para as elasticidades da população urbana foram encontrados resultados próximos aos encontrados por Paulsen (2012) utilizando densidade populacional ou dados de satélite para mensurar área urbana, e por McGrath (2005). Com relação as trabalhos de Spivey (2008) e Brueckener e Fansler (1983) os coeficientes são menores. Sobre a renda domiciliar a elasticidade são bem próximas as de Paulsen (2012) quando utilizou dados de satélite para medir a área urbana. Porém, em relação a todos os outros estudos mostram coeficientes muito menores ou muito maiores. Para o rendimento agrícola foi calculado apenas o ano de 2010 onde é possível verificar o coeficiente significativo, porém esta elasticidade está muito abaixo da medida em outros estudos.

Na tabela 6 as elasticidades calculadas utilizando nossa amostra de Regis com 100 mil ou mais habitantes também são apresentadas. Os coeficientes da variável de população urbana mudam muito pouco em relação aos resultados vistos para os AMCs, sendo bastante próximos dos trabalhos feitos anteriormente. Porém os valores da renda domiciliar mudam bastante, tendo uma queda de aproximadamente 0,5 para todos os anos censitários. Em relação aos estudos elencados na mesma tabela, este coeficiente é significativamente menor. O rendimento agrícola continua com coeficiente bastante inferior ao encontrado em trabalhos anteriores.

Os coeficientes do modelos em painel são apresentados para as amostras de AMCs, Regis com 100 mil habitantes ou mais e com 500 mil habitantes ou mais. Ao lado dos resultados obtidos neste trabalho, foram adicionados os resultados obtidos por Paulsen

Tabela 5 – Painel para amostra de Regis - Variável dependente é a Área Urbana em hectares

	100k ou mais			entre 100k e 500k			500k ou mais						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
Renda Domiciliar	0,160*** (0,0328)	0,263*** (0,0457)	0,323*** (0,0569)	0,320*** (0,0578)	0,0959*** (0,0162)	0,123*** (0,0164)	0,197*** (0,0243)	0,196*** (0,0244)	0,231* (0,130)	0,769*** (0,181)	0,879** (0,327)	0,885** (0,329)	
População	0,00868*** (0,000386)	0,00497*** (0,000811)	0,00500*** (0,000829)	0,00496*** (0,000838)	0,0123*** (0,00113)	0,00976*** (0,00114)	0,0108*** (0,00115)	0,0111*** (0,00114)	0,00817*** (0,000349)	0,00358*** (0,000713)	0,00365*** (0,000740)	0,00364*** (0,000746)	
Rendimento Agrícola (pa_pam)	-1,735 (16,14)	-16,36* (8,238)	-15,58 (9,175)	-15,15 (9,590)	-0,600 (8,085)	1,137 (6,348)	1,634 (5,970)	0,921 (5,560)	-9,577 (43,85)	-26,78 (20,02)	-26,21 (26,28)	-27,21 (28,50)	
% Pop subnormal				18,31 (16,53)				-12,69 (7,971)				17,77 (37,84)	
Constant	1,497*** (302,6)	2,458*** (176,3)	2,373*** (188,2)	2,359*** (187,0)	503,1*** (168,0)	790,5*** (154,4)	557,3*** (165,0)	527,7*** (160,0)	6,720*** (1,389)	12,167*** (856,5)	11,967*** (1,039)	11,871*** (1,005)	
Observations	1,071	1,071	1,071	1,071	936	936	936	936	135	135	135	135	
R-squared	0,633	0,671	0,677	0,678	0,750	0,754	0,773	0,775	0,636	0,721	0,731	0,732	
Number of cod_rgi	357	357	357	357	312	312	312	312	45	45	45	45	
EF Regis	NÃO	SIM	SIM	SIM	NÃO	SIM	SIM	SIM	NÃO	SIM	SIM	SIM	
EF Ano	NÃO	NÃO	SIM	SIM	NÃO	NÃO	SIM	SIM	NÃO	NÃO	SIM	SIM	

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

Tabela 6 – Elasticidades com coeficientes das estimações MQO e Painel

	MQO												Painel						
	AMCs		Regis (100k)		Paulsen (2012) - Densidade		Paulsen (2012) - Satélites		Spivey (2008)		McGrath (2005)		Brueckner e Fansler (1983)		Área Urbana		Paulsen (2012)		
	1990	2010	1990	2000	1990	2010	1980	1990	2000	2001	2000	1950-1990	1970	100k ou mais	Entre 100k e 500k	500k ou mais	Amostra Completa	Abaixo de 500k	Acima de 500k
População Urbana	0,56	0,60	0,67	0,68	0,68	0,65	0,63	0,62	0,63	0,58	0,91	0,76	1,08	0,33	0,62	0,28	0,80	0,98	0,83
Renda domiciliar	1,02	0,98	0,97	0,40	0,48	0,62	1,56	1,51	1,48	0,83	-0,39	0,33	1,49	0,25	0,27	0,19	0,39	0,20	0,31
Rendimento Agrícola	-	-	-0,00	-	-0,02	-	-0,22	-0,25	-0,24	-0,25	-0,03	-0,1	-0,23	-	-	-	-	-	-

Fonte: Elaboração própria

(2012) em suas estimações de painel para que possam ser comparados aos obtidos neste trabalho. Em cada coluna existem as diferentes amostras utilizados pelo autor em seu painel, a amostra completa, a amostra de regiões metropolitanas acima de 500 mil habitantes e abaixo de 500 mil habitantes.

O fato das elasticidades serem menores para o Brasil em ambos condicionantes pode ser relacionado com fatores particulares do país. Com relação à elasticidade da população, as elasticidades são menores em relação ao caso americano, variando entre 0,33 e 0,62, enquanto que são observados, entre 0,80 e 0,98 para o caso americano. Como defendido na teoria urbana, os custos de transportes fazem com que seja mais barato trocar acessibilidade por espaço nas franjas da cidade. O fato de o Brasil não estar equipado com infraestruturas urbanas de qualidade parece fazer com que o movimento de espraiamento seja retraído. Além disso, as ruins condições de mobilidade urbana podem colaborar com a concentração dos habitantes em espaços cada vez mais aglomerados. Estas podem ser forças que retraem a expansão da área urbana das cidades, refletindo assim elasticidades menores.

Com relação à elasticidade da renda, a variação está entre 0,14 e 0,27 para o caso brasileiro e variações entre 0,20 e 0,39 para o caso americano. A desigualdade de renda no Brasil e o menor nível de renda quando comparado aos Estados Unidos podem restringir a magnitude das elasticidades da renda para os múltiplos cortes populacionais utilizados. Considerando os dois fatores particulares das cidades brasileiras, incrementos de renda podem não resultar em aumentos consideráveis na demanda por espaço por parte das famílias.

Outra consideração importante diz respeito à falta de insignificância da elasticidade da renda agrícola. Um possível motivo para este resultado é a estrutura produtivas de muitas cidades brasileiras que dependem substancialmente da atividade agrícola. Especialmente no Centro - Oeste e Sul do país, cidades são dinamizadas graças a atividade agrícola, o que faz com que esta força tenha efeitos contrários ao esperado. Ou seja, o rendimento agrícola poderia possibilitar crescimento da área urbana devido a renda advinda da produtividade da terra.

Ainda com relação à elasticidades, os resultados sugerem que a elasticidade da população das Regis menores é maior em comparação às das Regis maiores, fato observado no trabalho de Paulsen (2012). Porém, ao contrário do observado no caso americano, a elasticidade de renda é maior em Regis menores. São sugeridas duas explicações para estes fenômenos. Primeiro, estes resultados podem ter relação com a diferença de tempo de *commuting* existente entre as Regis menores e maiores. Dado o custo de *commuting* nas maiores Regis, a expansão da área urbana seria dificultada devido ao custo de acessibilidade enfrentado pelos agentes. Os dados do Censo IBGE 2010 sobre *commuting* mais elevado nas Regis fortalecem nossas hipóteses. Segundo os dados, as Regis maiores tem tempo médio

de transporte mais longo, de 28,4 min, em relação as menores, com 20,8 min, corroborando com nossas hipóteses. Em segundo lugar, a desigualdade maior em Regis maiores faz com que o aumento por demanda de espaço na cidade ocasionado pelo crescimento de renda das famílias seja menor. Isto pode ocorrer devido ao fato de que esses incrementos de rendas são concentrados em poucas famílias. Os dados de índice de Gini para as Regis em 2010 indicam a existência de maior desigualdade nas Regis maiores. Este índice mede o nível de desigualdade de renda existente nas unidades de observação, quanto mais próximo de 1, maior a desigualdade. O Gini em Regis maiores é de 0,5473, maior quando comparada com a menores, que é de 0,5118. Esta evidência está, assim, em linha com nossa terceira hipótese para explicar a menor elasticidade da renda. A validade dessas hipóteses deve ser verificada de forma mais profunda em estudos futuros.

5 Considerações Finais

Os economistas formalizaram teorias acerca das forças relacionadas à dinâmica de desenvolvimento da área urbana das cidades. O modelo mais utilizado e consolidado na literatura é o desenvolvido por Alonso, Mills e Muth (Alonso, 1964; Mills, 1967; Muth 1969), intitulado como modelo AMM. Este modelo foi utilizado para verificar a validade da teoria desenvolvida por economistas na tentativa de explicar o espraiamento das cidades. Embora existam muitos estudos para casos internacionais, pouco se sabe com relação ao caso brasileiro. O presente trabalho tem como objetivo colaborar com evidências sobre a validade do modelo monocêntrico e sobre os condicionantes do crescimento da área urbana para o caso Brasileiro. Outras contribuições deste trabalho dizem respeito ao papel da moradia informal sobre a dinâmica de crescimento das cidades brasileiras e o uso de novos dados de satélite na aplicação o modelo monocêntrico. Considerando-se o caso Brasileiro, esta discussão é de suma importância tendo em vista a forte presença da moradia informal em muitas cidades brasileiras.

Os resultados do trabalho estão de acordo com o esperado segundo a literatura econômica. Foram encontrados em todas as especificações sinais positivos e significantes para a população e renda das famílias, indicando que o aumento dessas variáveis estaria correlacionado com maiores áreas urbanas. Para a variável de rendimento agrícola, os resultados indicam que a produtividade da terra tem correlação negativa com o desenvolvimento da área urbana das cidades, assim como apontado pela teoria e como indica as evidências de trabalhos passados. Embora na maior parte das especificações não tenha se revelado estatisticamente significativo. As elasticidades obtidas indicam que para o caso brasileiro a magnitude destas força é menor em relação ao estudos realizados para dados de cidades americanas. As causas destas magnitudes inferiores podem ter origem nas particularidades do Brasil, especificamente a desigualdade de renda e na precária condição da mobilidade urbana. Ainda sobre as elasticidades, foi observado que as elasticidades são menores para as Regis menores. O tempo de *commuting* e a desigualdade maior nas Regis maiores podem ser as causas destas elasticidades menores em Regis com 500 mil habitantes ou mais. Porém, estas possíveis causas devem ser investigados mais profundamente em trabalhos futuros.

Referências

ALONSO, William et al. Location and land use. Toward a general theory of land rent. Location and land use. Toward a general theory of land rent., 1964.

BRUECKNER, Jan K.; FANSLER, David A. The economics of urban sprawl: Theory and evidence on the spatial sizes of cities. The review of Economics and Statistics, p. 479-482, 1983.

BRUECKNER, Jan K. Urban sprawl: diagnosis and remedies. International regional science review, v. 23, n. 2, p. 160-171, 2000.

BURGESS, Robin et al. The political economy of deforestation in the tropics. The Quarterly Journal of Economics, v. 127, n. 4, p. 1707-1754, 2012.

DA MATA, Daniel et al. Determinants of city growth in Brazil. Journal of urban economics, v. 62, n. 2, p. 252-272, 2007.

DENG, Xiangzheng et al. Growth, population and industrialization, and urban land expansion of China. Journal of Urban Economics, v. 63, n. 1, p. 96-115, 2008.

DELINCE, Jacques et al. Handbook on remote sensing for agricultural statistics. 2017.

DONALDSON, Dave; STOREYGARD, Adam. The view from above: Applications of satellite data in economics. Journal of Economic Perspectives, v. 30, n. 4, p. 171-98, 2016.

ENGSTROM, Ryan; HERSH, Jonathan; NEWHOUSE, David. Poverty from space: using high-resolution satellite imagery for estimating economic well-being. 2017.

FUJITA, Masahisa. Urban economic theory: land use and city size. Cambridge university press, 1989.

FUJITA, Masahisa; THISSE, Jacques-François. Economics of agglomeration. Jour-

nal of the Japanese and international economies, v. 10, n. 4, p. 339-378, 1996.

GARCIA-LÓPEZ, Miquel-Àngel; PASIDIS, Ilias-Nikiforos; VILADECANS-MARSAL, Elisabet. Amphitheatres, cathedrals and operas: The role of historic amenities on suburbanization. 2018.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Divisão Regional do Brasil em Regiões Geográficas Imediatas e Regiões Geográficas Intermediárias. 2017.

HARARI, Mariaflavia. "Cities in bad shape: Urban geometry in India." Processed, Wharton School of the University of Pennsylvania (2016).

HENDERSON, J. Vernon; STOREYGARD, Adam; WEIL, David N. Measuring economic growth from outer space. *American economic review*, v. 102, n. 2, p. 994-1028, 2012.

HENDERSON, J. Vernon; REGAN, Tanner; VENABLES, Anthony J. Building the city: sunk capital, sequencing, and institutional frictions. 2016.

HODLER, Roland; RASCHKY, Paul A. Regional favoritism. *The Quarterly Journal of Economics*, v. 129, n. 2, p. 995-1033, 2014.

JAYACHANDRAN, Seema. Air quality and early-life mortality evidence from Indonesia's wildfires. *Journal of Human resources*, v. 44, n. 4, p. 916-954, 2009.

JEAN, Neal et al. Combining satellite imagery and machine learning to predict poverty. *Science*, v. 353, n. 6301, p. 790-794, 2016.

KEOLA, Souknilanh; ANDERSSON, Magnus; HALL, Ola. Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. *World Development*, v. 66, p. 322-334, 2015.

LOPEZ, Russ; HYNES, H. Patricia. Sprawl in the 1990s: measurement, distribution, and trends. *Urban affairs review*, v. 38, n. 3, p. 325-355, 2003.

MARTINS, Hilton; SILVEIRA NETO, Raul da Mota. A Inserção do migrante

rural no mercado de trabalho urbano no Brasil: uma análise empírica da importância dos setores informal e formal. *Estudos Econômicos (São Paulo)*, v. 42, n. 4, p. 731-771, 2012

MCGRATH, Daniel T. More evidence on the spatial scale of cities. *Journal of Urban Economics*, v. 58, n. 1, p. 1-10, 2005.

MILLS, Edwin S. An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, v. 57, n. 2, p. 197-210, 1967.

MUTH, Richard F. CITIES AND HOUSING; THE SPATIAL PATTERN OF URBAN RESIDENTIAL LAND USE. 1969.

PAULSEN, Kurt. Yet even more evidence on the spatial size of cities: Urban spatial expansion in the US, 1980–2000. *Regional Science and Urban Economics*, v. 42, n. 4, p. 561-568, 2012.

REIS, Eustáquio et al. Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000. Rio de Janeiro: Ipea/Dimac, 2008.

SPIVEY, Christy. The Mills—Muth model of urban spatial structure: Surviving the test of time?. *Urban Studies*, v. 45, n. 2, p. 295-312, 2008.

SILVA, Diego; ELHORST, J. Paul; SILVEIRA NETO, Raul da Mota. Urban and rural population growth in a spatial panel of municipalities. *Regional Studies*, v. 51, n. 6, p. 894-908, 2017.

STOREYGARD, Adam. Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa. *The Review of economic studies*, v. 83, n. 3, p. 1263-1295, 2016.

VILA NOVA, C., SILVEIRA NETO, R. (2014) Condicionantes da Área Urbana no Brasil: uma Análise Empírica. *Econtro da Associação Nacional de Estudos Regionais e Urbanos – ENABER*, 2014.

WHEATON, William C. A comparative static analysis of urban spatial structure. *Journal of Economic Theory*, v. 9, n. 2, p. 223-237, 1974.

Apêndice A - Testes de robustez

Nesta seção são apresentados os testes de robustez realizados. O objetivo desta seção é trazer maior solidez e consistência aos resultados mostrados anteriormente. No nosso exercício de robustez é utilizada outra medida de rendimento agrícola - *pt_lc* - na repetição dos mesmos modelos vistos anteriormente, MQO e painel.

Para construir a variável de rendimento agrícola por hectare utilizado nos exercícios de robustez foi utilizado o *land cover* de áreas de pastagem e agricultura e a combinação entre a pesquisas da produção agrícola e agropecuária nos AMCs. No numerador foi posto a soma da produção agrícola, retiradas da Pesquisa da Produção Agrícola Municipal (PAM), e da soma das produção agropecuária, retiradas da Pesquisa da Produção Agropecuária Municipal (PPM), ambas conduzidas pelo IBGE. No denominador utilizou-se os dados de *land cover* das áreas de pastagens ou agricultura em hectares, medida calculada pelo projeto MapBiomas através da categorização de imagens de satélite. Através dessa razão, foi construída a variável *pt_lc* que será utilizada nos testes de robustez. A variável dependente continua sendo o *land cover* da área urbana em hectares, e as variáveis explicativas continuam sendo a população urbana e a renda domiciliar anual média. São apresentados primeiramente os resultados do método de MQO e posteriormente os resultados dos painéis.

Utilizando a amostra de AMCs, os resultados de MQO diferem dos expostos anteriormente utilizando a outra variável de rendimento agrícola, *pa_pam*. Com relação as variáveis de população urbana e renda domiciliar, os resultados variam muito pouco, sendo todos os coeficiente positivos e fortemente significante, assim como visto utilizando *pa_pam*. No entanto, os coeficientes da nova variável de produtividade da terra tem mudanças significantes para alguns anos estudados. Para os anos de 1991 e 2000, não foi encontrado nenhum coeficiente negativo, mesmo com a introdução de controles de macro região. Este fato não segue o esperado e difere dos resultados encontrados quando utilizou-se *pa_pam*, onde foram encontrados coeficientes negativos com o uso de controles. Por outro lado, no ano censitário de 2010 o mesmo resultado foi observado anteriormente, com a uso de controles, a variável passa a ser negativamente correlacionada com a área urbana. A correlação com a moradia informal continua a mesma, sendo positiva e significante para todos os anos censitários.

Ao utilizar a amostra de Regis os resultados variam em relação ao rendimento agrícola. A população, renda e população residente em aglomerados subnormais continuam positivamente correlacionados com a área urbana em todas as especificações. Todas as relações são estatisticamente significantes, com exceção do coeficiente da moradia subnormal em 2010. O rendimento agrícola tem resultados melhores comparados aos vistos

na amostra de AMCs. Embora não seja possível encontrar a relação esperada para o ano de 1991, para os demais anos foram encontradas relações negativas quando utilizou-se controles. Este mesmo fato ocorre quando foi utilizada a variável de rendimento agrícola, *pa_pam*.

Tendo em vista os resultados expostos acima, os resultados para as diferentes formas de medir a produtividade da terra não urbana são bem semelhantes. O sentido das forças das variáveis explicativas são muito semelhantes para as amostras das Regis e AMCs, além disso, a magnitude dos coeficientes também são bem próximos. Portanto, os resultados utilizando os modelos de MQO mostram-se robustos para as duas formas de medir a produtividade da terra.

Por fim, é necessário analisar os resultados obtidos utilizando o método de painel com a variável *pt_lc*. Utilizando a amostra de AMCs os resultados são bem semelhantes para as duas variáveis de rendimento agrícola. População e renda continuam positivamente relacionadas, estatisticamente significante e com magnitudes bastante semelhantes. A população subnormal tem sinal positivo e não significante, não sendo afetada pelo tipo de variável de produtividade da terra. O rendimento agrícola, *pt_lc*, segue com o sinal negativo e significante em nossa especificação de painel. Em suma, todos os resultados para os AMCs estão de acordo com o indicado pela teoria e variam muito pouco com os diferentes tipos de mensuração da produtividade da terra.

Com relação as Regis, analisou-se os diferentes cortes feitos, 100 mil habitantes ou mais e 500 mil habitantes ou mais. Para as Regis com 100 mil ou mais habitantes os resultados seguem os sentidos vistos com a amostra de AMCs. A população, renda, e moradia informal tendo sinais positivos, sendo apenas a última estatisticamente insignificante. O rendimento agrícola é negativamente correlacionado e significativamente estatístico independente da forma de medir este rendimento. Os coeficientes variam pouco entre as especificações e independentemente da forma de medir a produtividade da terra. Com relação a amostra de Regis com 500 mil habitantes ou mais, os sinais sem mantêm os mesmo em relação ao corte de 100 mil habitantes. A única diferença está no rendimento agrícola, que continua com relação insignificante independente de qual variável de produtividade da terra utilizada.

Os resultados mostrados anteriormente mostram que os resultados são pouco sensíveis a mudança na forma de medir o rendimento da terra não urbana. Tanto no Painel, quanto utilizando o método MQO, os resultados não mudam substancialmente. Isso está evidenciado na constância dos sinais encontrados para os coeficientes da variáveis independentes e suas respectivas significâncias estatísticas. Além disso, a variabilidade da magnitude dos resultados também é baixa entre as especificações utilizadas. Estas evidências dão consistência aos resultados mostrados anteriormente para as amostras de AMCs e Regis.

OLS com a amostra de AMCs utilizando pt_lc - Variável dependente é a Área Urbana em hectares

Variáveis	(1) 1991	(2) 1991	(3) 1991	(4) 2000	(5) 2000	(6) 2000	(7) 2010	(8) 2010	(9) 2010
Renda Domiciliar	0,149*** (0,0157)	0,195*** (0,0237)	0,186*** (0,0216)	0,0803*** (0,00799)	0,109*** (0,0133)	0,105*** (0,0123)	0,0807*** (0,00879)	0,117*** (0,0169)	0,111*** (0,0146)
População	0,00935*** (0,000542)	0,00929*** (0,000526)	0,00922*** (0,000500)	0,00913*** (0,000463)	0,00908*** (0,000447)	0,00898*** (0,000428)	0,00853*** (0,000670)	0,00845*** (0,000650)	0,00833*** (0,000627)
Rendimento Agrícola (pt_lc)	0,151 (0,126)	0,131 (0,120)	0,0875 (0,0743)	1,156*** (0,421)	0,525 (0,408)	0,254 (0,401)	-0,555 (0,444)	-1,547** (0,687)	-1,540** (0,715)
% Pop subnormal			23,32** (9,621)			33,61** (16,16)			31,02** (14,48)
Constant	-174,3*** (27,12)	-249,8*** (62,74)	-232,7*** (59,05)	-159,5*** (25,53)	-247,5*** (64,79)	-227,1*** (61,20)	-238,9*** (36,55)	-405,2*** (103,9)	-363,2*** (90,70)
Observations	4,156	4,156	4,156	4,178	4,178	4,178	4,172	4,172	4,172
R-squared	0,913	0,917	0,918	0,920	0,922	0,923	0,897	0,900	0,902
Macro	NO	YES	YES	NO	YES	YES	NO	YES	YES

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

Painel - Amostra de AMCs utilizando pt_lc - Variável dependente é a Área Urbana em hectares

Variáveis	(1)	(2)	(3)	(4)
Renda Domiciliar	0,0166*** (0,00284)	0,0233*** (0,00443)	0,0210*** (0,00459)	0,0209*** (0,00455)
População	0,00785*** (0,000508)	0,00440*** (0,00112)	0,00440*** (0,00112)	0,00438*** (0,00112)
Rendimento Agrícola (pt_lc)	-0,0525*** (0,0151)	-0,0665*** (0,00903)	-0,0663*** (0,00918)	-0,0663*** (0,00946)
% Pop subnormal				2,641 (2,019)
Constant	161,6*** (33,80)	238,3*** (19,49)	246,7*** (20,17)	245,6*** (19,75)
Observations	12,506	12,506	12,506	12,506
R-squared	0,398	0,415	0,417	0,418
Number of Codigo	4,209	4,209	4,209	4,209
EF AMC	NÃO	SIM	SIM	SIM
EF Ano	NÃO	NÃO	SIM	SIM

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

OLS - Regis com 100 mil habitantes ou mais utilizando pt_lc - Variável dependente é a Área Urbana em hectares

Variáveis	(1) 1991	(2) 1991	(3) 1991	(4) 2000	(5) 2000	(6) 2000	(7) 2010	(8) 2010	(9) 2010
Renda Domiciliar	0,835*** (0,113)	0,825*** (0,177)	0,793*** (0,168)	0,570*** (0,0706)	0,625*** (0,128)	0,614*** (0,125)	0,617*** (0,0733)	0,687*** (0,153)	0,661*** (0,150)
População	0,0107*** (0,000287)	0,0108*** (0,000293)	0,0106*** (0,000223)	0,00999*** (0,000295)	0,0100*** (0,000299)	0,00982*** (0,000252)	0,00946*** (0,000463)	0,00948*** (0,000468)	0,00939*** (0,000459)
Rendimento Agrícola (pt_lc)	116,4 (90,24)	113,1 (94,27)	72,54 (80,60)	10,51 (15,08)	-9,119 (14,06)	-34,21* (20,39)	8,358 (21,09)	-17,26 (19,97)	-23,53 (18,86)
perc_sub			17,848* (9,661)			20,425* (12,015)		6,080 (6,282)	
Constant	-838,6*** (264,2)	189,7 (672,1)	173,2 (666,9)	-793,9*** (235,4)	-127,4 (783,6)	44,31 (774,2)	-1,514*** (341,2)	-679,4 (1,175)	-493,7 (1,155)
Observations	357	357	357	357	357	357	357	357	357
R-squared	0,955	0,959	0,961	0,958	0,961	0,962	0,948	0,951	0,951
Macro	NO	YES	YES	NO	YES	YES	NO	YES	YES

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

Painel - Regis com 100 mil habitantes ou mais utilizando pt_lc - Variável dependente é a Área Urbana em hectares

Variáveis	(1)	(2)	(3)	(4)
Renda Domiciliar	0,161*** (0,0329)	0,267*** (0,0457)	0,330*** (0,0574)	0,330*** (0,0582)
População	0,00868*** (0,000386)	0,00495*** (0,000807)	0,00499*** (0,000825)	0,00495*** (0,000835)
Rendimento Agrícola (pt_lc)	-2,944 (14,28)	-20,47** (8,802)	-20,08* (9,929)	-19,75* (10,28)
% Pop subnormal				18,21 (16,65)
Constant	1,500*** (304,4)	2,476*** (178,4)	2,389*** (190,2)	2,374*** (188,8)
Observations	1,071	1,071	1,071	1,071
R-squared	0,634	0,672	0,678	0,679
Number of cod_rgi	357	357	357	357
EF Regis	NÃO	SIM	SIM	SIM
EF Ano	NÃO	NÃO	SIM	SIM

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

Painel - Regis com população entre 100 e 500 mil habitantes ou mais utilizando pt_lc - Variável dependente é a Área Urbana em hectares

Variáveis	(1)	(2)	(3)	(4)
Renda Domiciliar	0,0958*** (0,0162)	0,123*** (0,0166)	0,197*** (0,0246)	0,195*** (0,0247)
População	0,0123*** (0,00113)	0,00976*** (0,00114)	0,0108*** (0,00115)	0,0111*** (0,00115)
Rendimento Agrícola (pt_lc)	-0,334 (7,577)	1,371 (6,128)	1,541 (5,802)	0,867 (5,450)
% Pop subnormal				-12,69 (7,978)
Constant	503,0*** (167,9)	789,7*** (154,9)	556,8*** (165,0)	527,4*** (160,2)
Observations	936	936	936	936
R-squared	0,750	0,754	0,773	0,775
Number of cod_rgi	312	312	312	312
EF Regis	NÃO	SIM	SIM	SIM
EF Ano	NÃO	NÃO	SIM	SIM

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria

Painel - Regis com 500 mil habitantes ou mais utilizando pt_lc - Variável dependente é a Área Urbana em hectares

Variáveis	(1)	(2)	(3)	(4)
Renda Domiciliar	0,229* (0,130)	0,770*** (0,181)	0,888** (0,325)	0,895** (0,328)
População	0,00819*** (0,000347)	0,00357*** (0,000708)	0,00365*** (0,000737)	0,00363*** (0,000743)
Rendimento Agrícola (pt_lc)	-12,09 (33,55)	-33,34* (18,69)	-33,61 (24,93)	-34,55 (26,71)
% Pop subnormal				19,11 (38,46)
Constant	6,721*** (1,389)	12,242*** (864,3)	12,017*** (1,042)	11,915*** (1,005)
Observations	135	135	135	135
R-squared	0,637	0,723	0,733	0,733
Number of cod_rgi	45	45	45	45
EF Regis	NÃO	SIM	SIM	SIM
EF Ano	NÃO	NÃO	SIM	SIM

Erros padrões robustos em parênteses

*** p<0,01, ** p<0,05, * p<0,1

Fonte: Elaboração própria