UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA
CURSO DE DOUTORADO EM ECONOMIA

LAÉRCIO DAMIANE CERQUEIRA DA SILVA

TAMANHO DOS CLUBES, FUNÇÕES DE CONGESTIONAMENTO E ECONOMIAS DE ESCALA NA PROVISÃO DE BENS PÚBLICOS LOCAIS NO BRASIL.

JOÃO PESSOA/ PB

2015

LAÉRCIO DAMIANE CERQUEIRA DA SILVA

TAMANHO DOS CLUBES, FUNÇÕES DE CONGESTIONAMENTO E ECONOMIAS DE ESCALA NA PROVISÃO DE BENS PÚBLICOS LOCAIS NO BRASIL.

Tese submetida ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba (PPGE/UFPB), como um dos requisitos para obtenção do título de Doutor em Economia.

LINHA DE PESQUISA: ECONOMIA REGIONAL E POLÍTICAS PÚBLICAS.

Orientadora: Profª. Drª. Maria da Conceição Sampaio de Sousa

JOÃO PESSOA/ PB

2015

S586t Silva, Laércio Damiane Cerqueira da.
Tamanho dos clubes, funções de congestionamento e economias de escala na provisão de bens públicos locais no Brasil / Laércio Damiane Cerqueira da Silva.- João Pessoa, 2015.
108f. : il.
Orientadora: Maria da Conceição Sampaio de Sousa
Tese (Doutorado) - UFPB/CCSA
1. Economia regional. 2. Provisão local. 3. Modelo de demanda. 4. Modelo de oferta. 5. Efeito *Threshold*. 6. Efeito congestionamento.

UFPB/BC CDU: 338.1(043)

LAÉRCIO DAMIANE CERQUEIRA DA SILVA

TAMANHO DOS CLUBES, FUNÇÕES DE CONGESTIONAMENTO E ECONOMIAS DE ESCALA NA PROVISÃO DE BENS PÚBLICOS LOCAIS NO BRASIL.

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba - (PPGE/UFPB), em cumprimento às exigências para obtenção do título de Doutor em Economia.
Submetida à apreciação da Banca Examinadora, e aprovada em 04 / 12 / 2015

BANCA EXAMINADORA

Profª. Drª. Maria da Conceição Sampaio de Sousa
(Orientadora – UFPB)

Profª. Drª. Mércia Santos da Cruz
(Examinadora Interna – UFPB)

Profº. Dr. Magno Vamberto Batista da Silva
(Examinador Interno – UFPB)

Prof. Dr. Guilherme Diniz Irffi
(Examinador Externo – UFC)

Profª. Drª. Rozane Bezerra de Siqueira
(Examinadora Externa – UFPE)

Dedico aos meus familiares, aos quais devo, em grande parte, as minhas conquistas, e o que sou hoje.

The only way to have real success in science...is to describe the evidence very carefully without regard to the way you feel it should be. If you have a theory, you must try to explain what's good about it and what's bad about it equally. In science you learn a kind of standard integrity and honesty.

(Richard Feynman)

AGRADECIMENTOS

A Deus, sobre todas as coisas e pessoas.

Aos meus amados, Raimundo Vieira e Solange Cerqueira, meus pais, exemplos de carinho, amor, decência, os quais deram apoio incondicional às minhas escolhas nessa jornada.

Aos meus irmãos, Laerte Cerqueira e Laerge Cerqueira, pela união, pelo apoio de todas as horas, por acreditarem em mim.

À minha orientadora, Professora Conceição, que me direcionou para a área de Setor Público, por ter aceitado de imediato o desafio de elaborar essa pesquisa, e que com extrema dedicação e responsabilidade, dosou as críticas com comentários de incentivos e transmitiu de forma hábil os seus conhecimentos para o resultado desta.

Aos professores Mércia Santos e Magno Vamberto, pelas sugestões feitas ao projeto apresentado na qualificação, e sobretudo por suas valorosas contribuições à minha vida acadêmica. Também aos professores Guilherme Irffi e Rozane Bezerra, pela disponibilidade em participar da avaliação deste estudo. É um enorme prazer tê-los fazendo parte desse trabalho.

Às colegas do Doutorado Emanuelle Alícia, Marcella Braga, e Patrícia Amarante pela convivência e amizade construídas ao longo do curso.

Ao colega de mestrado e amigo, Profº Paulo Roberto Freitas, pelo valoroso auxílio na aplicação econométrica.

Às Secretárias e Secretário do PPGE: Risomar, Miriam e Ricardo, por toda atenção, amizade e carinho que despendem sempre.

Por fim, a todos que de uma forma direta ou indireta contribuíram para este momento. Sugestões, críticas, palavras de incentivo, direcionamento me trouxeram até aqui. A todos vocês, indistintamente, o meu muito obrigado.

.

RESUMO

A tese versa, em três capítulos, sobre a provisão de bens públicos nos municípios brasileiros. O primeiro capítulo coloca em competição duas especificações alternativas, uma do lado da demanda e outra do lado da oferta, para um conjunto de dados de 5285 municípios brasileiros para 2010, a fim de testar a hipótese de que a relevância dos modelos que explicam o comportamento das despesas públicas locais pode depender do tamanho da população das jurisdições. A metodologia econométrica usa a combinação do Teste J, proposto por Davidson e MacKinnon (1981), com o *Threshold*, sugerido por Hansen (2000), como um método de triagem de dados capaz de testar se a resposta dos governos locais aos seus eleitores é estável entre os municípios de pequeno e de grande porte. Os resultados permitem inferir que o modelo do eleitor mediano aplica-se às cidades menores, com tamanho ótimo populacional variando entre 13604 e 13661 habitantes, dependendo da função de congestionamento utilizada. Não obstante, é rejeitado em 2293 municípios, para os quais o modelo orientado para a oferta apresenta melhor ajuste. No segundo capítulo, aplica-se os referidos modelos alternativos de determinação da provisão pública a fim de mensurar o impacto do tamanho das cidades sobre a qualidade dos serviços públicos - via análise da publicidade e da rivalidade dos bens – em 5268 municípios brasileiros baseado na Teoria dos Clubes. Para os dois modelos, os resultados apontam a existência da correlação espacial entre os municípios, sugerindo que os padrões de gastos em determinada localidade influencia nas decisões fiscais de vizinhos. Comprova-se a hipótese de que o tamanho dos clubes e o custo dos serviços públicos diferenciam os resultados da ação de provisão, corroborando com a literatura sobre o tema a qual defende que em jurisdições menores a proximidade dos eleitores com seus agentes públicos exerce maior pressão e controle sobre seus atos políticos, e que preferências individuais agregadas apresentam menor heterogeneidade e podem representar a comunidade como um todo. Este resultado gera um efeito congestionamento maior para as maiores cidades, o que vai de encontro à literatura internacional. Por fim, usando um conjunto de dados de localidades brasileiras que formam Consórcios Públicos Intermunicipais de serviços públicos, o terceiro capítulo visa estabelecer evidências empíricas para um fenômeno denominado por Oates (1988) como "efeito-zoo". De acordo com esse efeito, o nível mais baixo de despesas em cidades menores não deve ser explicado pela menor aglomeração, mas simplesmente pelo fato de que certas categorias de gastos - como um zoológico - precisam de um tamanho mínimo populacional para serem fornecidos. Destarte, a relação positiva entre o tamanho da população e o gasto local em serviços públicos não leva em conta o fato de que nos municípios mais populosos, a cesta de serviços ofertados é mais ampla e mais diversificada, e por essa razão as estimativas de congestionamento existentes tendem a ser superestimadas. Através da estatística espacial confirmou-se a hipótese do efeito zoo nos municípios brasileiros, e que a maior diversificação na gama de serviços expande a oferta de serviços em áreas vizinhas, justificando a interação entre os municípios como forma de se obter melhor e maior nível de serviços públicos para seus cidadãos. As evidências apontam que a intensidade do efeito depende do gradiente urbano-rural, sendo o fenômeno maior em áreas urbanizadas, devido às economias de escala mais substanciais nessas localidades.

PALAVRAS-CHAVE: Provisão local; Modelo de Demanda; Modelo de Oferta; Efeito *Threshold;* Efeito Congestionamento; Teoria dos Clubes; Efeito-zoo.

ABSTRACT

The thesis studies, in three chapters, on the provision of public goods in municipalities. The first chapter puts in competition two alternative specifications, one on the demand side and another on the supply side, to a set of 5285 municipalities data for 2010 in order to test the hypothesis that the relevance of the models that explain the behavior local government spending can depend on the size of the population of the jurisdictions. The econometric methodology uses a combination of Test J proposed by Davidson and MacKinnon (1981), with the threshold suggested by Hansen (2000), as a data screening method to test whether the response of local governments to their constituents is stable between small and large municipalities. Results show that the median voter model applies to smaller cities, with great population size ranging between 13604 and 13661 inhabitants, depending on the congestion function used. Nevertheless, it rejected in 2293 municipalities, for which the model-oriented supply fits better. In the second chapter, these alternative models of determination of public provision in order to measure the impact of the size of cities on the quality of public services applies - via analysis of advertising and rivalry of goods - in 5268 Brazilian municipalities based on the Theory of Clubs. For both models, the results point to the existence of spatial correlation between municipalities, suggesting that spending patterns in a given locality influence on the fiscal decisions of neighbors. Proves the hypothesis that the size of the clubs, and the cost of utilities differentiate the results of the provision of action, supporting the literature on the subject which argues that in smaller jurisdictions proximity of voters with its public officials exerts more pressure and control over their political acts, and that aggregate individual preferences have less heterogeneity and can represent the community as a whole. This result yields a larger congestion effect to the larger cities, which goes against the international literature. Finally, using a set of data that make up Brazilian localities Public Consortia Inter-municipal public services, the third chapter aims to establish empirical evidence for a phenomenon called by Oates (1988) as "zoo-effect." According to this effect, the lowest level of expenditure in smaller towns should not be explained by the smaller clusters, but simply by the fact that certain categories of spending - like a zoo - they need a minimum population size to be provided. Thus, the positive relationship between population size and local spending on public services does not take into account the fact that in the most populous cities, the basket of services offered is broader and more diverse, and for this reason, estimates of existing congestion tend to be overestimated. By spatial statistics confirmed the hypothesis of the zoo effect on municipalities, and that the more diversified the range of services expands the range of services in neighboring areas, which explains the interaction between the municipalities as a way to get better and higher level of public services to its citizens. Evidence shows that the intensity of the effect depends on the urban-rural gradient, with the biggest phenomenon in urban areas due to more substantial economies of scale in these locations.

KEYWORDS: Local Provision; Demand model; Supply model; Threshold effect; Congestion effect; Theory of Clubs; Zoo-effect.

LISTA DE TABELAS

Tabela 1.1: Estatísticas descritivas 24
Tabela 1.2: Resultados dos testes de linearidade para as subamostras 32
Tabela 1.3: Resultados das estimações do Modelo *Threshold* - Modelo B-D 36
Tabela 1.4: Resultados das estimações do modelo *Threshold – Camaraderie*.....37
Tabela 1.5: Resultados das estimações do Modelo *Threshold* – Modelo Flexível.....38
Tabela 1.6: Elasticidades e parâmetros de congestionamento por forma funcional..... 40
Tabela 2.1: Estatísticas descritivas 58
Tabela 2.2: Estimações por MQO do Modelo da Demanda – Por forma funcional..... 62
Tabela 2.3: Diagnóstico para dependência espacial do Modelo da Demanda – Por forma funcional.....62
Tabela 2.4: Estimações por MQO do Modelo da Oferta – Por forma funcional 63
Tabela 2.5: Diagnóstico para dependência espacial – Modelo da Oferta – Por forma funcional.....64
Tabela 2.6: Estimações do Modelo da Demanda – Método SAR/VI – Formulação Flexível 65
Tabela 2.7: Estimações do Modelo da Oferta – Método SAR/VI – Formulação Flexível. 69
Tabela 2.8: Elasticidades e parâmetros de congestionamento por forma funcional – Modelo da Demanda 71
Tabela 2.9: Elasticidades e parâmetros de congestionamento por forma funcional – Modelo da Oferta 72
Tabela 2.10: Comparação estatística dos modelos de demanda e oferta 73
Tabela 3.1: Distribuição de municípios que formam consórcios por regiões do Brasil 83
Tabela 3.2: Tipos de CPI's por região do Brasil 84
Tabela 3.3: Estatísticas descritivas 88
Tabela 3.4: Teste de I de *Moran Global* Bivariado 93
Tabela 3.5: Testes I de Moran e Multiplicador de Lagrange (ML) 95
Tabela 3.6: Resultados das estimações por modelos: MQO e SAR por VI 96

LISTA DE QUADROS

Quadro 1.1: Descrição das variáveis utilizadas no estudo..... 24
Quadro 2.1: Resumo descritivo das variáveis utilizadas no estudo 57

LISTA DE FIGURAS

Figura 1.1: Teste para a variável *threshold*: Modelo Borcherding-Deacon..... 33
Figura 1.2: Teste para a variável *threshold*: modelo *camaraderie*..... 34
Figura 1.3: Teste para a variável *threshold*: modelo flexível..... 34
Figura 2.1: Função de Borcheding-Deacon (B-D) 49
Figura 2.2: Função *"Camaraderie"* 50
Figura 3.1: Congestão em bens públicos locais 79
Figura 3.2: Número de Serviços Públicos e Tamanho da População 86
Figura 3.3: Mapa de *Clusters* e *Outliers* 94

LISTA DE SIGLAS E ABREVIATURAS

B-D - Borcherding e Deacon

BIM - Base de Informações Municipais

C1 - Clube 1

C2 - Clube 2

C3 - Clube 3

C4 - Clube 4

C5 - Clube 5

CMFP – Custo Marginal dos Fundos Públicos

CPIs - Consórcios Públicos Intermunicipais

DUM - *Dummy*

DUMCN - *Dummy* Consórcio

DUMCP - *Dummy* Capital

DUMP1 - *Dummy* de Alinhamento Político Estadual

DUMP2 - *Dummy* de Alinhamento Político Federal

DUMRG - *Dummy* Regional

FINBRA - Finanças Brasil

IBGE - Instituto Brasileiro de Geografia e Estatística

IDHM - Índice de Desenvolvimento Humano Municipal

LISA - *Local Indicator of Spatial Association*

ML - Multiplicador de Lagrange

MQO – Mínimos Quadrados Ordinários

MRLC - Modelo de Regressão Linear Clássico

MV - Máxima Merossimilhança

SAR - *Spatial Auto Regressive*

SEM - *Spatial Error Model*

STN - Secretaria do Tesouro Nacional

TMS – Taxa Marginal de Substituição

TSE - Tribunal Superior Eleitoral

VI - Variáveis Instrumentais

SUMÁRIO

INTRODUÇÃO 14
1.MODELOS DE DEMANDA E OFERTA NA PROVISÃO MUNICIPAL DE BENS PÚBLICOS LOCAIS: UMA ANÁLISE COM *THRESHOLD* 19
1.1.INTRODUÇÃO 19
1.2.EXPLICANDO AS DESPESAS PÚBLICAS LOCAIS: HIPÓTESES ALTERNATIVAS 21
1.2.1. O lado da demanda e a escolha do eleitor mediano 21
1.2.2. O lado da oferta e as escolhas representativas 22
1.3.DADOS 23
1.4.ESPECIFICAÇÕES ECONOMÉTRICAS 25
1.4.1. O teste *j e* o diagnóstico de modelos concorrentes 27
1.4.2. O efeito *threshold* e teste de identificação do parâmetro 28
1.4.3. Combinando o teste j com o modelo *threshold* 30
1.5. RESULTADOS 32
1.6. CONSIDERAÇÕES FINAIS 41
2. A QUALIDADE NA PROVISÃO LOCAL DE BENS PÚBLICOS NO BRASIL: UMA ANÁLISE DE CONGESTIONAMENTO BASEADA NA TEORIA DOS CLUBES 43
2.1. INTRODUÇÃO 43
2.2. UMA BREVE RESENHA SOBRE A TEORIA DOS CLUBES 44
2.3. O EFEITO DO TAMANHO DOS CLUBES SOBRE A QUALIDADE DA PROVISÃO PÚBLICA 47
2.4. PROVISÃO DE BENS PÚBLICOS NO CONTEXTO DE CLUBES: MODELOS DE DEMANDA E OFERTA 51
2.4.1. O modelo da demanda 51
2.4.2. O modelo da oferta 53
2.5. DADOS VARÁVEIS 55
2.6. ESPECIFICAÇÕES ECONOMÉTRICAS 58
2.7. RESULTADOS 61
2.7.1. Comparação estatística dos modelos estimados 73
2.8. CONSIDERAÇÕES FINAIS 74

3. INDIVISIBILIDADES NA PROVISÃO LOCAL DE BENS PÚBLICOS NO BRASIL: UMA ANÁLISE DAS EVIDÊNCIAS DO "EFEITO ZOO" 76
3.1. INTRODUÇÃO 76
3.2. INTUIÇÃO TEÓRICA DO EFEITO ZOO..... 78
3.2.1. A Literatura Empírica 80
3.3. O CONTEXTO INSTITUCIONAL BRASILEIRO E A DIVERSIDADE NO FORNECIMENTO DE SERVIÇOS PÚBLICOS..... 82
3.4. METODOLOGIA 85
3.4.1. Base de Dados e Estratégias de Estimação do Modelo 85
3.4.2. Análise Espacial 89
3.4.3. Modelos Econométricos Espaciais 91
3.5. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS 93
3.6. CONSIDERAÇÕES FINAIS 99
CONCLUSÃO 100
REFERÊNCIAS 103

INTRODUÇÃO

Os estudos sobre a determinação da provisão ótima de bens públicos passou a ser recorrente na literatura econômica. As análises iniciais partiram de uma perspectiva em equilíbrio parcial, propostas por Lindahl (1919) e revisadas adiante por Johansen (1963), nas quais as observações limitavam-se a considerar o mercado do bem público, ignorando as suas inter-relações com os demais setores da economia.

Considerando, por definição, a não-rivalidade e não excludência no consumo dos bens, e que a demanda da sociedade representava a soma vertical das demandas individuais, esta abordagem assumia a produção do bem a custos constantes no longo prazo, e a oferta infinitamente elástica ao nível de seu custo marginal. Assim, nesse modelo, a provisão ótima seria no nível em que o Custo Marginal Social fosse igual à Utilidade Marginal Social.

Samuelson (1954), em sua abordagem que considerava simultâneo os vários bens e mercados, concluiu que a soma das Taxas Marginais de Substituição (TMS) entre bens públicos e privados deveria ao menos equivaler à relação de preços entre esses bens[1] se a soma das TMS for maior que a relação de preços entre bens públicos e privados a economia estará sujeita a presença dos *free riders*.[2] Do contrário, se a soma das TMS for menor, ocorre uma ineficiência no sentido de Pareto, resultando em uma subprovisão dos bens. Buchanan (1965), com uma metodologia diferenciada, que se aplicava a bens intermediários entre os extremos "puros e privados", comprovou as mesmas condições de eficiência, atribuindo a alocação ótima de recursos, no sentido Paretiano da expressão, quando o resultado da análise é tal que os agentes tenham esgotado todas as vantagens de troca, considerando as dotações iniciais e já tendo definidos, previamente, os direitos de propriedade.

A crítica que sofre esses modelos de provisão ótima, é que nem sempre é possível recolher informações suficientes para se conhecer as verdadeiras demandas dos agentes, já que estes adotam comportamentos estratégicos que visam maximizar seus benefícios individuais. De fato, em análises empíricas, um dos problemas observados quando o Estado tem o encargo da provisão de certos bens e serviços é que as decisões de alocação no âmbito público são, por vezes, tecnicamente inviáveis, devido à assimetria informacional no sentido de que os consumidores não revelam as suas preferências pela escolha de um bem e o pagamento do seu respectivo preço. As decisões públicas se fazem então por meio de

---

[1] Esta é a condição de Equilíbrio Geral, também conhecida como Condição de Samuelson.

[2] Problema que permeia toda a análise de ação coletiva, este comportamento caracteriza-se como aquele em que devido as características de não-rivalidade e não-exclusividade dos bens públicos, um ou mais agentes usufruem de um certo benefício proveniente de um bem, sem que tenha contribuído para a obtenção deste.

processos políticos, como tentativa de resolução deste problema. Contudo, há uma percepção de que esses mecanismos não têm se mostrado tão eficientes.

Dentro desse contexto, no que se refere à provisão pública, outro conceito passa a ter grande relevância: a descentralização governamental. Conforme a Teoria do Federalismo Fiscal[3], no caso de bens públicos locais, a oferta descentralizada oferece algumas vantagens que geram ganhos de bem-estar social. Para esses bens, os governos locais podem oferecer quantidades que atendam melhor às demandas dos moradores de suas respectivas jurisdições, uma vez que conhecem melhor as preferências e as condições específicas de cada localidade. Isso proporcionaria um maior nível de bem-estar em relação à situação onde o mesmo bem fosse ofertado uniformemente pelo governo central em todas as regiões.

A literatura apresenta dois tipos de fatores que atuam a favor dos ganhos de desempenho com a provisão local dos bens e serviços. O primeiro envolve as questões já citadas, de assimetria de informações; os governos locais estariam mais pertos de suas bases eleitorais, tendo um melhor conhecimento das preferências ou demandas dos moradores e maiores conhecimentos dos custos desses bens. O outro argumento, de caráter político, ressalta a existência de constrangimentos que impedem o governo central de fornecer mais e melhores bens em uma dada jurisdição, tendo assim que existir certa homogeneidade de tratamento em escala nacional. Isto torna difícil para o governo central variar os níveis de bens e serviços públicos entre jurisdições.

Em síntese, há maximização da eficiência econômica quando os bens públicos locais são ofertados pelos entes governamentais mais próximos dos beneficiários. Com efeito, os governos estaduais, por exemplo, seriam responsáveis pela provisão de bens com preferências mais regionais e os governos municipais responderiam pelos bens cujo perfil da demanda fosse específico de uma localidade. Por sua vez, o governo central deveria prover bens cujos benefícios espalham-se por todo o território federativo ou cuja produção esteja sujeita a grandes economias de escala. Além disso, existiria maior *accountability*[4] por parte dos residentes em cada local, quando comparados com o caso de centralização.

A Constituição Federal de 1988 identifica esse tipo de preocupação quando considera os dados da população e renda *per capita* de uma localidade, como base para fixar quotas de transferências de recursos, que complementam as fontes de receitas próprias dos municípios e

---

[3] Para detalhes , ver Oates (1977).

[4]Termo usado na literatura da economia pública, que em português aproxima-se da palavra **responsabilização** ou **confiabilidade,** e remete à obrigação de membros de um órgão representativo de **prestar contas** a instâncias controladoras ou a seus representados.

proporcionam condições financeiras mais adequadas aos governos municipais para o efetivo cumprimento das suas responsabilidades.

Nesse contexto, os parâmetros da população podem ser entendidos, portanto, como características socioeconômicas básicas dos municípios que refletem as condições de demanda local. Parâmetros que devem ser considerados indicadores sintéticos representativos dos municípios do país, mas que influenciam diretamente a capacidade das cidades receberem recursos das esferas superiores de governo para a provisão de serviços públicos à comunidade local. As características socioeconômicas locais podem, assim, representar a demanda local por bens públicos. As estimativas de demanda para diversos serviços governamentais, revelam que as despesas são explicadas por diferenças em rendimentos *per capita*, taxa de urbanização, densidade demográfica, taxas de impostos, população, transferências de recursos intergovernamentais, nível educacional, entre outros fatores.

A partir de tais considerações, a literatura nacional tem focado primordialmente na análise de como o padrão de demanda local pode explicar o nível ou a composição da despesa pública local que, em última instância, representa os vários tipos de serviços públicos providos pelo governo local à sua comunidade. O problema, entretanto, é que a literatura tem considerado, *a priori*, uma relativa rigidez no lado da oferta de serviços públicos, como se as instituições públicas, os grupos de interesse e os gestores locais eleitos fossem agentes exógenos ao processo de decisão de fornecimento dos bens públicos locais. Deve-se levar em conta que, em níveis descentralizados de governo, os determinantes das despesas públicas podem caber em duas categorias de modelos – oferta e demanda - dependendo da natureza do referido processo.

Concomitantemente, desconsideram o impacto do tamanho dos clubes sobre a qualidade da provisão do serviço público e sobre o poder explicativo dos modelos que determinam as despesas locais. O foco dos estudos nas análises do efeito congestionamento pode não permitir a compreensão plena do impacto do tamanho populacional. Isto porque se, por um lado, o crescimento populacional das cidades pode reduzir a qualidade do bem público por meio do supracitado efeito, esse crescimento contribui para que possam beneficiar-se de economias de escala associadas à produção dos bens públicos locais.

Por fim, os estudos sobre economia pública local, que analisam como o tamanho populacional afeta o nível dos serviços governamentais provisionados, desconsideram que os níveis mais baixos de gastos em cidades menores não devem ser explicados pela menor aglomeração, mas simplesmente pelo fato de que certas categorias de despesas - como um zoológico - precisam de um tamanho populacional mínimo para serem fornecidas. Isto faz

com que nos municípios mais populosos, a cesta de serviços ofertados seja mais ampla e mais diversificada, e por essa razão, as estimativas de congestionamento existentes tendem a ser superestimadas. Esse fenômeno é conhecido na literatura como “Efeito Zoo” (OATES, 1988).

Embora existam no Brasil vários trabalhos que abordem os dois primeiros pontos acima mencionados, nenhum deles analisa, simultaneamente, a participação dos eleitores e dos gestores públicos na tomada de decisão no que diz respeito aos gastos dos governos referentes à provisão de serviços públicos, no âmbito municipal. Ademais, esses trabalhos não consideram o impacto do tamanho da cidade sobre a escolha dos modelos concorrentes - o modelo do eleitor mediano e o modelo de oferta - com o intuito de identificar um possível efeito limiar, a partir do qual a dominância de um modelo frente a outro seja bem estabelecida. Por fim, ao ignorar o “Efeito Zoo”, nenhuma das pesquisas leva em conta a questão do superdimensionamento do efeito congestionamento. Esta tese contribui para preencher essa lacuna a partir do desenvolvimento de estudos que contemplam essa agenda de pesquisa.

Dentro desse contexto, o primeiro capítulo, com base nos trabalhos de Josselin, Rocaboy e Tavéra (2007), e Breunig e Rocaboy (2008), confronta as duas especificações alternativas de provisão de bens públicos, uma do lado da demanda e outra direcionada à oferta, a fim de verificar se o tamanho da população de uma localidade em um mesmo nível de governo influencia a capacidade de resposta dos governos locais com relação às preferências do eleitorado. A metodologia econométrica utilizada nesta etapa do trabalho é a combinação do Teste J, proposto por Davidson e MacKinnon (1981), com o *Threshold*, sugerido por Hansen (2000), como um método capaz de testar a possibilidade de quebra do conjunto de dados, em regimes, e ao mesmo tempo a expectativa desses regimes serem gerados pelo modelo do eleitor mediano e/ou pelo modelo direcionado para a oferta.

O segundo capítulo mensura o impacto do tamanho das cidades sobre a qualidade dos serviços públicos - via análise de publicidade e rivalidade dos bens - nas municipalidades brasileiras sob o arcabouço da Teoria dos Clubes. Nesta fase do trabalho, a metodologia econométrica emprega, além do modelo clássico linear, a análise espacial, proposta por Anselin (1988). Adicionalmente, para comparação dos padrões, utilizou-se do Teste J (Davidson e Mackinnon, 1981) ao ponto de identificar o seu poder explicativo como determinantes para toda a gama de serviços providos pelos clubes.

Finalmente, considerando que muitos bens públicos são indivisíveis, e que a gama de serviços governamentais aumentam proporcionalmente à dimensão das jurisdições, o terceiro capítulo avalia a magnitude do viés do congestionamento nos municípios brasileiros,

buscando evidencias do que se conhece na literatura por "efeito zoo". Esse estágio da pesquisa utiliza-se da econometria espacial e terá como principais referências Oates (1988), Frère, Hammadou e Paty (2011).

Em linhas gerais, esta tese estuda os fatores associados à provisão de bens nos municípios brasileiros evidenciando as idiossincrasias que determinam a sua demanda, e o efeito, em termos de eficiência alocativa, da percepção dos agentes públicos quanto às peculiaridades no comportamento da sociedade local. Os resultados podem contribuir para uma discussão a respeito da organização do setor público, para provisão de serviços públicos mais baratos, melhor adequação do fornecimento às preferências dos cidadãos espacialmente heterogêneos e, por fim, a possibilidade de incentivos do governo para desenvolvimento da coordenação municipal em zonas mais densamente povoadas.

1. MODELOS DE DEMANDA E OFERTA NA PROVISÃO MUNICIPAL DE BENS PÚBLICOS LOCAIS: UMA ANÁLISE COM *THRESHOLD*

1.1 INTRODUÇÃO

A análise da determinação das despesas públicas constitui um tópico relevante nas discussões que envolvem as finanças dos governos locais. Na extensa literatura sobre o tema, enfoques distintos competem para explicar as escolhas fiscais desses governos. Do lado da demanda, o modelo de referência tradicional é o modelo do eleitor mediano, no qual as escolhas orçamentárias da comunidade baseiam-se nas preferências do eleitor mediano, desde que assegurada a competição política em um ambiente democrático. O agente político, escolhido por votação majoritária, tenderia a escolher uma posição que refletiria a mediana da distribuição das preferências dos eleitores.

Desde os trabalhos seminais de Borcherding e Deacon (1972) e Bergstrom e Goodman (1973), as características da demanda do eleitor mediano têm sido usadas por vários autores para explicar o comportamento dos gastos públicos locais.

No entanto, vários aspectos das escolhas fiscais – em particular aqueles ligados à tomada de decisão pelos gestores públicos - não são considerados pela teoria do eleitor mediano. Ora, estes agentes públicos podem agir de forma discricionária, privilegiando interesses próprios em detrimento daqueles da população. Os trabalhos de Wyckoff (1990), Dye e Mac Guire (1997), Ahmed e Greene (2000) exemplificam esse tipo de argumentação. Para esses autores, os gastos locais estão mais ligados ao comportamento burocrático dos governos locais, à ineficiência da estrutura tributária, à existência de grupos de interesse e aos controles de agenda política, fatores predominantemente de oferta.

A provisão de bens públicos seria melhor explicada por modelos de oferta, nos quais o processo político seria claramente explicitado. Dois modelos concorrentes tentam analisar o comportamento fiscal dos gestores locais: a hipótese de comportamento oportunista e o modelo partidário. No primeiro, os governantes utilizam a passagem no poder para elevar sua probabibilidade de reeleição (Nordhaus, 1975 ; Rogoff e Sibert, 1988 ; Besley e Case, 1995). No modelo partidário, os gestores públicos maximizam apenas o bem-estar dos seus eleitores sendo suas escolhas fiscais motivadas por objetivos ideológicos (Hibbs, 1977 ; Alesina, 1987). Nessa visão, elementos da oferta regeriam a provisão dos bens públicos locais, sendo as características do eleitor mediano irrelevantes para explicar as escolhas fiscais nas esferas inferiores de governo.

Quando testados separadamente, a evidência empírica disponível suporta tanto os modelos tradicionais, baseados na demanda, como os que privilegiam os aspectos da oferta. Ampliando-se o escopo da investigação, um ponto relevante é determinar qual dessas abordagens prevaleceria - e em que circunstâncias – quando testadas simultaneamente. Ademais, poder-se-ia investigar se a escolha do enfoque é influenciada pelo tamanho da população de uma dada jurisdição. Esse ponto é particularmente relevante no caso brasileiro, em que as comunidades são bastante heterogêneas, com grandes cidades coexistindo com uma miríade de pequenas municipalidades. Assim, por exemplo, é razoável supor que a capacidade de resposta dos governos locais às preferências do eleitorado seja maior em pequenos munícipios, onde a proximidade do gestor com seu eleitorado é maior e o controle social mais presente. A hipótese do eleitor mediano – em que os gestores locais buscam maximizar a utilidade desse eleitor – seria, nesse caso, o enfoque mais apropriado para explicar a provisão de bens públicos locais. Já nas maiores cidades, onde o anonimato contribui para desobrigar os representantes eleitos de um compromisso rigoroso com a escolha do eleitor mediano, a provisão de bens públicos pode ser feita de forma discricionária. Neste caso, o lado da oferta prevalece, e as escolhas fiscais adotadas serão aquelas que atendem aos interesses próprios dos políticos locais e burocratas do governo.

O objetivo deste trabalho é explicar o comportamento dos gastos públicos locais, nos municípios brasileiros, com base nas duas especificações previamente mencionadas, a saber, a hipótese do eleitor mediano (modelo de demanda) e o modelo de oferta, em que predominam os aspectos políticos. A partir de informações das finanças municipais e dos dados do Censo 2010, a metodologia da pesquisa combina o Teste *J*, proposto por Davidson e MacKinnon (1981) com o Modelo Limiar (*Threshold*), sugerido por Hansen (2000), para avaliar, simultaneamente, o poder explicativo das especificações propostas. Analisa-se, também, o impacto do tamanho da cidade sobre a escolha dos modelos concorrentes com o intuito de identificar um possível efeito limiar, a partir do qual a dominância de um modelo frente a outro seja bem estabelecida.

Este capítulo está organizado em cinco seções, além dessa introdução. As duas próximas seções descrevem, respectivamente, especificações alternativas de determinação da provisão de bens públicos, ou seja, o modelo do eleitor mediano e o modelo direcionado para a oferta. Na seção quatro, são apresentados os procedimentos metodológicos utilizados na pesquisa. Na quinta seção, analisam-se os resultados das estimações realizadas. Por fim, têm-se as considerações finais, onde estão dispostas as conclusões, sugestões para futuras pesquisas na área e limitações da pesquisa.

1.2 EXPLICANDO AS DESPESAS PÚBLICAS LOCAIS: HIPÓTESES ALTERNATIVAS

As subseções 1.2.1 e 1.2.2 descrevem sucintamente as duas vertentes teóricas utilizadas para explicar o comportamento da despesa pública nos municípios brasileiros.

1.2.1 O lado da demanda e a escolha do eleitor mediano

Considere I municipalidades (i = 1,..., I.). As preferências do eleitor mediano são dadas pela expressão $u_m(x_m, q)$, onde x_m é o consumo do bem privado e q representa o bem público local. A renda do eleitor mediano é y_m e sua base fiscal, b_m. Para uma alíquota τ, a restrição orçamentária do eleitor mediano pode ser escrita por $x_m + \tau b_m = y_m$. Os gastos públicos locais são denotados por E, e a base fiscal total é B. As receitas de transferências oriundas de outros níveis de governo são denotadas por S. Dessa forma, a restrição orçamentária total da jurisdição é $E = \tau B + S$. Finalmente, para uma população N, a função de congestionamento é expressa por $f(N)$ tal que $q = f(N)E$. Definindo-se as variáveis em termos *per capita*, tem-se que: $e = E/N$, $b = B/N$, e $s = \frac{S}{N}$. O problema do eleitor mediano torna-se, então:

$$\max_{x_m, q} u_m(x_m, q) \qquad sujeito\ a \quad x_m + \frac{q}{Nf(N)}\frac{b_m}{b} = y_m + s\frac{b_m}{b} \tag{1.1}$$

A condição de 1ª ordem associada ao problema acima dá o preço do serviço público (*tax-price*) para o eleitor mediano:

$$p_m = \frac{\frac{\partial u_m}{\partial q}}{\frac{\partial u_m}{\partial x_m}} = \frac{1}{Nf(N)}\frac{b_m}{b} \tag{1.2}$$

A demanda do eleitor mediano pelo bem público, q, sob a hipótese de log-linearidade, é:

$$q = c_1 p_m^{\varepsilon} y_m^{\beta_1} \left(s\frac{b_m}{b}\right)^{\beta_2} \tag{1.3}$$

Sendo c_1 uma constante; os coeficientes ε, β_1 e β_2 representam, respectivamente, a elasticidade preço, a elasticidade-renda, e a elasticidade das transferências. Substituindo a equação (1.2) em (1.3), tem-se:

$$q = c_1 \left(\frac{1}{Nf(N)} \frac{b_m}{b}\right)^{\varepsilon} y_m^{\beta_1} \left(s \frac{b_m}{b}\right)^{\beta_2} \tag{1.4}$$

Como q é uma variável não observável, usa-se $q = f(N)E = Nf(N)e$, obtêm-se:

$$e = c_1 [Nf(N)]^{-(1+\varepsilon)} \left(\frac{b_m}{b}\right)^{\varepsilon+\beta_2} y_m^{\beta_1} s^{\beta_2} \tag{1.5}$$

A equação acima, que representa a função estimável da despesa pública local *per capita*, mostra que essa despesa depende da renda do eleitor mediano (incluindo transferências), de sua capacidade contributiva e do fator de congestionamento, que caracteriza o uso do bem público.

1.2.2 O lado da oferta e as escolhas representativas

No modelo de oferta, os interesses próprios dos representantes eleitos prevalecem quando da determinação da provisão de bens públicos locais. Fatores típicos de oferta como entraves burocráticos, controle da agenda política, e complexidades fiscal e institucional concorrem para explicar os gastos públicos locais. Elementos de demanda, como renda e o *tax-price* do eleitor mediano não são relevantes para as escolhas municipais e, portanto, não aparecem na equação de gastos associados a este enfoque. A parcela fiscal agora é $(1/b)$ já que o representante não considera a capacidade contributiva dos indivíduos, mas apenas o seu nível médio, caracterizando uma medida de esforço fiscal uniforme.

Nesse modelo, o político eleito p maximiza sua função de popularidade $u_p(q, \tau)$ com $\partial u_p / \partial q > 0$ e $\partial u_p / \partial \tau < 0$. Como ele não leva em consideração a restrição orçamentária do eleitor mediano, então, o seu problema é:

$$\max_{q,\tau} u_p(q, \tau) \qquad sujeito\ a \qquad \frac{q}{Nf(N)} = \tau b + s \tag{1.6}$$

O *tax-price* com o qual o gestor público se confronta é:

$$p_p = \frac{\partial u_p / \partial q}{-\partial u_p / \partial \tau} = \frac{1}{Nf(N)} \frac{1}{b} \tag{1.7}$$

Assumindo novamente a log-linearidade, a função de gastos é:

$$q = c_2 p_p^{\varepsilon} s^{\beta_2} \tag{1.8}$$

Usando (1.7) e lembrando que $q = Nf(N)e$, então:

$$e = c_1 [Nf(N)]^{-(1+\varepsilon)} \left(\frac{1}{b}\right)^{\varepsilon} s^{\beta_2} \tag{1.9}$$

Esta é a função estimável de despesa pública local *per capita,* do lado da oferta. As equações (1.5) e (1.9) fornecem a justificativa teórica para as especificações econométricas que serão utilizas, expostas adiante, dos dois modelos concorrentes da provisão pública local. A diferença entre as abordagens vem da ausência no modelo de oferta da renda e da base tributária do eleitor mediano. As semelhanças nos dois modelos dizem respeito à influência de transferências governamentais, e à existência do efeito congestionamento.

1.3 DADOS

A amostra contém variáveis referentes a 5285 municípios brasileiros, do total de 5565, para o ano de 2010. A exclusão de algumas cidades deve-se à ausência de dados e a outros problemas detectados durante a coleta das informações. Além das variáveis básicas de cada modelo, foram adicionadas como variáveis explicativas, a área (A_i) de cada localidade, e *dummies* de capital ($DUMCP_i$), no sentido de captar, respectivamente, a distribuição desigual de terras entre municípios, a heterogeneidade dos dados, e a influência geográfica das sedes administrativas sobre o comportamento do gasto público municipal.

O Quadro 1.1, a seguir, apresenta a lista das variáveis utilizadas nesse estudo, bem como as suas descrições sintéticas.

Quadro 1.1: Descrição das variáveis utilizadas no estudo.

Variável	Notação	Descrição
Despesas locais com Serviços Públicos	e	Despesas totais *per capita* com os serviços públicos locais.
***Tax-price* ou Base Fiscal do Eleitor Mediano**	b_m/b_i	Razão entre renda mediana e renda média. Mede a sensibilidade da demanda em relação ao preço dos bens públicos.
Renda Mediana	y_{im}	Soma entre a renda mediana de cada município e o produto da multiplicação do *tax-price* pelo montante das transf. intergovernamentais recebidas.
Parcela fiscal do Modelo da Oferta	$1/b_i$	Razão entre o valor do imposto sobre a propriedade e a receita local.[5]
Transferências do governo	s_i	Participação das rendas provenientes de transf. intergovernamentais na renda do município.
População	N_i	Número de cidadãos em cada município analisado.
Área	A_i	Área de cada município.
***Dummy* Capital**	$DUMCP_i$	Assume o valor 1 se o município for capital de Estado; 0, de outra forma.

Fonte: Elaboração própria.

As principais fontes de informações advêm do Censo (2010) e da Base de Informações Municipais (BIM), coletadas junto ao Instituto Brasileiro de Geografia e Estatística (IBGE). Com relação às informações sobre as finanças públicas municipais – gastos, receitas e transferências - a fonte de informações é a Base de Finanças Municipais do Brasil, FIBRA, da Secretaria do Tesouro Nacional (STN). A Tabela 1.1 mostra as estatísticas descritivas das variáveis consideradas.

Tabela 1.1: Estatísticas descritivas

Variável	Observações	Média	Desv. Padrão	Mínimo	Máximo
e (R$)	5285	1504,32	709,34	127,69	10424,06
$y_m(R\$)$	5285	621,92	169,74	200,74	1800,30
b_m/b_i	5285	0,7128	0,1372	0,2645	1,2365
$1/b_i$	5285	0,4742	0,1740	0,026	0,8894
$s_i(R\$)$	5285	1367,76	845,84	148,91	12445,17
N_i	5285	34511,84	204811,10	1020	1,13e+07
$A_i(Km^2)$	5285	1525,77	5609,71	3,56	159533,25

Fonte: Elaboração própria a partir de dados da pesquisa.

[5] Seguindo a linha dos trabalhos de Guengant, Josselin e Rocaboy (2002), e Josselin, Rocaboy e Tavéra (2008).

Chama atenção a grande flutuação dos gastos correntes, os quais apresentam uma média de R$ 1.504,32 e das transferências entre os municípios analisados que oscilam entre um valor mínimo de R$ 148,91 e máximo de R$ 12.445,17 em termos *per capita*.

Feitas todas as considerações sobre a base de dados do trabalho e suas fontes, tem-se, a seguir, as especificações econométricas.

1.4 ESPECIFICAÇÕES ECONOMÉTRICAS

Lembrando que o subscrito *i* indexa as municipalidades, nas especificações econométricas, três formas funcionais distintas podem ser adotadas para a função de congestionamento:

i. Borcherding e Deacon (1972): $f(N) = N^{\gamma}, \; com. \; \gamma < 0.$ Note-se que nessa formulação, o parâmetro de congestionamento - elasticidade dos gastos em relação ao tamanho das cidades - é constante e igual a γ.

ii. Buchanan (1965): $f(N) = N^{\gamma'} e^{\gamma'' N}$. Esta função, conhecida como a função de *camaraderie*, permite que até certo ponto, o aumento do número de usuários eleve o nível das despesas públicas; depois desse limiar, predomina o efeito congestionamento que leva à redução desse nível; A elasticidade do tamanho da cidade é dada pela expressão:$\epsilon_{N=\gamma'+\gamma'' N}$;

iii. A forma flexível, proposta por Edwards (1990) é: $f(N) = \exp(\gamma' N + \gamma'' N^2 + \gamma''' N^3)$. A elasticidade do tamanho das cidades nessa formulação é $\epsilon_{N=}\gamma' N + 2\gamma'' N^2 + 3\gamma''' N^3$.

Substituindo as funções de congestionamento acima citadas em (1.5) e (1.9), as formas finais das equações do modelo de demanda (equações (1.10), (1.12) e (1.14)) e do modelo de oferta (equações (1.11), (1.13) e (1.15)) a serem estimadas são as seguintes:

Borcherding e Deacon (B-D):

Modelo do Eleitor Mediano – Regime (1)

$$ln(e_i) = K_1 + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + u_i \quad (1.10)$$

Modelo de Oferta - Regime (2)

$$ln(e_i) = K_2 + \varepsilon ln(1/b_i) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + v_i \quad (1.11)$$

***Camaraderie*:**

Modelo do Eleitor Mediano – Regime (1)

$$ln(e_i) = K_1 + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma' + 1) ln(N_i) - (\varepsilon + 1)\gamma''(N) + \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + u_i \quad (1.12)$$

Modelo de Oferta - Regime (2)

$$ln(e_i) = K_2 + \varepsilon ln\left(\frac{1}{b_i}\right) - (\varepsilon + 1)(\gamma' + 1) ln(N_i) - (\varepsilon + 1)\gamma''(N) + \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + v_i \quad (1.13)$$

Forma flexível:

Modelo do Eleitor Mediano – Regime (1)

$$ln(e_i) = K_1 + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1) ln(N_i) - (\varepsilon + 1)\gamma'(N) - (\varepsilon + 1)\gamma''(N^2) - (\varepsilon + 1)\gamma'''(N^3) + \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + u_i \quad (1.14)$$

Modelo de Oferta - Regime (2)

$$ln(e_i) = K_2 + \varepsilon ln(1/b_i) - (\varepsilon + 1) ln(N_i) - (\varepsilon + 1)\gamma'(N) - (\varepsilon + 1)\gamma''(N^2) - (\varepsilon + 1)\gamma'''(N^3) + \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + v_i \quad (1.15)$$

Resta, agora, testar os modelos (1.10) a (1.15) a fim de saber qual deles melhor se adequa às municipalidades brasileiras. Em particular, investiga-se a relação entre os níveis populacionais e a prevalência de uma dessas hipóteses – eleitor mediano e modelo de oferta – na explicação do dispêndio público local. Trata-se, pois, de determinar endogenamente o tamanho da população a partir do qual as preferências do eleitor mediano são ignoradas, e a provisão de bens públicos se dá conforme os interesses dos agentes políticos. Isto equivale a encontrar um limiar, definido em termos populacionais, que aponte a predominância de um modelo contra o seu concorrente. Para tal, desenvolveu-se uma estratégia de estimação, que combina o Teste J, proposto por Davidson e MacKinnon (1981), com um teste para instabilidade dos parâmetros e efeito limiar (*threshold effect*) sugerido por Hansen (2000). Abaixo, discute-se, brevemente, os referidos testes e a estratégia de combinação.

1.4.1 O teste *j e* o diagnóstico de modelos concorrentes

A ideia básica do Teste J *(J Test)*, proposto por Davidson e Mackinnon (1981), é testar um modelo frente a um ou mais concorrentes. Trata-se de um diagnóstico que permite ao pesquisador avaliar a adequação de um modelo a um conjunto de dados, auxiliando-o na escolha entre modelos alternativos. Supondo-se, por exemplo, o caso de duas hipóteses: $H0: y = X_0\hat{\theta}_0 + u_0$ e $H1: y = X_1\hat{\theta}_1 + u_1$, pode-se testar $contra\ H0$, por meio do seguinte procedimento:

a) Estima-se $H1$ e obtêm-se os valores ajustados ($\hat{y}_1 = X_1\hat{\theta}_1$);
b) Inclui-se $\hat{y}_1$como regressor em $H0$ e estima-se, então, o modelo ampliado $y = X_0\theta_0 + \lambda\hat{y}_1 + u$
c) Testa-se então a exclusão da variável adicionada. A rejeição da hipótese $\lambda = 0$ significa a rejeição de $H1$. Adota-se um procedimento análogo para se testar a hipótese $H0$ contra $H1$:
d) Estima-se $H0$ e obtêm-se os valores ajustados ($\hat{y}_0 = X_0\hat{\theta}_0$).
e) Inclui-se $\hat{y}_0$como regressor em $H1$ e estima-se então o modelo ampliado $y = X_1\theta_1 + \lambda'\hat{y}_0 + u$, em que $\lambda' = 1 - \lambda$.
f) Testa-se então a exclusão da variável adicionada. A rejeição da hipótese $\lambda' = 0$ implica a rejeição de $H0$.

Aqui, *H0* corresponde ao modelo do eleitor mediano, sendo o modelo de oferta a hipótese alternativa *H1*. A adequação do modelo ocorre quando a abordagem concorrente não melhora

o seu poder de explicação (DAVIDSON E MACKINNON, 1981, 1993). Note-se que essa abordagem pode levar à rejeição dos dois modelos, assim como a aceitação de ambos. Embora não desejáveis, esses resultados são informativos: se houver rejeição de todos os modelos, torna-se necessário encontrar outro que apresente melhor ajuste; caso nenhum deles seja rejeitado, os dados são compatíveis com ambas às hipóteses. (DAVIDSON E MACKINNON, 1993).

1.4.2 O efeito *threshold* e teste de identificação do parâmetro

Em análises econômicas, uma questão de interesse é investigar mudanças de regimes desencadeadas quando uma variável observada cruza um determinado limiar. Nesse âmbito, os modelos de regressão *threshold*, propostos por Howell Tong nos anos setenta e revisitados por Hansen (1996, 2000), vem sendo amplamente utilizados em economia. Embora simples, esses modelos capturam as dinâmicas e assimetrias particularmente relevantes para os problemas de identificação de modelos econômicos. Hansen (2000) desenvolveu um instrumental que permite dividir uma dada amostra, por meio de uma função indicadora, e testar a hipótese de uma formulação linear frente ao modelo de regressão *threshold*. O modelo assume que são gerados dados a partir de $\{y_i, x_i, N_i\}_{i=1}^{n}$, onde y_i são observações sobre a variável dependente, x_i é um vetor *px1* de variáveis independentes para cada unidade observada e (N_i) a variável *threshold*, que além de poder estar contida em x_i, divide a amostra em regimes, classes ou subgrupos. O modelo de Hansen é sumariado nas equações abaixo:

$$y_i = \theta_1' x_1 + e_i, \qquad N_i \leq \mu \tag{1.16}$$

$$y_i = \theta_2' x_1 + e_i, \qquad N_i > \mu \tag{1.17}$$

onde e_i é o termo de erro da regressão e μ o parâmetro que divide a amostra em dois grupos.

Definindo-se uma variável *dummy* $d_i(\mu) = I\{N_i \leq \mu\}$, em que $I\{.\}$ é a função indicadora, o modelo pode ser expresso em uma única equação. De fato, para $x_i(\mu) = x_i d_i(\mu)$, as equações (1.16) e (1.17) podem ser reescritas como:

$$y_i = \theta_2' x_1 + \delta_n' x_i(\mu) + e_i \tag{1.18}$$

onde $\delta_n' = \theta_2 - \theta_1$. Na forma matricial,

$$Y = X\theta + X_\mu \delta + e \tag{1.19}$$

Y e e são vetores (*n x 1*), e X e X_μ são matrizes (*n x m*).

Hansen (2000) propôs um procedimento que identifica os valores de μ e também computa estimativas para os parâmetros θ e δ por Mínimos Quadrados Ordinários (MQO). Definindo-se $S_n(\theta, \delta, \mu) = (Y - Y\theta - X_\mu \delta)'(Y - Y\theta - X_\mu \delta)$ como a função de soma de quadrados dos erros, por definição, os estimadores de MQO $\hat{\theta}, \hat{\delta}, \hat{\mu}$, minimizam esta soma. Supõe-se que μ é restrito a um conjunto limitado $\left[\underline{\mu}, \overline{\mu}\right] = \Gamma$.

Condicionando os parâmetros da equação (1.19) aos valores de μ, obtêm-se estimadores de Mínimos Quadrados Condicionados, $\hat{\theta}(\mu)$ e $\hat{\delta}_n(\mu)$, por meio da regressão de Y em $X_\mu^* = \left[X\ X_\mu\right]$. Desta feita, $S_n(\theta, \delta, \mu)$ escreve-se como:

$$S_n(\mu) = S_n\left(\hat{\theta}(\mu), \hat{\delta}(\mu), \mu\right) = Y'Y - Y'X_\mu^*(X_\mu^{*\prime} X_\mu^*)^{-1} X_\mu^{*\prime} Y \tag{1.20}$$

Para encontrar o valor de $\hat{\mu}$ que minimiza a equação (1.20), analisa-se os valores de μ, ou seja, sobre Γ, tal que $\hat{\mu} = \operatorname{argmin}_{\mu \epsilon \Gamma_n} S_n(\mu)$, onde $\Gamma_n = \left[\underline{\mu}, \overline{\mu}\right] \cap \{N_1, \dots, N_n\}$, e $\{N_1, \dots, N_n\}$ é a amostra de possíveis variáveis *threshold*. Após a determinação de $\hat{\mu}$, as estimativas de MQO de θ e μ são dadas por $\hat{\theta} = \hat{\theta}'(\hat{\mu})$ e $\hat{\delta} = \hat{\delta}(\hat{\mu})$.

A estimação do modelo confronta a hipótese nula de linearidade frente à especificação de existência do efeito *threshold*. No entanto, como o teste não tem uma distribuição assintótica padrão, e como o parâmetro μ não é identificado, Hansen (2000) sugeriu um procedimento *bootstrap* do Teste do Multiplicador de Lagrange (ML), consistente com heterocedasticidade, para tal análise. A sua estratégia tem como base a estatística de verossimilhança $LR_n(\mu)$, e a construção das regiões de confiança baseadas em $LR_n(\mu)$ é realizada pela definição de C como o nível de intervalo de confiança assintótico (C=0,95) e $c = c_\xi(C)$ como valor crítico no intervalo $\hat{\Gamma} = \{\mu: LR_n(\mu) \leq c\}$. Em termos gráfico, para encontrar a região $\hat{\Gamma}$, sob hipótese de homocedasticidade, grafa-se a taxa de verossimilhança $LR_n(\mu)$ contra μ, e traça-se uma linha horizontal em c. No caso da hipótese de homocedasticidade ser rejeitada, a estatística de verossimilhança pode ser definida como:

$$LR_n^*(\mu) = \frac{LR_n(\mu)}{\hat{\eta}^2} = \frac{S_n(\mu) - S_n(\hat{\mu})}{\hat{\sigma}^2 \hat{\eta}^2} \tag{1.21}$$

1.4.3 Combinando o teste j com o modelo *threshold*[6]

Para identificar a possibilidade de quebra da amostra em regimes e testar especificações alternativas para a determinação das despesas públicas, combinou-se o Teste J e o modelo *threshold*. Utilizou-se o Teste do Multiplicador de *Lagrange* (ML) consistente com a hipótese de heterocedasticidade, usando o *bootstrap* com 1000 replicações e analisou-se os valores-*p* estimados[7]. Testou-se ao mesmo tempo a possibilidade dos regimes serem gerados pelo modelo do eleitor mediano (HO) e/ou pelo modelo direcionado para a oferta ($H1$), com possíveis pontos de quebras entre eles. O teste conta com o auxílio da estatística *t* para a hipótese de $\lambda = 0$, na equação (1.22):

$$ln(e_i) = K_1 + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \beta_2 ln(s_i) + \quad (1.22)$$
$$\beta_3 ln(A_i) + \beta_4 DUMCP_i + \lambda ln(\hat{e}_i^S) + u_i$$

onde $ln(\hat{e}_i^S)$ é o vetor de valores ajustados a partir da regressão por Mínimos Quadrados Ordinários do modelo da oferta ($H1$). Assintoticamente, a estatística J tem distribuição normal (Davidson e MacKinnon, 1981), e como $I \to \infty$, se $H1$ é verdadeira, a probabilidade de que $\hat{\lambda}$ seja significantemente diferente de zero aproxima-se de 1. Em suma, testa-se as seguintes hipóteses:

$\boldsymbol{HO}$: O modelo do lado da demanda prevalece continuamente na amostra.

$\boldsymbol{H1}$: Abaixo de um limiar desconhecido μ, o modelo da demanda prevalece ($N_i \leq \mu$) enquanto que além deste ponto de quebra, o modelo da oferta prevalece ($N_i > \mu$).

Ou seja,

[6] Esta subseção explica a combinação dos testes com exposição, apenas, das equações a partir da forma funcional B-D, contudo, os mesmos procedimentos foram realizados com base nas especificações *camaraderie* e flexível.

[7] Como regra, sob a hipótese nula de que não há o efeito *threshold*, o *p-valor* computado deve permitir a rejeição da hipótese nula, e sugerir realmente o efeito limiar com base na população local. Caso contrário, identifica-se a linearidade da amostra, e a população é rejeitada como bom parâmetro para a análise.

$$\begin{cases} \boldsymbol{H0}: ln(e_i) = K_1 + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \\ \qquad \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + u_i & i = 1..I \\ \\ \boldsymbol{H1}: \begin{cases} ln(e_i) = K_1 + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \\ \qquad \beta_2 ln(s_i) + \beta_3 ln(A_i) + \beta_4 DUMCP_i + u_i & N_i \leq \mu \\ \\ ln(e_i) = K_2 + \varepsilon ln\left(\frac{1}{b_i}\right) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \beta_2 ln(s_i) + \beta_3 ln(A_i) \\ \qquad + \beta_4 DUMCP_i + v_i & N_i > \mu \end{cases} \end{cases}$$

A construção da técnica de estimação sugerida por Hansen (2000), passa pela definição da função indicadora $\psi_i(\mu) = \begin{cases} 0 \; para \; N_i \leq \mu \\ 1 \; para \; N_i > \mu \end{cases}$. Com isso, o modelo estimado para testar a hipótese HO contra $H1$ é:

$$ln(e_i) = K_1 + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \beta_2 ln(s_i) + \quad (1.23)$$
$$\beta_3 ln(A_i) + \beta_4 DUMCP_i + \psi \lambda ln(\hat{e}_i^S) + u_i$$

A última etapa define $S_1(\mu)$ como a soma dos quadrados dos resíduos com o nível *threshold* da variável N_i fixado em μ. Os estimadores de mínimos quadrados de μ são dados por $\hat{\mu} = \underset{\mu}{\operatorname{argmin}} \, S_1(\mu)$. Isto é importante para determinar a significância estatística do efeito *threshold.*

Então, a equação (1.23) testa a hipótese de linearidade, HO: $\lambda = 0$, contra a hipótese alternativa de efeito limiar. Como exposto na subseção (1.4.2), o teste não possui uma distribuição assintótica padrão, desta feita, fez-se uso do procedimento *bootstrap* para simular a distribuição do teste de razão de verossimilhança de HO: $LR_0 = (S_0 - S_1(\hat{\mu}))/\hat{\sigma}^2$, onde S_0 e $S_1(\mu)$ são os resíduos da soma dos quadrados sob $HO: \lambda = 0$ e $H1: \lambda \neq 0$, respectivamente; e $\hat{\sigma}^2$ é a variância dos resíduos sob $H1$. Hansen (1996, 2000) mostra que a melhor maneira para formar uma região de confiança para $\hat{\mu}$ é computar a "região de não rejeição" usando a estatística da razão da verossimilhança para testes em μ. Estes procedimentos foram utilizados neste estudo.

1.5 RESULTADOS

Para verificar se os modelos foram corretamente especificados e detectar possíveis problemas com as hipóteses do Modelo de Regressão Linear Clássico (MRLC), os seguintes testes foram realizados: teste de especificação do modelo (RESET), teste de normalidade (Jarque-Bera), teste de autocorrelação e o teste de heterocedasticidade de *White*. O Teste Jarque-Bera sugere a não-normalidade dos resíduos, contudo, Brooks (2002) ressalta que este tipo de ocorrência é frequente na modelagem baseada em dados econômicos, e que apelando para a teoria do limite central, a estatística teste seguirá assintoticamente a distribuição apropriada, mesmo na ausência da normalidade na distribuição dos erros. A hipótese de homocedasticidade foi rejeitada; por isso, as estimativas usam a estatística de *White*.

O teste do Multiplicador de Lagrange (*ML*), obtido por meio de 1000 passos de *bootstrap* e consistente para heterocedasticidade, computou um *p-valor* significante de 0,048, sugerindo a rejeição da hipótese nula de ausência do efeito *threshold*, e confirmando a divisão da amostra em dois grupos a partir do contingente populacional.

Testes adicionais foram realizados, para as duas subamostras (para as três formas de congestionamento) no intuito de descobrir outros possíveis pontos de quebra, contudo as estatísticas do *bootstrap* não foram significativas, conforme exposto na Tabela 1.2.

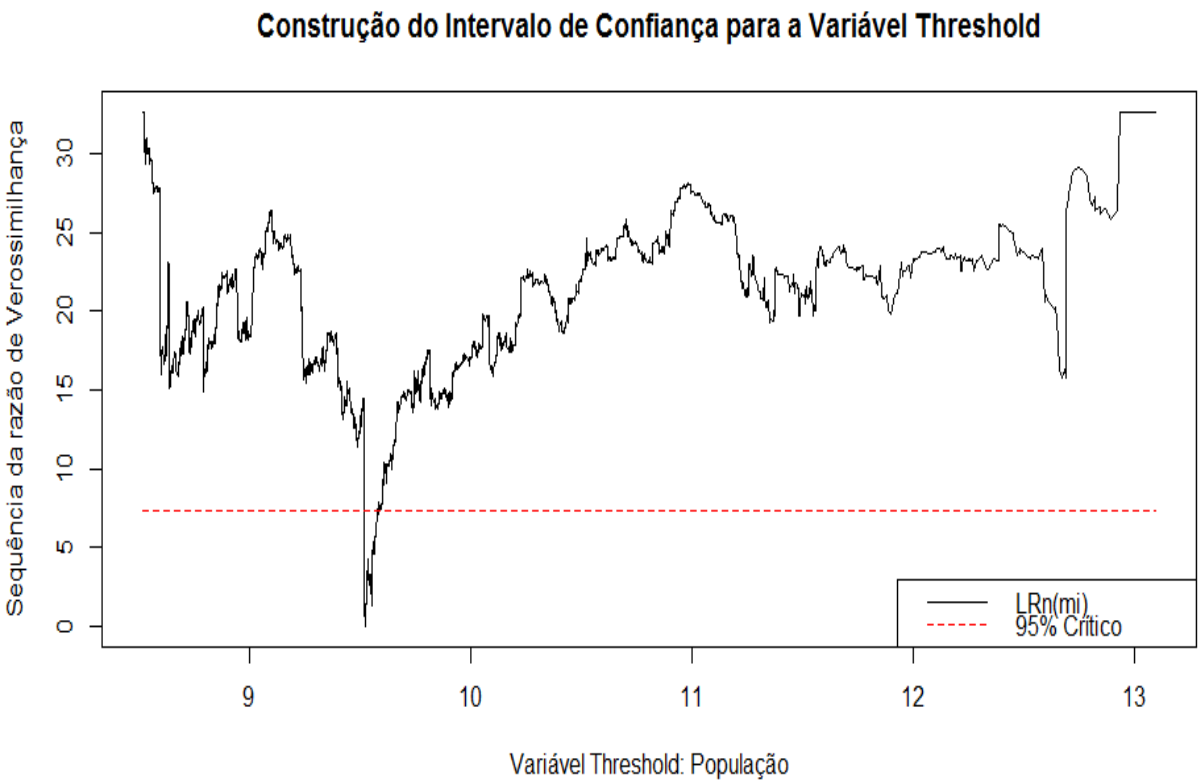
Tabela 1.2: Resultados dos testes de linearidade para as subamostras.

MODELO DA DEMANDA	MODELO DA OFERTA
Especificação B-D - drop if pop>13661 Threshold Estimado: 1228 Teste LM para não threshold: 6.907 **P-valor Bootstrap: .2841**	***Especificação B-D*** - drop if pop<=13661 Threshold Estimado: 169137 Teste LM para não threshold: 11.413 **P-valor Bootstrap: .1486**
Especificação Camaraderie - drop if pop>13604 Threshold Estimado: 1196 Teste LM para não threshold: 6.855 **P-valor Bootstrap: .1846**	**Especificação Camaraderie** - drop if pop<=13604 Threshold Estimado: 169101 Teste LM para não threshold: 11.318 **P-valor Bootstrap: .1624**
Especificação Flexível - drop if pop>13604 Threshold Estimado: 1196 Teste LM para não threshold: 6.851 **P-valor Bootstrap: .2011**	**Especificação Flexível** - drop if pop<=13604 Threshold Estimado: 169103 Teste LM para não threshold: 11.312 **P-valor Bootstrap: .1799**

Fonte: Elaboração própria a partir dos resultados do trabalho.

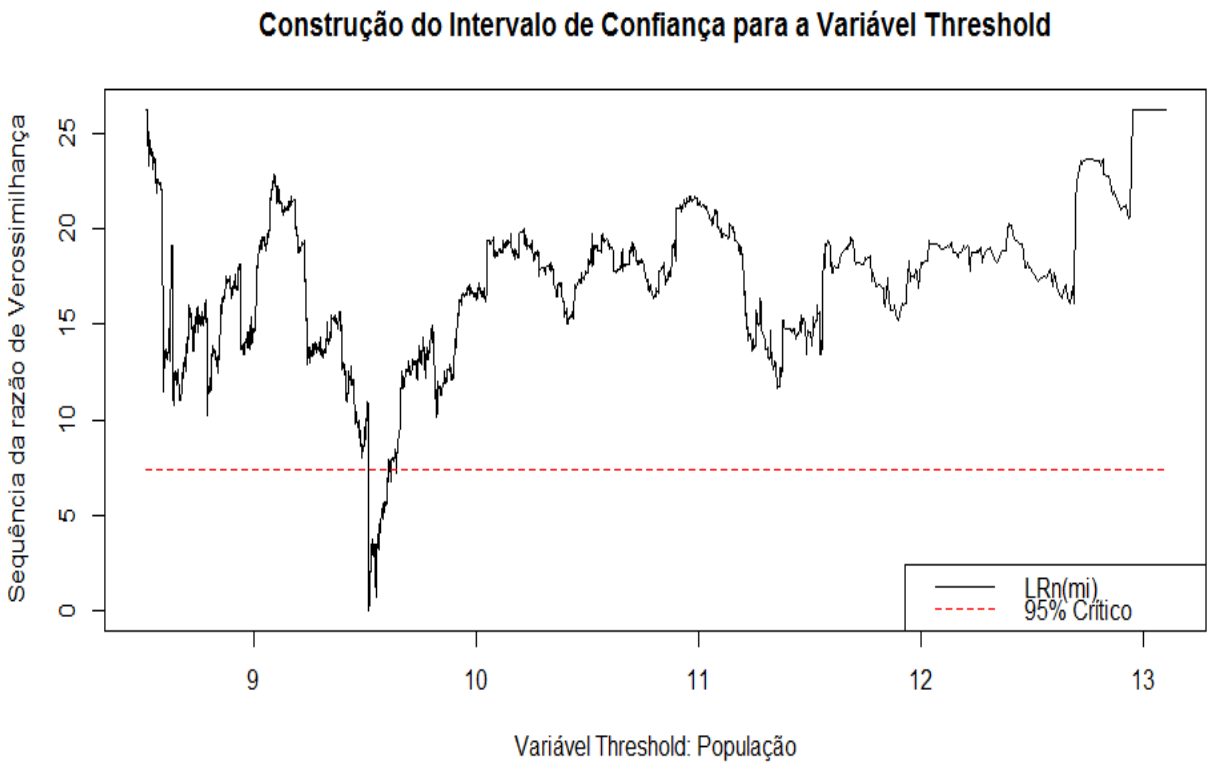
As Figuras 1.1 a 1.3 mostram o gráfico da estatística sequência da razão de verossimilhança normalizada $LR_n^*(\mu)$ em função da população local, para os três modelos considerados: Borcherding e Deacon, *camaraderie* e o modelo flexível. No modelo B-D, a estimativa por mínimos quadrados do parâmetro μ é o valor que minimiza o gráfico, e ocorre no ponto (em log) $\hat{\mu}$ = 9.5223, ou seja, o limiar corresponde a cidades com população local de 13661 habitantes. Note-se que até este patamar populacional, existem 2992 municípios, o que corresponde a 56,6% de toda a amostra. Também é possível visualizar o valor crítico assintótico de 7.35 (linha tracejada em vermelho) que define um intervalo para o limiar estimado com 95% de confiança. O conjunto [9.517384; 9.591513] define este intervalo, que em termos de população local representa [13594; 14640]. Para o modelo *camaraderie* (Figura 1.2), $\hat{\mu}$ = 9.518119, o que corresponde a uma população de 13604 habitantes, com intervalo de confiança [9.517384; 9.622707]. Já no modelo flexível, o limiar permanece o mesmo, alterando-se apenas o intervalo de confiança, que em termos de população local é [13594; 15103].

Figura 1.1: Teste para a variável *threshold*: Modelo Borcherding-Deacon.



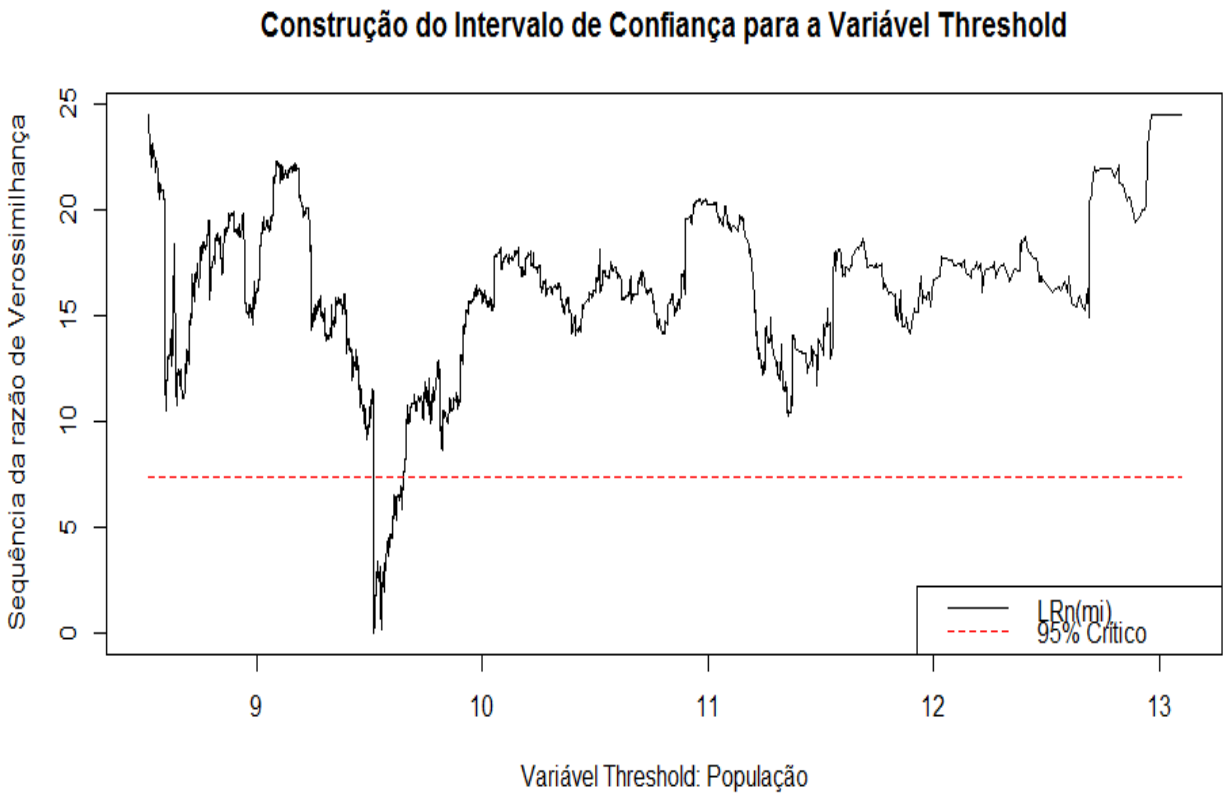
Fonte: Elaboração própria.

Figura 1.2: Teste para a variável *threshold*: modelo *camaraderie*.



Fonte: Elaboração própria.

Figura 1.3: Teste para a variável *threshold*: modelo flexível.



Fonte: Elaboração própria.

Os parâmetros também confirmam a adequação dos valores ajustados da equação do modelo da oferta frente ao modelo da demanda, com nível de 1% de significância (t=3,74*);* e a soma dos quadrados dos resíduos é 311,5 para o modelo da demanda, 492,21 para o modelo orientado a oferta e 267,01 para o modelo *threshold,* quando se utiliza a função de congestionamento B-D. As estatísticas referentes ao modelo *camaraderie* também confirmam a adequação dos valores ajustados da equação do modelo da oferta frente ao modelo da demanda, com nível de 1% de significância (t=3,18*)*. Já a soma dos quadrados dos resíduos: 386,24 para o modelo da demanda, 518,03 para o modelo da oferta e 307,55 para o modelo *threshold.* E no modelo flexível, (t=3,06*);* a soma dos quadrados dos resíduos é 391,33 para o modelo da demanda, 523,14 para o modelo de oferta e 309,80 para o modelo *threshold*.

Os resultados das estimações dos modelos analisados, expostos nas Tabelas 1.3 a 1.5, completam a análise. Nas três tabelas, a primeira coluna apresenta os coeficientes estimados do modelo completo (*threshold),* a segunda coluna mostra os coeficientes para o modelo do eleitor mediano e a terceira coluna contém os parâmetros estimados do modelo de oferta.

Note-se em primeiro lugar que os resultados obtidos são relativamente independentes da função de congestionamento escolhida. Em todos os modelos considerados, os resultados sugerem a rejeição de *H0,* e aceitação da hipótese alternativa *H1,* a qual prediz com base no limiar μ, que o modelo da demanda prevalece no intervalo ($N_i \leq \mu$), enquanto que além deste ponto de quebra ($N_i > \mu$) o modelo da oferta apresenta melhor desempenho. De outra forma, o modelo do eleitor mediano aplica-se às municipalidades, cuja população é inferior a 15000 habitantes; de fato, esse limiar varia entre 13604 (modelos *camaraderie* e flexível) e 13661 habitantes (modelo B-D), dentro dos intervalos de confiança computados. Acima desses limiares prevalece o modelo de oferta.

Ademais, o sinal e a significância do termo $\psi(N_i > \mu) ln(e_i^S)$ sugere que os modelos de oferta e demanda são complementares para explicar os gastos públicos locais per capita.

A partir desses resultados, pode-se fornecer explicações exploratórias para as especificidades no comportamento da despesa pública, a partir das duas modelagens. Corroborando estudos anteriores (Turnbull e Mitias, 1999), percebe-se um relevante aumento da complexidade e comprometimento da estrutura orçamental, quando o tamanho da cidade aumenta. No Brasil, orçamentos maiores, certamente, aumentam a ilusão fiscal e, possivelmente, geram uma sensação de afastamento das decisões que resulta em maior independência para o gestor municipal. Esta independência pode ser não intencional, como uma espécie de efeito de isolamento, ou estratégica (como em Denzau e Mackay, 1983).

Observando-se as Tabelas 1.3 a 1.5, percebe-se que todos os parâmetros são significativos e apresentam os sinais esperados. O valor estimado para o coeficiente da variável *tax price* - negativo e significante – é relativamente baixo, que sugere a existência de inelasticidade da demanda, o que caracteriza os bens públicos locais como bens essenciais para a população. Já no modelo da oferta, o coeficiente atrelado à parcela fiscal é positivo, o que vai ao encontro das predições da lei geral da oferta: a oferta de bens públicos está diretamente relacionada com o preço do serviço pago pela comunidade.

Tabela 1.3: Resultados das estimações do Modelo *Threshold* - Modelo B-D.

	Regressão do Modelo *Threshold*	**Regime 1** $\boldsymbol{N_i \leq 13661}$	**Regime 2** $\boldsymbol{N_i > 13661}$
Elasticidades			
$ln\left({b_{im}}/{b_i}\right)$	- 0,0829* (0,0146)	- 0,0613* (0,0172)	
$ln\left({1}/{b_i}\right)$			0,0841*(0,0174)
$ln(y_i)$	0,5011* (0,0790)	0,4613* (0,0945)	
$ln(s_i)$	0,5517* (0,0430)	0,5234* (0,0378)	0,4698*(0,1071)
$ln(N_i)$	- 0,1197 *(0,0264)	- 0,1422* (0,0336)	- 0,2932*(0,0834)
$ln(A_i)$	- 0,0721* (0,0149)	0,0112 (0,0201)	- 0,0533*(0,0109)
DUM_i	0,5921* (0,0028)	0,6048* (0,0071)	0,6927*(0,0055)
$\boldsymbol{\psi(N_i > 13661)ln(e_i^S)}$	0,0058 *(0,0001)		
Constante	2,0096*(0,0862)	2,3318* (0,1708)	1,6095*(0,0783)
R^2 Ajustado	0,7405	0,7194	0,8566
Nº de Observações	5285	2992	2293
Desvio Padrão da Regressão	0,1788	0,1522	0,1063

Fonte: Elaboração própria a partir dos resultados do trabalho. **Nota:** * significante ao nível de 1%. **Obs**. Os valores entre parênteses são os erros padrões.

No que concerne às transferências, os coeficientes atrelados a essa variável são positivos e significantes em todos os modelos testados. Ademais, tanto no modelo *threshold* como no modelo do eleitor mediano, as elasticidades do gasto em relação às transferências são superiores às elasticidades renda, o que aponta para a existência do efeito *flypaper*. De acordo com esse efeito, a separação entre a arrecadação dos municípios e seus gastos, viabilizada pelas transferências incondicionais, leva os contribuintes a subestimarem o custo real dos bens públicos locais, o que contribui para subestimá-lo, resultando em aumentos da demanda e dos gastos com esses serviços.

Note-se, por fim, que a maior sensibilidade das despesas com relação às transferências no Regime 1 corrobora estudos recentes (Buettner e Fabritz, 2011; Dahlby e Ferede, 2012). Estes estudos mostram que o impacto das transferências sobre os gastos públicos locais é maior nas pequenas jurisdições porque o custo marginal dos fundos públicos (CMFP) é mais elevado nessas localidades, que dispõem de bases tributárias mais restritas; nesse caso, as transferências viabilizam a provisão de serviços públicos, que não poderiam ser financiados por suas receitas próprias.

Tabela 1.4: Resultados das estimações do modelo *Threshold – Camaraderie.*

	Regressão do Modelo *Threshold*	**Regime 1** $N_i \leq 13604$	**Regime 2** $N_i > 13604$
Elasticidades			
$ln\left({b_{im}}/{b_i}\right)$	- 0,0804* (0,0129)	- 0,0593* (0,0157)	
$ln\left({1}/{b_i}\right)$			0,0864*(0,0181)
$ln(y_i)$	0,5008* (0,0715)	0,4590* (0,0879)	
$ln(s_i)$	0,5506* (0,0388)	0,5226* (0,0321)	0,4703*(0,1097)
$ln(N_i)$	- 0,1156 *(0,0201)	- 0,1448* (0,0320)	- 0,2937*(0,0848)
$ln(A_i)$	- 0,0707* (0,0125)	0,0113 (0,0199)	- 0,0534*(0,0109)
N	1,07E-07*(1,21E-08)	3,37E-08*(4,68E-09)	7,46E-09*(6,42E-10)
DUM_i	0,5921* (0,0028)	0,6077* (0,0068)	0,6930 *(0,0058)
$\boldsymbol{\psi(N_i > 13604)ln(e_i^S)}$	0,0063 *(0,0001)		
Constante	2,0113*(0,0913)	2,5678* (0,1926)	1,6314*(0,0814)
R^2 Ajustado	0,7478	0,7216	0,8598
Nº de Observações	5285	2969	2316
Desvio Padrão da Regressão	0,1813	0,1581	0,1089

Fonte: Elaboração própria a partir dos resultados do trabalho. **Nota:** * significante ao nível de 1%. **Obs.:** Os valores em parênteses são os erros padrões.

Já no modelo da oferta, que engloba os maiores municípios, a magnitude do parâmetro sugere que suas economias mais desenvolvidas têm maior potencial produtivo em atividades, como comércio, indústria e serviços, que são bases de incidência dos impostos municipais; consequentemente essas cidades são menos dependentes das transferências de outras esferas de governo. Como o financiamento dos gastos se dá, além dos repasses intergovernamentais, a partir do esforço fiscal sobre a própria base tributária, é plausível supor que em grandes cidades há maior volume de recursos próprios, capazes de gerar uma maior (ou melhor) provisão de serviços.

Tabela 1.5: Resultados das estimações do Modelo *Threshold* – Modelo Flexível

	Regressão do Modelo *Threshold*	**Regime 1** $N_i \leq 13604$	**Regime 2** $N_i > 13604$
Elasticidades			
$ln\left({b_{im}}/{b_i}\right)$	- 0,0802* (0,0126)	- 0,0592* (0,0155)	
$ln\left({1}/{b_i}\right)$			0,0866*(0,0183)
$ln(y_i)$	0,5002* (0,0719)	0,4589* (0,0867)	
$ln(s_i)$	0,5500* (0,0376)	0,5223* (0,0313)	0,4708*(0,1102)
$ln(N_i)$	- 0,3698 *(0,0198)	- 0,4185* (0,0314)	- 1,0867*(0,0834)
$ln(A_i)$	- 0,0706* (0,0123)	0,0111 (0,0195)	- 0,0532*(0,0107)
N	6,08E-06*(1,15E-07)	3,78E-05*(4,87E-06)	2,16E-05*(1,49E-06)
N^2	- 4,21E-11*(9,81E-12)	- 1,57E-11*(3,46E-12)	- 2,41E-10*(1,07E-11)
N^3	4,25E-14*(1,08E-15)	5,03E-14*(3,11E-15)	9,68E-15*(7,88E-16)
DUM_i	0,5938* (0,0032)	0,6096* (0,0074)	0,6938 *(0,0060)
$\boldsymbol{\psi(N_i > 13604)ln(e_i^S)}$	0,0069 *(0,0001)		
Constante	2,0205*(0,0920)	2,5835* (0,2004)	1,6514*(0,0935)
R^2 Ajustado	0,7507	0,7293	0,8640
Nº de Observações	5285	2969	2316
Desvio Padrão da Regressão	0,1821	0,1598	0,1104

Fonte: : Elaboração própria a partir dos resultados do trabalho. **Nota:** * significante ao nível de 1%. **Obs.:** Os valores em parênteses são os erros padrões.

Com relação à população, esta apresenta efeito negativo sobre as despesas públicas *per capita*, em todos os regimes, o que contraria a chamada "Lei de Brecht"[8], segundo a qual áreas densamente povoadas apresentam gastos públicos mais elevados. Não obstante, uma nova literatura (Wildasin, 1989; Reiter e Weicherieder, 1997) fundamenta esses resultados a partir da utilização das funções de congestionamento. Conforme essas abordagens, o impacto da população sobre a despesa em termos *per capita* depende da tecnologia de produção dos bens públicos. Dessa forma, devido à existência de economias de escala, é possível que a despesa *per capita* seja relacionada inversamente com a população, e em alguns casos, com a densidade demográfica.

Em linhas gerais, os resultados fortalecem o argumento de que em ambos os modelos, o comportamento das despesas públicas varia com a população. No Regime 1, nas formulações B-D e *camaraderie*, pode-se inferir que o crescimento de 10% no número de

[8] Para detalhes ver Reiter e Weicherieder (1997).

habitantes desses municípios, reduz a demanda *per capita* por bens e serviços em torno de 1,5%; já nas localidades inseridas no segundo regime, um aumento de 10% do contingente populacional ocasiona uma redução da oferta pública de aproximadamente 3%, em termos *per capita*. Note-se o maior efeito da variação populacional sobre as despesas no caso da especificação flexível, em decorrência do fato de que essa formulação não impõe nenhum perfil ao tamanho da população para análise do efeito congestionamento.

Com respeito à área do território, há influência negativa desta variável para o modelo combinado e para o Regime 2, e positiva para o modelo da demanda. A maior despesa nesses municípios inseridos no Regime 1 pode ser atribuída à presença de retornos crescentes de escala prevalecente em pequenos municípios. Por fim, no tocante aos efeitos de localização, o fato da cidade ser capital influencia fortemente os padrões de despesa é interessante. Este resultado confirma o papel exercido pelas capitais, que são os centros do poder político local, centralizando as pressões por maiores gastos públicos.

Outro ponto refere-se aos efeitos da urbanização: Como a população cresce, a gama de serviços para uma determinada função de gastos é fornecida por um número crescente de jurisdições sobrepostas; contudo, os municípios devem ter basicamente as mesmas prerrogativas, seja qual for o tamanho de sua população, já que é um pressuposto constitucional. No entanto, os políticos tomam as suas decisões em ambientes bastante diferentes, e proporcionais ao tamanho da comunidade em que foram eleitos. O que fica evidente, de fato, é a percepção de que à medida que os municípios crescem o gestor público se afasta cada vez mais das preferências do eleitor mediano, e passam a ofertar bens e serviços públicos que correspondam ao interesse da agenda política e da máquina estatal.

Completando as análises, a Tabela 1.6 apresenta as elasticidades-preço, as elasticidades-renda[9], os parâmetros de congestionamento e as elasticidades da qualidade do gasto público com respeito à população, para os três modelos testados. Note-se, em primeiro lugar, que as elasticidades preço, calculadas a partir dos coeficientes da parcela fiscal e das transferências são, respectivamente, $\varepsilon = \hat{\alpha}_2 - \hat{\alpha}_3 = -0{,}06 - 0{,}52 = -0{,}58$, para o Regime 1 (modelo de Demanda) , e $\varepsilon = \hat{\alpha}_2 = 0{,}08$ para o Regime 2 (Modelo de oferta), no modelo B-D. Em ambos os regimes, observa-se que tanto a demanda como a oferta dos gastos públicos

[9] O Teste de Hausman foi realizado para testar a possível endogeneidade da variável renda, e o *p*-valor computado de 0,157 permitiu a não rejeição da hipótese nula de exogeneidade. A estatística para o teste é $(\beta^{VI} - \beta^{MQO})'[V(\beta^{VI}) - V(\beta^{MQO})]^{-1}(\beta^{VI} - \beta^{MQO})$, onde $(\beta^{VI} - \beta^{MQO})$ é a diferença entre os parâmetros dos estimadores de variáveis instrumentais e de mínimos quadrados ordinários, e $[V(\beta^{VI}) - V(\beta^{MQO})]$ é a diferença entre a matriz de covariância dos estimadores VI e MQO. As variáveis utilizadas como instrumentos foram a distância em linha reta (Km) das cidades à capital de seu estado e a altitude dos municípios, em metros.

locais no Brasil são inelásticas; porém, essa inelasticidade é muito maior no modelo de oferta. Resultados similares foram obtidos para os modelos *camaraderie* e flexível.

As elasticidades-renda são positivas e inferiores à unidade, tanto no modelo *threshold* como no modelo do eleitor mediano. Essas estimativas encontram-se dentro dos padrões internacionais (Doi, 1998) e corroboram estudos anteriores (Mendes e Sampaio de Sousa, 2006) que atestam que o bem público agregado (*composite local public good*) é um bem normal e não um bem de luxo, conforme sugerido por alguns autores.

As elasticidades com respeito ao tamanho das cidades (*cities size elasticies*), ϵ_N, computadas a partir dos parâmetros de congestionamento, são negativas e inferiores à unidade, em todos os modelos considerados. Estes valores são claramente inferiores aos resultados encontrados na literatura (Reiter e Weichenrieder, 1997; Guengant, Josselin e Rocaboy, 2002), em que o valor desse parâmetro situa-se em torno de -1, implicando, assim, efeitos-congestionamento substanciais.

Tabela 1.6: Elasticidades e parâmetros de congestionamento por forma funcional.

Observações	**Regressão do Modelo *Threshold***	**Regime 1**	**Regime 2**
Borcherding-Deacon			
Elasticidades			
$Preço\ (\varepsilon)$	- 0,6346*	- 0,5847*	0,0841*
$Renda(\beta_1)$	0,5011*	0,4613*	
Elasticidade do Tamanho das Cidades			
ϵ_N	- 0,6724*	- 0,6576*	- 0,7295*
Camaraderie			
Elasticidades			
$Preço\ (\varepsilon)$	- 0,6310*	- 0,5819*	0,0864*
$Renda(\beta_1)$	0,5008*	0,4590*	
Elasticidade do Tamanho das Cidades			
ϵ_N	- 0,6898*	- 0,6547*	- 0,7295*
Flexível			
Elasticidades			
$Preço\ (\varepsilon)$	- 0,6302*	- 0,5815*	0,0866*
$Renda(\beta_1)$	0,5002*	0,4589*	
Elasticidade do Tamanho das Cidades			
ϵ_N	- 0,6687*	- 0,5917*	- 0,7126*
Mediana da população	11041	5800	26488

Fonte: Elaboração própria a partir de resultados do trabalho

O tamanho sub-ótimo da maioria das municipalidades brasileiras restringe a exploração de economias de escala, que caracteriza a produção de serviços públicos limitando, assim, a redução dos custos médios dos serviços públicos e o uso eficiente desses recursos. Não surpreendentemente, o valor estimado de ϵ_N no modelo de oferta – que se aplica aos municípios cuja população é superior a 13661 habitantes (13604 nos demais especificações) – é mais elevado, indicando que nessas cidades a rivalidade - portanto, o congestionamento - dos bens públicos locais é maior.

1.6 CONSIDERAÇÕES FINAIS

Neste trabalho testou-se a hipótese de que a partir de um dado limiar populacional, os gastos públicos locais não mais se norteiam pelas preferências do eleitor mediano, mas são determinados também de maneira discricionária pelos gestores eleitos, privilegiando interesses próprios em detrimento daqueles da população. Para tanto, utilizou-se a análise *threshold* combinada com o teste J para estimar duas especificações alternativas de indução dos gastos públicos – o modelo de demanda, baseado na hipótese do eleitor mediano (Regime 1), e o modelo de oferta (Regime 2), em que predominam motivações políticas. Ênfase especial foi dada à investigação do impacto do tamanho das cidades sobre a qualidade do gasto público, que leva em conta o efeito congestionamento. Três distintas funções de congestionamento foram testadas, a saber: parâmetro de congestionamento proposto por Borcherding-Deacon, a função *camaraderie,* e o modelo flexível, sugerido por Edwards (1990).

Os resultados confirmam a hipótese segundo a qual, o Regime 1 aplica-se às cidades menores; neste caso, as características do eleitor mediano determinam os gastos públicos. O tamanho ótimo da população, para o modelo da demanda, situa-se entre 13604 e 13661 habitantes, dependendo da função de congestionamento usada, o que atesta a robustez das estimativas face a essas especificações. Para cidades, cuja população é superior a esses limiares, prevalece o lado da oferta. Rejeitou-se, pois, a hipótese de que o modelo de demanda aplica-se a todos os níveis populacionais.

Com relação aos parâmetros estimados, as baixas elasticidades-preço mostraram que os serviços públicos são considerados mais essenciais, nos pequenos municípios. Nos maiores municípios, em que prevalece o Regime 2 (Oferta), a maior inelasticidade atesta que o financiamento dos gastos públicos é mais independente da parcela fiscal, já que estas cidades caracterizam-se pela existência de maior diversificação da base tributária. As elasticidades-

renda, positivas e inferiores à unidade, indicam que os gastos municipais se apresentam dentro dos padrões de serviços normais. Sobre as transferências, a magnitude dos parâmetros sugere uma alta dependência das localidades menores com relação a estes recursos para financiar a provisão de bens, contudo o desenvolvimento das municipalidades no contexto econômico possibilita maior base de incidência de tributos e consequentemente maior arrecadação, para dispêndio com verbas próprias.

As elasticidades com respeito ao tamanho das cidades (*cities size elasticities)* são negativas e inferiores à unidade, em todos os modelos considerados, indicando que os bens e serviços ofertados pelas jurisdições têm características de bem público. Não surpreendentemente, o valor estimado de $\epsilon_{N,}$ no modelo de oferta é mais elevado, indicando que nessas cidades a rivalidade dos bens públicos locais é maior.

De uma forma geral, os resultados do trabalho sugerem que à medida que em cidades maiores, os representantes eleitos se afastam de um compromisso rigoroso com a escolha do eleitor mediano, e passam a prover bens públicos de forma discricionária, no sentido de maximizar o seu interesse. O desenvolvimento natural da pesquisa envolve a introdução de variáveis específicas que caracterizem impacto direto na função demanda e oferta da provisão pública, inseridas em cada regime, e partindo do estudo de parâmetros que possam descrever a qualidade dos serviços públicos fornecidos à sociedade brasileira, via análise do congestionamento, considerando clubes de municípios – pelo tamanho de sua população – classes de despesas, bem como a interação espacial entre os municípios, ou possíveis efeitos *spillovers.*

2. A QUALIDADE NA PROVISÃO LOCAL DE BENS PÚBLICOS NO BRASIL: UMA ANÁLISE DO CONGESTIONAMENTO BASEADA NA TEORIA DOS CLUBES

2.1 INTRODUÇÃO

No capítulo anterior foram tratados aspectos empíricos que permitem estabelecer algumas relações determinantes para as atribuições municipais na provisão de serviços públicos. Confirmou-se a hipótese de que o nível e exclusividade dos bens públicos fornecidos localmente dependem do tamanho da população, e que o eleitor mediano perde poder representativo nas decisões sobre políticas de gastos públicos em grandes cidades. A justificativa pode estar no fato de que, nessas localidades, os gestores estão distantes do seu eleitorado, e assim, não observam as preferências da comunidade, além de enfraquecer o poder de barganha da população e fiscalização do erário público.

Esses resultados são importantes e geram discussões a respeito da qualidade do serviço público fornecido pelos governos locais, não só em grandes cidades onde supostamente a hipótese do eleitor mediano é menos relevante e a política pública discricionária, mas também em pequenas comunidades a partir do seu desenvolvimento, fonte de migração e adição de novos usuários dos bens fornecidos.

Baseado nisso, o objetivo principal deste capítulo é aplicar os modelos de provisão pública a fim de mensurar o impacto do tamanho das cidades sobre a qualidade dos serviços públicos - via análise da publicidade e da rivalidade dos bens - nas municipalidades brasileiras no âmbito da Teoria dos Clubes. Diferentes estimações são realizadas. As primeiras referem-se ao modelo da demanda por serviços públicos para o conjunto de 5268 municípios brasileiros em 2010; em seguida, as estimativas consideram o modelo da oferta; e por fim, as avaliações levam em conta a competição exógena entre os dois modelos para análise do melhor ajuste da determinação da provisão pública, referente a intervalos populacionais.

Entende-se que os estudos atuais sobre o congestionamento podem não permitir a compreensão plena desses efeitos. Se for considerado, por exemplo, o crescimento das cidades, elas podem, em algumas fases do seu desenvolvimento, não apresentar congestionamento, mas pelo contrário, beneficiar-se de economias de escala derivadas da existência dos bens públicos locais. A análise no contexto da teoria dos clubes permite simular a entrada de novos membros a fim de trazer uma visão mais consistente do funcionamento do mercado, para sustentação ou não da publicidade dos bens.

Salienta-se que esta versão apresenta modelos ampliados com a inserção de variáveis específicas a cada caso, adaptando elementos teóricos, empíricos e institucionais. Foram consideradas as três diferentes formas funcionais capazes de captar o efeito congestionamento (ou escala) na provisão de serviços em resposta ao seu crescimento, além dos efeitos *spillovers*, para captação da interdependência ou assimetria de relações espaciais.

A metodologia econométrica emprega, além do modelo clássico linear, a análise espacial, proposta por Anselin (1988), a qual vem sendo amplamente utilizada na área de estudos regionais. Adicionalmente, compara-se esses padrões através do Teste J (Davidson e Mackinnon, 1981) ao ponto de identificar o seu poder explicativo como determinantes para toda a gama de serviços providos pelos clubes.

Para tal, além dessa introdução, o capítulo está dividido em sete seções. Na segunda seção tem-se uma breve resenha sobre a Teoria dos Clubes. A seguir, na terceira e quarta seção, apresentam-se a modelagem teórica utilizada a partir de diferentes formas funcionais que captam os efeitos de congestionamento dos clubes sobre a provisão de bens. Na quinta seção estão expostas informações sobre as variáveis utilizadas; Na seção seis, têm-se os procedimentos metodológicos utilizados na pesquisa. Na sétima são analisados os resultados das estimações realizadas. Por fim, têm-se as considerações finais, onde estão dispostas as conclusões.

2.2 UMA BREVE RESENHA SOBRE A TEORIA DOS CLUBES

Um clube é um grupo voluntário que compartilha benefícios mútuos entre seus membros referentes a um ou mais procedimentos característicos dos sócios, ou dos serviços providos entre eles. Apesar da maioria de artigos econômicos analisarem os clubes a partir da obra "*An Economic Theory of Clubs*", de Buchanan (1965), a origem de sua teoria pode ser traçada a partir de Pigou (1920) e Knight (1924) e seus estudos sobre a possibilidade de uso de duas rotas alternativas por parte de uma comunidade, sendo uma congestionada de boa qualidade e outra não congestionada, porém de qualidade ruim. Ao sugerir pedágios na estrada congestionada, os autores essencialmente buscavam resolver um problema de clubes, uma vez que o pedágio restringiria os usuários e, assim, determinaria "o tamanho ótimo da associação" para a rodovia congestionada[10].

[10] Mishan (1971) mostrou que as soluções de Pigou (1920) e Knight (1924) são idênticas, sendo que Pigou considerou os custos marginais de congestionamento excluindo o conceito de aluguel, enquanto que Knight usou

Outros pioneiros sobre o modelo intuitivo de clubes foram Tiebout (1956), cuja hipótese "*voting with the feet*" (votando com os pés) atentava em mostrar como o tamanho jurisdicional de um governo local poderia determinar as decisões voluntárias de mobilidade (ou associação) dos membros da comunidade, e Wiseman (1957), que no contexto de bem privado, analisou o princípio dos clubes para compartilhamento de custos entre usuários de serviços públicos. Com efeito, Tiebout (1956) e Wiseman (1957) foram os primeiros pesquisadores a focarem na lógica de partilha dos custos por clubes supondo que a despesa por usuário cairia em conformidade ao tamanho da associação.

Contudo, o mais influente sobre as investigações iniciais da teoria é mesmo Buchanan (1965). Partindo da tentativa de preencher a lacuna dos estudos Samuelsonianos dos bens públicos e privados, ele foca em clubes que partilham um bem público parcialmente rival, caracterizado por benefícios excludentes, e adiante promove a primeira preposição analítica das condições de oferta e associações derivadas para clubes que compartilhavam bens. Identificou economias de escala a serem exploradas na provisão de serviços locais, e distinguiu os clubes de acordo com as características do uso dos serviços compartilhados. Clubes Inclusivos seriam aqueles que compartilhavam bens públicos puros e não requeriam restrições sobre o tamanho, enquanto que os Clubes Exclusivos seriam os que compartilhavam bens públicos impuros e requeriam restrições de tamanho devido à superlotação ou congestionamento. Esses passaram a ser chamados de "Bens de Clubes".

Segundo o autor, esses bens são caracterizados diferentemente dos bens públicos puros, por uma série de razões. Primeiro, porque o uso de um bem de clube é voluntário onde os participantes devem se juntar para receber seus benefícios, os quais são retidos para não membros, e apesar dos custos, os sócios percebem um ganho líquido pela adesão. Já para os bens públicos puros, todos dentro da gama de serviços distribuídos recebem automaticamente o benefício do bem ou o custo pela prestação. Em segundo lugar, o número ideal de participantes do clube que compartilham os bens é finito, ao contrário dos bens públicos puros, onde todos podem ser acomodados sem externalidades ou *crowding effect*. Os clubes são, portanto, coletivos exclusivos.

Em terceiro lugar, para os bens de clube, a disposição de adesão para os não membros deve ser abordada: Formar vários clubes, onde a população é dividida em clubes idênticos que não se sobrepõem (Pauly 1967; McGuire 1974), ou formar um único clube, onde os indivíduos que ficam de fora não consomem o bem? (HELPMAN E HILLMAN, 1977).

---

os custos médios incluindo a possibilidade de aluguel. Contudo, sob as condições de mercado competitivo as duas análises convergiam.

Em quarto lugar, os bens de clube devem possuir um mecanismo de exclusão que é praticamente sem custos, contudo, os membros que não cooperam não recebem os seus benefícios. Em contraste, a exclusão não é possível nem desejável para um bem público puro.

Em quinto, ao contrário de bens públicos puros, os bens de clube são muitas vezes providos com a internalização do processo de congestionamento, e eficientemente através de acordos institucionais alternativos, e por fim, envolvem uma dupla decisão - a escolha da provisão e o tamanho de seus membros.

Os bens de clube não representam todos os bens entre os extremos polares dos privados e públicos puros. Não há escala linear entre esses dois extremos, na medida em que as propriedades de caráter público realmente variam ao longo de duas dimensões - a rivalidade e a possibilidade de exclusão. Assim, os bens públicos que são não rivais, mas excludentes, não são bens do clube, porque o custo marginal de usuários adicionais é zero. Como tal, a exclusão não alcança a eficiência (SANDLER, 2004). Os bens públicos que estão sujeitos à aglomeração, mas não podem ser excluídos também não são bens de clube, o que igualmente é válido para bens públicos cujo custo supera exclusão de quaisquer ganhos de eficiência, de modo que a exclusão não se justifica sob realidades tecnológicas atuais.

Assim, os bens de clube não abrangem todos os bens públicos impuros, onde os benefícios são parcialmente excludentes e/ou parcialmente rivais. A questão fundamental é que existe um mecanismo de exclusão suficientemente barato para cobrar dos usuários pelo congestionamento que seu uso provoca. O autor imaginou, então, clubes em que seus membros possuem um arranjo institucional próprio para a provisão de bens que estão sujeitos a alguma rivalidade na forma de efeito *crowding* ou congestionamento, que simplesmente representa uma detração na quantidade ou qualidade dos bens em decorrência do aumento de sua utilização pelos participantes desse clube.

A Teoria dos clubes ocupa um lugar de destaque na literatura sobre escolha pública para um número de razões. Pode servir como base teórica para um projeto jurisdicional, enfatizando que os bens públicos podem, em circunstâncias chaves, ser providos de modo que não implique necessariamente o fornecimento pelo governo. A teoria indica que o tamanho dos clubes ou o número de participantes é uma escolha exógena que não é independente da decisão de associar-se. E por último, o estudo dos clubes pode ser aplicado a uma incrível variedade de situações que incluem impactos dos clubes em praticamente todos os campos da economia.

Sem embargos, o refinamento dos estudos sobre o tema tem centrado suas análises na otimização dos clubes, nas formas institucionais alternativas para a formação, na composição

de seus membros, no tipo de mecanismo de exclusão, e financiamento dos clubes, e sobretudo na função de congestionamento. Contribuições como as de Craig (1987), Edwards (1990), McGreer e MacMillan (1993), Means e Mehay (1995), e Reiter e Weichenrieder (1997) fortaleceram cada vez mais as análises, estabelecendo-as a partir da extensão do debate associada a outros tópicos importantes da economia pública. Exemplos incluem a análise do imposto sobre a propriedade como uma taxa de congestionamento (Wilson, 1997); a qualidade do serviço público quando a população dos clubes é heterogênea (Glazer a Niskanen, 1997); e a competição fiscal entre clubes na presença de congestionamento (Matsumoto, 2000).

Tais fatos constatam que, de fato, Buchanan (1965) habilmente lançou as bases para a teoria em uma estrutura simples, mas eficiente, partindo de um problema no qual os agentes pertencentes ao clube deveriam maximizar uma função que dependia do consumo do bem privado e coletivo, propondo formas de análise que passavam, sobretudo, por determinar o tamanho ideal do clube, dado a quantidade de um bem público.

Com efeito, esta pesquisa baseia-se no entendimento que as condições de fornecimento público dependem do tamanho do clube correspondente e que os mecanismos de provisão são processos políticos em que a demanda e a oferta são consideradas. Desta feita, fundamenta suas análises em um modelo teórico que incorpora tais predições.

2.3 O EFEITO DO TAMANHO DOS CLUBES SOBRE A QUALIDADE DA PROVISÃO PÚBLICA

Baseados na Teoria dos Clubes, o modelo teórico aqui apresentado, foi proposto no intuito de analisar o impacto do crescimento das cidades sobre a qualidade dos serviços públicos locais, com base na publicidade dos bens. Partem da suposição da existência de um clube qualquer de tamanho N que provém um bem público Z com qualidade q aos seus membros. Expõem que se, para um dado nível de Z, se a variação do tamanho do clube for positiva ($\Delta^{+}N$) ela certamente induzirá também uma variação na qualidade dos serviços. Se $q > 0$, evidencia efeitos de economia de escala e boa prestação pública, do contrário, se $q < 0$, significa que o efeito *crowding* (congestionamento) impacta a provisão do serviço. Formalmente, para um dado clube, o efeito do tamanho pode ser formalmente expresso como:

$$q = g(Z, N) \tag{2.1}$$

onde q reflete a qualidade dos bens públicos consumidos pelos membros de uma comunidade, Z é a composição dos bens os quais representam a quantidade total de serviços providos por um clube, e N o número de usuários simultâneos desse bens. Assume-se log-linearidade em q, para posteriores análises econométricas, e para atender a suposição de proporcionalidade entre q e Z[11] usada por Borcheding e Deacon (1972) e Bergstron e Goodman (1973). Então:

$$q = f(N)Z \tag{2.2}$$

A diferenciação total da equação (2.2), $dq = Zf'(N)dN + f(N)dZ$, dividida por q ou $f(N)Z$ dá:

$$\frac{dq}{q} = \frac{f'(N)N}{f(N)}\frac{dN}{N} + \frac{dZ}{Z} \tag{2.3}$$

Definindo a elasticidade do tamanho do clube (ηN) como uma medida da variação da qualidade de serviços públicos no caso da entrada de usuários adicionais ao clube, mantendo simultaneamente a quantidade total dos mesmos serviços, tem-se $\eta N = \partial q/\partial N . N/q$. Se $\eta N = 0,$ o crescimento do clube não provoca efeitos sobre a qualidade dos serviços; caso $\eta N > 0$, fica caracterizado o efeito escala, e se $\eta N < 0$, denota efeito congestionamento sobre a provisão. Como $\partial q/\partial N = f'(N)Z$ e $q = f(N)Z$, a elasticidade do tamanho do clube pode ser escrita como:

$$\eta N = \frac{f'(N)N}{f(N)} \tag{2.4}$$

Combinando a equação (2.3) com a (2.4), tem-se:

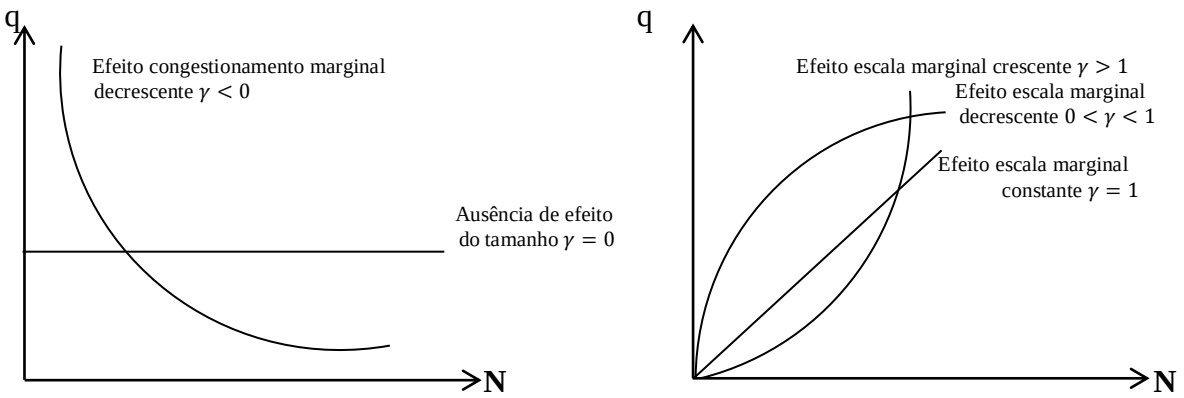
$$\frac{dq}{q} = \eta N\frac{dN}{N} + \frac{dZ}{Z} \tag{2.5}$$

ou seja, a variação relativa na qualidade dos serviços públicos depende não só do nível da provisão (Z), mas também da variação relativa dos membros do clube e da elasticidade do tamanho do clube.

[11] Para detalhes, ver Reiter e Weichenrieder (1997).

Como a especificação da elasticidade do tamanho do clube está condicionada a forma funcional f, as funções de congestionamento podem ser ampliadas para permitirem uma visão mais consistente do funcionamento do mercado, em diferentes fases do desenvolvimento de um clube, para sustentação ou não da publicidade dos bens. As três formas para captação do efeito *crowding out* e publicidade do bem usualmente consideradas na literatura são expostas a seguir. A primeira é a de Borcheding-Deacon (1972), a qual pode descrever o congestionamento marginal ou alternativamente o efeito escala marginal decrescente, constante, ou crescente, a partir da intensidade do uso do bem, conforme a Figura 2.1.

Figura 2.1: Função de Borcheding-Deacon (B-D)



Fonte: Guengant, Josselin e Rocaboy (2002).

Nesse caso, a função $f(N) = N^{\gamma}$ pode ser expressa através da forma:

$$q = N^{\gamma} Z \tag{2.6}$$

onde a elasticidade do tamanho do clube é dada por uma constante expressa por:

$$\eta N = \frac{\partial q}{\partial N}\frac{N}{q} = \gamma N^{\gamma-1} Z \frac{N}{N^{\gamma} Z} = \gamma \tag{2.7}$$

Sendo $q = N^{\gamma} Z$; $\frac{\partial q}{\partial N} = \frac{\gamma}{N} q$; $\frac{\partial^2 q}{\partial N^2} = \frac{\gamma(\gamma-1)}{N^2} q$, a natureza do efeito do tamanho do clube depende do valor do parâmetro γ.

Outra forma funcional, $f(N) = N^{\gamma'} e^{\gamma'' \boldsymbol{N}}$, foi fornecida por Buchanan (1965), e sugere que pode haver um intervalo onde o aumento do número de usuários que compartilham bens no clube, aumenta, até certo ponto, os benefícios de seus membros do clube e gera uma

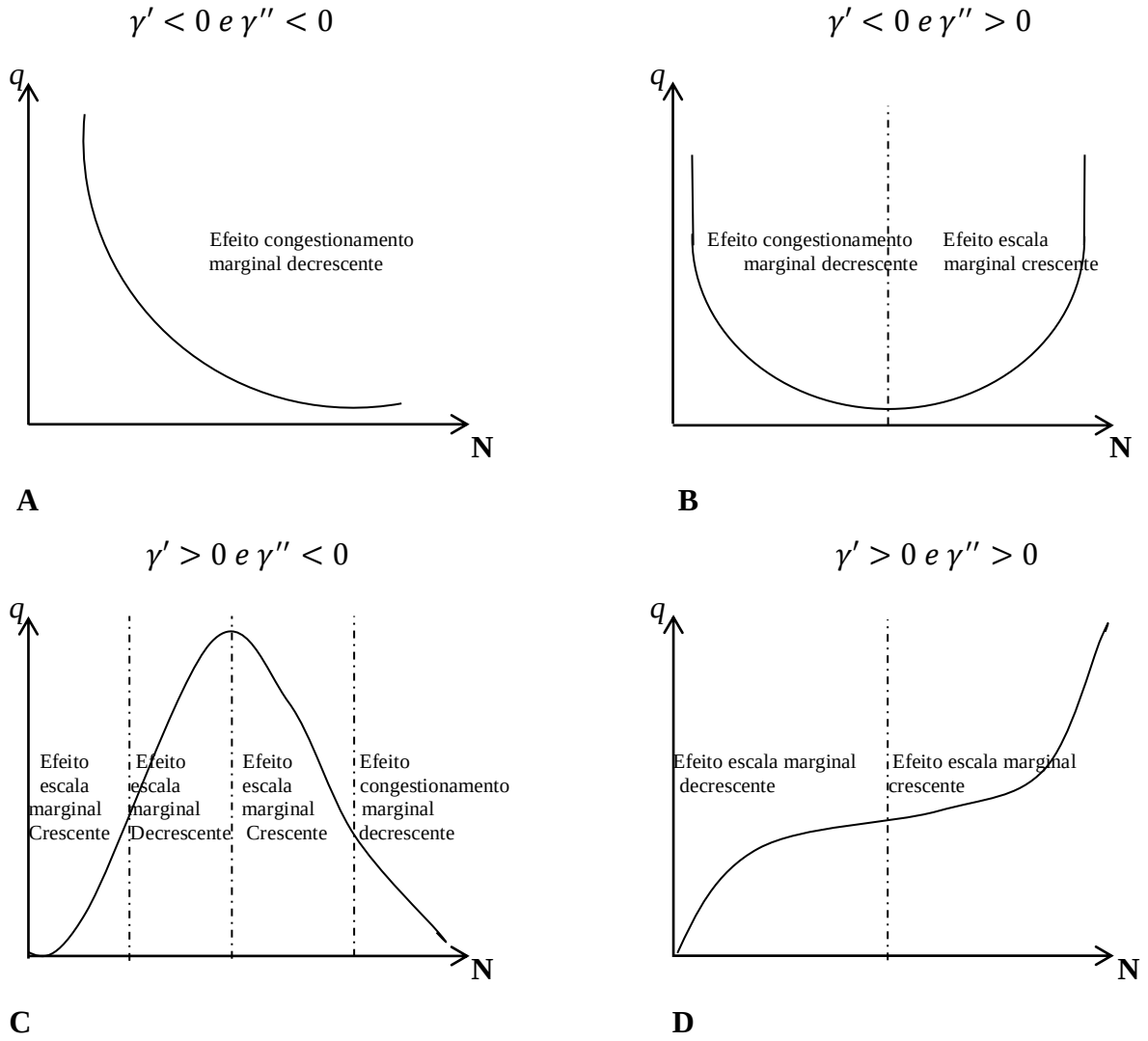
camaradagem entre eles. Este é o chamado efeito "*camaraderie*". Contudo, após este ponto, o processo se reverte e os benefícios gerados são corroídos pelo congestionamento, como exposto na Figura 2.2. A função "*camaraderie*" pode ser escrita como:

$$q = N^{\gamma'} e^{\gamma'' \boldsymbol{N}} Z \tag{2.8}$$

onde a elasticidade do tamanho do clube é:

$$\eta N = \frac{\partial q}{\partial N}\frac{N}{q} = [(\gamma' N^{\gamma'-1} e^{\gamma''} N) + (\, e^{\gamma'' N} \gamma'' N^{\gamma'})] z \frac{N}{N^{\gamma'} e \gamma''^{N} Z} = \gamma' + \gamma'' N \tag{2.9}$$

Figura 2.2: Função "*Camaraderie*"



Fonte: Guengant, Josselin e Rocaboy (2002).

Sendo $q = N^{\gamma'} e^{\gamma'' N} Z$; $\frac{\partial q}{\partial N} = \frac{\gamma' + \gamma'' N}{N} q$; $\frac{\partial^2 q}{\partial N^2} = [\left(\frac{\gamma'}{N} + \gamma''\right)^2 - \frac{\gamma'}{N^2}\Big] q$; e $\frac{\partial^2 q}{\partial N^2} = 0$ para $N = -\frac{\gamma' \pm \sqrt{\gamma'}}{\gamma'}$, com $\gamma' > 0$. Então, a natureza do efeito do tamanho do clube depende do valor dos parâmetros γ' e γ''.

Destaca-se que a formulação de Borcheding-Deacon também descreve o comportamento "*camaraderie*", se $\gamma > 0$. (Ver Fig.2.1)

E por fim, a forma flexível, desenvolvida por Edwards (1990), o qual não impõe nenhum perfil ao tamanho do clube para análise do efeito congestionamento:

$$q = e^{(\gamma' N + \gamma'' N^2 + \gamma''' N^3)} Z \tag{2.10}$$

Com a elasticidade do tamanho do clube denotado por:

$$\eta N = \frac{\partial q}{\partial N} \frac{N}{q} = \gamma' N + 2\gamma'' N^2 + 3\gamma''' N^3 \tag{2.11}$$

Então, o efeito marginal pode ser positivo ou negativo, de acordo com os valores dos parâmetros da população.

Apresentadas as possíveis formas funcionais do efeito do tamanho do clube sobre a qualidade dos serviços públicos, parte-se para a formalização dos modelos de decisão dos gastos, os quais representam o contexto de determinação da provisão no Brasil em função da demanda e em função da oferta.

2.4 PROVISÃO DE BENS PÚBLICOS NO CONTEXTO DOS CLUBES: MODELOS DE DEMANDA E DE OFERTA

2.4.1 O modelo da demanda

Este modelo baseia-se na teoria do eleitor mediano, definida no âmbito da democracia representativa, onde os governos locais, representadas por i = 1,...., I, assumem o compromisso de atender as preferências do eleitor mediano m. Assume-se que o consumo individual deste eleitor define-se como um numerário x. As jurisdições provêm um bem público local de qualidade q. Assim, a função utilidade do eleitor mediano é $u_m(x_m, q)$, onde sua renda é representada por y_m e a base fiscal por b_m. Para uma alíquota τ, a restrição

orçamentária individual do eleitor mediano pode ser escrita por $x_m + \tau b_m = y_m$. A tecnologia de produção do bem é representada por uma função custo $C(Z)$, e a base fiscal total é *B*. As receitas de transferências *lump-sum* oriundas de outros níveis de governo são denotadas por *S*. (Em termos *per capita*, tem-se $b = B/N$, e $s = G/N$). Dessa forma, a restrição orçamentária total da jurisdição é $C(Z) = \tau B + S$, o que supõe a necessidade de equilíbrio do orçamento por parte do governo local. E por fim, mais uma restrição é válida, $q = f(N)Z$, que expressa os mecanismos de efeitos do tamanho dos clubes.

Assim, o problema de maximização da utilidade do eleitor mediano, combinada as três restrições anteriores é:

$$\underset{x_m, q}{\text{Max}}\, u_m(x_m, q) \tag{2.12}$$

$$Sujeito\ a \qquad x_m + \frac{C(f(N)^{-1}q)}{N}\frac{b_m}{b} = y_m + s\frac{b_m}{b}$$

A resolução do problema descreve o preço do serviço público (*tax-price*) para o eleitor mediano:

$$\frac{\frac{\partial u_m}{\partial q}}{\frac{\partial u_m}{\partial x_m}} = C_Z' \frac{f(N)^{-1}}{N}\frac{b_m}{b} = p_m \tag{2.13}$$

onde C_Z' denota o custo marginal de produção do bem público local. O primeiro termo da equação descreve as condições de produção C_Z', combinadas com as características do consumo coletivo $\frac{f(N)^{-1}}{N}$. O segundo termo representa a influência direta da taxação sobre as escolhas do eleitor mediano.

No tocante à estimação do modelo descrito, supõe-se uma função demanda definida por $q = f(p_m, y_m, s)$ caracterizada por elasticidades do conjunto de variáveis que influenciam a demanda sob a hipótese de log-linearidade. A assume-se a função correspondente a seguinte forma:

$$q = c_1(p_m)^{\varepsilon}(y_m)^{\beta_1}(s)^{\beta_2} \tag{2.14}$$

onde os coeficientes ε, β_1 e β_2, respectivamente, representam a elasticidade-preço da demanda (*tax-price)*, a elasticidade-renda, e a elasticidade das transferências. c_1 é uma constante.

Como q é uma variável não observável, usa-se a uma função estimável de gastos governamentais como *proxy*. Adicionalmente, supõe-se que os custos marginais e médios de produção são iguais e constantes dentro da comunidade $(C(Z)/Z = C'_Z)$. Então, como $q = f(N)Z$:

$$q = \frac{f(N)C(Z)}{C'_Z} = c_1(p_m)^{\varepsilon}(y_m)^{\beta_1}\left(s\frac{b_m}{b}\right)^{\beta_2} \tag{2.15}$$

Substituindo (13) e (15), e considerando a função em termos *per capita* tem-se que:

$$\frac{C(Z)}{N} = e = c_1(C'_Z)^{(1+\varepsilon)}[Nf(N)]^{-(1+\varepsilon)}\left(\frac{b_m}{b}\right)^{\varepsilon+\beta_2}(y_m)^{\beta_1}(s)^{\beta_2} \tag{2.16}$$

Esta equação representa a função estimável de despesa pública local *per capita*, e descreve um quadro quando o objetivo é demonstrar um processo de democracia direta, onde o comportamento dos representantes eleitos está "vinculado" a características da demanda.

2.4.2 O modelo da oferta

Este modelo repousa sobre o lado da oferta das despesas públicas locais, e a ideia que o "auto-interesse" do político eleito pode afastá-lo das preferências do eleitor mediano, e assim, seu comportamento é livre para escolher as políticas que pretende empreender. O representante político não teria que levar em conta as características relevantes do eleitor mediano; no entanto, suas políticas devem ser relevantes no caso do governo local apresentar aspirações, como uma reeleição, por exemplo. Por outro lado, o eleitor gostaria de assegurar que essas ações levarão ao nível desejado de serviços públicos, com o menor custo possível.

Em termos de formalização do modelo, (y_m) e (p_m) não influenciam a equação da oferta, e a parcela fiscal agora é (1/*b*) já que o representante não considera a distribuição das bases tributárias individuais, mas apenas o seu nível médio, caracterizando uma medida de esforço fiscal uniforme.

O político eleito assume que deve maximizar uma função social $W_s(q,\tau)$ com $\partial W_s/\partial q > 0$ e $\partial W_s/\partial \tau < 0$. Como o eleitor mediano não é relevante, a sua restrição orçamentária não é considerada. Neste modelo, considera-se a função custo diferenciável, $C(Z)$ que representa a tecnologia de produção do bem, e em termos de notação para Z, a função de congestionamento é $q = f(N)Z$. Dessa forma o problema de maximização do representante eleito pode ser escrito por:

$$\underset{q,\tau}{\text{Max}}\, W(q,\tau) \tag{2.17}$$

$$Sujeito\ a \qquad C(Z) = \tau B + S, \quad q = f(N)Z$$

onde as restrições podem ser combinadas, supondo-se que o custo marginal é igual ao custo médio de produção:

$$\frac{f(N)^{-1}C_Z'}{N}q - b\tau = s \tag{2.18}$$

Então, o *tax-price* implícito é:

$$\frac{\partial W/\partial q}{-\partial W/\partial \tau} = C_Z'\frac{f(N)^{-1}}{N}\frac{1}{b} = p_s \tag{2.19}$$

Considerando a função oferta definida como $q = f(p_s, s)$, tem-se:

$$q = c_2(p_s)^{\varepsilon}(s)^{\beta_2} \tag{2.20}$$

em que ε é a elasticidade-preço da oferta e β_2 a elasticidade das transferências. Usando (2.19) e rearrumando a equação (2.20) em termos *per capita*, tem-se:

$$\frac{C(Z)}{N} = e = c_2(C_Z')^{(1+\varepsilon)}[Nf(N)]^{-(1+\varepsilon)}\left(\frac{1}{b}\right)^{\varepsilon}(s)^{\beta_2} \tag{2.21}$$

Esta é a função estimável de despesa pública local *per capita*, do lado da oferta.

Diante do exposto, as equações (2.16) e (2.21) fornecem a justificativa teórica para as especificações econométricas, que estão expostas adiante, dos dois modelos concorrentes da provisão pública local.

2.5 DADOS E VARIÁVEIS

As estimações centram-se em variáveis relacionadas a 5268 municípios brasileiros. As principais fontes de informações necessárias para o desenvolvimento deste estudo são referentes ao Censo 2010. Ressalta-se que a variável dependente dos modelos considera o valor das despesas correntes municipais em termos *per capita*. Além das variáveis explicativas básicas das hipóteses alternativas, foram introduzidas na análise parâmetros de controle e *dummies* específicas para cada modelo, para levar em conta elementos locacionais e características dos municípios.

No modelo da demanda, foi incorporada uma matriz Ω de características socioeconômicas comumente utilizadas na literatura sobre o eleitor mediano, tais como Percentual de Pessoas com menos de 18 anos; Percentual de Pessoas com mais de 65 anos; Esperança de Vida e Índice de Desenvolvimento Municipal da Educação– IDHM/Educação, e *dummies* Regionais (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) e de Participação em Consórcio Municipal.

Já no modelo direcionado a oferta, a matriz Π representa a inserção de algumas variáveis representativas de elementos institucionais e políticos determinantes para a provisão governamental. O Número de Funcionários Públicos Municipais; O Salário Médio dos Funcionários Públicos da localidade; além das *dummies* Regionais, de Alinhamento Político com o Governo do seu Estado e Alinhamento Político com o Governo Federal.

É importante destacar que, apesar da suposição de que o custo marginal de produção é constante dentro das comunidades, as especificações requerem a sua estimativa, já que esta é uma variável própria de cada localidade quando o conjunto de municípios é considerado. Encontrar então uma *proxy* que se adeque corretamente a análise é uma tarefa extra do trabalho. Guengant, Josselin e Rocaboy (2002), expõem que uma saída seria a utilização de dados sobre a eficiência dos serviços das comunidades, contudo, a insuficiência de dados, e a não consideração da eficiência técnica no escopo da pesquisa delinearam outra solução para o problema.

Estudos internacionais, por exemplo, Turnbull e Chang (1998); Sanz e Velazquez (2002) e Sampaio de Sousa, Cribari Neto e Stosic (2005), e nacionais como Mendes e

Sampaio de Sousa (2006) entre outros, tratam fatores geográficos como relevantes para explicação das despesas, ao ponto em que o efeito dessas variáveis pode (ou não) limitar a exploração de economias de escala que caracterizam a produção de serviços públicos. Destarte, a variável *Área* é utilizada como forma de capturar a distribuição desigual de terras entre os municípios, identificando a variação de custos em decorrência da densidade de cada território. Assim, $(C_Z' = A_i^{\alpha})$, em que o parâmetro α reporta a elasticidade-área.

Na prática, o uso da área como parâmetro também ajuda a supor o que acontece com as pequenas comunidades quando sua população diminui e os custos fixos dos serviços não diminuem na mesma proporção, já que a relativa rigidez na prestação de alguns serviços públicos (como infra-estrutura, por exemplo) revela-se um encargo orçamental para muitos gestores.

Expostas todas as informações a respeito das variáveis utilizadas no trabalho, salienta-se que esses dados foram, em sua maioria, coletadas junto ao Instituto Brasileiro de Geografia e Estatística (IBGE). Com relação aos dados relacionados às finanças públicas, como os gastos públicos, as receitas e transferências, a fonte de informações é a Base de Finanças Municipais do Brasil, FINBRA, da Secretaria do Tesouro Nacional (STN), e sobre os partidos políticos as informações são provenientes do Tribunal Superior Eleitoral (TSE).

O Quadro 2.1, a seguir, apresenta a lista das variáveis utilizadas em cada modelo, bem como as suas descrições sintéticas.

Quadro 2.1: Resumo descritivo das variáveis utilizadas no estudo

Variáveis	Notação	Descrição
Despesas locais com Serviços Públicos	$\frac{C(Z)}{N} = e$	Despesas totais *per capita* com os serviços públicos locais.
Renda Mediana	y_{im}	Soma entre a renda mediana de cada município e o produto da multiplicação do *tax-price* pelo montante das transf. intergovernamentais recebidas.
***Tax-price* ou Base Fiscal do Eleitor Mediano**	b_m/b_i	Razão entre renda mediana e renda média. Mede a sensibilidade da demanda em relação ao preço dos bens públicos.
Parcela fiscal do Modelo da Oferta	$1/b_i$	Razão entre o valor do imposto sobre a propriedade e a receita local.[12]
Transferências do governo	s_i	Participação das rendas provenientes de transf. intergovernamentais na renda do município.
População	N_i	Número de cidadãos em cada município analisado.
Área	A_i	Área de cada município.
População menor que 18 anos	$Pop18_i$	Percentual de Pessoas com menos de 18 anos
População maior que 65 anos	$Pop65_i$	Percentual de Pessoas com mais de 65 anos
Esperança de vida	Esp_i	Indicador utilizado como *proxy* de demanda por serviços de saúde nos municípios.
Índice de Desenvolvimento Humano Municipal – IDH/Educação	IDH_i	Indicador utilizado como *proxy* de demanda por serviços em educação fundamental nos municípios.
Número de funcionários	Fun_i	Número de Funcionários Públicos Municipais.
Salário Público	Sal_i	O Salário Médio dos Funcionários Públicos da localidade
***Dummy* Regional (N, NE, CO, SE, S)**	$DUMRG_i$	Assume o valor 1 se o município for localizado na Região; 0, de outra forma.
***Dummy* Consórcio**	$DUMCN_i$	Assume o valor 1 se o município participa de coordenação intermunicipal de serviços públicos; 0, caso contrário.
***Dummy* de Alinhamento Político Estadual**	$DUMP1_i$	Assume o valor 1 se o prefeito do município for aliado político do Governador do Estado; 0, de outra forma.
***Dummy* de Alinhamento Político Federal**	$DUMP2_i$	Assume o valor 1 se o prefeito do município for aliado político do Presidente da República; 0, de outra forma.

Fonte: Elaboração própria.

A Tabela 2.1 mostra as estatísticas descritivas das variáveis analisadas nesse estudo.

[12] Seguindo a linha dos trabalhos de Guengant, Josselin e Rocaboy (2002), e Josselin, Rocaboy e Tavéra (2008).

Tabela 2.1: Estatísticas descritivas

Variável	Observações	Média	Desv. Padrão	Mínimo	Máximo
e (R$)	5268	1504,32	709,34	127,69	10424,06
$y_m(R\$)$	5268	621,92	169,74	200,74	1800,30
b_m/b_i	5268	0,7128	0,1372	0,2645	1,2365
$1/b_i$	5268	0,4742	0,1740	0,026	0,8894
$s_i(R\$)$	5268	1367,76	845,84	148,91	12445,17
N_i	5268	34511,84	204811,10	1020	1,13e+07
$A_i(Km^2)$	5268	1525,77	5609,71	3,56	159533,25
$IDH - Ed$	5268	0,66	0,07	0,42	0,86
$Pop18$ (%)	5268	25,11	4,97	7,91	58,50
$Pop65$ (%)	5268	12,2	3,26	2,61	29,39
Esp	5268	73,16	2,66	65,3	78,64
Fun	5268	806,67	4663,06	2	203464
$Sal(R\$)$	5268	1.307,97	654,67	300,00	27.000,00

Fonte: Elaboração própria a partir de dados da pesquisa.

Algumas considerações são importantes ao se observar a Tabela 2.1. Destaca-se a enorme discrepância entre o número mínimo e máximo de funcionários públicos municipais, bem como a diferença salarial proveniente dos governos municipais.

Feitas todas as considerações sobre a base de dados do trabalho e suas fontes, tem-se, a seguir, as especificações econométricas.

2.6 ESPECIFICAÇÕES ECONOMÉTRICAS

Considerando as formas $f(N)$ apropriadas, as regressões de cada modelo foram realizadas da seguinte forma:

Borcherding e Deacon (B-D):

Modelo do Eleitor Mediano

$$ln(e_i) = K_1 + (\varepsilon + 1)\alpha ln(A_i) + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2) ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma + 1) ln(N_i) + \beta_2 ln(s_i) + \beta_j ln(\Omega_i) + \sum_{l=2}^{n} \beta_l \, DUM_i + u_i \quad (2.22)$$

Modelo de Oferta

$$ln(e_i) = K_2 + (\varepsilon + 1)\alpha ln(A_i) + \varepsilon ln\left(\frac{1}{b_i}\right) - (\varepsilon + 1)(\gamma + 1)ln(N_i) + \beta_2 ln(s_i) + \beta_j ln(\Pi_i) + \sum_{l=2}^{n} \beta_l DUM_i + v_i \quad (2.23)$$

***Camaraderie*:**

Modelo do Eleitor Mediano

$$ln(e_i) = K_1 + (\varepsilon + 1)\alpha ln(A_i) + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2)ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)(\gamma' + 1)ln(N_i) - (\varepsilon + 1)\gamma''(N) + \beta_2 ln(s_i) + \beta_j ln(\Omega_i) + \sum_{l=2}^{n} \beta_l DUM_i + u_i \quad (2.24)$$

Modelo de Oferta

$$ln(e_i) = K_2 + (\varepsilon + 1)\alpha ln(A_i) + \varepsilon ln\left(\frac{1}{b_i}\right) - (\varepsilon + 1)(\gamma' + 1)ln(N_i) - (\varepsilon + 1)\gamma''(N) + \beta_2 ln(s_i) + \beta_j ln(\Pi_i) + \sum_{l=2}^{n} \beta_l DUM_i + v_i \quad (2.25)$$

Forma flexível:

Modelo do Eleitor Mediano

$$ln(e_i) = K_1 + (\varepsilon + 1)\alpha ln(A_i) + \beta_1 ln(y_{im}) + (\varepsilon + \beta_2)ln\left(\frac{b_{im}}{b_i}\right) - (\varepsilon + 1)ln(N_i) - (\varepsilon + 1)\gamma'(N) - (\varepsilon + 1)\gamma''(N^2) - (\varepsilon + 1)\gamma'''(N^3) + \beta_2 ln(s_i) + \beta_j ln(\Omega_i) + \sum_{l=2}^{n} \beta_l DUM_i + u_i \quad (2.26)$$

Modelo de Oferta

$$ln(e_i) = K_2 + (\varepsilon + 1)\alpha ln(A_i) - \varepsilon ln\left(\frac{1}{b_i}\right) - (\varepsilon + 1)ln(N_i) - (\varepsilon + 1)\gamma'(N) - (\varepsilon + 1)\gamma''(N^2) - (\varepsilon + 1)\gamma'''(N^3) + \beta_2 ln(s_i) + \beta_j ln(\Pi_i) + \sum_{l=2}^{n} \beta_l DUM_i + v_i \quad (2.27)$$

com $K_1 = (\varepsilon + 1)ln(c_1)$ e $K_2 = (\varepsilon + 1)\ln(c_2)$.

Cabe também ressaltar a possibilidade de se captar efeitos espaciais entre os dados analisados devido à existência de alguma relação entre as estruturas de eficiência municipal em diferentes pontos no espaço. Como ressaltam Mendes e Sampaio de Sousa (2006), quanto

menor a área onde os municípios estão localizados, maior a probabilidade de correlação geográfica. Dessa forma, pela peculiaridade dos dados municipais brasileiros, aos modelos empíricos utilizados nesse estudo foi incorporada uma matriz *W*, *n x n*, com a finalidade de controlar a existência de efeitos vizinhança. A construção dessa matriz apoiou-se nos critérios de proximidade, e foi definida considerando a contiguidade e/ou distância geográfica, segundo uma métrica estabelecida.

A fim de contornar uma possível arbitrariedade na escolha da matriz, realizou-se o procedimento proposto por Baumont (2004), que tem por escopo capturar o máximo de dependência espacial. O procedimento é feito em três passos: 1) Estima-se o modelo linear clássico de regressão linear; 2) Testa-se os resíduos desse, com base no *I de Moran* (estatística capaz de mensurar a autocorrelação espacial contida numa variável); e 3) Seleciona-se a matriz de pesos que tenha gerado o mais alto valor do teste de *I de Moran,* e que seja estatisticamente significativo.

A partir das estimações por MQO, foram realizados testes para autocorrelação espacial a partir do cálculo das estatísticas de *I de Moran* e LISA (*Local Indicator of Spatial Association*), que são indicadores que permitem estimar o quanto o valor observado de um atributo numa área é correlacionado com os valores desta mesma (técnica univariada) ou de outras variáveis (técnica multivariada), nas localizações vizinhas[13]. Dada à confirmação da autocorrelação espacial, testes diagnósticos com base no princípio do Multiplicador de Lagrange (*ML*), permitem captar como a dependência espacial toma forma[14], e sugerem qual especificação gera inferências confiáveis[15] a respeito dos determinantes da qualidade dos serviços públicos nos clubes, já que sob evidências de dependência espacial o estimador de MQO não é indicado.

Sob tais aspectos, os modelos de regressão estimados no trabalho - em sua formulação geral - podem ser expressos por:

$$Y = \rho WY + X\beta + u\,, \qquad \text{com } u = \lambda Wu = e \qquad (2.28)$$

[13] Fornecem um único valor como medida de associação espacial para todo o conjunto de dados. Valores iguais a 0 (zero) indicam que não há correlação espacial. Já valores maiores ou menores que 0 (zero), indicam autocorrelação espacial positiva ou negativa, respectivamente. (ANSELIN, 1995).

[14] Conforme Almeida (2012), o procedimento procura discriminar se a autocorrelação se manifesta na forma de defasagem espacial ou erro espacial.

[15] O procedimento é simples: Sendo significantes as inferências das estatísticas ML_{err} e ML_{lag}, faz-se uso de suas versões robustas. Se RML_{err} for mais significante que RML_{lag} estima-se o modelo SEM, inversamente, se RML_{err} for menos significante que RML_{lag}, estima-se o modelo SAR. (ANSELIN *et al*, 1996).

no qual ρ é o coeficiente que mede a autocorrelação espacial, ou seja, que capta os efeitos da interação entre um município com seus vizinhos; *W (nxn)* é a matriz de pesos espaciais, e o parâmetro λ capta a correlação espacial entre os erros, e e é um novo termo de erro.

Sem embargos, tais procedimentos geraram estimativas para descrições a respeito dos parâmetros das elasticidades dos tamanhos dos clubes por forma funcional. De um modo geral, mensurou-se a variação relativa de q, dado uma quantidade Z de bens públicos locais, devido uma variação relativa da população, tanto pela hipótese da demanda como determinante da provisão pública, quanto da oferta.

2.7 RESULTADOS

Inicialmente, salienta-se que a matriz utilizada na estimação de todos os modelos foi a de pesos espaciais binários, construída a partir da ideia de vizinhança baseada na contiguidade, na qual duas localidades são vizinhas caso partilhem de uma fronteira em comum. A ideia é que duas regiões contíguas possuem uma maior interação espacial. (ALMEIDA, 2012). De acordo com este conceito, é atribuído um valor unitário na matriz a dois municípios vizinhos; caso contrário, atribui-se um valor nulo. Formalmente:

$$W_{ij} = \begin{cases} 1 & se\ i\ e\ j\ são\ contíguos \\ 0 & se\ i\ e\ j\ não\ são\ contíguos \end{cases} \tag{2.29}$$

Note-se que a matriz escolhida (critério de Baumont (2004)) foi a do tipo “rainha” (*queen*), a qual considera pontos de fronteiras e vértices entre os municípios como contíguos.

Os modelos apresentados na seção anterior foram estimados por MQO considerando, inicialmente, os 5268 municípios com populações diversas, no qual foram realizados os testes sobre a existência da autocorrelação espacial. As Tabelas 2.2 e 2.3 apresentam, respectivamente, os resultados das estimações do modelo da demanda, por formulação funcional, e os resultados da análise da autocorrelação a partir dos resíduos da regressão. As Tabelas 2.4 e 2.5 reportam aos mesmos parâmetros, porém, referindo-se ao modelo da oferta.

Tabela 2.2: Estimações por MQO do Modelo da Demanda – Por forma funcional

	Formulação Borcheding-Deacon	**Formulação *Camaraderie***	**Formulação Flexível**
$ln\left({b_{im}}/{b_i}\right)$	- 0,0928* (0,0121)	- 0,0911* (0,0118)	- 0,0905* (0,0110)
$ln(y_i)$	0,5031* (0,0509)	0,5028* (0,0487)	0,5013* (0,0481)
$ln(s_i)$	0,5421* (0,0447)	0,5398* (0,0413)	0,5377* (0,0404)
$ln(N_i)$	- 0,1201* (0,0323)	- 0,1192* (0,0301)	- 1,0906* (0,1287)
$ln(A_i)$	- 0,0242** (0,0117)	- 0,0245**(0,0119)	- 0,0246* (0,0121)
$ln(Pop18_i)$	0,0201***(0,0103)	0,0198***(0,0101)	0,0197 (0,0101)
$ln(Pop65_i)$	- 0,0267** (0,0128)	- 0,0258** (0,0128)	- 0,0256** (0,0125)
$ln(Esp_i)$	- 0,2305** (0,1146)	- 0,2298** (0,1127)	- 0,2297** (0,1125)
$ln(IDH_i)$	0,1808**(0,0881)	0,1808**(0,0880)	0,1807**(0,0880)
N		1,76E-07*(3,81E-09)	3,22E-06*(2,19E-08)
N^2			- 4,65E-11*(1,18E-12)
N^3			5,71E-14*(2,24E-15)
$Região\ CO$	0,1722* (0,0361)	0,1704*(0,0398)	0,1695*(0,0409)
$Região\ N$	0,0103* (0,0032)	0,0101*(0,0028)	0,0098*(0,0030)
$Região\ NE$	0,2513* (0,0804)	0,2499*(0,0554)	0,2472*(0,0598)
$Região\ S$	0,1783* (0,0511)	0,1755*(0,0528)	0,1723*(0,0501)
$DUMCM_i$	- 0,4121*(0,1223)	- 0,4116*(0,1218)	- 0,4116*(0,1209)
Constante	3,1403*(0,0594)	3,0167*(0,0453)	2,9571*(0,0185)
R^2 Ajustado	0,7730	0,7793	0,7811
Critério Akaike info:	947,31	948,36	949,27
Critério de Shwarz	1002,04	1002,68	1003,21
Log *likelihood*	-688,74	-688,11	-687,55

Fonte:: Elaboração própria a partir dos resultados do trabalho. **Nota:** * significante ao nível de 1%; ** significante ao nível de 5%; *** significante ao nível de 10%. **Obs.:** Os valores em parênteses são os erros padrões.

Tabela 2.3: Diagnóstico para dependência espacial do Modelo da Demanda – Por forma funcional

	Formulação Borcheding-Deacon	**Formulação *Camaraderie***	**Formulação Flexível**
$I\ de\ Moran$	0,2156 (0,0001)	0,2175 (0,0001)	0,2184 (0,0001)
ML_{err}	77,41 (0,0000)	77,45 (0,0000)	77,92 (0,0000)
ML_{lag}	24,55 (0,0000)	24,58 (0,0000)	24,66 (0,0000)

RML_{err}	15,73 (0,0632)	15,86 (0,0708)	15,97 (0,0771)
RML_{lag}	44,13 (0,0000)	44,26 (0,0000)	44,57 (0,0000)
ML_{SARMA}	96,21 (0,0000)	96,38 (0,0000)	96,55 (0,0000)

Fonte: Elaboração própria com base nos dados da pesquisa.
Obs: Os valores entre parênteses representam a probabilidade da estatística.

Tabela 2.4: Estimações por MQO do Modelo da Oferta – Por forma funcional

	Formulação Borcheding-Deacon	**Formulação *Camaraderie***	**Formulação Flexível**
$ln\left({}^{1}/_{b_i}\right)$	0,1038* (0,0289)	0,1045* (0,0301)	0,1047* (0,0309)
$ln(s_i)$	0,4977* (0,0671)	0,5003* (0,0736)	0,5011* (0,0732)
$ln(N_i)$	- 0,3014 *(0,0743)	- 0,3045 *(0,0621)	- 1,1049 *(0,1788)
$ln(A_i)$	- 0,0462* (0,0144)	- 0,0485* (0,0156)	- 0,0485* (0,0157)
$ln(Fun_i)$	0,1633** (0,0796)	0,1646** (0,0801)	0,1653** (0,0796)
$ln(Sal_i)$	0,2108** (0,1003)	0,2113** (0,1009)	0,2114** (0,1009)
N		6,04E-07*(2,53E-08)	8,21E-06*(3,12E-07)
N^2			- 3,46E-11*(1,52E-12)
N^3			5,88E-14*(5,21E-16)
$Região\ CO$	0,1297* (0,0311)	0,1283*(0,0298)	0,1254*(0,0296)
$Região\ N$	- 0,0233* (0,0072)	- 0,0225*(0,0074)	- 0,0220*(0,0069)
$Região\ NE$	0,2133* (0,0314)	0,2128*(0,0301)	0,2111*(0,0298)
$Região\ S$	0,1411* (0,0308)	0,1402*(0,0302)	0,1399*(0,0304)
$DUMP1_i$	0,2321* (0,0422)	0,2335* (0,0436)	0,2338* (0,0439)
$DUMP2_i$	0,1291* (0,0261)	0,1302* (0,0274)	0,1303* (0,0276)
Constante	4,0362*(0,1367)	3,2731* (0,1649)	3,5420*(0,1832)
R^2 Ajustado	0,8112	0,8147	0,8196
Critério Akaike info:	1078,54	1086,12	1098,06
Critério de Shwarz	1267,11	1275,84	1287,22
Log *likelihood*	-507,98	-518,43	-523,36

Fonte:: Elaboração própria a partir dos resultados do trabalho. **Nota:** * significante ao nível de 1%; ** significante ao nível de 5%; *** significante ao nível de 10%. **Obs.:** Os valores em parênteses são os erros padrões.

Tabela 2.5: Diagnóstico para dependência espacial – Modelo da Oferta – Por forma funcional

	Formulação Borcheding-Deacon	**Formulação *Camaraderie***	**Formulação Flexível**
$I\ de\ Moran$	0,1831 (0,0001)	0,1888 (0,0001)	0,1915 (0,0001)
ML_{err}	51,34 (0,0000)	52,17 (0,0000)	52,81 (0,0000)
ML_{lag}	37,12 (0,0000)	37,79 (0,0000)	38,06 (0,0000)
RML_{err}	21,19 (0,0812)	22,35 (0,0586)	22,68 (0,0901)
RML_{lag}	24,88 (0,0000)	25,76 (0,0000)	25,99 (0,0000)
ML_{SARMA}	84,05 (0,0000)	85,31 (0,0000)	85,67 (0,0000)

Fonte: Elaboração própria com base nos dados da pesquisa.
Obs: Os valores entre parênteses representam a probabilidade da estatística.

Importa observar na Tabela 2.3 (para o modelo da demanda) e na Tabela 2.5 (para o modelo da oferta) que o índice *I de Moran* confirma a presença de autocorrelação espacial, o que mostra a relevância do efeito vizinhança na distribuição espacial de despesas públicas locais. O parâmetro de correlação positivo indica que os níveis de gastos tendem a se espalhar para localidades vizinhas (efeito *spillover).* Aplicando-se então a regra de decisão, detectou-se a presença de dependência no coeficiente relacionado à variável explicativa defasada espacialmente, o que consequentemente garantiu a estimação, por forma funcional e clubes, do modelo SAR pelo método de variáveis instrumentais[16].

Tal procedimento é correto, pois, dada a natureza da dependência espacial, a presença da defasagem espacial da variável dependente (WY) equivale à introdução de uma variável endógena num sistema de equações simultâneas (ANSELIN, 1992). Pensou-se num conjunto de instrumentos[17] para a defasagem espacial da variável dependente que apresente as propriedades significativas da relevância e da exogeneidade[18].

As estimações então, para a demanda e para a oferta, foram realizadas a partir de uma divisão da amostra em 6 clubes: O Clube 1 (C1) contém cidades com população de até 5.000 habitantes; o Clube 2 (C2), cidades com população entre 5.001 à 10.000 habitantes; o Clube 3 (C3) envolve municípios de porte médio, com população entre 10.001 à 50.000; o Clube 4 (C4) concentra localidades com população entre 50.001 à 100.000 habitantes, o Clube 5 (C5)

[16] A não normalidade dos resíduos sugere a estimação por VI.
[17] Os instrumentos utilizados foram as variáveis explicativas defasadas espacialmente.
[18] Conforme Almeida (2012), o cumprimento da propriedade da relevância requer que esses instrumentos precisem estar correlacionados com a defasagem espacial da variável dependente a fim de que apresentem capacidade preditiva. Já a propriedade da exogeneidade exige que tal conjunto de instrumentos não esteja correlacionado com o termo de erro.

concentra localidades com população acima de 100.001 habitantes; por fim o Clube Brasil contém os 5268 municípios utilizados na pesquisa, com diversas populações.

As Tabelas 2.6 e 2.7 apresentam, respectivamente, as estimações do modelo da demanda e da oferta, por tamanho de clube. Como a maioria dos valores dos coeficientes é semelhante entre as especificações, em sua magnitude e significâncias dos parâmetros, a título de simplificação a tabela e os comentários serão reportados considerando apenas a forma flexível. Não obstante, as Tabelas 2.8 e 2.9 farão referências aos parâmetros e às elasticidades do tamanho dos clubes considerando as três formas funcionais.

Os resultados estimados para o modelo da demanda mostram que as elasticidades-preço apresentam comportamento esperado sendo negativas e significantes, em todos os casos, corroborando a importância dos preços na determinação da procura pelos serviços públicos no Brasil. A magnitude dos coeficientes sugere que o aumento da parcela tributária dos bens e serviços públicos provoca uma queda na demanda, contudo, o valor estimado abaixo dos encontrados na literatura empírica indica que o consumo é inelástico, caracterizando os bens como básicos ou essenciais para a população.

Entre clubes, algumas particularidades merecem destaque: em C1, o parâmetro demonstra menor sensibilidade das pequenas localidades ao custo de provisão do serviço; isso permite inferir que os menores municípios apresentam carência ou maior necessidade de provisão governamental. Essa pouca sensibilidade ainda persiste nos clubes de médio porte, C2 e C3; entretanto, em C4 e principalmente em C5, percebe-se um maior impacto do *tax-price* sobre a demanda por bens e serviços, o que pode ser explicado pelo fato que grandes clubes a base tributária é mais diversificada, além de refletir o fato de que as cidades maiores são mais integradas na economia de mercado.

As elasticidades-renda no modelo do eleitor mediano são positivas e inferiores a unidade, dentro da média dos estudos internacionais confirmando, entre clubes no Brasil, que os gastos públicos locais baseiam-se em padrões de bens normais, e que o maior desenvolvimento das localidades tem efeito crescente pelos serviços providos.

Tabela 2.6 - Estimações do Modelo da Demanda – Método SAR/IV – Formulação Flexível.

	Clube 1	**Clube 2**	**Clube 3**	**Clube 4**	**Clube 5**	**Brasil**
$ln\left({b_{im}}/{b_i}\right)$	- 0,0417* (0,0101)	- 0,0522* (0,0160)	- 0,0624* (0,0189)	- 0,0858* (0,0297)	- 0,1135* (0,0323)	- 0,0932* (0,0253)
$ln(y_i)$	0,4501* (0,1025)	0,4581* (0,0932)	0,4727* (0,0809)	0,5029* (0,0921)	0,5482* (0,1103)	0,5095* (0,1028)

$ln(s_i)$	0,6044* (0,1483)	0,6000* (0,1432)	0,5875* (0,1280)	0,5615* (0,1529)	0,5118* (0,0448)	0,5457* (0,1152)
$ln(N_i)$	- 0,1465* (0,0277)	- 0,1420* (0,0223)	- 0,1313* (0,0297)	- 0,1125* (0,0316)	- 0,1031* (0,0369)	- 0,1231* (0,0407)
$ln(A_i)$	0,0136*** (0,0067)	0,0123** (0,0050)	0,0086*** (0,0041)	- 0,0060* (0,0015)	- 0,0415* (0,0132)	- 0,0308** (0,0145)
$ln(Pop18_i)$	0,0250** (0,0122)	0,0123*** (0,0062)	0,0101 (0,0094)	- 0,0322** (0,0148)	- 0,0561** (0,0267)	0,0262*** (0,0132)
$ln(Pop65_i)$	0,0534** (0,0230)	0,0501*** (0,0253)	- 0,0065** (0,0027)	- 0,0364*** (0,0182)	- 0,0589* (0,0201)	- 0,0320*** (0,0163)
$ln(Esp_i)$	- 0,3154* (0,1045)	- 0,3084* (0,1129)	- 0,2550** (0,1237)	- 0,2271** (0,1086)	- 0,1317*** (0,0664)	- 0,2350*** (0,1179)
$ln(IDH_i)$	0,1914** (0,0908)	0,1995*** (0,1013)	0,1604 (0,0836)	0,1483*** (0,0754)	0,1303*** (0,0665)	0,1514** (0,0738)
N	6,39E-05* (1,32E-06)	3,49E-05* (1,01E-06)	8,54E-06* (3,11E-07)	3,15E-06* (2,17E-08)	8,23E-07* (1,63E-08)	4,28E-08* (5,21E-10)
N^2	- 3,19E-10* (1,02E-11)	- 4,21E-10* (1,23E-11)	- 5,12E-10* (2,27E-11)	-1,22E-12* (4,18E-14)	-4,22E-13* (2,38E-15)	- 7,18E-12* (1,09E-13)
N^3	1,51E-15* (1,02E-17)	8,51E-15* (5,23E-16)	2,24E-14* (1,27E-15)	5,16E-17* (1,18E-18)	7,69E-18* (2,34E-19)	5,88E-14* (1,09E-15)
$Região\ CO$	0,1631** (0,0760)	0,1693** (0,0667)	0,1505** (0,0649)	0,1573* (0,0489)	0,1877** (0,0878)	0,1421* (0,0442)
$Região\ N$	0,0427*** (0,0213)	0,0453*** (0,0228)	0,0577** (0,0286)	- 0,1184** (0,0545)	- 0,1573* (0,0437)	0,0184** (0,0088)
$Região\ NE$	0,1834* (0,0502)	0,1841** (0,0717)	0,1935* (0,0592)	0,2065* (0,0636)	0,2803* (0,0923)	0,2311** (0,0808)
$Região\ S$	0,1502* (0,0464)	0,1527* (0,0473)	0,1739* (0,0502)	0,1878* (0,0469)	0,2026** (0,0980)	0,1624** (0,0641)
$DUMCM_i$	- 0,4827* (0,1165)	- 0,4801* (0,1515)	- 0,4538* (0,1103)	- 0,4061* (0,1421)	- 0,3785* (0,1110)	- 0,4154* (0,1236)
$Constante$	3,9538* (0,4187)	4,0225* (0,3471)	3,6421* (0,6728)	4,8742* (0,2984)	3,0984* (0,1821)	5,2109* (0,1798)
$\boldsymbol{\rho}$	0,4784* (0,0821)	0,4801* (0,0911)	0,3024* (0,1023)	0,2809* (0,0723)	0,2007* (0,8977)	0,3988* (0,1024)
$Pseudo\ R^2$	75,94	74,78	78,03	73,89	74,23	78,96
Nº de Obs.	1209	1141	2324	318	276	5268

Fonte: Elaborado pelo autor a partir de resultados do trabalho.
Nota: * significante ao nível de 1%, ** significante ao nível de 5%,*** significante ao nível de 10%.
Obs.: Os valores em parênteses são os erros padrões.

No que concerne às transferências, os valores e sinais das elasticidades comprovam sua relevância para a determinação da demanda de bens públicos. A análise dos coeficientes entre clubes (C1 e C5) sugere que as municipalidades de menor porte são mais dependentes das transferências que suas congêneres maiores, cuja base tributária é mais ampla e mais diversificada. Os coeficientes positivos e significantes atrelados à variável área com sinal negativo indicam que há existência de economias de escala na provisão de bens públicos locais. O crescimento populacional e, possivelmente, das atividades econômicas provocam flutuações na demanda pelos serviços públicos locais.

O próximo passo das interpretações, ainda considerando a demanda como determinante dos gastos, foca na especificação da *proxy* para o custo marginal utilizada no trabalho. O critério da área (densidade) é significante, e tem impacto positivo sobre as despesas públicas em pequenas localidades (C1, C2, C3) e negativos para os demais clubes. Os resultados dos parâmetros permitem sugerir que pequenas localidades não conseguem explorar as economias de escala na produção dos serviços, e assim trabalham com custos elevados, reduzindo a eficiência no uso dos recursos. Nos municípios maiores, a maior concentração populacional tende a diminuir o custo marginal do serviço público e, portanto, contribui para a redução do gasto *per capita* municipal (SAMPAIO DE SOUSA, CRIBARI NETO E STOSIC, 2005).

Por fim, com relação a variável população, a relação mostra-se inversa para todos os padrões de municípios, e apresenta efeito crescente, indicando que quanto menores forem as cidades, maior é a despesa pública decorrente do aumento populacional, o que pode estar atrelado à limitação e problemas de ganhos de escala na provisão dos bens locais com o aumento o número de consumidores, particularmente em menores clubes. Uma observação importante a se destacar, é que o desenvolvimento dos clubes, em termos populacionais, possibilita num determinado ponto menor grau de congestão, propiciando benefícios de economias de escala derivadas da existência dos bens públicos locais.

A parcela da população com até 18 anos de idade apresenta efeito positivo sobre as localidades relativas à C1 e C2 e para o Brasil como um todo, contudo, negativo para os demais clubes, Como esta *proxy* representa a demanda potencial por educação fundamental e média nos municípios, infere-se que este resultado está associado ao fato de que em grandes clubes/localidades, o maior poder aquisitivo eleva a procura pela rede privada, reduzindo a demanda pela educação pública.

Inferência análoga acontece quando se observa os parâmetros relativos à população com mais de 65 anos de idade. Esta classe é considerada relativamente mais “próspera” nos

grandes clubes, podendo consumir substitutos privados dos bens públicos. O coeficiente positivo e significante para os pequenos clubes sugere que essa faixa etária nessas localidades são menos prósperas e assim demandam maior nível de serviços governamentais, sobretudo naqueles referentes à saúde pública.

Com relação à variável esperança de vida ao nascer, os coeficientes negativos e significantes mostram que essa variável tem efeito redutor dos gastos em todos os clubes; este resultado provavelmente reflete melhores condições de serviços públicos sanitários, saúde e saneamento, principalmente nos menores municípios. Isto implica que políticas que elevem a expectativa de vida local reduzem a demanda por gastos públicos nesses clubes.

Já o impacto do IDHM é positivo e significante para quase todos os clubes (exceção de C3), e crescente entre os clubes, revelando que o efeito dessa variável é maior quanto menores forem as cidades, possivelmente em decorrência da maior carência dessas localidades por serviços de educação.

Enquanto isso, o fato de uma jurisdição localizar-se fora da região Nordeste contribui para elevar os gastos públicos locais. A maior demanda por serviços do governo decorre do fato de que em muitas cidades nordestinas, os serviços públicos são os únicos disponíveis para a população pobre (MENDES E SAMPAIO DE SOUSA, 2006). Note-se por fim que, com exceção da região Norte, não somente o gasto adicional é mais elevado na maioria das regiões quando comparadas com a região Sudeste, mas ele cresce com o tamanho das municipalidades. Isto se explica porque o aumento da demanda pelos serviços analisados está relacionado às características fortemente urbanas existentes nos grandes centros concentrados nessa região.

Já o coeficiente associado à consorcialização municipal (DUMCN) - negativo, significante e decrescente com o tamanho do clube - indica que a cooperação pode ser uma solução para o problema da "fragmentação municipal" para os pequenos municípios. Isto porque os benefícios da consorciação, ao permitir explorar as economias de escala substanciais na provisão dos bens públicos locais, permite reduzir os gastos e elevar o grau de publicidade desses serviços.

Com relação ao modelo da oferta (Tabela 2.7), os resultados aqui descritos corroboram as predições teóricas. As estimativas do modelo apresentam diferenças com relação à análise anterior: primeiro que os coeficientes da elasticidade preço são positivos, o que vai ao encontro das predições básicas da lei geral da oferta.

Tabela 2.7- Estimações do Modelo da Oferta – Método SAR/IV – Formulação Flexível.

	Clube 1	Clube 2	Clube 3	Clube 4	Clube 5	Brasil
$ln\left({}^{1}/_{b_i}\right)$	0,0501* (0,0138)	0,0524* (0,0132)	0,0995* (0,0268)	0,1133* (0,0249)	0,1515* (0,0418)	0,0940* (0,0108)
$ln(s_i)$	0,5471* (0,1610)	0,5355* (0,1528)	0,4911* (0,1375)	0,4701* (0,1233)	0,4558* (0,1070)	0,5027* (0,1420)
$ln(N_i)$	- 1,0502* (0,1280)	- 1,0524* (0,1822)	- 1,0994* (0,1798)	- 1,1134* (0,1918)	- 1,1516* (0,1865)	- 1,0941* (0,1958)
$ln(A_i)$	- 0,0194*** (0,0098)	- 0,0179** (0,0060)	- 0,0124*** (0,0059)	- 0,0101* (0,0036)	- 0,0087* (0,0023)	- 0,0124** (0,0048)
$ln(Fun_i)$	0,0503** (0,0237)	0,0594*** (0,0301)	0,0801* (0,0246)	0,1048** (0,0429)	0,1318** (0,0527)	0,0879** (0,0357)
$ln(Sal_i)$	0,0691** (0,0273)	0,0725*** (0,0367)	0,1036** (0,0478)	0,1385*** (0,0703)	0,1857* (0,0411)	0,1170*** (0,0586)
N	2,79E-05* (1,12E-06)	6,92E-05* (2,19E-06)	4,78E-07* (1,93E-08)	1,78E-06* (1,29E-07)	3,48E-07* (1,25E-08)	3,74E-05* (1,17E-06)
N^2	- 8,19E-10* (4,31E-11)	- 1,87E-10* (3,54E-12)	- 9,37E-11* (4,71E-12)	-7,84E-12* (3,36E-13)	- 6,19E-13* (4,13E-14)	- 4,50E-11* (2,04E-12)
N^3	4,55E-12* (3,77E-14)	2,04E-13* (1,06E-15)	4,29E-14* (2,63E-15)	8,52E-16* (4,54E-17)	4,37E-17* (1,88E-18)	9,84E-14* (3,54E-15)
$Região\ CO$	0,1154** (0,0446)	0,1191* (0,0395)	0,1394** (0,0592)	0,1473* (0,0622)	0,1801** (0,0894)	0,1377* (0,0429)
$Região\ N$	0,0378** (0,0177)	0,0399** (0,0184)	- 0,0438** (0,0208)	- 0,0975* (0,0299)	- 0,1073* (0,0343)	- 0,0417** (0,0196)
$Região\ NE$	0,1588** (0,0744)	0,1513** (0,0723)	0,1708* (0,0567)	0,1798* (0,0541)	0,2611** (0,0870)	0,2109** (0,0736)
$Região\ S$	0,1321* (0,0438)	0,1332** (0,0526)	0,1443* (0,0475)	0,1561* (0,0469)	0,1877** (0,0838)	0,1301** (0,0610)
$DUMP1_i$	0,4085** (0,1835)	0,4132** (0,1964)	0,4513** (0,2180)	0,4781** (0,2305)	0,5128** (0,2065)	0,3741* (0,1025)
$DUMP2_i$	0,2957* (0,0763)	0,3015* (0,0893)	0,3318** (0,1480)	0,3610* (0,1114)	0,4084* (0,0870)	0,2678* (0,0727)
$Constante$	4,9378* (0,4106)	4,7416* (0,3244)	4,2615* (0,3814)	6,0215* (0,4211)	6,1826* (0,4718)	5,0124* (0,2847)
$\boldsymbol{\rho}$	0,4023* (0,0790)	0,4466* (0,0821)	0,3519* (0,0614)	0,3106* (0,0821)	0,2819* (0,0784)	0,4176* (0,1031)
$Pseudo\ R^2$	80,32	79,95	82,04	78,71	79,11	83,15
Nº de Obs.	1209	1141	2324	318	276	5268

Fonte: Elaborado pelo autor a partir de resultados do trabalho.
Nota: * significante ao nível de 1%, ** significante ao nível de 5%,*** significante ao nível de 10%.
Obs.: Os valores em parênteses são os erros padrões.

A análise entre os clubes mostra que em comunidades com menor população há menor sensibilidade da oferta pública com relação à variação da parcela tributária paga. A justificativa pode estar no baixo nível de esforço fiscal destes municípios, quando comparado aos maiores clubes. Essa inferência pode ser fortalecida através da observação dos parâmetros relativos às elasticidades das transferências, que se comportam em sentido contrário às elasticidades-preço. Isso comprova a incapacidade das jurisdições menores (C1 e C2) em ofertar os serviços públicos com receita própria, oriunda da arrecadação municipal.

Verifica-se que as cidades de médio (C3 e C4), e principalmente de grande porte (C5), onde suas economias são mais desenvolvidas, há maior potencial no comércio, indústria e serviços, que são bases de incidência tributária dos municípios. De fato, a literatura internacional, como Turnbull (1992), e a nacional (Costa, 2013), comprova que além do tamanho da população afetar positivamente o esforço fiscal praticado pelas jurisdições, as transferências governamentais apresentam impacto negativo sobre o esforço fiscal.

Já os parâmetros de custos reportam às mesmas considerações do modelo da demanda: De uma forma geral, para esta abordagem, revelam que a exploração de economias de escala para os pequenos clubes se faz necessária para a obtenção de níveis adequados de eficiência na oferta dos serviços do governo, haja vista que o consumo adicional do serviço ofertado proporciona um custo marginal na sua provisão. Os parâmetros referentes à variável população apresentam-se inversamente relacionados às despesas públicas, e assim como no modelo anterior são decrescentes conforme o aumento do tamanho dos clubes. Contudo, para esta abordagem, é maior a magnitude dos coeficientes, sugerindo maiores economias de escala a serem exploradas quando comparado ao modelo da demanda.

Os coeficientes associados às demais variáveis de controle apresentam resultados particulares que ratificam suas especificidades na função oferta de bens públicos. O número de funcionários públicos municipais e o salário público, como esperado, contribuem para elevar as despesas públicas.

As *dummies* regionais apresentaram os mesmos padrões do modelo da demanda, contudo, com magnitude menor para cada clube e região, permitindo inferir sobre um *gap* entre os fatores de demanda e a disposição da oferta por serviços públicos com base nos fatores regionais. As *dummies* políticas também se mostram significativas para a oferta de serviços públicos municipais, são crescentes conforme o tamanho dos clubes e sugerem que o alinhamento político, tanto na esfera estadual, quanto na federal (este com menor impacto) é fundamental para a provisão de bens locais.

As Tabelas 2.8 e 2.9 mostram as elasticidades com respeito ao tamanho dos clubes e os parâmetros de congestionamento, para os modelos de demanda e de oferta[19]. Estes coeficientes mensuram a variação relativa da qualidade q das despesas - dada uma determinada quantidade Z do bem público - devido a uma variação relativa da população.

Tabela 2.8: Elasticidades e parâmetros de congestionamento por forma funcional – Modelo da Demanda

Observações	**Clube 1**	**Clube 2**	**Clube 3**	**Clube 4**	**Clube 5**	**Brasil**
Borcherding-Deacon						
Elasticidade do Tamanho do Clube						
$\gamma = \eta_N$	- 0,5848*	- 0,5897*	-0,6201*	- 0,6788*	-0,7238*	-0,6569*
Camaraderie						
Parâmetro de Congestionamento						
γ						
γ'	-0,5855*	-0,5905*	-0,6242*	-0,6793	-0,7244*	-0,6581*
γ''	-2,86E-08*	-1,42E-07*	-1,72E-07*	-2,01E-09*	-2,41E-11*	-7,67E-08*
Elast. Tamanho dos Clubes						
η_N	- 0,5854*	-0,5902*	-0,6211*	-0,6791*	- 0,7243*	-0,6572*
Flexível						
Parâmetro de Congestionamento						
γ						
γ'	-1,81E-04*	-1,00E-04*	-2,44E-05*	-8,93E-06*	-2,20E-06*	-1,19E-07*
γ''	9,01E-10*	1,21E-09*	1,46E-09*	3,46E-12*	1,13E-12*	1,99E-11*
γ'''	-4,47E-15*	-2,45E-14*	-6,40E-14*	-1,46E-16*	-2,05E-17*	-1,63E-13*
Elast. Tamanho dos Clubes						
η_N	-0,5996*	-0,6017*	- 0,6280*	-0,6861*	-0,7287*	-0,6625*
Mediana da população	3436	6907	18161	66044	186240	11089

Fonte: Elaboração própria a partir de resultados do trabalho
Nota: * significante ao nível de 1%.

[19] Ressalta-se que para o modelo B-D a elasticidade do tamanho do clube é o próprio parâmetro de congestionamento. Para as formas *camaraderie* e flexível, a elasticidade foi calculada pela mediana da população.

Tabela 2.9: Elasticidades e parâmetros de congestionamento por forma funcional – Modelo da Oferta

Observações	Clube 1	Clube 2	Clube 3	Clube 4	Clube 5	Brasil
Borcherding-Deacon						
Elasticidade do Tamanho do Clube						
$\gamma = \eta_N$	- 0,5897*	- 0,5919*	-0,6348*	- 0,6924*	-0,7418*	-0,7249*
Camaraderie						
Parâmetro de Congestionamento						
γ						
γ'	-0,5901*	-0,5921*	-0,6348*	-0,6930	-0,7418*	-0,7255*
γ''	-6,09E-09*	-1,21E-09*	-2,47E-11*	-3,78E-12*	-5,54E-11*	-7,80E-10*
Elast. Tamanho dos Clubes						
η_N	- 0,5900*	-0,5920*	-0,6348*	-0,6930*	- 0,7418*	-0,7254*
Flexível						
Parâmetro de Congestionamento						
γ						
γ'	-2,66E-05*	-5,98E-05*	-4,35E-07*	-1,60E-06*	-3,13E-07*	-3,38E-05*
γ''	7,80E-10*	1,78E-10*	8,52E-11*	7,05E-12*	5,57E-13*	4,14E-11*
γ'''	-4,33E-12*	-2,13E-13*	-3,90E-14*	-7,67E-16*	-3,80E-17*	-8,97E-14*
Elast. Tamanho dos Clubes						
η_N	-0,6001*	-0,6062*	- 0,6528*	-0,7068*	-0,7555	-0,7320*
Mediana da população	3436	6907	18161	66044	186240	11089

Fonte: Elaboração própria a partir de resultados do trabalho
Nota: * significante ao nível de 1%.

Como se esperava para os dados brasileiros, o tamanho dos clubes influencia a qualidade dos bens por meio do efeito congestionamento. Para todos os clubes, os parâmetros estimados situam-se entre zero e um, e apontam que o grau de publicidade não é constante, que o efeito é crescente, menor em pequenos clubes - ou cidades de menor porte.

A justificativa para os resultados apresentados, como expõem Mendes e Sampaio de Sousa (2006), está no fato de que, além dos vários serviços públicos em pequenos municípios serem subsidiados, as indivisibilidades que limitam a provisão de serviços em pequenas localidades concentram suas provisões em grandes cidades, e assim maiores despesas refletem não apenas o custo do congestionamento, mas também o fato que os clubes maiores oferecem um conjunto mais complexo de serviços, quando comparado aos clubes de pequeno porte.

2.7.1 Comparação estatística dos modelos estimados

A hipótese testada nesta etapa foi que a gama de serviços providos pelas municipalidades, para todos os tamanhos de clubes, atendem às preferências da demanda, ou a política discricionária da oferta. A partir da semelhança entre os resultados dos parâmetros aqui estimados, decidiu-se fazer uma comparação estatística entre os dois modelos para identificar o poder explicativo destes como determinantes da provisão pública, por tamanho de clubes. As comparações foram feitas com base na forma flexível.

A estratégia utilizada envolve o Teste J, proposto por Davidson e Mackinnon (1981), o qual consiste na construção de uma combinação convexa de abordagens competitivas para testar a significância estatística (a estatística t é adequada) de um modelo $Y = X_1\beta_1 + \varepsilon$, frente ao outro, $Y = X_2\beta_2 + \varepsilon$ (onde X descreve as variáveis independentes e β o vetor de parâmetros). Nesse caso, tem-se $Y = (1 - w)X_1\beta_1 + wX_2\hat{\beta}_2 + \varepsilon$ (no qual, $X_2\hat{\beta}_2$ é o vetor de valores ajustados de Y do modelo da oferta) para testar $w = 0$. Depois se testa $w' = 0$ $Y = w'X_1\hat{\beta}_1 + (1 - w')X_2\beta_2 + \varepsilon$. Se w é estatisticamente diferente de 0 e w' não o é, o modelo da oferta tem maior poder explicativo do que o modelo da demanda. Se os parâmetros w e w' são simultaneamente significativos ou não significativos, então o teste é ambíguo e não dá qualquer informação sobre o poder explicativo dos modelos.

Para as várias especificações de clubes, os *t*-valores correspondentes aos testes são dados pela Tabela 2.10.

Tabela 2.10. Comparação estatística dos modelos de demanda e oferta

	Clube 1	Clube 2	Clube 3	Clube 4	Clube 5
Especificação da Hipótese					
Oferta/Demanda	30,29	28,12	15,54	1,55	-1,74
Demanda/Oferta	-1,46	-1,69	-4,01	19,03	11,15

Fonte: Elaborado pelo autor a partir de resultados do trabalho.

A constatação é que para pequenos municípios as características da demanda se adequam melhor aos dados na determinação da provisão pública. No caso de C1 e C2, a estatística *t* justifica a total significância desta abordagem, quando apresenta a rejeição da hipótese de ajuste do modelo da oferta frente à demanda, mas não rejeita o contrário. Já para o caso de C3, os parâmetros apontam uma boa adequação entre os dois modelos, o que sugere, para esse conjunto de municípios, que tanto o modelo da demanda, quanto o modelo da oferta se adequam como determinantes dos serviços públicos no Brasil. Contudo, a magnitude do

parâmetro mostra que o modelo da demanda ajustado proporciona um melhor resultado quando comparado ao modelo da oferta, do que no caso contrário, ao nível de 1% de significância.

Já em C4 e C5, os dados apresentam definitivamente o melhor ajuste do modelo da oferta frente à demanda. Ou seja, para clubes maiores, a política pública direcionada pelo lado da oferta prevalece como determinante da provisão pública.

Estes resultados vão em direção contrária aos estudos da literatura internacional, como Hayes e Slotje (1987), e Guengant, Josselin e Rocaboy (2002), os quais encontraram maior poder explicativo do modelo da oferta para pequenos municípios. Não obstante, os resultados aqui apresentados parecem consistentes, pois, apesar de não haver na literatura nacional trabalho semelhante para comparação, corroboram com as predições teóricas do Federalismo Fiscal brasileiro. De uma forma geral, mostram que à medida que os clubes crescem, os representantes eleitos se afastam do compromisso com a escolha do eleitor mediano. Os gestores passam a prover bens públicos (com menor qualidade) de forma discricionária, no sentido de maximizar seus interesses, o que impacta na magnitude do efeito *crowding*, gerando um congestionamento maior para os clubes maiores.

2.8 CONSIDERAÇÕES FINAIS

Buscou-se aplicar os modelos alternativos - demanda e oferta - de provisão pública a fim de mensurar o impacto do tamanho das cidades sobre a qualidade dos serviços públicos, via análise de publicidade e rivalidade dos bens, nas municipalidades brasileiras sob o arcabouço da Teoria dos Clubes, com o a análise espacial, proposta por Anselin (1988).

Além das variáveis básicas, no modelo da demanda, foi incorporada uma matriz Ω de características socioeconômicas normalmente utilizadas na literatura sobre o eleitor mediano, tais como Percentual de Pessoas com menos de 18 anos; Percentual de Pessoas com mais de 65 anos; Esperança de Vida e Índice de Desenvolvimento Municipal – IDHM/Educação, e *dummies* regionais e de Participação em Consórcio Municipal.

Já no modelo direcionado a oferta, uma matriz Π representou a inserção de algumas variáveis representativas de elementos institucionais e políticos determinantes para a provisão governamental. O Número de Funcionários Públicos Municipais; O Salário Médio dos Funcionários Públicos da localidade; além das *dummies* regionais; de Alinhamento Político com o Governo do seu Estado e Alinhamento Político com o Governo Federal.

Para os dois modelos, os resultados apontam para existência da correlação espacial entre os municípios brasileiros, sugerindo que os padrões de gastos em determinada localidade influencia nas decisões fiscais de vizinhos. Com respeito aos coeficientes, em sua maioria, as variáveis se apresentaram significativas e determinantes para a provisão pública local, e comprovam que à medida que as cidades se desenvolvem, os gestores eleitos se afastam do compromisso "firmado" com o eleitor mediano, e passam a prover bens públicos com menor qualidade, já que visam apenas maximizar o seu interesse.

Na prática, estes resultados indicam que os mecanismos de gastos locais para a provisão de bens são influenciados tanto pela participação e "força política" do eleitor, quanto pela discricionariedade ou autonomia do gestor na execução da política pública.

Comprova-se a hipótese referendada por alguns autores, como Turnbull e Mitias (1999), Turnbull e Geon (2006), que o tamanho dos clubes, e o custo dos serviços públicos diferenciam os resultados da ação de provisão, corroborando com a literatura sobre o tema a qual defende que em jurisdições menores a proximidade dos eleitores com seus agentes públicos exerce maior pressão e controle sobre seus atos políticos, e que preferências individuais agregadas apresentam menor heterogeneidade e podem representar a comunidade como um todo. Não obstante, em localidades maiores prevalece a ideia de Wyckoff (1990), Dye e Mac Guire (1997), que os representantes eleitos são relativamente livres para escolher as políticas que vão empreender, devido ao menor poder de barganha da sociedade como um todo, para exigir boa aplicação dos recursos governamentais.

Este resultado gera um congestionamento maior para as maiores clubes, o que vai de encontro à literatura internacional. Uma avaliação detalhada permite entender que as indivisibilidades características do fornecimento de certos bens e serviços limitam a sua provisão aos grandes centros urbanos, assim, os maiores gastos refletem não só um custo de congestionamento, mas adicionalmente, o fato de que os maiores clubes provém um conjunto mais diversificado e complexo de serviços, em comparação aos pequenos clubes. Este fenômeno é conhecido na literatura por "Efeito Zoo".

3. INDIVISIBILIDADES NA PROVISÃO LOCAL DE BENS PÚBLICOS NO BRASIL: UMA ANÁLISE DAS EVIDÊNCIAS DO "EFEITO ZOO"[20]

3.1 INTRODUÇÃO

A partir dos trabalhos seminais de Borcherding e Deacon (1972) e Bergstrom e Goodman (1973), a avaliação da publicização na provisão de bens públicos tem ganhado enorme destaque na literatura. A abordagem padrão envolve a geração de estimativas das elasticidades preço e renda da demanda, a partir de uma função multiplicativa que usa como principal argumento uma *proxy* representativa do tamanho da comunidade. Essa abordagem permite a mensuração do congestionamento desses bens decorrente da variação populacional, os quais sugerem que os governos locais fornecem serviços à sociedade que apresentam maior ou menor grau de rivalidade no consumo, assim como acontece com os bens privados.

Especificamente, a relação entre o tamanho da população e o consumo do bem, tem sido estudada sob a forma $Z = N^{\gamma}q$, na qual Z indica o número de unidades de serviços providos por uma localidade, q o nível dos bens públicos consumidos por um indivíduo ou população, N o tamanho da população e γ representa o parâmetro de congestionamento. A partir de tal formulação, a literatura econométrica possibilitou o cálculo das estimativas do parâmetro para a gama de serviços locais providos, as quais indicam que se $\gamma = 0$, Z é um bem público puro, do qual cada consumidor captura exatamente a quantidade disponível. Valores maiores para o parâmetro refletem o efeito *crowding out*, sendo que $\gamma = 1$ indica congestionamento, e apresenta propriedade de rivalidade absoluta, de modo que cada membro da comunidade irá capturar tão somente a 'enésima' fração de Z (ou seja, a constatação aqui é que o consumo de um indivíduo interfere no consumo do outro, como um bem privado).

A literatura sobre economia pública local, desde então, tem trabalhado no sentido de analisar como a aglomeração populacional impacta sobre o nível dos serviços governamentais providos, e como as municipalidades exploram economias de escala que caracterizam a produção desses serviços. Em linhas gerais, os estudos internacionais, incluindo-se ai os mais recentes como Reiter e Weicherieder (1997); Baudry, Leprince e Moreau (2002), e Guengant, Josselin e Rocaboy (2002) apontam para o supercongestionamento e a pouca relevância de efeitos de escala no consumo dos serviços locais. Já na literatura nacional, (Mendes e

[20] Este capítulo resultou no estudo de Silva e Sampaio de Sousa (2015a), apresentado no Encontro de Economia Baiana 2015, e Silva e Sampaio de Sousa (2015b) aceito no Encontro da Associação Brasileira de Economia Regional e Urbana (Enaber) 2015.

Sampaio de Sousa, 2006; Araújo e Siqueira, 2014) os coeficientes estimados variam, com valores abaixo da unidade, e indica certa publicidade (*publicness*) dos serviços locais no Brasil, e economias de escala a serem exploradas.

Não obstante a importância das análises realizadas desde então, Oates (1988) criticara a metodologia utilizada para computo dos parâmetros de congestionamento, a partir da observação de que muitos bens públicos são indivisíveis, e expôs que os estudos sobre o tema não consideram que a variedade de serviços prestados em cidades maiores excede a dos menores municípios e a oferta pública cresce com o tamanho das localidades. A conclusão é que, apesar disto, o nível mais baixo de gasto em cidades menores não deve ser explicado pela menor aglomeração, mas simplesmente pelo fato de que certas categorias de despesas - como um zoológico - precisam de um tamanho mínimo populacional para serem fornecidas. Este fenômeno ele passou a chamar de "efeito zoo".

A ideia base do efeito, de acordo com Oates (1988), é que a "primeira unidade" da produção dos bens pode requerer um dispêndio substancial de tal modo que não se torne desejável proporcionar o bem à população até que esta atinja o tamanho para o qual a soma das taxas marginais de substituição seja igual (ou superior) ao custo da primeira unidade. Isto exposto, fica claro que a produção dos serviços especializados muitas vezes exige escala de provisão incompatível com a demanda da população residente em pequenos municípios. A consequência é a busca, pela população, de serviços de maior qualidade, ofertados somente em cidades maiores mais próximas. Sem embargo, a relação positiva entre o tamanho da população e o gasto local em serviços públicos não leva em conta o fato de que nos municípios mais populosos, a cesta de serviços ofertados é mais ampla e mais diversificada, e por essa razão, as estimativas de congestionamento existentes tendem a ser superestimadas.

Apesar da clara intuição, a literatura ainda carece de evidências sobre tal efeito. Este trabalho então, visando contribuir com a literatura, testa a existência do “efeito zoo” e avalia a magnitude do viés do congestionamento, a partir de um conjunto de dados de 2903 municípios brasileiros que formam grupos voluntários criados no intuito de resolver o problema da "fragmentação municipal" no país, e de permitir o financiamento e gestão coletiva de alguns bens públicos locais. Com outras palavras, utilizando-se de econometria espacial, o estudo concentra-se nos Consórcios Públicos Intermunicipais e na gama de serviços por eles providos baseados em benefícios das economias de escala, internalização de repercussões negativas na produção, e dificuldades de coordenação, permitindo assim um nível de decisão superior ao que seria provido individualmente pelo município.

As questões levantadas sobre o "efeito zoo" podem contribuir para um debate sobre a organização ótima do setor público, no que diz respeito ao provimento de serviços públicos menos caros, melhor adequação da provisão as preferências dos cidadãos espacialmente heterogêneos, e por fim, a possibilidade de incentivos do governo para desenvolvimento da cooperação entre os municípios locais em zonas mais densamente povoadas.

Este artigo está organizado da seguinte forma: A seção 3.2 discute a fundamentação teórica do "efeito zoo", seguida de pequena literatura empírica sobre o tema. A próxima seção aborda o contexto institucional brasileiro a respeito dos Consórcios Públicos Intermunicipais, a seguinte seção apresenta a metodologia econométrica utilizada no trabalho, adiante, são apresentados os resultados das estimativas, e por fim as considerações finais.

3.2 INTUIÇÃO TEÓRICA DO EFEITO ZOO

A base intuitiva, fundamentada por Oates (1988), a respeito da indivisibilidade na provisão dos bens públicos associada ao tamanho populacional, e consequente análise viesada dos parâmetros de congestionamento parte de um contexto muito simples, contudo pertinente, desenhado a partir da consideração de uma atividade local, "Parques e Recreação" a qual fornece dois bens: (1) *Playgrounds* locais (Z_1), providos para todas as localidades; e (2) Zoológicos (Z_2) providos somente para localidades onde a população excede um nível crítico, N^*. Então:

$Z_1^* = q_1 N^\gamma$, onde $q_1 > 0$ para todas as localidades;

$Z_2^* = q_2 N^\gamma$, onde $q_2 = 0$ para localidades nas quais $N \leq N^*$, e

$q_2 > 0$ para localidades nas quais $N > N^*$.

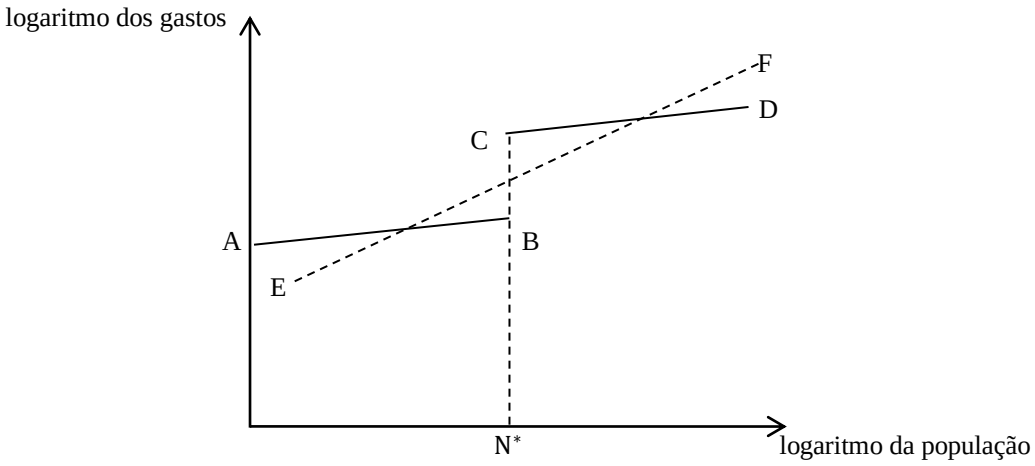
Assume-se, por simplicidade, que γ é igual para ambos os bens.

O que o autor argumenta, a partir de tal formulação, é que a condução dos estudos econométricos não faz distinção entre q_1 e q_2, pelo contrário, usa-se os dados agregados das despesas que representam a atividade "Parques e Recreação". A natureza do problema então torna-se aparente, como observada na Figura 3.1.

A relação entre os gastos com serviços públicos e o tamanho da população, na forma logarítmica, pode ser observada conforme representada na figura pela curva descontínua ABCD. A inclinação dos segmentos AB e CD indica a elasticidade gastos-população (α) com cada um dos bens locais, e a descontinuidade em N^* caracteriza a situação em que localidades que apresentam $N > N^*$são providas tanto de *playgrounds*, como no zoológico. Contudo, se

for ignorada esta distinção e realizada uma estimação da relação log-linear entre os gastos e toda classe de atividades e tamanho da população, usando os pontos observados ao longo da linha dos dois segmentos, obtém-se uma curva estimada com inclinação semelhante à linha pontilhada EF.

Figura 3.1: Congestão em bens públicos locais



Fonte: Oates (1988)

Em resumo, considerando que a verdadeira elasticidade da população para a provisão de cada serviço local, α, é a inclinação das linhas dos segmentos AB e CD, as elasticidades estimadas que desconsideram as indivisibilidade dos bens representam a inclinação EF, ou seja, um coeficiente superestimado. Como o parâmetro de congestionamento γ é uma transformação linear positiva da elasticidade da população, a sua mensuração também estará superestimada.

Com efeito, nesta simplificação Oates (1988) utiliza-se de uma hipótese básica envolvendo uma única atividade e descontinuidade definida nas funções de gastos, entretanto argumenta que na prática cada ampla classe de serviços locais engloba inúmeras funções e subfunções, e que aumentam com o tamanho da população. Não obstante, relata que seria possível incorporar o "efeito zoo" às análises de uma forma mais realista, a partir de uma abordagem geral que postula a verdadeira relação descontínua na provisão pública.

Em consonância com o modelo do eleitor mediano, o autor desenvolve um quadro em que as despesas locais E estão positivamente relacionados com o nível individual de serviços Z e com a gama de serviços prestados R:

$$E = f(Z, R) \tag{3.1}$$

por suposição, Z e *R* crescem em função da população local *N*, ou seja:

$$Z = g(N) \text{ e } R = h(N) \text{ com } g'(N) > 0 \text{ e } h'(N) > 0 \tag{3.2}$$

Mais especificamente, Z é definido na literatura como $L=TL/N^{\gamma}$, onde TL é o nível total de serviços públicos providos em uma localidade com população de tamanho N, e γ é o parâmetro de congestionamento (também chamado de "parâmetro de capturalidade") tal que $[\gamma = \varepsilon_{POP}/(1 + \varepsilon_{PRICE})]$ onde ε_{POP} é a elasticidade da população com relação aos gastos, e ε_{PRICE} a elasticidade preço-demanda.

A derivada total pode ser escrita como:

$$\frac{dE}{dN} = \frac{\partial E}{\partial Z}\frac{dZ}{dN} + \frac{\partial E}{\partial R}\frac{dR}{dN} > 0 \tag{3.3}$$

onde os termos do lado direito da equação são positivos. Pensando na mensuração destas variáveis em termos de logaritmo, a elasticidade da população com relação aos gastos para um dado nível de oferta pública pode ser representada por:

$$\varepsilon_{POP} = \frac{\partial E}{\partial Z}\frac{dZ}{dN} = \frac{dE}{dN} - \frac{\partial E}{\partial R}\frac{dR}{dN} = \hat{\varepsilon}_{POP} - \frac{\partial E}{\partial R}\frac{dR}{dN} \tag{3.4}$$

onde o termo $\frac{\partial E dR}{\partial R dN}$ incorpora o "efeito zoo" e ε_{POP} a elasticidade da população considerando tal efeito. Já $\hat{\varepsilon}_{POP}$, é a elasticidade estimada quando o efeito é ignorado, como por exemplo, em Borcherding e Deacon (1972), e Bergstrom e Goodman (1973). Consequentemente, sem considerarem a presença de tal fenômeno, esses estudos empíricos supervalorizam sistematicamente a extensão das economias de escala no setor público. De outra forma, suas estimações não consideraram a relação entre a gama de serviços públicos e o tamanho da localidade, ou seja, $R = h(N)$.

Apesar do argumento de Oates (1988) ser bastante intuitivo e gerar consequências importantes sobre a concepção de modelos econométricos para análise do efeito, o mesmo sofre com a falta de evidências empíricas. As poucas referências estão relatadas a seguir.

3.2.1 A Literatura Empírica

Antes da formalização do tema, Schmandt e Stephens (1960), tentaram inferir a existência do “Efeito Zoo”, quando buscaram explicar o número de serviços prestados por alguns municípios do condado de *Milwaukee*, nos Estados Unidos. Como resultado, eles destacam a relação positiva esperada entre o tamanho da população dos municípios e da gama de serviços públicos que fornecem. Desde então, muitos trabalhos surgiram no intuito de estimar o grau de congestionamento dos bens públicos, contudo, sem apresentar evidências do efeito.

Mais recentemente, Frère, Hammadou e Paty (2011), usando um conjunto de dados de comunidades francesas, ou seja, de governos locais que reúnem vários municípios vizinhos, a fim de gerir coletivamente alguns produtos locais, visaram estabelecer as evidências empíricas para este fenômeno. Utilizando-se da econometria espacial, os resultados estimados favorecem a ideia proposta por Oates (1988), de que a gama de serviços providos por comunidades maiores é mais diversificada do que a oferecida pelas menores. Afirmaram que, em decorrência do “efeito-zoo”, o desenvolvimento de uma comunidade poderia ser considerado como uma boa maneira de aumentar a variedade de serviços públicos locais fornecidos. E que o efeito é menos intenso nas áreas rurais, sugerindo que economias de escala podem existir em áreas urbanas, sendo benéficas, devido à internalização de *spillovers* (ou transbordamentos).

Na literatura nacional, apesar de muitos trabalhos utilizarem a abordagem do eleitor mediano na análise da determinação das despesas locais, e o grau de congestionamento na provisão pública, poucos enfatizam a relevância do fenômeno. Contudo, ressalta-se o trabalho de Mendes e Sampaio de Sousa (2006), os quais argumentam com base nos resultados do estudo que maiores despesas públicas em cidades grandes não refletem significativamente o custo do congestionamento, mas ao fato de que há maior diversificação dos serviços nessas localidades do que em cidades de menor porte; e Sampaio de Sousa e Piola (2010), que mensuram o índice de eficiência dos gastos com saúde nas municipalidades brasileiras, e aludiram à existência do “efeito-zoo” como uma das possíveis causas para a análise relativa ao aumento da ineficiência em cidades maiores.

Diante do exposto, entende-se que a análise desse efeito para o Brasil, dará uma visão crítica sobre os custos dos serviços nas municipalidades, e um melhor entendimento do problema na diversificação no fornecimento de bens públicos locais e perdas de benefício marginal pelo consumo de um bem quando o número de consumidores aumenta. Dessa forma, o presente trabalho trata de um fenômeno muito maior que o simples caso da diversidade de serviços públicos fornecidos pelos municípios brasileiros.

3.3 O CONTEXTO INSTITUCIONAL BRASILEIRO E A DIVERSIDADE NO FORNECIMENTO DE SERVIÇOS PÚBLICOS

Após a Constituição Federal de 1988 (CF/88), a descentralização política e fiscal proporcionou maior autonomia às jurisdições na condução de suas políticas públicas, na qual os governos locais passaram a decidir a melhor forma de aplicação e gerência dos recursos destinados ao atendimento das necessidades da população, sem interferência do governo central. Concomitante a esse processo de descentralização, verificou-se no país um aumento significativo no número de municípios criados desde então, e consequente impacto nas contas públicas induzido pelas regras na distribuição de recursos necessários ao provimento de bens a sociedade.

A fim de superar essa atomização de jurisdições e fazer frente ao aumento dos gastos de governos locais, os municípios passaram a se organizar na forma de arranjos associativos, chamados Consórcios Públicos Intermunicipais (CPIs), para a realização de ações que pudessem recobrar escalas produtivas e financeiras, e assim incrementar a qualidade dos serviços públicos prestados à sua população.

Ordenados pela Lei nº 11.107/2005, os inúmeros CPIs existentes no Brasil apresentam benefícios diretos, pois possibilitam o agrupamento de pequenos municípios, agregando um número suficiente de habitantes, para que estes possam ser contemplados com diversos programas dos governos, em diversas áreas, que seriam destinados apenas a municípios com número maior de habitantes, possibilitando melhor capacidade de alocação dos governos municipais, ampliação do atendimento aos cidadãos, eficiência na utilização dos recursos, poder de barganha dos municípios junto às organizações privadas e outras esferas do poder público. Em suma, o consorciamento propicia o aumento da eficiência e qualidade dos serviços providos, pois exige escala de produção compatível com a demanda da população correspondente.

Segundo dados do Instituto Brasileiro de Geografia e Estatística (IBGE, 2011), dos 5565 municípios brasileiros, 52% participam de CPIs concentrados em todas as regiões do país, sendo que 44% dos consórcios foram formados em localidades com população de até 50.000 habitantes, o que sinaliza o fato de que o tamanho subótimo de grande parte dos municípios brasileiros restringe a exploração de economias de escala que caracterizam a produção dos serviços, e limitam a redução dos custos e uso eficiente dos recursos, e por isso a coordenação entre pequenos municípios para a oferta de serviços governamentais tende a

otimizar a escala de operação, cortar despesas e proporcionar a expansão dos serviços disponíveis àquele grupo de jurisdições.

Os dados agrupados da Tabela 3.1 apresentam a distribuição de municípios que formam consórcios em todo o Brasil, por faixa populacional e por região do país.

Tabela 3.1 - Distribuição de municípios que formam consórcios por regiões do Brasil

Faixa Populacional	Norte	Nordeste	Sudeste	Sul	Centro-Oeste	Total
Até 5 000	4	59	281	343	38	725
De 5 001 a 10 000	5	98	276	222	37	638
De 10 001 a 20 000	16	168	257	194	60	695
De 20 001 a 50 000	20	149	201	117	24	511
De 50 001 a 100 000	11	50	71	45	4	181
De 100 001 a 500 000	9	17	75	36	-	137
Mais de 500 000	-	5	9	2	-	16
Total	**65**	**546**	**1 170**	**959**	**163**	**2 903**

Fonte: Elaborado pelo autor com base nas informações do IBGE - Base Munic 2011.

Os dados do IBGE (2011) indicam ainda que 11[21] políticas públicas mobilizam a organização dos consórcios no Brasil, sendo que a cooperação encontrada mais frequentemente na articulação entre os municípios é a da saúde, com 40,63% do total de arranjos associativos, seguida pelos consórcios de meio ambiente e saneamento básico, com a ocorrência de 14%. No entanto, a perspectiva e predominância de algum tipo de cooperação dos municípios dependem do equilíbrio entre duas forças contraditórias imanentes à gestão das cidades, quais sejam, as forças centrífugas, representadas pelas demandas específicas da população local e o receio dos gestores de perder o controle o os dividendos das decisões de política pública que dão origem ao individualismo local, e por outro lado, as forças centrípetas, vinculadas aos limites da capacidade fiscal e aos problemas que transbordam as fronteiras de um único município (NIERO, 2011).

Com efeito, a maior ocorrência de consórcios de saúde, educação, saneamento básico e desenvolvimento urbano concentra-se na faixa de municípios com população entre 5.000 e 50.000 habitantes, o que não é excepcional, pois são nesses municípios que as demandas sociais mais se intensificam perante o poder público, e onde a capacidade financeira dessas localidades são mais deficientes. Chama atenção o fato da maioria das CPIs nessas áreas estarem concentradas nas regiões mais desenvolvidas do país, o que pode ser justificado pela maior capacidade de articulação política da gestão em formar as associações.

[21] Cultura, Desenvolvimento Social, Desenvolvimento Urbano, Educação, Emprego e Trabalho, Habitação, Meio Ambiente, Saneamento Básico, Saúde, Transporte, Turismo.

A Tabela 3.2 apresenta a quantidade e percentual de municípios por tipos de CPIs entre regiões do país.

Tabela 3.2 - Tipos de CPI's por região do Brasil

CPI's	Norte	Nordeste	Sudeste	Sul	Centro-Oeste	Qtde	Total(%)
Saúde	25	312	974	865	112	2288	40,63%
Meio Ambiente	14	125	275	225	65	704	12,50%
Turismo	5	48	214	163	26	456	8,10%
Saneamento Básico	14	156	113	129	14	426	7,57%
Desenvolv. Urbano	12	100	170	93	27	402	7,14%
Educação	12	86	95	73	14	280	4,97%
Cultura	10	60	94	70	14	248	4,40%
Habitação	8	95	63	61	14	241	4,28%
Desenvolvimento Social	8	66	75	59	24	232	4,12%
Transporte	5	52	79	48	27	211	3,75%
Emprego e Trabalho	3	48	46	36	10	143	2,54%
Total	**116**	**1148**	**2198**	**1822**	**347**	**5631**	**100%**

Fonte: Elaborado pelo autor com base nas informações do IBGE - Base Munic 2011.

Sem embargos, essas estruturas “supra-municipais” coexistentes com a estrutura municipal, baseadas no princípio da voluntariedade, visam atender as competências atribuídas à sua gestão delimitando as fronteiras de seu poder de decisão às políticas que não se aplicariam eficientemente em qualquer outra unidade do governo isoladamente. Esta forma de cooperação tem sido amplamente promovida pelo Estado, inclusive por meio de incentivos financeiros, como a solução para o problema da "fragmentação municipal". Espera-se que as associações permitam a redução das despesas públicas, e limitem as desigualdades fiscais entre seus membros, e consequentemente possibilitem economias de escala substanciais na provisão, para maior publicidade dos bens públicos.

Este cenário constitui, então, a melhor configuração para a análise do efeito zoo, no sentido que evidencia as limitações dos governos, quando fragmentados, em alcançar o equilíbrio entre as forças de demanda e oferta do serviço público, e a necessidade de se coordenar espacialmente para que a utilização dos recursos sejam bem focalizados, e que atinjam “quórum” mínimo para torná-lo eficiente.

Para capturar todos os impactos específicos desta configuração, foi elaborada uma estratégia metodológica, exposta a seguir, com base na análise exploratória espacial, capazes de estimar os determinantes da gama de serviços públicos nas localidades brasileiras e o comportamento dos governos municipais em resposta à heterogeneidade de sua população.

3.4 METODOLOGIA

3.4.1 Base de Dados e Estratégias de Estimação do Modelo

As variáveis e *proxies* utilizadas neste estudo foram utilizadas com base em dados censitários de 2010, de informações sobre o perfil dos municípios brasileiros - Base Munic 2011, relativo a temas da administração pública municipal e articulações intrainstitucionais, ambos fornecidos pelo Instituto Brasileiro de Geografia e Estatística-IBGE; também do FINBRA, Finanças Brasil, disponibilizados pela Secretaria do Tesouro Nacional do Ministério da Fazenda (STN), no que tange às funções de despesas públicas.

Com bases nessas informações, a ideia básica para testar a existência e magnitude do "efeito zoo" é estimar o impacto do tamanho da população (N) sobre a gama de serviços públicos (R) prestados em 2903 municípios que declararam participar de algum tipo de CPI. Para tal, Assume-se que a gama de serviços depende, também, de algumas características sócioeconômicas exógenas (X) das comunidades que formam o consórcio *i*.

As estimações então partem de:

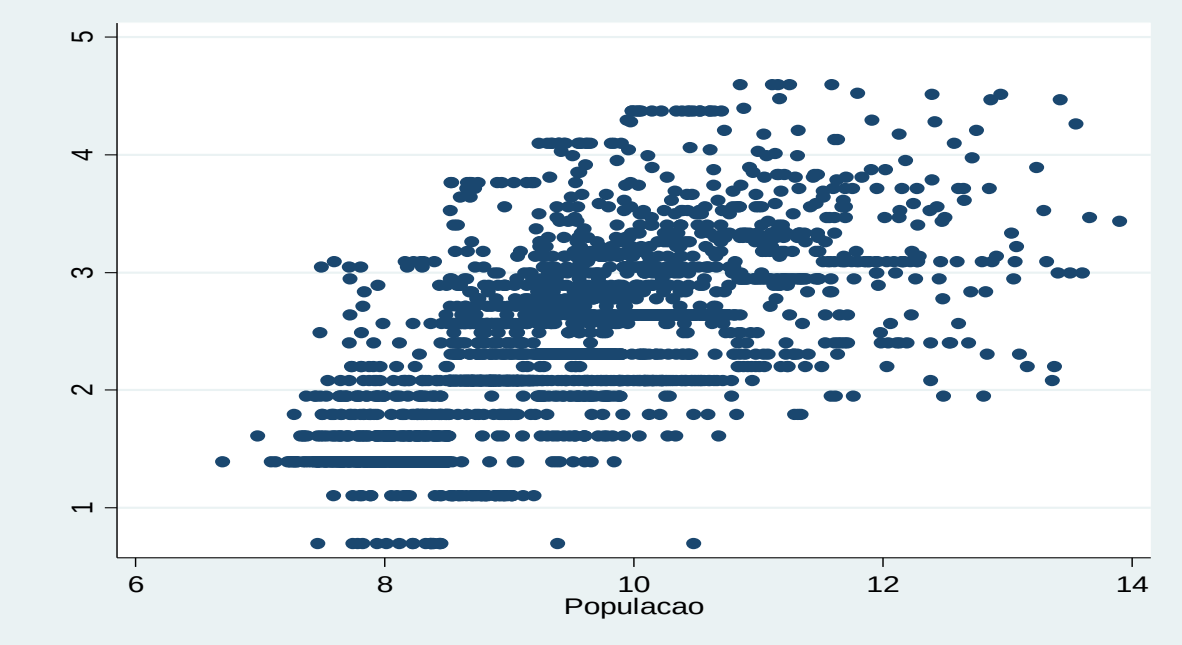
$$R = h(N, X) \tag{3.5}$$

Mais precisamente, (R) é o número de serviços providos por cada localidade consorciada. Para cada consórcio, estão listados os municípios membros e com base em dados municipais do IBGE de 2011, todas as competências prestadas de despesas públicas. N é a população total atendida por cada CPI. Para um consórcio agrupando *n* municípios (j) com população Pop_j, N é definido por:

$$N = \sum_{j=1}^{n} Pop_j \tag{3.6}$$

A Figura 3.2 representa o número de competências fornecidas pelos consórcios com relação ao tamanho populacional (*log*), expondo a correlação entre as duas variáveis. Percebe-se que o número de competências dos serviços governamentais cresce à medida que o contingente populacional aumenta.

Figura 3.2: Número de Serviços Públicos e Tamanho da População



Fonte: Elaboração própria a partir de dados da pesquisa.

Por fim, as *proxies* representativas das características socioeconômicas exógenas (X) consideradas são:

- **O Índice Herfindahl *(H)***, que leva em conta a composição, ou concentração de cada consórcio em termos de população, isto é, usando a mesma notação da Equação (3.6):

$$H = \sum_{j=1}^{n} [Pop_j/N]^2 \tag{3.7}$$

Este índice oscila entre $1/n$ e 1. O elevado valor deste implica que o grupo inclui um pequeno número de municípios densamente povoados, enquanto que um valor baixo corresponde a um grande número de pequenos municípios. O sinal esperado para o coeficiente associado é incerto, já que o número de serviços fornecidos pelos consórcios depende em parte da escolha das jurisdições envolvidas a fim de manter um serviço a nível municipal. Assim, eles vão decidir entre perder o poder político sobre um determinado serviço (e arriscar enfraquecer os laços com o eleitorado), ou aproveitar as vantagens das economias de escala e aumentar a coordenação com as políticas municipais vizinhas.

- ***Dummy* urbana (*DU*)** para captação das especificidades espaciais dos consórcios com relação às áreas urbanas e rurais. Esta *dummy* apresenta valor 1 para municípios localizados em áreas urbanas e 0 para municípios em áreas rurais, e permite checar se a

relação entre o fornecimento de serviços públicos e o tamanho da população difere ao longo do "gradiente urbano-rural". Também, é possível interagir esta *proxy* com o tamanho da população para se descobrir onde os efeitos *spillovers* são comumente predominantes, ou seja, o "efeito zoo" é mais intenso em áreas urbanas ou em áreas rurais? Em suma, se o agrupamento de municípios permite a internalização de *spillovers*, as localidades urbanas tenderão a transferir mais competências para os CPI's do que seus vizinhos rurais, *ceteris paribus*. Além disso, seguindo Josselin, Guengant e Rocaboy (2008), se localidades com menor taxa de urbanização apresentam também menor contingente populacional, é fato que os cidadãos, nesse caso, exercem maior controle sobre as ações do governo, isto é, o modelo de demanda predomina sobre o da oferta na determinação da provisão governamental, consequentemente, a fim de preservar a forte ligação com as preferências dos munícipes rurais, os gestores consideram essencialmente o poder de decisão daqueles com respeito aos serviços públicos locais, e assim e "efeito zoo" torna menos intenso.

- **Área (*A*)** total de cada município que forma o CPI. Esta variável é bastante utilizada em trabalhos do gênero como *proxy* para captação de efeitos de escala. Mais especificamente, ela mede o impacto da densidade populacional sobre o número de serviços prestados pela comunidade. Em comunidades com densidade populacional relativamente baixa (ou seja, para um dado nível da população, a área de superfície é relativamente grande), os ganhos de economias de escala serão tão pequenos que os municípios serão mais propensos a reter o poder de decisão e transferir alguns serviços para os consórcios.

- **(*Des*) é a taxa de desemprego** de cada município. O impacto esperado desta variável sobre o número de competências é incerto. Por um lado, se a existência da associação de municípios é vista como uma solução para os desequilíbrios no mercado de trabalho local, espera-se um impacto positivo. Por outro lado, os municípios com elevada taxa de desemprego, em termos relativos, podem preferir atender o poder de decisão da demanda nesta sensível área, a fim de manter uma relação estreita com os eleitores.

- ***(G)* é o Índice de Gini**, que mensura as desigualdades entre os municípios que formam um consórcio. Baseado na distribuição de renda municipal, este indicador de heterogeneidade tende a 0 quando há completa igualdade de renda e alcança 1 quando a desigualdade atinge seu nível máximo. Assume-se que quadros de grande heterogeneidade entre os municípios fornecem incentivos para a provisão ou ampliação de alguns serviços públicos (como os sociais e de desenvolvimento) em nível de município. Assim, espera-se um sinal positivo para este indicador de desigualdade.

- ***(Pop18)*** **e** ***(Pop65)*** denotam respectivamente, para cada consórcio, o percentual da população com idade inferior a 18 anos e acima de 65 anos. Os sinais esperados para esses coeficientes são incertos. Supõe-se que se o peso dessas pessoas nas localidades é relativamente alto, os municípios tendem a transferir mais competências relacionadas a essa faixa populacional (ou correspondentes à sua demanda por serviços públicos locais particularmente elevados), a fim de diminuir custos de produção baseados em economias de escala, ou para melhorar a qualidade de os serviços públicos. No entanto, os municípios também podem querer manter estas competências por razões eleitorais.

A Tabela 3.3 apresenta as estatísticas descritivas das variáveis utilizadas:

Tabela 3.3: Estatísticas descritivas

Variável	Observações	Média	Desv. Padrão	Mínimo	Máximo
R	2903	6,18	9,02	1	94
N_i	2903	30497,36	91550,89	811	238564
$A_i(Km^2)$	2903	945,60	2669,55	3,56	76351,79
H	2903	0,54	0,11	0,09	0,85
G	2903	0,48	0,06	0,28	0,78
Des	2903	5,63	3,26	0,01	21,19
$Pop65$ (%)	2903	12,2	3,71	2,85	29,38
$Pop18$ (%)	2903	23,09	4,13	11,52	42,91

Fonte: Elaboração própria a partir de dados da pesquisa.

Ressalta-se que o valor máximo referente a gama de serviços prestados considera a soma das competências associadas a todas as funções de gastos, declaradas pelos municípios consorciados.

A partir da *log*-transformação das variáveis citadas, a análise dos determinantes da gama de serviços públicos providos pelas municipalidades brasileiras, foi realizada com base no seguinte modelo:

$$\ln(R_i) = \beta_0 + \beta_1 \ln(N_i) + \beta_2 DU_i \ln(N_i) + \beta_3 \ln(A_i) + \beta_4 \ln(Des_i) + \qquad (3.8)$$

$$\beta_5 \ln(Pop18_i) + \beta_6 \ln(Pop65_i) + \beta_7 \ln(H_i) + \beta_8 \ln(G_i) + \beta_9 DU_i + \varepsilon_i$$

Sem embargos, deve-se levar em conta na elaboração do modelo à existência de externalidades espaciais no fornecimento de serviços públicos, pois entende-se que as autoridades locais interagem umas com as outras, com maior facilidade de comunicação para as comunidades bem próximas às adjacentes.

3.4.2 Análise Espacial

A análise espacial é um tipo de análise exploratória que faz uso de dados georeferenciados para tentar descrever e visualizar padrões espaciais nos dados, visando descrever as distribuições espaciais e identificar, principalmente, localidades atípicas (*outliers* espaciais), agrupamentos de observações semelhantes (*clusters*) e formas de heterogeneidade espacial (Anselin, 1988). A estrutura da distribuição é incorporada através de uma matriz de ponderação espacial, W, com elementos w_{ij}, onde o índice *ij* corresponde ao vizinho *i* da observação *j*. A presença de zeros na matriz de pesos indica a ausência de interação espacial entre as observações.

Usualmente, a presença de autocorrelação espacial nos dados é medida por estatísticas globais como *I de Moran* (*Moran's I)* e *C de Geary* (*Geary's C*), sendo o primeiro o mais utilizado. Conforme Anselin (1995), o Índice Global de *Moran* (I), é o indicador de autocorrelação espacial global, que permite estimar o quanto o valor observado de um atributo numa área é correlacionado com os valores desta mesma (técnica univariada) ou de outras variáveis (técnica multivariada), nas localizações vizinhas, fornecendo um único valor como medida de associação espacial para todo o conjunto de dados. No referido trabalho, o primeiro passo da análise espacial do efeito zoo, parte da técnica multivariada[22], para detecção da existência de autocorrelação espacial global entre a gama de serviços públicos providos e o tamanho da população. Para tal, calcula-se a estatística *I de Moran* para estas duas variáveis padronizadas, z_1 e z_2, da seguinte forma:

$$I^{z_1 z_2} = \frac{n}{S_0} \frac{z_1 W z_2}{z_1 z_1} \tag{3.9}$$

em que *n* é o número de observações, *z* representa os valores das variáveis de interesse, e S_0 é igual à operação $\sum\sum w_{ij}$, significando que todos os elementos da matriz de pesos espaciais *W* devem ser somados.

Os valores de *I de Moran* normalmente variam entre 1 e -1, e sua interpretação é expressa por três possibilidades:

- Se o *I de Moran* for próximo de 1 (um) positivamente, significa que existe uma correlação espacial positiva, com os valores das áreas vizinhas similares entre si. Quanto mais perto de 1, maior a autocorrelação espacial, sendo o valor 1 atribuído a uma autocorrelação positiva (direta) perfeita;

[22] Para detalhes, ver Anselin, Syabri e Smirnov (2003).

- Se o *I de Moran* for próximo de um negativamente, ocorre uma não similaridade dos valores das áreas vizinhas. Quanto mais perto de -1, maior é a autocorrelação espacial, sendo o valor -1 atribuído a uma autocorrelação negativa (inversa) perfeita;
- Se o *I de Moran* for zero ou aproximadamente zero não há evidências de autocorrelação espacial.

O índice considera duas hipóteses: a) a primeira que corresponde à hipótese nula (H_0) admite independência espacial; b) a segunda hipótese é a alternativa (H_1) admite a existência de dependência espacial.

Não obstante, esse índice ignora a existência de padrões locais de dependência, o que poderia mascarar sobre a existência de autocorrelação espacial nos dados.

Desta forma, em análises desse tipo, faz-se também necessário à utilização das estatísticas LISA[23] que são indicadores locais de associação espacial. Da mesma forma que se pôde conseguir um parâmetro de autocorrelação espacial global no contexto bivariado, é possível também conseguir uma medida de autocorrelação espacial local multivariada. Nesse caso, a LISA expressa o grau de associação linear (positiva ou negativa) entre o valor para uma variável em uma dada localização i e a média de uma outra variável nas locações vizinhas (ANSELIN, SYABRI E SMIRNOV, 2003). Então:

$$I_i^{Z_1 Z_2} = Z_{1i} W Z_{2i} \tag{3.10}$$

Com o *I de Moran* local é possível decompor o coeficiente de autocorrelação bivariada em quatro categorias: A autocorrelação espacial positiva expressa valores similares (altos ou baixos) à média dos vizinhos para uma determinada observação, evidenciando a presença de *clusters* de altos (alto-alto) e/ou baixos (baixo-baixo) valores para as variáveis. Por outro lado, a autocorrelação espacial negativa indica valores significativamente mais altos (ou baixos) para a observação quando comparados à média dos vizinhos (alto-baixo ou baixo-alto), evidenciado a presença de observações atípicas, chamadas *outliers* espaciais.

Econometricamente, a autocorrelação espacial ocorre quando os valores da variável dependente ou da (s) independente (s), do termo de erro, ou até mesmo, simultaneamente, em mais de um tipo de variável em um local estão correlacionados com os valores correspondentes nas observações de localidades próximas, e assim sendo, a literatura busca implantar nas análises de regressão linear a influência dessas características.

[23] *Local Indicator of Spatial Association*

Tradicionalmente, em modelos de regressão múltipla uniequacional, as variáveis explicativas utilizadas devem explicar o comportamento de uma dada variável dependente. Contudo, constatada a existência de autocorrelação espacial, a não inclusão de uma variável que capture esse efeito proporciona estimativas inconsistentes ou ineficientes. A dificuldade na realização de modelagens espaciais é a identificação do tipo de transbordamento presente e qual o modelo mais apropriado para a captura desse processo.

3.4.3 Modelos Econométricos Espaciais

Na literatura, existem várias ferramentas espaciais globais, locais e mistos, que especificam modelos apropriados para cada caso. Com respeito à realização deste trabalho, sob a justificativa de que o alcance do transbordamento reflete para todas as regiões da área em estudo, e não apenas para seus vizinhos diretos e os vizinhos de segunda ordem, os modelos considerados para a pesquisa são: Modelo de Defasagem Espacial (ou SAR – *Spatial Auto Regressive*), o Modelo de Erro Autorregressivo Espacial, ou Modelo de Erro Espacial (SEM – *Spatial Error Model*).

A utilização do SAR parte da predição que a autocorrelação espacial pode ser captada por meio da estimativa de um coeficiente relacionado com a variável dependente defasada espacialmente, expresso na forma:

$$Y = \rho WY + X\beta + u \tag{3.11}$$

em que Y é um vetor (*nx1*) de observações da variável dependente; β é um vetor de coeficientes a serem estimados; X é uma matriz com (*nxk*) variáveis explicativas; e u é um vetor de erros. ρ é um coeficiente autorregressivo espacial, o qual capta os efeitos da interação entre um município com seus vizinhos; *W (nxn)* é a matriz de pesos espaciais, sendo o termo *WY* o vetor de *lags* espaciais. Desta forma, considerando o SAR, a Equação (3.8) pode ser estimada por:

$$\ln(R_i) = \beta_0 + \rho W \ln(R_{-i}) + \beta_1 \ln(N_i) + \beta_2 DU_i \ln(N_i) + \beta_3 \ln(A_i) + \beta_4 \ln(Des_i) \tag{3.12}$$

$$+ \beta_5 \ln(Pop18_i) + \beta_6 \ln(Pop65_i) + \beta_7 \ln(H_i) + \beta_8 \ln(G_i) + \beta_9 DU_i + \varepsilon_i$$

Por outro lado, a utilização do SEM ocorre quando as variáveis omitidas, e incluídas nos termos de erro são autocorrelacionadas espacialmente. Neste caso, os resíduos da equação possuem uma estrutura autorregressiva, e diferentemente do SAR não apresentam a variável resposta como uma função direta dos seus *lags* espaciais. Desta feita, pelo SEM, estima-se a Equação (3.8) pela forma:

$$\ln(R_i) = \beta_0 + \beta_1 \ln(N_i) + \beta_2 DU_i \ln(N_i) + \beta_3 \ln(A_i) + \beta_4 \ln(Des_i) + \qquad (3.13)$$

$$\beta_5 \ln(Pop18_i) + \beta_6 \ln(Pop65_i) + \beta_7 \ln(H_i) + \beta_9 DU_i + \varepsilon_i$$

Sendo que neste caso, $\varepsilon_i = \lambda W_\varepsilon + v_i$, W é a mesma matriz da Equação (3.11) e v_i é o vetor de erros bem-comportados. A dependência do erro neste caso ocorre provavelmente quando ε inclui variáveis omitidas correlacionadas espacialmente, como as políticas públicas das localidades, ou quando há choques comuns não observáveis que afetam os municípios vizinhos, como impactos relacionados aos custos de prestação de bens públicos locais (REVELLI, 2005).

Anselin (1999) expõe que conforme seja o processo de autocorrelação espacial, as consequências de ignorá-la nos modelos de regressão dependem da hipótese alternativa utilizada. Se a hipótese alternativa for o modelo de defasagem espacial, o estimador de MQO será viesado e inconsistente. Por outro lado, caso a hipótese alternativa seja o modelo de erro espacial, o estimador dos mínimos quadrados ordinários é não viesado, contudo ineficiente. A solução então é a técnica de estimação utilizando Variáveis Instrumentais (VI) ou por Máxima Merossimilhança (MV).

Para a especificação e determinação do modelo que melhor se ajusta aos dados de uma pesquisa, existem alguns testes que têm papel essencial em econometria espacial. O mais conhecido é teste o "*I de Moran*" para verificação de autocorrelação espacial nos resíduos, previamente discutido. Esse índice, contudo, apresenta baixo poder explicativo na ausência de normalidade dos erros. Outros testes bastante utilizados são: o teste difuso de Kelejian & Robinson, os testes focados Multiplicador de Lagrange para os erros (ML_{err}), para o coeficiente da variável defasada (ML_{lag}) e para a presença de autocorrelação espacial simultânea nos erros e na variável dependente, ou seja, uma especificação SARMA, além dos testes robustos à presença de outras especificações, com maior complexidade computacional (ALMEIDA, 2012).

O procedimento adotado nesta pesquisa segue todas essas etapas expostas acima, e os resultados do trabalho estão expostos na seção seguinte.

3.5 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Primeiramente, através do *I de Moran* Global Bivariado, buscou-se evidências estatísticas da presença de autocorrelação espacial entre a gama de serviços públicos e o contingente populacional numa dada região. Como era esperado, o resultado positivo e significativo, Tabela 3.4, permite rejeitar a hipótese nula de aleatoriedade espacial do fornecimento de bens públicos entre municípios consorciados, e sugere a associação de circunvizinhos no intuito de atingirem certo contingente populacional e receberem um maior número de competências, ou programas governamentais.

Tabela 3.4: Teste de *I* de *Moran Global* Bivariado

Indicador	Valor do índice	p-valor
Gama de serviços públicos providos (R)/população	**0,6281**	0,001

Fonte: Elaboração própria a partir de dados da pesquisa.

Entretanto, como bem expõe Anselin (1995), o indicador *I de Moran Global* pode ser uma medida insuficiente na identificação de padrões locais espaciais, como os *clusters* ou *outliers* locais significantes. Sendo assim, para identificar a existência de *clusters* ou *outliers* localmente significativos, foi utilizado o indicador *I de Moran Local*, ou LISA, que testa a hipótese de inexistência de associação espacial, comparando os valores de cada localização com os valores de seus vizinhos, possibilitando identificar a existência de aglomerações, sujeita ao nível de significância estatística utilizada no teste de autocorrelação espacial.

Dentro desta perspectiva, foi elaborado o **Moran Map -** Figura 3.1 - como forma de visualização dos municípios para os quais os valores do indicador LISA foram considerados significantes.

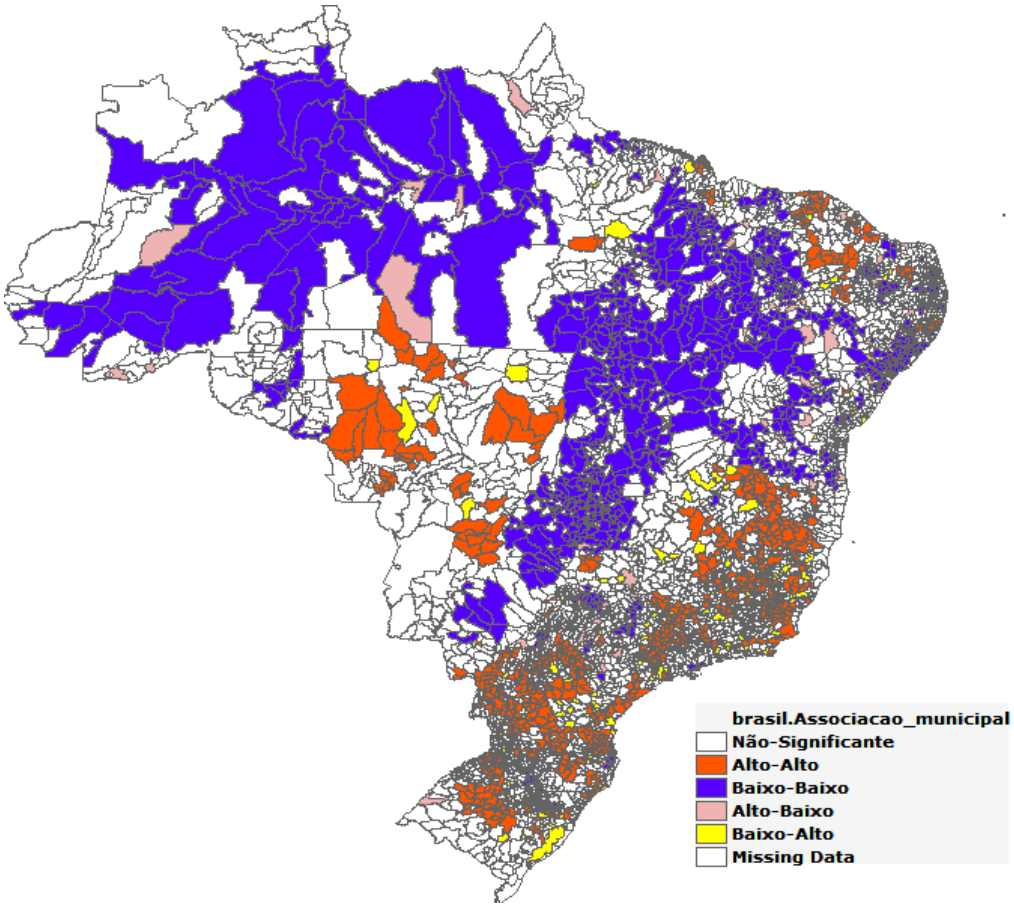
Refere-se a um mapa temático que visa classificar as áreas em quatro possibilidades de associação espacial, sendo **Alto-Alto (AA)**, encontradas no mapa com a cor vermelha representando correlação espacial positiva, apresentam municípios com maior número de competências públicas, associados a localidades com maior contingente populacional.

Baixo-Baixo (BB), também representando tipos de correlação positiva, encontram-se no mapa na cor azul, e identifica municípios com baixos índices de serviços circundados por

municípios também com pequeno número de residentes; **Alto-Baixo (AB),** encontradas no mapa na cor rosa, representando tipos de correlação espacial negativa, denominados *outliers*, identifica municípios com alta gama de serviços públicos rodeados por municípios com baixo contingente populacional; e **Baixo-Alto (BA),** em cor amarela no mapa, também representam correlação espacial negativa (*outliers*) e identificam municípios com pequeno número de competências governamentais vizinhos a municípios com grande população.

Pela Figura 3.3, é possível identificar que os *clusters* do tipo Alto-Alto, localizam-se na grande maioria nas regiões Sul e Sudeste do país. Com efeito, pode-se considerar esta a primeira evidência do efeito zoo, ao ponto em que a associação dessas localidades, em sua maioria, com grande contingente populacional se dá pela maior demanda desses consórcios por uma gama de serviços mais complexos e com maior diversidade, os quais seriam difíceis ou impossíveis de serem alcançados se essas jurisdições estivessem fragmentadas.

Figura 3.3 – Mapa de *Clusters* e *Outliers*



Fonte: Elaboração própria com base nos dados da Pesquisa

Já nas regiões Norte e Nordeste do país encontram-se as coordenações municipais do tipo Baixo-Baixo, o que realmente explicita a carência de serviços públicos nessas localidades, e justifica a associação entre elas no sentido de receberem melhor e maior nível de programas governamentais condizentes com a população e com características socioeconômicas dos consórcios. Também são nessas regiões que podem ser visualizadas os poucos municípios considerados *outliers* Alto-Baixo, enquanto que os do tipo Baixo-Alto estão dispersos em sua maioria no Sudeste-Sul. É importante destacar que a distribuição espacial das associações Alto-Alto e Baixo-Baixo indica haver um efeito transbordamento entre municípios de estados vizinhos, uma vez que alguns *clusters* são formados por municípios pertencentes a outras unidades federativas.

A sequência da estratégia metodológica se deu por meio da estimação da Equação (3.8) usando MQO (resultados apresentados na Tabela 3.6, colunas 1 e 2), e a partir dos resíduos da regressão foram feitos todos os diagnósticos[24], e os testes do *I de Moran* e Multiplicador de Lagrange (ML) para dependência espacial, semelhante ao procedimento de Frère, Hammadou e Paty (2011). O *I de Moran* confirmou a hipótese geral de presença de autocorrelação espacial do modelo, o que sugere que a inferência baseada em uma especificação não espacial é inválida para os dados da provisão de serviços públicos, e a comparação dos níveis de significância dos testes ML_{err} e ML_{lag}, e suas versões robustas RML_{err} e RML_{lag} indicou como essa dependência toma forma, e qual modelo tem especificação mais confiável[25]. Os resultados estão expostos na Tabela 3.5.

Tabela 3.5 – Testes I de Moran e Multiplicador de Lagrange (*ML*)

	I de Moran	ML_{err}	ML_{lag}	RML_{err}	RML_{lag}
(1)	0,3851	168,41	147,21	72,14	42,16
	(0,0001)	(0,0000)	(0,0000)	(0,0086)	(0,0000)
(2)	0,4233	176,13	157,23	79,22	45,11
	(0,0001)	(0,0000)	(0,0000)	(0,0091)	(0,0000)

Fonte: Elaboração própria com base nos dados da pesquisa.
Obs: (i) Os valores entre parênteses representam a probabilidade da estatística. (ii) Os resultados da linha (1) correspondem a estimação da coluna 1, e da linha (2) refere-se a estimação da coluna 2, da Tabela 2.5. (iii) Foram realizados testes com as matrizes de contiguidade e de proximidade geográfica e a que gerou melhores estatísticas foi a matriz com base no decaimento da distância euclidiana entre as regiões consorciadas.

[24] O Teste de White rejeita a hipótese de homocedasticidade, ao passo em que o teste *condicional number* não indica problema de multicolinearidade, e o Teste Jarque-Bera mostra que os erros não são considerados normais.
[25] A regra de decisão é simples: Sendo significantes as inferências das estatísticas ML_{err} e ML_{lag}, faz-se uso de suas versões robustas. Se RML_{err} for mais significante que RML_{lag} estima-se o modelo SEM, inversamente, se RML_{err} for menos significante que RML_{lag}, estima-se o modelo SAR. (ANSELIN *et al*, 1996).

Aplicando-se a regra de decisão, detectou-se a presença de dependência no coeficiente relacionado à variável explicativa defasada espacialmente, o que consequentemente possibilitou a estimação do modelo SAR por Variáveis Instrumentais (VI)[26] - já que o método não prescinde do requisito da propriedade da normalidade do erro aleatório - no sentido de gerar coeficientes consistentes e não viesados. A Tabela 3.6 apresenta os resultados da estimação por MQO (colunas 1 e 2) e considerando a autocorrelação espacial na variável defasada espacialmente (coluna 3 para a análise sem a utilização da interação *dummy* urbana*população, e coluna 4 considerando tal interação), como foi sugerido pelos testes realizados acima.

Tabela 3.6: Resultados das estimações por modelos: MQO e SAR por VI

Variáveis Explicativas	**MQO (1)**	**MQO (2)**	**SAR (3)**	**SAR (4)**
População	0,2211*	0,1431*	0,2046*	0,1522*
	(0,0000)	(0,0000)	(0,0000)	(0,0000)
População**DU*	-	0,0818*	-	0,0897*
	-	(0,0000)	-	(0,0000)
População -18	- 0,0157***	- 0,0256**	- 0,0153**	- 0,0211**
	(0,0078)	(0,0122)	(0,0072)	0,0104
População +65	0,0561***	0,0573***	0,0621***	0,0422**
	(0,0281)	(0,0287)	(0,0312)	(0,0203)
Ìndice *Herfindhal*	- 0,0212***	- 0,0167***	- 0,0173***	- 0,0181**
	(0,0106)	(0,0099)	(0,0086)	(0,0084)
Taxa de Desemprego	- 0,0159	- 0,0083	- 0,0106	- 0,0093
	(0,0978)	(0,1051)	(0,0123)	(0,0813)
Área	0,0143**	0,0171**	0,0154**	0,0163**
	(0,063)	(0,0082)	(0,0065)	(0,0071)
Coeficiente de Gini	0,0368**	0,0312***	0,0431***	0,0356**
	(0,0175)	(0,0157)	(0,0021)	(0,0151)
Dummy Urbana	- 0,0073	- 0,6711*	- 0,0088*	- 0,6033*
	(0,0038)	(0,0000)	(0,0031)	(0,0000)
Constante	1,9131*	2,1131*	1,8783*	2,2704*
	(0,0000)	(0,0000)	(0,0000)	(0,0000)
ρ	-	-	0,5833*	0,6481*
			(0,0000)	(0,0000)
Estatísticas				
Coef. Determinação - R^2	0,2676	0,3154	0,3832	0,4009
Multicolinearidade	87,21	79,98	-	-
Teste de White	461,28*	448,21*	-	-
Teste Jarque-Bera	113,64	97,54	-	-
Log likelihood	-	-	- 202,33*	-189,54*

Fonte: Elaboração própria a partir dos resultados da pesquisa. **Obs:** (i) Todas as variáveis estão em logaritmo natural; (ii) Erro padrão entre parênteses. **Notas:** * significante ao nível de 1%.** significante ao nível de 5%.*** significante ao nível de 10%.

[26] Variáveis explicativas defasadas espacialmente foram utilizadas como instrumentos.

Ressalta-se, pelo maior coeficiente de determinação (Pseudo R^2), e maior valor para o critério *Lik* que a análise com a interação da *dummy* urbana introduzida no modelo apresenta melhor poder explicativo das variáveis explicativas sobre a dependente, assim as análises dos resultados serão reportadas com base nos coeficientes desse modelo[27].

Como se esperava, a maioria das variáveis exógenas ao modelo é estatisticamente significante (com exceção da taxa de desemprego), ou seja, as características socioeconômicas são relevantes para a determinação do número de serviços prestados nos consórcios.

O parâmetro espacial ρ positivo indica uma autocorrelação espacial global positiva, e traduz que a maior (menor) diversificação na gama de serviços públicos nas regiões vizinhas expande (retrai) a oferta de serviços em áreas adjacentes. Isto fundamenta a formação de associações municipais em busca de um número maior, ou melhor, nível de competências públicas para seus habitantes, e de certo modo esboça o efeito zoo.

Os transbordamentos entre municípios coordenados levam em conta os ajustes das suas escolhas com relação às competências exercidas frente a seus vizinhos, com base em duas circunstâncias: A complementaridade ou continuidade de serviços. Se houver efeitos colaterais substanciais, municípios que prestam serviços públicos que ainda não foram fornecidos pelos vizinhos, e vice-versa, associam-se a fim de diversificar a gama de serviços locais capazes de atender todos os seus cidadãos. Porém, se não existem efeitos colaterais substanciais, as intermunicipalidades fornecem serviços públicos que já são ofertados pelos vizinhos, a fim de assegurar a continuidade dos serviços prestados localmente, no intuito de atender a um maior contingente possível. De uma forma ou de outra, busca-se equilibrar as forças de demanda e oferta deste serviço, tornando-o menos oneroso e proporcionando a melhor focalização possível. Em outras palavras, a provisão pública torna-se menos ineficiente.

Considerando a questão principal da pesquisa para confirmação do efeito zoo, que é a relação entre a gama de serviços públicos locais e tamanho da população, o coeficiente referente à variável população é positivo e altamente significativo, sugerindo que a série de bens públicos providos pelas associações intermunicipais é uma função crescente do tamanho de sua população. Portanto, a reunião de cidadãos de vários municípios vizinhos em consórcios pode promover uma expansão dos serviços públicos locais disponíveis, mediante o fornecimento de novos bens públicos indivisíveis.

---

[27] Salienta-se que o modelo estimado não deve apresentar autocorrelação espacial em seus resíduos. Se isso ocorre, o modelo não captura todo o processo de dependência espacial e é inadequado. Desta forma, realizada tal avaliação, a estatística do teste de validade do modelo não indicou dependência espacial para os resíduos do modelo SAR.

O coeficiente relativo à interação *dummy* urbana*população, gerado no sentido de captar o impacto do efeito no gradiente urbano-rural, também apresenta sinal esperado, positivo e significante, e sugere que o fenômeno é maior em áreas bastante urbanizadas, devido às economias de escala mais substanciais nessas localidades, e pelo fato de que as coordenações pouco urbanizadas e/ou tipicamente rurais reúnem pequenos municípios que preferem uma gestão municipal, alinhada com as preferências de seus cidadãos (JOSSELIN, GUENGANT e ROCABOY, 2008).

A magnitude do efeito zoo nas áreas urbanas pode ser visualizada a partir do resultado positivo e significante (1,6671) do valor resultante da soma do coeficiente representativo da *dummy* urbana com o valor do intercepto (BRAMBOR, CLARK e GOLDER, 2005). Resumidamente, a variedade dos serviços providos por coordenações intermunicipais mais desenvolvidas excede a das menos desenvolvidas, ou seja, confirma-se a hipótese da existência do efeito zoo nas grandes municipalidades brasileiras.

Com relação às variáveis de controle do modelo, os coeficientes referentes a munícipes com menos de 18 anos e mais de 65 anos são significantes, negativo e positivo, respectivamente, denotando que quanto maior for à faixa etária dos cidadãos dentro das coordenações municipais, maior e mais diversificada será a gama de serviços locais.

Quanto ao Ìndice *Herfindhal*, o valor do parâmetro indica um grande número de pequenas cidades densamente povoadas, e assim sendo, quanto menor o município, menor é a gama de serviços públicos prestados e maiores são as responsabilidades transferidas para um grupo consorciado (LEPRINCE e GUENGANT, 2002). Com relação à área, o impacto positivo sobre o número de serviços prestados pelas comunidades, expõe os ganhos de economias de escala dos municípios, os quais se tornam propensos a manter o poder de decisão sobre gastos e não transferir alguns serviços para os consórcios.

Com respeito ao Índice de Gini, o coeficiente positivo expõe que os quadros de grande heterogeneidade entre os municípios fornecem incentivos para a provisão ou ampliação de serviços públicos em nível de município.

Em suma, os resultados apresentados corroboram as predições teóricas de Oates (1988) de acordo com as quais a gama de serviços públicos em comunidades maiores é mais ampla e diversificada, o que confirma empiricamente a existência do efeito zoo nos municípios brasileiros. As implicações destes resultados são muito importantes, pois como consequência do efeito, a coordenação entre jurisdições seria uma boa maneira de aumentar a variedade de serviços públicos locais. Assim, uma vez que a estrutura governamental entre as associações municipais são maiores do que seus municípios membros, individualmente, elas

são capazes de fornecer conjuntamente serviços adicionais que cada município seria incapaz de fornecer por conta própria. Ao mesmo tempo, destaca-se que a intensidade do efeito zoo depende do gradiente urbano-rural, pois é menos intenso em áreas menos urbanizadas, sugerindo que as economias mais substanciais de escala podem existir em áreas urbanas.

Por fim, em termos de política pública, os resultados confirmam a relevância dos incentivos do governo para desenvolver a cooperação entre os municípios locais em áreas mais densamente povoadas, devido à internalização de *spillovers*.

3.6 CONSIDERAÇÕES FINAIS

Uma vez que muitos bens públicos são indivisíveis, a gama de serviços governamentais devem aumentar proporcionalmente à dimensão das jurisdições, o que está em linha com um fenômeno descoberto e definido por Oates (1988) como efeito zoo. Usando um conjunto de dados de localidades brasileiras que formam Consórcios Públicos Intermunicipais de serviços públicos, ou seja, governos locais que reúnem vários municípios vizinhos, a fim de gerenciar coletivamente alguns produtos locais, este artigo visa estabelecer a evidência empírica para esse fenômeno no Brasil.

Através da estatística I de *Moran* Global Bivariado confirmou-se a hipótese de correlação espacial do modelo, ao passo que o *LISA* apontou que os municípios com maior número de competências públicas estão espacialmente correlacionados aos municípios com maior contingente populacional, formando *clusters* que se localizam em sua maioria nas regiões Sul e Sudeste do Brasil, e que os municípios com menor número de competências estão associados às jurisdições com menor número de consumidores do serviço público.

Para análise dos determinantes da gama de serviços públicos entre os municípios consorciados, o modelo *Spatial Auto Regressive,* ou Modelo de Defasagem Espacial foi estimado por Variáveis Instrumentais, e os resultados da estimativa favorecem a ideia proposta por Oates (1988): O parâmetro de defasagem espacial (ρ) positivo e significante indica que a maior diversificação na gama de serviços públicos nas regiões vizinhas expande a oferta de serviços em áreas vizinhas, justificando a formação de interação entre os municípios como forma de se obter melhor e maior nível de serviços públicas para seus cidadãos.

O parâmetro da população mostra-se diretamente correlacionado ao número de competências providas pelas municipalidades, sugerindo que a série de bens públicos providos pelas associações é uma função crescente do tamanho de sua população, e que o

agrupamento de consumidores de vários municípios vizinhos pode promover uma expansão dos bens existentes e a provisão de novos bens públicos indivisíveis, o que comprova empiricamente a existência do efeito zoo nos municípios brasileiros. A análise da interação *dummy* urbana e população evidência que a intensidade do efeito depende do gradiente urbano-rural, sendo o fenômeno maior em áreas urbanizadas, na magnitude de 1,6671, devido às economias de escala mais substanciais nessas localidades.

Através dos parâmetros referentes à população nova e idosa, infere-se que localidades com maior faixa etária provêm uma gama de competências governamentais mais diversificadas, ao passo que as características socioeconômicas justificam a provisão dos serviços prestados pelos consórcios, o que pode supostamente acontecer em decorrência do grau de integração dos municípios ser mais sensível as preferências dos tomadores de decisão no âmbito dos consórcios, do que as características da jurisdição.

Conclui-se então que os resultados apresentados corroboram com as predições teóricas de Oates (1988), de que a gama de serviços públicos em comunidades maiores são mais amplas e diversificadas do que nas menores, ou seja, existe o efeito zoo nos municípios brasileiros. Estes resultados sugerem que políticas públicas de incentivo a coordenação municipal devem ocorrer, sobretudo em áreas desenvolvidas e densamente povoadas no sentido de proporcionar bens públicos que atinjam toda a população demandante, fornecidas com o menor custo para o Estado. Em comparação com a gestão municipal, de forma fragmentada, esse fenômeno ainda poderia ser incentivado pelas economias de escala e maiores receitas fiscais através de um aumento das taxas dos impostos locais, consequência de uma menos intensa concorrência fiscal local (CHARLOT, PATY e PIGUET, 2009).

Este foi o primeiro passo da análise do fenômeno no Brasil. A extensão da pesquisa terá como objetivo obter formas para estimar o grau de congestionamento dos bens públicos locais, tendo em conta o efeito zoo, partindo também da visualização de consórcios específicos de serviços, levando em consideração suas características fiscais correspondentes.

CONCLUSÃO

Este trabalho estudou, em três capítulos, os fatores associados à provisão de bens nos municípios brasileiros evidenciando as idiossincrasias que determinam a sua demanda e a oferta, no sentido de contribuir para uma discussão a respeito da organização do setor público, com fornecimento de serviços governamentais menos custosos ao erário, melhor adequação da provisão frente aos desejos dos cidadãos espacialmente heterogêneos, e por fim, a

possibilidade de incentivos do governo para desenvolvimento da coordenação municipal em zonas mais densamente povoadas.

Buscou-se abordar temáticas ainda não exploradas para o caso brasileiro. A motivação para o primeiro artigo foi a observação de que a literatura nacional tem focado primordialmente na análise de como o padrão de demanda local pode explicar o nível da despesa pública, considerando uma relativa rigidez no lado da oferta de serviços públicos, como se as instituições públicas, e os gestores locais eleitos fossem agentes exógenos no processo. Dessa forma, baseado nas premissas do federalismo fiscal, confrontou-se duas especificações alternativas que explicam o comportamento das despesas públicas locais, uma do lado da demanda e outra do lado da oferta, para um conjunto de dados de 5285 municípios brasileiros para 2010.

A metodologia econométrica utilizada nesta etapa do trabalho foi a combinação do Teste J, proposto por Davidson e MacKinnon (1981), com o *Threshold*, sugerido por Hansen (2000), como um método de triagem de dados capaz de testar se a resposta dos governos locais aos seus eleitores é estável entre os municípios de pequeno e de grande porte. Especificamente, testou-se a possibilidade de quebra, do conjunto de dados, em regimes e ao mesmo tempo a expectativa dos regimes serem gerados pelo modelo do eleitor mediano e/ou pelo modelo direcionado para a oferta, com possíveis pontos de quebras entre eles.

Os resultados permitem inferir que o modelo do eleitor mediano pode ser aplicado às cidades menores, com tamanho ótimo populacional variando entre 13604 e 13661 habitantes, dependendo da função de congestionamento utilizada. Na formulação B-D, por exemplo, o limiar corresponde a cidades com população local de 13661 habitantes. Até este patamar populacional, são consideradas 56,6% de toda a amostra. Para o modelo *camaraderie*, o limiar corresponde a uma população de 13604 habitantes, e no modelo flexível, o limiar permanece o mesmo, alterando-se apenas o intervalo de confiança, que em termos de população local é [13594; 15103].

De uma forma geral, os resultados, fundamentados adicionalmente por meio das estimações dos parâmetros de cada modelo e do coeficiente de congestionamento dos bens, sugerem que à medida que as cidades crescem, os representantes eleitos se afastam de um compromisso rigoroso com a escolha do eleitor mediano, e passam a prover bens públicos de forma discricionária, no sentido de maximizar o seu interesse.

A partir da validação dos modelos do eleitor mediano e da oferta, o Capítulo 2 analisou a qualidade na provisão dos serviços públicos, via efeito congestionamento dos bens, em 5268 municípios brasileiros, tendo como referencial a Teoria dos Clubes. Neste Capítulo,

os modelos foram ampliados com a inserção de variáveis específicas a cada caso, adaptando elementos teóricos, empíricos e institucionais. Como as análises supõem clubes que recebem novos usuários dos serviços, foram consideradas as diferentes formas funcionais capazes de captar o efeito congestionamento e/ou escala no fornecimento de serviços em resposta ao seu crescimento, além dos efeitos transbordamentos, para captação da interdependência ou assimetria de relações espaciais.

A metodologia econométrica empregou, além do modelo clássico linear, a análise espacial, proposta por Anselin (1988), a qual vem sendo amplamente utilizada na área de estudos regionais. Adicionalmente, para comparação dos padrões, utilizou-se do Teste J (Davidson e Mackinnon, 1981) ao ponto de identificar o seu poder explicativo como determinantes para toda a gama de serviços providos pelos clubes.

Para os dois modelos, os resultados apontam para existência da correlação espacial entre os municípios, sugerindo que os padrões de gastos em determinada localidade influenciam as decisões fiscais de vizinhos. Comprova-se a hipótese de que o tamanho dos clubes e o custo dos serviços públicos diferenciam os resultados da ação de provisão, corroborando a literatura sobre o tema a qual defende que em jurisdições menores a proximidade dos eleitores com seus agentes públicos exerce maior pressão e controle sobre seus atos políticos, e que preferências individuais agregadas apresentam menor heterogeneidade e podem representar a comunidade como um todo. Este resultado gera um efeito congestionamento maior para as maiores cidades, o que vai de encontro à literatura internacional.

Uma avaliação detalhada deste resultado permitiu inferir sobre as indivisibilidades características do fornecimento de certos bens e serviços, as quais limitam a sua provisão aos grandes centros urbanos; mostrou-se, assim, que os maiores gastos refletem não só um custo de congestionamento, mas adicionalmente, o fato de que os maiores clubes provêm um conjunto mais diversificado e complexo de serviços, em comparação aos pequenos clubes. Este fenômeno é conhecido na literatura por “Efeito Zoo”.

Com base nesta observação, o terceiro capítulo procurou estabelecer evidências empíricas para o “efeito-zoo” nos municípios brasileiros usando um conjunto de dados de localidades brasileiras que formam Consórcios Públicos Intermunicipais de serviços públicos. Segundo esse efeito, o nível mais baixo de despesas em cidades menores não deve ser explicado pela menor aglomeração, mas simplesmente pelo fato de que certas categorias de gastos - como um zoológico - precisam de um tamanho mínimo populacional para serem fornecidos.

Para análise dos determinantes da gama de serviços públicos entre os municípios consorciados, estimou-se o modelo o Modelo de Defasagem Espacial (*Spatial Auto Regressive*), por Variáveis Instrumentais. Os resultados das estimativas favorecem a ideia proposta por Oates (1988) de que a maior diversificação na gama de serviços públicos nas regiões vizinhas expande a oferta de serviços em áreas vizinhas, justificando a interação entre os municípios como forma de se obter maiores e melhores níveis de serviços públicos para seus cidadãos. Confirmou-se a hipótese da existência do efeito zoo nos municípios brasileiros e as evidências de que a intensidade do efeito depende do gradiente urbano-rural, sendo o fenômeno mais relevante em áreas urbanizadas, devido às economias de escala mais substanciais nessas localidades.

Entende-se que ainda existe a necessidade de uma série de estudos adicionais para fundamentar e respaldar os resultados aqui estimados. Assim, a extensão natural desta pesquisa envolve a introdução de parâmetros específicos que captem o efeito direto na função demanda e oferta da provisão pública, em cada regime, com o uso de dados em painel, o que permitiria uma avaliação dinâmica na determinação das despesas. Além disso, o uso dos modelos de Regressão Quantílica permitiria a divisão endógena dos clubes no estudo da qualidade dos gastos para fornecimento de bens e serviços à sociedade brasileira. Por fim, pretende-se estimar o grau de congestionamento dos serviços governamentais locais, considerando o efeito zoo, em consórcios específicos de serviços, como saúde e educação.

REFERÊNCIAS

AHMED, S.; GREENE, K.V. Is the median voter a clear-cut winner? Comparing the median voter theory and competing theories in explaining local government spending. **Public Choice**, 105: 207–230, 2000.

ALESINA, A. Macroeconomic Policy in a Two-Party System as a Repeated Game. Quarterly Journal of Economics, 102, 651-78, 1987.

ALMEIDA, E. S. **Econometria Espacial Aplicada.** Campinas: Alínea Editora, 2012.

ANSELIN, L. **Spatial Econometrics:** methods and models**.** Boston: Kluwer Academic. 1988.

_____. **SpaceStat tutorial**. Mimeo, University of Illinois, 1992.

_____. Local indicators of special association – LISA. **Geographical Analysis**, v. 27, n. 2, p. 93-115, 1995.

_____. Interactive techniques and exploratory spatial data analysis. In: LONGLEY, P. A. et al. (Ed). **Geographic information system**: principles, techniques, management and applications. Nova York: Jonh Wiley, p.251-264, 1999.

_____; BERA, A.K.; FLORAX, R.; YOON, M.J. Simple diagnostic tests for spatial dependence, **Regional Science and Urban Economics**, Vol.26, n.l, p.77-104,1996.

_____; SYABRI, I.; SMINORV, O. **Visualizing multivariate spatial correlation with dynamically linked windows**. University of Illinois, 2003, (Mimeo).

ARAUJO, J. M.; SIQUEIRA, R. B. Demanda por gastos públicos locais: Evidências dos efeitos de ilusão fiscal no Brasil. 42º Encontro Nacional de Economia, **Anais**...42º Anpec, Natal-RN, 2014.

BAUDRY, M.; LEPRINCE, M.; MOREAU, C. Préférences révélées, bien public local et electeur median: tests sur données françaises. **Économie et Prévision**, v. 156, p. 125-145, 2002.

BAUMONT, C. **Spatial Effects in housing price models**: do house prices capitalize urban development policies in the agglomeration of Dijon (1999)? Mimeo., Université de Bourgogne, 2004.

BERGSTROM, T.C; GOODMAN, R.P. Private Demands for Public Goods. **The American Economic Review**, 63(3), 280-296, 1973.

BESLEY, T.; CASE, A. Incumbent behavior: vote seeking, tax setting and yardstick competition. **The American Economic Review**, 85(1): 25– 45, 1995.

BORCHERDING, T.E.; DEACON, R.T. The Demand for the Services of Non-Federal Governments. **The American Economic Review**, 62(5), 891-901, 1972.

BRAMBOR, T.; CLARK, W.R.; GOLDER, M. Understanding interaction models: Improving empirical analyses", **Political Analysis**, Vol. 14, pp.63-82, 2005.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF, Senado, 2014.

BREUNIG, R.; ROCABOY, Y. Per-capita expenditures and population size: a non-parametric analysis using French data. Public Choice, 136: 429-445, 2008.

BROOKS, C. **Introductory econometrics for finance.** 1st ed. Cambridge: Cambridge University Press, 2002.

BUCHANAN, J. An economic theory of clubs. **Economica**, 33: 1–14, 1965.

BUETTNER, T.; FABRITZ, N. Responses to grants and the marginal cost of public funds. **Mimeo,** 2011.

CHARLOT S.; PATY, S.; PIGUET, V. The Effects of Fiscal Cooperation on Local Taxation: The French Case. **Working paper**. 2005.

COSTA, F. R.; **O impacto das transferências constitucionais sobre o comportamento fiscal dos municípios brasileiros.** Tese (Doutorado em Economia) - Universidade Federal do Ceará – UFC, Fortaleza, 2013.

CRAIG, S.G. The impact of congestion on local public goods production. **Journal of Public Economics**, 32: 331-353, 1987.

DAHLBY, B.; FEREDE, E. The Stimulative Effects of Intergovernmental Grants and the Marginal Cost of Public Funds. **International Tax and Public Finance**, Vol. 19 (6), 844-883, 2012.

DAVIDSON, R.; MACKINNON, J.G. Several tests for model specification in the presence of alternative hypotheses. **Econometrica,** 49: 551–565, 1981.

_____; _____. **Estimation and inference in econometrics**. Oxford University Press, 1993.

DENZAU A.; MACKAY, R. Gate keeping and monopoly power of committes: an analysis of sincere and sophisticated behavior. **American Journal of Political Science**, 27:740-761, 1983.

DOI, T. **New evidence on the median voter in Japan**. Tokio: Institute of Social Science, 1998.

DYE, R.F.; MCGUIRE, T.J. The effect of property tax limitation measures on local government fiscal behaviour. **Journal of Public Economics,** 66: 469–487, 1997.

EDWARDS, J.Y.H. Congestion function specification and the 'publicness' of local public goods. **Journal of Urban Economics**, 27: 80–96, 1990.

FRÈRE, Q.; HAMMADOU, H.; PATY, S. Is there a "zoo effect" in French Local Governments? **Working Paper Cesaer**. 2011.

GLAZER, A.; NISKANEN, E. Why voters may prefer congested public clubs. **Journal of Urban Economics,** 65: 37-44, 1997.

GUENGUANT, A, JOSSELIN, J. M, ROCABOY, Y. Effects of club size in the provision of public goods. Network and congestion effects in the case of the French municipalities. **Papers in Regional Science,** 81: 443–460, 2002.

HANSEN, B.E. Inference when a nuisance parameter is not identified under the Null hypothesis. **Econometrica**, 64:413-430, 1996.

_____. Sample splitting and threshold estimation. **Econometrica**, 68: 575–603, 2000.

HAYES, K.; SLOTTJE, D.J. Measures of publicness based on demographic scaling. **Review of Economics and Statistics,** 69: 713–718, 1987.

HELPMAN, E.; HILLMAN, A. L. Two remarks on optimal club size. **Economica** *44*, 293–296. 1977.

HIBBS, D. Political Parties and Macroeconomic Policy. **American Political Science Review**, Chicago, vol. 71, p. 1467-1487, 1977.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTCA - (IBGE). Disponível em: <http://www.ibge.gov.br/home/estatistica/>. Acesso em: 21 set 2014.

JOHANSEN, L. Some notes on the Lindhal theory of determination of public expenditures. **International economics review**, 4, 3, p 346-358, 1963.

JOSSELIN, J.M.; ROCABOY, Y.; TAVERA, C. The Influence of Population Size on the Relevance of Demand or Supply Models for Local Public Goods: Evidence from France. **Papers in Regional Science**, 88 (3), 563-574, 2008.

KNIGHT, F. H. Some Fallacies in the Interpretation of Social Cost. **The Quarterly Journal of Economics**, vol. 38, Nº 4, 582-606. Aug. 1924.

LEPRINCE, M.; GUENGANT, A. Interactions fiscales verticales et reaction des communes a la cooperation intercommunale, **Revue Economique**, Vol.53, No.3, pp.525-535, 2002.

LINDAHL, E. Positive Losung, die Gerechtigkeit der Besteuerung, Lund, reprinted as "Just Taxation - A Positive Solution" in R.A. Musgrave and A.T. Peacock (eds.), **Classics in the Theory of Public Finance**, London: Macmillan,1919.

MACMILLAN, M.L.; WILSON, R.W.; ARTHUR, L.M. The publicness of local public goods: evidence from Ontario municipalities. **Canadian Journal of Economics**, v.14, p. 596-608, 1981.

MATSUMOTO, M. A tax competition analysis of congestible public inputs. **Journal of Urban Economics,** 48: 242-259, 2000.

MCGREER, E.; MCMILLAN, M.L. Public output demands from alternative congestion functions. **Journal of Urban Economics**, 33: 95-114, 1993.

MCGUIRE, M. C. Group segregation and optimal jurisdictions. **Journal of Political Economy** *82*, 112–132. 1974.

MEANS, T.S.; MEHAY, S.L. Estimating the publicness of local government services: Alternative congestion function specifications. **Southern Economic Journal**, 61: 614-627, 1995.

MENDES, C.C.; SAMPAIO DE SOUSA, M.C. Estimando a Demanda por Serviços Públicos nos Municípios Brasileiros. **Revista Brasileira de Economia**, v. 60, n.3, p.281-296, 2006.

_____;_____. Demand for Local Public Schooling in Brazilian Municipalities: A Median Voter Approach. **Urban Economics**, 2010.

MISHAN, E. J. The Postwar Literature on Externalities: An Interpretative Essay. **Journal of Economic Literature,** Vol. 9, Nº 1, 1-28. Mar 1971.

NIERO, J. C. C.; **Gestão das Relações Governamentais:** fatores condicionantes da propensão a cooperar por parte dos governos locais. 2011. 215 f. Dissertação (Mestrado) - Curso de Mestrado em Administração, Universidade Municipal de São Caetano do Sul, São Caetano do Sul, 2011.

NORDHAUS, W. The Political Business Cycle. **Review of Economic Studies**, New York, vol. 42, p. 169-190, 1975.

OATES, W. E. **Federalismo fiscal**. Madri: Instituto de Estudios de Administración Local, 1977.

_____, W.E. On the Measurement of Congestionin the Provision of Local Public Goods. **Journal of Urban Economics**, 24 (1), 85-94, 1988.

PAULY, M. V. Clubs, commonality, and the core: An integration of game theory and the theory of public goods. **Economica,** *34*, 314–324. 1967.

PIGOU, A. C. **The Economics of Welfare**. Palgrave Macmillan. London. 1920.

PONCE, C**. Determinantes del gasto público local:** teoria y evidencia empírica para las municipalidades de la província de Córdoba**.** [S.l,;s.n], Mimeo, 1998.

REITER, M.; WEICHENRIEDER, A. Public goods, club goods and the measurement of crowding. **Journal of Urban Economics**, 24: 84-94, 1997.

REVELLI, F. On spatial public finance empirics. **International Tax and Public Finance**, Vol.12, n. 4, pp.475-492, 2005.

ROGOFF, K.; SIBERT, A. **Elections and macroeconomic policy cycles.** Review of economic studies. no 55, páginas 1-16, 1988.

SAMPAIO DE SOUSA, M.C.; CRIBARI NETO, F.; STOSIC, B.D. Explaining DEA technical efficiency scores in an outlier corrected environment: the case of public services in the Brazilian municipalities. **Brazilian Review of Econometrics,** v. 25, n.2, p.289-315, 2005.

_____; PIOLA, M.G.S. Mensurando a Eficiência dos Gastos com Saúde nas Municipalidades. In: SAMPAIO DE SOUSA *et al*. (Orgs) **Economia Pública Brasileira,** Brasília, 2010.

SAMUELSON, P. A. The pure theory of public expenditures. **Review of economics and statistics**, 36, 4, pp. 387-389, 1954.

SANDLER, T. **Global collective action.** New York: Cambridge University Press. 2004.

SANZ, I.; VELÁZQUEZ, F.J. Determinants of the composition of government expenditure by functions. Madrid: European Economy Group, Working paper 13, 2002.

SCHMANDT, H.J.; STEPHENS, G.R. Measuring Municipal Output. **National Tax Journal**, 13 (4), 369-375, 1960.

SECRETARIA DO TESOURO NACIONAL – STN/FINBRA. Disponível em: <http://www.tesouro.fazenda.gov.br/estados_municipios/index.asp>. Acesso em: 23 maio 2014.

SILVA, L.D.C.; SAMPAIO DE SOUSA. A coordenação intermunicipal e as indivisibilidades dos bens públicos locais: uma análise da economia baiana. **Anais.**..XI Encontro de Economia Baiana. Salvador, 2015.

_____;_____. Indivisibilidades na Provisão Local de Serviços Públicos no Brasil: Uma análise das evidências do "efeito Zoo". **Anais...**Encontro da Associação Brasileira de Economia Regional e Urbana (Enaber), 2015.

TIEBOUT, C. M. A pure theory of local expenditures. **Journal of Political Economy,** 64, 416–424. 1956.

TONG, H. On a threshold model. In **Pattern Recognition and Signal Processing** (Edited by C. H. Chen), NATO ASI Series E: Applied Sec. No. 29, Sijthoff & Noordhoff, Amsterdam, 575-586,1978.

TRIBUNAL SUPERIOR ELEITORAL – TSE. Disponível em http://www.tse.jus.br/. Acesso em: 26 abril 2014.

TURNBULL, G.K.; The overspending and flypaper effects of fiscal illusion. **Journal of Urban Economics**, 44: 1-26, 1992.

_____, G.K.; CHANG, C. The median voter according to GARP. **Southern Economic Journal,** 64: 1001-10, 1998.

_____, G.K.; MITIAS, P.M. The median voter model across levels of government. **Public Choice**, 99: 119–138, 1999.

_____ ; GEON, G. Local government internal structure, external constraints and the median voter. **Public Choice,** 129: 487–506, 2006.

WILDASIN, D.E. Demand estimation for public goods: distortionary taxation and other sources of bias. Regional Science and Urban Economics, v. 19, p. 353-79, 1989.

WILSON, J.D. Property taxation, congestion, and local public goods. **Journal of Public Economics,** 64: 207-217, 1997.

WISEMAN, J. The Theory of Public Utility Pricing: An Empty Box, **Oxford Economic Papers**. 1957.

WYCKOFF, P. The simple analytics of slack maximising bureaucracy. **Public Choice,** 67: 35–47, 1990.