

"A FEA e a USP respeitam os direitos autorais deste trabalho. Nós acreditamos que a melhor proteção contra o uso ilegítimo deste texto é a publicação online. Além de preservar o conteúdo motiva-nos oferecer à sociedade o conhecimento produzido no âmbito da universidade pública e dar publicidade ao esforço do pesquisador. Entretanto, caso não seja do interesse do autor manter o documento online, pedimos compreensão em relação à iniciativa e o contato pelo e-mail bjbfea@usp.br para que possamos tomar as providências cabíveis (remoção da tese ou dissertação da BDTD)."

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E
CONTABILIDADE
DEPARTAMENTO DE ECONOMIA

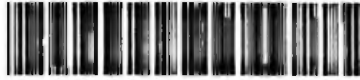
**Como as pessoas votam? Um estudo econométrico dos
determinantes do voto no município de São Paulo**

Sérgio Conti

Orientador: Prof. Dr. Naércio Aquino Menezes Filho

SÃO PAULO, SETEMBRO DE 2004

DEDALUS - Acervo - FEA



20600026845

**Como as pessoas votam? Um estudo econométrico dos
determinantes do voto no município de São Paulo**

Sérgio Conti Macedo

Dissertação apresentada à Faculdade de
Economia, Administração e Contabilidade
da Universidade de São Paulo para
obtenção do título de mestre em Economia

Orientador: Prof. Dr. Naércio Aquino Menezes Filho

SÃO PAULO

2004

FICHA CATALOGRÁFICA

Elaborada pela Seção de Publicações e Divulgação do SBD/FEA/USP

Macedo, Sérgio de Conti

Como as pessoas votam?: um estudo econométrico dos determinantes do voto no município de São Paulo / Sérgio de Conti Macedo. -- São Paulo, 2004.

xxx f.

Dissertação (Mestrado) – Universidade de São Paulo, 2004
Bibliografia.

1. São Paulo - Eleições 2. Eleição municipal 3. Econometria
I. Faculdade de Economia, Administração e Contabilidade da USP II. Título.

CDD – 324.98161

AGRADECIMENTOS

A meu orientador, Prof. Naércio Aquino Menezes Filho, por suas contribuições fundamentais na definição do objeto de estudo da dissertação e por ter me dado os meios para que eu realizasse e concluísse esta longa, mas edificante empreitada.

Aos membros de minha banca de qualificação – Prof. Eduardo Haddad e Prof. Haroldo Torres – pelas importantes recomendações referentes aos campos da econometria espacial e da ciência política e pelas sugestões de extensões do presente estudo.

Aos professores do departamento de economia e de ciência política da Universidade de São Paulo, que me estimularam na escolha de um tema de dissertação que unisse os interesses destas duas importantes áreas do conhecimento.

A Lara Mesquita, que contribuiu decisivamente para a realização deste estudo com suas sugestões, seu tempo, mas, sobretudo, sua paciência comigo. A Solange Kileber, que me ajudou e me acompanhou durante três semestres no desenvolvimento de meus conhecimentos como econometrista.

A todos os meus colegas de trabalho, mas em especial a Denise de Pasqual e Andres Rojas, que colaboraram na conclusão desta dissertação.

Aos financiadores desta empreitada - Conselho Nacional de Pesquisa (CNPQ), Fipe, minha mãe e meus verdadeiros amigos José Guilherme Pantano e Paschoal Chemin Paione - sem os quais a realização deste estudo teria deixado de ser algo gratificante para se tornar uma penitência.

A meus amigos de longa data, Adriano Pitoli, Felipe Mattar e Maurício Fernandes, com quem tenho aprendido e crescido junto nos últimos dez anos.

A minha afilhada Isadora Santos, que é muito pequena para saber o que é uma dissertação de economia e, se tudo correr bem, não seguirá os passos de seu pai e de seu padrinho e fará algo melhor da vida que não Economia.

E principalmente a minha mãe, a quem devo todo o apoio na concretização de mais esta etapa de minha vida.

ÍNDICE

RESUMO	i
ABSTRACT	ii
1. INTRODUÇÃO	8
2. REVISÃO BIBLIOGRÁFICA	11
3. METODOLOGIA	19
3.1. MODELO DE GOODMAN.....	19
3.2. MODELO DE KING.....	20
3.2.1. <i>Descrição Formal do Modelo</i>	24
4. DESCRIÇÃO DOS DADOS	36
4.1. VARIÁVEIS DEMOGRÁFICAS: ELEIÇÃO DE 1996	36
4.2. VARIÁVEIS DEMOGRÁFICAS: ELEIÇÃO DE 2000	40
4.3. VARIÁVEIS ELEITORAIS: ELEIÇÃO DE 1996	49
4.4. VARIÁVEIS ELEITORAIS: ELEIÇÃO DE 2000	51
5. RESULTADOS	53
5.1. ESTIMAÇÃO: ELEIÇÃO DE 1996	53
5.2. ESTIMAÇÃO: ELEIÇÃO DE 2000	54
5.3. ELEIÇÃO DE 1996: MODELO DE GOODMAN	54
5.4. ELEIÇÃO DE 1996: MODELO DE KING.....	56
5.5. ELEIÇÃO DE 2000: MODELO DE GOODMAN	62
5.6. ELEIÇÃO DE 2000: MODELO DE KING.....	65
6. CONCLUSÃO	79
7. BIBLIOGRAFIA	82

RESUMO

Este trabalho busca estudar os determinantes sociais do voto nas eleições para prefeito de 1996 e 2000, realizadas no município de São Paulo. Para alcançar tal objetivo são utilizadas duas metodologias estatísticas distintas: um modelo linear simples – aqui denominado Modelo de Goodman – e um modelo bayesiano desenvolvido pelo cientista político Gary King.

Inicialmente, são derivados os fundamentos estatísticos e matemáticos dos modelos de Goodman e de King, para que, finalmente, sejam apresentados os resultados das estimações realizadas para três variáveis demográficas no pleito de 1996 e para sete variáveis demográficas no pleito de 2000, através de mapas espaciais do município de São Paulo.

ABSTRACT

This essay seeks to study the social determinants of the vote in the elections for mayor in the city of São Paulo that occurred in 1996 and 2000. To reach such a target, two statistical methodologies are applied: a simple classical linear model – the so called Goodman model – and a bayesian model developed by Prof. Gary King.

In the first part of the essay, the statistical and mathematical foundations of these two models are presented. In the last part, the estimation results for three social variables in the 1996 election and seven social variables in the 2000 election are presented. The estimated parameters are presented in the form of spatial maps.

1. Introdução

Esta dissertação tem por objetivo estudar a questão dos determinantes sociais do voto. O estudo dos determinantes sociais do voto é importante na medida em que os resultados obtidos pela análise podem trazer respostas a uma gama de questões sociológicas, históricas, econômicas, geográficas e de ciência política. No âmbito sociológico, a compreensão de como os entes sociais registram os seus votos individuais é importante insumo no estudo de aspectos da cidadania e – dinamicamente – no estudo de como as alterações nas condições de vida e de educação afetam as escolhas por dirigentes e partidos. Na ótica dos historiadores, o estudo dos determinantes sociais do voto pode elucidar questões e corroborar teorias sobre movimentos históricos envoltos em controvérsia. Por exemplo, nos últimos dez anos, com a organização, na Alemanha, de uma série de arquivos eleitorais do período entre guerras, uma série de estudos passou a tentar explicar uma grande controvérsia histórica: quais foram os grupos sociais que garantiram a ascensão ao poder do Nazismo na Alemanha dos anos trinta.

No âmbito econômico, a literatura de “nova economia política” busca analisar – entre inúmeras outras questões – o impacto de alterações no nível de renda, emprego e escolaridade sobre o ciclo eleitoral. A Geografia estuda como os agentes interagem no espaço. Conseqüentemente, a análise da dispersão espacial do voto no espaço e suas inter-relações com os diversos sub-grupos sociais sob análise constitui tradicional objeto de interesse dos geógrafos. Por fim, a ciência política tem uma longa tradição no estudo dos determinantes sociais do voto, iniciada nos primeiros anos do século XX, em virtude do interesse em se entender os impactos políticos da implantação do sufrágio feminino nos Estados Unidos.

Os partidos políticos, por sua vez, têm particular interesse em conhecer como os eleitores votam. Atualmente, os partidos políticos no Brasil têm adotado um comportamento mais profissional em sua busca pelo voto, principalmente nas grandes cidades. Dentro desta nova estratégia, os partidos têm buscado analisar os resultados de eleições passadas com o objetivo de “conhecer” melhor o seu eleitorado (saber quem são seus eleitores e onde os mesmos estão localizados) e quais são as características dos grupos de eleitores que apresentam rejeição ao partido ou ao candidato em questão. De

posse destas informações o partido define uma estratégia ótima de discurso e de propaganda para cada sub região em que a próxima eleição se realizará.

O objetivo específico desta dissertação é analisar os determinantes sociais do voto dos eleitores do município de São Paulo nas eleições para prefeito, em segundo turno, realizadas em 1996 e 2000.

Na eleição realizada em 1996, em segundo turno, o candidato Celso Pitta (PPB)¹ obteve 61% dos votos válidos, contra 39% dos votos válidos obtidos pela candidata petista Luiza Erundina. Já nas eleições municipais de 2000, a candidata Marta Suplicy (PT) obteve 58,5% dos votos válidos, batendo o candidato Paulo Maluf (PPB) que totalizou 41,5% dos votos válidos, segundo dados finais do TRE.

Para analisar os determinantes sociais do voto foram considerados dois modelos teóricos alternativos: (1) um modelo linear simples, denominado, na literatura de ciência política, como modelo de Goodman (1953) e (2) um modelo estatístico bayesiano desenvolvido por King (1997), doravante denominado Modelo de King. Ambos os modelos se aplicam a apenas uma cross-section. Desta forma, cada um dos modelos será ajustado às duas eleições sob análise, de modo que serão estimados quatro modelos estáticos para efeito de análise dos determinantes sociais do voto no município de São Paulo.

Os dados demográficos utilizados neste estudo foram disponibilizados pelo IBGE e incluem os dados da Contagem 1996 do IBGE das variáveis sexo, idade e escolaridade (anos concluídos de estudo) e dados do Censo 2000, que incluem as variáveis sexo, idade, raça, religião, escolaridade (anos concluídos de estudo), renda (todas as fontes de renda) e origem migratória, por área de ponderação².

A dissertação está dividida da seguinte forma. No segundo capítulo é apresentada uma breve revisão bibliográfica da literatura. No terceiro capítulo, a metodologia dos modelos utilizados nesta dissertação é detalhada. Os dados eleitorais e demográficos utilizados na análise são descritos estatística e espacialmente no capítulo 4. O capítulo 5

¹ Atualmente, o então Partido Progressista Brasileiro (PPB) denomina-se apenas Partido Progressista (PP).

² Segundo a 'Documentação dos Microdados da Amostra' do Censo Demográfico 2000, Área de Ponderação "constitui uma unidade geográfica, formada por um agrupamento mutuamente exclusivo de setores censitários, para a aplicação dos procedimentos de calibração das estimativas com as informações conhecidas para a população como um todo".

traz a estimação e os resultados dos modelos econométricos, enquanto que o último capítulo apresenta os comentários finais e recomendações de pesquisas futuras.

2. Revisão Bibliográfica

O objetivo deste capítulo é fazer uma breve revisão da literatura relacionada a escolhas dos eleitores e características demográficas. Inicialmente são apresentados alguns estudos mais gerais no campo da “nova economia política”. Na sequência é apresentada uma revisão bastante detalhada de uma vertente da literatura que estuda a associação entre variáveis eleitorais a variáveis demográficas par a par: a literatura de “Inferência Ecológica”. Por fim, são apresentados estudos que analisam eleições consecutivas.

A literatura de nova economia política tem se desenvolvido bastante desde o início da década de 1970. Entre os seus diversos objetos de estudo destacam-se as teorias de “*political business cycles*”, segundo as quais os ciclos econômicos seriam afetados diretamente pelos calendários políticos eleitorais. Esta vertente de estudos analisa a existência de ciclos políticos oportunistas (os quais podem ser racionais ou irracionais) e ciclos partidários racionais. Nesta linha de literatura destacam-se autores como Nordhaus (1975), Rogoff (1990) e Alesina (1987). De modo geral, tais autores buscam uma associação entre os ciclos políticos e as políticas econômicas dos governos.

Outra linha de estudo importante dentro da nova economia política é a que busca relacionar o desempenho eleitoral de determinado partido – ou candidato – às características demográficas do eleitorado. O fato da informação sobre os votos ser secreta determina uma divisão fundamental nesta literatura, entre os estudos que não consideram os problemas teóricos existentes em se calcular regressões a partir de dados agregados para se obter parâmetros individuais e estudos que consideram estas dificuldades.

No primeiro grupo, destaca-se Simmons (1965), que usa técnicas de regressão múltipla para combinar informações de censo com resultados eleitorais no Canadá. Miller, Raab e Brito (1974) analisam a relação na Grã-Bretanha de variáveis como ocupação, sexo, religião, educação, renda, região de residência com a distribuição de votos. Os autores salientam a importância de variáveis qualitativas (como hábitos, impacto da mídia, etc) e espaciais em seus resultados. Sprague (1976), por sua vez, analisa o comportamento político eleitoral no estado do Texas (EUA) através de um modelo de regressão múltipla por distritos eleitorais.

O segundo grupo, por sua vez, considera a importância de se levar em consideração a dificuldade existente em se estimar regressões a partir de dados agregados (proporções) para se estimar parâmetros individuais. Este problema, que na literatura de ciência política possui longa tradição, é conhecido como o problema da “Inferência Ecológica”.

O interesse dos cientistas políticos em relação à Inferência Ecológica é, pelo menos, tão remoto quanto o estudo realizado por William Ogburn e Inez Goltra (1919). Neste estudo, Ogburn e Goltra (1919) analisam, através do cálculo de correlações parciais, qual teria sido o impacto da introdução do sufrágio feminino sobre a dinâmica política da cidade de Portland, Oregon. Os autores tinham o interesse de testar a hipótese de que a introdução do sufrágio feminino por políticos progressistas aumentaria o conservadorismo do eleitorado, de modo a, no médio prazo, diminuir o poder dos mesmos progressistas que introduziram o sufrágio feminino³.

Apesar de Ogburn & Goltra (1919) já terem salientado a dificuldade existente na estimação de parâmetros individuais a partir de dados agregados, foi apenas Robinson (1950), quem formalizou as condições nas quais correlações realizadas a partir de dados agregados equivaleriam matematicamente a correlações baseadas em observações individuais. Em função disso, Robinson (1950) alertava para que nunca se tentasse inferir parâmetros individuais a partir de dados agregados, evitando assim a chamada falácia ecológica. Em suas palavras, “...*the purpose of this paper will have been accomplished if it prevents the future computation of meaningless correlations and stimulates the study of similar problems with the use of meaningful correlations between the properties of individuals*”.

O impacto do artigo de Robinson (1950) - e de sua recomendação - sobre a literatura de Inferência Ecológica não poderia ter sido maior⁴. Pesquisadores dos mais diversos ramos das ciências sociais passaram a realizar estudos preferivelmente com base em dados individuais, em uma época que estes tipos de dados passaram a ser disponibilizados de forma mais ampla. Por outro lado, em um segundo momento, a comunidade estatística passou a devotar grande atenção e esforço no intuito de solucionar o problema da Inferência Ecológica.

³ Os autores trabalhavam com esta hipótese baseados na experiência do impacto da introdução do sufrágio feminino na Noruega, anos antes.

⁴ Segundo King (1997), uma medida do impacto causado pelo artigo “Ecological Correlations and the Behavior of Individuals” de Robinson é o fato do mesmo ter sido citado em mais de 800 artigos acadêmicos entre 1950 e 1997.

Goodman (1953, 1959) apresentou dois artigos que tinham por objetivo tornar mais claras as condições sob as quais seria possível realizar inferências ecológicas, uma vez que, apesar das recomendações feitas por Robinson (1950), alguns cientistas sociais continuavam a publicar artigos em que parâmetros individuais eram estimados a partir de dados agregados

Paralelamente aos desenvolvimentos de Goodman (1953), Duncan e Davis (1953) apresentaram um método matemático, posteriormente denominado método de “bounds”, que tinha por objetivo ser um instrumento adicional na análise de resultados de correlação ecológica. Segundo os autores, antes de aceitar os resultados de uma inferência ecológica como estimativas de parâmetros individuais, o pesquisador poderia adotar o procedimento por eles desenvolvido⁵.

Nos anos que se seguiram aos artigos clássicos de Robinson (1950), Goodman (1953) e Duncan & Davis (1953) a literatura se dividiu entre os autores que buscavam uma solução para o problema da Inferência Ecológica e os autores que não acreditavam ser possível inferir parâmetros individuais a partir de dados agregados e priorizavam a obtenção de dados individuais.

Entre os autores céticos em relação ao futuro das pesquisas no campo da Inferência Ecológica, destacam-se Achen & Shively (1995), que afirmavam que a busca de uma solução para o problema da Inferência Ecológica era “infrutífera”.

Por outro lado, entre os autores que continuavam a buscar soluções para o problema da Inferência Ecológica, duas agendas de pesquisa diferentes passaram a ser exploradas: de um lado os pesquisadores que tentaram buscar uma solução para o problema da Inferência Ecológica analisando apenas uma “cross section” - Freedman et al.(1991), King (1997), Cho e Anselin (2002) - e, de outro, pesquisadores que passaram a estudar Inferência Ecológica de Painel - Penubarti & Schuessler (1998) e Pelzer, Eisinga & Franses (2002).

Na agenda de pesquisa que busca uma solução para o problema da Inferência Ecológica analisando apenas uma “cross section”, destacam-se Freedman, Klein, Sacks, Smyth & Everett (1991), os quais propuseram o Modelo de Vizinhança Linear como alternativa

⁵ O método de “bounds” foi proposto tomando como base os desenvolvimentos realizados por Udny Yule no estudo de inferências a partir de dados incompletos.

ao Modelo de Goodman (1953), cujas hipóteses os autores consideravam muito restritivas⁶.

King (1997) apresentou um engenhoso modelo estatístico bayesiano, que possibilita não apenas a estimação dos parâmetros de interesse agregados (para uma cidade como um todo, por exemplo), mas também para quaisquer micro-regiões de interesse (tais como bairros, áreas de ponderação, etc).⁷ O modelo proposto por King (1997) incorpora desenvolvimentos já existentes na literatura, tal como o método determinístico de “bounds” de Duncan & Davis (1953), além de se basear em hipóteses estatísticas específicas.

O impacto da obra de King sobre a literatura de Inferência Ecológica somente foi inferior ao impacto causado pelo artigo de Robinson, 47 anos antes. Não apenas cientistas políticos, mas estudiosos pertencentes aos mais diversos ramos das ciências sociais passaram a se posicionar contra ou a favor o modelo proposto por King.

Um dos primeiros autores a se posicionar em relação à obra de King foi Cho (1998), que apresentou uma criteriosa análise da solução proposta por King para o problema da Inferência Ecológica. Segundo Cho (1998), o modelo proposto por King (1997) é, sem dúvida, uma contribuição significativa à literatura de Inferência Ecológica. Entretanto, Cho (1998) critica a falta de estatísticas de teste efetivas que tornem possível diagnosticar o atendimento, ou não, das hipóteses fundamentais do modelo de King (1997)⁸.

Entre os críticos do modelo de King destaca-se Freedman et al. (1998), que, em uma série de artigos publicados entre 1998 e 1999, questionou a robustez dos estimadores do modelo de King⁹.

King (1999) argumenta que a questão não é achar um conjunto de dados a partir dos quais o modelo estatístico em questão teria uma performance ruim – o que pode ocorrer em qualquer área da estatística – mas, em suas palavras, considerar “...*how much*

⁶ Entretanto, segundo King (1997), as hipóteses presentes no Modelo de Vizinhança Linear eram tão ou mais restritas que as incluídas no modelo de Goodman (1953) e os resultados das estimações realizadas pelo primeiro modelo se mostraram menos satisfatórias do que as obtidas pelo segundo.

⁷ O modelo de King será descrito e apresentado detalhadamente no Capítulo 3 desta dissertação.

⁸ Segundo Cho, a principal hipótese do modelo de King é a que assume que os regressores não são correlacionados com os parâmetros de interesse.

⁹ Os principais artigos produzidos por Freedman et al. em sua controvérsia com King entre 1998 e 1999 foram: “On ‘Solutions’ to the Ecological Inference Problem”; “Response to King’s Comment”; “Review of A Solution to the Ecological Inference Problem, by G. King” e “Ecological Inference and the Ecological Fallacy”.

*information a model uses*¹⁰, *how often one can recognize when the model is inappropriate, and whether the approach suggests other models in these instances*". Deste modo, King (1999) destaca que seu modelo deve ser aplicado não de modo mecânico, mas contando com a "sensibilidade" do pesquisador, a qual pode ser incorporada no modelo através de 'priors' estatísticas na estimação.

Mais recentemente, Anselin & Cho (2002) criticaram a importância marginal dada pelo modelo de King (1997), especificamente, à violação da hipótese de independência espacial e aos efeitos espaciais em geral¹¹. Os autores abordam o problema da Inferência Ecológica considerando explicitamente o problema de efeitos espaciais (autocorrelação e heterogeneidade espacial) e como diferentes modelagens destes efeitos afetam os resultados do modelo de King.¹² Segundo Anselin & Cho (2002), os resultados obtidos anteriormente - que indicavam a possibilidade de se relaxar a hipótese de independência espacial sem conseqüências muito graves - eram exageradamente otimistas. Segundo os autores, se os dados do estudo em questão apresentarem elevada autocorrelação espacial os parâmetros estimados permanecerão não-viesados, mas serão ineficientes.

Entre os autores que contribuíram para o aprimoramento do Método de King, estão Mattos & Veiga (2000), que desenvolveram um modelo que utiliza um algoritmo de convergência diferente¹³ do utilizado por King em seus programas EI e EzI¹⁴. O objetivo de Mattos & Veiga (2000) não é apresentar um modelo alternativo ao desenvolvido por King, mas ampliar o conhecimento estatístico em torno do problema de estimação de parâmetros individuais a partir de dados agregados visando o desenvolvimento de testes estatísticos mais formais e de orientações práticas mais adequadas à estimação do modelo.

Ainda na agenda de pesquisa que busca uma solução para o problema da Inferência Ecológica analisando apenas uma "cross section", destacam-se os desenvolvimentos

¹⁰ King argumenta que seu modelo é superior ao modelo de Vizinhança apresentado por Freedman et al. por ter consolidado duas linhas de pesquisa anteriormente concorrentes na literatura, sejam elas, a do Modelo de Goodman e a do Método de "Bounds" de Duncan & Davis. Adicionalmente, King argumenta que outra vantagem adicional de seu modelo é a possibilidade de se inserir conhecimentos qualitativos como "prior" durante a estimação.

¹¹ King (1999) argumenta que uma das áreas de pesquisa no campo da Inferência Ecológica que deveria ser desenvolvida era a área relativa à hipótese espacial de seu modelo.

¹² King utiliza em seu modelo uma estrutura espacial unidimensional análoga a problemas de séries de tempo unidirecionais, na forma: $\beta_i = \delta\beta_{i-1} + (1-\delta)u_i$.

¹³ Algoritmo EM.

¹⁴ Estes são os dois programas disponibilizados por King no site <http://gking.harvard.edu/stats.shtml> para estimar seu modelo. EI é o programa usado para estimar o modelo de King estendido e modelos com tabelas maiores do que 2x2, enquanto que EzI é o programa utilizado na estimação de modelos mais simples, mas que, em compensação, possui mais instrumentos gráficos e de visualização.

recentes na busca de estimadores de método de momentos mais rápidos (Lewis, 1998) e na derivação de estimadores não paramétricos (Rivers, 1998).

Penubarti & Schuessler (1998) foram os precursores da agenda de pesquisa que busca uma solução para o problema da Inferência Ecológica através da análise de painéis ou pseudopainéis. A estratégia analítica dos autores foi utilizar o modelo desenvolvido por King (1997) – o qual foi originalmente desenvolvido para apenas uma “cross section” – para estimar um modelo dinâmico de 2 períodos em que se estuda o apoio ao Presidente americano Ronald Reagan, conforme a estrutura abaixo¹⁵:

Tabela 1
Modelo de King Aplicado à
“Cross Sections” Repetidas

		Opinião em t_2		Total
		Aprova	Não Aprova	
Opinião em t_1	Aprova	β^{11}_i	$1 - \beta^{11}_i$	A^1_i
	Não Aprova	β^{01}_i	$1 - \beta^{01}_i$	$1 - A^1_i$
Total		A^2_i	$1 - A^2_i$	

Ao contrário do modelo original de King, no qual existem P micro-regiões de análise, nesta abordagem, trabalha-se com i “perfis” de eleitores, sendo que um exemplo de perfil poderia ser “homem branco, sem ensino superior completo, de classe média, filiado ao partido democrata”.¹⁶ O objetivo da estimação, neste caso, é observar a variação do apoio ao Governo Reagan, em 2 períodos, para cada um dos perfis de eleitores, dado que somente é possível observar a aprovação agregada ao governo.

Na mesma agenda de pesquisa que busca uma solução para o problema da Inferência Ecológica através da análise de painéis e pseudopainéis, Pelzer, Eisinga & Franses (PE&F) apresentaram, em uma série de 4 artigos produzidos entre 2001 e 2002, um modelo de transição de Markov para a estimação de probabilidades de transição binárias (de entrada e de saída) em nível individual.

Em seus artigos, PE&F apresentaram aplicações empíricas dos mais diversos campos de interesse: da análise da evolução da penetração de computadores nos lares holandeses,

¹⁵ Os autores analisam dois momentos do Governo Reagan: o primeiro entre Set/86 e Jan/87, no qual a aprovação ao governo Reagan caiu 15 pontos percentuais; e o segundo entre Out/87 e Jan/88 durante o qual a aprovação ao Governo Reagan permaneceu constante.

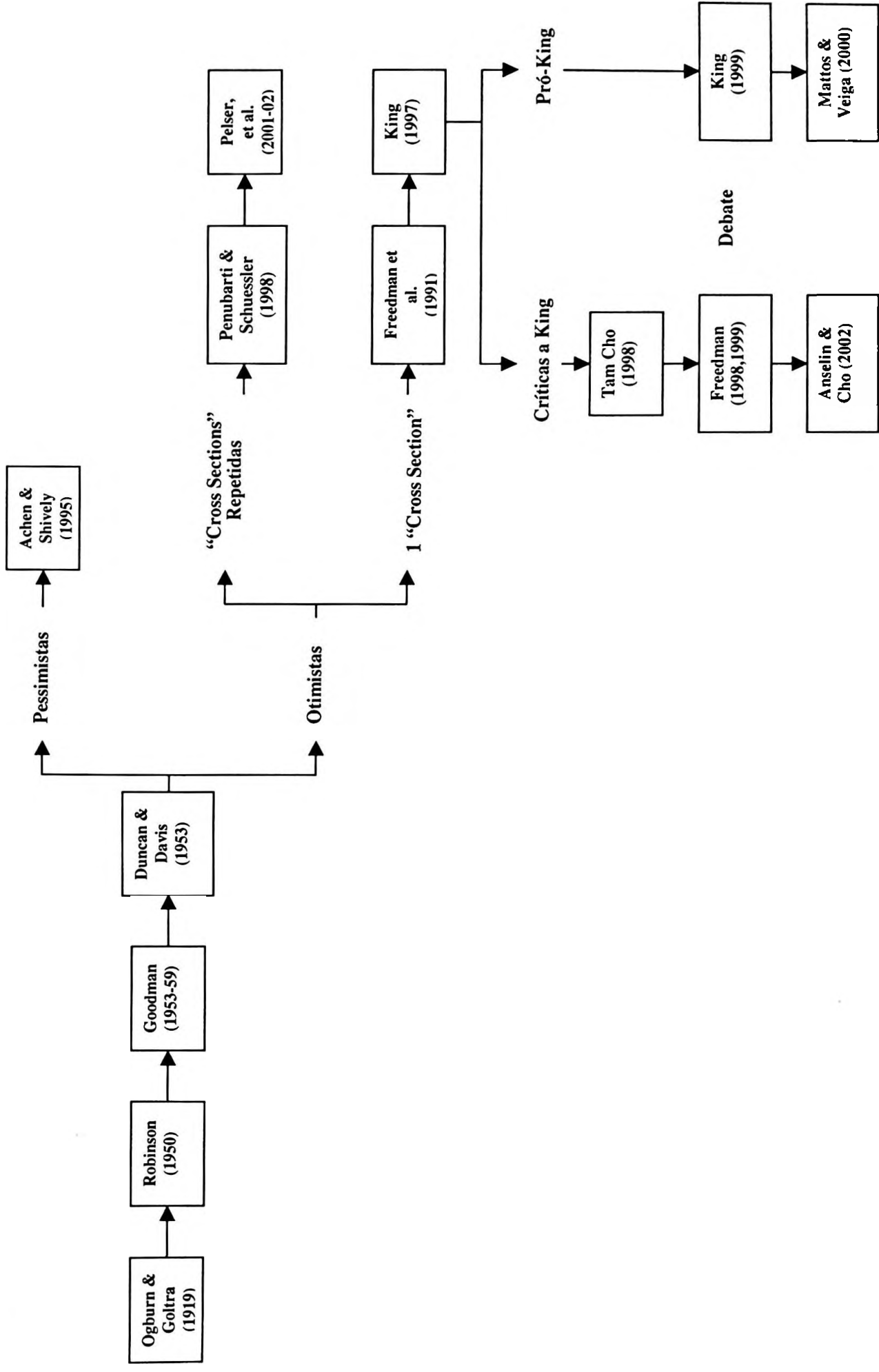
¹⁶ Na aplicação apresentada no artigo, os autores trabalham com 5 variáveis demográficas – sexo, educação, raça, classe social e filiação partidária – e obtém, deste modo, 72 perfis demográficos para efetuar a análise.

passando por um estudo das intenções de voto nas eleições presidenciais americanas de 1976, até chegar à análise da participação das mulheres no mercado de trabalho da Holanda.

Nesta dissertação não serão estimados modelos de Inferência Ecológica através de técnicas de painel e pseudopainel, sendo estimados apenas modelos de Inferência Ecológica para uma única “cross section”, individualmente.

A seguir, é apresentada uma esquematização gráfica dos desenvolvimentos da literatura de Inferência Ecológica.

Figura 1: Desenvolvimentos da Literatura de Inferência Ecológica



3. Metodologia

3.1. Modelo de Goodman

Goodman (1953) apresentou um modelo de Inferência Ecológica para a análise de “cross sections” individuais que se baseia em uma tabela de entradas na forma:

Tabela 2
Modelo de Goodman
Raça vs. Candidato A

		Candidato A		Total
		Vota	Não Vota	
Raça	Negro	β_i^b	$1 - \beta_i^b$	X_i
	Branco	β_i^w	$1 - \beta_i^w$	$1 - X_i$
Total		T_i	$1 - T_i$	

onde:

T_i = % de votos obtidos pelo candidato A em cada uma das P micro-regiões¹⁷;

X_i = % de pessoas que possuem uma determinada característica demográfica (neste caso, pertencer à raça negra) em cada uma das P micro-regiões;

β_i^b = % de pessoas negras que votaram no candidato A em cada uma das micro-regiões;

β_i^w = % de pessoas brancas que votaram no candidato A em cada uma das micro-regiões;

O problema de Inferência Ecológica ocorre quando o pesquisador tem interesse em estimar os parâmetros da parte interior (parâmetros β_i^b e β_i^w) de uma tabela como a apresentada acima a partir dos dados totais das linhas e colunas (variáveis T_i e X_i). A forma linear padrão do problema da Inferência Ecológica é dada pela seguinte equação:

$$T_i = \beta_i^b X_i + \beta_i^w (1 - X_i)$$

Obviamente, esta forma funcional não é identificada no modo irrestrito como se apresenta, uma vez que existem duas vezes mais parâmetros a serem estimados do que observações¹⁸.

Para solucionar este problema de identificação, Goodman apresenta um conjunto de três hipóteses:

¹⁷ Uma micro-região pode ser um bairro, uma área de ponderação, uma seção eleitoral, etc.

¹⁸ Deseja-se estimar dois parâmetros (β_i^b e β_i^w) para cada uma das P micro-regiões.

1. Parâmetros constantes: os parâmetros de cada micro-região (β_i^b e β_i^w) devem ser constantes e iguais aos parâmetros agregados B^b e B^w ;¹⁹
2. Modelo linear: Goodman trabalha com uma “identidade contábil” linear a ser estimada por Mínimos Quadrados Ordinários (MQO) que deve valer para cada uma das P micro-regiões;

$$T_i = B^b X_i + B^w (1 - X_i)$$

3. Homoscedasticidade.

O modelo de Goodman (1953) foi aplicado, separadamente, para os dados das eleições, em segundo turno, para a prefeitura de São Paulo realizadas em 1996 e 2000.

Na estimação do modelo de Goodman (1953) para a “cross section” de 1996 foram relacionados, individualmente, os percentuais de votos obtidos pela então candidata Luiza Erundina (PT) em 398 áreas de ponderação do município de São Paulo a três diferentes variáveis demográficas, sejam elas: sexo, idade e escolaridade.

Já na estimação da “cross section” de 2000, o modelo de Goodman (1953) foi aplicado relacionando, individualmente, os percentuais de votos obtidos pela então candidata Marta Suplicy (PT) em 412 áreas de ponderação do município de São Paulo às seguintes variáveis demográficas: sexo, idade, raça, religião, escolaridade, rendimentos totais e origem migratória²⁰. Cabe salientar, mais uma vez, que o modelo de Goodman utiliza os dados de todas as áreas de ponderação disponíveis para estimar apenas 2 parâmetros agregados (para o município de São Paulo como um todo).

3.2. Modelo de King

King (1997) apresentou uma engenhosa contribuição para a solução do problema da Inferência Ecológica. Duas principais inovações são associadas ao modelo proposto por King (1997). A primeira reside no fato de seu modelo incorporar, pela primeira vez, dois veios até então antagônicos da literatura de Inferência Ecológica: o método de ‘bounds’ de

¹⁹ Goodman (1959, pág. 612) argumenta que “...in practice, the actual values of the parameters will not be constant, but it may be the case that the average $E(\beta_i^b | X)$ of the values of β_i^b , for populations with the same proportion X of Negroes, is constant [i.e., $E(\beta_i^b | X)$ is the same for different values of X] and the average $E(\beta_i^w | X)$ of the values of β_i^w , for populations with the same X value, is also constant. In this situation, the main assumption of linear regression analysis, $E(T | X) = A + BX$, holds true, where $A = E(\beta_i^w | X)$ and $B = E(\beta_i^b | X) - E(\beta_i^w | X)$ ”.

²⁰ Os 7 modelos foram rodados utilizando o software EzI disponibilizado por King no site <http://gking.harvard.edu/stats.shtml>, o qual não apenas estima o modelo proposto por King (1997), mas também estima o modelo de Goodman (1953).

Duncan & Davis (1953) e o modelo de Goodman (1953). A segunda é a possibilidade que o modelo apresenta de estimar não apenas os parâmetros agregados de interesse – para a cidade como um todo -, mas também os parâmetros associados às micro-regiões. Adicionalmente, o modelo de King (1997) soluciona três críticas frequentemente endereçadas ao modelo de Goodman (1953), sejam elas:

1. Que os parâmetros estimados pelo modelo de Goodman (1953), muitas vezes, não possuem significado prático, uma vez que estes, por estarem fora do intervalo entre zero e um, não representam probabilidades;
2. Que a violação da forte hipótese de parâmetros de interesse constantes entre as micro-regiões gera viés nos parâmetros estimados;
3. Que o modelo de Goodman (1953) não é robusto à existência de heteroscedasticidade nas observações.

Desta forma, o modelo de King (1997) é capaz de estimar parâmetros que sempre estarão contidos no intervalo (0,1) – de modo a sempre representar probabilidades ou proporções -, além de relaxar a forte hipótese de parâmetros de interesse constantes e de ser robusto à presença de heteroscedasticidade.

O modelo proposto por King (1997) se divide em 2 estágios: no primeiro, os dados são ajustados a um modelo parametrizado do modo mais conveniente possível. Isso torna o ajustamento dos dados e a estimação processos diretos, apesar dos parâmetros obtidos nesta primeira etapa não serem os parâmetros de interesse. No segundo estágio, os resultados obtidos no estágio anterior são utilizados, sem necessidade de nenhuma hipótese adicional, para se obter os parâmetros de interesse (parâmetros agregados e para cada uma das micro-regiões).

O modelo padrão proposto por King (1997)²¹ se baseia em uma matriz de entradas 2x2 padrão, tal como a exemplificada na Tabela 3, abaixo.

²¹ Adicionalmente, King apresenta outros dois modelos, um em que se pode trabalhar com tabelas de dimensões maiores do que 2x2 (doravante denominado modelo de King Generalizado) e outro em que são introduzidas covariadas para identificar o modelo (doravante denominado modelo de King Estendido). Trabalharemos, no presente estudo apenas com o modelo de King Padrão, uma vez que tanto o modelo Generalizado quanto o Estendido ainda não são consensuais na literatura e continuam sendo desenvolvidos.

Tabela 3
Modelo de King
Religião vs. Candidato A

		Candidato A		Total
		Vota	Não Vota	
Religião	Protestante	β_i^p	$1 - \beta_i^p$	X_i
	Católica	β_i^c	$1 - \beta_i^c$	$1 - X_i$
Total		T_i	$1 - T_i$	

King (1997) fundamenta seu modelo em hipóteses determinísticas e estatísticas. As hipóteses determinísticas do modelo são as seguintes:

1. A “identidade contábil” $T_i = \beta_i^b X_i + \beta_i^w(1 - X_i)$ deve ser válida para todas as micro-regiões.
2. A cada micro-região é aplicado o método de ‘bounds’, limitando matematicamente os valores que β^b e β^w podem assumir, de modo a agregar mais informação ao processo de estimação.

$$\beta_i^b \in [L_i^b, U_i^b] \text{ e } \beta_i^w \in [L_i^w, U_i^w]$$

onde L = lower bound e U = upper bound

$$L_i^b = \max\left\{0, \frac{[T_i - (1 - X_i)]}{X_i}\right\} \geq 0 \quad U_i^b = \min\left(\frac{T_i}{X_i}, 1\right) \leq 1$$

$$L_i^w = \max\left[0, \frac{(T_i - X_i)}{X_i}\right] \geq 0 \quad U_i^w = \min\left[\frac{T_i}{(1 - X_i)}, 1\right] \leq 1$$

As hipóteses estatísticas do modelo, por sua vez, são as seguintes:

1. Regressores não-estocásticos: $X_i; i=1, 2, \dots, P$ são variáveis determinísticas;
2. $(\beta_i^b, \beta_i^w); i=1, \dots, P$ são gerados por uma função de distribuição normal bivariada truncada em $[0,1] \times [0,1] \in \mathfrak{R}^2$, condicional em X_i , na forma²²;

²² Ao invés de assumir que os parâmetros das micro-regiões são constantes assume-se apenas que estes parâmetros têm “algo em comum”.

$$P(\beta_i^b, \beta_i^w) = TN(\beta_i^b, \beta_i^w | B, \Sigma)$$

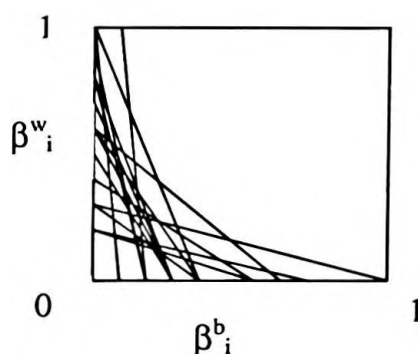
onde

$$B = \begin{pmatrix} B^b \\ B^w \end{pmatrix} \text{ e } \Sigma = \begin{pmatrix} \sigma_b^2 & \sigma_{bw} \\ \sigma_{bw} & \sigma_w^2 \end{pmatrix}$$

3. Médias Constantes: o vetor B não depende de X_i e $(1 - X_i)$, $i=1, \dots, P$.²³
4. Independência Espacial: a variável aleatória condicional $T_i | X_i$ é independente entre seções.²⁴

A intuição por trás do modelo proposto por King (1997) é utilizar ao máximo as informações disponíveis em cada micro-região, através da aplicação do método de “bounds” para limitar ao máximo os valores que os parâmetros β_i^b e β_i^w podem assumir. King (1997) sugere que seja utilizado um quadrado com arestas iguais à unidade como instrumento gráfico de visualização para facilitar a compreensão intuitiva de seu modelo. Na Figura 2, abaixo, cada linha representa os valores que os parâmetros β_i^b e β_i^w podem assumir para cada micro-região após a aplicação do método de “bounds”.

Figura 2
“Tomography Plot”
Limites por Micro-região



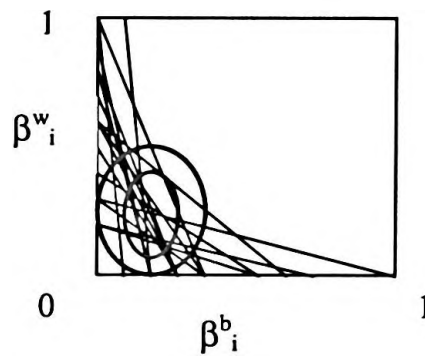
Os parâmetros β_i^b e β_i^w podem estar situados em qualquer um dos infinitos pontos pertencentes ao P-ésimo segmento de reta associado a uma micro-região específica. Para

²³ A hipótese de independência da média dos parâmetros em relação a X_i é conhecida na literatura de Inferência Ecológica como a hipótese que garante a ausência de viés de agregação. Esta é a hipótese mais importante do modelo de King (1997).

²⁴ King argumenta que a violação desta hipótese não gera consequências graves a seu modelo. Anselin & Cho (2002) discutem com mais detalhe as consequências da violação da hipótese de independência espacial. Segundo as simulações de Monte Carlo empreendidas por Anselin & Cho (2002), a existência de autocorrelação espacial não gera viés nos parâmetros estimados, mas tem efeitos sérios sobre a precisão dos mesmos. Segundo os autores, “The effect (of spatial autocorrelation) is limited to the precision of the estimator, since these forms of spatial correlation, in and of themselves, do not induce bias”.

se obter as estimativas de (β^b_i, β^w_i) para cada uma das P micro-regiões é feita a sobreposição de uma função de distribuição normal bivariada truncada agregada²⁵ (para o município como um todo) sobre os P segmentos de reta contidos na Figura 2, conforme figura abaixo.

Figura 3
Limites por Micro-região sobrepostos pela Distribuição Normal Bivariada Truncada



Para cada segmento de reta individualmente, assume-se que o ponto (β^b_i, β^w_i) se situa sob a região em que a função de distribuição concentra maior massa. Deste modo, é possível determinar as posições, para cada um dos segmentos de reta (i.e., micro-regiões), em que as estimativas pontuais (β^b_i, β^w_i) têm maior probabilidade de estarem situadas.

3.2.1. Descrição Formal do Modelo

3.2.1.1. Parâmetros de Interesse e Parâmetros de Parcial Interesse

Uma distinção inicial a ser feita entre o modelo de King (1997) e o modelo de Goodman (1953) – e que terá conseqüências importantes na derivação matemática do modelo de King (1997) mais à frente – diz respeito aos parâmetros agregados de interesse que cada um dos autores considera. Goodman (1953) estima e utiliza apenas os parâmetros agregados e não-ponderados B^b e B^w , dados pelas fórmulas abaixo:

$$B^b = \frac{1}{P} \sum_{i=1}^P \beta^b_i \qquad B^w = \frac{1}{P} \sum_{i=1}^P \beta^w_i$$

King (1997), por sua vez, calcula os parâmetros de interesse agregados ponderados pelo número de indivíduos pertencentes à categoria demográfica em questão em cada uma das

²⁵ Uma hipótese razoavelmente forte por trás da utilização de uma função de distribuição normal bivariada truncada – e que deve ser verificada e, se não confirmada, alterada – é a existência de apenas uma moda na distribuição dos dados. Segundo King, este normalmente é o caso nas aplicações práticas.

micro-regiões. Desta forma, evita-se o que na literatura de Inferência Ecológica é conhecido como “Efeito Manhattan”²⁶.

$$B^b = \frac{\sum_{i=1}^p N_i^b \beta_i^b}{N^b} \quad B^w = \frac{\sum_{i=1}^p N_i^w \beta_i^w}{N^w}$$

onde $N^b = \sum_{i=1}^p N_i^b$, ou seja, o total de indivíduos pertencentes à categoria demográfica em questão (neste caso, negros) em todas as micro-regiões e

$N^w = \sum_{i=1}^p N_i^w$, ou seja, o total de indivíduos *não* pertencentes à categoria demográfica em questão (neste caso, negros) em todas as micro-regiões.

Dado que o modelo de King (1997) se propõe a estimar parâmetros agregados de interesse ponderados, a esperança e a variância dos parâmetros por micro-regiões não serão parâmetros de interesse, mas apenas aproximações dos mesmos a serem utilizadas durante o processo de ajustamento dos dados (denominados, em função disto, parâmetros de parcial interesse).

$$E \begin{pmatrix} \beta_i^b \\ \beta_i^w \end{pmatrix} = \begin{pmatrix} B^b \\ B^w \end{pmatrix} = B$$

$$Var \begin{pmatrix} \beta_i^b \\ \beta_i^w \end{pmatrix} = \begin{pmatrix} \sigma_b^2 & \sigma_{bw} \\ \sigma_{bw} & \sigma_w^2 \end{pmatrix} = \Sigma$$

A relação entre os parâmetros das micro-regiões e sua média agregada não ponderada é:

$$\begin{aligned} \beta_i^b &= B^b + \varepsilon_i^b \\ \beta_i^w &= B^w + \varepsilon_i^w \end{aligned}$$

As quais, substituídas na identidade contábil, mostram a importância deste vetor de parâmetros não ponderados no processo de estimação.

²⁶ Denomina-se “Efeito Manhattan” na literatura de Inferência Ecológica o problema de viés nos parâmetros estimados causado pela ausência de ponderação das micro-regiões de interesse por seu tamanho, ou, mais especificamente, pelo número de indivíduos pertencentes a uma determinada característica demográfica que habitam esta micro-região. O nome “Efeito Manhattan” tem origem no fato de que, ao contrário dos outros quatro distritos de Nova York, Manhattan não possui uma desagregação espacial por área de ponderação significativa, o que acaba viesando qualquer espécie de estudo de Inferência Ecológica a ser realizado sem ponderação para o município de Nova York.

$$T_i = \beta_i^b X_i + \beta_i^w (1 - X_i)$$

$$T_i = (B^b + \varepsilon_i^b) X_i + (B^w + \varepsilon_i^w) (1 - X_i)$$

$$= B^b X_i + B^w (1 - X_i) + \varepsilon_i$$

$$\text{onde } \varepsilon_i = \varepsilon_i^b X_i + \varepsilon_i^w (1 - X_i)$$

Como $E(\varepsilon_i | X_i) = 0$, tem-se que

$$E(T_i | X_i) = B^b X_i + B^w (1 - X_i)$$

Por outro lado, as variâncias σ_b^2 e σ_w^2 e a covariância σ_{bw} afetam os dados agregados através da função variância condicional, a qual é heteroscedástica e possui a seguinte forma quadrática:

$$\begin{aligned} \text{Var}(T_i | X_i) &= \text{Var}(\varepsilon_i | X_i) \\ &= \sigma_b^2 X_i^2 + \sigma_w^2 (1 - X_i)^2 + \sigma_{bw} 2X_i (1 - X_i) \\ &= (\sigma_w^2) + (2\sigma_{bw} - 2\sigma_w^2) X_i + (\sigma_b^2 + \sigma_w^2 - 2\sigma_{bw}) X_i^2 \end{aligned}$$

A importância das derivações matemáticas da esperança e da variância condicional acima não deve ser negligenciada, uma vez que, sem elas, não seria possível se fazer Inferência Ecológica. Nas palavras de King, *“the exact procedure for finding evidence of individual level effects in these aggregate data comes from the fact that the expected values of β_i^b and β_i^w (B) appear in the mean function of the aggregate variable T_i and the variances and covariances (Σ) turn up in the variance function of T_i . The fact that these parameters reveals themselves by affecting specific features of aggregate data is what makes Ecological Inference possible”*.

3.2.1.2. Abordagem Bayesiana: Visão Geral

Formalmente, King (1997) utiliza uma abordagem estatística bayesiana tradicional, da forma:

$$P(\Psi | T) = \alpha \cdot P(\Psi) \cdot P(T | \Psi)$$

onde $P(\Psi | T)$ = Distribuição Posterior

$P(\Psi) = \text{'Prior'}$

$P(T|\Psi) = \text{Função Densidade de Probabilidade.}$

$T = \text{total de votos obtidos por um candidato específico.}$

e $\Psi = \{B^b, B^w, \sigma_b, \sigma_w, \rho\} = \{B, \Sigma\}$ é o vetor de parâmetros de parcial interesse

O objetivo final da estimação é obter a distribuição posterior dos parâmetros de cada uma das P micro-regiões. Para alcançar tal objetivo utiliza-se como 'prior' na regra de Bayes a distribuição posterior dos parâmetros agregados, que é obtida em um primeiro estágio de estimação.

3.2.1.3. Abordagem Bayesiana: Truncamento das Micro-regiões

Inicialmente, por se tratar de um problema de Inferência Ecológica, deriva-se a função de distribuição do parâmetro por micro-região β^b_i , condicional à variável conhecida T_i e ao vetor de parâmetros desconhecidos $\overset{\cup}{\Psi}$ ²⁷. O objetivo final do modelo de King é obter a distribuição posterior por micro-região $P(\beta^b_i | T_i)$, a qual é, apenas, condicionada à variável conhecida T_i .

King (1997) assume que os parâmetros de interesse de cada micro-região (β^b_i) são descritos por uma função de distribuição normal truncada univariada, na forma:

²⁷ $\overset{\cup}{\Psi} = \left\{ \overset{\cup}{B}^b, \overset{\cup}{B}^w, \overset{\cup}{\sigma}_b, \overset{\cup}{\sigma}_w, \overset{\cup}{\rho} \right\}$. King utiliza este vetor de parâmetros para ajustar os dados do modo mais eficiente possível no primeiro estágio de seu modelo. Conforme salientado anteriormente, estes não são os parâmetros de interesse, mas, apenas, parâmetros de parcial interesse. A notação \cup sobre o vetor de parâmetros Ψ indica que os parâmetros são não-truncados. Esta reparametrização é necessária em função da impossibilidade de se derivar uma forma algébrica explícita para os parâmetros truncados. Nas etapas intermediárias do modelo, King utiliza três parametrizações diferentes para o vetor $\overset{\cup}{\Psi}$: $\overset{\cup}{\Psi}$, que é o vetor de parâmetros não truncados; Ψ , que é o vetor de parâmetros truncados e Φ , que é uma reparametrização específica de $\overset{\cup}{\Psi}$ que é utilizada para facilitar a convergência numérica de um algoritmo a ser descrito mais à frente.

$$P\left(\beta_i^b \mid T_i, \check{\Psi}\right) = TN\left(\beta_i^b \mid \check{B}^b + \frac{\omega_i}{\sigma_i^2} \varepsilon_i, \check{\sigma}_b^2 - \frac{\omega_i^2}{\sigma_i^2}\right) = N\left(\beta_i^b \mid \check{B}^b + \frac{\omega_i}{\sigma_i^2} \varepsilon_i, \check{\sigma}_b^2 - \frac{\omega_i^2}{\sigma_i^2}\right) \frac{1(\beta_i^b)}{S\left(\check{B}, \check{\Sigma}\right)}^{28}$$

onde

$$\omega_i = \check{\sigma}_b^2 X_i + \check{\sigma}_{bw}(1 - X_i)$$

$$\varepsilon_i = T_i - \check{B}^b X_i - \check{B}^w(1 - X_i)$$

$$\sigma_i^2 = \text{Var}(T_i \mid X_i) = \check{\sigma}_b^2 X_i^2 + \check{\sigma}_w^2(1 - X_i)^2 + \check{\sigma}_{bw} 2X_i(1 - X_i)$$

$$S\left(\check{B}, \check{\Sigma}\right) = \int_{\max\left(0, \frac{T_i - (1 - X_i)}{X_i}\right)}^{\min\left(1, \frac{T_i}{X_i}\right)} N\left(\beta_i^b \mid \check{B}^b + \frac{\omega_i}{\sigma_i^2} \varepsilon_i, \check{\sigma}_b^2 - \frac{\omega_i^2}{\sigma_i^2}\right) d\beta^b$$

A função de distribuição normal truncada univariada é obtida a partir de uma função de distribuição normal univariada, na qual aplica-se uma função indicador $1(\beta_i^b)$ que assume o valor 1 se β_i^b estiver contido nos limites definidos pelo método de “bounds” e igual a zero, caso contrário. A função normalizadora $S\left(\check{B}, \check{\Sigma}\right)$ é utilizada, então, para fazer com

que a área remanescente da distribuição dentro dos limites definidos pelo método de “bounds” tenha integral igual a 1 – propriedade, esta, comum a funções densidade de probabilidade. Desta forma, condicionar a função de distribuição de β_i^b em relação à variável T_i equivale a aplicar o método de “bounds” de Duncan & Davis (1953) em nível de micro-região.

²⁸ O valor esperado e a variância presentes em $P\left(\beta_i^b \mid T_i, \check{\Psi}\right)$ possuem termos de correção dados por

$$\check{E}\left(\beta_i^b \mid T_i, \check{\Psi}\right) = \check{B}^b + \frac{\omega_i}{\sigma_i^2} \varepsilon_i$$

$$\check{\text{Var}}\left(\beta_i^b \mid T_i, \check{\Psi}\right) = \check{\sigma}_b^2 - \frac{\omega_i^2}{\sigma_i^2}$$

A correção para o valor esperado pode ser interpretada como um termo para a correção do erro descrito na equação $\beta_i^b = B^b + \varepsilon_i^b$, de modo a melhorar o cálculo do valor esperado. Analogamente, a correção para a variância pode ser vista como uma medida direta do benefício, para cada seção, de condicionar β_i^b em relação a T_i , uma vez que $\check{\text{Var}}\left(\beta_i^b \mid \check{\Psi}\right) = \check{\sigma}_b^2$.

Neste estágio do método de King (1997), utiliza-se apenas a distribuição para um dos parâmetros de interesse (β^b_i , por exemplo), uma vez que, a partir do momento que este parâmetro for conhecido, β^w_i poderá ser facilmente obtido através da identidade contábil.

$$\beta^w_i = \frac{T_i - \beta^b_i X_i}{1 - X_i}$$

3.2.1.4. Abordagem Bayesiana: Estimação da Função de Distribuição Agregada

Conforme descrito anteriormente, o objetivo final do modelo de King (1997) é estimar $P(\beta^b_i | T_i)$ a partir de $P(\beta^b_i | T_i, \check{\Psi})$, sendo que $\check{\Psi}$ é um vetor desconhecido de parâmetros agregados. Para alcançar este objetivo, King (1997) propõe um método de estimação da função de distribuição posterior agregada, $P(\check{\Psi} | T)$, a qual condensa toda informação disponível sobre o vetor de parâmetros agregados $\check{\Psi}$, baseado na seguinte regra de Bayes:

$$P(\check{\Psi} | T) \propto P(\check{\Psi}) \cdot P(T | \check{\Psi})$$

onde $P(T | \check{\Psi}) = L(\check{\Psi} | T) = \text{função de verossimilhança}$

$$P(\check{\Psi}) = \text{"prior" e}$$

$$P(\check{\Psi} | T) = \text{distribuição posterior agregada}$$

A função de verossimilhança apresentada acima é indireta, pois se baseia apenas nos dados marginais (i.e., ecológicos). Se fosse possível observar β^b_i e β^w_i , a função de verossimilhança seria exatamente uma normal bivariada truncada e $\check{\Psi}$ poderia ser estimado calculando as médias, variâncias e a correlação entre β^b_i e β^w_i .

A função de verossimilhança apresentada por King (1997) é proporcional a uma função densidade de probabilidade na seguinte forma:

$$L(\check{\Psi}|T) \propto \prod_{X_i \in (0,1)} P(T_i|\check{\Psi}) = \prod_{X_i \in (0,1)} N(T_i|\mu_i, \sigma_i^2) \frac{S(\check{B}, \check{\Sigma})}{R(\check{B}, \check{\Sigma})}^{29}$$

onde

$N(T_i|\mu_i, \sigma_i^2)$ = função de distribuição normal univariada com média linear e variância quadrática em X_i , dadas por:

$$\mu_i = \check{B}^b X_i + \check{B}^w (1 - X_i)$$

$$\sigma_i^2 = \text{Var}(T_i|X_i) = (\sigma_w^2) + (2\sigma_{bw} - 2\sigma_w^2)X_i + (\sigma_b^2 + \sigma_w^2 - 2\sigma_{bw})X_i^2$$

e

$$S(\check{B}, \check{\Sigma}) = \int_{\max\left\{0, \frac{T_i - (1 - X_i)}{X_i}\right\}}^{\min\left\{1, \frac{T_i}{X_i}\right\}} N\left(\beta_i^b \left| \check{B}^b + \frac{\omega_i}{\sigma_i^2} \varepsilon_i, \sigma_b^2 - \frac{\omega_i^2}{\sigma_i^2}\right.\right) d\beta^b$$

²⁹ Para se obter a função de verossimilhança acima, parte-se de uma função de distribuição normal bivariada e truncada $P(\beta_i^b, \beta_i^w | \check{\Psi})$ e obtêm-se $P(\beta_i^b, T_i | \check{\Psi})$, uma vez que $\beta_i^w = (T_i - \beta_i^b X_i) / (1 - X_i)$.

O resultado também é uma função de distribuição truncada bivariada normal:

$$P(\beta_i^b, T_i | \check{\Psi}) = \frac{1(T_i)1(\beta_i^b)}{R(\check{B}, \check{\Sigma})} N\left(\beta_i^b, T_i \left| \check{B}^b, \mu_i, \check{\sigma}_b^2, \check{\sigma}_i^2, \omega_i, I(\check{\sigma}_b, \sigma_i)\right.\right).$$

Fatorando esta função de distribuição conjunta em sua função marginal e condicional, tem-se:

$$P(\beta_i^b, T_i | \check{\Psi}) \propto \frac{1(T_i)}{R(\check{B}, \check{\Sigma})} N(T_i|\mu_i, \sigma_i^2) N\left(\beta_i^b \left| \check{B}^b + \frac{\omega_i}{\sigma_i^2} \varepsilon_i, \sigma_b^2 - \frac{\omega_i^2}{\sigma_i^2}\right.\right) 1(\beta_i^b).$$

Por fim, calcula-se a integral em relação à β_i^b e de modo a obter a função de verossimilhança de interesse:

$$P(T_i | \check{\Psi}) \propto \int_{\max\left\{0, \frac{T_i - (1 - X_i)}{X_i}\right\}}^{\min\left\{1, \frac{T_i}{X_i}\right\}} P(\beta_i^b, T_i | \check{\Psi}) d\beta^b \propto \frac{1(T_i)}{R(\check{B}, \check{\Sigma})} N(T_i|\mu_i, \sigma_i^2) S(\check{B}, \check{\Sigma})$$

$$R\left(\overset{\cup}{\mathbf{B}}, \overset{\cup}{\Sigma}\right) = \int \int N\left(\beta^b, \beta^w \mid \overset{\cup}{\mathbf{B}}, \overset{\cup}{\Sigma}\right) d\beta^b d\beta^w$$

O fator $S\left(\overset{\cup}{\mathbf{B}}, \overset{\cup}{\Sigma}\right) / R\left(\overset{\cup}{\mathbf{B}}, \overset{\cup}{\Sigma}\right)$ é a diferença entre a função de verossimilhança apresentada acima e a distribuição normal. Quando o truncamento das micro-regiões for pequeno, esta razão será igual a 1 e a função de verossimilhança será igual à distribuição normal com uma média linear em relação a X, mas com variância quadrática.

Para facilitar a maximização da função de verossimilhança – especificamente, a maximização da integral dupla contida em $R\left(\overset{\cup}{\mathbf{B}}, \overset{\cup}{\Sigma}\right)$, adota-se uma segunda reparametrização visando aproximar os parâmetros aos de uma distribuição normal padrão. Como o algoritmo de cálculo da máxima verossimilhança é invariante a reparametrizações, a reparametrização apresentada abaixo não irá alterar a localização do ponto de máximo na escala de $\overset{\cup}{\Psi}$.

$$\Phi_1 = \frac{\overset{\cup}{\mathbf{B}}^b - 1/2}{\overset{\cup}{\sigma}_b^2 + 1/4} \quad \Phi_2 = \frac{\overset{\cup}{\mathbf{B}}^w - 1/2}{\overset{\cup}{\sigma}_w^2 + 1/4} \text{ }^{30}$$

$$\Phi_3 = \ln\left(\overset{\cup}{\sigma}_b\right) \quad \Phi_4 = \ln\left(\overset{\cup}{\sigma}_w\right) \text{ }^{31}$$

$$\Phi_5 = \frac{1}{2} \ln\left(\frac{1 + \overset{\cup}{\rho}}{1 - \overset{\cup}{\rho}}\right) \text{ }^{32}$$

Uma vez feita a reparametrização, a função de distribuição posterior agregada $P(\Phi|T)$ é estimada em um processo de 2 estágios condizente com a regra de Bayes já apresentada, e aqui repetida (apenas substituindo $\overset{\cup}{\Psi}$ por Φ em função da reparametrização).

$$P(\Phi|T) \propto P(\Phi).P(T|\Phi)$$

$$\text{onde } P(T|\Phi) = L(\Phi|T)$$

³⁰ Esta reparametrização é feita para compensar a correlação que normalmente existe entre $\overset{\cup}{\mathbf{B}}^b$ e $\overset{\cup}{\sigma}_b^2$.

³¹ Esta reparametrização tem a vantagem de não ser limitada e de ser mais simétrica.

³² Esta é a transformação Z de Fisher.

No primeiro estágio, a função de verossimilhança é maximizada, de modo a se obter as estimativas pontuais de $\hat{\Phi}$ e $\hat{Var}(\Phi)$, ou seja, a média e a variância dos parâmetros de parcial interesse reparametrizados.

No estágio seguinte, dois procedimentos alternativos podem ser utilizados na estimação de $P(\Phi|T)$.

1. Utilizar uma distribuição normal com parâmetros $\hat{\Phi}$ e $\hat{Var}(\Phi)$ para aproximar a distribuição posterior agregada $P(\Phi|T)$, i.e., considerar $P(\Phi)=1$.
2. Utilizar o algoritmo de “Importance Sampling”: neste algoritmo a distribuição normal $\Phi \approx N\left(\Phi|\hat{\Phi}, \hat{Var}(\hat{\Phi})\right)$ é utilizada como uma primeira tentativa para o cálculo da distribuição posterior agregada, seguida por sucessivas aproximações até alcançar $P(\Phi|T)$.³³

Na prática o software desenvolvido e disponibilizado por King em <http://gking.harvard.edu/stats.shtml>, segue o procedimento de “importance sampling”.

3.2.1.5. Abordagem Bayesiana: Estimação dos Parâmetros por Micro-região

³³ O algoritmo de “importance sampling” é calculado, na prática, através das seguintes etapas:

1. Retira-se uma amostra aleatória dos parâmetros Φ usando a aproximação normal. Este procedimento é feito através da maximização da função de verossimilhança - de modo a se obter $\hat{\Phi}$ e $\hat{Var}(\hat{\Phi})$ - e da extração aleatória de um valor de Φ de uma distribuição normal com parâmetros $\hat{\Phi}$ e $\hat{Var}(\hat{\Phi})$.
2. Reparametriza-se os valores simulados do vetor de parâmetros Φ nas versões não-truncadas dadas por $\hat{\Psi}$.
3. Utiliza-se o “importance sampling” para melhorar a aproximação da distribuição posterior pela normal através de 2 passos:
 - i. Calcula-se a razão de importância, que é a razão entre o valor da função de verossimilhança e da aproximação normal avaliadas no vetor de parâmetros simulados e normalizada entre zero e um para as sucessivas extrações aleatórias.
 - ii. Aceite o valor extraído $\hat{\Psi}$ com probabilidade igual à razão de importância. Ou seja, o valor extraído é aceito se for maior do que um valor retirado aleatoriamente de uma distribuição uniforme dentro do intervalo [0,1]. Se o valor extraído for rejeitado, retorna-se a etapa 1.

Conforme dito anteriormente, o objetivo final do modelo de King (1997) é estimar a distribuição posterior por micro-região $P(\beta_i^b|T)$, a qual é condicionada apenas à variável conhecida T, a partir da função de distribuição $P(\beta_i^b|T, \Phi)$, a qual é condicionada à variável conhecida T e ao vetor de parâmetros desconhecidos Φ .

Matematicamente, a função de distribuição posterior por micro-região $P(\beta_i^b|T)$ pode ser obtida através do cálculo da expressão abaixo:

$$\begin{aligned} P(\beta_i^b|T) &\propto \int_{-\infty}^{\infty} P(\beta_i^b, \Phi|T) d\Phi \\ &\propto \int_{-\infty}^{\infty} P(\Phi|T) P(\beta_i^b|T, \Phi) d\Phi \\ &\propto \int_{-\infty}^{\infty} P(\Phi) \prod_{i=1}^P N(T_i|\mu_i, \sigma_i^2) \frac{S\left(\overset{\cup}{\mathbf{B}}, \overset{\cup}{\Sigma}\right)}{R\left(\overset{\cup}{\mathbf{B}}, \overset{\cup}{\Sigma}\right)} \cdot TN\left[T_i \left| \overset{\cup}{\mathbf{B}}_i + \frac{\omega_i}{\sigma_i^2} \varepsilon_i, \sigma_i^2 - \frac{\omega_i^2}{\sigma_i^2}\right.\right] d\Phi \end{aligned}$$

Como a solução algébrica da equação acima não é trivial, King (1997) propõe que $P(\beta_i^b|T)$ seja obtido através de simulações baseadas na seguinte regra de Bayes.

$$P(\beta_i^b|T) \propto \int_{-\infty}^{\infty} P(\Phi|T) P\left(\beta_i^b|T, \overset{\cup}{\Psi}\right) d\overset{\cup}{\Psi}$$

onde a distribuição posterior agregada $P(\Phi|T)$ é utilizada como “prior” no cálculo da distribuição posterior por micro-região $P(\beta_i^b|T)$.

As estimativas pontuais de β_i^b , para cada uma das micro-regiões, são obtidas a partir do cálculo das médias de N valores de β_i^b retirados aleatoriamente de $P(\beta_i^b|T)$. Os desvios-padrão estimados de β_i^b , para cada uma das micro-regiões, são calculados, analogamente, com base nos erros-padrão de N valores de β_i^b retirados de $P(\beta_i^b|T)$.

Para se extrair os valores de β_i^b de $P(\beta_i^b|T)$, utiliza-se um algoritmo que, em linhas gerais, é composto pelos seguintes passos:

1. Extrai-se um valor de Φ da distribuição $P(\Phi|T)$, denominado $\overset{\circ}{\Phi}$.

2. Reparametrização Φ na forma não truncada, de modo a se obter Ψ .
3. Insere-se Ψ em $P(\beta_i^b | T, \Psi)$ e retira-se um valor de β_i^b aleatoriamente de $P(\beta_i^b | T, \Psi)$ ³⁴.

As etapas adotadas no algoritmo acima equivalem a retirar um valor de β_i^b diretamente da distribuição posterior por micro-região $P(\beta_i^b | T)$.

Conforme descrito anteriormente, este procedimento deve ser repetido N vezes para cada uma das P micro-regiões de interesse, de modo que seja possível calcular a média e o desvio-padrão de β_i^b através das seguintes fórmulas:

$$\hat{\beta}_i^b = \frac{1}{K} \sum_{k=1}^K \beta_i^{b(k)*}$$

onde $\beta_i^{b(k)*}$ é a k-ésima simulação.

$$SE(\beta_i^b) = \sqrt{\frac{1}{K} \sum_{k=1}^K \left(\beta_i^{b(k)*} - \hat{\beta}_i^b \right)^2}$$

Por fim, uma vez calculados os valores de β_i^b para cada uma das micro-regiões, os valores dos parâmetros de interesse β_i^w são calculados através da identidade contábil.

$$\beta_i^w = \frac{T_i - \beta_i^b X_i}{1 - X_i}$$

3.2.1.6. Abordagem Bayesiana: Estimação dos Parâmetros Agregados Ponderados

Os parâmetros agregados de interesse são calculados a partir da média dos parâmetros por micro-região - obtidos na seção anterior - ponderados pelo número de indivíduos

³⁴ Na prática, insere-se Ψ (o vetor de parâmetros obtido através do algoritmo de “importance sampling” descrito na nota de rodapé anterior) na distribuição posterior condicional de β_i^b , $P(\beta_i^b | T, \Psi)$, e retira-se

um valor de β_i^b através dos seguintes passos:

- i. Calcula-se σ_i^2 , ω_i e ϵ_i através das fórmulas apresentadas na página 23.
- ii. Substitui-se estes valores na distribuição posterior normal condicional truncada e retira-se um valor de β_i^b .

detentores de uma dada característica demográfica qualquer (como pertencer à raça negra, por exemplo) em cada uma das P micro-regiões, conforme fórmula abaixo.

$$B^b = \frac{\sum_{i=1}^p N_i^b \beta_i^b}{N^b}$$

onde $N^b = \sum_{i=1}^p N_i^b$, ou seja, o total de indivíduos pertencentes à categoria demográfica em questão (neste caso, negros) em todas as micro-regiões.

Este procedimento deve ser repetido K vezes, de modo a haver K valores calculados para os parâmetros agregados de interesse. Finalmente, as estimativas pontuais e os desvios-padrão dos parâmetros de interesse agregados são calculados de modo similar aos parâmetros das micro-regiões, ou seja, através da média dos K valores calculados e do erro-padrão destes K valores.

$$\hat{B}^b = \frac{1}{K} \sum_{k=1}^K B^{b(k)}$$

$$SE(\hat{B}^b) = \sqrt{\frac{1}{K} \sum_{k=1}^K (B^{b(k)} - \hat{B}^b)^2}$$

O modelo de King (1997) foi aplicado, analogamente ao modelo de Goodman (1953), aos dados das eleições realizadas em 1996 e 2000 para a prefeitura de São Paulo, em segundo turno. Na primeira “cross section”, foram relacionados, individualmente, os percentuais de votos obtidos pela então candidata Luiza Erundina (PT) na eleição de 1996, em segundo turno, a 3 variáveis demográficas, sejam elas: sexo, idade e escolaridade. Na segunda “cross section”, por sua vez, o modelo de King (1997) foi aplicado relacionando os percentuais de votos obtidos pela então candidata Marta Suplicy (PT) em 412 áreas de ponderação do município de São Paulo às variáveis demográficas: sexo, idade, raça, religião, escolaridade, rendimentos totais e origem migratória.

O modelo proposto por King (1997) estima os parâmetros agregados ponderados (para o município como um todo) e os parâmetros para cada uma das micro-regiões de interesse (neste caso, para cada uma das áreas de ponderação de São Paulo).

4. Descrição dos dados

Os determinantes sociais do voto serão analisados, no presente estudo, através de dois modelos específicos: o modelo de Goodman (1953) e o modelo de King (1997). Para estimar estes modelos serão utilizadas variáveis demográficas dos eleitores (tais como sexo, renda e raça, por exemplo) e dados eleitorais por micro-região (i.e, o percentual de votos obtidos por um determinado candidato em uma determinada micro-região na eleição ocorrida no ano t).

Os modelos de King (1997) e de Goodman (1953) serão estimados, individualmente, para as eleições de 1996 e 2000, com base em variáveis demográficas e eleitorais disponíveis para estes anos para o município de São Paulo. Nesta seção, estas variáveis serão descritas e será feita uma breve apresentação espacial destas mesmas variáveis através de instrumentos gráficos.

4.1. Variáveis Demográficas: Eleição de 1996

Os dados demográficos utilizados para estimar os modelos de Inferência Ecológica referentes à eleição municipal de 1996 em São Paulo foram obtidos na Contagem 1996 do Instituto Brasileiro de Geografia e Estatística (IBGE)³⁵. Entre as diversas variáveis demográficas contidas na Contagem 1996 do IBGE, foram selecionadas as variáveis sexo, idade e escolaridade (em anos de estudo completos) por área de ponderação.

Para possibilitar a estimação do modelo de Goodman e do modelo de King as variáveis demográficas a serem analisadas foram divididas de forma binária.³⁶ Com relação à variável demográfica sexo, utilizamos o percentual de pessoas do sexo feminino por área de ponderação, visando captar o impacto do voto feminino na votação obtida pela ex-prefeita Luiza Erundina.

A variável idade considerou o percentual de pessoas por área de ponderação que têm idade entre 16 e 39 anos (média de idade dos eleitores do município de São Paulo). O objetivo

³⁵ A Contagem do IBGE possui uma metodologia distinta da utilizada nos Censos Demográficos. Enquanto o Censo Demográfico é construído a partir de uma combinação de dados amostrais e populacionais, a Contagem Demográfica é um procedimento intercensitário baseado em uma amostra mais restrita da população. O objetivo da Contagem, segundo o IBGE, é fornecer "...informações para a revisão e atualização das estimativas de população no período intercensitário e para a criação da base cadastral do próximo censo".

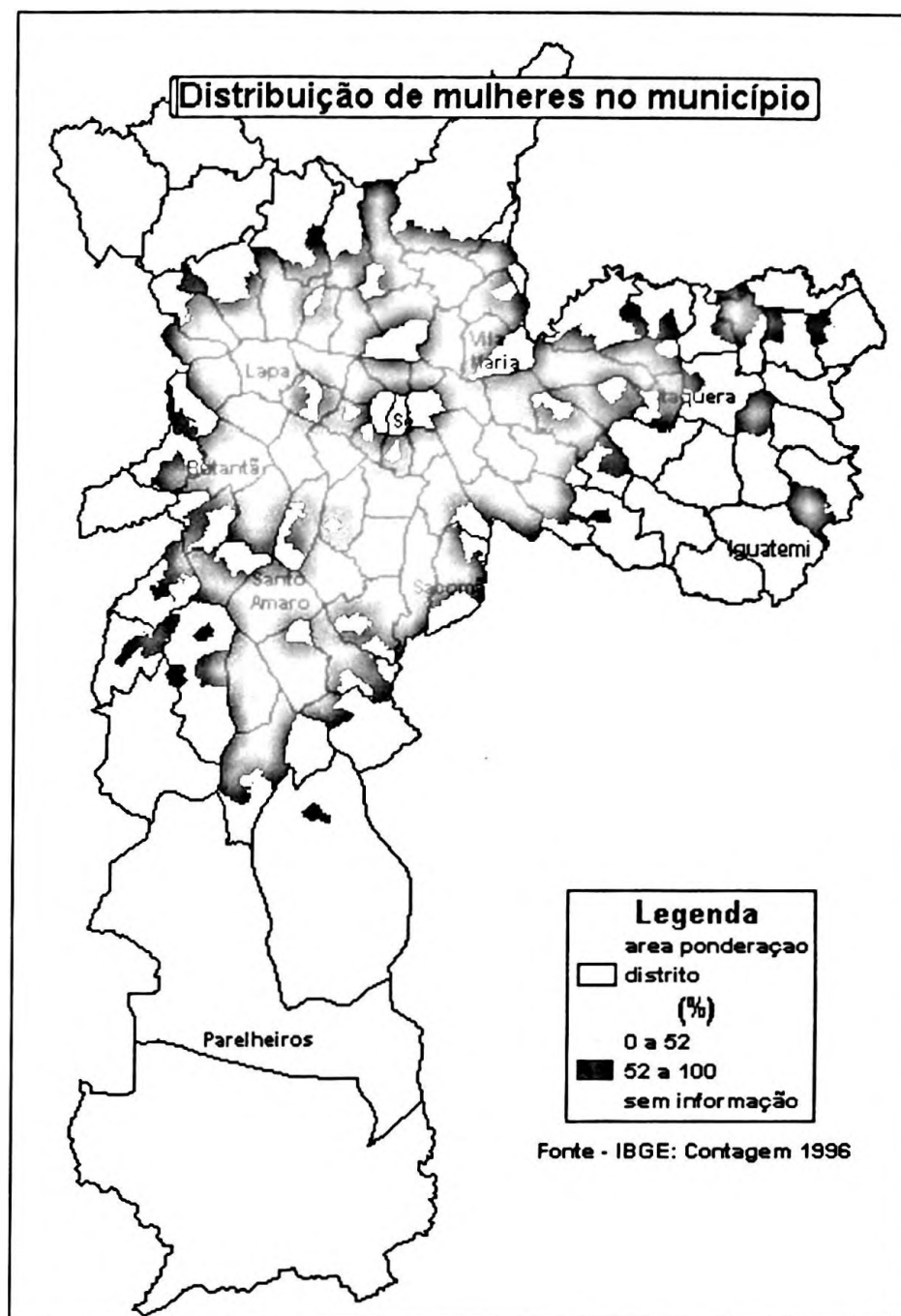
³⁶ Apesar da arbitrariedade aparente destas divisões binárias, estas são função do interesse investigativo do pesquisador.

era tentar corroborar uma hipótese de que os eleitores mais jovens tendem a ter uma simpatia maior por partidos de esquerda, enquanto que eleitores mais maduros teriam uma preferência por candidatos de perfil mais conservador.

A variável demográfica escolaridade foi dividida de modo análogo à variável idade, considerando o percentual de pessoas por área de ponderação que está abaixo da média da escolaridade do eleitorado, que, neste caso, é de 8 anos. Este critério tenta captar um apoio supostamente maior dos eleitores menos instruídos aos candidatos de esquerda.

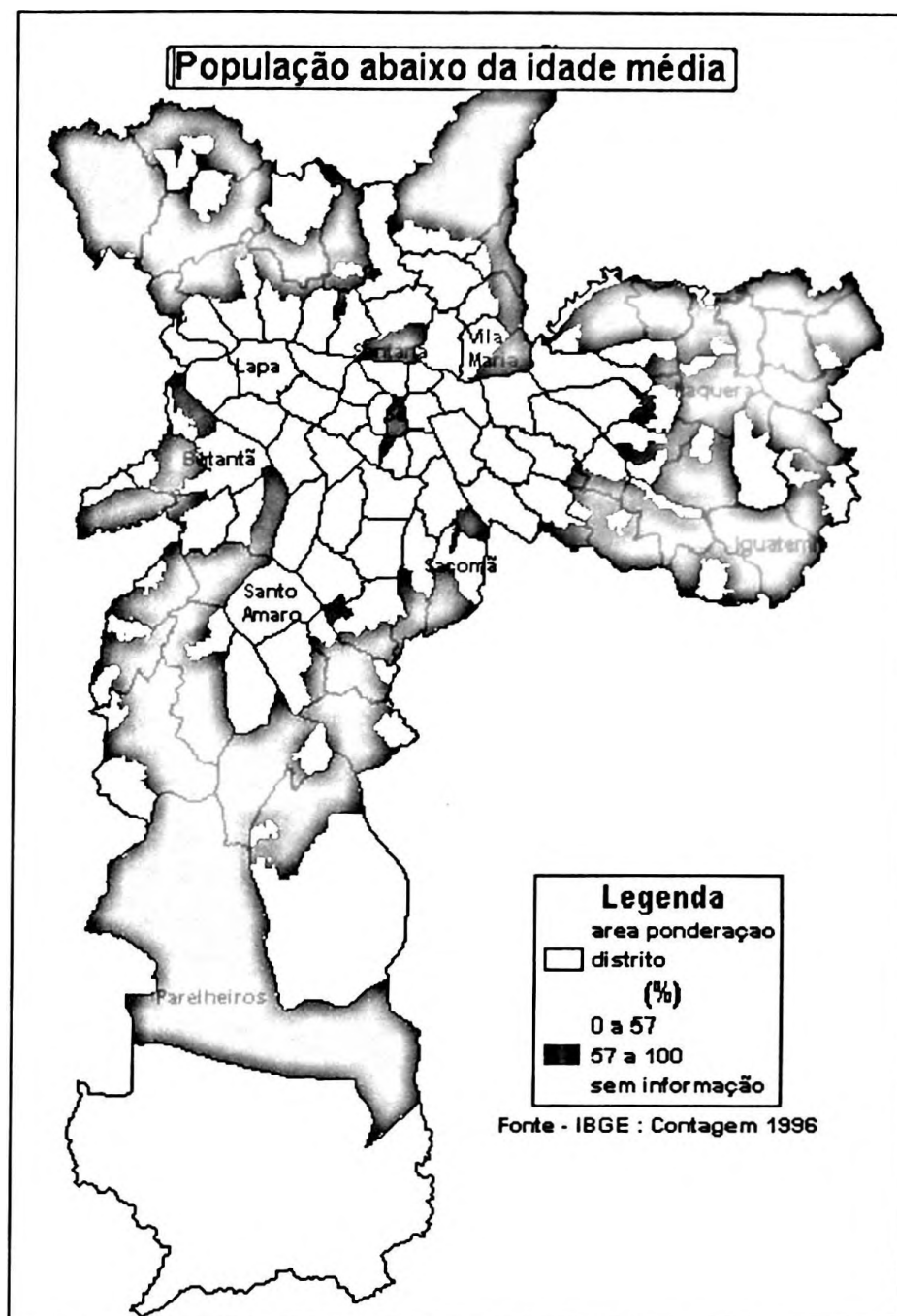
Abaixo, são apresentados os gráficos espaciais produzidos para as variáveis sexo, idade e escolaridade para o município de São Paulo, em 1996, de acordo com as definições acima. Na figura 4, são destacadas as áreas de ponderação que possuem um percentual de mulheres acima da média. Na figura 5, salienta-se as áreas de ponderação no mapa do município de São Paulo em que existe uma maior concentração de indivíduos “jovens”, com idade entre 16 e 39 anos. Enquanto que, no último gráfico, destacam-se as áreas de ponderação em que existe uma maior concentração de indivíduos cuja escolaridade é inferior à média do município.

Figura 4 – Distribuição de mulheres no município³⁷



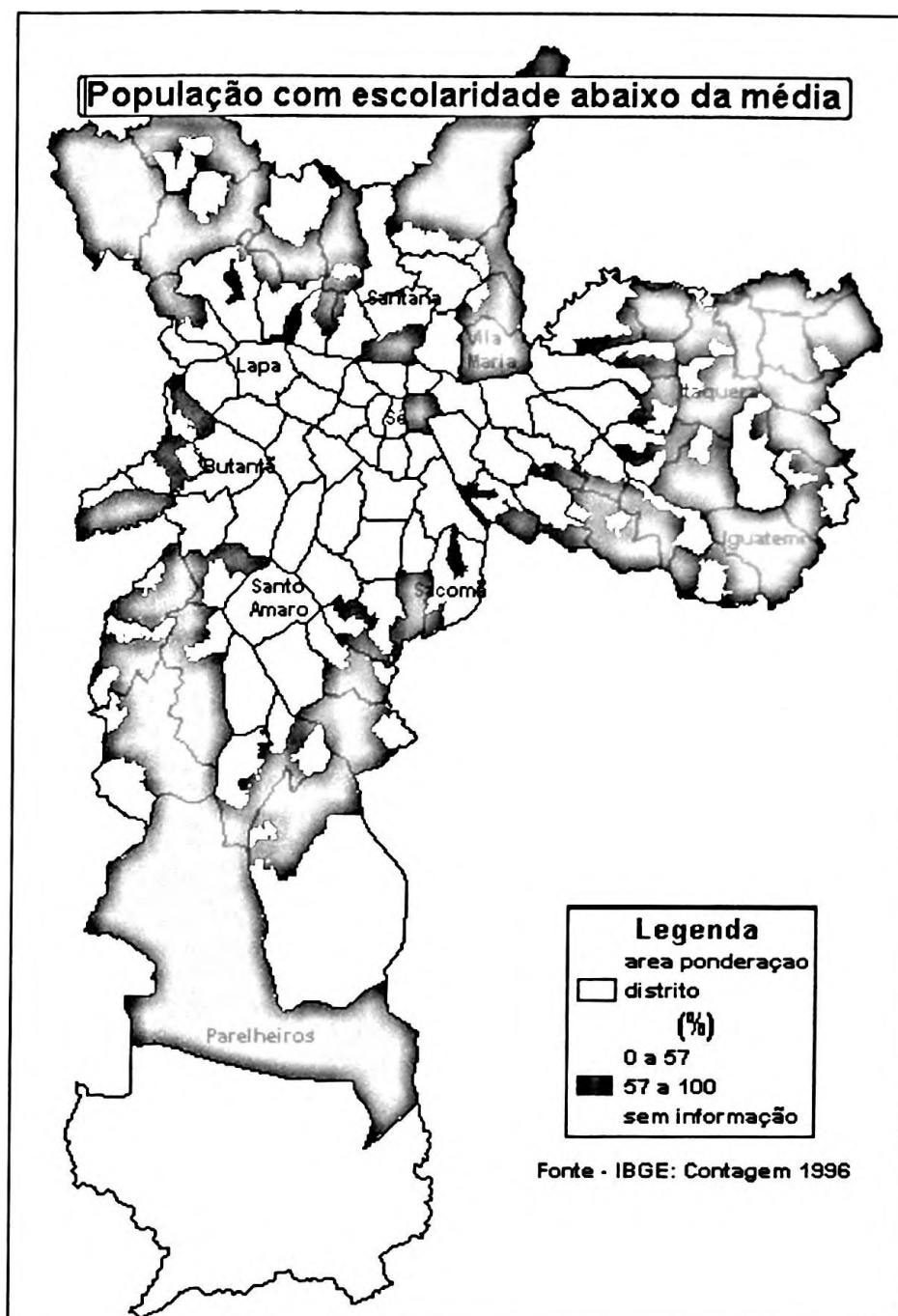
³⁷ Áreas de ponderação com % de mulheres abaixo (cor amarela) e acima (cor marrom) da média do município.

Figura 5 – Distribuição da população por idade³⁸



³⁸ A média de idade da população maior de 16 anos no município de São Paulo na Contagem de 1996 do IBGE foi de 39 anos. A Figura 5 mostra as áreas de ponderação em que há uma maior (cor marrom) ou uma menor (cor amarela) concentração de indivíduos “jovens”.

Figura 6 – Distribuição da população com escolaridade abaixo da média do município³⁹



As figuras acima mostram uma concentração da população com idade e escolaridade inferior à média do município em regiões periféricas, enquanto que se observa uma concentração maior de mulheres na região central da cidade.

4.2. Variáveis Demográficas: Eleição de 2000

Os modelos referentes à eleição de 2000 se basearam em dados disponibilizados pelo IBGE no Censo Demográfico de 2000, por área de ponderação. Foram utilizadas sete

³⁹ A escolaridade média no município de São Paulo era de oito anos em 1996. A Figura 6, acima, mostra as áreas de ponderação com uma maior concentração de indivíduos com escolaridade abaixo desta média.

variáveis demográficas: sexo, idade, raça, religião, escolaridade, renda e origem migratória. Novamente, estas variáveis foram definidas de modo binário, com o objetivo de estimar os modelos padrões de King (1997) e de Goodman (1953).

As variáveis demográficas sexo, idade e escolaridade foram definidas de modo análogo ao feito anteriormente para os dados da Contagem 1996. Para a variável raça, por sua vez, foi adotado como critério o percentual da população que é negra, parda ou indígena. Desta forma, brancos e amarelos foram mantidos como base de comparação. Este critério se justifica, neste caso, pelo interesse em observar um apoio supostamente maior de minorias a candidatos de esquerda.

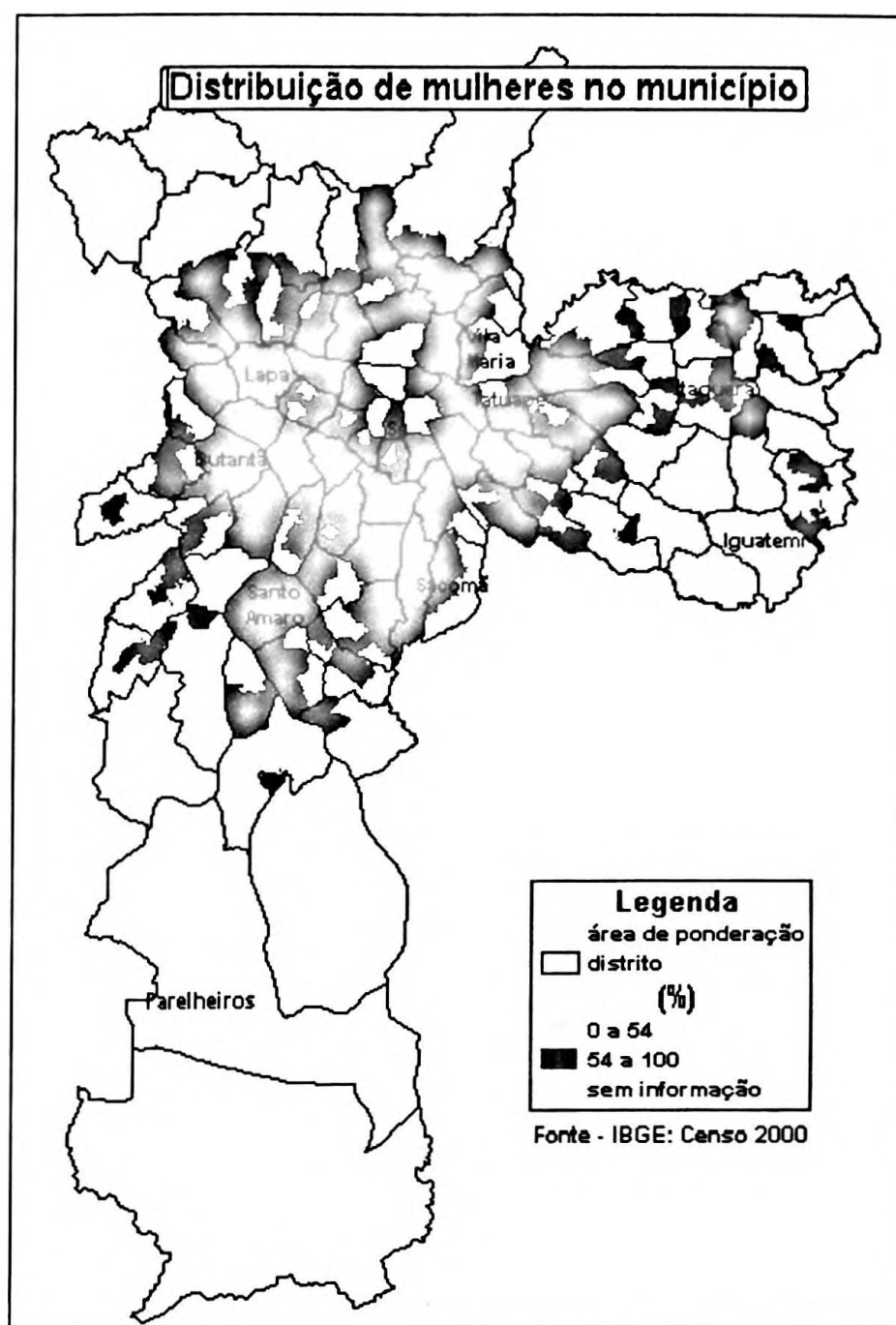
A variável demográfica religião, por outro lado, possui inúmeras classificações, mas escolheu-se trabalhar apenas com as categorias amplas “Católica” (que inclui, entre outras, a igreja Católica Apostólica Romana, Católica Ortodoxa e Ortodoxa Cristã) e “Evangélica” (que inclui igrejas como a Igreja Universal do Reino de Deus, Igreja Evangélica Batista, Igreja Evangélica Presbiteriana, etc). Para o estudo foi utilizado o percentual da população, por área de ponderação, que afirmou ser seguidor da Igreja Evangélica, visando captar o apoio dos evangélicos na eleição da candidata Marta Suplicy para a prefeitura da cidade.

A variável demográfica renda (categoria Total dos Rendimentos do IBGE) considera a renda total das pessoas, em Reais, tendo como base Julho/2000. Esta variável foi dividida de forma análoga às variáveis escolaridade e idade, ou seja, considerou o percentual de pessoas por área de ponderação que está abaixo da renda média da população (de R\$835,00 para o município de São Paulo). Novamente, o objetivo desta classificação foi tentar captar o apoio dos eleitores mais pobres da periferia paulistana à candidata de esquerda Marta Suplicy (PT).

Por fim, a variável demográfica origem migratória considerou o percentual das pessoas, por área de ponderação, que são imigrantes no município de São Paulo. Evidentemente, o objetivo desta variável é captar o apoio dos imigrantes, supostamente de extratos sociais menos favorecidos, à então candidata Marta Suplicy.

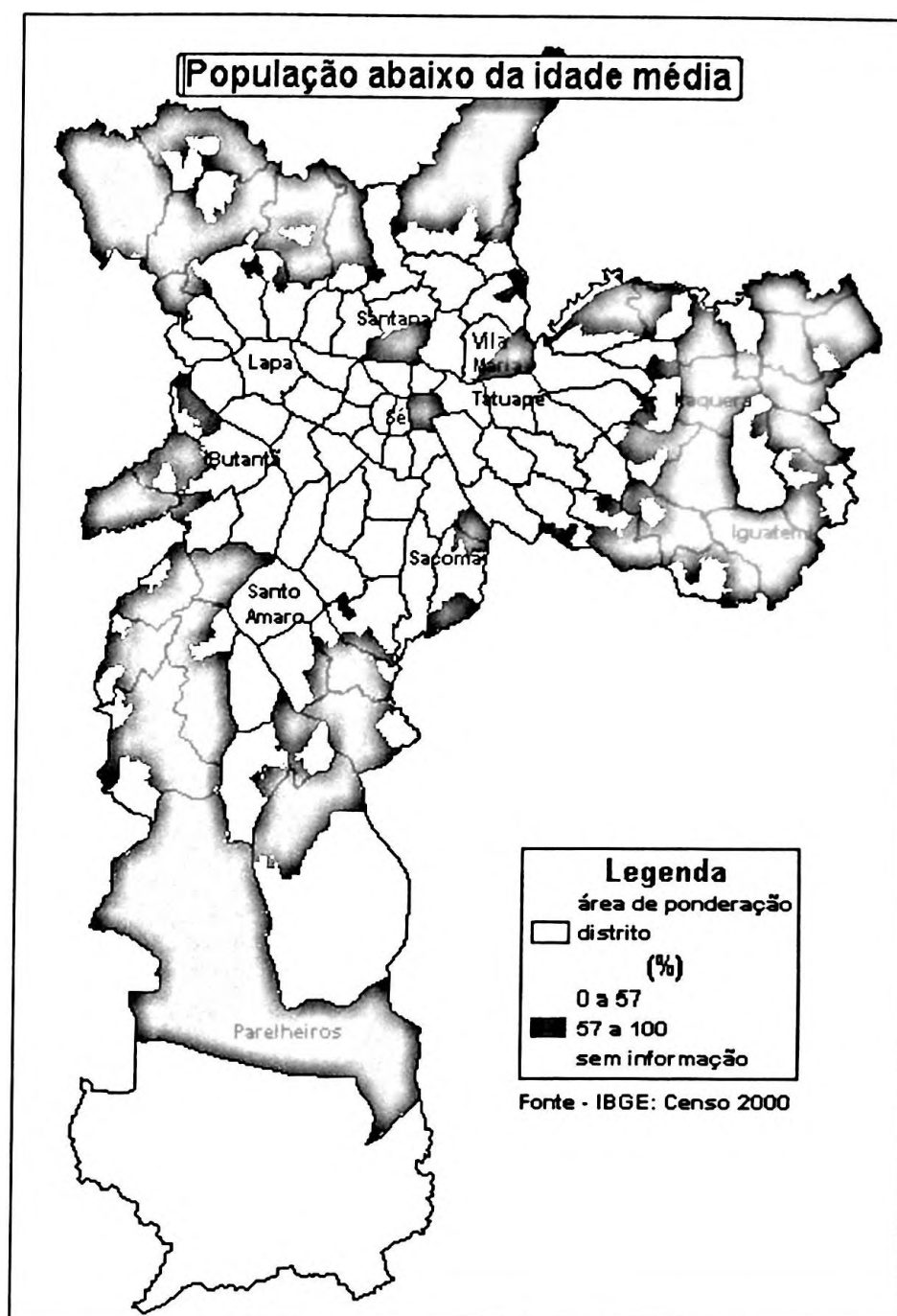
A seguir, são apresentados os gráficos espaciais para o município de São Paulo das sete variáveis demográficas de interesse, conforme definidas acima.

Figura 7 – Distribuição de mulheres no município⁴⁰



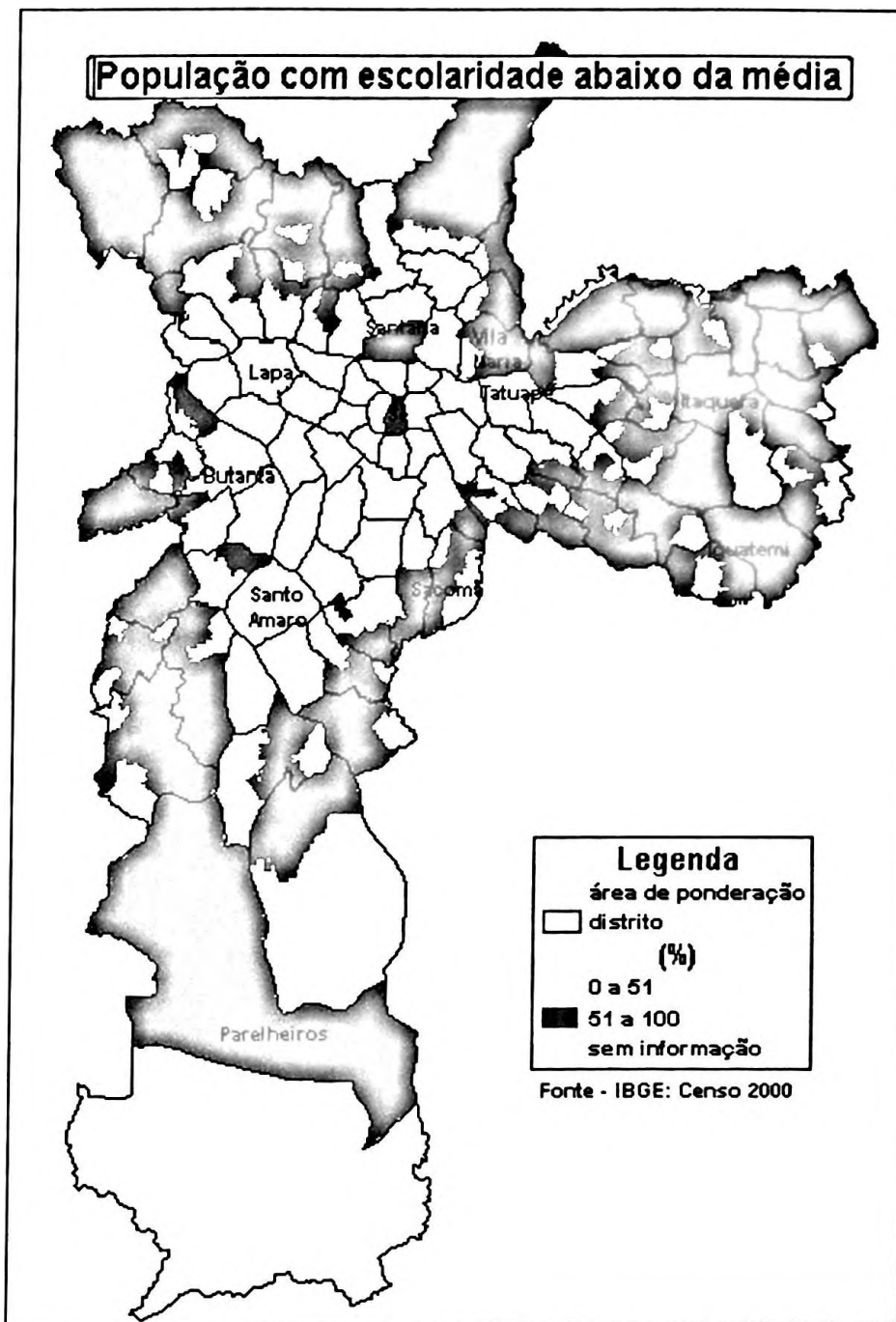
⁴⁰ Áreas de ponderação com % de mulheres abaixo (cor amarela) e acima (cor marrom) da média do município.

Figura 8 – Distribuição da população por idade⁴¹



⁴¹ A Figura 8 mostra as áreas de ponderação em que há uma maior (cor marrom) ou uma menor (cor amarela) concentração de indivíduos “jovens”.

Figura 9 – Distribuição da população com escolaridade abaixo da média do município⁴²



⁴² A Figura 9, acima, mostra as áreas de ponderação com uma maior concentração de indivíduos com escolaridade abaixo da média no ano 2000.

**Figura 11 – Distribuição da população por religião –
% de protestantes por área de ponderação**

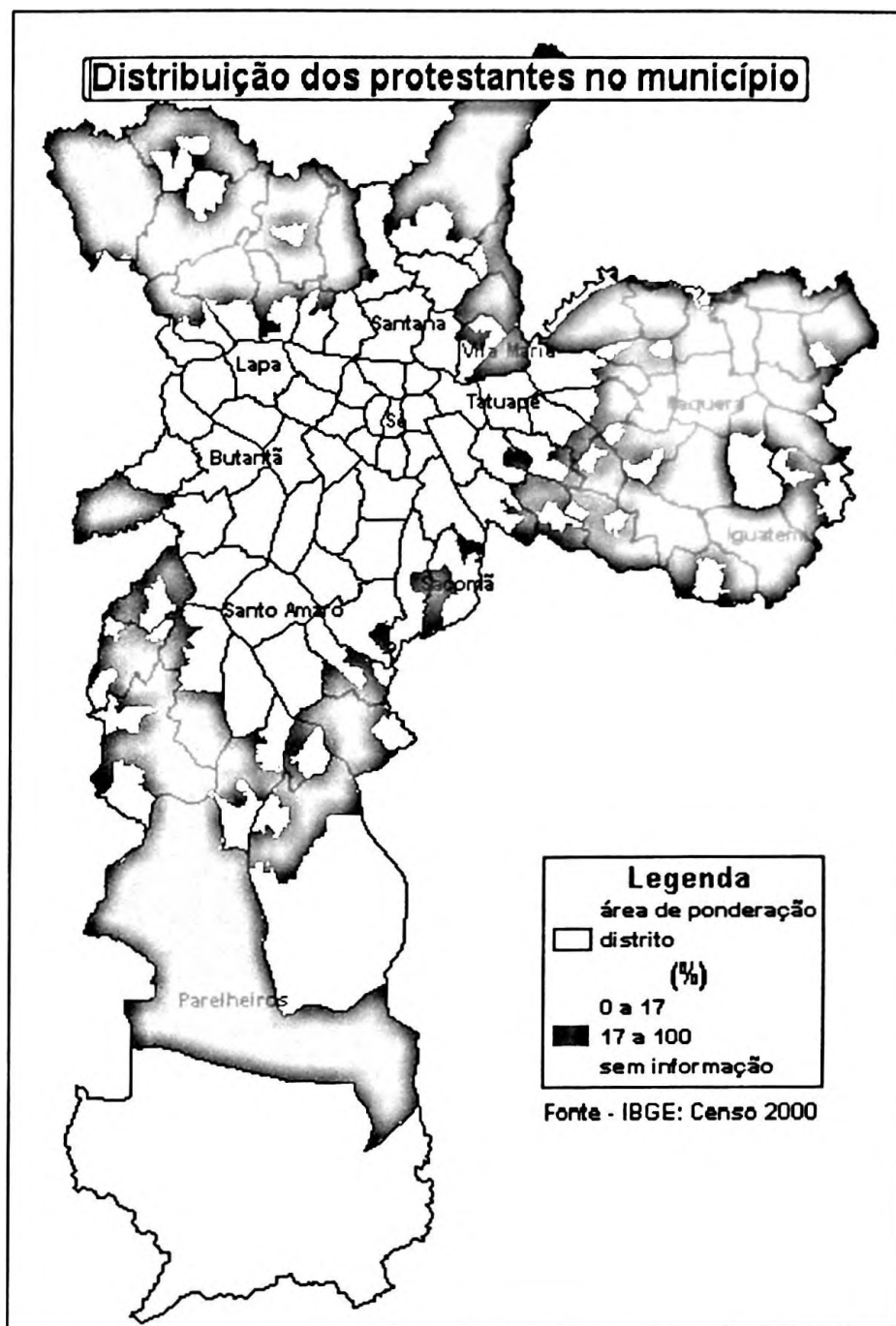
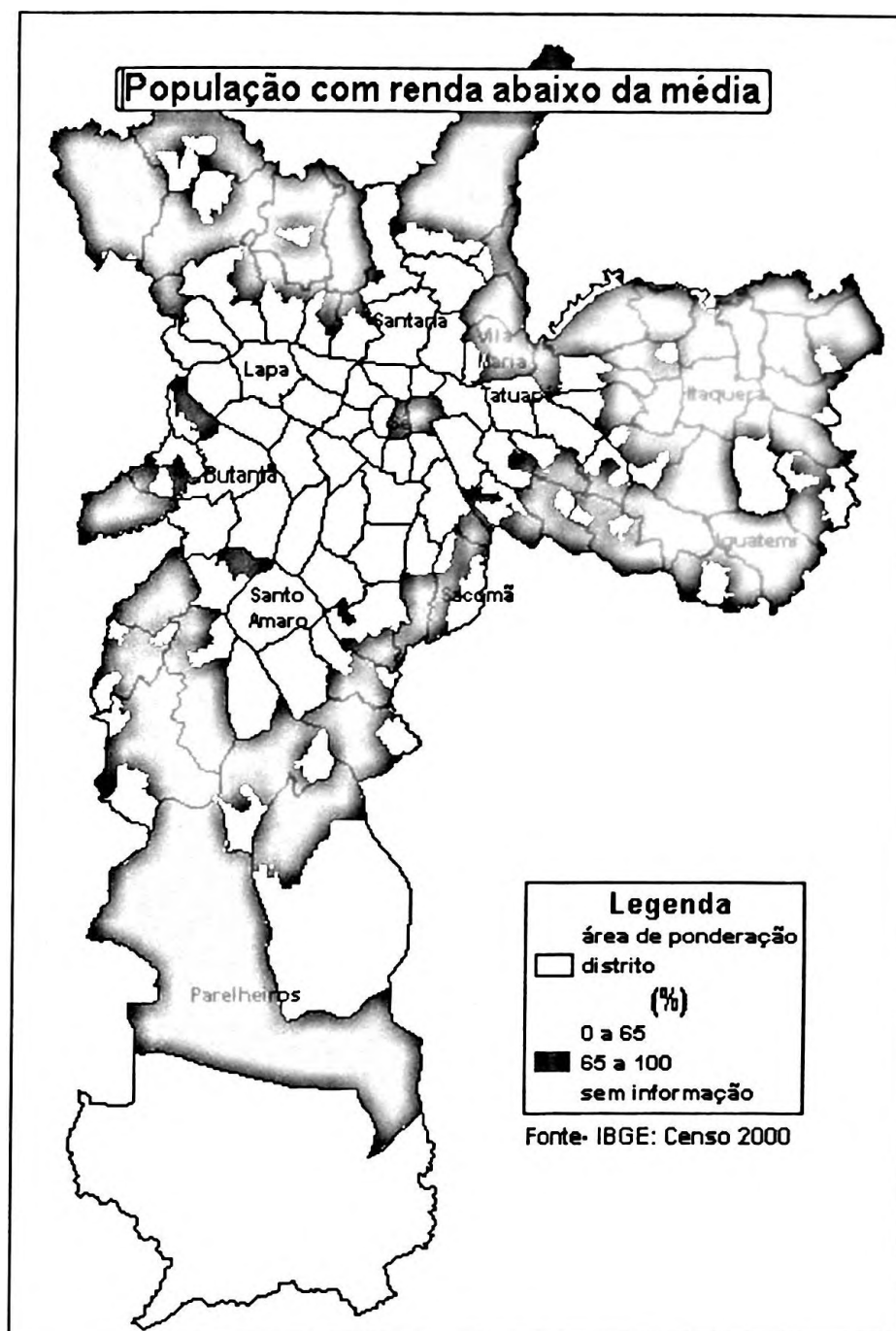


Figura 12 – Distribuição da população por renda – % da população com renda inferior à média do município⁴³



⁴³ A Figura 12 mostra as áreas de ponderação com uma maior concentração de indivíduos com renda abaixo da média do município de São Paulo no ano 2000.

município de São Paulo se localizam espacialmente em regiões de periferia e em bairros centrais próximos à praça da Sé e à região da avenida Paulista.

4.3. Variáveis Eleitorais: Eleição de 1996

Os dados referentes às eleições, em segundo turno, para a prefeitura do município de São Paulo ocorridas em 1996 e 2000 foram disponibilizados pelo TRE-SP, por seção eleitoral.

Os modelos a serem estimados nesta dissertação têm como característica comum o fato de serem binários. Dada esta restrição, decidiu-se pela utilização de variáveis eleitorais relativas ao segundo turno das eleições realizadas no município de São Paulo nos anos de 1996 e 2000. Entre as críticas que podem ser endereçadas a esta escolha, estão os argumentos de que, (1) os eleitores, no segundo turno, não revelam suas verdadeiras preferências pelos candidatos e – associado a isso – (2) que, no segundo turno eleitoral, prevalece o “voto útil” dos eleitores. A alternativa, neste caso, seria trabalhar com dados do primeiro turno, seja dividindo os partidos políticos que disputaram o primeiro turno entre “direita” e “esquerda” ou dividindo os partidos que disputaram o primeiro turno de modo polarizado, como, por exemplo, “PT” e “outros partidos”. Apesar de estas duas alternativas também se colocarem, a opção pelos dados eleitorais de segundo turno pareceu mais natural pelo caráter binário do modelo e deste turno da eleição. Adicionalmente, existe uma longa tradição em ciência política que argumenta que os eleitores sempre agem estrategicamente ao votar, de modo que, em nenhum momento os eleitores “revelam suas preferências” de modo efetivo, seja no segundo turno de uma eleição ou mesmo em um primeiro turno eleitoral. Deve-se destacar que a questão do voto estratégico é mais sensível em (1) eleições majoritárias - como as eleições para prefeito - do que em eleições proporcionais, (2) em sistemas em que existem listas conjuntas (no caso de eleições proporcionais) e (3) em coalizões partidárias, conforme destacado por Cox (1999).

Para se viabilizar a comparação entre os dados eleitorais e demográficos foi necessário compatibilizar as unidades espaciais de análise em um procedimento em dois estágios. Primeiramente foram agregados os dados das seções eleitorais em locais de votação (uma escola, por exemplo); no segundo estágio foi feita uma sobreposição espacial dos diversos locais de votação e das áreas de ponderação existentes no município de São Paulo com o

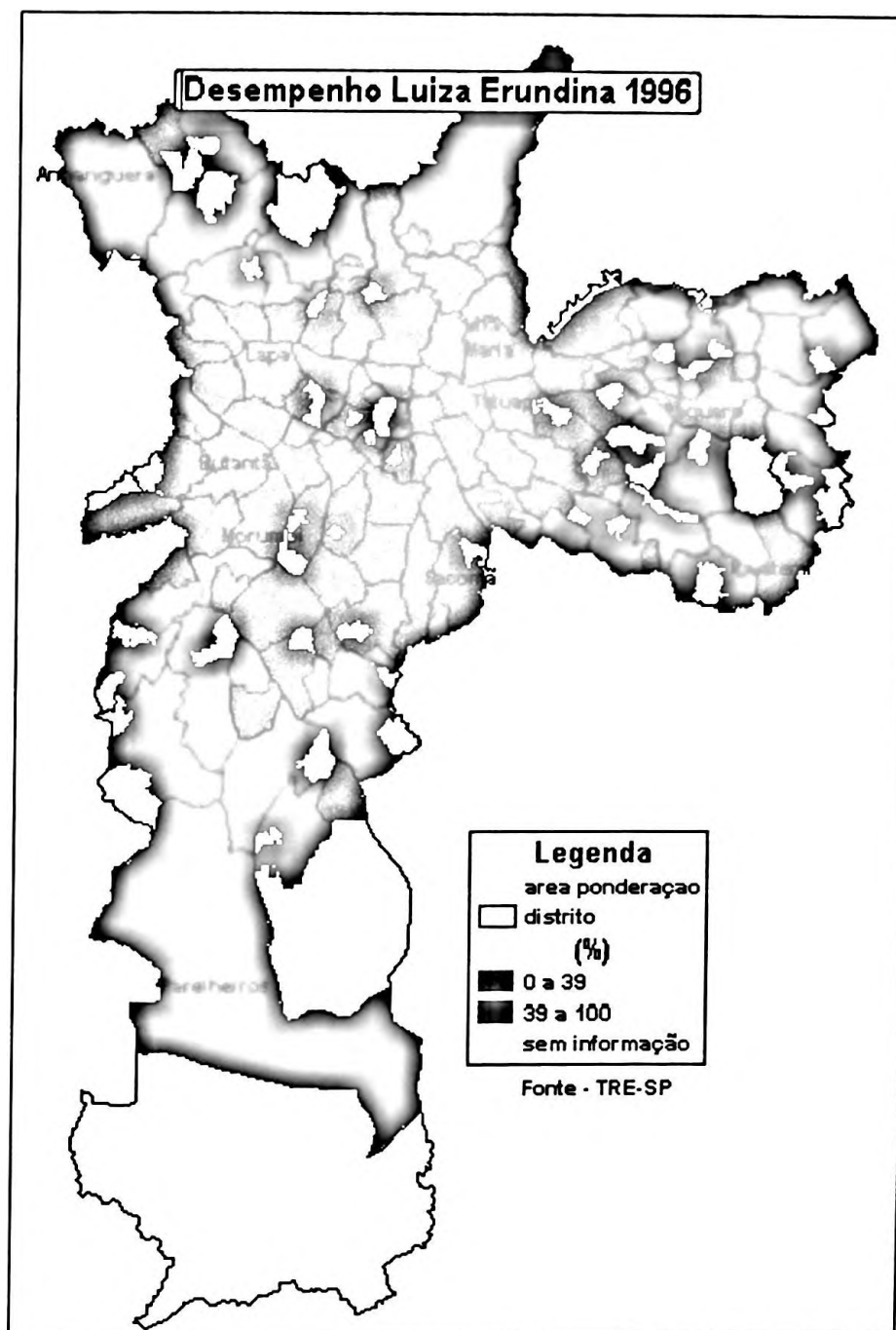
objetivo de agregar os dados eleitorais em uma unidade básica de análise: a área de ponderação⁴⁴.

O procedimento descrito acima gerou dados demográficos e eleitorais para 398 áreas de ponderação diferentes situadas em 94 distritos do município de São Paulo em 1996. Desta forma, existem casos em que um distrito está associado a apenas uma área de ponderação (bairro da Sé, por exemplo), assim como casos em que até 9 áreas de ponderação estão contidas em um mesmo distrito (caso de Sapopemba). Este procedimento considera, por hipótese, que os domicílios eleitorais das pessoas estão localizados nas mesmas áreas de ponderação onde estas residem, ou seja, casos em que os eleitores votam em locais distantes de suas residências são tratados como exceção.

Abaixo, é apresentado o gráfico espacial das áreas de ponderação do município de São Paulo em que a então candidata Luiza Erundina obteve uma performance eleitoral superior à sua média.

⁴⁴ Para a realização deste procedimento de agregação espacial contamos com a colaboração dos analistas do Centro de Estudos da Metrópole (CEM-CEBRAP).

Figura 14 – Distribuição dos votos obtidos por Luiza Erundina (PT) nas eleições de 1996 – Áreas de Ponderação com votação superior e inferior à média da candidata



4.4. Variáveis Eleitorais: Eleição de 2000

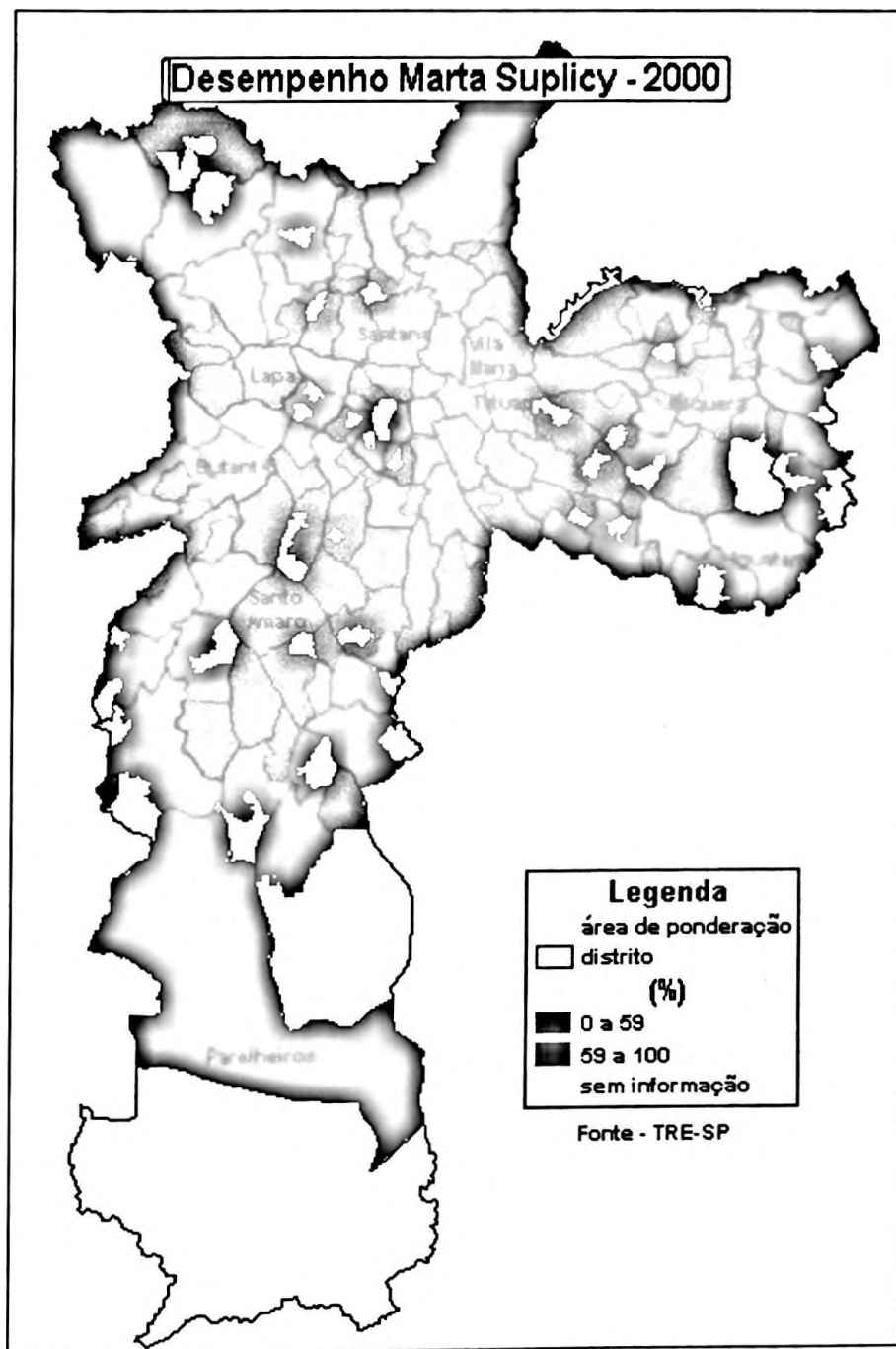
Os dados referentes à eleição de 2000 foram compatibilizados espacialmente de modo análogo aos dados eleitorais de 1996. Esta compatibilização, por sua vez, gerou 412 áreas de ponderação nos mesmos 94 distritos do município de São Paulo⁴⁵.

O gráfico espacial das áreas de ponderação do município de São Paulo em que a prefeita Marta Suplicy recebeu uma votação superior à sua média de votos é apresentado abaixo.

⁴⁵ Cabe destacar que consideramos apenas as áreas de ponderação nas quais existiam pelo menos um local de votação no ano 2000. O total de áreas de ponderação no município era de 456 em 2000.

Ao contrário do ocorrido nas eleições de 1996, a candidata do Partido dos Trabalhadores alcançou uma penetração muito mais significativa na área central do município, mantendo, todavia, uma votação bastante elevada em seus tradicionais redutos eleitorais na periferia.

Figura 15 – Distribuição dos votos obtidos por Marta Suplicy (PT) nas eleições de 2000 – Áreas de Ponderação com votação superior e inferior à média da candidata



5. Resultados

A estimação do modelo de Goodman (1953) e do modelo de King (1997) foi realizada, para cada uma das duas “cross sections” individualmente, através do software *EzI* disponibilizado por King no site <http://gking.harvard.edu/stats.shtml>. Este software possui a vantagem de não apenas apresentar as estimativas do modelo de Goodman (1953) para efeito de comparação de resultados, mas de também apresentar mais ferramentas de diagnóstico de resultados que as disponíveis no software *EI*, também disponibilizado por King⁴⁶.

5.1. Estimação: Eleição de 1996

Na primeira “cross section” analisada, o modelo de King (1997) e o modelo de Goodman (1953) foram aplicados a 3 modelos que relacionaram o percentual de votos obtidos pela candidata derrotada na eleição de 1996 Luiza Erundina, individualmente, às variáveis demográficas sexo (percentual de eleitoras), idade (percentual de eleitores entre 16 e 39 anos, i.e., eleitores jovens) e escolaridade (percentual de eleitores com escolaridade inferior a 8 anos, i.e., escolaridade abaixo da média do município). Para realizar a inferência foram utilizados dados eleitorais e demográficos de 398 áreas de ponderação – de um total de 456 no município – em que existia ao menos um local de votação⁴⁷.

No primeiro modelo, que associa a variável demográfica sexo ao voto na candidata Luiza Erundina foram necessárias 70 interações para alcançar a convergência. Este foi o modelo que levou mais tempo para convergir entre todos os modelos estimados tanto para a eleição 1996 quanto para a eleição de 2000. No segundo modelo, que relaciona o percentual de votos obtidos pela candidata petista à variável demográfica idade, a convergência ocorreu após 52 interações. Finalmente, o terceiro e último modelo, que associa a variável demográfica escolaridade (em anos de estudo completos) à proporção de votos obtida pela candidata Luiza Erundina apresentou convergência após 54 interações.

⁴⁶ O software *EI* possibilita a estimação dos modelos de King Estendido e Generalizado e possui, sem dúvida, mais flexibilidade. Todavia, para os propósitos deste primeiro estudo, o pacote *EzI* parece ser amplamente satisfatório.

⁴⁷ Em 1996, existiam 58 áreas de ponderação no município de São Paulo que não dispunham de locais de votação. Na presente dissertação só consideramos as áreas de ponderação que possuíam locais de votação.

5.2. Estimação: Eleição de 2000

Nesta segunda “cross section”, 7 modelos foram estimados através do método de Goodman (1953) e do método de King (1997).

No primeiro modelo, que relaciona o percentual de votos obtidos pela candidata Marta Suplicy ao percentual de eleitoras por área de ponderação, a convergência ocorreu após 67 interações. O modelo que relaciona a variável idade à proporção de votos obtidos pela candidata Marta Suplicy convergiu após 18 interações. O terceiro modelo, que associa a variável demográfica raça ao voto na candidata Marta Suplicy convergiu após 14 interações. Este foi modelo que mais rapidamente convergiu e – conforme será observado na próxima seção – apresentou as estimativas mais robustas. Por outro lado, o modelo que associa a variável demográfica religião à proporção de votos obtidos pela candidata petista apresentou convergência após 34 interações. O quinto modelo, que associa a variável demográfica escolaridade (em anos de estudo completos) à proporção de votos obtida pela candidata Marta Suplicy apresentou convergência após 23 interações. O modelo que relaciona a variável demográfica renda (categoria Total dos Rendimentos do IBGE) à proporção de votos obtidos pela candidata Marta Suplicy convergiu após 18 interações. Finalmente, o sétimo modelo, que relaciona a variável demográfica imigração (percentual de eleitores que não nasceram no município de São Paulo por área de ponderação) à proporção de votos obtida pela candidata petista apresentou convergência após 27 interações.

Notadamente, observa-se que a estimação dos modelos referentes à eleição de 1996 foi mais morosa do que a estimação dos modelos associados à eleição de 2000. As conseqüências desta constatação serão analisadas em maior detalhe na próxima seção.

5.3. Eleição de 1996: Modelo de Goodman

O Modelo de Goodman (1953), conforme descrito na seção 3.1., se propõe a estimar os parâmetros agregados não ponderados B^b e B^w a partir dos dados disponíveis para as áreas de ponderação. Os resultados das estimações dos parâmetros realizadas para os 3 modelos previamente descritos referentes à eleição de 1996 são apresentados na Tabela 4, abaixo, juntamente com seus respectivos desvios-padrão .

Tabela 4
Eleição de 1996
Resultados do Modelo de Goodman

Variável (X_i)	Parâmetros Agregados	
	B^b	B^w
Sexo	-0.7117 (0.0891)	1.5896 (0.0974)
Idade	0.7349 (0.0154)	-0.0786 (0.0205)
Escolaridade	0.6468 (0.0142)	0.0487 (0.0183)

Analisando os resultados das estimações realizadas através do modelo de Goodman (1953) para as variáveis sexo, idade e escolaridade, observa-se que apenas o terceiro modelo, que associa o percentual de votos obtidos pela candidata Luiza Erundina ao percentual de eleitores com escolaridade inferior à média municipal, apresentou estimativas de parâmetros contidas no intervalo (0,1) e, portanto, passíveis de interpretação analítica. Os parâmetros estimados para este terceiro modelo – ambos estatisticamente significativos – parecem corroborar as explicações apresentadas após as eleições de 1996 para a derrota da candidata Luiza Erundina. Os valores estimados para B^b e B^w mostram que, apesar de quase 65% dos eleitores de baixa escolaridade terem votado na candidata petista, esta foi derrotada no pleito de 1996 por não ter conseguido repetir esta mesma performance no nicho de eleitores de escolaridade superior à média. Neste grupo, apenas 4,9% dos eleitores votaram em Luiza Erundina, de modo que 95,1% dos votos válidos foram direcionados para o candidato eleito Celso Pitta.

O primeiro modelo, que relaciona a proporção de votos obtidos pela candidata petista Luiza Erundina à proporção de eleitoras por área de ponderação, gerou os resultados mais inconsistentes entre todos os modelos estimados neste estudo. Tanto as estimativas obtidas para o parâmetro B^b , quanto as estimativas obtidas para o parâmetro B^w ficaram fora do intervalo (0,1) e, em função disso, não possuem interpretação lógica. O parâmetro B^w , especificamente, foi estimado em 1,59, enquanto que o parâmetro B^b foi estimado em – 0,71.

O segundo modelo, que associa a variável demográfica idade (percentual de eleitores com idade entre 16 e 39 anos) ao percentual de votos obtidos pela candidata Luiza Erundina em 1996, também apresentou estimativas de parâmetros fora do intervalo entre zero e um. Enquanto o parâmetro B^w foi estimado em –0,08 pelo modelo de Goodman (1953), o

parâmetro B^b foi estimado em 0,73. Uma análise dos limites superior e inferior calculados através do método de “bounds” de Duncan & Davis (1953) para o parâmetro B^b confirma o viés dos parâmetros estimados pelo modelo de Goodman (1953) neste caso específico, uma vez que, segundo o método de “bounds”, a estimativa de B^b deveria estar contida no intervalo (0,07; 0,67).

5.4. Eleição de 1996: Modelo de King

O modelo de King (1997) foi estimado para as mesmas variáveis demográficas analisadas pelo modelo de Goodman (1953) na sub-seção anterior. A Tabela 5, abaixo, mostra os parâmetros agregados não ponderados estimados pelo modelo de King (1997) e seus respectivos desvios padrão.

*Tabela 5
Eleição de 1996
Resultados do Modelo de King*

Variável	Parâmetros Agregados	
	B^b	B^w
Sexo	0,0974 (0,0084)	0,6847 (0,0092)
Idade	0,6693 (0,0003)	0,0062 (0,0004)
Escolaridade	0,6387 (0,0064)	0,0552 (0,0078)

Os seis parâmetros agregados estimados pelo modelo de King (1997) se situaram dentro do intervalo entre zero e um – como era de se esperar por construção – e apresentaram elevada significância estatística.

King (1997) apresenta uma série de instrumentos gráficos de análise de resultados. Os três principais instrumentos gráficos utilizados por King (1997) na análise dos resultados de seu modelo são apresentados abaixo e utilizados para verificar a robustez das estimativas acima apresentadas.

1. *Gráfico de Dispersão das Variáveis T_i e X_i* : Neste gráfico é mostrada a dispersão dos dados originais, por área de ponderação, e a reta de regressão estimada pelo modelo de Goodman (i.e., Mínimos Quadrados Ordinários) que melhor se ajusta a estes dados. Estes dados estão contidos em um gráfico quadrangular com arestas de tamanho igual à unidade.

2. *Gráfico Comparativo dos Modelos de Goodman e King*: este gráfico mostra a dispersão dos dados originais (variáveis T_i e X_i) através de círculos de diferentes tamanhos, círculos estes proporcionais ao tamanho de cada uma das áreas de ponderação. Adicionalmente, são mostradas as retas de regressão estimadas pelo modelo de Goodman (1953) e pelo modelo de King (1997), assim como um intervalo de confiança de 80% para a reta inferida através deste último modelo.
3. *Gráfico de Tomografia – “Tomography Plot”*: este gráfico é o mesmo apresentado na Figura 3, na página 24, quando da descrição intuitiva do modelo de King (1997). Este gráfico relaciona os parâmetros β_i^b e β_i^w por micro-região (neste caso, por área de ponderação) através de retas truncadas em um quadrado de arestas iguais à unidade. Sobre as P áreas de ponderação truncadas sobrepõe-se uma função de distribuição normal bivariada truncada (em forma de curvas de nível), a qual é utilizada para determinar a localização mais provável dos P pares de parâmetros de interesse relacionados às P retas.

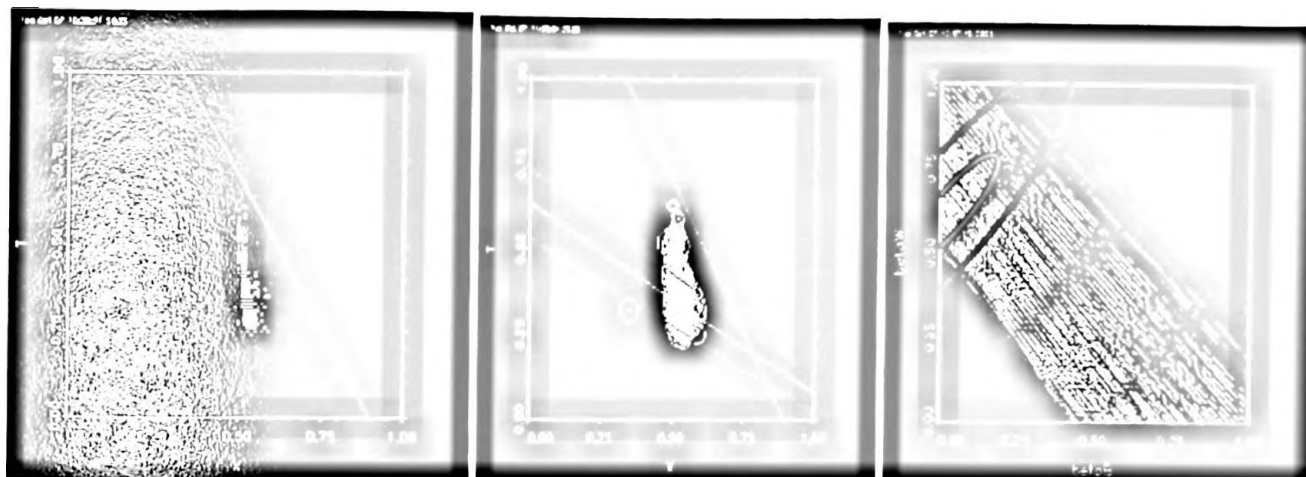
O terceiro gráfico, o “tomography plot”, é o instrumento mais importante apresentado por King (1997) na análise dos dados. A localização da função de distribuição normal bivariada truncada depende da concentração de linhas no gráfico e a precisão da localização da mesma é maior quanto maior for a quantidade de linhas cruzadas no gráfico. Desta forma, uma quantidade pequena de micro-regiões – ou áreas de ponderação – pouco informará sobre a exata localização da função de distribuição normal bivariada truncada, enquanto que uma grande quantidade de linhas gráficas paralelas pode limitar até um certo ponto as possibilidades de localização da função de distribuição normal bivariada truncada, mas não trará muita certeza quanto à exata localização da mesma ao longo destas retas paralelas.

Assim, os dados mais informativos – ou ainda, os dados que possibilitam um melhor estudo de inferência ecológica – são aqueles cujos gráficos de tomografia possuem um grande número de linhas que se concentram e se cruzam em algum ponto dentro do quadrado de arestas iguais à unidade.

Os parâmetros estimados para o primeiro modelo - que relaciona o percentual de votos obtidos pela candidata Luiza Erundina ao percentual de eleitoras por área de ponderação – indicam que apenas 9,4% dos eleitores do sexo feminino votaram na candidata Luiza Erundina no pleito de 1996, enquanto que 68,5% dos eleitores do sexo masculino votaram

na petista. Abaixo são apresentados os gráficos gerados pelo software EzI referentes à variável demográfica sexo.

Figura 16
Instrumentos de Diagnóstico
Luiza Erundina vs. Sexo



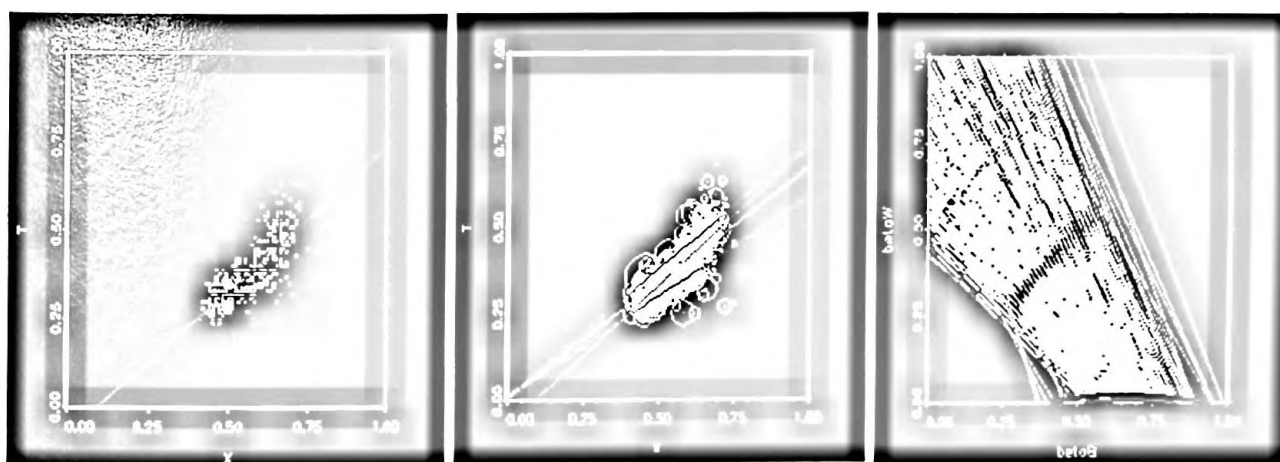
Ao analisar o primeiro gráfico, observa-se que os dados apresentam muito pouca variação e a estimação pelo modelo de Goodman (1953) gera parâmetros fora do intervalo entre zero e um.⁴⁸ O segundo gráfico, por sua vez, mostra a diferença significativa existente entre os parâmetros estimados pelo modelo de King (1997) e pelo modelo de Goodman (1953). O intervalo de confiança de 80% calculado com base nas estimativas obtidas pelo modelo de King (1997) mostra que os parâmetros estimados pelos dois modelos não se ajustam bem aos dados. A peculiaridade dos dados da variável sexo – e a conseqüente dificuldade em se fazer inferência ecológica a partir dos mesmos – pode ser observada no terceiro gráfico, no qual pode-se observar um grande número de retas quase que perfeitamente paralela. Este tipo de gráfico de tomografia salienta a dificuldade em se localizar precisamente a posição da função de distribuição normal bivariada truncada e a incerteza associada aos parâmetros estimados pelo modelo de King (1997). Conseqüentemente, os parâmetros estimados pelo modelo de King (1997) para a variável demográfica sexo não devem ser considerados para efeito de análise.

O segundo modelo, que relaciona os votos obtidos pela candidata Luiza Erundina à variável demográfica idade, apresentou os seguintes parâmetros agregados estimados pelo modelo de King: $B^b = 0,67$ e $B^w = 0,006$. Ou seja, 67% dos eleitores com idade entre 16 e 39 anos votaram em Luiza Erundina, enquanto que menos de 1% dos eleitores com idade superior a 39 anos votaram na petista. Os instrumentos de diagnóstico apresentados abaixo

⁴⁸ Para chegar a esta conclusão deve-se estender a reta até o ponto em que esta cruze os eixos verticais.

para a variável idade mostram que estes dados possuem maior variação do que os dados referentes à variável sexo e que as estimativas obtidas pelo modelo de King (1997) e pelo modelo de Goodman (1953) são bastante próximas. O intervalo de confiança de 80% construído para os parâmetros estimados pelo método de King (1997) parece indicar uma qualidade de ajuste razoável em relação aos dados. O gráfico de tomografia, por sua vez, indica que o parâmetro B^w deve se situar muito próximo de zero, o que confirmaria a suspeita de que os eleitores de maior idade possuem um perfil mais conservador. Deve-se destacar que - principalmente a análise do "tomography plot" - indica uma maior cautela em relação aos parâmetros estimados, uma vez que poucas linhas deste gráfico se cruzam, o que indica uma maior imprecisão das estimativas, segundo King (1997).

Figura 17
Instrumentos de Diagnóstico
Luiza Erundina vs. Idade

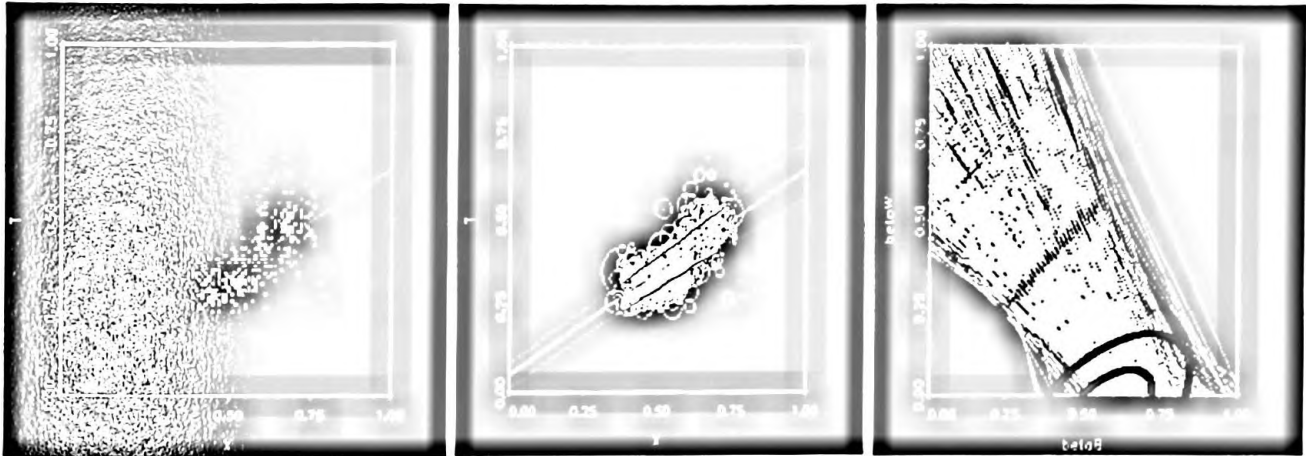


Por fim, os parâmetros agregados estimados pelo modelo de King (1997) para o modelo que relaciona o percentual de votos obtido pela ex-prefeita de São Paulo Luiza Erundina à variável demográfica escolaridade foram muito semelhantes aos previamente estimados pelo modelo de Goodman (1953).

Entre os três modelos que relacionam as variáveis sexo, idade e escolaridade ao voto, este último foi o que apresentou maior variação nos dados, conforme é possível observar abaixo no primeiro gráfico à esquerda. Em função disso, o modelo de Goodman (1953) e o modelo de King (1997) ajustaram-se bem aos dados e apresentaram resultados quase idênticos, conforme é possível observar no segundo gráfico. O intervalo de confiança de 80%, por sua vez, confirma o bom ajuste obtido para o modelo de escolaridade. Finalmente, o gráfico de tomografia mostra uma grande concentração de linhas se

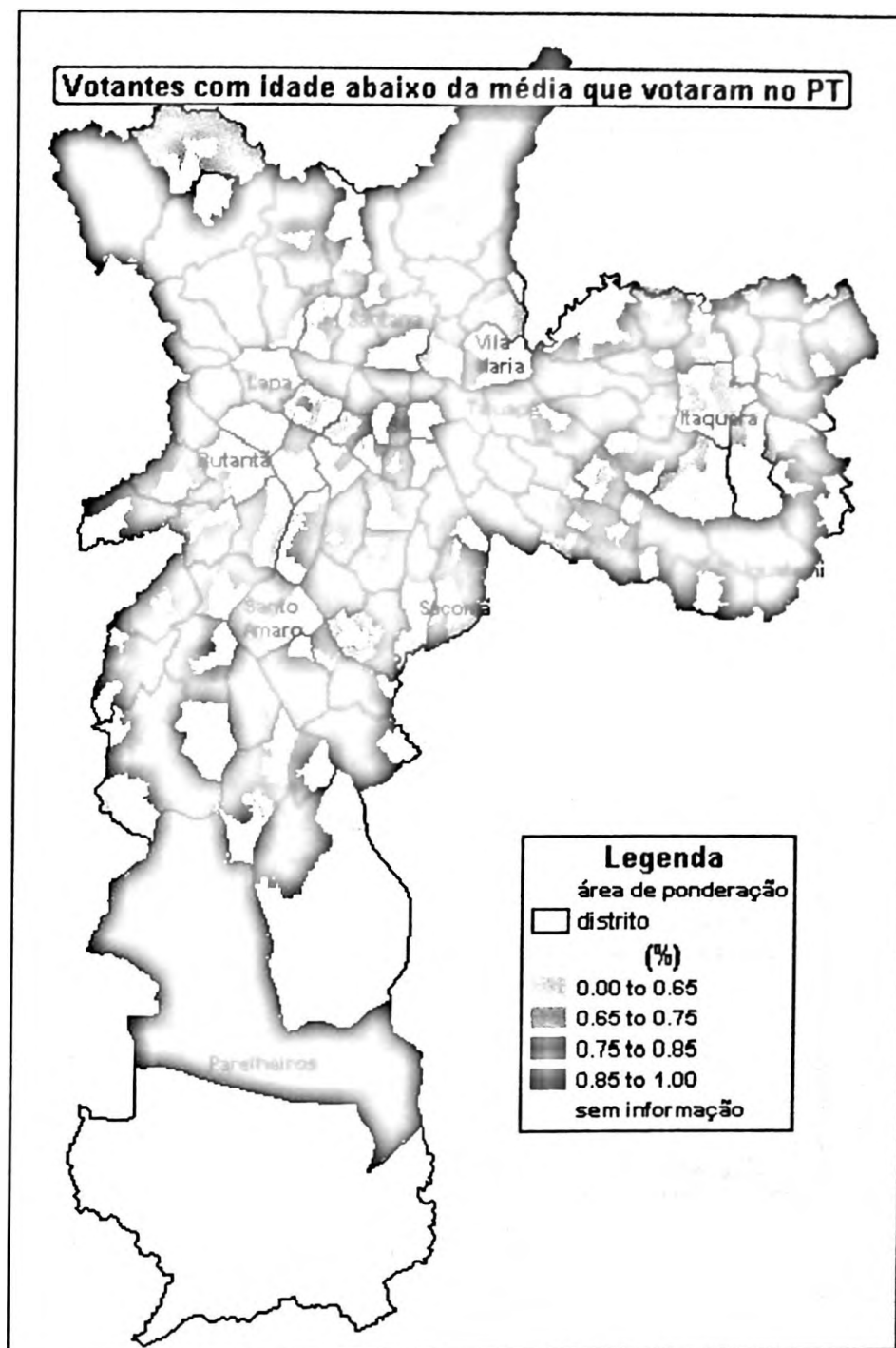
cruzando em sua parte inferior direita, o que acaba por gerar maior conforto em relação aos parâmetros estimados.

Figura 18
Instrumentos de Diagnóstico
Luiza Erundina vs. Escolaridade



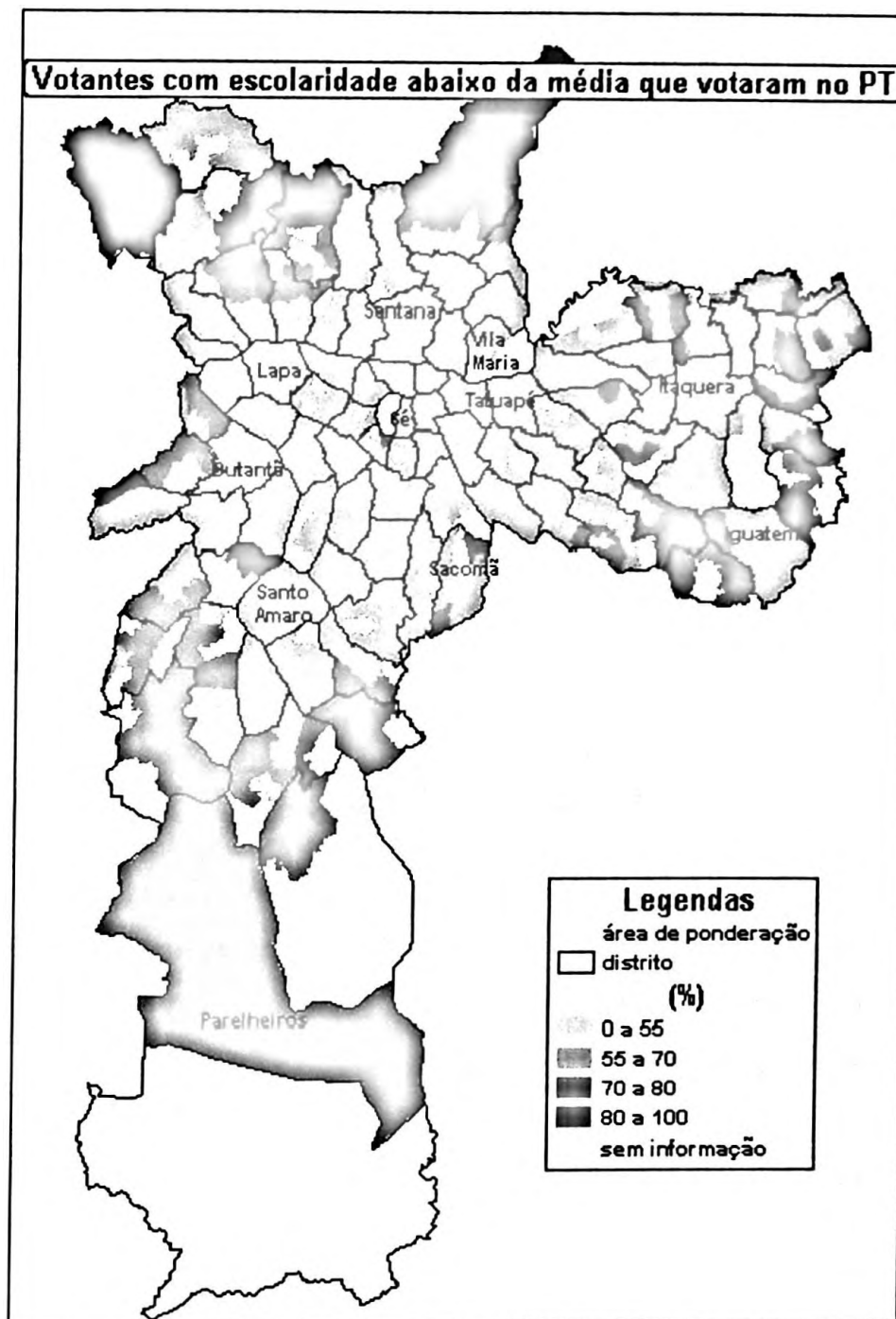
Uma das principais vantagens do modelo desenvolvido por King (1997) em relação aos modelos anteriores como o de Goodman (1953) é o fato do primeiro estimar não apenas os parâmetros agregados B^b e B^w , mas também os parâmetros por micro-região (neste caso, por área de ponderação). A seguir, são apresentados os resultados das estimações de dois dos três modelos aqui desenvolvidos (para as variáveis idade e escolaridade) por área de ponderação. Como os resultados do modelo desenvolvido para a variável sexo não foram estatisticamente satisfatórios, os mesmos não serão apresentados espacialmente.

Figura 19 – Percentual de indivíduos com idade entre 16 e 39 anos que votaram em Luiza Erundina (PT) em 1996 – por área de ponderação



A Figura 19 mostra que a elevada votação alcançada por Luiza Erundina em 1996 entre os indivíduos com idade entre 16 e 39 anos foi distribuída de modo bastante equilibrada no espaço. Deve-se destacar, entretanto, que as votações mais elevadas obtidas pela candidata petista neste nicho de indivíduos se localizaram em regiões de periferia do município.

Figura 20 – Percentual de indivíduos com escolaridade inferior a 8 anos que votaram em Luiza Erundina (PT) em 1996 – por área de ponderação



A Figura 20 mostra que a candidata petista Luiza Erundina recebeu, na eleição de 1996, uma maior concentração de votos de indivíduos com escolaridade inferior a oito anos nas regiões de periferia. O distrito em que a candidata petista recebeu o menor percentual de votos de indivíduos com escolaridade abaixo da média do município foi a Vila Maria, um conhecido reduto malufista.

5.5. Eleição de 2000: Modelo de Goodman

A Tabela 6, abaixo, apresenta os parâmetros agregados estimados B^b e B^w (parâmetros para o município de São Paulo como um todo) e seus respectivos desvios-padrão para cada

um dos sete modelos propostos para analisar a eleição para a prefeitura de São Paulo realizada em 2000, em segundo turno.

Tabela 6
Eleição de 2000
Resultados do Modelo de Goodman

<i>Variável (X_i)</i>	<i>Parâmetros Agregados</i>	
	B^b	B^w
Sexo	0.2354 (0.0535)	1.0067 (0.0623)
Idade	0.7728 (0.0125)	0.3558 (0.0162)
Raça	0.7431 (0.0111)	0.5315 (0.0049)
Religião	0.8970 (0.0367)	0.5341 (0.0074)
Escolaridade	0.6654 (0.0083)	0.5160 (0.0084)
Renda	0.6374 (0.0056)	0.5053 (0.0096)
Imigração	0.7762 (0.0129)	0.4158 (0.0123)

Ao analisar os resultados do primeiro modelo, que relaciona o percentual de votos obtidos pela candidata Marta Suplicy ao percentual de eleitoras por área de ponderação, observa-se que o Modelo de Goodman (1953) gerou resultados que não são passíveis de interpretação, uma vez que B^w está fora do intervalo (0,1) e não pode ser interpretado como uma proporção. Em função disso, fica impossível fazer qualquer análise, neste momento, sobre a relação existente entre o voto feminino e a eleição da então candidata Marta Suplicy.

Os outros seis modelos estimados, por outro lado, apresentaram parâmetros estimados contidos no intervalo entre zero e um e, assim, possuem valor analítico. Adicionalmente, observa-se que todos os doze parâmetros estimados são estatisticamente significantes, sendo que o modelo que apresentou menor nível de significância foi o que relaciona a variável demográfica idade ao voto.

Os resultados obtidos na estimação do segundo modelo, que relaciona a variável idade à proporção de votos obtidos pela candidata Marta Suplicy, indicam uma grande concentração dos votos dos eleitores mais jovens (entre 16 e 39 anos) na candidata de esquerda Marta Suplicy. Segundo o Modelo de Goodman (1953), a proporção de “jovens”

que votaram na candidata Marta Suplicy é de 0.77, enquanto que o percentual de eleitores com idade superior a 39 anos que votaram na candidata petista foi de apenas 35% na eleição de 2000.

O modelo que associa a variável demográfica raça ao voto na candidata Marta Suplicy indicou uma conotação racial importante na eleição municipal de 2000. Nesta eleição, a proporção de negros, pardos e indígenas que votaram na candidata Marta Suplicy foi de 0.74 (sendo que apenas 26% dos eleitores pertencentes a estas etnias votaram no candidato Paulo Maluf). Por outro lado, o modelo indica que o voto dos brancos e amarelos foi muito mais dividido entre a candidata petista e o candidato pepebista, já que o percentual de brancos e amarelos que votaram na candidata Marta Suplicy foi de 53%.

O quarto modelo, que associa a variável demográfica religião à proporção de votos obtidos pela candidata petista, apresentou um resultado bastante informativo. À primeira vista, duas hipóteses poderiam ser feitas com relação ao apoio dos eleitores evangélicos a uma candidata de um partido esquerdista como Marta Suplicy: (1) os eleitores evangélicos poderiam possuir um perfil mais conservador e ter resistência a votar em uma candidata de um partido de esquerda; (2) por outro lado, como a maior parte dos eleitores evangélicos reside em bairros da periferia – onde o Partido dos Trabalhadores é historicamente forte – seria razoável assumir que Marta Suplicy teria tido uma significativa votação entre os evangélicos. Os parâmetros estimados através do modelo de Goodman (1953) mostraram que os eleitores evangélicos apoiaram fortemente a candidata petista (90% dos evangélicos votaram em Marta Suplicy), enquanto que os eleitores católicos ficaram bastante divididos (53% votaram em Marta Suplicy e 47% em Paulo Maluf).

O modelo que associa a variável demográfica educação (em anos de estudo completos) à proporção de votos obtida pela candidata Marta Suplicy apresentou um resultado menos contrastante do que os modelos anteriores. A proporção de eleitores com menor escolaridade (com escolaridade menor do que a média do eleitorado, i.e., menor do que 8 anos) que votou na candidata Marta Suplicy, segundo o Modelo de Goodman (1953), foi de 0,66. Este resultado parece ser condizente com o fato da candidata petista possuir um apoio maior nos bairros de periferia, onde a escolaridade média é menor. Por outro lado, o parâmetro estimado para o extrato do eleitorado com maior nível educacional indica uma divisão quase que perfeita entre os candidatos Marta Suplicy e Paulo Maluf (51% dos votos para a petista e 49% para o pepebista).

O sexto modelo, que associa a variável demográfica renda (categoria Total dos Rendimentos do IBGE) à proporção de votos obtidos pela candidata Marta Suplicy confirma a importância de uma variável que foi destacada pelos cientistas políticos após a eleição como um dos principais fatores que explicaram a eleição de Marta Suplicy no pleito municipal de 2000: a maior penetração da candidata petista entre os eleitores de classe média e alta. Na eleição de 2000, o percentual de eleitores com renda superior à renda média do eleitorado (R\$835,00) que votaram na candidata Marta Suplicy foi de 51% (contra 49% dos votos dos eleitores deste extrato de renda para Paulo Maluf). Por outro lado, a proporção de eleitores com renda menor do que R\$835,00 que votaram na então candidata Marta Suplicy foi de 0,64 na eleição de 2000.

Finalmente, o último modelo, que relaciona a variável demográfica imigração (percentual de eleitores que não nasceram no município de São Paulo por área de ponderação) à proporção de votos obtida pela candidata petista mostrou que os imigrantes apoiariam de modo significativo a candidata Marta Suplicy. Os parâmetros estimados pelo Modelo de Goodman (1953) mostraram que quase 78% dos imigrantes votaram na candidata do Partido dos Trabalhadores, enquanto que apenas 22% votaram em Paulo Maluf. Por outro lado, entre os eleitores nascidos no município de São Paulo, apenas 42% votaram na candidata Marta Suplicy.

5.6. Eleição de 2000: Modelo de King

A Tabela 7, abaixo, apresenta os parâmetros agregados ponderados estimados pelo Modelo de King (1997) - B^b e B^w - e seus respectivos desvios-padrão para cada um dos sete modelos anteriormente estimados através do método de Goodman (1953). Todos os quatorze coeficientes estimados apresentaram elevada significância estatística.

Tabela 7
Eleição de 2000
Resultados do Modelo de King

Variável	Parâmetros Agregados	
	B ^b	B ^w
Sexo	0.3059 (0.0165)	0.9123 (0.0194)
Idade	0.7700 (0.0094)	0.3569 (0.0115)
Raça	0.7379 (0.0077)	0.5305 (0.0027)
Religião	0.8697 (0.0238)	0.5340 (0.0042)
Escolaridade	0.6600 (0.0073)	0.5138 (0.0068)
Renda	0.6321 (0.0053)	0.5039 (0.0089)
Imigração	0.7807 (0.0097)	0.4069 (0.0088)

Os parâmetros estimados para o primeiro modelo, que relaciona o percentual de votos obtidos pela candidata Marta Suplicy ao percentual de eleitoras por área de ponderação, diferentemente dos estimados pelo método de Goodman (1953), ficaram limitados dentro do intervalo (0,1), como era de se esperar por construção. Os parâmetros agregados (para o município de São Paulo como um todo) para a variável sexo foram estimados em 0,31 e 0,91, i.e., 31% das mulheres e 91% dos homens votaram na candidata Marta Suplicy (enquanto que 69% delas votaram em Paulo Maluf e apenas 9% dos homens votaram no pepebista).

Abaixo, são apresentados os gráficos gerados para a variável sexo pelo software EzI para a eleição de 2000.

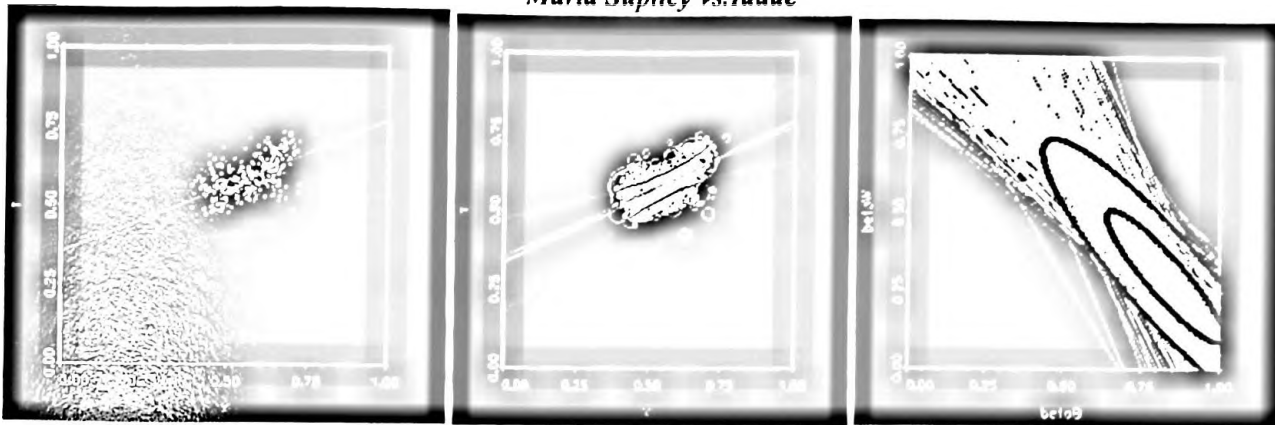
Figura 21
Instrumentos de Diagnóstico
Marta Suplicy vs. Sexo



O primeiro gráfico à esquerda, que mostra o ajuste dos dados originais ao modelo de Goodman (1953), salienta a falta de variação da variável demográfica sexo por área de ponderação. Este problema já havia sido detectado nos dados da Contagem de 1996 e foram confirmados para os dados do Censo 2000. O gráfico que compara as retas de regressão obtidas pelo modelo de King (1997) e pelo modelo de Goodman (1953) apresentou uma menor diferença entre os dois modelos estimados do que os modelos estimados para a variável demográfica sexo na “cross section” de 1996. Entretanto, o intervalo de confiança de 80% construído com base nas estimativas obtidas pelo modelo de King (1997) mostra que tanto o modelo de Goodman (1953) quanto o modelo de King (1997) não se ajustam bem aos dados. Novamente, a peculiaridade da variável demográfica sexo pode ser observada no gráfico de tomografia, o qual, neste caso, é composto, quase que unicamente, por retas paralelas. Em função disso, os parâmetros estimados para a variável sexo devem ser descartados para efeito de análise, assim como ocorreu com o modelo análogo estimado para as eleições de 1996.

O segundo modelo, que associa a variável demográfica idade à proporção de votos obtidos pela então candidata Marta Suplicy, apresentaram parâmetros muito próximos tanto para o modelo de Goodman (1953) como para o modelo de King (1997). Segundo ambos modelos, 77% dos eleitores “jovens” votaram na candidata do Partido dos Trabalhadores, enquanto que 36% dos eleitores com idade superior a 39 anos votaram na candidata Marta Suplicy.

Figura 22
Instrumentos de Diagnóstico
Marta Suplicy vs. Idade

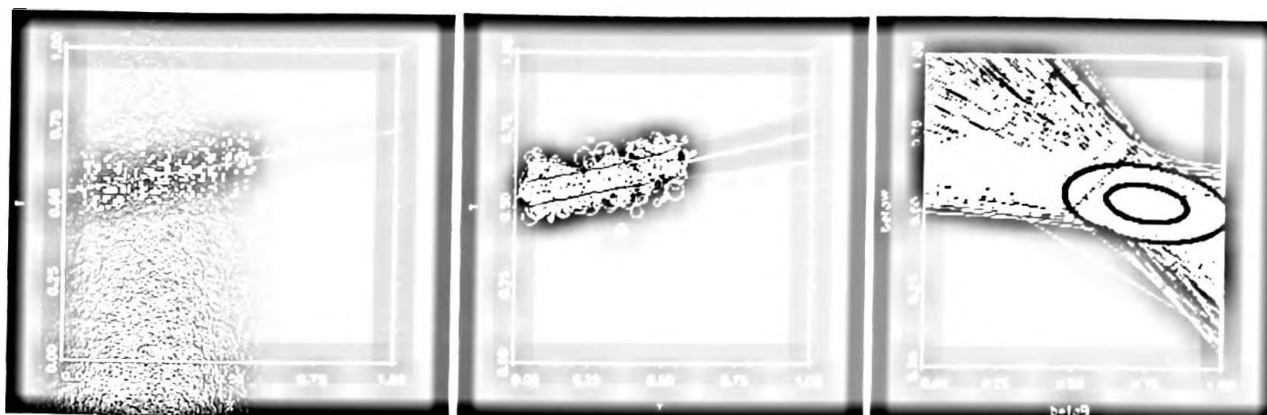


Os gráficos de diagnóstico acima mostram que as retas de regressão dos modelos de King (1997) e de Goodman (1953) são bastante próximas e que o intervalo de confiança de 80% indica um bom ajuste dos dois modelos aos dados originais. Adicionalmente, o gráfico de tomografia contém um grande número de retas de regressão se cruzando, o que possibilita localizar a posição da função de distribuição normal bivariada truncada usada na estimação do modelo de King (1997) com menor incerteza.

O modelo que associa a variável demográfica raça ao percentual de votos obtidos pela candidata Marta Suplicy no pleito municipal de 2000 apresentou os seguintes parâmetros agregados estimados: $B^b=0,74$ e $B^w=0,53$. Estes resultados são muito próximos às estimativas obtidas pelo modelo de Goodman (1953) e indicam que, enquanto negros, pardos e indígenas concentraram seus votos de modo acentuado na candidata Marta Suplicy, brancos e amarelos dividiram seus votos quase que igualmente entre a candidata do Partido dos Trabalhadores e o pepebista Paulo Maluf.

Os instrumentos gráficos de diagnóstico apresentados abaixo mostram que o modelo inferido para a variável demográfica raça possui o melhor ajuste entre todos os modelos estimados nesta dissertação. O intervalo de confiança de 80% construído a partir dos resultados das estimativas do modelo de King (1997), e apresentado no gráfico central acima, mostra que os parâmetros estimados se ajustam bem aos dados. Adicionalmente, o gráfico de tomografia, à esquerda, possui um grande número de retas de regressão se cruzando, o que confirma a qualidade dos parâmetros estimados pelo modelo de King (1997). Por fim, cabe ressaltar que o modelo estimado para a variável demográfica raça foi, entre todos, o modelo que demandou o menor número de interações para obter a convergência na etapa de estimação (18 interações).

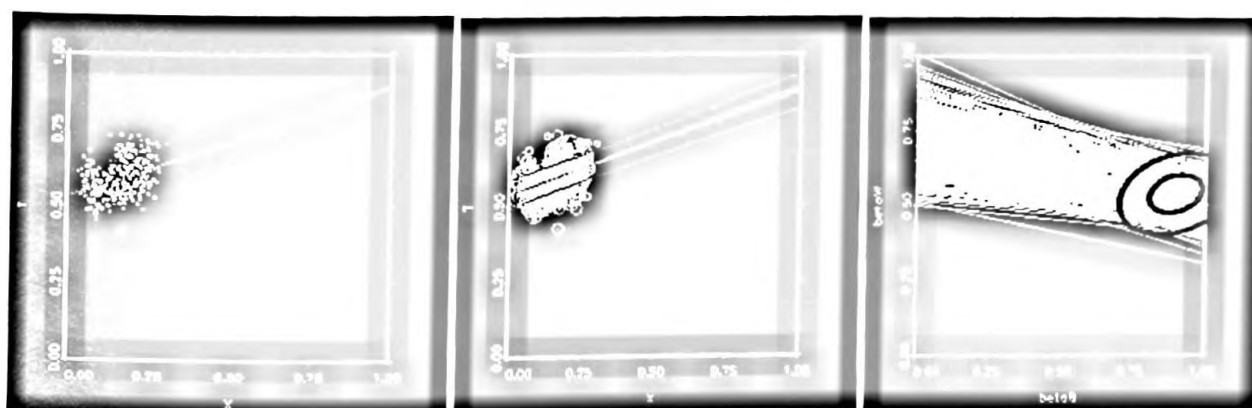
Figura 23
Instrumentos de Diagnóstico
Marta Suplicy vs. Raça



O quarto modelo, que relaciona a variável demográfica religião (percentual de eleitores que se declararam membros de igrejas cristãs protestantes) ao percentual de votos obtidos por Marta Suplicy, indicou que 87% dos eleitores protestantes votaram na candidata petista, enquanto que apenas 53% dos eleitores católicos votaram em Marta Suplicy.

Os instrumentos gráficos de diagnóstico apresentados abaixo destacam a pequena variação existente nos dados originais, assim como a similaridade das estimativas obtidas pelos modelo de King (1997) e de Goodman (1953).

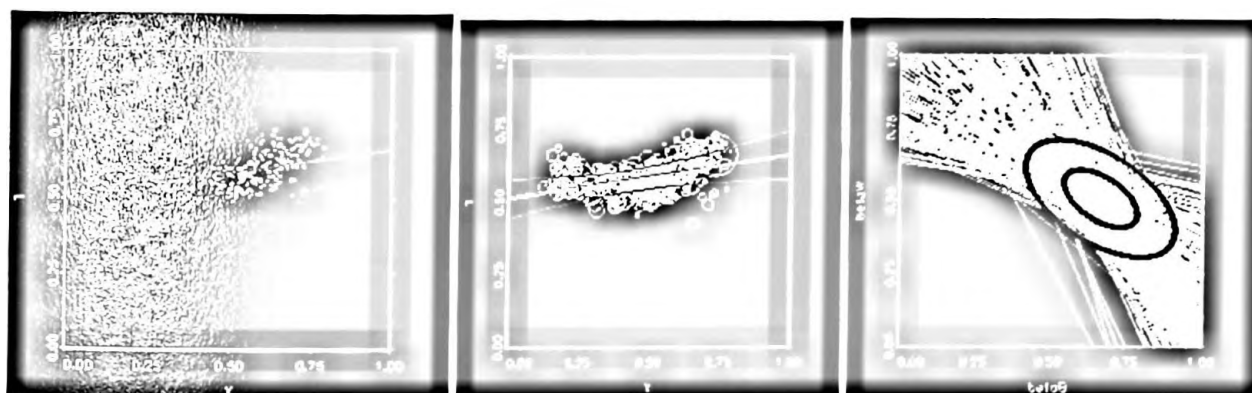
Figura 24
Instrumentos de Diagnóstico
Marta Suplicy vs. Religião



Apesar da similaridade das estimativas obtidas pelos modelos de King (1997) e de Goodman (1953), a análise do gráfico de tomografia – que dispunha de um grande número de retas de regressão paralelas –, a pequena variação dos dados e a grande quantidade de pontos fora do intervalo de confiança de 80% no gráfico central apresentado acima indicam problemas no tocante à qualidade dos ajustes obtidos tanto pelo modelo de Goodman (1953), quanto pelo modelo de King (1997). Por estes motivos, os parâmetros estimados para a variável religião não devem ser considerados para efeito de análise.

Os parâmetros estimados para o modelo que combina a variável demográfica escolaridade ao total de votos obtidos pela candidata petista Marta Suplicy na eleição municipal de 2000 indicam que 66% dos eleitores com escolaridade inferior à média do município (i.e., com menos que 8 anos completos de estudo) votaram na candidata do Partido dos Trabalhadores, enquanto que 51% dos eleitores com escolaridade superior à média votaram na candidata eleita. Estes resultados são bastante próximos aos obtidos previamente pelo modelo de Goodman (1953).

*Figura 25
Instrumentos de Diagnóstico
Marta Suplicy vs. Escolaridade*

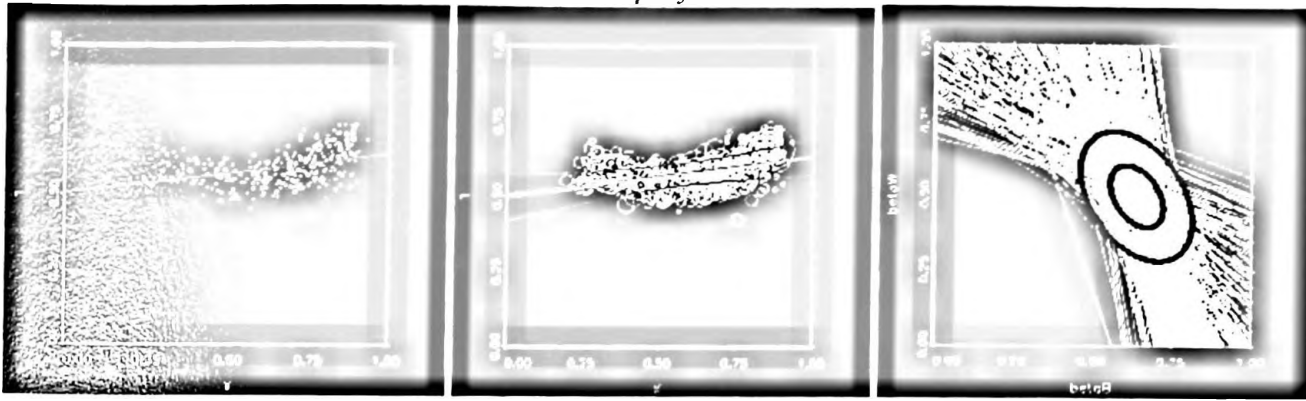


A análise dos instrumentos gráficos de diagnóstico acima dispostos mostra um ajuste bastante bom para este quinto modelo. Os gráficos mostram a similaridade entre os parâmetros estimados pelos modelos de King (1997) e de Goodman (1953), a grande variabilidade dos dados e as boas evidências sobre a localização da função de distribuição normal bivariada truncada no gráfico de tomografia.

O penúltimo modelo – que associa a variável demográfica renda ao voto na candidata Marta Suplicy – indicou que 63% dos eleitores com renda inferior à média municipal de R\$835,00 votaram na candidata petista, enquanto que 50% dos eleitores das classes mais abastadas votaram na candidata petista. Estes resultados estão em linha com os obtidos previamente através do método de Goodman (1953) e indicam a maior penetração da candidata do Partido dos Trabalhadores no eleitorado de classe média.

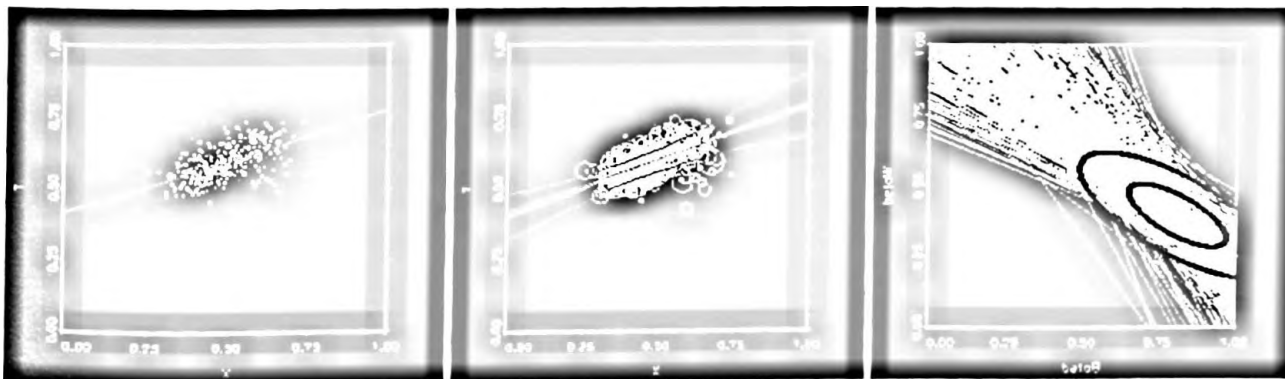
Novamente, os instrumentos gráficos de análise indicam um ajuste bastante bom para o modelo em questão, conforme se pode ver abaixo.

Figura 26
Instrumentos de Diagnóstico
Marta Suplicy vs. Renda



Finalmente, o último modelo, que associa a variável demográfica imigração – total de eleitores que não nasceram no município de São Paulo – à proporção de votos obtida pela então candidata Marta Suplicy na eleição municipal de 2000 apresentou um ajuste relativamente bom, mas, sem dúvida, inferior aos obtidos para as variáveis demográficas escolaridade e renda. Os parâmetros de agregados de interesse foram estimados pelo modelo de King (1997) em 0,78 e 0,41, resultados estes bastante similares aos obtidos pelo modelo de Goodman (1953). Assim, 78% dos eleitores imigrantes votaram na candidata eleita Marta Suplicy, enquanto que apenas 41% dos eleitores nascidos em São Paulo votaram na prefeita eleita.

Figura 27
Instrumentos de Diagnóstico
Marta Suplicy vs. Migração



Na Tabela 8, abaixo, são apresentados os parâmetros estimados pelos modelos de King (1997) e Goodman (1953), para cada uma das sete variáveis demográficas de interesse analisadas para a “cross section” de 2000. Entre todas, apenas a variável demográfica sexo, cujas estimativas obtidas através do modelo de Goodman (1953) não estavam contidas no intervalo (0,1), apresentou diferenças significativas entre os 2 modelos aplicados.

Tabela 8
Eleição de 2000
Comparação de Resultados dos
Modelos de Goodman e King

Variável	Parâmetro	Modelo	
		Goodman	King
Sexo	B ^b	0.24	0.31
	B ^w	1.01	0.91
Idade	B ^b	0.77	0.77
	B ^w	0.36	0.36
Raça	B ^b	0.74	0.74
	B ^w	0.53	0.53
Religião	B ^b	0.90	0.87
	B ^w	0.53	0.53
Escolaridade	B ^b	0.67	0.66
	B ^w	0.52	0.51
Renda	B ^b	0.64	0.63
	B ^w	0.51	0.50
Imigração	B ^b	0.78	0.78
	B ^w	0.42	0.41

A análise pormenorizada realizada acima através dos instrumentos de diagnóstico indicou que os modelos que apresentaram um melhor ajuste aos dados são os que consideraram as variáveis demográficas renda, escolaridade e raça. Já os modelos que associaram ao voto as variáveis demográficas idade e migração apresentaram ajustes não tão satisfatórios, de modo que seus parâmetros estimados devem ser considerados com maior cautela. Finalmente, os modelos estimados para as variáveis demográficas sexo e religião não apresentaram ajustes minimamente satisfatórios e – por isso – devem ser descartados para efeito de análise.

Na Tabela 9, abaixo, são apresentadas as estatísticas descritivas dos dados demográficos. Observa-se que as variáveis demográficas que possuem menor variação - em pontos percentuais entre o percentual máximo e o percentual mínimo – são precisamente as variáveis que possuem um ajuste de menor qualidade aos dados.

Tabela 9
Eleição de 2000
Estatísticas Descritivas dos Dados
por Área de Ponderação

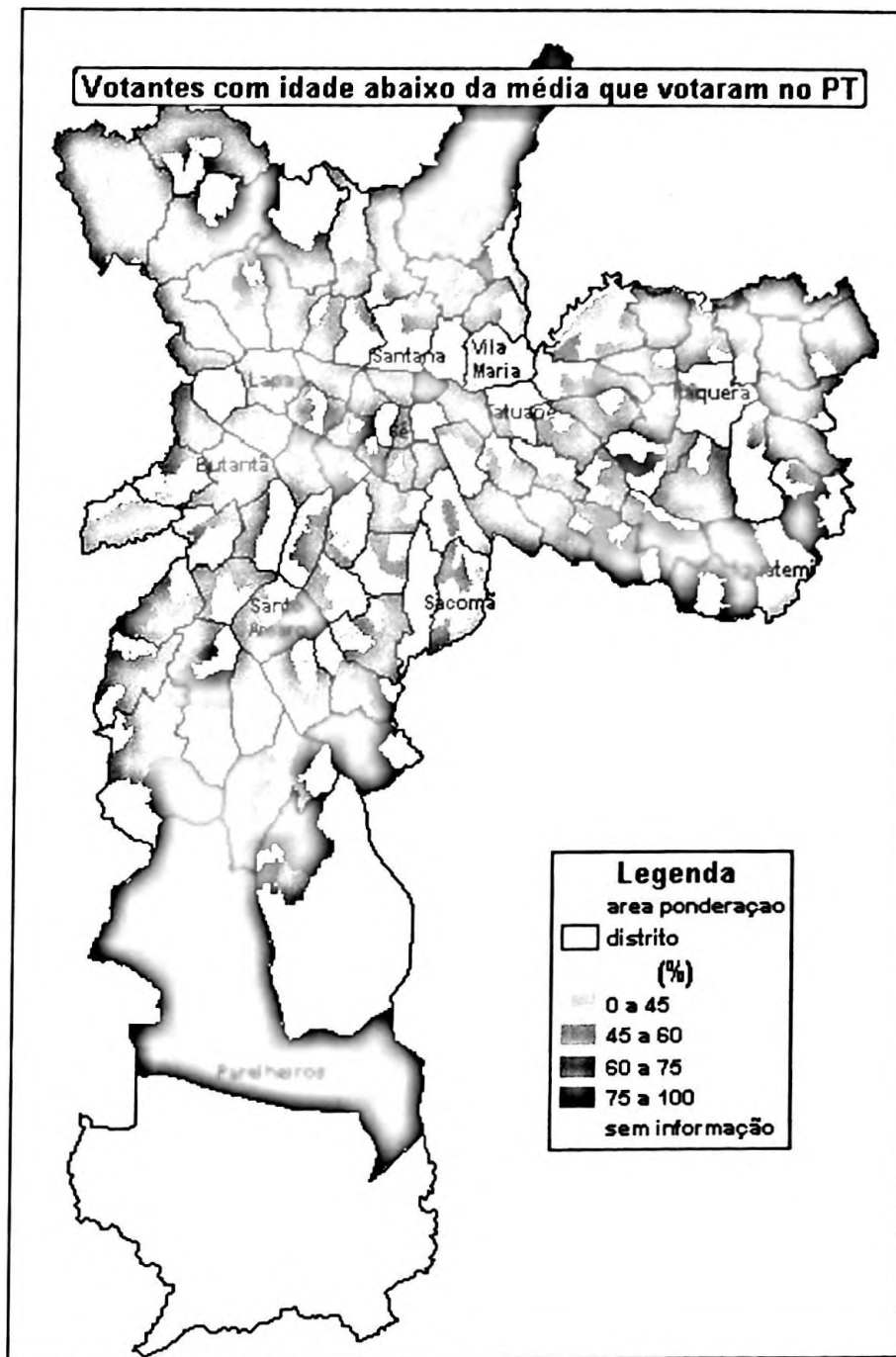
Percentuais	Média	Desvio Padrão	Min.	Max.
% Votos para Marta Suplicy	0,5916	0,0566	0,43	0,73
% Sexo Feminino	0,5382	0,0230	0,50	0,60
% Eleitores entre 16 e 39 anos	0,5651	0,0797	0,39	0,77
% Negros/Pardos/Indígenas	0,2839	0,1521	0,04	0,60
% Religião Evangélica	0,1581	0,0593	0,03	0,30
% Menos de 8 anos de Estudo	0,5055	0,1591	0,13	0,79
% Renda Menor que a Média	0,6523	0,1781	0,20	0,93
% Imigrantes	0,4876	0,0914	0,26	0,78

Adicionalmente, são apresentadas as estimativas dos parâmetros β^b_i e β^w_i , por área de ponderação, para cinco das sete variáveis demográficas de interesse aqui analisadas. As estimativas dos parâmetros β^b_i e β^w_i foram obtidas para cada uma das 412 áreas de ponderação do município de São Paulo às quais era possível associar pelo menos um local de votação na eleição de 2000.⁴⁹

Novamente, não serão apresentados os gráficos espaciais das variáveis que não apresentaram bons ajustes segundo os critérios disponíveis, neste caso, os gráficos espaciais das variáveis sexo e religião.

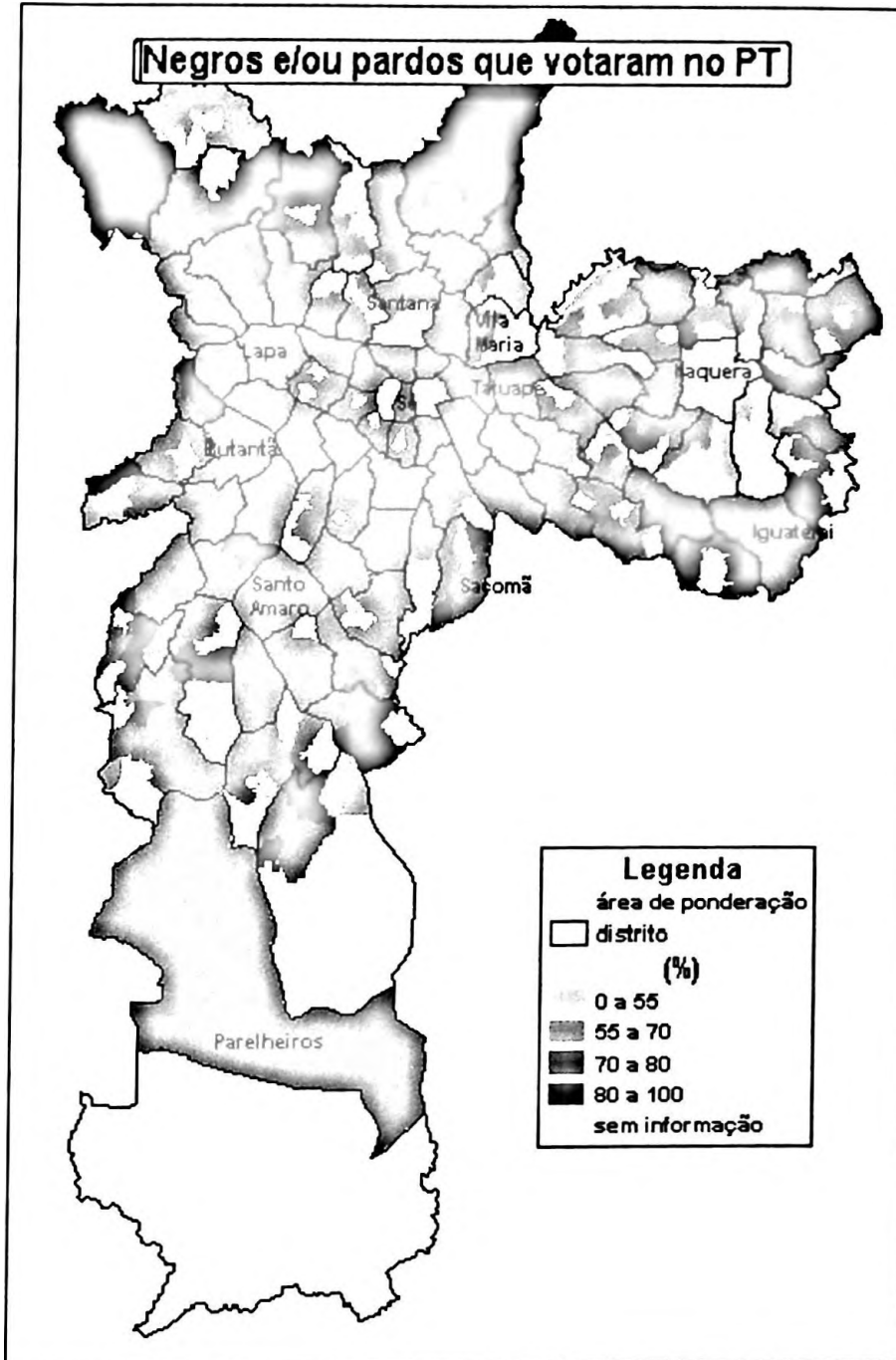
⁴⁹ O município de São Paulo possui no total, 456 áreas de ponderação. Entretanto, neste estudo sobre análise de votação, desconsideramos as áreas de ponderação nas quais não existiam locais de votação na eleição de 2000 e 1996.

Figura 28 – Percentual de indivíduos com idade entre 16 e 39 anos que votaram em Marta Suplicy (PT) em 2000 – por área de ponderação



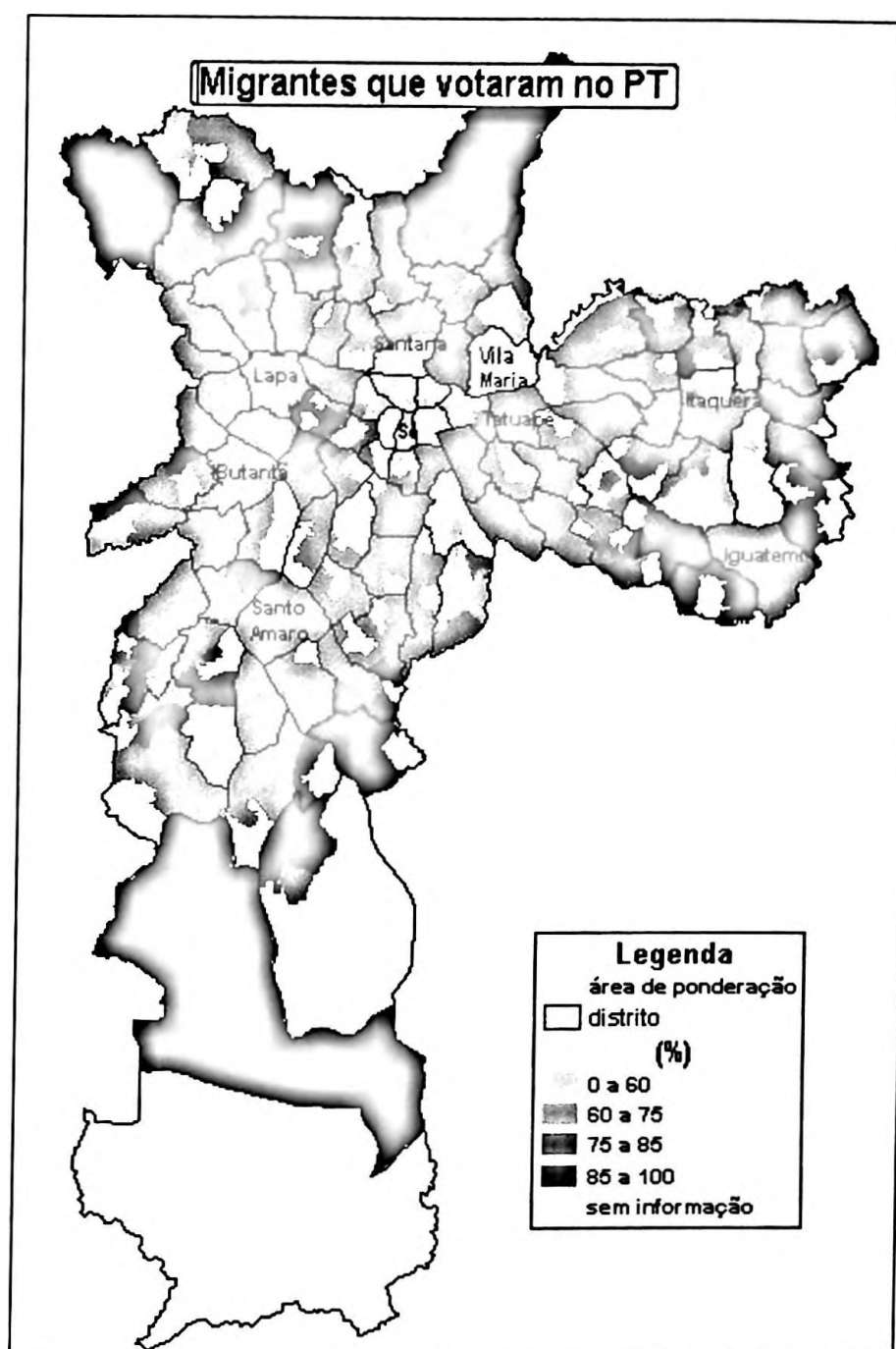
Assim como ocorrido nas eleições municipais de 1996 (ver Figura 19), a distribuição de eleitores com idade entre 16 e 39 anos que votaram no PT é bastante equilibrada espacialmente, apesar de se observar uma concentração ligeiramente maior de votos desta categoria de eleitores em regiões periféricas do município.

Figura 29 – Percentual de indivíduos negros, pardos ou indígenas que votaram em Marta Suplicy (PT) em 2000 – por área de ponderação



Assim como no caso da variável demográfica idade, no caso da variável “raça”, é possível observar uma distribuição espacial bastante equilibrada dos eleitores negros, pardos ou indígenas que votaram na candidata petista Marta Suplicy nas eleições municipais de 2000. Merecem destaque os votos desta categoria de eleitores na parte leste da Vila Maria, onde se observou a menor votação de negros, pardos e indígenas em Marta Suplicy do município.

Figura 32 – Percentual de indivíduos não nascidos em São Paulo que votaram em Marta Suplicy (PT) em 2000 – por área de ponderação



Finalmente, as estimativas do voto em Marta Suplicy dos eleitores não nascidos em São Paulo mostram uma razoável distribuição espacial dos votos no município. Merece destaque a região central de São Paulo, a qual apresenta uma elevada concentração de imigrantes e uma reduzida votação na candidata petista.

6. Conclusão

Este trabalho tem por objetivo estudar os determinantes sociais do voto para prefeito, no município de São Paulo, a partir dos dados das eleições de 1996 e 2000, em segundo turno, e de dados demográficos da Contagem 1996 e do Censo Demográfico 2000.

Para alcançar tal objetivo, foram utilizadas duas metodologias distintas: um modelo linear clássico (aqui denominado Modelo de Goodman) e um modelo estatístico bayesiano, este último desenvolvido por King (1997). Após uma breve apresentação dos fundamentos estatísticos e matemáticos dos modelos de Goodman e de King, foram estimados dez modelos binários distintos, que relacionavam sempre uma variável eleitoral a uma variável demográfica.

Os modelos estimados para a eleição de 1996 buscaram explicar a votação obtida pela candidata petista derrotada Luiza Erundina através de três variáveis demográficas distintas: sexo, idade e escolaridade. Estas variáveis foram obtidas da Contagem 1996 do IBGE. Já os modelos para a eleição de 2000 analisaram os votos obtidos pela então candidata petista Marta Suplicy e as variáveis demográficas sexo, idade, raça, religião, escolaridade, renda e origem migratória.

Entre os dez modelos estimados apenas cinco apresentaram resultados cujos instrumentos de análise disponíveis apontaram como robustos: os modelos para as variáveis escolaridade e idade nas eleições de 1996 e para as variáveis raça, escolaridade e renda nas eleições de 2000. As estimativas das variáveis demográficas idade e origem migratória na eleição de 2000 apresentaram características estatísticas razoáveis, enquanto que os modelos para a variável religião na eleição de 2000 e da variável demográfica sexo nas duas eleições não apresentaram ajustes minimamente aceitáveis.

Entre os resultados obtidos destaca-se a proximidade existente entre os parâmetros estimados pelo modelo linear de Goodman e os parâmetros obtidos pelo sofisticado modelo de King. As estimativas destes modelos apenas diferem de modo significativo quando os parâmetros estimados pelo modelo linear se situam fora do intervalo (0,1), ou quando os dados apresentam pequena variação no espaço.

Apesar dos parâmetros agregados estimados pelo modelo de King não trazerem muitas novidades em relação aos parâmetros estimados pelo modelo linear de Goodman, o mesmo possui o benefício de possibilitar a estimativa de parâmetros por micro-regiões – neste caso, áreas de ponderação do município de São Paulo. E os resultados obtidos são

bastante interessantes. Os parâmetros estimados para a variável idade, por exemplo, indica que os eleitores mais jovens votaram sistematicamente mais nas candidatas petistas de 1996 e 2000 do que nos candidatos Celso Pitta e Paulo Maluf, independentemente do local do município em que moravam.

Adicionalmente, os parâmetros estimados para as variáveis escolaridade e renda na eleição de 2000 evidenciam a existência de uma região de reduzida penetração dos candidatos petistas no eixo que parte do centro da capital em direção à zona leste.

Os parâmetros estimados para a variável demográfica “raça” relativos à eleição de 2000 indicam que a elevada aceitação da candidata Marta Suplicy entre os eleitores negros, pardos e indígenas independe da localização espacial destes grupos sociais.

Entre os parâmetros agregados estimados, merecem destaque os dados relativos à origem migratória na eleição de 2000, que indicam que, entre os eleitores nascidos no município de São Paulo, apenas 42% votaram em Marta Suplicy no segundo turno. Ou seja, os parâmetros estimados indicam que os eleitores aqui nascidos elegeriam Paulo Maluf, não Marta Suplicy para a prefeitura naquele ano.

Os dados de escolaridade e renda, por sua vez, mostram um aumento da penetração da candidata petista no conjunto de eleitores de renda e escolaridade superior à média do município, o que – muito provavelmente – garantiu a eleição da prefeita Marta Suplicy no pleito de 2000. Deve-se destacar que o resultado eleitoral do PT na população com escolaridade maior que a média do município na eleição anterior foi bastante pífio. O modelo para a eleição de 1996 estimou que a então candidata Luiza Erundina obteve apenas 5,5% dos votos dos eleitores que possuíam um maior nível educacional. Em 2000, este valor alcançou 51%. Finalmente, merece destaque a elevada votação obtida pela candidata Marta Suplicy entre os eleitores negros, pardos e indígenas e entre os eleitores com idade entre 16 e 39 anos.

Evidentemente, a inferência de parâmetros individuais a partir de dados agregados não pode ser realizada de modo inadvertido. A ausência de estatísticas confiáveis de avaliação dos parâmetros estimados e a simplicidade da hipótese de autocorrelação espacial embutida no modelo de King são pontos de fragilidade deste tipo de estimação que não podem deixar de ser consideradas por um leitor menos atento.

Por outro lado, as fragilidades que podem caracterizar as estimações dos parâmetros aqui obtidos têm como contrapartida a existência de um amplo campo de extensões que podem

ser aplicadas na busca do aprimoramento e da corroboração dos resultados obtidos nesta primeira investida no campo da Inferência Ecológica. Uma primeira extensão do trabalho aqui apresentado seria a aplicação de modelos do tipo HASS, muito comuns no campo da economia regional e urbana, os quais poderiam ser usados para checar a consistência dos parâmetros estimados através de modelos de inferência ecológica. Também parece natural tentar estender esta primeira análise binária realizada por meio do software EzI para modelos mais amplos e multivariados a serem estimados por meio do software EI, também desenvolvido por King e disponível em <http://gking.harvard.edu/stats.shtml>. Nesta iniciativa, dever-se-ia levar em conta novas formas funcionais para tentar modelar eventuais problemas de autocorrelação espacial que estivessem presentes nos dados. Finalmente, os modelos dinâmicos baseados em cadeias de Markov desenvolvidos por Pelzer *et al* (2002) para “cross sections” repetidas também parecem ser um fértil campo de análise na investigação dos determinantes sociais do voto em eleições consecutivas.

7. Bibliografia

ACHEN, C and W. SHIVELY (1995), "Cross-Level Inference", University of Chicago Press.

ALESINA, A. (1987), "Macroeconomic Policy in a Two-Party System as a Repeated Game", *Quarterly Journal of Economics*, n°102, 663-688.

ANSELIN, L. and W. TAM CHO (2002), "Spatial Effects and Ecological Inference", *Political Analysis*, 10, 276-297.

ANSELIN, L. (1988), "Spatial Econometrics: Methods and Models", Boston: Kluwer Academic Publishers.

COX, G. (1999), "Electoral Rules and Electoral Coordination", *Annual Reviews of Political Science*, 2:145-61.

CROSS, P. and C. MANSKI (1999), "Regressions, Long and Short", *mimeo*, Northwestern University.

DUNCAN, O. and B. DAVIS (1953), "An Alternative to Ecological Correlation", *American Sociological Review*, 18, 665-666.

FREEDMAN, D., S. KLEIN, J. SACKS, C. SMYTH, and C. EVERETT (1991), "Ecological Regression and Voting Rights", *Evaluation Review*, 15, 673-711.

FREEDMAN, D., S. KLEIN, M. OSTLAND, and M. ROBERTS (1998), "On "Solutions" to the Ecological Inference Problem", *mimeo*, University of California at Berkeley.

FREEDMAN, D., S. KLEIN, M. OSTLAND, and M. ROBERTS (1998), "Response to King's Comment", *Journal of the American Statistical Association*, 94, 355-357.

FREEDMAN, D., S. KLEIN, M. OSTLAND, and M. ROBERTS (1998), "Review of A Solution to the Ecological Inference Problem, by G. King", *Journal of the American Statistical Association*, 93, 1518-1522.

FREEDMAN, D. (1999), "Ecological Inference and the Ecological Fallacy", *mimeo*, University of California at Berkeley.

GOODMAN, L. (1953), "Ecological Regressions and Behavior of Individuals", *American Sociological Review*, 18, 663-664.

GOODMAN, L (1959), "Some Alternatives to Ecological Correlation", *American Journal of Sociology*, 64, 610-624.

GREENE, W. (2003), *Econometric Analysis*, Prentice Hall – 5th Edition.

KING, G. (1997), *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton University Press.

KING, G. (1999), "The Future of Ecological Inference Research: A Comment on Freedman et.al.", *Journal of the American Statistical Association*, 94, 352-355.

LEWIS, J. (1998), "Method-of-Moment Estimators for Kings's Ecological Inference Model", artigo apresentado no "Annual Meeting of the American Political Science Association".

MATTOS, R. and A. VEIGA (2000), "Estimating King's Ecological Inference Normal Model via the EM Algorithm", *mimeo*, Universidade Federal de Juiz de Fora.

MILLER, W. L., G. RAAB and K. BRITTO (1974), "Voting Research and the Population Census 1918-71: Surrogate Data for Consituency Analysis", *Journal of the Royal Statistical Society*, Vol. 137, Issue 3, 384-411.

MOFFITT, R. (1993), "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross Sections", *Journal of Econometrics*, 59, 99-123.

NORDHAUS, W. (1975), “*The Political Business Cycle*”, Review of Economic Studies: 42 169-190.

OGBURN, W. and I. GOLTRA (1919), “How Women Vote: A Study of an Election in Portland, Oregon”, *Political Science Quarterly*, 34, 413-433.

O’LOUGHLIN, J. (2000), “Can King’s Ecological Inference Method Answer a Social Scientific Puzzle: Who Voted for the Nazi Party in Weimar Germany?”, *mimeo*, University of Colorado at Boulder.

PELZER, B., R. EISINGA and P. FRANSES (2001), “Estimating Transition Probabilities from a Time Series of Independent Cross Sections”, *Statistica Neerlandica*, 55, 249-262.

PELZER, B., R. EISINGA and P. FRANSES (2002), “Inferring Transition Probabilities from Repeated Cross Sections”, *Political Analysis*, 10, 113-133.

PELZER, B., R. EISINGA and P. FRANSES (2002), “Ecological Panel Inference in Repeated Cross Sections”, *mimeo*, Erasmus University Rotterdam.

PELZER, B. and R. EISINGA (2002), “Bayesian Estimation of Transition Probabilities from Repeated Cross Sections”, *Statistica Neerlandica*, 56, 23-33.

PENUBARTI, M. and A. SCHUESSLER (1998), “Inferring Micro- from Macrolevel Change: Ecological Panel Inference in Surveys”, *mimeo*, University of California at Los Angeles.

RIVERS, D. (1998), “A Nonparametric Method of Ecological Inference”, artigo apresentado no “Annual Meeting of the American Political Science Association”.

ROBINSON, W. S. (1950), “Ecological Correlations and the Behavior of Individuals”, *American Sociological Review*, 15, 351-357.

ROGOFF, K. (1990), "Equilibrium Political Budget Cycles", *The American Economic Review*: n°80, 21-36.

SIMMONS, J. (1965), "Voting Behavior and Social-Economic Characteristics: The Middlesex East Federal Election" *The Canadian Journal of Economics and Political Science*, Vol. 33, 389-400.

SPRAGUE, J. (1976), "Estimating a Boudon Type Contextual Model: some Practical and Theoretical Problems of Measurement" *Political Methodology* 3, 333-353.

ROSS, S. M. (1996), "Stochastic Processes". San Diego CA: Academic Press.

TAM CHO, W. K. (1999), "Iff the Assumption Fits... A Comment on the King Ecological Inference Solution", *Political Analysis*, 7, 143-163.