

**Renoir Vieira**

**Teoria dos Jogos aplicada à Análise de Eventos Políticos**

Renoir Vieira

## **Teoria dos Jogos aplicada à Análise de Eventos Políticos**

Monografia apresentada ao curso de Ciências  
Econômicas, como requisito parcial para obtenção do grau  
de Bacharel do Insper – Ibmec São Paulo.

Orientador: Prof. Dr. Carlos Melo – Insper – Ibmec São Paulo

São Paulo

2011

## Sumário

1. Introdução	4
2. Metodologia	7
2.1 Modelo de Preferência do Eleitor	8
2.2 Coleta de Dados	9
2.3 Estimação dos Modelos Econométricos	12
2.4 Motivação para Estimação dos Modelos da Classe Nested Logit	16
2.5 Teste para Evidência de Comportamento estratégico dos legisladores	19
3. Análise Preliminar dos Dados	21
4. Resultados	29
5. Conclusões	44
6. Referências	46

## 1. INTRODUÇÃO

A política tem papel fundamental no funcionamento das sociedades modernas. É através do sistema político que crenças, valores e interesses são coordenados, para que decisões coletivas sejam alcançadas. Os sistemas políticos diferem quanto ao peso que atribuem às preferências de grupos ou indivíduos na formação da decisão coletiva.

Eventos políticos emergem do processo de acomodação das crenças, valores e interesses individuais em torno da decisão coletiva. Estes eventos têm consequências não apenas sociais, mas também econômicas. Por exemplo, as recentes revoltas populares no mundo árabe aumentam a incerteza sobre o fornecimento de combustíveis minerais e contribuem para aumentar o preço destes combustíveis no mercado internacional.

Assumindo que os agentes são racionais e têm preferências em relação aos resultados, e que, além disso, gozam de liberdade para fazer suas escolhas políticas; considerando, ainda, que há indiferença em relação a quaisquer características pessoais ou psicológicas dos candidatos, pode-se aproximar o comportamento do sistema político de um mercado de *commodities* em que políticos vendem posições sobre diferentes temas aos eleitores, como é apontado por Tsai e Hsu (2005).

Nessas condições, o processo de decisão do eleitor pode ser modelado de maneira similar ao processo de decisão do consumidor que é amplamente estudado em microeconomia.

O Modelo do eleitor mediano, desenvolvido por Black (1948) e Downs (1957), é uma primeira aproximação. No modelo de Downs (1957), tudo se passa como se os eleitores estivessem distribuídos ao longo de um espectro político e a votação acontecesse após os candidatos escolherem simultânea e independentemente qual seu posicionamento ao longo do espectro político. Assumindo que esse espectro possa ser representado pelo intervalo contínuo  $[0,1]$ , em que 0 é extremamente liberal e 1 extremamente conservador, por exemplo. Supondo ainda que a função  $F(x)$  descreve a distribuição acumulada dos eleitores ao longo do espectro político, em uma eleição com dois

candidatos, eles têm incentivos para escolher seu posicionamento político de modo que  $F(x^*) = 1/2$  e, portanto, os candidatos no equilíbrio escolhem o mesmo posicionamento político.

Assumindo que as preferências dos eleitores apresentam um único ponto de máximo, os eleitores elegerão o candidato, cuja cesta de bens e serviços públicos ofertados mais se aproximar da cesta demandada pelo eleitor mediano em um sistema eleitoral majoritário. Menezes, Saiani e Zoghbi (2010), que testam as evidências empíricas do modelo do eleitor mediano para os municípios brasileiros, a partir de dados das eleições municipais de 2000, encontram evidências que suportam a validade do modelo.

A metodologia aplicada por Menezes *et al.* (2010) baseia-se em estimar a demanda por bens públicos. Assume-se que a demanda do eleitor mediano é igual aos gastos do município. O erro da regressão entre a demanda estimada para o eleitor mediano e os gastos do município é utilizado como erro de percepção do governante em relação a demanda do eleitor mediano. Quanto menor o erro do candidato, menor a chance de o candidato se reeleger. Os autores propõem uma formulação Heckit já que existe problema de viés de auto-seleção, uma vez que o candidato só decide tentar a reeleição se acredita ter probabilidade razoável de se reeleger.

Uma aproximação mais geral é apresentada por Myerson e Weber (1993), que propõem um modelo em que o comportamento do eleitor depende de suas preferências e de sua percepção acerca de viabilidade do candidato. Quando há um número grande de candidatos concorrendo para o mesmo cargo, o eleitor pode não votar em seu candidato preferido, se este for percebido como inviável.

O aumento da abstenção no segundo turno das eleições pode ser uma evidência deste comportamento. Em 2010, por exemplo, o candidato José Serra acabou derrotado por uma diferença de 11 pontos percentuais, mais do que o previsto por qualquer uma das pesquisas pré-eleitorais, inclusive acima da margem de erro das pesquisas. Uma possível explicação para esse fato está relacionada à expectativa do eleitor de que o candidato seria derrotado,

portanto, o custo de votar relacionado a ir até o posto de votação, esperar na fila, e votar se tornou maior do que o benefício esperado, dado que o candidato era visto como inviável.

Dada a relevância social e econômica do processo político, é natural que os agentes tenham incentivos para tentar antecipar e até mesmo influenciar o curso de eventos políticos relevantes. Restringiu-se até aqui a discussão a análise de eventos eleitorais por serem ilustrativos, contudo, outros eventos como a aprovação de uma legislação ou o posicionamento de um país, diante de uma demanda internacional também têm muita relevância.

No caso de previsão do resultado de eleições, é comum o uso de levantamento de intenção de voto. A principal motivação para uso desse método é o fato de ser um método simples, bem conhecido e que depende de dados que podem ser levantados com qualquer frequência. Este método apresenta diversas falhas, como apontado por Kamakura, Mazzon, e Bruyn (2006). Esses ressaltam que os eleitores podem: i) apresentar comportamento estratégico enviando sinais para os candidatos por meio de sua intenção de voto declarada, ii) não ter as decisões de voto claras até enfrentarem a decisão real na hora da votação e iii) sentirem-se envergonhados de revelarem sua real intenção de voto, em especial, quando o levantamento é feito presencialmente. O levantamento de intenção de voto pode inferir a preferência declarada, contudo, esta nem sempre é a preferência revelada. Além disso, este método se baseia em amostra relativamente pequena que pode não ser diversa o suficiente para capturar corretamente a heterogeneidade dos eleitores.

Na previsão do resultado de votação de legislação ou do posicionamento de um país diante da comunidade internacional, geralmente, considera-se as opiniões de especialistas que se baseiam no conhecimento prévio dos atores envolvidos, de seu comportamento anterior e do seu possível envolvimento com a situação que se quer analisar. O principal ponto fraco deste método é o fato de não produzir uma estrutura analítica baseada em um arcabouço teórico consistente, sobre a qual o evento em questão possa ser analisado como discutido por Mesquita (1984).

Mesquita (1984) argumenta que nesse último caso a construção de um modelo de teoria dos jogos que leve em consideração a estrutura de incentivos na qual os agentes estão inseridos é a melhor maneira de gerar previsões baseadas em um modelo internamente consistente e capaz de gerar um arcabouço analítico que permite analisar diversos cenários comparáveis.

O objetivo deste trabalho é colocar duas questões. A primeira diz respeito à previsibilidade de um evento político específico: resultados de eleições com dois turnos em que o resultado do primeiro turno é conhecido com certa segmentação antes do segundo turno, em um esforço parecido com o desenvolvido por Kamakura *et al.* (2006). A segunda questão que se coloca é se com base nas preferências relevadas pelo eleitor, através da votação, é possível prever o resultado de outros eventos, tais como o resultado da tentativa de aprovação de certas matérias pelo governo.

## 2. METODOLOGIA

A metodologia deste trabalho está dividida em duas partes: a primeira consiste em i) propor um modelo teórico para a preferência do eleitor, ii) coletar os dados relevantes para estimação das diversas especificações possíveis do modelo, iii) estimar diversas especificações possíveis, a saber *logit multinomial*, e *nested logit factor model* e iv) comparação dos resultados.

A segunda parte consiste em buscar evidências para a hipótese de que os legisladores votam de acordo com a percepção que formam sobre as preferências dos eleitores em relação a um *mix* de políticas. Para testar essa hipótese analisaremos um caso recente: a votação do Novo Código Florestal. Cerca de 60 legisladores votaram contra a aprovação do Novo Código Florestal, alguns desses contrariando a recomendação de seus próprios partidos. Será proposto um modelo de escolha discreta para tentar capturar a probabilidade de o legislador votar a favor do Código Florestal, dadas variáveis como partido ao qual é filiado, estado pelo qual foi eleito e a porcentagem de votos que recebeu a candidata Marina Silva nas cinco cidades onde o

legislador recebeu maior votação. A metodologia de teste será discutida na seção 2.6.

## 2.1 Modelo de Preferência do Eleitor

Supõe-se que a preferência do eleitor dependa de características pessoais como idade, sexo, renda, escolaridade, estado onde mora, além da oferta de serviços públicos.

Uma especificação possível para o modelo de preferência do eleitor é dada por

$$U_{ij} = \alpha_j + \beta_{ij}(idade_i) + \beta_{ij}sexo_i + \beta_{ij} \log(renda_i) + \beta_{ij}escolaridade_i + \beta_{ij}estado_i + \beta_{ij}ORC_i \quad (1)$$

onde,  $U_{ij}$  é a preferência pelo candidato  $j$  do individuo  $i$ ,  $\alpha_j$  é o capital político do candidato  $j$  entre todos os eleitores, a idade é a idade medida em anos de vida do individuo, sexo é uma *dummy* com valor 1 para sexo masculino, renda é a renda medida em Reais, escolaridade é medida em anos de estudo, estado é uma *dummy* com valor 0 para São Paulo e ORC é o orçamento per capita do município, usada como uma *proxy* da oferta de serviços públicos.

Mas, de modo geral, não se tem a votação de cada um dos eleitores. Como foi descrito anteriormente, os dados foram coletados por cidade, caso em que o modelo proposto terá que ser modificado para acomodar essa dificuldade.

Assim,

$$U_{ij} = \alpha_j + \beta_{ij} \log(idade_i) + \beta_{ij}sexo_i + \beta_{ij} \log(renda_i) + \beta_{ij}escolaridade_i + \beta_{ij}estado_i + \beta_{ij}ORC_i \quad (2)$$

nesse caso,  $U_{ij}$  é a preferência pelo candidato  $j$  na cidade  $i$ ,  $\alpha_j$  é o capital político do candidato  $j$  entre todos os eleitores, idade é a pirâmide de idade, sexo é o número de homens menos o número de mulheres, renda é a pirâmide de renda, escolaridade é a pirâmide de escolaridade, estado é uma *dummy*



com valor zero para São Paulo e *ORC* é o orçamento per capita do município usada como uma *proxy* da oferta de serviços públicos.

As leis eleitorais brasileiras impedem o acesso aos dados de votação por eleitor de modo que as especificações (1) e (2), não podem ser estimadas diretamente. Nesse trabalho foi feita a escolha de utilizar dados por município uma vez que para os municípios brasileiros existem dados tanto sobre a votação quanto sobre as características sociais e demográficas. Os dados de votação estão disponíveis em um nível ainda mais segmentado, existem dados disponíveis por seção eleitoral, contudo, há uma dificuldade em se obter dados tão segmentados sobre características do eleitorado.

## 2.2 Coletas de Dados

A coleta de dados foi realizada durante a campanha presidencial de 2010, quando o autor participou da campanha do candidato José Serra na área de inteligência.

A base de dados é composta por dados coletados junto à base *IBGE@Cidades* que compreende uma vasta coleção de dados reunidos por diversas fontes oficiais e agrupados por município. Os dados foram reunidos e classificados por meio de um *software* escrito pelo autor em linguagem *JAVA*.

Além destes dados, foram reunidos dados de votação por candidato à presidência por município. Estes dados são disponibilizados pelo TSE em formato “*txt*”.

A Tabela 1 descreve as variáveis disponíveis no banco de dados.

Nome	Breve Descrição
COD_MUN	Código do Município IBGE
MUNICIPIO	Nome do Município
ESTADO	Estado
EST_POP_2009	População Estimada em 2009
AREA	Área do Município
ELEITORADO_2006	Número de Eleitores em 2006

PIB_P_CAPITA_2007	PIB per Capita em 2007
MATR_ENS_FUND_2009	Número de Matrículas no Ensino Fundamental em 2009
DOC_ENS_FUND_2009	Número de Ens. Fund. Docentes em 2009
DOC_ENS_MED	Número de Docentes Ensino Médio 2009
EST_SUS_2005	Número de Estabelecimentos do SUS em 2005
NASC_VIVOS_2008	Nascidos vivos em 2008
REC_ORC_REAL_2008	Receita Orçamentária Real em 2008
DESP_ORC_REAL_2008	Despesa Orçamentária Real em 2008
FPM_2008	Recursos Fundo de Municípios em 2008
EMP_LOCAIS_2008	Número de empresas locais em 2008
PES_OCU_2008	Pessoal ocupado em 2008
LAT	Latitude
LONG	Longitude
RESIDENTES_2001	Residentes em 2001
RESIDENTES_10_MAI_2001	Residentes mais de 10 anos de idade em 2001
M_RES_10_MAI_2001	Mulheres residentes mais de 10 anos de idade 2001
H_RES_10_MAI_2001	Homens residentes mais de 10 anos de idade 2001
RES_10_MAI_SEM_INST	Residentes mais de 10 anos de idade sem instrução
RES_REN_1_SAL	Residentes com renda até um salário mínimo
RES_REN_1_2_SAL	Residentes com renda entre um e dois salários mínimos
RES_REN_2_3_SAL	Residentes com renda entre dois e três salários mínimos
RES_REN_3_5_SAL	Residentes com renda entre três e cinco salários mínimos
RES_REN_5_10_SAL	Residentes com renda entre cinco e dez salários mínimos
RES_REN_10_20_SAL	Residentes com renda entre dez e vinte salários mínimos
RES_REN_20_SAL	Residentes com renda entre vinte ou mais salários mínimos
RES_REN_SEM	Residentes mais de 10 anos de idade sem renda.

RES_REN_COM	Residentes mais de 10 anos de idade com renda
HOM_C_REND	Homens Residentes com mais de 10 anos de idade com renda
MUL_C_REND	Mulheres Residentes com mais de 10 anos de idade com renda
REND_MED	Renda média residentes com mais de 10 anos de idade
H_REND_MED	Renda média dos homens com mais de 10 anos de idade
M_REND_MED	Renda média dos mulheres com mais de 10 anos de idade
DOM_PERM_2001	Domicílios Permanentes em 2001
DOM_PERM_C_REND	Domicílios Permanentes com renda em 2001
PES_RES_10_1_EST	Pessoas com mais de 10 anos com até ano de escolaridade
PES_RES_10_1_3_EST	Pessoas com mais de 10 anos com entre um e três anos de escolaridade
PES_RES_10_4_7_EST	Pessoas com mais de 10 anos com entre quatro e sete anos de escolaridade
PES_RES_10_8_10_EST	Pessoas com mais de 10 anos com entre oito e dez anos de escolaridade
PES_RES_10_11_14_EST	Pessoas com mais de 10 anos com entre onze e quatorze anos de escolaridade
PES_RES_10_15_EST	Pessoas com mais de 10 anos com 15 anos ou mais de escolaridade
PES_RES_0_3_ANOS	Pessoas residentes com idade entre 0 e 3 anos
PES_RES_4_ANOS	Pessoas residentes com 4 anos de idade
PES_RES_5_6_ANOS	Pessoas residentes com idade entre 5 e 6 anos
PES_RES_7_9_ANOS	Pessoas residentes com idade entre 7 e 9 anos

PES_RES_10_14_ANOS	Pessoas residentes com idade entre 10 e 14 anos
PES_RES_15_17_ANOS	Pessoas residentes com idade entre 15 e 17 anos
PES_RES_18_19_ANOS	Pessoas residentes com idade entre 18 e 19 anos
PES_RES_20_24_ANOS	Pessoas residentes com idade entre 20 e 24 anos
PES_RES_25_29_ANOS	Pessoas residentes com idade entre 25 e 29 anos
PES_RES_30_39_ANOS	Pessoas residentes com idade entre 30 e 39 anos
PES_RES_40_49_ANOS	Pessoas residentes com idade entre 40 e 49 anos
PES_RES_50_59_ANOS	Pessoas residentes com idade entre 50 e 59 anos
PES_RES_60_64_ANOS	Pessoas residentes com idade entre 60 e 64 anos
PES_RES_65_69_ANOS	Pessoas residentes com idade entre 65 e 69 anos
PES_RES_70_74_ANOS	Pessoas residentes com idade entre 70 e 74 anos
PES_RES_75_79_ANOS	Pessoas residentes com idade entre 75 e 79 anos
PES_RES_80_ANOS	Pessoas residentes com 80 anos ou mais
BOL_FAM	Número de famílias no programa Bolsa Família
SERRA	Votos nominais no candidato José Serra
DILMA	Votos nominais na candidata Dilma Roussef
MARINA	Votos nominais na candidata Marina Silva
OUTROS	Votos nominais em outros candidatos
VOTOS_VALIDOS_2010	Quantidade de Votos Válidos em 2010

Tabela 1. Listagem das Variáveis incluídas no Banco de Dados

Além desta base, foi necessário levantar dados de votação por cidade para os deputados eleitos em 2010. Esta base será utilizada para testar a evidência de que os deputados votam certos temas com base da percepção que forma sobre as preferências dos eleitores, a partir do resultado das eleições. Estes dados podem ser obtidos junto ao TSE e foram coletados

utilizando uma versão modificada do *software* desenvolvido para coletar os dados de votação dos candidatos a presidência.

### 2.3 Estimação dos Modelos Econométricos

Os modelos econométricos estimados serão: i) *logit* multinomial, ii) *nested logit*, iii) *nested logit factor model* na formulação proposta por Kamakura et al (2006).

#### 2.3.1 Modelo *Logit* Multinomial

O modelo *Logit* Multinomial é um modelo de escolha discreta que permite que a variável dependente represente mais do que duas possibilidades discretas, ao contrário do modelo *Logit* tradicional em que a variável dependente deve ser binária.

Seja  $y$  a variável resposta e assumindo que esta possa assumir valores como  $\{1, 2, \dots, J\}$  em que  $J$  é um número inteiro, seja  $X$  uma coleção de variáveis explicativas que pode conter, por exemplo, educação, salário, renda, idade, entre outros. Denotamos  $(X_i, y_i)$  como uma realização aleatória da população.

De maneira similar ao modelo *Logit* tradicional, o interesse é estimar como mudanças nos elementos de  $X$  afetam a probabilidade da resposta,  $P(y = j|X), j = 0, 1, 2, \dots, J$ .

Seja  $X$  um vetor  $1 \times K$  de modo que o primeiro seja um vetor de um. A resposta do modelo multinomial é dada por

$$P(y = j|X) = \frac{\exp(X\beta_h)}{1 + \sum_{h=1}^J \exp(X\beta_h)}, \quad j = 1, \dots, J \quad (3)$$

onde  $\beta_h$  é um vetor  $K \times 1$ .

A estimação do modelo pode ser feita pela maximização da função log-verossimilhança. A função log-verossimilhança para cada observação  $i$  na amostra é dada por

$$l_i(\beta) = \sum_{j=0}^J 1[y_i = j] \log(p_j(X_i, \beta))$$

(4)

onde  $1[y_i = j]$  é uma função indicadora que assume valor 1 quando  $y_i = j$ .

### 2.3.2 Modelo *Nested Logit*

O modelo *Logit* Multinomial apresenta uma limitação importante, que se refere à hipótese de independência das alternativas irrelevantes. Em termos práticos, assumir essa hipótese implica que a relação de probabilidade entre duas alternativas é independente de uma terceira alternativa. Portanto, adicionar uma terceira ou modificar as características de uma terceira alternativa não tem qualquer impacto sobre a razão de probabilidade de duas alternativas. No modelo *Logit* Multinomial a razão de probabilidade entre duas alternativas é dada por,

$$\frac{p_j(X, \beta)}{p_h(X, \beta)} = \exp(x(\beta_j - \beta_h))$$

(5)

Essa hipótese pode ser muito restritiva, em particular quando existem alternativas muito similares, o que frequentemente é o caso em eleições majoritárias, por exemplo, em que existem muitos candidatos.

Uma maneira de relaxar a hipótese de independência das alternativas irrelevantes é especificar um modelo do tipo *Nested Logit*. Esse modelo permite que certas alternativas sejam reunidas em grupos de acordo com sua similaridade. A primeira hierarquia consiste em determinar em que grupo  $y$  está e depois a qual alternativa  $y$  corresponde dentro de cada grupo. A probabilidade de  $y$  pertencer a um grupo é dada por,

$$q_s = \frac{[\alpha_s (\sum_{j \in G} \exp(\rho_s^{-1} X_j \beta))]^{\rho_s}}{[\sum_{k=1}^S \alpha_k (\sum_{j \in G} \exp(\rho_k^{-1} X_j \beta))^{\rho_k}]}$$

(6)

em seguida, queremos modelar a probabilidade de  $y$  ser igual a uma determinada alternativa dado que  $y \in G_s$ , esta probabilidade é dada por,

$$p_s = \exp(\rho_s^{-1} X_j \beta) / [\sum_{h \in G_s} \exp(\rho_s^{-1} X_h \beta)]$$

(7)

Este modelo pode ser estimado pela maximização da função log-verossimilhança. A função log-verossimilhança para cada observação  $i$  na amostra é dada por,

$$l_i(\beta, \alpha, \rho) = \sum_{s=1}^S \left[ 1[y \in G_s] \left( \log(q_s(X_i; \beta, \alpha, \rho)) \right) + \sum_{j \in G_s} 1[y = j] \log(p_{sj}(X_i; \beta, \rho_s)) \right] \quad (8)$$

### 2.3.3 Nested Logit Factor Model

Kamakura *et al.* (2006) propõe um modelo para permitir que o resultado final de eleições de dois turnos sejam previstos a partir do resultado do primeiro turno. Assumindo que os eleitores não votam estrategicamente no primeiro turno e que as preferências são aproximadamente constantes entre um turno e outro, é possível prever o resultado da eleição.

O modelo proposto por Kamakura *et al.* (2006) permite a estimação de fatores latentes que levam em consideração a heterogeneidade das preferências dos eleitores em diferentes cidades. Os autores assumem que a preferência dos eleitores na cidade  $i$  depende de características dos candidatos e dos eleitores, a especificação do modelo é

$$U_{ij} = a_j + \beta_j X_i + \gamma_j Z_i + \varepsilon_{ij} \quad (9)$$

onde,

- $\alpha_j$  é representa o intercepto da regressão que está relacionado ao apelo do candidato junto ao eleitorado
- $X_i$  são características demográficas observáveis para a cidade  $i$
- $\beta_j$  é o vetor de coeficientes ligados a características geográficas que determinam o apelo do candidato em relação a essas características demográficas (idade, renda, educação, etc.)
- $Z_i$  é um vetor de variáveis latentes que captura desvios não observados nas preferências dos eleitores da cidade  $i$  em relação a

média da população. Os autores assumem que  $Z_i$  tem distribuição normal padrão.

- $\gamma_j$  é um vetor de pesos dos fatores  $Z_i$  para o candidato  $j$ , esse vetor representa o posicionamento do candidato  $j$  no espaço latente de heterogeneidade dos eleitores.
- $\varepsilon_{ij}$  são componentes randômicos de utilidade identicamente distribuídos com distribuição valor-extremo para dos candidatos e municípios.

Como o modelo permite heterogeneidade nas preferências, o modelo relaxa a hipótese de independência das alternativas irrelevantes, como argumentado por Kamakura *et al.* (2006).

O modelo pode ser estimado usando a função de log-verossimilhança simulada, Kamakura *et al.* (2006), que é dada por

$$\begin{aligned}
 & l(\theta|Y) \\
 &= \sum_i^N \left( \log \sum_{t=1}^T \left[ \frac{\exp(a_0 + \beta_0 X_i + \gamma_0 Z_{it})}{\exp(a_0 + \beta_0 X_i + \gamma_0 Z_{it}) + \exp(\delta \log(\sum_{j=1}^J \exp(a_j + \beta_j X_i + \gamma_j Z_{it})))} \right]^{y_{it}} \right) \\
 &\times \prod_{j=1}^J \left[ \frac{\exp(\delta \log(\sum_{j=1}^J \exp(a_j + \beta_j X_i + \gamma_j Z_{it})))}{\exp(a_0 + \beta_0 X_i + \gamma_0 Z_{it}) + \exp(\delta \log(\sum_{j=1}^J \exp(a_j + \beta_j X_i + \gamma_j Z_{it})))} \right. \\
 &\left. - \log(T) \right]^{y_{ij}}
 \end{aligned}
 \tag{10}$$

A estimação dos parâmetros é dada pela maximização da função log-máximo verossimilhança acima. Os detalhes técnicos podem ser encontrados em Train (2003).

Os *softwares* estatísticos normalmente empregados não fazem essa estimação. O autor entrou em contato com Kamakura que sugeriu a implementação de um código proprietário para estimação dos parâmetros. Este



código foi desenvolvido em *MATLAB* utilizando algoritmos de otimização padrão com gradientes analíticos e hessiano numérico.

#### 2.4 Motivação para Estimação dos Modelos da Classe *Nested Logit*

Supondo que em uma eleição de dois turnos três candidatos estejam concorrendo a um cargo no poder executivo; assumindo que, por exemplo, o resultado do primeiro turno tenha sido iguais aos resultados descritos na Tabela 2.

Candidato	Votação
A	60%
B	30%
C	10%

Tabela 2. Exemplo de resultado de primeiro turno em uma eleição de dois turnos.

Nesse caso, haveria segundo turno entre os candidatos A e B. Dado que tenha sido estimados dois modelos para explicar as preferências dos eleitores um modelo *Logit* Multinomial e um modelo *Nested Logit* tradicional.

A previsão para o resultado do segundo turno dada pelo modelo *Logit* Multinomial consiste em distribuir os 10% de votos do eleitor C, entre os candidatos A e B numa proporção de dois para um, respectivamente. Essa é uma consequência da hipótese de independência das alternativas irrelevantes.

Mas, supondo que o candidato C seja mais parecido com o candidato B, ou seja, os eleitores do candidato C tem mais afinidade com o candidato B do que com o candidato A, nesse caso, seria natural que os votos de C migrassem para o candidato B, numa proporção maior do que para o candidato A. Supondo que 80% dos eleitores de C preferem o candidato ao candidato A, nesse caso, os 10% de votos de C se distribuiriam entre A e B em uma proporção de um para quatro, respectivamente.

O modelo *Nested Logit* é mais completo, pois estima também a proporção na qual os votos serão distribuídos, caso uma alternativa seja

eliminada, e por isso mesmo é mais útil para prever o resultado de uma eleição de dois turnos a partir dos resultados do primeiro turno.

## **2.5 Teste para Evidência de Comportamento estratégico dos legisladores**

Recentemente, foi votado na Câmara dos Deputados o Novo Código Florestal. Este tem dispositivos que pretendem flexibilizar a legislação anterior. Um exemplo disso, é a flexibilização quanto à forma de cálculo para determinação da área de Reserva Legal que agora pode incluir todas as áreas de preservação permanente que anteriormente não faziam parte do cálculo de Reserva Legal. As áreas de preservação permanente incluem, por exemplo, áreas próximas a rios e nascentes e anteriormente.

Como citado anteriormente, alguns legisladores votaram contra a aprovação do Novo Código Florestal, em certos casos contrariando as recomendações de seus próprios partidos. Propõe-se duas explicações para esse comportamento: i) o legislador votou conforme sua consciência e convicção a respeito da matéria, ou ii) o legislador votou considerando a percepção que tem a cerca das preferências de seus eleitores.

Argumenta-se que a o caso em que o legislador vota considerando a percepção que tem a cerca das preferências de seus eleitores é mais comum. Essa hipótese pode ser testada utilizando a estimação de um modelo *Logit*.

### **2.6.1 Especificação do Modelo**

Supondo que os políticos e os cidadãos participam de um jogo dinâmico em que os eleitores votam nos políticos que propõem um *mix* de políticas mais próximo daquele preferido, existe uma geração inicial de políticos que ocupam os cargos de legisladores na primeira rodada. Os legisladores, então, decidem o *mix* de políticas que irão apoiar. Na segunda rodada os eleitores tendo observado as posições de cada um dos legisladores em relação às diversas matérias, escolhem aqueles que irão se reeleger. Ainda na segunda rodada, os eleitores recebem o *payoff* das políticas implementadas pelos políticos que é

dado pela distância entre as preferências do eleitor e as políticas implementadas, e os políticos recebem *payoff* 1 se são reeleitos e 0 se não são reeleitos. Na terceira rodada os políticos observam a votação dada pelos eleitores a cada candidato e decidem suas posições em relação ao *mix* de políticas que vão implementar. As rodadas dois e três são repetidas indefinidamente.

Se os políticos, de fato, observarem o resultados das eleições para formarem suas percepções do *mix* de política demandado pelos eleitores, então, a votação das matérias feita pelos legisladores é influenciada pela votação que os candidatos associados a certos temas tiveram nas cidades.

### 2.6.2 Estratégia de Teste

A candidata Marina Silva está bastante associada ao tema da sustentabilidade e meio ambiente. O partido ao qual a candidata era associada à época da eleição é o Partido Verde, ambos o partido e a candidata eram contrários à aprovação do Novo Código Florestal. Assumindo que o voto na candidata Marina Silva não é um voto baseado nas características pessoais da mesma e sim um voto no *mix* de políticas defendidas por ela, é razoável pensar que nas cidades onde a candidata teve maior votação, o eleitor valoriza o *mix* de políticas defendidas por ela e, portanto, é também contrário a aprovação do Novo Código Florestal.

Assim, a estratégia de teste consiste em estimar um modelo *Logit* para modelar a probabilidade de um legislador  $k$  votar a favor da aprovação do Novo Código Florestal. Serão incluídas como variável explicativa uma *dummy* para levar em consideração partido ao qual o legislador é filiado, uma *dummy* para considerar o estado pelo qual o legislador foi eleito e a votação da candidata Marina Silva nas cinco cidades onde o legislador obteve mais votos. O modelo assume a seguinte forma,

$$V_k = \Phi(\rho_0 + \rho_1 Estado_k + \rho_2 Partido_k + \rho_3 Marina_{kj}) + \epsilon_k \quad (12)$$

onde,

- $Estado_k$  é o estado pelo qual o legislador  $k$  foi eleito
- $Partido_k$  é o partido pelo qual o legislador
- $Marina_{kj}$  é a proporção de votos da candidata Marina Silva no  $j$ -ésimo Estado em que o legislador  $k$  obteve mais votos

Se os legisladores levam em consideração o resultado das eleições para formar suas percepções sobre o *mix* de políticas preferidas pelos eleitores, então, os coeficientes estimados para as variáveis associadas à proporção de votos da candidata Marina Silva nas cinco cidades onde o legislador  $k$  obteve maior votação devem ser negativos e conjuntamente significantes, o que implica em dizer que o fato de o legislador ter sua base eleitoral em cidades onde o *mix* de políticas representado pela candidata Marina Silva diminui a probabilidade de o legislador votar a favor da aprovação do Novo Código Florestal.

### 2.6.3 Modelo *Logit*

O modelo *logit* tem o objetivo de modelar a decisão discreta binária, esse modelo estima a relação entre certas variáveis explicativas e a probabilidade de escolha de cada alternativa.

A forma geral de um modelo logit é dada por

$$Prob(y = 1|X) = G(X\beta) \equiv p(X) \quad (13)$$

onde  $\beta$  é o vetor de coeficientes associados as variáveis explicativas reunidas em  $X$ . A função  $G(\cdot)$  é a função de transferência, esta função garante que as probabilidades previstas estejam dentro de intervalo  $[0,1]$ . No caso do modelo *Logit* a função  $G(\cdot)$  assume a forma

$$G(X\beta) = \exp(X\beta) / (1 + \exp(X\beta)) \quad (14)$$

Os coeficientes  $\beta$  podem ser estimados pela maximização da função de log-verossimilhança definida por

$$l_i(\beta) = y_i \log(G(X_i\beta)) + (1 - y_i)\log(G(X_i\beta)) \quad (15)$$

para cada observação  $i$  na amostra. Os softwares econométricos normalmente utilizados são capazes de estimar os coeficientes  $\beta$ .

#### 2.6.4 Teste de Hipóteses

Dado o modelo definido pela equação (12), o teste de hipótese a ser conduzido é representado por,

$$H_0: \rho_3 = 0$$

$$H_A: \rho_3 \neq 0$$

onde  $\rho_3$  é o coeficiente estimado para a proporção de votação da candidata Marina Silva no Estado em que o legislador foi eleito votação.

Será utilizado o teste de  $t$  em que o modelo é estimado conforma especificação eq.12 e é utilizada a estatística da por

$$t = \frac{\rho_3}{S_{\rho_3}}$$

que segue uma distribuição t-Student.

A interpretação do teste é simples, se o coeficiente estimado for significativo, existem evidências de que os legisladores consideram o resultado da eleição para formação de suas percepções a cerca das preferências dos eleitores nas localidades onde mantém suas bases eleitorais. Caso o coeficiente não seja significantes existem evidências contrárias à hipótese de que os legisladores formam suas percepções a partir dos resultados eleitorais.

### 3. ANÁLISE DE DADOS

Os dados reunidos foram organizados e foi gerada uma análise descritiva dos dados, assim como a estimação inicial da primeira especificação dos modelos propostos.

Os resultados desta primeira estimação estão de acordo com o esperado e servirão de base à proposta de especificação do modelo final a ser estimado.

### **3.1 Análise Descritiva dos Dados**

A base de dados reunida contém um número bastante grande de variáveis, como descrito na seção 2.2, assim um estudo preliminar dos dados é essencial para selecionar as variáveis relevantes ao problema e garantir a parcimônia do modelo a ser estimado.

A análise descritiva dos dados é aqui implementada de duas maneiras: i) a análise descritiva estatística dos dados quantitativos associados a porcentagem de votação dos candidatos à presidência por cidade assim como as características socioeconômicas dos municípios; e ii) a geolocalização dos dados de votação de modo a descrever qualitativamente as características espaciais de distribuição de votos por candidato para tanto o autor utilizou o serviço *SpatialKey* que é um serviço em nuvem para geolocalização de dados.

### **3.2 Análise Descritiva Qualitativa dos Dados**

Como descrito na seção 3.1, optou-se por fazer uma análise descritiva qualitativa dos dados por meio da geolocalização dos dados de distribuição espacial da população, distribuição de município, distribuição espacial dos votos válidos e a distribuição espacial da votação por candidato.

A motivação para esta análise é verificar qualitativamente se existe uma tendência sistemática no padrão de votação favorecendo ou não algum dos candidatos em determinada região.

A primeira consideração que se deve fazer é que existe uma distribuição bastante concentrada da população ao longo da costa. Além disso, existe uma clara concentração populacional na região Sul e Sudeste (Fig. 01).



Fig.01 – Distribuição Espacial da População Brasileira  
(vermelho indica maior concentração populacional)

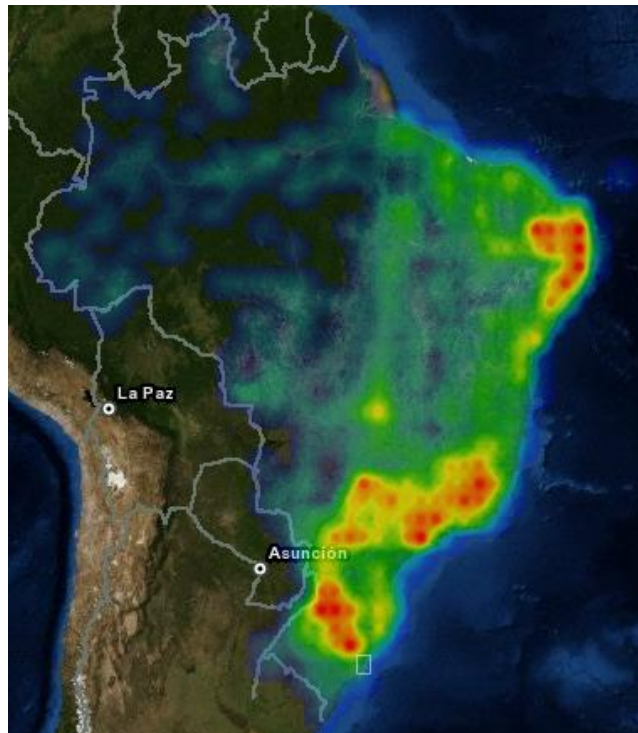


Fig.02 – Distribuição Espacial dos Municípios Brasileiros  
(vermelho indica maior concentração de municípios)

A distribuição espacial de municípios segue um padrão bastante similar ao padrão de distribuição da população (Fig.01) com concentração de municípios ao longo da costa e nas regiões Sul e Sudeste (Fig.02).

Uma variável bastante importante na determinação do resultado eleitoral é a quantidade de votos válidos. O número de votos válidos determina a base para o cálculo do percentual de votos obtidos por cada candidato. A quantidade de votos válidos é determinada pelo número de votos destinados a candidatos concorrentes para determinado cargo, ou seja, são computados apenas votos não nulos ou brancos.

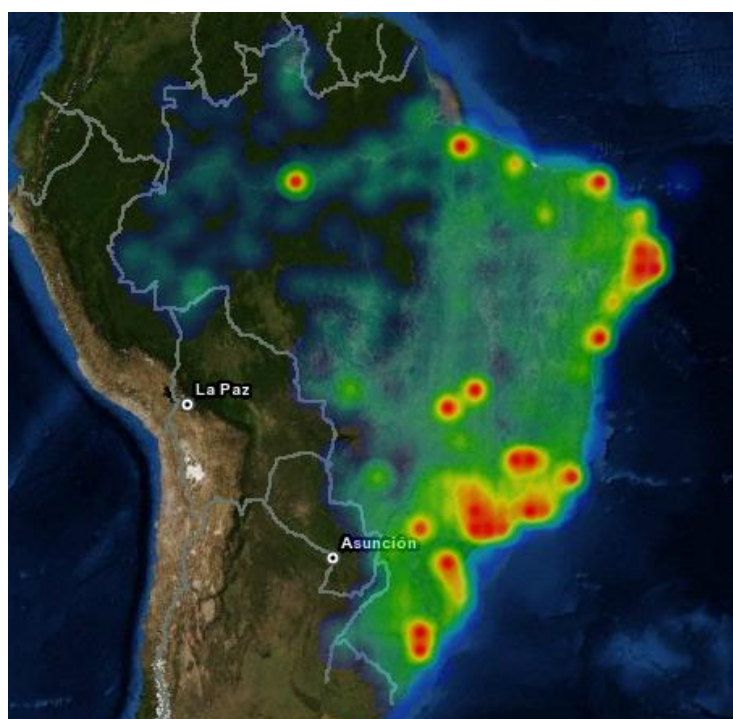


Fig.03 – Distribuição Espacial dos Votos Válidos na Eleição Presidencial  
(vermelho indica maior concentração de votos válidos)

É interessante notar que a distribuição de votos válidos é praticamente idêntica à distribuição populacional com uma diferença apenas de escala, o que leva naturalmente a hipótese de que os eleitores dispostos a anular, votar em branco ou que não poderão comparecer a votação estão distribuídos proporcionalmente entre os municípios.



A distribuição de votos por candidato por candidato visa verificar a tendência sistemática de votação em um determinado candidato espacialmente como citado anteriormente. Procurou-se adotar a mesma escala para que os mapas entre os candidatos fossem comparáveis. Para construção dos mapas adotou-se a distribuição da votação percentual dos candidatos em relação aos votos válidos renormalizados para considerar apenas os candidatos José Serra, Dilma Rouseff e Marina Silva, excluindo, portanto, candidatos menos expressivos.

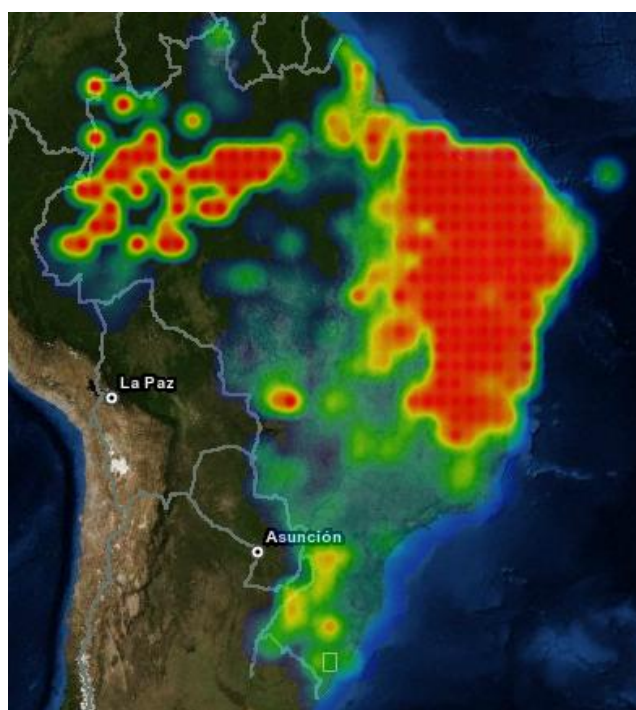


Fig.04 – Distribuição Espacial da Porcentagem de Votos em Dilma Rouseff  
(vermelho indica maior concentração de porcentagem de votos)

É interessante notar que a distribuição da porcentagem de votos na candidata Dilma Rouseff é bastante concentrada nas regiões Norte e Nordeste (Fig.04)

Esse resultado é ainda mais convincente quando se leva em consideração a distribuição de votos percentuais do candidato José Serra (Fig. 05).



Fig.05 – Distribuição Espacial da Porcentagem de Votos em José Serra  
(vermelho indica maior concentração de porcentagem de votos)

Fica claro que existe uma diferenciação muito clara da base eleitoral dos candidatos regionalmente. O candidato José Serra obteve maior porcentagem de votos nas regiões Sul, Sudeste e Centro-Oeste (Fig.05)

Por outro lado, a base eleitoral da candidata Marina Silva parece estar mais dispersa espacialmente, com exceção do estado do Rio de Janeiro (Fig.06).

Um dos objetivos deste trabalho é verificar se as características socioeconômicas regionais são capazes de explicar a diferença expressiva de porcentagem de votos obtida pelos candidatos. É possível verificar, mesmo que apenas qualitativamente que existe uma segmentação regional bem definida que diferencia os candidatos José Serra e Dilma Rouseff, principalmente.



Fig.06 – Distribuição Espacial da Porcentagem de Votos em Marina Silva  
(vermelho indica maior concentração de porcentagem de votos)

### 3.3 Estatística Descritiva dos Dados

Variável	Num. Obs	Média	Desvio Padrão	Mínimo	Máximo
População Estimada 2009	5482	34.565,29	203.245,60	837,00	11.000.000,00
Eleitorado em 2006	5482	22.705,57	141.284,20	834,00	7.953.144,00
PIB per Capita	5482	9.239.547,00	10.661,95	1.566,00	239.506,00
Matriculas Ensino Fundamental 2009	5482	5.712.551,00	29.074,29	126,00	1.587.501,00
Matriculas Ensino Médio 2009	5466	1.509.913,00	8.778.919,00	27,00	462.777,00
Docentes Ensino Fundamental 2009	5482	2.728.342,00	1.263,63	11,00	70.544,00
Docentes Ensino Médio 2009	5466	9.097.274,00	4.882.441,00	2,00	25.609,00
Estabelecimentos do SUS	5477	985.503,00	1.925.693,00	1,00	572,00
Nascidos Vivos em 2008	5041	6.095.193,00	3.448.808,00	1,00	186.093,00
Receita Orçamentaria Realizada 2008	4972	50.400.000,00	392.000.000,00	2.229.594,00	23.300.000.000,00
Despesa Orçamentaria Realizada 2008	4972	41.400.000,00	354.000.000,00	1.627.760,00	21.000.000.000,00
Fundo de Participação dos Municípios 2008	4968	9.352.792,00	16.800.000,00	1.267.556,00	530.000.000,00
Empresas Locais Ativas 2008	5482	9.007.962,00	8.242.829,00	4,00	520.533,00
Pessoal Ocupado 2008	5482	8.064.872,00	87.709,41	8,00	5.241.615,00
Latitude	5482	1.649.007,00	8.287.717,00	0,00	336.911,00
Longitude	5482	4.601.174,00	6.501.425,00	0,00	728.923,00
Pessoas residentes - resultados da amostra - municípios vigentes em 2001	5478	30.670,44	187.081,60	795,00	10.400.000,00
Pessoas residentes - 10 anos ou mais de idade - municípios vigentes em 2001	5478	24.728,11	156.187,20	587,00	8.727.317,00
Mulheres residentes - 10 anos ou mais de idade - municípios vigentes em 2001	5478	12.659,61	82.958,10	285,00	4.620.804,00

Variável	Num. Obs	Média	Desvio Padrão	Mínimo	Máximo
Homens residentes - 10 anos ou mais de idade - municípios vigentes em 2001	5478	12.068,50	73.242,46	302,00	4.106.513,00
Pessoas residentes - frequência à creche ou escola - municípios vigentes em 2001	5478	9.645.188,00	55.875,82	146,00	2.992.928,00
Pessoas residentes - 10 anos ou mais de idade - sem instrução e menos de 1 ano de estudo - municípios vigentes em 2001	5478	2.703.697,00	8.001.013,00	5,00	407.936,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - até 1 salário mínimo - municípios vigentes em 2001	5478	4.431.287,00	14.042,59	70,00	510.498,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - mais de 1 a 2 salários mínimos - municípios vigentes em 2001	5478	3.469.101,00	18.949,06	62,00	909.290,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - mais de 2 a 3 salários mínimos - municípios vigentes em 2001	5477	1.774.845,00	14.422,64	5,00	863.224,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - mais de 3 a 5 salários mínimos - municípios vigentes em 2001	5476	1.843.309,00	17.445,93	2,00	1.054.122,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - mais de 5 a 10 salários mínimos - municípios vigentes em 2001	5446	1.670.981,00	18.773,18	2,00	1.116.318,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - mais de 10 a 20 salários mínimos - municípios vigentes em 2001	5252	7.424.888,00	9.736.474,00	1,00	558.277,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - mais de 20 salários mínimos - municípios vigentes em 2001	4779	4.645.049,00	6.960.405,00	1,00	386.938,00
Pessoas residentes - 10 anos ou mais de idade - rendimento nominal mensal - sem rendimento - municípios vigentes em 2001	5478	10.432,25	59.593,19	201,00	3.328.650,00
Pessoas residentes - 10 anos ou mais de idade - com rendimento - municípios vigentes em 2001	5478	14.295,86	96.882,33	258,00	5.398.667,00
Homens residentes - 10 anos ou mais de idade - com rendimento - municípios vigentes em 2001	5478	8.231.563,00	52.023,80	117,00	2.927.649,00
Mulheres residentes - 10 anos ou mais de idade - com rendimento - municípios vigentes em 2001	5478	6.064,29	44.906,78	75,00	2.471.018,00
Rendimento nominal - pessoas residentes - 10 anos ou mais de idade - com rendimento - médio mensal - municípios vigentes em 2001	5478	3.709.212,00	173.651,00	111,96	3.630,07
Rendimento nominal - homens residentes - 10 anos ou mais de idade - com rendimento - médio mensal - municípios vigentes em 2001	5478	4.335.612,00	2.250.792,00	99,69	4.893,84
Rendimento nominal - mulheres residentes - 10 anos ou mais de idade - com rendimento - médio mensal - municípios vigentes em 2001	5478	2.705.452,00	106.713,00	97,43	1.694,37
Domicílios particulares permanentes - resultados da amostra - municípios vigentes em 2001	5478	8.087.683,00	53.530,52	227,00	2.984.416,00
Domicílios particulares permanentes - com rendimento domiciliar - municípios vigentes em 2001	5478	7.715,03	51.540,86	142,00	2.868.383,00
Famílias residentes - domicílios particulares - municípios vigentes em 2001	5478	8.717.319,00	56.770,10	232,00	3.131.389,00
Pessoas residentes - 10 anos ou mais de idade - sem instrução e menos de 1 ano de estudo - municípios vigentes em 2001	5478	2.703.697,00	8.001.013,00	5,00	407.936,00
Pessoas residentes - 10 anos ou mais de idade - 1 a 3 anos de estudo - municípios vigentes em 2001	5478	4.662,10	18.238,69	62,00	961.283,00
Pessoas residentes - 10 anos ou mais de idade - 4 a 7 anos de estudo - municípios vigentes em 2001	5478	8.470,73	49.744,18	163,00	2.828.813,00
Pessoas residentes - 10 anos ou mais de idade - 8 a 10 anos de estudo - municípios vigentes em 2001	5478	3.804.839,00	30.101,28	6,00	1.694.161,00
Pessoas residentes - 10 anos ou mais de idade - 11 a 14 anos de estudo - municípios vigentes em 2001	5478	3.796.772,00	35.781,37	0,00	1.964.427,00
Pessoas residentes - 10 anos ou mais de idade - 15 anos ou mais de estudo - municípios vigentes em 2001	5478	1.073.362,00	15.055,28	0,00	838.239,00
Pessoas residentes - 0 a 3 anos - municípios vigentes em 2001	5478	2.350.385,00	12.734,86	42,00	709.056,00
Pessoas residentes - 4 anos - municípios vigentes em 2001	5478	604.113,00	3.129,58	5,00	171.383,00
Pessoas residentes - 5 e 6 anos - municípios vigentes em 2001	5478	1.214.573,00	6.254.062,00	3,00	340.774,00

Variável	Num. Obs	Média	Desvio Padrão	Mínimo	Máximo
Pessoas residentes - 7 a 9 anos - municípios vigentes em 2001	5478	1.773.258,00	8.937.174,00	25,00	487.017,00
Pessoas residentes - 10 a 14 anos - municípios vigentes em 2001	5478	3.129.108,00	16.041,16	48,00	885.453,00
Pessoas residentes - 15 a 17 anos - municípios vigentes em 2001	5478	1.935.302,00	10.431,71	15,00	574.366,00
Pessoas residentes - 18 a 19 anos - municípios vigentes em 2001	5478	1.303.438,00	7.641.339,00	19,00	418.436,00
Pessoas residentes - 20 a 24 anos - municípios vigentes em 2001	5478	2.914.184,00	18.312,84	40,00	1.022.482,00
Pessoas residentes - 25 a 29 anos - municípios vigentes em 2001	5478	2.499.744,00	16.606,23	40,00	946.536,00
Pessoas residentes - 30 a 39 anos - municípios vigentes em 2001	5478	4.566.649,00	30.083,06	104,00	1.699.367,00
Pessoas residentes - 40 a 49 anos - municípios vigentes em 2001	5478	3.484.108,00	24.188,42	74,00	1.352.029,00
Pessoas residentes - 50 a 59 anos - municípios vigentes em 2001	5478	2.263,46	15.459,00	36,00	855.640,00
Pessoas residentes - 60 a 64 anos - municípios vigentes em 2001	5478	8.343.047,00	5.561.819,00	2,00	302.554,00
Pessoas residentes - 65 a 69 anos - municípios vigentes em 2001	5478	6.479.158,00	4.478.952,00	0,00	238.133,00
Pessoas residentes - 70 a 74 anos - municípios vigentes em 2001	5478	5.024.257,00	3.652,39	0,00	197.599,00
Pessoas residentes - 75 a 79 anos - municípios vigentes em 2001	5478	323.481,00	2.313.808,00	0,00	122.219,00
Pessoas residentes - 80 anos e mais - municípios vigentes em 2001	5478	3.239.783,00	2.177.049,00	0,00	112.502,00
VOTOS EM SERRA	5482	5.949,71	41.290,25	60,00	2.677.318,00
VOTOS EM DILMA	5482	8.594.498,00	46.324,26	237,00	2.532.225,00
VOTOS EM MARINA	5482	3.550.517,00	29.506,86	20,00	1.334.140,00
VOTOS VALIDOS	5482	18.094,72	114.013,50	606,00	6.543.683,00
PORCENTAGEM SERRA	5482	33,7%	14,1%	2,5%	75,7%
PORCENTAGEM DILMA	5482	55,8%	15,8%	15,9%	95,2%
PORCENTAGEM MARINA	5482	10,5%	6,2%	0,9%	43,1%

Tabela 03 – Estatística Descritiva dos Dados

## 4. RESULTADOS

### 4.1 Resultados Esperados

Desde a eleição do ex-presidente Luís Inácio Lula da Silva em 2002, a disputa de poder se polarizou em torno de dois partidos o Partido dos Trabalhadores (PT) e o Partido da Social Democracia Brasileira (PSDB). Em 2006, por exemplo, esses partidos detinham 29% das cadeiras na Câmara dos Deputados. Esses partidos vinham alternando o poder tanto na esfera federal como em estados importantes como São Paulo e Minas Gerais. É natural que nessas condições o PT e o PSDB tenham uma base de eleitores e simpatizantes bem consolidada e provavelmente os eleitores tem características os distinguem quanto a preferência partidária, o que deve ficar evidente nos dados.

O Partido Verde (PV), por outro lado, é um partido de menor expressão. O partido tem como principal bandeira o meio ambiente e o desenvolvimento sustentável. Em 2006, enquanto PSDB e PT detinham 29% das cadeiras na

Camara dos Deputados com cerca de 149 deputados ao todo, o PV tinha apenas 15 deputados ou 2,5% das cadeiras. Parece natural que o partido tenha um perfil de eleitorado mais jovem, urbano e identificado com causas ambientais. É possível, também, que os eleitores não tenham características tão marcadas que os separe claramente do grupo de eleitores do candidato José Serra ou da candidata Dilma Rouseff.

É importante notar aqui que a tentativa apontar quais são os candidatos mais semelhantes em termos de identificação ideologica e eleitorado tem importancia prática primeiro porque a aplicação da metodologia aqui proposta depende de criar estimar modelos que dependem de um classificação *a priori* quanto a semelhança entre as alternativas possíveis, segundo porque a resposta para essa questão tem muitas implicações, por exemplo, para saber se um político perdedor no primeiro turno deve ou não apoiar um dos candidatos que continua na disputa.

Existe um debate bastante intenso acerca de se as políticas governamentais, especialmente as de distribuição de renda, introduzidas pelo ex-presidente Lula tem ou não impacto sobre a votação da candida apoiada por ele. É bastante intuitivo pensar que o eleitor beneficiado pelas políticas governamentais deve ter uma inclinação por votar em um candidato associado a continuidade da politica que o beneficia, de modo que se espera observar um percentual maior de votação em cidades que são mais beneficiadas por programas como o “Luz para Todos” e o “Bolsa Família”.

## 4.2 Estimação dos Modelos de Escolha Discreta do Eleitor

### 4.2.1 Estimação do Modelo Multinomial Logit

Seguindo a metodologia descrita na seção 2.3.1, inicia-se a investigação estimando um modelo Multinomial *Logit*. Esse modelo funciona bem como uma primeira aproximação apesar de impor algumas restrições anteriormente discutidas.

Optou-se por estimar o modelo introduzindo a seguinte legenda: cidades em que a maior votação percentual foi obtida pelo candidato José Serra foram marcadas como 0, cidades em que a maior votação foi obtida pela candidata Dilma Rousseff foram marcadas como 1, por último, cidades em que a candidata Marina Silva obteve maior votação foram marcadas como 2.

No processo de estimação de modelos de escolha discreta é normalmente necessário escolher uma das alternativas como base, essa escolha não tem qualquer impacto sobre os resultados sendo apenas uma questão de escolha de referencial. Adotou-se aqui a escolha da candidata Dilma como referência, a interpretação dos resultados seguirá esta mesma convenção.

Logit Multinomial I		
José Serra		
	Coef	p-Valor
Polação Estimada 2009	8,03E-06	0,196
PIB per Capita 2007	0,0000374	0,000
Num. Est. Do SUS	-0,0016447	0,754
Receita Orçamentária 2008	-7,2E-10	0,879
Despesa Orçamentaria 2008	-3,03E-09	0,622
Empresas Locais	0,0007077	0,000
Pessoal Ocupado 2008	-0,0000312	0,002
Domicilios Permantes 2001	0,0025819	0,000
Domicilios Permantes com Renda 2001	0,0026135	0,000
Constante	-1,342533	0,000
Dilma Rouseff		
	Coef	p-Valor
BASE		
Marina Silva		
	Coef	p-Valor
Polação Estimada 2009	1,38E-06	0,920
PIB per Capita 2007	0,000465	0,000
Num. Est. Do SUS	0,0166674	0,202
Receita Orçamentária 2008	3,33E-09	0,620
Despesa Orçamentaria 2008	1,28E-09	0,885
Empresas Locais	0,000290	0,009
Pessoal Ocupado 2008	-0,000030	0,025
Domicilios Permantes 2001	-0,003087	0,000
Domicilios Permantes com Renda 2001	0,003171	0,000
Constante	-6,365334	0,000

Tabela 4 – Estimação Modelo Multinomial *Logit* variáveis em nível

O modelo Multinomial *Logit* não é um linear de modo que os coeficientes estimados não são iguais aos efeitos marginais, para facilitar a interpretação dos resultados devemos calcular os efeitos marginais.



Logit Multinomial I - Marginal Effects	
José Serra	
	dy/dx
Polação Estimada 2009	1,37E-06
PIB per Capita 2007	6,34E-06
Num. Est. Do SUS	-0,000304
Receita Orçamentária 2008	-1,28E-10
Despesa Orçamentaria 2008	-5,21E-10
Empresas Locais	0,0001208
Pessoal Ocupado 2008	-5,30E-06
Domicilios Permantes 2001	-0,000438
Domicilios Permantes com Renda 2001	0,0004433
Dilma Rousef	
	dy/dx
BASE	
Marina Silva	
	dy/dx
Polação Estimada 2009	-6,52E-09
PIB per Capita 2007	9,09E-08
Num. Est. Do SUS	0,0000526
Receita Orçamentária 2008	1,10E-11
Despesa Orçamentaria 2008	7,92E-12
Empresas Locais	-6,43E-08
Pessoal Ocupado 2008	-4,94E-08
Domicilios Permantes 2001	-5,90E-06
Domicilios Permantes com Renda 2001	-5,90E-06

Tabela 5 – Efeitos Marginais do Modelo Multinomial *Logit*

É interessante notar, por exemplo, que o candidato José Serra tem maior chance de vencer em municípios em que o PIB per Capita é maior o mesmo o corre com a candidata Marina Silva e ainda com maior intensidade. Isso fica claro quando comparamos os efeitos marginais, um aumento de R\$ 1.000, no PIB per capita da cidade causa em média um aumento de 63.4 pontos base na probabilidade de o candidato José Serra ser o vencedor e um aumento de 100 pontos bases na probabilidade da candidata Marina Silva ser a vencedora.

O candidato José Serra é muito identificado com a temática da saúde por ter sido ministro da saúde durante o governo do presidente Fernando Henrique Cardoso. O candidato por diversas vezes explorou essa relação apontando a suposta piora do sistema de saúde desde o final do seu partido em 2002. Tentou-se capturar o possível efeito do mau atendimento usando como proxy o número de estabelecimentos do SUS em cada município, se espera que municípios com menor número de estabelecimentos do SUS tenham mais propensão a encarar a temática da saúde como mais relevante o que seria vantajoso para o candidato José Serra. A estimação do modelo mostra que na verdade a relação não é significativa ou pelo menos não pode ser capturada por essa *proxy*, um modelo adicional será proposta de modo a tentar capturar melhor esse efeito, caso ele exista.

O número de empresas locais tem relação bastante significativa com o a votação dos candidatos. No caso do candidato José Serra, a cada 100 empresas registradas no município aumenta em 120 pontos base a probabilidade de o candidato ser o vencedor no município. No caso da candidata Marina Silva a cada 100 empresas registradas no município diminui a probabilidade de a candidata ser a vencedora no município em 0.05 pontos base, ou seja, praticamente não há impacto.

Embora o impacto seja bastante pequeno, é interessante que a relação entre a probabilidade de a candidata Marina Silva ser a vencedora no município seja negativo, isso colabora para evidenciar a existência de um embate entre a atividade produtiva e a questão ambiental.

Vale a ressalva de que o modelo foi especificado com as variáveis em nível o que pode gerar distorções. Uma maneira de tentar solucionar o problema é estimar o modelo em termos das variáveis *per capita*, notando que o PIB *per Capita* já está condizente com essa especificação.

Logit Multinomial II		
José Serra		
	Coef	p-Valor
PIB per Capita 2007	8,32E-06	0,041
Num. Est. Do SUS per Capita	-319,51	0,013
Receita Orçamentária 2008 per Capita	-5,37E-04	0,005
Despesa Orçamentaria 2008 per Capita	7,64E-04	0,003
Empresas Locais per Capita	43,28	0,000
Pessoal Ocupado 2008 per Capita	0,91	0,047
Domicilios Permantes sem Renda em 2001 per Capita	-50,72	0,000
Constante	-1,708	0,000
Dilma Rouseff		
	Coef	p-Valor
BASE		
Marina Silva		
	Coef	p-Valor
PIB per Capita 2007	7,02E-06	0,638
Num. Est. Do SUS per Capita	-3980,584	0,008
Receita Orçamentária 2008 per Capita	-0,001748	0,076
Despesa Orçamentaria 2008 per Capita	0,0029539	0,017
Empresas Locais per Capita	47,75094	0,000
Pessoal Ocupado 2008 per Capita	2,817277	0,045
Domicilios Pers sem Renda em 2001 per Capita	-53,061970	0,269
Constante	-6,252389	0,000

Tabela 6 – Estimação Modelo Multinomial *Logit* variáveis *per Capita*

A estimação do modelo Multinomial *Logit* com as variáveis *per capita* leva a uma imagem um pouco diferente daquela obtida pelo modelo estimado com as variáveis em nível.

Embora, de modo geral, a direção das correlações não mude muito, fica claro que algumas variáveis que não pareciam ser significantes passaram a ser significantes quanto aplicada a transformação de nível para per capita, isso é bastante intuitivo a medida que, por exemplo, um número maior de estabelecimentos do SUS será sistematicamente observado a medida que a cidade é maior, por outro lado, a população de uma cidade é melhor atendida a medida que aumenta a relação de estabelecimentos do SUS por habitante (menor lotação, melhor atendimento, menor tempo de espera, entre outros).

José Serra	
	dy/dx
PIB per Capita 2007	1,32E-06
Num. Est. Do SUS per Capita	-44,25
Receita Orçamentária 2008 per Capita	-0,0000829
Despesa Orçamentaria 2008 per Capita	1,17E-04
Empresas Locais per Capita	6,8416
Pessoal Ocupado 2008 per Capita	0,1414
Domicilios Permantes sem Renda em 2001 per Capita	-8,0231

Dilma Rousef	
	dy/dx
BASE	

Marina Silva	
	dy/dx
PIB per Capita 2007	8,60E-09
Num. Est. Do SUS per Capita	12,4397
Receita Orçamentária 2008 per Capita	-4,78E-06
Despesa Orçamentaria 2008 per Capita	8,32E-06
Empresas Locais per Capita	0,0813
Pessoal Ocupado 2008 per Capita	0,0076
Domicilios Permantes sem Renda em 2001 per Capita	-0,0858

Tabela 7 – Efeitos Marginais do Modelo Multinomial *Logit*

Analisando os efeitos marginais, percebe-se que as relações encontradas são bastante compatíveis com o esperado. De fato, se consideras as variáveis *per capita*, o candidato José Serra tem maior change de ser o vencedor em cidades em que a relação Número de Estabelecimentos do SUS *per Capita* é baixa, a candidata Marina Silva perde votos a medida que essa relação.

Outras relações que ficam mais evidentes nessa especificação do modelo são as relações associadas a Receita Orçamentária *per capita* e número de Domicílios Permanentes sem Renda *per capita*, nota-se que em municípios com maior receita orçamentária *per capita* os candidatos José Serra e Marina Silva tem votações piores, evidenciando o efeito esperado de que maior receita orçamentária esta associada a satisfação com o *status quo* e

portanto, a um desejo de continuidade. Assim, os candidatos José Serra e Marina Silva perdem votos em cidades em que a Receita Orçamentária per Capita é maior.

Em relação ao número de Domicílios Permanentes sem Renda *per capita*, existe um debate tanto na mídia quanto na academia sobre o impacto das políticas sociais introduzidas no governo do presidente Luis Inácio Lula da Silva tanto sobre variáveis sociais como educação, melhoria nos índices de igualdade, mortalidade infantil e também quanto ao impacto eleitoral dessas políticas. Considerando o número de domicílios permanentes sem renda pode-se notar que os candidatos Marina Silva e em especial José Serra tendem a perder votos em cidades em que há um número maior de domicílios permanentes sem renda. Uma vez que uma grande parte desses domicílios é atendida por algum tipo de programa social como o “Luz para Todos”, “Vale Gás” e “Bolsa Família”, essa é uma evidência de que os programas sociais têm impacto eleitoral favorecendo a candidata Dilma Roussef.

Esses resultados, em conjunto, corroboram com a hipótese de que o eleitor é racional e consciente a medida que as relações estimadas mostram que o eleitor escolhe em média o candidato mais alinhado com seus interesses. Em cidades com maior número de empresas *per capita* o candidato José Serra tende a ser melhor votado o que é razoável uma vez que o candidato é teoricamente mais alinhado com políticas pró-mercado, em cidades com maior número de domicílios sem renda *per capita* a candidata Dilma Roussef é leva vantagem uma vez que é mais alinhada com a manutenção de política sociais introduzidas pelo seu partido. Em cidades em que o número de estabelecimentos do SUS *per capita* é menor, o candidato José Serra tende a ser mais bem votado o que é compatível com o esperado uma vez que o candidato é bastante identificado com a temática da saúde.

#### 4.2.2 Estimação do Modelo Nested Logit

Foi estimado um modelo *Nested Logit*, assumindo que a candidata Marina Silva pertence ao mesmo aninhamento que o candidato Jose Serra. Essa opção se deve ao fato de que o candidato Jose Serra ganhou votos no segundo turno da eleição o que evidencia que foi capaz de atrair eleitores da candidata Marina Silva.

Nested Logit		
Campo Político	Coef	p-Valor
PIB per Capita 2007	1,66E-04	0,010
Num. Est. Do SUS per Capita	15,51	0,740
Receita Orçamentária 2008 per Capita	-5,34E-02	0,650
Despesa Orçamentaria 2008 per Capita	-8,30E+01	0,983
Empresas Locais per Capita	-9,00	0,341
Pessoal Ocupado 2008 per Capita	0,91	0,047
Domicilios Permantes sem Renda em 2001 per Capita	-50,72	0,124

José Serra		
	Coef	p-Valor
PIB per Capita 2007	0,8315	0,041
Num. Est. Do SUS per Capita	0,3885	0,000
Receita Orçamentária 2008 per Capita	-0,8680	0,820
Despesa Orçamentaria 2008 per Capita	-1,0475	0,047
Empresas Locais per Capita	0,5205	0,900
Pessoal Ocupado 2008 per Capita	0,3626	0,047
Domicilios Permantes sem Renda em 2001 per Capita	0,5212	0,000
Constante	-0,5638	0,030

Dilma Rouseff		
	Coef	p-Valor
BASE		

Marina Silva		
--------------	--	--

	Coef	p-Valor
PIB per Capita 2007	9,96E-01	0,066
Num. Est. Do SUS per Capita	- 0,03417	0,008
Receita Orçamentária 2008 per Capita	0,16576	0,076
Despesa Orçamentaria 2008 per Capita	- 0,27958	0,000
Empresas Locais per Capita	50,06638	0,000
Pessoal Ocupado 2008 per Capita	2,72742	0,045
Domicilios Permantes sem Renda em 2001 per Capita	0,33548	0,450
Constante	- 5,36500	0,000

Tabela 8 – Efeitos Marginais do Modelo Multinomial *Logit*

#### 4.2.3 Estimação do Modelo Nested Factor Logit

A implementação do modelo Nested Logit Factor model não convergiu. Muitas estratégias de estimação foram tentadas, incluindo a estimação dos fatores por PCA, *Principal Component Analysis*. Contudo, nenhuma das tentativas produziu resultados satisfatórios. De modo que a análise deste modelo não pode ser feita adequadamente.

Essa alternativa de especificação fica em aberto como uma possível extensão para o trabalho futuramente.

#### 4.3 Comparação dos Modelos de Escolha Discreta do Eleitor

#### 4.4 Modelo de Escolha do Político

O objetivo dessa seção é verificar a hipótese de que os políticos ao aprovarem ou reprovarem uma matéria votam estrategicamente alinhados com que eles percebem como sendo o a posição que lhes fará ganhar mais votos. Em outras palavras, o político maximiza sua utilidade à medida que consegue se manter no poder, para se manter em um sistema democrático votos são necessários sendo assim, o agente político maximiza sua utilidade a medida que conquista votos.

Se existem evidencias de que o eleitor vota maximizando sua utilidade, é de se esperar que o agente político racional com o objetivo de se manter no

poder tente direcionar sua atuação de modo a garantir que suas ações sejam compatíveis com a maximização da utilidade da maioria do eleitorado. Um caso particular dessa dinâmica é dado pelo modelo do eleitor mediano de Downs (1957). O argumento é bastante convincente, se os eleitores são agente heterogêneos e, portanto, não existe um mix de políticas que maximize o bem-estar de todos os eleitores, então, a estratégia do agente político racional é ter ações compatíveis com o interesse do eleitor mediano de modo que sua atuação agrada exatamente metade do eleitorado e não desagrade mais do que a metade do eleitorado.

Menezes *et al.* (2010) mostram que existem evidências favoráveis ao modelo do eleitor mediano para o caso brasileiro. A metodologia aplicada por Menezes *et al.* (2010) assume a perspectiva do eleitor, a perspectiva adotada aqui é a do agente político.

A especificação aqui proposta está baseada na hipótese de que o eleitorado é grande e diverso o suficiente para que o agente político opere sobre informação imperfeita. Nesse caso, o agente político racional tenta interpretar as sinalizações dadas pelo eleitor e atualiza suas posições à medida que mais informações se tornam disponíveis.

Uma tentativa de verificar evidências desse processo é verificar a votação do Código Florestal. Essa matéria é bastante identificada com a candidata Marina Silva, entre os deputados do seu partido o Partido Verde (PV), 100% dos deputados votaram contra o Código Florestal. Em estados em que a candidata Marina Silva teve maior votação é de se esperar que o eleitor esteja relativamente mais preocupado com questões ambientais, sendo assim, os deputados desses estados devem estar relativamente mais inclinados a votar contra a aprovação do Código Florestal.

Uma análise preliminar dos dados aponta na direção de que a relação entre probabilidade de votação contra o Código Florestal e a votação da candidata Marina Silva existe. O estado do Rio de Janeiro, onde a candidata Marina Silva teve maior votação, tem 46 deputados, dos quais 45 estavam presentes no dia da votação, desses 11 deputados votaram contra a aprovação



do Código Florestal, ou seja, quase 25% dos deputados do Rio de Janeiro votaram contra a aprovação do Código Florestal.

O estado do Maranhão, onde a candidata Marina Silva teve menor votação, tem 18 deputados, desses 14 deputados estavam presentes no dia da votação dos quais apenas 2 deputados votaram contra o Código Florestal, ou 14% dos deputados.

Esses exemplos mostram uma relação qualitativa entre a votação da candidata Marina Silva e o comportamento dos deputados na votação do Código Florestal. A maneira de investigar se essa relação é estatisticamente significativa é estimar um modelo *Logit*, conforme descrito na seção 2.6.3.

#### 4.4.1 Modelo *Logit* Estimado para Votação do Código Florestal

Seguindo a metodologia proposta na seção 2.6 e brevemente discutida na seção anterior, estimou-se o seguinte modelo

$$Prob(y = 1|X) = G(\beta_0 + \beta_1VotMarina + \sum_i^N \beta_{i+1}Est_i + \sum_j^K \beta_{j+N+1}Part_j)$$

(16)

onde *VotMarina* é a votação da candidata Marina Silva no estado em que o deputado foi eleito, *Est* é uma *dummy* para Estado pelo qual o deputado foi eleito, e *Part* é o partido ao qual o deputado é filiado.

Logit - Votação Código Florestal		
	Coef	p-Valor
Log(Votação Marina Silva)	-2.01740	0.002
Maranhão	-2.59096	0.014
Rio Grande do Sul	1.45377	0.038
Sergipe	-1.69626	0.100
Piauí	-1.93339	0.063
Rio Grande do Norte	-1.45209	0.000
PT	-2.73819	0.000
PPS	-1.46663	0.078
PDT	-1.38812	0.014
Constante	-0.35496	0.730

Tabela 8 – Efeitos Marginais do Modelo Multinomial *Logit*

A tabela 8 mostra os coeficientes estimados para o modelo final, inicialmente utilizou-se *dummies* para todos os estados assumindo São Paulo como sendo a constante e para todos os partidos com representantes da Câmara Federal assumindo o PV como base. As variáveis não significantes estatisticamente foram sendo iterativamente eliminadas do modelo até que se chegasse à especificação final.

Uma vez que o modelo *logit* é não linear os coeficientes não refletem o efeito marginal. Pode-se fazer uma interpretação mais direta em termos dos efeitos marginais dispostos na tabela 9.

Logit - Votação Código Florestal	
	dy/dx
Log(Votação Marina Silva)	-0.15787
Maranhão	-0.45653
Rio Grande do Sul	-0.19002
Sergipe	-0.24817
Piauí	-0.30134
Rio Grande do Norte	-0.19696
PT	-0.42095
PPS	-0.19830
PDT	-0.17909

Tabela 9 – Efeitos Marginais do Modelo *Logit* para a Votação do Código Florestal

Os efeitos marginais mostram que a probabilidade de um deputado eleito por um estado em que a candidata Marina Silva teve relativamente mais votos votar a favor do Código Florestal é menor. A interpretação dos efeitos marginais para as *dummies* é menos interessante à medida que as *dummies* não variam em um espaço contínuo.

#### 4.4.2 Teste de Hipóteses

Na seção 2.6, propõe-se um teste para verificar a existência de correlação parcial significativa entre a probabilidade de voto a favor do Código

Florestal e a votação da candidata Marina Silva no estado pelo qual o legislador foi eleito.

Nesse caso, o teste de hipótese é bem direto uma vez que se deve testar apenas um coeficiente estimado, o coeficiente ligado a variável *Log(Votação Marina Silva)*.

$$H_0: \beta_{Marina} = 0$$

$$H_A: \beta_{Marina} \neq 0$$

Portanto, quer-se testar se o efeito de uma votação maior da candidata Marina Silva no estado é zero ou se tem impacto na decisão do legislador.

Como discutido na seção 2.6.4, a estatística do teste segue uma distribuição t-Student. Levando em conta o número de graus de liberdade e adotando nível de confiança de 5%, chega-se ao valor crítico para a estatística de 1,75 aproximadamente.

$$t_{obs} = \frac{\beta}{\sigma_{\beta}} = \frac{-2.0174}{0,64006} \cong 3.15$$

Uma vez que o valor observado da estatística é 3.15, existe base estatística para rejeitar a hipótese nula de que a votação da Marina no Estado pelo qual o legislador não tem impacto sobre a decisão de voto a respeito do Código Florestal em favor da hipótese alternativa de que existe algum impacto.

Esse resultado colabora com a hipótese de que os políticos levam em consideração a sinalizações dadas pelos eleitores durante o processo eleitoral. Além disso, o fato de o sinal do coeficiente estimado ser negativo mostra que os legisladores são racionais, tentam compatibilizar suas atuações com a posição do eleitorado nos estados pelos quais são eleitos.

## 5. CONCLUSÃO

Os modelos estimados assim como a análise qualitativa dos dados colaboram com a tese de que tanto o eleitor quanto os políticos são em média racionais e tomam suas decisões de modo a maximizarem seu bem-estar.

Ficar evidente, que existe grande correlação entre a decisão do eleitor e o mix de políticas a ele ofertada através do processo eleitoral. Os modelos estimados demonstram que também o político também observa e se adéqua ao desejo do eleitor buscando assim maximizar o numero de votos que poderá receber e, portanto, atingindo seu objetivo de se perpetuar no poder.

O modelo Multinomial Logit não difere tão significativamente dos resultados obtidos para o que evidencia que os candidatos não eram similares o suficiente para que o aninhamento fosse necessário para explicação da decisão de voto.

Nesse sentido, parece que a candidata Marina Silva tomou a decisão correta e racional ao não se alinha a nenhum dos candidatos no segundo turno. Além disso, parece que os candidatos, de modo geral, perceberam seus pontos fortes e exploraram ao longo da campanha. A candidata Dilma utilizando sua propaganda para falar sobre os programas sociais do governo e os progressos recentes no campo econômico. O candidato Jose Serra, por outro lado, explorou o tema da saúde com o qual é bastante identificado e os dados mostram que de fato isso tem alguma importância na explicação do voto.

Muitas vezes se coloca em questão se a decisão do eleitor de votar em um candidato que o beneficia, por exemplo, com políticas de distribuição de renda, parece claro por essa análise que o eleitor que faz esse cálculo de votar em no político que defende políticas mais compatíveis com seu objetivo de maximização de utilidade está fazendo uma escolha racional e autointeressada. Esse talvez seja o tipo de voto mais consciente, uma vez que o eleito está de fato calculando o quanto tem a ganhar ou perder em vários dos possíveis *outcomes*.

Encontra-se evidências de que o modelo do eleitor mediano é válido e não apenas isso, que os políticos e os eleitores tendem a tomar decisões em um sistema dinâmico em que o político está sempre atualizando sua atuação para corresponder à expectativa do eleitor.

Por fim, essas evidências colaboram para a ideia de que o problema de compatibilização de incentivos se resolve nesse caso a medida que os políticos são avaliados pelos eleitores a cada eleição, e por isso os políticos tendem a prescrever políticas que maximizam o bem-estar do eleitorado, não por altruísmo ou espírito público, o que alguns até podem considerar, mas principalmente visando maximizar sua própria utilidade.

## 6. REFERÊNCIAS

- [1] ABRAMSON, P. R.; ALDRICH, J. H.; PAOLINO, P.; ROHDE, D. W.; “Sophisticated” Voting in the 1988 Presidential Primaries. **The American Political Science Review**, v. 86, n. 1, Mar. 1992.
- [2] BLACK, J. H.; SPENCER, B. J.; **Theory of Committees and Elections**. Cambridge: Cambridge University Press, 1958.
- [3] DOWNS, A.; **An Economic Theory of Democracy** NewYork: Harper and Row, 1957.
- [4] HARSANYI, J. C.; Games with incomplete information played by bayesian players. **Management Science**, v. 14, n. 3, Nov. 1967.
- [4] KAMAKURA, W. A.; MAZZON, J. A.; BRUYN, A. D.; Modeling Voter Choice to Predict the Final Outcome of Two-Stage Elections. **Int. Journal of Forecasting**, v. 22, n. 2, 2006.
- [5] MENEZES, R. T.; SAIANI, C. S.; ZOGHBI, A. P.; Demanda Mediana por Serviços Públicos e Desempenho Eleitoral: Evidencias do Modelo do Eleitor Mediano para os Municípios Brasileiros. **Est. Econ.**, v. 41, n. 1, Jan-Mar. 2011.
- [5] MESQUITA, B. B.; Forecasting Policy Decisions: An Expected Utility Approach to Post-Khomeini Iran. **PS**, v. 17, n. 2, Spring. 1984.
- [6] MESQUITA, B. B.; **The Predictioneer's Game: Using the Logic of Brazen Self-Interest to See and Shape the Future**. NewYork: Random House Trade Paperbacks, Oct. 2010.
- 5] MYERSON, R.; WEBER, R. A Theory of Voting equilibria. **Amer. Polit. Sci. Rev.**, v. 87, pag. 151-166, 1993.
- [11] NEWMAN, D.; BRIDGES, B.; North Korean Nuclear Weapons Policy: An Expected Utility Study. **Pacific Focus**, v. 9, n. 2, 1994.

[12] SELTEN, R; Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragerträgeit. **Zeitschrift für die gesamte Staatswissenschaft**, v. 121, 1965.

[13] TSAI, C.; HSU, Y.; Expected-Utility Voting in Taiwan's Legislative Elections: 1995, 1998, and 2001. **International Political Science Review**, v. 26, n. 4, Out. 2005.

[6] WOOLDRIDGE, J. M.; **Econometric Analysis of Cross Section and Panel Data**. London: The MIT Press, 2001.