

**- Título de la ponencia:**

Comportamento eleitoral, dados individuais (survey) e dados agregados: proposta metodológica e formatação de modelo hierárquico para o estudo do voto econômico.<sup>1</sup>

**- Autora**

Iara Lima Vianna

**- Afiliación institucional**

Universidade Federal de Minas Gerais (UFMG)

**- Correo electrónico**

[iaravianna18@hotmail.com](mailto:iaravianna18@hotmail.com)

**- Resumen**

Pesquisadores das correntes clássicas do Comportamento Eleitoral (Psicológica, Psicossociológica e Teoria da Racionalidade) se dedicam a investigar quais fatores influenciam na decisão do voto do eleitor. A ideia básica do voto econômico, enquadrado na Teoria da Escola Racional, é a lógica de punição-recompensa: quando a economia vai bem, os eleitores premiam o incumbente com o seu voto; já quando a econômica está mal, os eleitores punem o mandatário direcionando seu voto aos candidatos desafiantes.

Em estudos com dados agregados, em geral, considera-se crescimento do PIB per capita, variação da inflação e variação do desemprego (variáveis independentes) e sua influência em resultados eleitorais (variável dependente). Uma dificuldade em constante discussão é o problema estatístico denominado como falácia ecológica: a inferência de comportamento individual (voto) através de resultados agregados.

Encorajado por este dilema, o objetivo deste trabalho foi desenvolver análise para discutir e elucidar os benefícios do uso de modelos multiníveis e propor um modelo para análise do chamado voto econômico. No modelo hierárquico proposto, optou-se por inserir duas variáveis agregadas (crescimento real do PIB e nível de desocupação por estado) e uma variável individuais (avaliação e percepção a respeito da inflação) na explicação da decisão eleitoral (voto) dos cidadãos. O modelo final ajustado inseriu variáveis associadas a fatores econômicos de forma a operacionalizar a discussão; contudo, a intenção é a de que seja um modelo genérico, a ser adaptado e adequado ao contexto, esfera política e momento econômico que o pesquisador interessado desejar.

**- Tres palabras clave**

Comportamento eleitoral; voto econômico; modelos hierárquicos.

---

<sup>1</sup> Trabajo presentado en el Quinto Congreso Uruguayo de Ciencia Política, “¿Qué ciencia política para qué democracia?”, Asociación Uruguaya de Ciencia Política, 7-10 de octubre de 2014.

## 1 - Introdução

O estudo do comportamento político-eleitoral, enquanto área de pesquisa científica-acadêmica, é perpassado e fundamentado por três grandes teorias clássicas: a Sociológica, a Psicossociológica (ou Psicológica) e a Teoria da Racionalidade. As duas primeiras trabalham com os chamados fatores de longo prazo para explicação do voto; sendo exemplos de algumas de suas principais variáveis de estudo a classe social ou a religiosidade (para o enfoque sociológico) e a identificação partidária (para o psicológico). A terceira teoria, por outro lado, está centrada em fatores de curto prazo, como a avaliação do desempenho do governo, sobretudo através de suas políticas econômicas; em alguns casos, ao se tratar do voto racional, fala-se também da imagem dos candidatos e os temas debatidos durante as campanhas eleitorais (Ratto e Montero, 2013).

É inegável a importância das teorias do voto econômico para a compreensão da lógica operada pelo cidadão para decidir seu voto. Inúmeras pesquisas explanam teórica e empiricamente a relevância e a significância dos achados que demonstram a relação entre dados econômicos (avaliação e percepção individual e/ou mensuração de taxas agregadas) e a escolha do eleitor entre premiar o incumbente (votando pela sua reeleição ou no candidato apoiado por sua coligação, considerado como seu sucessor) ou puni-lo, votando em algum dos candidatos desafiantes.

Essas teorias têm buscado, ao longo do tempo, investigar e testar empiricamente quais fatores de fato determinam (e com que intensidade influenciam) o comportamento dos eleitores. Inúmeros trabalhos foram desenvolvidos com a finalidade de defender maior eficiência explicativa de uma ou de outra corrente: alguns, essencialmente, confrontando e exaltando os pontos fortes e as fragilidades de uma ou mais correntes, outros direcionando seus esforços à realização de análises e testes empíricos de dados e resultados eleitorais (circunscritos a um tempo e espaço específicos), e muitos fazendo ambas as atividades.

O presente trabalho não tem nenhum dos objetivos citados acima, a finalidade deste paper se foca fundamentalmente em trazer à tona uma discussão metodológica importante para os estudiosos do voto econômico (e para pesquisadores de várias outras áreas do conhecimento): a análise e inferência estatística multinível.

A ideia básica do voto econômico, enquadrado na Teoria da Escola Racional, é a lógica de punição-recompensa: quando a economia vai bem, os eleitores premiam o

incumbente com o seu voto; já quando a econômica está mal, os eleitores punem o mandatário direcionando seu voto aos candidatos desafiantes. Em estudos sobre voto econômico com dados individuais (micro dados), coletados em pesquisa de opinião (survey), pesquisadores analisam o impacto da percepção e/ou avaliação da situação econômica (variáveis independentes) na intenção de voto do eleitor: mandatário/incumbente ou desafiante (variável dependente). Já em estudos com uso de dados agregados (macro dados), em geral, considera-se como condicionantes de resultados eleitorais (variável dependente) três principais fatores: o crescimento do PIB per capita, variação da inflação e variação do desemprego (variáveis independentes).

Estudos sobre voto econômico baseados em micro dados (individuais) mensuram percepções e avaliações que os indivíduos fazem de aspectos econômicos (preocupação com a inflação e com o desemprego, avaliação da economia, etc.). Já estudos com dados macro (agregados) consideram em suas análises taxas e dados econômicos oficiais das unidades de análise (estados, países, regiões) estudo. Em pesquisas com macro dados, um tópico em constante discussão é o problema estatístico denominado como falácia ecológica: a inferência de comportamento individual (voto) através de resultados agregados (Soares e Terron, 2008).

Em estudos sobre voto econômico que compreendam variações entre seus elementos em mais de um nível (análise multinível), considera-se ser importante a inclusão e análise de dados agregados, de forma a captar e considerar no modelo analítico dados individuais dos eleitores bem como as variações dos fatores econômicos entre as estruturas nas quais estão inseridos tais eleitores<sup>2</sup>. Ou seja, um modelo de avaliação dos efeitos do voto econômico no comportamento eleitoral não deve se debruçar apenas sobre uma ou outra perspectiva (macro vs. micro), ele precisa considerar em seu cálculo a existência de variação individual e de variação estrutural.

Os dados a respeito do processo eleitoral em diversos países, incluindo o Brasil, sugerem que, sem alguma forma de articulação entre variáveis de tipo macro e variáveis que dizem respeito às características individuais dos eleitores, não é possível sustentar uma teoria suficientemente ampla e complexa do fenômeno eleitoral. Neste sentido, somente uma articulação entre sujeito da ação e estrutura seria capaz de dar conta do comportamento dos eleitores. O problema é como articular de forma fecunda essas duas dimensões. (Castro, 1992:18)

---

<sup>2</sup> Um exemplo aplicado da necessidade deste tipo de análise multinível seriam estudos nacionais nos quais há variação de mensuração dos fatores de interesse entre os estados confederados, afinal taxa de desemprego, PIB per capita e taxa de inflação podem apresentar variações diferentes (e bem distantes, em alguns casos) entre um estado confederado e outro.

Encorajado por este dilema, o objetivo do trabalho aqui apresentado foi desenvolver análise que com vistas a discutir e elucidar os benefícios do uso de modelos hierárquicos, bem como propor, ao final da exposição, um possível modelo de análise multinível a ser usado por estudiosos do voto econômico. Neste modelo multinível, buscou-se contemplar através de variáveis agregadas (crescimento do PIB real e percentual de desocupação/taxa de desemprego nos estados brasileiros) e de variável individual (avaliação e percepção a respeito da inflação) fatores considerados como aqueles que mais influenciam na decisão do voto dos cidadãos, segundo a teoria do voto econômico.

Devido à sua construção multinível, essa abordagem hierárquica é caracterizada por não assumir o pressuposto de independência entre as observações das unidades pertencentes a um agregado, admitindo violação da hipótese de erros independentes e identicamente distribuídos (i.i.d), produto da dependência entre os indivíduos de um mesmo agregado.

Além disso, a possibilidade oferecida pelos modelos hierárquicos de que a variabilidade da variável resposta seja explicada através de variáveis preditoras incluídas em diferentes níveis, permite evitar o problema das falácias ecológica e/ou atomística, que consiste em analisar dados em um nível e formular conclusões em outro nível (Raudenbush e Bryk, 2002). Conforme será melhor explicado a seguir, a falácia ecológica é o erro que ocorre quando o pesquisador interpreta dados agregados em nível individual, confundindo efeito individual com efeito agregado; por outro lado, a falácia atomística é o problema de se extrair conclusões em nível hierárquico agregado baseadas em análise efetuada no nível individual (Hox, 2002).

...

O trabalho está organizado da seguinte forma: inicialmente, faz-se breve apresentação sobre as três grandes teorias clássicas no estudo do comportamento político-eleitoral (correntes Sociológica, Psicossociológica e a Teoria da Racionalidade), dando esperado destaque à terceira corrente, por esta ser a teoria evocada e discutida no artigo. Em seguida, o modelo hierárquico (ou multinível) é exposto, discutido e defendido enquanto possibilidade sofisticada de análise de dados em mais de um nível. Por fim, propõe-se um modelo multinível para análise do voto econômico, a ser avaliado, adaptado, melhorado e empiricamente testado por pesquisadores e curiosos da área. Vale ressaltar que não é feita explanação sobre como operar e executar modelos hierárquicos em softwares de análise estatística.

## 2 - Comportamento eleitoral e as três correntes clássicas

Conforme destaca Ruiz (2013:25), “após décadas de sucessivos estudos sobre comportamento eleitoral, parece óbvia a afirmação de que são muitos os fatores que condicionam o voto”. Serão apresentadas, a seguir, as três grandes teorias clássicas no estudo do comportamento político-eleitoral (correntes Sociológica, Psicossociológica e a Teoria da Racionalidade). Destacando-se que as duas primeiras são reconhecidas por se centrarem em fatores de longo prazo (classe social, religião, ideologia, etc.); enquanto a última trabalha com fatores de curto prazo, itens que podem ser apontados como ‘conjunturais’ ou ‘momentâneos’ (fruto do momento vigente), tais como avaliação do desempenho do governo (ou do governante incumbente) e avaliação econômica.

Há de se ressaltar serão pontuados brevemente os principais conceitos e definições a respeito das duas primeiras correntes (Sociológica e Psicossociológica). Maior aprofundamento teórico será dado à corrente que é foco da análise deste trabalho: a Teoria da Racionalidade (ou o chamado voto econômico).

### *As teorias que avaliam o ‘longo prazo’ – a corrente Sociológica e a Psicossociológica*

A Corrente Sociológica coloca como centrais na explicação do comportamento eleitoral os contextos sociais-políticos e a interação social nos quais o indivíduo está inserido, levando em conta em suas análises fatores como: ocupação, região de residência (urbana/rural), situação econômica e social (classe trabalhadora, burguesia, etc), religião, sexo, idade, escolaridade, entre outros. Nessa perspectiva, segundo coloca Castro (1992), seria necessário analisar a natureza das relações, a inserção dos indivíduos em grupos sociais, bem como o interesse de classe desses grupos.

Conforme destacado por Antunes (2010), o modelo sociológico de comportamento eleitoral foi, inicialmente, apresentado em três trabalhos essenciais: *The People’s Choice* (Lazarsfeld, Berelson, & Gaudet, 1944); *Voting* (Berelson, Lazarsfeld, & McPhee, 1954) e *Personal Influence* (Katz & Lazarsfeld, 1955); desenvolvidos no âmbito da Universidade de Columbia (EUA).

A pesquisa de painel conduzida por Lazarsfeld et al. (1944) entrevistou sete vezes, ao longo de sete meses, cerca de 600 eleitores; que foram então segmentados em

três grupos: aqueles que decidiram seu voto antes do começo da campanha, aqueles cuja decisão foi tomada durante as convenções partidárias e os que tomaram sua decisão já em um estágio avançado das campanhas

Os resultados sugeriram que o efeito da exposição às mídias na decisão eleitoral é mínimo; decisivo e significativo mesmo na explicação do voto é o grupo social do qual os eleitores fazem parte. Ou seja, para Lazarsfeld et al. (1944), segundo escrevem Ratto e Montero (2013), grande parte dos eleitores apresentava uma predisposição política ou partidária ancorada em sua identidade social, uma pré-disposição que seria imune aos investimentos e mensagens das campanhas. Na pesquisa realizada por Lazarsfeld et al. (1944), a associação entre o comportamento eleitoral e os grupos sociais aos quais pertenciam os indivíduos era tão forte que seria possível explicar as escolhas eleitorais utilizando apenas 'status socioeconômico', 'religião' e 'área de residência', fatores que definem o Índice de Predisposição Política utilizado na pesquisa. (Antunes, 2010).

Diante de tal achado, no final de seu livro, Lazarsfeld et al. (1944:69, apud Ratto, Montero, 2013) ressaltam que embora a significância dos resultados tenha sido reveladora e recebido destaque no estudo que desenvolveram, novas e aprofundadas investigações seriam necessárias para fortalecer a área (tradução livre).

Muitas pesquisas e estudos foram realizados com vistas a investigar o papel de fatores e características sociais na decisão eleitoral dos cidadãos. Estudos sobre as eleições federais canadenses encontraram que as características socioeconômicas são um importante preditor de escolha dos eleitores (McGrane, 2007). Segundo André Blais (2005 *apud* McGrane, 2007), enquanto algumas características socioeconômicas, como classe, tornaram-se menos importante ao longo do tempo, seria impossível compreender as recentes eleições federais canadenses sem olhar para a região, religião e etnia como importantes determinantes do comportamento eleitoral. Outros estudos também realizados no Canadá descobriram que Católicos e cidadãos com origens étnicas não europeias votavam no Partido Liberal, enquanto protestantes e eleitores com origens étnicas europeias estariam mais propensos a votar nos Partidos Conservador, Aliança Canadense ou Novo Partido Democrático.

Ratto e Montero (2013) também desenvolveram pesquisa cujo foco de uma das vertentes de análise do comportamento eleitoral dos argentinos se dedica a examinar a influência dos fatores sociológicos 'classe social' e 'religião'. Os achados apontam que os fatores sociais de longo prazo analisados (classe social e religião) apresentaram baixa significância explicativa do voto presidencial em 2007 na Argentina.

Na pesquisa realizada no Canadá, religião é um importante preditor do comportamento eleitoral; já na pesquisa argentina, religião não apresenta grande significância. É interessante observar que dependendo da localidade, da conjuntura e contexto político um ou outro atributo socioeconômico pode ganhar ou perder relevância explicativa do voto. Nota-se então que ocorre dentre os fatores de uma mesma corrente algo semelhante com o que também costuma acontecer entre as correntes teóricas: em determinado momento político-conjuntural e local um ou outro atributo (ou teoria) apresenta maior significância na explicação do comportamento eleitoral.

...

A Corrente Psicossociológica se centra na crença de que o comportamento eleitoral dos indivíduos é definido pelos valores e motivações psicológicas que estes carregam consigo. Entender o comportamento dos eleitores nessa perspectiva psicossociológica é entender a mente do indivíduo, e nesse sentido, elementos como socialização, educação, valores adquiridos no âmbito familiar, entre outros, influenciarão a maneira do indivíduo pensar e, conseqüentemente, a sua decisão de voto

Segundo Veiga (2001), a corrente Psicossociológica assume que, conforme defende a teoria Sociológica, as características socioeconômicas influenciam sim o comportamento eleitoral, mas se propõe a sanar as limitações impostas por essa primeira teoria (Sociológica): chamando a atenção para a fragilidade dessa corrente em explicar porque em determinados momentos os indivíduos acabam por contrariar o que seria esperado deles em função de seu meio e destacando que a análise puramente externa tende também a negligenciar o raciocínio desenvolvido pelo indivíduo em uma tomada de decisão.

A Universidade de Michigan (EUA) revelou os precursores da teoria psicológica (ou psicossociológica). Campbell e sua equipe de pesquisadores são os autores de *The American Voter* (1964), trabalho cujo principal foco de análise é o indivíduo e suas motivações psicológicas para definir seu voto; segundo estes estudiosos, para entender política seria necessário entender a mente dos indivíduos (Santos, 2001). Curiosamente, o interesse de Angus Campbell e Robert Kahn, membros na época do Survey Research Center da Universidade de Michigan, era conduzir uma pesquisa a respeito de atitudes a respeito de política externa e ao final das entrevistas, para determinar o grau de interesse e as orientações políticas dos respondentes perguntavam se eles planejavam votar na próxima eleição presidencial e para o qual partido

(Campbell e Kahn, 1952 *apud* Bartels, 2008). Após a eleição, estimulados pelo fracasso muito divulgado do Instituto Gallup para predizer os resultados – Campbell e Kahn decidiram entrevistar novamente os mesmos respondentes da pesquisa anterior, para averiguar a confirmação das intenções de voto, para registrar as características pessoais, comportamentais e demográficas dos eleitores republicanos e democratas, e para avaliar a influência de fatores vários psicológicos, sociológicos e políticos sobre a decisão do voto (Campbell e Kahn, 1952 *apud* Bartels, 2008). Conforme destacada Bartels (2008), muito provavelmente Campbell e Kahn não poderiam imaginar que este seria o início de um dos projetos de pesquisa de mais longa duração na história acadêmica das ciências sociais.

Segundo Figueiredo (1988, p. 20 *apud* SANTOS, 2008), Campbell elaborou uma ideia na qual o voto iria para o partido ou candidato que melhor estivesse alinhado com valores do eleitor, e a existência de um pensamento linear por parte dos eleitores permitiria a possibilidade de se prever a direção do voto. Contudo, em estudos realizados na mesma época, Converse (1964), estudioso também considerado como expoente da psicossociológica, aponta que seriam poucos aqueles que agem através de um pensamento linear, com alto grau de adesão política e forte identificação partidária. A maior parcela do eleitorado seria composta pelos eleitores com baixo envolvimento político e com pouco interesse por essas questões. Nesse sentido, Figueiredo (1988, p. 21 *apud* SANTOS, 2008) aponta que a questão que se coloca é a de que “se o comportamento dos indivíduos é função dos sistemas atitudinais frente ao mundo e se esses sistemas não são coerentemente estruturados, então não é possível prever nada”.

Converse aponta duas possibilidades para se resolver este dilema: 1) verificar o grau de centralidade que certas questões ligadas ao mundo da política têm no dia a dia do cidadão, e 2) averiguar o grau de motivação dos eleitores em relação à política; ambos os pontos são colocados por considerar-se que as atitudes dos cidadãos possuem um grau de elasticidade para a mobilização política. Assim, os fatores e motivações conjunturais, mais do que os fatores estruturais, impactariam no comportamento dos eleitores em relação às opções políticas. (FIGUEIREDO *apud* SANTOS, 2008).

Em suma, além de evidenciar as limitações da corrente sociológica – na qual a decisão do voto é baseada apenas em análises dos grupos sociais aos quais pertencem os eleitores – a corrente psicossociológica propõe uma abordagem baseada na atitude política dos cidadãos e em sua “estruturação ideológica”; estruturação ideológica sendo



entendida como a relação que o eleitor faz entre as opiniões sobre as diversas questões em discussão na arena política e a decisão do voto. (Veiga, 2001).

### *A corrente do 'curto prazo' – a Teoria da Escolha Racional*

Baseada no individualismo metodológico, conforme destaca Camargos (1999), a perspectiva da Corrente/Teoria da Escolha Racional aponta que “os indivíduos são racionais e agem intencionalmente, procurando, antes de tudo, maximizar ou otimizar seus ganhos” (p. 01); sendo expoente de uma lógica na qual a tomada de decisão política decorre do fato de que eleitor é "movido por razões egoístas" e "vota no partido que ele acredita que lhe proporcionará mais benefícios do que qualquer outro" (Figueiredo, 1991, p.37)

É no âmbito desta terceira Corrente que se encontram os estudos que buscam associar a percepção do eleitor em relação à economia e a direção do seu voto, o chamado voto econômico. De acordo com a Teoria da Escolha Racional, também estudada por Samuel Popkin (1991), os eleitores buscam reduzir os custos (no sentido de ‘esforços’) da tomada de decisão político-eleitoral usando atalhos informacionais, para tomar decisões racionais. No caso do voto econômico, os assuntos e dados referentes à economia seriam então o atalho escolhido pelo eleitor para decidir seu voto.

A teoria do voto econômico defende que um dos fatores que certamente explicam a decisão do voto é a avaliação que os eleitores fazem do governo atual, e que esta avaliação está fortemente associada à avaliação que estes mesmos cidadãos fazem da economia. Essa teoria, em sua descrição mais simples e seminal, conforme coloca Camargos (1999), é aquela que considera que os eleitores agirão sempre como juízes, independentemente das situações institucionais ou econômicas sob as quais estejam vivendo, fazendo o seguinte cálculo: “se a avaliação do governante é positiva, vota na situação; sendo negativa, vota na oposição”. (Key 1966; Kramer 1971; Fiorina 1981 e Guseh 1996P; *apud* Camargos, 1999).

Lewis-Beck e Stegmaeir (2007) descrevem que a ideia básica do voto econômico é a lógica de punição-recompensa (pontuada brevemente na introdução deste trabalho). Segundo esses autores, quando a economia vai bem, os eleitores premiam o incumbente com o seu voto; por outro lado, quando a econômica está mal, os eleitores punem o incumbente conferindo seu voto ao candidato desafiante.

Segundo Fernandes e Fernandes (2012), os fatores econômicos podem influenciar o comportamento eleitoral de três maneiras: 1) uma melhora nas condições

econômicas aumenta a probabilidade de voto no partido que é percebido como responsável pela mudança, 2) um melhor desempenho econômico tende a indicar uma capacidade administrativa do partido no poder, e 3) uma economia que vai bem tende a afetar positivamente os votos de eleitores que não se identificam partidariamente (Virmani, 2004 e Anderson e Morgan, 2011).

Um fator importante a ser averiguado na abordagem do voto econômico é a capacidade do governo influenciar de fato o desempenho da economia (Fernandes e Fernandes, 2012). Duch e Stevenson (2008 *apud* Fernandes e Fernandes, 2012) estudaram 18 democracias e mostraram que existe uma importante relação entre percepção econômica e a escolha do voto, sobretudo em sistemas que concentram o poder de decisão política.

Outro estudo que merece destaque é o de Fair (1978 *apud* Fernandes e Fernandes, 2012), que desenvolveu um modelo para prever o resultado das eleições presidenciais estadunidenses, através do qual se constatou que a decisão do voto dependeu principalmente dos eventos econômicos do último ano. O estudo de Fair (1978) aponta que os presidentes estadunidenses são reeleitos quando há crescimento econômico, o gasto federal está sob controle e uma importante guerra foi evitada.

Embora haja convergência na crença e afirmação de que fatores econômicos influem fortemente no resultado eleitoral, alguns aspectos são alvo de discordância entre os estudiosos do voto econômico. O primeiro diz respeito à direção do cálculo feito pelos eleitores: o voto econômico é retrospectivo (eleitores pensam no desempenho econômico do atual governo nos últimos anos) ou prospectivo (eleitores olham para o futuro projetando suas expectativas econômicas entre crença na prosperidade ou no declínio)? Key (1966) e Fiorina (1981) argumentam que o eleitorado julga a economia retrospectivamente para decidir seu voto. Já Downs (1957) sugere que os eleitores usam a lógica de que precisam mirar o futuro, escolhendo o candidato que irá governá-lo nos próximos anos, e comparando as performances que espera dos partidos competidores.

O segundo aspecto de desavença entre pesquisadores do voto econômico é a dimensão e abrangência do cenário econômico que os cidadãos consideram para decidir seu voto: os eleitores votam pensando em seu próprio bolso (*pocketbook voting*) ou considerando a situação econômica no país (*sociotropic voting*)? A teoria do 'próprio bolso' acredita que se a situação econômica pessoal ou da família piorou (retrospectiva) ou se creem que ela estará pior (prospectiva), os eleitores irão punir o incumbente; mas se essa situação melhorou (ou consideram que irá melhorar), irão premiar, através do

voto, o atual governante. Já na teoria ‘sociotrópica’ a crença é a de que se os eleitores acham que a economia do país cresceu (ou irá melhorar no próximo ano), votarão na continuidade do incumbente; mas se a economia nacional está pior (ou o cenário é de declínio econômico), os cidadãos irão punir o incumbente, dando seu voto para o desafiante.

Para contemplar o maior número de pesquisadores e mediar esses dois aspectos controversos, pesquisas de *survey* realizadas em inúmeros países democráticos buscam então incluir em seu leque de questões quatro perguntas básicas, fruto das quatro combinações possíveis entre a direção do cálculo<sup>1</sup> e a abrangência de cenário econômico considerada pelos eleitores: 1) Pensando no último ano, você diria que a economia do país está melhor, pior ou igual? (retrospectiva + sociotrópica); 2) Pensando no último ano, você diria que a sua situação financeira pessoal está melhor, pior ou igual? (retrospectiva + ‘próprio bolso’); 3) Pensando no próximo ano, você acredita que a economia do país estará melhor, pior ou igual? (prospectiva + sociotrópica); 4) Pensando no próximo ano, você acredita que a sua situação financeira pessoal estará melhor, pior ou igual? (prospectiva + ‘próprio bolso’). Além dessas quatro perguntas, que permitem a avaliação da influência de cada dimensão na decisão do voto, as pesquisas de *survey* costumam também, claro, incluir outras questões sobre temas políticos, de forma a permitir que os pesquisadores possam mensurar o impacto que a economia tem na decisão do voto em relação a outros fatores (Lewis-Beck e Stegmaeir, 2007).

Conforme já fora citado na introdução deste paper, em geral, para mensurar o impacto da situação econômica, em estudos com dados agregados, pesquisadores consideram o crescimento do PIB per capita, a variação da inflação e a variação do desemprego (variáveis independentes) e os resultados eleitorais (variável dependente). Já em estudos com dados individuais, coletados em pesquisa de opinião (*survey*), pesquisadores analisam o impacto da percepção e/ou avaliação da situação econômica (variáveis independentes) na intenção de voto do eleitor: mandatário/incumbente ou desafiante? (variável dependente).

A existência e o uso desses dois tipos de fontes de dados (dados agregados VS. dados individuais), abre espaço para um outro aspecto controverso, um dilema ainda não superado e em constante discussão pelos estudiosos do voto econômico, o problema

da falácia ecológica<sup>3</sup>: ao usar dados agregados a pesquisa realmente está conseguindo captar e tratar do que os eleitores pensam a respeito da economia em seu nível individual e pessoal? Ou seria essa uma relação passível de ser apontada como meramente espúria? (Lewis-Beck e Stegmaier, 2007). O presente trabalho tem como objetivo justamente propor o uso de um tipo de modelo de análise estatística que possibilita a combinação de ambos os formatos (macro e micro) de dados: os modelos hierárquicos, ou também chamados de modelos multiníveis.

### **3 - O modelo multinível hierárquico**

*Do que se trata, e quais seus fundamentos*

Luke (2004) inicia seu livro destacando uma frase de Dewey (1931) que cabe ser realçada de entrada aqui também. John Dewey disse que “se aventuraria a afirmar que a falácia mais difundida do pensamento filosófico seria a negligência do contexto” (tradução livre).

Inúmeras são as situações e desenhos de pesquisas nos quais indivíduos são influenciados por seu contexto. Por exemplo, é conhecido que a probabilidade de desenvolvimento de depressão é influenciada por estresse social. O desenvolvimento na infância é afetado por um leque de condições: dieta alimentar, quantidade de estímulos recebidos do ambiente em que se desenvolve, qualidade do relacionamento com a mãe, e outros. (Luke, 2004). Em estudos organizacionais, pesquisadores devem, por exemplo, investigar como as características do local de trabalho e a centralização do poder de decisão influenciam na produtividade do trabalhador; sendo os trabalhadores e as empresas unidades de análises e a mensuração feita em dois níveis (Raudenbush e Bryk, 2002). Na área de estudos de educação as análises multinível são muito usadas (Rios Neto, César e Gonçalves, 2011; Fernandes e Barbosa, 2012); afinal os alunos, além de terem seu sucesso educacional influenciado por características pessoais e fatores de background escolar familiar, são agrupados em várias dimensões, tendo o seu resultado e alcance educativo fortemente afetado pela sala de aula em que estudam, pela escola em que estão matriculados, pelo sistema educacional vigente, pelo estado e/ou país em que residem.

---

<sup>3</sup> Soares e Terron (2008) escrevem que a inferência de comportamento individual através de resultados e dados agregados é um problema estatístico denominado como falácia ecológica.

Todos estes exemplos servem para ilustrar a afirmativa de que características ou processos que ocorrem em um nível mais alto de análise podem ter relação com processos observados no nível mais baixo. Segundo Luke (2004), sabendo-se que os constructos explicativos de um fenômeno podem então operar em diferentes níveis, para avaliar adequadamente as relações existentes é preciso técnicas e ferramentas de análise especializadas. Os modelos multiníveis e seus softwares operacionais (SPSS, SAS, R, etc.) se propõem a esta tarefa.

Um modelo multinível é um modelo estatístico aplicado a dados coletados em mais de um nível (Luke, 2004: 8). A modelagem estatística multinível tem sido desenvolvida ao longo das últimas décadas, por diferentes áreas de estudos, e recebido várias nomenclaturas diferentes, tais como modelos lineares hierárquicos (Raudenbush e Bryk, 2002), modelos de coeficientes aleatórios (Longford, 1993), modelos de efeitos mistos (Pinheiro e Bates, 2000), modelos com estruturas de covariância (Muhtén, 1994), os modelos de curva de crescimento (Mc Ardle e Epstein, 1987). Neste trabalho, serão usadas ora a nomenclatura modelo multinível, ora modelo hierárquico.

O argumento mais simples para então defender o uso de modelos multiníveis é o seguinte: em função do fato de muitos fenômenos que estudamos serem multiníveis em sua natureza, devemos então usar teorias e técnicas analíticas que também sejam multiníveis. Se isso não for feito, o pesquisador poderá incorrer em sérios problemas de análise e inferência (Luke, 2004: 4).

Os principais exemplos de problemas advindos da má interpretação de dados são a falácia ecológica (*ecological fallacy*) e a falácia atomística (*atomistic fallacy*). A primeira é a falha em se atribuir a indivíduos resultados observados nos grupos (em dados agregados). A segunda é o inverso: fazer inferências para grupos usando informações de nível individual. Essas falácias são problemas de inferência, e não de mensuração. Cabe deixar claro que caracterizar um coletivo (*high level*) usando informações obtidas a partir de seus membros (*lower level*) é perfeitamente admissível; os tipos de falácias descritas acima acontecem quando as relações observadas em um determinado nível são inadequadamente imputadas a ocorrer da mesma forma em algum outro nível, inferior ou superior (Luke, 2004:6).

Devido a seu caráter multinível, a relação existente as variáveis predictoras incluídas em um ou outro nível, a abordagem hierárquica não assume o pressuposto de independência entre as observações das unidades pertencentes a um grupo agregado; admitindo então a violação da hipótese de erros independentes e identicamente

distribuídos (i.i.d), uma vez que estes podem acabar sendo produtos da dependência (homogeneidade) entre os indivíduos de um mesmo grupo.

### *O modelo formal e suas notações*

Os modelos de regressão multiníveis/hierárquicos são conhecidos por incorporarem naturalmente a estrutura hierárquica ou de agrupamento da população em estudo. Cada um dos níveis na estrutura hierárquica é representado por um sub-modelo, que expressa o relacionamento entre as variáveis dentro de um determinado nível, além de apontar também como os fatores em um nível influenciam as relações que ocorrem em outros níveis.

Um modelo de regressão multinível ajustado em dois níveis, conforme o que será proposto neste trabalho, coloca os fatores individuais como a unidade do nível 1, identificado pelo subscrito  $i$ , e as dimensões agregadas (grupos) como unidade do nível 2, identificada pelo subscrito  $j$ .

Aplicando-se a explicação de imediato ao caso do estudo de comportamento eleitoral da corrente do voto econômico (de forma a ilustrar a explanação e desde já relatar a formatação do modelo a ser proposto), considerar-se-á então na modelagem a contagem de  $j$  grupos (estados federativos brasileiros), cada um deles com  $i$  eleitores.

Neste sentido, a variável dependente  $Y_{ij}$ , que representará o voto do eleitor (incumbente vs. desafiantes), será calculada em função da dimensão  $i$  e da dimensão  $j$ , que mensuram respectivamente a percepção do eleitor em relação a algum fator econômico, como, por exemplo, inflação (fator individual) e as taxas de desemprego e crescimento do PIB por estado (fatores agregados/coletivos).

Para analisar de forma linear (em apenas um nível) a relação descrita acima, entre decisão do voto e os assuntos econômicos, o modelo de regressão clássico poderia ser definido da seguinte forma:

$$Y_i = \beta_0 + \beta_1 X1_i + \beta_2 X2_i + \beta_3 X3_i + e_i$$

- Onde  $Y_i$  é a variável resposta do eleitor  $i$ ;

-  $\beta_0$  é o intercepto (média global da variável resposta quando o valor das variáveis explicativas são nulos);

-  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  são os coeficientes de inclinação, que dizem respeito ao efeito das variáveis explicativas X1, X2, e X3 inseridas no modelo (avaliação individual da inflação, taxa de desemprego e crescimento do PIB, por estado, respectivamente) sobre a variável dependente/resposta toda vez que as variáveis independentes são acrescidas em uma unidade;

- E  $e_i$  é o erro aleatório do modelo, o resíduo não explicado pelo modelo.

O modelo de regressão clássico, como visto acima, possui apenas um intercepto. Já no modelo multinível, ou mais especificamente o modelo de dois níveis (*The Basic Two-Level Multilevel Model*) que será proposto neste trabalho, tanto o intercepto quanto os coeficientes de inclinação são considerados como variáveis aleatórias que variam entre as estruturas de análise; no caso deste estudo, de estado para estado.

Considerando-se então que o objetivo de um modelo multinível é estimar valores de uma variável resposta/dependente com base em uma função de variáveis explicativas em mais de um nível, tem-se que o modelo para uma análise em dois níveis pode ser descrito da seguinte forma:

$$\text{Nível 1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij}$$

$$\text{Nível 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}$$

Como a análise hierárquica tem sua estrutura de modelagem semelhante à regressão linear, pode-se substituir as equações do nível 2 no modelo de equação do nível 1. Dessa forma, a equação para dois níveis pode ser escrita da seguinte maneira:

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + u_{0j} + (\gamma_{10} + \gamma_{11}W_j + u_{1j}) X_{ij} + e_{ij} ; \text{ ou:}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}W_j + \gamma_{11}W_jX_{ij} + [u_{0j} + u_{1j} X_{ij} + e_{ij}]$$

A notação dos modelos hierárquicos apresenta duas diferenças em relação ao modelo de regressão linear clássico: convencionalmente adota símbolos diferentes para

designar seus coeficientes e conta com a presença de dois números ou letras subscritos em seus elementos.

O símbolo  $\gamma$  (gama) entra no lugar do tradicional  $\beta$  (coeficiente de regressão). Os coeficientes de regressão hierárquicos (os gamas) têm dois subscritos em vez de um: o primeiro representa o número da variável do nível micro (primeiro nível) e o segundo indica o número da variável do nível macro (segundo nível).

Por exemplo,  $\gamma_{21}$  é o efeito da primeira variável do nível macro sobre a segunda variável do nível micro. Além de proporcionar a obtenção de melhores estimativas para os parâmetros, o modelo hierárquico permite analisar como as variáveis independentes medidas no nível 2 modificam o efeito das variáveis independentes medidas no nível 1 (Radenbush e Byrk, 2002).

O valor 0 (zero) do subscrito em um coeficiente de regressão hierárquico significa o intercepto; ou seja,  $\gamma_{00}$  é o valor médio esperado da variável dependente quando as independentes são nulas, e  $\gamma_{01}$  é o efeito da primeira variável do nível macro sobre o intercepto (quando a variável do nível micro é nula).

Em uma regressão hierárquica onde são considerados dois níveis utiliza-se uma letra para indicar uma variável explicativa do nível micro (letra X no caso aqui exemplificado) e outra para denotar variáveis explicativas do nível macro (letra W).

As variáveis mensuradas no nível micro têm dois subscritos: “i” indica o número do indivíduo (eleitor) e “j” representa o número da unidade do nível macro (estado). Já as variáveis de nível macro tem um único subscrito, pois dizem respeito apenas ao nível estrutural da análise. A idéia básica da modelagem multinível é que a variável dependentetem um aspecto individual assim como aspecto de grupo. Ficando determinado, entretanto, que esta variável resposta sempre será medida no nível mais desagregado, pois o modelo linear hierárquico é um modelo para explicar o que acontece no nível mais baixo, mais detalhado.

Sobre os elementos de erro  $u_{0j}$ ,  $u_{1j}$  e  $e_{ij}$ , tem-se que a diferença da equação do modelo hierárquico linear para um modelo de regressão usual é que este modelo multinível contém mais de um termo de erro: um (ou mais) para cada nível (Snijders e Bosker, 1999).



#### 4 – Operacionalização do modelo multinível aplicada ao estudo do voto econômico

O modelo apresentado na seção anterior expõe notação padrão para uma análise em que há apenas uma variável micro ( $X_{ij}$ ) e uma variável macro ( $W_j$ ). Quando há mais variáveis a serem consideradas na análise pretendida pelo modelo, seja no nível micro ou macro (como no caso do modelo que será aqui proposto), é preciso incluir este novo elemento na equação preditiva.

Em geral, são três os fatores econômicos indicados como preditores da decisão do voto, segundo a teoria do voto econômico:

- 1) avaliação ou preocupação em relação à inflação;
- 2) crescimento econômico (crescimento do PIB real por estado, anual ou acumulado, divulgado por institutos de estatísticas oficiais);
- 3) nível de desemprego (percentual de desocupação, por estado ou por município, divulgado por instituto de estatísticas oficiais) – em alguns estudos, é comum também avaliar a preocupação dos eleitores, em nível micro (individual), com relação à possibilidade de ficarem desempregados, de forma a mensurar qual a percepção e avaliação que fazem sobre a estabilidade econômica a qual estão expostos.

Sendo assim, uma proposta de modelo aplicado ao estudo do voto econômico que incluísse os três fatores citados teria que ter em sua equação três variáveis independentes: uma delas inserida em nível micro (avaliação/percepção da inflação) e as outras duas em nível macro (taxas estaduais de crescimento econômico e desemprego).

*Variáveis independentes:*














$X_{ij}$  = avaliação/percepção sobre a inflação;

$W_j$  = crescimento econômico (crescimento real do PIB por estado  $j$ );















$Z_j$  = desemprego (nível de desocupação/desemprego por estado  $j$ ).

Cabe falar um pouco sobre cada um dos fatores econômicos incluídos como variáveis explicativas da análise. Sabe-se que não há mensuração e disponibilização de taxas agregadas em nível estadual para medição de inflação; os institutos de estatísticas oficiais costumam trabalhar com o IPCA (Índice Nacional de Preços ao Consumidor Amplo) e o INPC (Índice Nacional de Preços ao Consumidor) para algumas regiões metropolitanas do país. Neste sentido, não seria possível trabalhar esta informação

(inflação) em nível agregado, nos restando então a possibilidade de analisá-la apenas enquanto fator de avaliação/percepção individual dos cidadãos-eleitores. Cabe mencionar que, além disso, inflação de preços é uma dimensão mais próxima da avaliação do cidadão, que tem propriedade para opinar a respeito do valor de compra de itens que costuma adquirir no seu dia-a-dia, tais como alimentos, artigos de higiene, gasolina, etc. Por outro lado, acredita-se que avaliar crescimento econômico e desemprego poderia demandar mais custo de informação e ser mais complicado para o cidadão. Assim, estas dimensões devem então ser incluídas na análise como fatores agregados por dois motivos: por demandarem maior custo de informação e por terem seus níveis significativamente desiguais de estado para estado<sup>4</sup>, como pode ser visto nas tabelas a seguir:

| Unidade federativa   | Crescimento real anual do PIB - 2010 | Taxa de desocupação/ desemprego* - 2010 |
|--|--------------------------------------|---|
|  Acre                 | 14,70%                               | 8,0%                                    |
|  Alagoas            | 15,70%                               | 10,7%                                   |
|  Amapá              | 11,60%                               | 11,9%                                   |
|  Amazonas           | 20,50%                               | 9,7%                                    |
|  Bahia              | 12,60%                               | 10,9%                                   |
|  Ceará              | 18,50%                               | 7,7%                                    |
|  Distrito Federal   | 14,00%                               | 8,2%                                    |
|  Espírito Santo     | 23,00%                               | 7,3%                                    |
|  Goiás              | 13,90%                               | 6,3%                                    |
|  Maranhão           | 13,50%                               | 8,7%                                    |
|  Mato Grosso        | 4,00%                                | 6,3%                                    |
|  Mato Grosso do Sul | 19,60%                               | 6,2%                                    |
|  Minas Gerais       | 22,40%                               | 6,8%                                    |

4 Há pesquisas que trabalham com a avaliação individual (micro) que o eleitor faz do crescimento econômico e com avaliação/preocupação em relação ao desemprego. A decisão de avaliar estes fatores em dimensão agregada é uma escolha metodológica e estratégica do modelo proposto neste trabalho.

|   |        |       |
|---|--------|-------|
|  Pará                | 33,30% | 9,2%  |
|  Paraíba             | 11,20% | 8,6%  |
|  Paraná              | 14,30% | 5,0%  |
|  Pernambuco          | 21,30% | 11,1% |
|  Piauí               | 15,90% | 7,9%  |
|  Rio de Janeiro      | 15,00% | 8,5%  |
|  Rio Grande do Norte | 15,90% | 9,9%  |
|  Rio Grande do Sul   | 16,90% | 4,9%  |
|  Rondônia            | 16,40% | 5,5%  |
|  Roraima             | 13,30% | 7,7%  |
|  Santa Catarina    | 17,40% | 3,8%  |
|  São Paulo         | 15,00% | 7,6%  |
|  Sergipe           | 21,00% | 10,3% |
|  Tocantins         | 18,30% | 7,4%  |

\*A taxa de desemprego é calculada a partir do número de pessoas maiores de 10 anos economicamente ativas e desocupadas, segundo o CENSO de 2010.

Falando agora sobre a variável resposta, tem-se que a variável dependente das análises de decisão do voto costuma ser operacionalizada em formato binário, sendo atribuído um numeral para o voto no incumbente e outro para o voto em qualquer dos candidatos desafiantes; essa operacionalização permite que seja mensurado individualmente se o eleitor vota pela reeleição ou sucessão do *status quo* ou se deseja a mudança. Em geral, atribui-se o valor 1 para todos os casos em que ocorre o fenômeno de interesse do estudo (voto no incumbente, no caso deste trabalho) e 0 para os casos contrários. Sabe-se que variáveis do tipo binária (também chamadas de *dummy*) demandam condições e distribuições<sup>5</sup> específicas para operarem. Considerando então

5 Uma vez que o comparecimento eleitoral é medido de forma dicotômica (0 e 1), cabe utilizar um modelo generalizado para variáveis com distribuição de Bernoulli (Raudenbush & Bryk, 2002)

variáveis respostas binárias, os modelos hierárquicos lineares se tornam inapropriados, pois a suposição de normalidade feita por grande parte dos modelos é violada, assim como a de homocedasticidade dos erros (variância não é homogênea). Para resolver o problema de variáveis que não são distribuídas normalmente, a solução é incluir uma transformação e escolher a distribuição de erros apropriada, não necessariamente a normal (Raudenbush e Bryk, 2002). Os modelos estatísticos que realizam essa tarefa são chamados modelos lineares generalizados. O método de estimação nos modelos lineares generalizados é o procedimento de máxima verossimilhança que usa a inversa da função de ligação para obter a variável resposta.

Além disso, o valor predito para a variável binária, se observado como a probabilidade do evento ocorrer, deverá situar-se entre 0 e 1. Essa definição permite que a análise seja feita da seguinte forma: a função de regressão é a probabilidade da variável dependente ser igual a 1, dado certo valor X; e o coeficiente de regressão ( $\beta$ , na notação de regressão clássica, ou  $\gamma$  em modelos hierárquicos) mede a variação na probabilidade de que a variável resposta Y seja igual a 1 (ou seja, a variação da probabilidade de que o indivíduo vote no incumbente, dados certos valores de X – variáveis independentes). O uso dos modelos multiníveis lineares não é capaz de enquadrar o resultado ao intervalo de 0 a 1, podendo atingir qualquer valor real.

Nesse sentido, o modelo proposto neste estudo, considerando que a variável resposta tem mensuração binária, será o modelo *logit* hierárquico (ou modelo hierárquico logístico), no qual o valor predito é convertido no logaritmo de chance de sucesso (chance de que o fenômeno de interesse aconteça; voto no incumbente), denominado de *log-odds*, e com a seguinte equação abaixo:

$$\eta_{ij} = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right)$$

Onde  $\eta_{ij}$  é o logaritmo da chance de sucesso (chance de que o fenômeno em estudo aconteça) para o indivíduo i, e  $\pi_{ij}$  é a probabilidade de sucesso para o indivíduo i.

Em suma, tem-se então que a variável dependente em um modelo logit hierárquico para estudo do voto econômico será o logaritmo da chance de o eleitor ter votado no incumbente ou em seu sucessor (apresentar variável resposta 1 = voto na situação, candidato à reeleição ou no candidato apoiado pelo atual mandatário/candidato

da continuidade). Essa operação de ‘logaritmizar’ uma variável pode (e deve) ser feita com o auxílio de softwares estatísticos.

Na análise hierárquica é conveniente considerar primeiro o modelo sem variáveis explicativas no segundo nível, e o intercepto ( $\beta_0$ ) tendo efeito aleatório. Como já fora dito, o intercepto é considerado como o valor médio esperado da variável dependente quando as variáveis independentes são iguais a zero. O intercepto varia entre as unidades  $j$ , grupos do nível 2 (Gelman e Hill, 2007). Esse modelo sem as variáveis explicativas é chamado de não-condicional, ou modelo hierárquico nulo, e possibilita identificar a contribuição de cada nível para a variação total através do coeficiente de correlação intraclasse, que pode ser estimado a partir do modelo nulo (Snijders e Bosker, 1999).

O modelo nulo apresenta as seguintes especificações para o primeiro e segundo nível, respectivamente:

$$(1) \eta_{ij} = \beta_{0j} + e_{ij}$$

Onde  $i = 1 \dots n$  são as unidades do nível 1, neste caso eleitores;

$j = 1 \dots J$  são as unidades do nível 2, neste caso estados;

$\beta_{0j}$  é o resultado médio para a  $j$ -ésima unidade; e

$e_{ij}$  é o efeito aleatório associado ao nível 1

$$(2) \beta_{0j} = \gamma_{00} + u_{0j}$$

Onde  $u_{0j}$  é o efeito aleatório associado ao nível 2, e

$\gamma_{00}$  é a média do resultado para todos os indivíduos.

O coeficiente de correlação intraclasse  $\rho$ , de acordo com Raudenbush e Bryk (2002), indica a proporção da variância explicada pela estrutura de agrupamento da população. Como a variância se decompõe em dois componentes independentes:  $\sigma_e^2$  (variância do erro  $e_{ij}$  do nível 1) e  $\sigma_u^2$  (variância do erro  $u_{0j}$  do nível 2), o coeficiente de correlação intraclasse para um modelo hierárquico linear pode ser obtido em função desses dois componentes:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2}$$

A interpretação deste coeficiente é a seguinte: a correlação intraclasse é uma proporção da variância no nível do grupo comparada com a variância total. A pergunta que se coloca então é se a incorporação de mais de um nível hierárquico se justifica do ponto de vista econométrico, isto é, se a inclusão do segundo nível aumenta o poder de explicação da variabilidade dos dados do modelo.

Contudo, o cálculo do coeficiente de correlação intraclasse da forma como colocado acima se aplica apenas aos modelos multiníveis lineares. Em modelos não lineares, como os logísticos, essa fórmula não é útil porque a variância no nível individual é heterocedástica, ou seja, não é constante ao longo do tempo (Raudenbush & Bryk, 2002, p. 298). Nesse caso, há que se recorrer então à alternativa proposta por Snijders e Bosker (1999), que trata a variável dependente como uma medida contínua latente seguindo uma distribuição logística. Assim, a fórmula para estimar o coeficiente de correlação passa a ser:

$$\rho = \frac{\sigma_{uo}^2}{\sigma_{uo}^2 + \pi^2/3}$$

Após identificar então os modelos nulos, tem-se o próximo e último passo na trajetória de construção de um modelo hierárquico logístico: a inclusão das variáveis independentes do nível 1, primeiramente, a fim de compreender a variabilidade associada a este nível; e em seguida a inclusão das variáveis explicativas de nível 2.

Finalmente então aplicando a análise ao modelo proposto por este trabalho, a variável independente associada ao nível 1 (nível individual ou nível micro) será apenas uma: avaliação/preocupação com relação à inflação (aumento do preço dos produtos).

O modelo com a inclusão desta variável explicativa de nível 1 ficaria da seguinte forma:

$$\eta_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + e_{ij} \quad (1)$$

- Onde os subscritos i e j indexam, respectivamente, eleitores e estados confederados brasileiros nos quais residem estes eleitores;
- $\beta_{0j}$  é o intercepto;
- $\beta_{1j}$  é o coeficiente de regressão, ou parâmetro a ser estimado pelo modelo; e

- $X_{1ij}$  é uma variável dummy para a avaliação/preocupação com relação à inflação do eleitor  $i$  no estado  $j$
- $\epsilon_{ij}$  é o termo de erro aleatório (randômico) associado ao nível 1.
- Nesse modelo, a equação do nível 2 ainda possui a forma apresentada pelo modelo nulo ( $\beta_{0j} = \gamma_{00} + u_{0j}$ ).

Após estimar esse modelo de nível 1, e verificar que a variância é significativamente diferente de zero, passa-se então a incluir, passo a passo, as variáveis explicativas do segundo nível. A inclusão e a significância estatística das variáveis explicativas de nível 2 buscará identificar se a propensão do eleitor votar na reeleição do incumbente ou em seu sucessor (variável resposta = 1) varia no contexto estrutural (estado) em que o eleitor está localizado, devido a características econômicas (crescimento do PIB real e nível de desocupação) específicas de cada estado:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{1j} + \gamma_{02}W_{2j} + u_{0j} \quad (2)$$

- Onde  $W_{1j}$  é o crescimento real do PIB por estado  $j$ ; e
- $W_{2j}$  é a taxa de desocupação/nível de desemprego por estado  $j$ .

Além dos fatores econômicos de nível 1 e nível 2, pode-se adicionar também ao modelo, tanto para o nível individual como para o nível agregado, variáveis chamadas de ‘variáveis controle’. Exemplos de variáveis controle para o nível 1 seriam a inclusão no modelo de variáveis como idade, a renda, a escolaridade, o sexo, a preferência partidária, e outros atributos individuais que o pesquisador julgar necessário serem ‘controlados’, ou mantidos constantes, na análise. No nível agregado, poderiam ser exemplos de variáveis controle a região geográfica, o partido do governador do estado, o tamanho do colégio eleitoral de cada estado, entre outros fatores associados ao nível 2.

Após elucidar os fundamentos e os benefícios do uso de modelos multiníveis, e detalhar toda sua lógica de operacionalização aplicando a análise ao estudo do voto econômico, substituindo-se então o modelo (2) no modelo (1) e recapitulando as definições colocadas em separado para os termos das equações de nível 1 e nível 2, tem-

se que o modelo *logit* hierárquico para estudo do voto econômico proposto por este trabalho seria o seguinte:

$$\eta_{ij} = \underbrace{\beta_{0j}}_{\gamma_{00} + \gamma_{01} W_{1j} + \gamma_{02} W_{2j} + u_{0j}} + \beta_{1j} X_{1ij} + e_{ij}$$

$$\eta_{ij} = \gamma_{00} + \gamma_{01} W_{1j} + \gamma_{02} W_{2j} + u_{0j} + \beta_{1j} X_{1ij} + e_{ij}$$

- Onde  $\eta_{ij}$  é o logaritmo da chance de sucesso (chance de que o eleitor vote no incumbente ou em seu sucessor) para o indivíduo  $i$ ;
- $\gamma_{00}$  é a média do resultado para todos os indivíduos;
- $\gamma_{01}$  e  $\gamma_{02}$  são os coeficientes de regressão das variáveis de nível 2;
- $W_{1j}$  é o crescimento real do PIB por estado  $j$ ;
- $W_{2j}$  é a taxa de desocupação/nível de desemprego por estado  $j$ ;
- $u_{0j}$  é o erro aleatório associado ao nível 2
- $\beta_{1j}$  é o coeficiente de regressão, ou parâmetro a ser estimado pelo modelo; e
- $X_{1ij}$  é uma variável *dummy* para a avaliação/preocupação com relação à inflação do eleitor  $i$  no estado  $j$
- $e_{ij}$  é o termo de erro aleatório (randômico) associado ao nível 1.

O modelo recomendado se baseia na inclusão de algumas das muitas possibilidades de variáveis referentes a fatores econômicos que podem estar associadas à decisão do voto. A discussão e o objetivo se centraram mais na operacionalização metodológica de um modelo para estudo do voto econômico do que na discussão e fundamentação das variáveis em si; acredita-se haver possibilidade, ou mais que isso, necessidade, de inclusão de outros fatores econômicos explicativos do comportamento eleitoral dos cidadãos.



## 5 - Considerações finais

O trabalho aqui apresentado teve como objetivo central a proposição de um modelo quantitativo de análise da teoria do chamado voto econômico; teoria na qual acredita-se que o eleitor decide seu voto com base em fatores e avaliações que faz da economia, decidindo pela manutenção do incumbente (ou pela eleição de seu sucessor, quando o mandatário não disputa a reeleição) quando a economia vai bem, e optando por algum dos candidatos desafiantes quando os fatores econômicos estão negativos.

A decisão foi pela proposição de um modelo hierárquico multinível, pois este é um tipo de modelagem estatística que permite análise de dados que estejam arranjados em mais de um nível, como é o caso dos assuntos econômicos, assumindo que a variável dependente é medida no nível mais desagregado (decisão do voto), mas que as variáveis independentes podem ser mensuradas e analisadas em mais de um nível (indivíduos, grupos, bairros, cidades, estados, regiões, países, etc.). Além disso, por ser um modelo em que a variável resposta assume dois valores (1 ou 0, voto ou não-voto no incumbente), optou-se por utilizar um modelo hierárquico logístico (logit), de forma a contemplar as definições e necessidades específicas de modelagem e de distribuição quando se pretende o uso de uma variável binária (*dummy* ou dicotômica).

As variáveis incluídas no modelo ajustado e proposto são fruto de ponderação e opção da autora, não sendo obrigatoriamente fixas, imutáveis e questionáveis. A proposição de um modelo hierárquico teve mais a finalidade de abrir a discussão e reflexão para o uso desse tipo de modelagem em estudos sobre o voto econômico. Inúmeras são as variáveis que podem ser estudadas e incluídas em um modelo de estudo sobre o efeito de fatores econômicos no voto, sendo alguns exemplos a avaliação da econômica em geral, a preocupação com o desemprego, a expectativa de avanço econômico para o próximo ano, as taxas de juros praticadas em determinado período de tempo, entre outros. Além de alterações nas variáveis incluídas, o pesquisador pode também usar o modelo para pesquisar esferas diferentes, tais como comprar países dentro de algum espectro de estudo. O modelo recomendado é apenas exatamente o que se propõe a ser: um modelo. Adequações e adaptações podem, e devem, ser feitas; de acordo com o objeto que se deseje estudar.

Este trabalho figura como um esforço elementar de se trazer para o ‘mundo’ do comportamento eleitoral a discussão metodológica sobre modelos hierárquicos. Muito ainda precisa ser feito.

Espera-se que a discussão e a proposição metodológica desenvolvidas neste trabalho sirvam de impulso para que o uso de modelos multiníveis ganhe espaço entre os estudiosos do comportamento eleitoral em geral (não só os ‘adeptos’ à teoria do voto econômico, mas também analistas da teoria sociológica e psicológica), para que novas análises sejam feitas e adaptações sejam indicadas, para que o modelo sugerido seja testado empiricamente em diferentes cenários e contextos, e melhorado em sua construção.

## Referências bibliográficas

ANTUNES, R. (2010). Theoretical models of voting behavior. **Revista Exedra**, (4):145-170. Coimbra.

BARTELS, L. M. (2010). The Study of Electoral Behavior. In: **The Oxford Handbook of American Elections and Political Behavior**. 239-261. Oxford University Press.

CAMARGOS, M. (1998). Do bolso para as urnas – a influência da economia na escolha entre Fernando Henrique e Lula nas eleições de 1998. **Dissertação** apresentada ao **Instituto Universitário de Pesquisas do Rio de Janeiro** como requisito parcial para obtenção do grau de Mestre em Ciência Política. Rio de Janeiro.

CAMARGOS, M. (2013). Refinando a teoria do voto econômico - América Latina revisitada. In: TELLES, H; MORENO, A. **Comportamento Eleitoral e Comunicação Política na América Latina**. Belo Horizonte: Editora UFMG.

CASTRO, M. M. M. (1992). Sujeito e estrutura no comportamento eleitoral. **Revista Brasileira de Ciências Sociais**, nº 20, vol. 01.

FERNANDES, G.; FERNANDES, I. (2012) **O voto econômico no Brasil: evidências com dados em painel e em multinível**. Disponível em: [http://www.cienciapolitica.org.br/wp-content/uploads/2014/04/14\\_7\\_2012\\_12\\_54\\_29.pdf](http://www.cienciapolitica.org.br/wp-content/uploads/2014/04/14_7_2012_12_54_29.pdf)

FIGUEIREDO, M. (1991) **A Decisão do Voto: Democracia e Racionalidade**. São Paulo: Sumaré/Anpocs.

GELMAN, A., HILL, J. (2007) **Data Analysis using Regression and Multilevel/Hierarchical Models**. 2 ed. New York: Cambridge University Press.

HOX, J. (2002). **Multilevel Analysis: Techniques and Applications**. New Jersey: Lawrence Erlbaum Associates, Inc.

LEWIS-BECK, M. S; STEGMAIER, M. (2007). Economic models of voting, in: Russell Dalton and Hans-Dieter Klingemann (eds). **The Oxford Handbook of Political Behavior**. pp. 518-537.

LUKE, A. D. (2004) **Multilevel modeling. Quantitative applications in the social science**. Sage Publications, Inc. Thousand Oaks, California.

McGRANE, D. (2007) **Socio-Economic Determinants of Voting Behaviour in Canadian Provincial Elections from 1988 to 2006**. Paper apresentado na Conferência Annual da Associação Canadense de Ciência Política. Em Saskatoon, Saskatchewan, Canadá.

POPKIN, S.(1991) **The Reasoning Voter**. Chicago: The University of Chicago Press.

RATTO, M. C; MONTERO, J. R. (2013) O Comportamento Eleitoral dos Argentinos - Algumas chaves das eleições presidenciais de 2007. In: TELLES, H; MORENO, A. **Comportamento Eleitoral e Comunicação Política na América Latina**. Belo Horizonte: Editora UFMG.

RAUDENBUSH, S. W., BRYK, A. S. (2002) **Hierarchical linear models: applications and data analysis methods**. 2.ed. Londres, Nova Deli: Sage.

SANTOS, S. A. (2011) Comportamento do voto: reeleições dos prefeitos em capitais brasileiras – 2008. **Dissertação de Mestrado** apresentada ao Programa de Pós-Graduação em Ciência Política da **Universidade Federal do Paraná**, como requisito para obtenção do título de Mestre em Ciência Política.

SNIDERMAN, P. & LEVENDUSKY, M. (2009). An Institutional Theory of Political Choice. In: DALTON, R. & KLINGEMANN, H. (eds.). **The Oxford Handbook of Political Behavior**. Oxford: Oxford University.

SNIJDERS, T. A., BOSKER, R.J. (1999) **An introduction to basic and advanced multilevel modeling**. 1 ed. London, Thousand Oaks, California, New Delhi: Sage.

SOARES, G. A. D. e TERRON, S. L. (2008). A Geografia Eleitoral da Reeleição de Lula no Brasil: explorando conceitos, métodos e técnicas de análise geoespacial”. Artigo apresentado no V Congresso Latino Americano de Ciência Política – ALACIP, São José da Costa Rica.

TELLES, H; MORENO, A. (2013) **Comportamento Eleitoral e Comunicação Política na América Latina**. Belo Horizonte: Editora UFMG.

VEIGA, L. F. (2001). Em busca de razões para o voto: o uso que o homem comum faz do horário eleitoral. **Tese de Doutorado** apresentada ao **Instituto Universitário de Pesquisas do Rio de Janeiro**, como requisito parcial para a obtenção do grau de Doutor em Ciência Política.