

UNIVERSIDADE FEDERAL DO CEARÁ – UFC
CURSO DE PÓS-GRADUAÇÃO EM ECONOMIA - CAEN

FRANCISCO GILDEMIR FERREIRA DA SILVA

**MODELANDO VALOR DE TEMPO DE VIAGEM PARA MODOS
CONCORRENTES POR DIFERENTES MODELOS LOGIT: O QUE
SE GANHA E O QUE SE PERDE?**

FORTALEZA – CEARÁ

2011

Francisco Gildemir Ferreira da Silva

MODELANDO VALOR DE TEMPO DE VIAGEM PARA MODOS
CONCORRENTES POR DIFERENTES MODELOS LOGIT: O QUE SE
GANHA E O QUE SE PERDE?

Tese submetida à Coordenação do Curso de Pós-Graduação em Economia da Universidade Federal do Ceará – CAEN/UFC, como requisito parcial para obtenção do Título de Doutor em Economia.

Orientador: Prof. PhD. Sergio Aquino DeSouza

FORTALEZA – CEARÁ

2011

MODELANDO VALOR DE TEMPO DE VIAGEM PARA MODOS CONCORRENTES POR DIFERENTES MODELOS LOGIT: O QUE SE GANHA E O QUE SE PERDE?

Tese apresentada em cumprimento às exigências do Doutorado em Economia da Universidade Federal do Ceará- UFC, para obtenção do diploma de Doutor em Economia área de Concentração: Organização Industrial e Econometria Aplicada.

Aprovada em: ____/____/____

BANCA EXAMINADORA

Prof. PhD. Sergio Aquino DeSouza
Universidade Federal do Ceará

Prof(a). PhD. Carlos Felipe Grangeiro Loureiro
Universidade Federal do Ceará

Prof(a). PhD. Ricardo Antônio de Castro Pereira
Universidade Federal do Ceará

Prof(a). PhD. Ricardo Brito Soares
Universidade Federal do Ceará

Prof(a). PhD. Augusto Marcos Sena
Universidade de Fortaleza

AGRADECIMENTOS

Em uma tese, alguns anos de nossa vida são transformados em esforço na busca de conhecimento, nesse momento, temos consciência que aprendemos muito e que temos muito a aprender, mas que o que se tem de conhecimento adquirido ou por adquirir não foi nem será fruto individual e sim que precisa de um conjunto de fatores e pessoas, o principal entre estes é Deus e a ele agradeço por ter me dado à possibilidade de nascer, viver e me transformar continuamente. Entretanto, não tem como deixar de lado os companheiros da vida, pois a tese é o resultado da abdicção de momentos e ganho de outros com pessoas queridas dentro de nossas vidas, nos fazendo questionar sobre as coisas, nos fazendo correr e motivando, a saber, mais e nos compreendendo dia a dia. No rol de pessoas que fazem parte deste processo de busca constante pelo saber tenho gratidão imensa aos meus avôs (Manuel Vitor de Goes e Livina Ferreira), pois me deram carinho e indiretamente me puseram a questionar e observar o mundo ao passearem comigo em tardes longas e de observação constante do meu entorno. Seguidamente tenho minha família direta: mãe (Vânia), irmãos (Laudenir e Germana) e sobrinhos (Lucas, Luana, Gabriel, Arthur, Leo, e Alan) que me acompanharam em toda a vida ou que surgiram em momentos distintos enchendo-me de apreço e de motivação de ter mais do que dinheiro para lhes dar, pois isso não é importante quando se é querido e se dá carinho e amor.

A CAPES pelo apoio financeiro com a bolsa de doutorado.

Aos companheiros, colegas, amigos e conhecidos que, de uma forma ou de outra, torceram por mim ou não. Com certeza as energias que emanaram dessas pessoas me fizeram mais forte frente aos desafios, mais sábio perante as decisões e mais alegre ao enfrentar uma vida cheia de altos e baixos. Em particular a equipe do LEP – Laboratório de Estudos da

Pobreza na pessoa do professor Flávio Ataliba por disponibilizar o material do laboratório para desenvolvimento de cálculos que em outros *desktops* não seria possível. Aos membros do LEP, Dércio Nonato Chaves de Assis, Valdemar Rodrigues de Pinho Neto, Arley Rodrigues Bezerra e José Iranildo da Silva Araújo que participaram dos momentos em conversas sobre a vida e sobre os cálculos intermináveis dos computadores.

Aos colegas que entraram comigo no CAEN nos idos de 2008, em destaque ao Sérgio, Pedro Ivo, Rafael Carneiro e Aline Menezes. Companheiros pacientes e dedicados a viver a vida aprendendo e rindo muito.

Aos professores do CAEN que suportaram um engenheiro a procura de aprender economia, em particular ao atual coordenador João Mário de França que sempre esteve disposto a me amparar e ao Sr. José Raimundo Carvalho por acreditar que eu poderia participar de um doutorado como o CAEN e fazer uma tese em economia. Destaco ainda os professores: Maurício Benegas, Ricardo Antônio de Castro Pereira, Paulo de Melo Jorge Neto, Roberto Tatiwa, Ivan Castelar. Com eles aprendi um pouco da economia e tenho certeza que quero aprender mais.

Aos meus orientadores, em destaque a José Ramalho Torres que me iniciou na pesquisa em tecnologia de concreto nos idos de 1996, ao professor Carlos Felipe Granjeiro Loureiro, grande doutor que me fez questionar o que eu achava inquestionável e me fez crer que eu poderia fazer sempre melhor. Ao Sr. Carlos Henrique Rocha, por ter me iniciado nesta lancinante ciência chamada economia, antes do Senhor imaginava que economia era apenas os livros de história mal contada sobre como se gasta e ganha dinheiro.

Por fim, ao orientador dessa tese Sr. Sérgio Aquino DeSouza por acreditar em mim em cada instante e mesmo não orientando diretamente em alguns momentos, fazendo com que eu refletisse intensamente sobre o problema que abracei. Aos co-orientadores de Lille-França, Sr. Hubert Jayet e Hakim Hammadou que me puseram a pensar na teoria do valor do tempo e como estruturar meus modelos econométricos.

E a parte mais arrepiadora dos agradecimentos destaca à minha esposa (Liana), PENSE NA MULHER PACIENTE E AMÁVEL! Sem você os dias seriam menos felizes e a vida seria imensamente mais dura, com você posso garantir que não é só necessário saber, mas também amar e amar muito, mas muito mesmo.

RESUMO

Economistas desde 1965 debruçam-se sobre uma forma de como mensurar o valor do tempo. O trabalho de Tuong e Hensher (1985) uniu teoria ao empirismo, entretanto, com o passar dos anos os métodos diversos de estimação por modelos de escolha discreta não foram utilizados para a medição do valor de tempo de viagem. Esta tese mensura o valor de tempo de viagem dos usuários de transporte intermunicipal no Ceará utilizando de quatro modelos: não estruturado tipo McFadden e os outros três, conforme proposto em Tuong e Hensher (1985a), para Becker, DeSerpa e a proposta de Tuong e Hensher via expansão de Taylor. Adicional ao modelo *logit* tradicional, faz-se estimações das mesmas formas funcionais utilizando da metodologia desenvolvida por Morikawa (1989) e depois a de estimações de máxima verossimilhança simulada e bayesiana para modelo *logit* com fatores aleatórios apresentada em Train (2003). O resultado indica que se devem utilizar os métodos de estimação e formas funcionais distintas com parcimônia, pois dependendo dos parâmetros obtidos, podem ser encontrados valores para tempo de viagem consideravelmente diferente. Por outro lado, as formas funcionais podem ser boas para previsão tal como discutido acima. Uma escolha entre métodos de estimação sugere o enriquecimento como o que menos distorce as estimativas do *logit* padrão. Uma escolha de forma funcional indica o uso do modelo de Becker ou do DeSerpa por conta da sua pequena oscilação independentemente dos métodos de estimação utilizados. Já uma escolha combinada entre método de estimação e forma funcional sugere o modelo de Becker e DeSerpa com Máxima Verossimilhança Simulada, tanto para obter valores de tempo de viagem como para prever eventos.

Palavras-Chave: Modelos de Escolha Discreta; Valor do Tempo de Viagem; Transporte de Passageiros.

ABSTRACT

Economists since 1965 search a way to measure the value of time. This paper addresses empirical results from a theoretical background proposed in the particular case of transport. The Tuong and Hensher (1985) paper unified the theory and empirical measurement however, over the years various methods for estimation of discrete choice models were not used to measure the values of travel time. This thesis measures the value of travel time of the intercity transport users in Ceará using four models: unstructured type McFadden and the other three, as proposed in Tuong and Hensher (1985a) (Becker, and DeSerpa and Tuong Hensher via Taylor expansion). Additional to the traditional logit model, it estimates the same models using the methodology developed by Morikawa (1989) and simulated maximum likelihood estimation and Bayesian to *logit* models like in Train (2003). The result indicates that it should use the estimation methods and different functional forms sparingly, because depending on the parameters obtained can be found values for travel time considerably different. On the other hand, the functional forms can be good for prediction. A choice of estimation methods such as enrichment suggests less distortion from standard model. A choice of functional form indicates the use the model of Becker or DeSerpa because of their small oscillation regardless of the estimation methods used. Already a combined choice of estimation method and functional form suggests Becker or DeSerpa with Simulated Maximum Likelihood estimation to obtain values of travel time and to forecasts.

Keywords: Discrete Choice Models; Travel Time Values; Passenger Transportation.

SUMÁRIO

APRESENTAÇÃO	1
CAPÍTULO 2. – TEORIA DO VALOR DO TEMPO	4
2.1. INTRODUÇÃO	4
2.2. MODELO TÉORICO DE VALOR DO TEMPO	6
2.2.1. <i>Modelo DeSerpa</i>	6
<i>Considerações sobre a demanda no modelo de DeSerpa</i>	9
<i>Modelos sucessores ao de DeSerpa e o empirismo</i>	10
<i>Considerações a cerca dos aspectos inerentes a teoria de Valor de Tempo e a atividade transportes</i>	13
2.3. MODELO EMPÍRICO PARA MENSURAR VALOR DO TEMPO	15
2.2.1. <i>Preâmbulo histórico da medição do valor de tempo</i>	15
2.2.2. <i>Modelo Empírico de Tuong e Hensher</i>	16
CONSIDERAÇÕES FINAIS	17
CAPÍTULO 3. – MODELOS LOGIT E MÉTODOS DE ESTIMAÇÃO COM ENRIQUECIMENTO E COM FATORES ALEATÓRIOS.....	19
3.1. INTRODUÇÃO	19
3.2. MODELO LOGIT: HISTÓRICO E TEORIA.....	20
3.2.1. <i>Introdução</i>	20
3.2.2. <i>Formalização do Modelo Logit</i>	21
3.2.3. <i>Problemática do Modelo Logit</i>	24
3.3. FONTES DE DADOS PARA MODELOS DE ESCOLHA DISCRETA.....	26
3.4. ESTRATÉGIAS DE ESTIMAÇÃO CONJUNTA COM FONTES DE DADOS DISTINTAS	28
3.5. MODELOS COM PARÂMETROS ALEATÓRIOS.....	33
3.5.1. <i>Estimação por Máxima Verossimilhança Simulada</i>	34
3.5.2. <i>Princípios da Estimação Bayesiana</i>	35
CONSIDERAÇÕES FINAIS	37
CAPÍTULO 4. –DESCRIÇÃO DA BASE DE DADOS DO TRANSPORTE DE PASSAGEIROS DO CEARÁ.....	39
4.1. INTRODUÇÃO.....	39
4.2. FONTES DE DADOS	41
4.2.1. <i>Generalidades dos dados de Preferência Revelada e Declarada</i>	41
4.2.2. <i>Considerações sobre a heterogeneidade na amostra da pesquisa declarada</i>	45
CONSIDERAÇÕES FINAIS	47
CAPÍTULO 5. – METODOLOGIA E RESULTADOS.....	48
5.1. PROCEDIMENTO METODOLÓGICO ADOTADO	48
5.2. DESCRIÇÃO DO MODELO E RESULTADOS DAS ESTIMAÇÕES	48
5.2.1. <i>Estratégias de Estimação</i>	48
<i>Considerações sobre as formas funcionais</i>	48
<i>Considerações sobre as utilidades e a estimação com modelo logit simples</i>	49
<i>Considerações sobre as utilidades e a estimação com junção de fontes de dados</i>	50
<i>Considerações sobre a utilidade e a estimação por logit com fatores aleatórios</i>	51

5.2.2. Resultados da estimação do logit simples	51
5.2.3. Resultados da Estimação do logit com enriquecimento de dados	53
5.2.4. Resultados da estimação do logit com parâmetros aleatórios.....	56
5.3. MEDIÇÃO DOS VALORES DE TEMPO DE VIAGEM.....	58
5.3.1. Valor de Tempo de Viagem pela estimação do logit simples.....	59
5.3.2. Valor de Tempo de Viagem pela estimação do logit com enriquecimento	60
5.3.3. Valor de Tempo de Viagem pela estimação do logit com parâmetros aleatórios	61
5.4. COMPARAÇÃO ENTRE OS MODELOS.....	62
CONSIDERAÇÕES FINAIS	64
CAPÍTULO 6. – CONCLUSÕES	65
APÊNDICE	68
REFERÊNCIAS BIBLIOGRÁFICAS	73

LISTA DE TABELAS

Tabela 2.1: Evolução dos modelos aplicados a mensurar o valor de tempo de viagem, ampliado de Jara-Díaz (2001).....	11
Tabela 4.1: Evolução das viagens no STRIP-CE (Sistema de Transporte Intermunicipal de Passageiros do Ceará) no período de 1997 a 2004.	40
Tabela 4.2: Níveis dos atributos de Tempo de Viagem.....	43
Tabela 4.3: Níveis dos atributos Custo da Tarifa	43
Tabela 4.5: Residência dos respondentes por renda.	47
Tabela 5.1: Modelos estimados. Os valores abaixo do parâmetro são os desvios padrões e os dois asteriscos apontam o parâmetro significativa a 5% de nível de significância..	52
Tabela 5.2: Comparação de modelos simples e com enriquecimento dos dados os valores abaixo do parâmetro representam os desvios padrões e os asteriscos indicam parâmetros significantes a 5% de nível de significância	54
Tabela 5.3: Resultados do modelo com os dados de preferência revelada os asteriscos indicam parâmetros significantes a 5% de nível de significância.	55
Tabela 5.4: Comparação de modelos com parâmetros aleatórios e estimados com diferentes procedimentos. Os valores abaixo do parâmetro representam os desvios padrões e os dois asteriscos apontam o parâmetro significativa a 5% de nível de significância.	57
Tabela 5.5: Valor de Tempo de viagem para os diferentes modelos em reais por hora.....	59
Tabela 5.6: Valor de Tempo de viagem para os diferentes níveis de renda no modelo McFadden em reais por hora.	60
Tabela 5.7: Valor de Tempo para os diferentes modelos em reais por hora.	60
Tabela 5.8: Valor de Tempo para os diferentes modelos em reais por hora.	61
Tabela 5.9: Comparação entre diferentes métodos de estimação para diferentes modelos do Valor de Tempo de viagem.	63

LISTA DE FIGURAS

Figura 4.1: Evolução da participação do transporte clandestino nas viagens radiais 1997 – 2004 (Fonte: PDOTIP-CE 2004).....	41
Figura 4.2: Exemplo de cartão de PD. •	43
Figura 4.3 –Municípios que mais produzem e atraem viagens no Ceará. (Fonte: PDOTIP-CE, 2004).....	46

LISTA DE SIGLAS E ABREVIATURAS

PDOTIP	Plano Diretor e Operacional do Transporte Intermunicipal de Passageiros do Estado do Ceará - 2005
PD	Pesquisa de Preferência Declarada
PR	Pesquisa de Preferência Revelada
STRIP-CE	Sistema de Transporte Intermunicipal de Passageiros do Ceará

APRESENTAÇÃO

A teoria do valor de tempo teve origem no trabalho de Becker (1965) indicando, inclusive, à aplicabilidade da teoria a medição de valores de tempo em transportes. Posteriormente outros autores (DeSerpa, 1972; Evans, 1972; McDonald, 1983; Jara-Diaz, 2001; entre outros) desenvolveram diferentes abordagens resultando em diferentes formas funcionais para entendimento do fenômeno, norteados pela percepção distinta do valor de uma atividade e dos elementos pertencentes à utilidade dos usuários de transportes.

A parte a abordagem teórica, a medição estruturada do valor de tempo só ocorreu empiricamente com o trabalho de Tuong e Hensher (1985), em destaque para a medição do valor de tempo de viagem. No trabalho de Tuong e Hensher (1985) é indicado que a estimação de valor de tempo por Mínimos Quadrados de formas funcionais log-linear ou log geram estimativas falaciosa (McGillivray, 1970; e DeSerpa, 1972) e Train e McFadden (1978) já haviam proposto a estimação por modelos de escolha discreta tipo *logit* para utilidades de transportes. Saliente-se que o trabalho de Train e McFadden (1978) desconsidera a existência do problema de IIA - Independência das Alternativas Irrelevantes¹ e que este problema foi solucionado ou mitigado com o desenvolvimento de modelos alternativos, tais como: Nested Logit, Mixed Logit apresentados em Train (2003).

Por outro lado, complementares ao problema de IIA, os dados utilizados para estimação possuem problemas decorrentes de respostas tendenciosas, tal como apresentadas em McFadden (2000a) e Morikawa (1989), e que podem resultar em estimações viesados. Nesta linha, muitos trabalhos foram desenvolvidos na intenção de reduzir o viés de estimação, proveniente dos problemas inerentes as fontes de dados, destacando os trabalhos de Morikawa

¹ IIA da sigla em inglês de Independency of Irrelevant Alternative.

(1989), Ben-Akiva e Morikawa (1990), Morikawa, Ben-Akiva, e Yamada (1991), Hensher e Bradley (1993), Hensher, Louviere, e Swait (1999) e Brownstone e Train (1999).

Com objetivo de contribuir para o entendimento dos ganhos de estimação de valor de tempo de viagem com diferentes formas de estimação, esta tese objetiva estudar formas funcionais e de estimação de valor de tempo de viagem por modelos *logit*, concluindo com a prospecção do modelo mais aplicável para mensurar valor de tempo de viagem independentemente do método de estimação.

De forma complementar na tese far-se-á um estudo pormenorizado da teoria do valor de tempo proposta por Becker (1965) e das diferentes formas funcionais subsequentemente propostas por estudiosos do tema e aplicáveis a mensuração do valor do tempo de viagem. Neste relato da teoria do valor de tempo serão feitas considerações sobre a proposta estruturada de Tuong e Hensher (1985) para mensurar valor de tempo de viagem e que norteará o processo de estimação. Paralelamente, será conduzida uma descrição da aplicabilidade da teoria a atividade transportes.

Far-se-á, ainda, uma prospecção de diferentes métodos de estimação de modelos *logit* incorporando heterogeneidade ou enriquecendo a estimação com diferentes fontes de dados, aplicando tais métodos as propostas funcionais de Tuong e Hensher (1985). Far-se-á, também, uma comparação dos modelos estimados utilizando a previsibilidade e os resultados das estimações para apontar o melhor modelo a ser utilizado na mensuração do valor do tempo de viagem.

Concomitante, motivado pelo fato do transporte clandestino (irregular) ter se intensificado do final da década de 1990 ao início de 2000 no Brasil e retratado em parte no estudo de Golub *et.al.* (2009), aplicado no contexto interurbano da região metropolitana do Rio de Janeiro,

assim como estudos desenvolvidos para o PDOTIP-CE (2005), far-se-á o estudo empírico com os dados de escolha entre o modo regular e o clandestino do estado do Ceará.

Em suma, esta tese irá mensurar o valor de tempo de viagem dos usuários de transporte intermunicipal no Ceará utilizando-se de quatro modelos: não estruturado tipo McFadden e os outros três, conforme proposto em Tuong e Hensher (1985a), para Becker, DeSerpa e a proposta de Tuong e Hensher com expansão de Taylor. Adicional ao modelo *logit* tradicional, far-se-á estimações dos mesmos modelos utilizando da metodologia desenvolvida por Morikawa (1989) e o modelo com fatores aleatórios proposto em Train (1998) com estimações de máxima verossimilhança simulada e bayesiana. Para consecução do objetivo, o trabalho será dividido em mais cinco capítulos que abordaram os aspectos teóricos da medição do valor do tempo, abordagem empírica, a problemática do transporte clandestino no estado do Ceará, a metodologia adotada no trabalho, seguida de resultados e conclusões.

CAPÍTULO 2. – TEORIA DO VALOR DO TEMPO

2.1. INTRODUÇÃO

O problema central na economia é alocar recursos escassos, entretanto, os modelos que tentavam mimetizar a decisão dos usuários negligenciavam o recurso tempo. Os economistas passaram a tentar entender como o recurso tempo era utilizado para a decisão e como insumo produtivo em meados do século passado. O modelo de alocação de tempo pensado por Becker (1965) é a primeira tentativa de medir qual seria o valor de tempo para as pessoas e como elas utilizavam este tempo em atividades. No modelo o valor do tempo é igual ao valor pago por unidade de tempo trabalhada.

Tomando como base Becker (1965), várias críticas surgiram a este modelo e que enriqueceram o debate acadêmico. Johnson (1966), por entender que tempo trabalhado pode gerar prazer, criticou os resultados de Becker (1965) sugerindo que o tempo deveria compor a função de utilidade. Adicionalmente Oort (1969) postulou que tempos de viagens também compunham variáveis de escolha do indivíduo, surgindo à concepção que tempo de viagem é uma componente importante na utilidade do indivíduo. Então, seguindo as críticas acima destacadas, o modelo de Becker (1965) desconsiderava que as atividades geravam utilidade ou desutilidade, destacando entre elas a viagem. Esta postura foi sanada em 1971 com o trabalho de DeSerpa.

DeSerpa (1971) inseriu tempo na função utilidade de Becker e restrições que relacionam os bens produzidos aos tempos das atividades, portanto, agregando as críticas de Johnson (1966) e de Oort (1969). O modelo de DeSerpa (1971) resultou em três diferentes conceitos de valor de tempo: valor de tempo como recurso, valor de tempo alocado a uma atividade e valor de tempo ganho em uma atividade. De forma alternativa, Evans (1972) adota a função de utilidade

dependente apenas dos tempos despendidos nas atividades e que as restrições sejam apenas restrições de tempo. Com sua abordagem ele concluiu que a taxa marginal de utilidade da renda por indivíduo pode ser igual à zero para indivíduos que ganham dinheiro mais rápido que possam gastar.

Duas décadas se passaram em que surgiram poucos estudos teóricos de valor de tempo, entretanto, com uma busca contínua e intensa de mensurar empiricamente o que os modelos teóricos apontavam. De fato, as primeiras tentativas de mensuração empírica foram desenvolvidas em paralelo com o trabalho de Becker (1965), a exemplo de Beesley (1965) e o método de transferência de preço apresentado em Ortúzar e Willumsen (2001). Entretanto, somente após Train e McFadden (1978) é que se pode ter uma ideia empírica dos valores reais do valor do tempo de uma atividade. Eles utilizam da ideia de utilidade aleatória e identificam que mudanças em níveis salariais afetarão a atração por diferentes modos. Por fim, Hensher e Tuong (1985a), adotando de modelos estruturados, verificam o que a teoria já apontava.

Uma vez que ficou a lacuna aberta nos trabalhos teóricos sobre a idiosincrasia das decisões individuais, os esforços teóricos da década de 90 culminaram com trabalhos em que tentavam sanar a falha, resultando em trabalhos tal como Mackie *et. al.* (2001) e Jara-Díaz (2002) e que, apoiados no trabalho de Train e McFadden (1978), apontam um valor subjetivo para o tempo, portanto o tempo não seria apenas fator do salário que as pessoas ganham, mas da pessoa *per se*.

Nesta mesma linha deve-se destacar o trabalho de Blayac e Causse (2001), pois eles demonstram a existência de uma relação entre o valor de tempo de viagem, preço e nível de renda, corroborando, de certo modo, a ideia de transporte coletivo como um bem inferior onde o efeito renda influencia no valor do tempo. Por seu turno, Jiang e Morikawa (2004) estudaram as alternativas de realocação do tempo poupado indicando uma relação dinâmica do valor do tempo poupado com a variação do valor do lazer e do trabalho para cada indivíduo. Os autores

ainda indicam que as substituições de tempos de viagem causam mudanças no padrão de consumo e a mudança de padrões de consumo já havia sido alvo de digressões por parte de DeSerpa (1971) quando alertou sobre a realocação de tempo. Como o modelo de DeSerpa foi o que deu o salto e contribuiu cabalmente para a fundamentação empírica, então, na sequência apresentaremos este em sua formulação matemática.

2.2. MODELO TEÓRICO DE VALOR DO TEMPO

DeSerpa (1971) contribuiu com o entendimento do valor de tempo de uma viagem como importante para economia. Nesta seção é apresentado o seu modelo.

2.2.1. Modelo DeSerpa

O modelo de DeSerpa diferencia de Becker ao incorporar uma relação entre o tempo e os bens consumíveis e produzidos (inadequação no problema de maximização abaixo descrito). Esta inequação indica uma relação cruzada entre as atividades e tempos exercidos em cada atividade.

$$\text{Max } U = u(X_1, \dots, X_n, T_1, \dots, T_n) \quad (2.7)$$

$$\text{s.t. } \sum_{i=1}^n P_i X_i = I_f \rightarrow \lambda \quad (2.8)$$

$$\sum_{i=1}^n T_i = \tau \rightarrow \mu \quad (2.9)$$

$$T_i \geq a_i X_i \rightarrow \kappa_i \quad (2.10)$$

Onde, T_i := Tempo exercido na atividade i ;

P_i := Preço da atividade i ;

X_i := Custo do Bem i ;

I_f := Restrição orçamentaria;

τ : = Restrição de tempo;

a_i : = Ponderador entre o tempo exercido em uma atividade i e o valor do produto obtido na atividade;

k_i : = Utilidade marginal do tempo poupado para a atividade i ;

μ : = Utilidade marginal da renda;

λ : = Utilidade marginal do tempo.

Relaxar a i ésima restrição de consumo de tempo é equivalente a poupar tempo no consumo da i ésima atividade, logo κ_i é interpretado como utilidade marginal do tempo poupado. Soluciona-se o problema de maximização conforme o seguinte lagrangeano:

$$\ell = u(X_1, \dots, X_n, T_1, \dots, T_n) + \lambda(I_f - \sum_{i=1}^n P_i X_i) + \mu(\tau - \sum_{i=1}^n T_i) + \kappa_i(a_i X_i - T_i)$$

$$k_i \geq 0, \quad i = 1, \dots, n \quad \mu, \lambda > 0$$

Exigem-se para a maximização as seguintes condições:

$$\frac{\partial U}{\partial x_i} = \lambda p_i + \kappa_i a_i \quad i = 1, \dots, n \quad (2.11)$$

$$\frac{\partial U}{\partial t_i} = \mu - \kappa_i \quad i = 1, \dots, n \quad (2.12)$$

$$\kappa_i(T_i - a_i X_i) = 0 \quad i = 1, \dots, n \quad (2.13)$$

Onde, $\frac{\mu}{\lambda}$ representa a taxa marginal de substituição do tempo por dinheiro ou valor do tempo

no conceito de Becker (1965); $(\frac{\mu}{\lambda} - \frac{\kappa_i}{\lambda})$ representa a taxa marginal da alocação do tempo; e

$\frac{\kappa_i}{\lambda}$ representa o valor de tempo poupado. Portanto os valores do tempo e do tempo poupado são:

$$\frac{\frac{\partial U}{\partial \tau}}{\lambda} = \frac{\mu}{\lambda}; \quad (2.14)$$

e

$$\frac{\kappa_i}{\lambda} = \frac{\mu}{\lambda} - \frac{\frac{\partial U}{\partial T_i}}{\lambda} \quad (2.15)$$

Observa-se que se $\kappa_i = 0$, então se tem a inversão do tempo de *commodity* para recurso de uma atividade, pois o indivíduo deseja passar mais tempo que o necessário para consumir o bem X_i . Já se o valor de tempo poupado for positivo, então existiria a possibilidade de realocação para outra atividade de maior valor que a do tempo poupado. Além disso, vale notar que o valor do tempo como recurso é decorrente do fato que o tempo é escasso, portanto é um custo de oportunidade de tempo positivo, não fazendo sentido adquirir mais tempo, mas poupar o tempo e realocá-lo.

Frise-se que no modelo de DeSerpa (1971) o lazer constitui o tempo destinado a consumir um bem cujo tempo poupado nele seja zero e todos os outros bens são considerados bens intermediários. Entretanto, a definição não se adequa ao tempo de viagem, pois $(\frac{\mu}{\lambda})$ só seria igual ao valor de tempo de viagem se as jornadas de trabalho fossem um bem de lazer o que não é real, pois o trabalho pode ser prazeroso tal como o lazer, mas não plenamente prazeroso. $(\frac{\mu}{\lambda})$ em Johnson (1966) é o valor do tempo em trânsito e o valor poupado deste tempo em trânsito seria o preço do tempo.

Considerações sobre a demanda no modelo de DeSerpa

A demanda por uma atividade i no modelo de DeSerpa (1971) seria:

$$X_i = X_i(p_1, \dots, p_n, a_1, \dots, a_n, I_f, \tau) \quad i = 1, \dots, n \quad (2.16)$$

DeSerpa (1971) trata o efeito preço e renda para seu modelo conforme a equação (2.17) e (2.18) para os bens de lazer e para os bens intermediários, respectivamente. Levando em consideração o sistema de equações $Hy=c$, onde H é uma matriz negativa semi-definida das derivadas totais das restrições (2.8) a (2.10), y é o vetor de diferenciais dos bens, atividades e utilidades marginais e c é o vetor de diferencial de valores para cada atividade. D_{ij} é a determinante menor de H ao retirar a i -ésima linha e a j -ésima coluna e quando $i=j$ tem-se a equação (2.17). Nota-se que quando a dimensão tempo é relevante, não é aplicável o teorema neoclássico de substituição.

$$\frac{\partial X_i}{\partial p_i} + X_i \frac{\partial X_i}{\partial I_f} = \lambda \frac{D_{ii}}{D} \leq 0 \quad i = 1, \dots, n \quad (2.17)$$

No caso de irrelevância da dimensão tempo, tem-se a equação (2.18), onde a elasticidade de qualquer bem com respeito ao preço é igual em valor absoluto, mas em sinal contrário a elasticidade renda da demanda por aquele bem.

$$\sum_{i=1}^n p_i \left[\frac{\partial X_i}{\partial p_i} + X_i \frac{\partial X_i}{\partial I_f} \right] = \sum_{i=1}^n p_i \frac{\partial X_i}{\partial p_i} + I_f \frac{\partial X_i}{\partial I_f} \quad i = 1, \dots, n \quad (2.18)$$

Além disso, pela proposta de DeSerpa (1971), não há justificativas para desenhar curvas decrescentes de demanda inelástica e medindo o preço do tempo de viagem tem-se a equação (2.19) que é mensurável se a função demanda por viagem for função dos preços da viagem e dos custos da viagem.

$$\frac{\kappa_i}{\lambda} = \frac{\frac{\partial X_i}{\partial a_i} - X_i \left[\frac{\partial(\frac{X_i}{\lambda})}{\partial p_i} + X_i \frac{\partial(\frac{\kappa_i}{\lambda})}{\partial I_f} \right]}{\frac{\partial X_i}{\partial p_i}} \quad (2.19)$$

Se X_i é função do custo do tempo, preço próprio e renda, então $\frac{\partial X_i}{\partial a_i}$ e $\frac{\partial X_i}{\partial p_i}$ são coeficientes de uma regressão simples. Entretanto, funções log-linear, conforme proposta por DeSerpa, ou log-linear com preço cheio, proposta por Granou (1986) não são adequadas à estimação destes parâmetros. A primeira porque exige derivadas segundas do termo entre colchetes da equação 2.19 ser zero ou negligenciáveis, sendo impossível no processo de maximização empírica. Já a segunda forma funcional, falha por gerar um viés de estimação, estimando preços de tempo de viagem constante, estando em desacordo com a teoria.

Modelos sucessores ao de DeSerpa e o empirismo

Os modelos sucessores a proposta de DeSerpa (1971) caracterizam-se por analisar diferentes restrições ao modelo e por incorporar a renda e o ambiente socioeconômico dos indivíduos, destacando os trabalhos de Blayac e Causse (2001) e de Jiang e Morikawa (2004).

Train e McFadden (1978) se assemelham a proposta de DeSerpa (1971), portanto havendo uma adequação á estimação via modelos *logit* aos parâmetros econômicos relatados em DeSerpa, sendo o problema de Train e McFadden (1978) a vertente discreta do modelo de Becker (1965).

Jara-Díaz (2002) adota as premissas de McFadden (1978) e chega à relação subjetiva do valor do tempo, algo, desconhecido por DeSerpa. Blayac e Causse (2001) assumem restrições de tempo com níveis de renda no modelo de DeSerpa e chegam à conclusão que rendas distintas implicam valores distintos para tempo. Para visualização de todos os modelos desenvolvidos

até o momento amplia-se abaixo, na Tabela 2.1., a tabela presente no trabalho de Jara-Díaz (2001).

Tabela 2.1: Evolução dos modelos aplicados a mensurar o valor de tempo de viagem, ampliado de Jara-Díaz (2001).

Autor	Modelo	Valor do Tempo
Becker (1965)	$Max U = U(Z_1, \dots, Z_n)$ $s.t. \quad \sum_{i=1}^n P_i X_i = wW + I_f \quad \rightarrow \quad \lambda$ $\sum_{i=1}^n T_i = \tau - W \quad \rightarrow \quad \mu$ $Z_i = f_i(X_i, T_i)$	$\frac{\mu}{\lambda} = w$
Johnson (1966)	$Max U = U(L, W, G)$ $s.t. \quad G = wW \quad \rightarrow \quad \lambda$ $\tau = L + W \quad \rightarrow \quad \mu$	$\frac{\mu}{\lambda} = w + \frac{\partial U}{\partial W} = \frac{\partial U}{\partial L}$
Oort (1969)	$Max U = U(L, W, t, G)$ $s.t. \quad G + c = wW \quad \rightarrow \quad \lambda$ $\tau = L + W + t \quad \rightarrow \quad \mu$	$\frac{\mu}{\lambda} = \frac{\partial U}{\partial L}$ $-\frac{dU}{dt} = w + \frac{\partial U}{\partial W} - \frac{\partial U}{\partial t}$
DeSerpa (1971)	$Max U = u(X_1, \dots, X_n, T_1, \dots, T_n)$ $s.t. \quad \sum_{i=1}^n P_i X_i = I_f \quad \rightarrow \quad \lambda$ $\sum_{i=1}^n T_i = \tau \quad \rightarrow \quad \mu$ $T_i \geq a_i X_i \quad \rightarrow \quad \kappa_i$	$\frac{\mu}{\lambda} = \frac{\partial U}{\partial L}$ $\frac{\kappa_i}{\lambda} = \frac{\mu}{\lambda} - \frac{\partial U}{\partial T_i}$

Evans (1972)	$\text{Max } U = U(T_1, \dots, T_n)$ $\text{s.t. } \sum_{i=1}^n w_i T_i \geq 0 \rightarrow \lambda$ $\tau - \sum_{i=1}^n T_i = 0 \rightarrow \mu$ $T_i - \sum_{\forall j \neq i} b_{ij} T_j \geq 0 \rightarrow \kappa_i$	$\frac{\kappa_i}{\lambda} = \frac{\mu}{\lambda} - \frac{\partial U}{\partial T_i} - w_i$ $\frac{\mu}{\lambda} = \frac{\partial U}{\partial L} + w_L$
Small (1982)	$\text{Max } U = U(G, L, W, s)$ $\text{s.t. } G + c(s) = I_f + wW \rightarrow \lambda$ $L + t(s) = \tau - W \rightarrow \mu$ $F(s, W; w) = 0 \rightarrow \nu$	$\frac{\mu}{\lambda} = w + \frac{\partial U}{\partial W} - \nu \frac{\partial F}{\partial W}$
Gronau (1986)	$\text{Max } U = U(Z_1, \dots, Z_n, Z_w)$ $\text{s.t. } \sum_{i=1}^n P_i X_i + P_w X_w = I(Z_w) + I_f \rightarrow \lambda$ $\sum_{i=1}^n T_i + W = \tau \rightarrow \mu$ $Z_i = f_i(X_i, T_i)$ $Z_w = f_w(X_w, W)$	$\frac{\mu}{\lambda} = w + \frac{\partial U}{\partial W} - P_w \frac{\partial X_w}{\partial W}$ $Z_w = W$ $I(Z_w) = wW$
MVA consultancy group (1986)	$\text{Max } U = U(x, q, t_1, \dots, t_m)$ $\text{s.t. } px + \sum_{i=1}^n d_i p_i = R \rightarrow \lambda$ $q + \sum_{i=1}^n d_i t_i = T \rightarrow \mu$ $t_i \geq \bar{t}_i$	$\frac{\kappa_i}{\lambda} = \frac{\partial V_i}{\partial p_i}$
Jiang e Morikawa (2004)	$\text{Max } U = U(p_1, t_1, t_i)$ $\text{s.t. } p_1 + p_i(t_i) = wt_w \rightarrow \lambda$ $t_1 + t_w + t_i = T \rightarrow \mu$ $t_i \geq \bar{t}_i$	$\frac{\mu}{\lambda} = \frac{1}{\lambda} \frac{\partial u}{\partial t_1}$ $\frac{\kappa_i}{\lambda} = \frac{\mu}{\lambda} - \frac{1}{\lambda} \frac{\partial u}{\partial t_i} + p_i'$

Tuong e Hensher (1985a e b) interpretaram os resultados de DeSerpa e Becker e apresentaram a abordagem empírica que será utilizada nesta tese. A mensuração dos efeitos empiricamente é feito no trabalho de Tuong e Hensher (1985b) por meio de um experimento e de uma estimação com modelo *logit* para uma utilidade indireta e assumindo que $k_i = f(t_i, p_i)$, ou seja, que existe uma relação entre tempo e custo. Observa-se a desagregação na estimação uma vez que não se estima demandas diretas, mas utilidades indiretas.

Por fim, Hammadou e Jayet (2002) utilizam a ideia presente em Tuong e Hensher (1985b) e o modelo proposto por Blayac e Causse (2001) para mensurar empiricamente o valor de tempo para usuários de transportes franceses.

Considerações a cerca dos aspectos inerentes a teoria de Valor de Tempo e a atividade transportes

Oort (1969) postulou, criticando o trabalho de Becker (1965), que tempos de viagens também compunham variáveis na utilidade do indivíduo, entretanto, Becker (1965) não negligenciava esta possibilidade, mas diferentemente de Oort (1969), assumia que a utilidade para uma viagem era diferente da utilidade que outra atividade gerava.

Com vistas a aceitar a possibilidade de Oort (1969), DeSerpa (1971) adotou uma utilidade genérica para os agentes econômicos, onde todos os tempos exercidos em diferentes atividades eram fatores de decisão. Isso tornava complexo o tratamento e mensuração de qualquer utilidade, uma vez que fazia necessário o tempo de todas as atividades e todos os bens gerados por atividade para estimar uma função utilidade direta ou indireta. Tratava-se, portanto não de mensurar a utilidade para um agente econômico decorrente de uma viagem, mas a utilidade do agente viver.

De toda sorte, retomando o pensamento de Becker (1965), poderia ser adotado a hipótese de separabilidade da utilidade de DeSerpa (1971). Portanto, a proposta de DeSerpa (1971) seria um somatório de utilidades onde um subproblema de maximização a condição de maximização de uma utilidade da tarefa n poderia ser feita sem perda de generalidade, dado a linearidade do somatório. Isso era condição necessária para o processo de estimação de utilidades de viagens tal como proposto em Oort (1969).

Dada a possibilidade de separação da utilidade, a abordagem de Tuong e Hensher (1985) e posteriores estimações de funções utilidade de viagem para mensurações de valores de tempo de viagem a partir de dados de viagens tal como são pesquisados em transportes foi possível. Portanto, assumia-se que a utilidade de uma viagem é diferente da utilidade de trabalhar que por sua vez é diferente da utilidade de praticar um esporte e assim sucessivamente para n atividades exercidas pelos agentes econômicos.

A abordagem dada pela MVA Consultancy Group (1986) é a mais aplicada para mensuração de valor de tempo de viagem, mas converge para o conceito de valor de tempo poupado de uma atividade proveniente da abordagem de DeSerpa (1971). Abrantes e Wardman (2011) fazem uma excelente descrição dos trabalhos de mensuração de valor de tempo de viagem. No Brasil os trabalhos são escassos e o trabalho mais significativo foi o de Arruda (1996). No geral, o valor de tempo de viagem estimado em trabalhos de transportes é aplicado para realizar análises custo-benefício da instalação de uma infraestrutura de transportes, considerando ganhos de operação, mas raramente utilizado para mensurar competição entre modos.

Para completar o estudo da teoria do valor de tempo e da mensuração do valor do tempo de viagem faz-se uma descrição da abordagem empírica no item que segue.

2.3. MODELO EMPÍRICO PARA MENSURAR VALOR DO TEMPO

Nesta seção apresentam-se a abordagem estruturada para mensurar do valor de tempo e as formas estruturais que serão utilizadas no trabalho.

2.2.1. Preâmbulo histórico da medição do valor de tempo.

DeSerpa (1971) deu um salto quantitativo na abordagem teórica identificando diferentes componentes do valor de tempo, entre eles o de ganho em tempo em uma atividade. Com esta abordagem, teóricos de transportes passaram a analisar os ganhos em tempo de viagem empiricamente, notadamente, Train e McFadden (1978), Tuong e Hensher (1985a e b), Bates (1987), Jara-Díaz (2000) e Hammadou e Jayet (2002). Anteriormente, Beesley (1965) desenvolveu o gráfico de Beesley que tinha a mesma intenção de mensurar empiricamente o valor do tempo.

Train e McFadden (1978) utilizam uma estrutura de utilidade tipo Cobb-Douglas com variáveis em bens e lazer e encontram que o tempo de viagem depende dos ganhos dos usuários. Os autores fazem estimativas da utilidade indireta obtida do problema de maximização e encontram valores do ponderador de lazer entre 0.7 e 1 do valor do salário dos usuários de transporte.

Tuong e Hensher (1985a) apontam as diferenças entre as abordagens de Becker (1965) e de DeSerpa (1972) e fazem a ligação do modelo teórico ao empírico pela utilidade indireta. Eles ressaltam que ganhos com tempo de viagem tem diferentes interpretações no modelo de DeSerpa (1971) dependendo do modo de transporte escolhido e da qualidade do transporte.

Tuong e Hensher (1985a) interpretam os resultados em um modelo de escolha discreta via modelos de alocação de tempo. Entretanto Bates (1987) critica o trabalho deles por apresentar interpretações distorcidas da teoria para o modelo empírico principalmente na concepção de

realocação do tempo poupado para o lazer. De qualquer forma as abordagens de Tuong e Hensher (1985a) complementadas pela crítica de Bates (1987) serviram como arcabouço para todos os trabalhos empíricos estruturados de valor de tempo. Neste escopo será apresentada a abordagem empírica de Tuong e Hensher (1985a).

2.2.2. Modelo Empírico de Tuong e Hensher

Até o trabalho de Tuong e Hensher (1985b) não foi feita ligação estreita entre os modelos teóricos e a sua estimação. Tuong e Hensher (1985b) propõem a estrutura da função utilidade indireta conforme a equação 2.20 para o modelo de Becker (1965) e 2.21 para o de DeSerpa (1971).

$$V_i = -\lambda x_i - \mu t_i \quad (2.20)$$

$$V_i = -\lambda x_i - (\mu - k_i)t_i \quad (2.21)$$

Onde: V_i := nível da utilidade indireta associada à escolha da viagem i ;

x_i, t_i := são o custo e o tempo associada à viagem i ;

λ, μ := parâmetros a serem estimados que retratem utilidades marginais da renda e do tempo respectivamente;

k_i := multiplicador de lagrange associado ao consumo de tempo restrito a tecnologia e que retrata a utilidade marginal do tempo poupado.

A relação $\frac{\mu}{\lambda}$ obtido da equação 2.20 representa o preço sombra associado ao tempo da viagem.

No segundo modelo, o ganho em tempo, específico de cada alternativa i , é representado pela

relação $\frac{k_i}{\lambda}$. Alternativamente, dada à impossibilidade de identificação de μ em 2.21, Tuong e

Hensher (1985b) assumem que $k_i = f(t_i, p_i)$ e fazendo uma aproximação por expansão de série de Taylor na equação 2.21, chega-se à expressão 2.22 para a utilidade indireta.

$$V_i = -\lambda x_i - (\mu - \bar{k})t_i + (\alpha + \beta t_i^2 + \gamma x_i t_i) \quad (2.22)$$

$$\text{Com } \alpha = -\left(\frac{\partial k}{\partial t}\right)_i \bar{t} - \left(\frac{\partial k}{\partial x}\right)_i \bar{x}; \quad \beta = \left(\frac{\partial k}{\partial t}\right)_i; \quad \gamma = \left(\frac{\partial k}{\partial x}\right)_i$$

Na abordagem de Tuong e Hensher (1985b) o valor do tempo é calculado da seguinte forma:

$$VOT = \left(\begin{array}{c} \frac{\partial V}{\partial t_i} \\ \frac{\partial V}{\partial x_i} \end{array} \right) \Bigg|_{V_i = const}$$

$$VOT = \frac{(\mu - \bar{k})}{\lambda} - \frac{(\gamma x_i + 2\beta t_i)}{\lambda}$$

Com esta abordagem, pode-se fazer uma distinção entre diferentes viagens, tempo andando, tempo parado, etc. e especificar o valor de tempo para cada uma das subatividades.

CONSIDERAÇÕES FINAIS

O tempo tem seu valor tendo relação direta com os aspectos inerentes a atividade executada e o nível de renda dos indivíduos. Os aspectos teóricos da tentativa de mensurar o insumo tempo surgiram em 1965 e continuam intrigando os teóricos até a atualidade, existindo problemas a serem sanados, principalmente do ponto de vista empírico, tais como: como as características das pessoas influenciam no valor do tempo? Níveis diferentes de renda influenciam no valor de tempo e como? O local de habitação afeta na percepção de quanto vale o tempo?, entre outros. A aplicação de métodos de medição mais fundamentada surgiu, somente, após a década de 1980, entretanto, sem considerar aspectos idiossincráticos dos indivíduos e apresentando

apenas fatos estilizados para as perguntas em aberto na teoria. Os arcabouços teóricos e empíricos acima relatados servirão de base para toda a tese dando não apenas o ferramental, mas a compreensão de como, para que e o que modelar no problema que será tratado no capítulo 4.

CAPÍTULO 3. – MODELOS LOGIT E MÉTODOS DE ESTIMAÇÃO COM ENRIQUECIMENTO E COM FATORES ALEATÓRIOS.

3.1. INTRODUÇÃO

A estimação de parâmetros de um modelo econométrico recai no parâmetro estimado representar bem o parâmetro populacional. Uma vez estimado, enseja-se que os estimadores tenham no mínimo três propriedades: ausência de viés, eficiência e consistência. A ausência de viés implica que o valor esperado do estimador é igual ao parâmetro populacional. A eficiência implica na variância do parâmetro ser a menor possível, ou seja, ser tão próximo quanto possível do limite de Cramer-Rao. Por fim, a consistência é uma propriedade assintótica dos estimadores e que implica em uma convergência em probabilidade do estimador para o parâmetro populacional.

Infelizmente, as propriedades somente serão atendidas se os dados satisfizerem outros pré-requisitos destacando erros de medidas mínimos que quase nunca são obtidos, pois estes podem ser inerentes ao processo de aquisição dos dados. Por exemplo, ao se questionar um indivíduo sobre a sua escolha entre uma ponte que não foi construída e outra que já está lá, o respondente pode ser tendencioso, dependendo da forma como a questão foi imposta ou pelo fato dele imaginar que pode ou não interferir na decisão da construção da nova ponte. Por outro lado, se perguntarmos um indivíduo sua percepção de meios de transportes ele pode viesar sua resposta por desconhecer as alternativas que poderiam existir.

Felizmente, estes problemas podem ser solucionados melhorando as estimativas, atendendo as propriedades desejadas dos estimadores total ou parcialmente. Logo, utilizar todos os dados com parcimônia é válido, mesmo com os problemas indesejados, notoriamente problemas de medida das variáveis, pois eles podem melhorar a mensuração dos estimadores e também conter informações relevantes referentes ao processo de decisão dos indivíduos.

Neste sentido, este capítulo apresenta formas de mensurar utilidade indireta de transportes conforme descrito no capítulo anterior, utilizando uma combinação de duas fontes de dados ou incorporando as idiosincrasias dos indivíduos por meio de uma estimação com fatores aleatórios. Posteriormente, calculam-se os valores de tempo para os modelos teóricos relatados no capítulo 2 e de acordo com a teoria do valor de tempo. Para tanto, dividiu-se o capítulo em mais cinco seções que apresentam as formas de estimar modelos de escolha discreta tipo *logit* e seus problemas, a solução de alguns dos problemas via modelagem com fontes de dados distintas e a redução do problema de IIA por meio de modelos *logit* com parâmetros aleatórios.

3.2. MODELO LOGIT: HISTÓRICO E TEORIA.

Nesta seção será apresentado o histórico dos modelos de maximização de utilidade aleatória, formalizado o modelo *logit* binário e apresentado suas limitações.

3.2.1. Introdução.

Os modelos de escolha discreta passaram por uma evolução partindo do conceito de utilidade aleatória presente no trabalho de Thurnstone (1927), tendo como ponto ápice o axioma de Luce (1963). Os modelos atuais descrevem de forma mais realista as diferenças entre as demandas ou utilidades de indivíduos que demanda produtos ou serviços.

McFadden (2000b) indica a vantagem do método por poder ser associado à teoria econômica, principalmente na adequação da abordagem de Lancaster (1966) reduzindo a dimensionalidade no tratamento de utilidades, não mais por produtos, mas por atributos.

Os modelos de escolha discreta tipo *logit* passaram por uma evolução do processo de estimação as formas de aquisição dos dados. Para o primeiro caso o principal foco era relaxar a hipótese de distribuição IID e conseqüente redução do problema de IIA- Independência das Alternativas Irrelevantes. Isso propiciou modelos que podem descrever comportamentos mais heterogêneos entre indivíduos ou entre grupos de indivíduos, com destaque aos modelos: Multinomial Logit, Nested Logit e Mixed Logit (Train, 2003). Na atualidade os modelos de escolha discreta tipo *logit* espacial se propõem a descrever heterogeneidade entre indivíduos de regiões distintas tal como o Mixed Logit, mas associando a estrutura espacial ao modelo (Klier e McMillen, 2006; Bhat e Guo, 2002; LeSage, 1999; e Pinkse e Slade, 1998)

As formas de aquisição de dados e a ausência de dados confiáveis ou em quantidade desejada para uma convergência dos estimadores incentivou muitos trabalhos em pesquisas de demanda por viagem com o uso de dados desagregados tais como: Morikawa (1989), Ben-Akiva e Morikawa (1990), Morikawa, Ben-Akiva, e Yamada (1991), Hensher e Bradley (1993), Hensher, Louviere, e Swait (1999) e Brownstone e Train (1999). Como dito, todos os modelos surgiram do modelo *logit*. Este será apresentado formalmente a seguir.

3.2.2. Formalização do Modelo Logit

A função de utilidade que representa o grau de preferência de uma alternativa para certo indivíduo é definida em função dos valores dos atributos das alternativas e das características socioeconômicas do indivíduo, tal como na equação 3.1.

$$U_{in} = U(z_{in}, S_n) \quad (3.1)$$

em que: U_{in} := Utilidade de uma alternativa i para um indivíduo n ;

$z_{in} :=$ Vetor dos atributos relevantes da alternativa i ;

$S_n :=$ Vetor das características socioeconômicas do indivíduo n ;

Segundo Ben-Akiva e Lerman (1985), dado que a utilidade não pode ser prevista com total certeza, então se trata a utilidade como uma variável aleatória, formada por uma componente determinística e outra aleatória que reflete as “irrationalidades” da escolha do indivíduo. Esta abordagem tem como base trabalhos de psicologia, inicialmente propostos por Thurstone em 1927 sobre utilidades aleatórias, onde se presume que os indivíduos não escolhem com 100% de certeza as coisas, mas com uma idiossincrasia presente no termo erro. Dessa forma, a utilidade de uma alternativa i para um indivíduo n (U_{in}) pode ser representada pela seguinte expressão:

$$U_{in} = V_{in} + \varepsilon_{in} \quad (3.2)$$

Com: $U_{in} :=$ Utilidade global de uma alternativa i para um indivíduo n ;

$V_{in} :=$ Componente sistemática da utilidade de uma alternativa i para um indivíduo n ;

$\varepsilon_{in} :=$ Componente aleatória da utilidade de uma alternativa i para um indivíduo n .

A escolha de uma alternativa i em relação à outra j se dá pela comparação entre duas utilidades tal que $U_{in} > U_{jn}, \forall i \neq j$. Como a utilidade possui um componente determinístico e outro aleatório, a escolha do indivíduo se dará ao satisfazer a inequação descrita abaixo.

$$\begin{aligned} \Pr(y = j \mid \mathbf{z}, \mathbf{x}) &= \Pr[U_{in} > U_{jn} \forall i \neq j \mid \mathbf{z}, \mathbf{x}] \\ &= \Pr[V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn} \forall i \neq j \mid \mathbf{z}, \mathbf{x}] \\ &= \Pr[\varepsilon_{in} - \varepsilon_{jn} > V_{jn} - V_{in} \forall i \neq j \mid \mathbf{z}, \mathbf{x}] \\ &= \Pr[\tilde{\varepsilon} > -\tilde{V} \forall i \neq j \mid \mathbf{z}, \mathbf{x}] \end{aligned}$$

Luce (1959) demonstrou, assumindo a parte determinística como uma forma linear descrita na equação 3.3 e a distribuição de probabilidade dos termos aleatórios valor extremo do tipo II

para ε_{ij} ($f(\varepsilon_{ij}) = \exp(-\exp(\varepsilon_{ij}))$), que a probabilidade de escolha tem a forma funcional da equação 3.4.

$$V_{in} = \beta_0 + \beta_1 x_{in1} + \beta_2 x_{in2} + \beta_3 x_{in3} + \dots + \beta_k x_{ink} \quad (3.3)$$

Onde: x_{ink} : Atributo k da alternativa i para o indivíduo n;

β_0 : Constante específica da alternativa e representa o efeito de escolha da alternativa que não está incluída nos efeitos dos atributos;

β_k : Peso relativo do atributo x_{ink} na composição da função utilidade.

$$P_n(i) = \frac{e^{V_{in}}}{\sum_{j \in A(n)} e^{V_{jn}}} \quad (3.4)$$

em que: $P_n(i)$: Probabilidade de escolha de uma alternativa i por um indivíduo n;

$A(n)$: Conjunto de alternativas j disponíveis para o indivíduo n;

V_{in} : Utilidade sistemática de uma alternativa i para um indivíduo n;

V_{jn} : Utilidade sistemática de uma alternativa j para um indivíduo n;

A estimação do modelo na forma binária é feita com a maximização da função 3.5. Considerando uma amostra com N observações e definindo $y_{in} = 1$ (se, na observação n, foi escolhida a alternativa i) ou $y_{in} = 0$ (caso contrário).

$$LL = \sum_{n=1}^N (y_{in} \log(P_{ni})) + (1 - y_{in}) \log(1 - P_{ni}) \quad (3.5)$$

Onde: LL : logaritmo da função de verossimilhança;

P_{ni} : Probabilidade do indivíduo n escolher a alternativa i, expressa em (3.4);

N : tamanho da amostra.

As estatísticas para avaliar a qualidade de estimação dos modelos da família *logit*, conforme Ben-Akiva e Lerman (1985) são as seguintes:

- $L(0)$: valor da função logarítmica de verossimilhança quando todos os parâmetros são zero;
- $L(c)$: valor da função logarítmica de verossimilhança quando somente a constante específica da alternativa é incluída. Isto corresponde ao caso onde a probabilidade de escolha é função apenas da fração de amostra que escolheu a determinada alternativa;
- $L(\beta)$: valor máximo da função logarítmica de verossimilhança;
- $-2(L(0) - L(\beta))$: estatística utilizada para testar a hipótese de que todos os parâmetros são zero; é assintoticamente distribuída como χ^2 com k graus de liberdade, onde k é igual ao número de parâmetros estimados;
- $-2(L(c) - L(\beta))$: estatística utilizada para testar a hipótese nula de que todos os parâmetros são zero; é assintoticamente distribuída como χ^2 com $k - J + 1$ graus de liberdade, onde J é o número de alternativas;
- ρ_{zero}^2 : mede a fração do valor de verossimilhança explicado pelo modelo, definido como

$$\left[1 - \left(\frac{L(\beta)}{L(0)} \right) \right].$$

Os valores de ρ_{zero}^2 dependerão do tipo de modelo a ser construído. Essa medida é mais adequada na comparação de duas especificações desenvolvidas com o mesmo conjunto de dados;

3.2.3. Problemática do Modelo *Logit*

O modelo *logit*, no entanto, apresenta sérios problemas conceituais, que se revelam claramente a partir do cálculo das elasticidades. Na literatura de transportes, este problema foi salientado pelo paradoxo ônibus azul-ônibus vermelho, onde dois substitutos próximos geram resultados improváveis no mercado².

A solução mais simples para esta restrição do modelo *logit* é oferecida pelo modelo *logit* agrupado, onde o pesquisador define a priori os agrupamentos. Outra solução, bem mais sofisticada se refere ao modelo *logit* com coeficientes aleatórios, onde a utilidade marginal pelos atributos características, incluindo preços, varia de consumidor para consumidor.

O modelo *logit* agrupado supõe que a decisão de compra dos consumidores é racionalizada por meio de uma árvore de decisão onde cada nível representa escolhas tomadas dentro de grupos e subgrupos de alternativas. Para o caso de transporte pode se supor que os usuários de transportes tomam decisões sequenciais e antes de decidir se utilizará um ônibus azul ou vermelho, primeiro ele decidirá se utilizará transporte coletivo ou individual.

Uma das principais justificativas para se utilizar um *logit* agrupado é o reconhecimento (pelo menos para a importância de se testar) de que as variâncias (e desvios-padrão) dos componentes do erro aleatório na expressão da utilidade são diferentes entre grupos de alternativas no conjunto de escolha dos consumidores.

A propriedade de Independência de Alternativas Irrelevantes (IIA) presente em modelos *logit* multinomiais convencionais é corrigida parcialmente utilizando *logit* agrupado ou um modelo com coeficientes aleatórios. Entretanto, a propriedade permanece válida para as alternativas existentes dentro dos subgrupos mais restritos, mas geralmente não será entre os diferentes grupos e subgrupos.

² Para verificar o problema ver Wooldridge 2002, PP 501-502.

3.3. FONTES DE DADOS PARA MODELOS DE ESCOLHA DISCRETA

Qualquer que seja a forma de modelar a escolha dos usuários de transporte, a sua concepção é fundamentada, principalmente, no planejamento de coleta de dados de preferência declarada (PD) e/ou preferência revelada (PR). Os dados de PR representam as escolhas já realizadas por usuários, portanto, representa o comportamento real de escolha. Os dados de PD fornecem informações a respeito da escolha do indivíduo diante de um conjunto de alternativas previamente definidas, hipotéticas ou não. Morikawa (1989) discute a estabilidade e a validade dos dados, dividindo-os nos seguintes itens:

1. Questionamentos referentes às decisões em dados de PD e PR. Na confecção dos questionários o pesquisador recai em quatro possibilidades;
 - a. “*Prominence Hypothesis*” onde se considera que o respondente escolhe observando um ou poucos atributos no seu processo de avaliação da escolha, desconhecendo que existe o efeito compensado entre os atributos que pode ser incorporado na escolha;
 - b. “*Policy-Response Bias*” onde o respondente presume que poderá se beneficiar da resposta que der (Dooley (1981) e Suzuki *et. al.* (1986) *apud* Morikawa, 1985);
 - c. Inércia das preferências relativa à situação onde o respondente não consegue compreender como o novo produto influenciaria na sua vida (Nelson (1979) *apud* Morikawa, 1985); e
 - d. “*Justification Bias*” onde o respondente pode querer justificar comportamentos passados na sua resposta.
2. Descrição imperfeita das alternativas:

- a. Omissão de variáveis para simplificação dos questionários;
 - b. Ao associar a situação com uma imagem pode estar perdendo a percepção do respondente frente à situação vivida e o presente na imagem;
3. Omissão de situações restritas:
- a. O respondente pode, consciente ou inconscientemente, ignorar suas restrições pessoais, uma vez que o questionário se aplica a uma hipótese.

Como descrito acima, os resultados podem ser viesados da forma como os questionários foram aplicados. Fora isso, cada uma das fontes de dados possui vantagens e limitações. Os dados de PR possuem sérios aspectos limitantes, podendo-se destacar (Ortúzar e Willuemssem, 1997):

- A existência de altas correlações entre os atributos, que impede a estimação isolada dos efeitos dos atributos;
- A dificuldade de estimar variáveis qualitativas; e
- A não possibilidade de avaliar alternativas que ainda não atuam no mercado.

Dessa forma, as lacunas formadas pelas limitações dos dados de PR podem ser preenchidas pelas vantagens advindas no uso de dados de PD, dentre as quais é possível destacar as seguintes:

- Permitem o controle dos valores dos atributos através de projetos ortogonais, que permitem a estimação dos efeitos de cada atributo isoladamente;
- Permitem a análise de variáveis qualitativas, com o alcance de resultados satisfatórios;
- Permitem a análise de alternativas que ainda não existem no mercado atual.

Por outro lado, a principal limitação dos dados de PD consiste no fato de não refletirem o comportamento atual do mercado, característica peculiar dos dados de PR. Por isso a literatura recomenda (ver Swait, Louviere e Williams (1994) e Hensher, Louviere e Swait (2000)),

sempre que viável a modelagem com dados de PR e de PD conjuntamente, pois assim seria possível ao mesmo tempo unir as vantagens e diminuir as limitações de cada fonte de dados.

Em casos de estudos de previsão de demanda, especificamente, a modelagem conjunta de dados de PD e PR é condição *sine qua non* para a obtenção de cenários consistentes. Em estudos de análise *trade-off*, que permitem a determinação da importância relativa dos atributos, é aconselhável somente a utilização de dados de PD (Swait, Louviere e Williams, 1994 e Hensher, Louviere e Swait, 2000).

Os métodos de estimação com dados conjuntos partem do princípio de explorar, por um lado, as estimações dos efeitos isolados de cada atributo da função de utilidade, obtidas com os dados de PD, e por outro, informações acerca da divisão atual de mercado, obtidas com os dados de PR. Esses métodos também são chamados de “métodos de enriquecimento de dados”. Morikawa (1989), Ben-Akiva e Morikawa (1990), Morikawa, Ben-Akiva e Yamada (1991) propõem uma forma de juntar diferentes amostras (dados de PD e PR). Na seção que segue são apresentadas as estratégias de estimação.

3.4. ESTRATÉGIAS DE ESTIMAÇÃO CONJUNTA COM FONTES DE DADOS DISTINTAS

Existem algumas formas de estimar modelos com a junção de dados, cujas citações são encontradas em Swait, Louviere e Williams (1994) e Hensher, Louviere e Swait (2000), e que tomaram como ponto inicial o trabalho de Morikawa (1989). Para a validade de qualquer método de estimação levantam-se as seguintes hipóteses:

1. O processo de geração do dado das fontes de dados utilizado possui os mesmos parâmetros para os atributos comuns, modificados apenas por um parâmetro de escala diferente para cada fonte de dados;
2. O termo erro das duas amostras tem distribuição de valor extremo II e é IID;

3. Os atributos de “*Trade-offs*” em amostras de PR são deficientes, mas os outros atributos melhor representam aspectos de equilíbrio de mercado, ou seja, inerentes aos indivíduos;
4. Os atributos de equilíbrio de mercado em amostras de PD são deficientes, mas os outros atributos se aplicam a capturar aspectos de “*Trade-offs*”, ou seja, inerentes ao objeto escolhido.

Tomando as hipóteses acima, temos duas possibilidades: utilizar os atributos de “*Trade-off*” como ligação entre as duas amostras, mesmo sabendo da deficiência deste na amostra de PR ou utilizar os atributos de “*Trade-off*” dos dados de PD e os de equilíbrio de PR e estimar em uma função conjunta. Para tanto, tomam-se os parâmetros de escala como relativos e modela-se a partir da hipótese dois, logo abaixo, assumindo G^* uma função *logit*.

Para os dois casos podem-se estimar sequencialmente ou simultaneamente os parâmetros, com a vantagem para o último caso de gerar um estimador com melhor eficiência. Hensher, Louviere e Swait (2000) propõem dois procedimentos para o caso sequencial: uma estimativa com uma procura do fator de escala e outro com um modelo *logit* aninhado. O primeiro caso é resumido nas equações 3.6 e 3.7 cuja matriz de dados tem a estrutura apresentada na equação 3.7.

$$LL = \sum_{n \in PR \cup SP} \sum_{i \in C_n} (y_{in}) \log(P_{ni}(Q(\lambda^{SP}))/\tau) \quad (3.6)$$

$$Q(\lambda^{SP}) = \begin{bmatrix} I^{PR} & 0 & X^{PR} & Z & 0 \\ 0 & \lambda^{PD} I^{PD} & \lambda^{PD} X^{PD} & 0 & \lambda^{PD} W \end{bmatrix} \quad (3.7)$$

Onde I^{PR} e I^{PD} são matrizes identidade; X^{PR} e X^{PD} matrizes de atributos de “*Trade-off*”; Z e W atributos de equilíbrio. A forma funcional para P_{ni}^* dependerá da estrutura do erro, no

caso da hipótese que aponta para valor extremo I, então se tem um modelo *logit*, conforme apresentado no segundo capítulo.

O algoritmo para a estimação segue os seguintes passos:

1. Escolhe-se um espectro de valores de λ^{PD} (lista de λ^{PD});
2. Escolhe-se um valor de λ^{PD} ;
3. Constrói-se a matriz 3.2;
4. Maximiza-se a função de verossimilhança (3.1) e guarda o resultado;
5. Faz-se os passos 1 a 5 repetidamente para a lista de valores de λ^{PD} ;
6. Comparam-se os valores da função de verossimilhança maximizada e escolhe o modelo que apresentou o menor valor.

Para o caso de estimação via estrutura aninhada, define-se uma árvore virtual onde, em um ramo, colocam-se os dados de PR e, no outro, os de PD. Como se supõe que os parâmetros são escalonados, então se restringe o parâmetro de PR como unitário, dessa forma tem-se os parâmetros estimados por uma estrutura aninhada. Alternativamente, Morikawa (1989) propõe sua estrutura sequencial que consiste nos seguintes passos:

1. Estima-se o modelo com os dados de PD e obtém-se $\mu\beta^*$ e $\mu\gamma^*$ onde μ é o parâmetro de escala;
2. Constrói-se a variável $V = \mu\beta' X^{PR}$;
3. Estima-se o modelo com os dados de PR com a variável V para obter os parâmetros de escala e das variáveis W, conforme equação 3.8.

$$U_{in} = \lambda V + \alpha W + \varepsilon_{in} \quad (3.8)$$

4. Calcula-se $\mu = 1/\lambda$, $\beta = \mu\beta^* / \mu e$.

Caso deseje-se melhorar a precisão, multiplicam-se os dados de X e Z do PD por μ , empilham-se os dados de PR e PD e estima-se novamente.

Por fim, para o caso simultâneo, o método proposto por Morikawa (1989) consiste em fazer uma estimação via máxima verossimilhança utilizando a equação (3.9). Na proposta, existe uma variável explicativa indicando que a amostra é de uma fonte ou de outra. Formalmente tem-se a função de verossimilhança da junção dos dados de PD e PR, atendendo as Hipóteses 1 a 4.

- **Hipótese 1:** $G_{PD}(\mu_x(x, z/\theta, \gamma)) = G_{PR}((x, y/\theta, \delta))$
- **Hipótese 2:** A função $f(*)$ é $C(2)$.
- **Hipótese 3:** A função $f(*)$ é $C(2)$, dado que é uma função multiplicativa de duas funções $C(2)$.
- **Hipótese 4:** A função $f(*)$ admite uma transformação monotônica.

Assim, a função união dos dados será:

$$f((x, z, y/\theta, \gamma, \delta, \mu) = (G_{PD}(\mu_x(x, z/\theta, \gamma))^d) * (G_{PR}((x, y/\theta, \delta))^{(1-d)}) \quad (3.9)$$

Onde:
$$d = \begin{cases} 1, & \text{se a amostra for de PD} \\ 0, & \text{caso contrário} \end{cases}$$

$x, z, y :=$ Variáveis explicativas;

$\theta, \gamma, \delta, \mu :=$ Parâmetros a estimar.

Note que a função união, trata-se de uma função contínua em partes, onde, a continuidade em todo o espaço é garantida no ponto de quebra dado pela variável *dummy*, onde os dados de PD são complementados pelos dados de PR. A função densidade de escolha, para o conjunto de escolha $E = \{0, 1\}$ é dado pela seguinte função de verossimilhança.

$$f((x, z, y / \theta, \gamma, \delta, \mu) = [(G_{PD}(\mu_x(x, z / \theta, \gamma))^d * (G_{PR}((x, y / \theta, \delta))^{(1-d)})]^{(E_1)} * \\ * [(G_{PD}(\mu_x(x, z / \theta, \gamma))^d * (G_{PR}((x, y / \theta, \delta))^{(1-d)})]^{(1-E_1)}$$

Logaritizando a função de verossimilhança tem-se:

$$\log(f((x, z, y / \theta, \gamma, \delta, \mu)) = E_1 * [d * \log((G_{PD}(\mu_x(x, z / \theta, \gamma))) + (1 - d) * \log((G_{PR}((x, y / \theta, \delta))))] + \\ (1 - E_1) * [d * \log((G_{PD}(\mu_x(x, z / \theta, \gamma))) + (1 - d) * \log((G_{PR}((x, y / \theta, \delta))))]$$

Assumindo verdadeira a hipótese acima se tem:

$$\log(f((x, z, y / \theta, \gamma, \delta, \mu)) = [d * \log((G_{PD}(\mu_x(x, z / \theta, \gamma))) + (1 - d) * \log((G_{PR}((x, y / \theta, \delta))))]$$

Sendo a função $G(*)$ contínua, diferenciável e monótona, então se tem uma solução interior. As variáveis z e y são comuns ou não às amostras, e a retirada de uma das variáveis em uma das amostras pode ser um artifício para a correção de multicolinearidade. Assumindo $G(*)$ como *logit*, então se tem a continuidade e solução interior, podendo fazer a maximização.

Deve-se ter em mente que a presença de heterocedasticidade será inevitável pela natureza dos dados diferentes, portanto, deve-se tratar o problema inerente a estimação. Por outro lado, para atender a hipótese de que os dados têm a mesma função geradora dos dados, deve-se fazer um teste tal como proposto em Hensher, Louviere e Swait (2000) e que consiste em fazer uma estimativa para cada fonte de dados e outra conjunta obtendo a estatística descrita na equação 3.10 e testar para uma distribuição qui-quadrada com $(|\beta| - 1)$ graus de liberdade.

$$-2((L^{PR} + L^{PD}) - L^{Conjunto}) \quad (3.10)$$

Os métodos acima descritos são uma das soluções para os problemas do modelo *logit* padrão. Na próxima seção apresenta-se a estrutura de utilidades para estimar modelos *logit* com parâmetros aleatórios.

3.5. MODELOS COM PARÂMETROS ALEATÓRIOS

Diferentemente do modelo clássico a proposta de estimação com parâmetros aleatórios, tenta captar diferenças em gostos, por conta da distribuição de probabilidade dos parâmetros β_n . Poduje (2002) e Train (2003) indicam que a distribuição de β_n dependerá dos pressupostos do analista referente ao comportamento do parâmetro. Tem-se a especificação para a utilidade indireta, conforme McFadden (1984) *apud* Poduje (2002) e Ben-Akiva, *et. al.* (2001), dada pela equação 3.12 e 3.13.

$$U_{in} = \beta_n x_{ink} + F_{in} T \zeta_{in} + \varepsilon_{in} \quad (3.11)$$

$$Cov(U_{in}) = F_{in} T T' F_{in} + (g / \mu^2) I_{J_n} \quad (3.12)$$

Onde: F_{in} := Matriz de fatores com parâmetros fixos e/ou desconhecidos; T := Matriz triangular inferior tal que $T T' = cov(T \zeta_{in})$; ζ_{in} := vetor de variáveis aleatórias com média zero e variância unitária; ε_{in} : Componente aleatória da utilidade de uma alternativa i para um indivíduo n iid com distribuição *Gumbel*, cujo parâmetro de escala $\mu > 0$ e a variância g / μ^2 , com g sendo a variância da distribuição *Gumbel* padrão ($\pi^2/6$).

A partir de F_{in} têm-se várias possibilidades de estimação além da proposta de parâmetros aleatórios, destacando os modelos com heterogeneidade não observada onde várias estruturas na variância e covariância das observações são propostas a exemplo de Morikawa (1989), Ben-Akiva e Morikawa (1990), Morikawa, Ben-Akiva, e Yamada (1991), Hensher e Bradley (1993), Hensher, Louviere, e Swait (1999) e Brownstone e Train (1999) ou então para estruturas autoregressivas tal qual Bhat e Guo (2002), LeSage (1999) e Pinkse e Slade (1998).

A estimação desta proposta segue a seguinte sequência:

1. Tome a estrutura de 3.12 e faça $\omega_{ink} = F_{ink} T \zeta_{ink} + \varepsilon_{ink}$ um erro composto por uma distribuição *Gumbel* ε_{ink} e outra proporção $F_{ink} T \zeta_{ink}$ definida pelo analista;
2. Assuma $\beta_n \sim f(b, \Sigma)$. Onde b é o vetor de médias populacionais dos parâmetros e Σ é a matriz de covariância destes na população. Note que o que se entende é que os indivíduos possuem diferentes β_n .
3. Para estimar o procedimento pode ser por Máxima Verossimilhança Simulada ou por Inferência Bayesiana. Faz-se a explicação de cada método a seguir.

Para este a distribuição Normal é a mais utilizada por ter uma variabilidade ampla dos parâmetros e uma especificação log-Normal quando se pretende restringir o sinal do parâmetro, mas a distribuição uniforme e a triangular podem também ser utilizadas.

3.5.1. Estimação por Máxima Verossimilhança Simulada.

Tome-se uma amostra Y_n de escolhas feitas pelo indivíduo n . A probabilidade condicional do indivíduo escolher uma sequência $Y_n=1$ se $U_{ink} > U_{jnk}, \forall j \neq i$ para parâmetros β_n fixos é expressa pela multiplicação de T_n funções *logit* tal como a equação 3.14. Com g_{ink} igual a um se $Y_n=1$ e zero em caso contrário.

$$L_i(y_n / \beta_n) = \prod_{t=1}^T \left[\frac{e^{(V_{in}^{(t)})}}{\sum_{j \in A(h)} e^{(V_{jn}^{(t)})}} \right]^{g_{ink}} \quad (3.14)$$

Pelo fato de β ser desconhecida, a probabilidade para a sequência de escolhas é feita pela integral da equação 3.15.

$$P(y_n) = \int_{-\infty}^{+\infty} L_i(y_n / \beta_n) f(\beta_n / b, \Sigma) d\Sigma \quad (3.15)$$

Com a função distribuição 3.15, então a função de log-verossimilhança simulada para estimação é dada pela equação 3.16 e que adota uma expressão a qual se simula a função 3.15 R vezes, pois esta não possui uma forma analítica fechada.

$$SL(b, \Sigma) = \sum_{n=1}^N \ln SP_n(y_n) \quad (3.16)$$

O modelo de máxima verossimilhança simulada é um estimador não viesado da probabilidade (Bhat, 2000; Train, 1999; e Hajivasilou e Ruud, 1993). O procedimento de estimação envolve simular a equação 3.15 a partir da distribuição condicional 3.17, expressa segundo a regra de Bayes, tendo em vista a sequência de escolhas Y_n dado os parâmetros populacionais b e Σ .

$$h(\beta_n | y_n, b, \Sigma) = \frac{P_n(y_n | \beta_n) f(\beta_n | b, \Sigma)}{P_n(y_n | b, \Sigma)} \quad (3.17)$$

A esperança condicional de β_n é calculada sobre todo o domínio da distribuição de β_n e a sua estimação por simulação de R observações de β_n^r feita conforme a equação 3.18, onde se extrai β_n^r da função densidade populacional $f(\beta_n | b, \Sigma)$, ponderada pela probabilidade condicionada *logit* de cada observação β_n^r .

$$SE(\beta_n | y_n, b, \Sigma) = \frac{\sum_{r=1}^R \beta_n^r P_n(y_n | \beta_n^r)}{\sum_{r=1}^R P_n(y_n | \beta_n^r)} \quad (3.18)$$

Posteriormente maximiza-se a equação 3.16 obtendo o valor de cada parâmetro β_n . Alternativamente, Revelt e Train (1998) utilizam o algoritmo de *Metropolis-Hastings* para obter simulações da distribuição de 3.17, para depois estimar os parâmetros.

3.5.2. Princípios da Estimação Bayesiana.

Segundo Poduje (2002) a estimação bayesiana permite aproveitar melhor as informações obtidas na amostragem. Para fazer a estimação segue-se o procedimento abaixo descrito e compilado do trabalho de Poduje (2002) *apud* Allenby e Ginter (1995) e *apud* Lenket. *al.* (1996).

1. Considere os parâmetros variáveis aleatórias e defina uma função *a priori* da distribuição desses;
2. Utilizando o teorema de probabilidades condicionais de Bayes onde:

$$\text{a. } P(\beta, y) = \frac{P(y|\beta)}{P(\beta)} ;$$

$$\text{b. } P(\beta, y) = \frac{P(\beta|y)}{P(y)} ; \text{ e}$$

$$\text{c. } P(\beta, y) = \frac{P(y|\beta)P(\beta)}{P(y)} .$$

obtenha uma distribuição de probabilidade $P(\beta, y)$ (tal como em 3.19) para os parâmetros e proporcional a $P(y|\beta)P(\beta)$, onde $P(y)$ é fixo por depender das observações Y_n e, portanto apresenta a soma das probabilidades igual a 1.

$$P(\beta|y) \propto P(y|\beta)P(\beta) \quad (3.19)$$

Onde: $P(\beta)$:= distribuição *a priori*;

$P(y|\beta)$:= verossimilhança dos dados; e

$P(\beta|y)$:= distribuição *a posteriori*.

Supondo $\beta_n \sim f(\beta_n|b, \Sigma)$ com a média populacional b segue uma distribuição específica e Σ segue outra distribuição, obtendo-se a expressão 3.20.

$$h(\beta_n) \propto \prod_{n=1}^N L_i(y_n / \beta_n) f(\beta_n / b, \Sigma) \gamma(b) \psi(\Sigma) \quad (3.20)$$

3. Iterativamente utilizando valores β_n e Σ faça:
 - a. Gere um valor para b ;
 - b. Utilizando valores β_n e b gere Σ de uma distribuição *Gamma* invertida;
 - c. Utilizando valores Σ e b gere valores de β_n utilizando o algoritmo de *Metropolis-Hastings* (amostragem de *Gibbs*);
 - d. Repita os passos anteriores até uma trajetória estável de β_n e b . Pontuje (2002) sugere 100.000 ou 200.000 iterações.
4. Faz-se a etapa 3 várias vezes retirando a cada 10 vezes o valor médio da distribuição *a posteriori* correspondente ao estimador bayesiano do parâmetro.
5. Para cada iteração calcula-se a função de verossimilhança dos parâmetros estimados até o momento. Esta função será utilizada para calcular as estatística de razão de verossimilhança e ρ^2 .

Para exposição das propriedades da estimação bayesiana ver Pontuje (2002). Train (2003) salienta que a estimativa bayesiana é mais vantajosa computacionalmente, gerando parâmetros mais robustos.

CONSIDERAÇÕES FINAIS

Enriquecimento de dados, tal como proposto em Morikawa (1989), pode gerar ganhos em estimação pontual, mas dependerá da função geradora dos dados das diferentes fontes serem compatível conforme salientado em Hensher, Louviere e Swait (2000), ou seja, os seus dados

para resultarem em uma boa combinação e em informações mais acertadas devem ter propriedades semelhantes. Por outro lado, os modelos com fatores aleatórios são importantes para descrever a idiosincrasia dos usuários, entretanto, a escolha do método de estimação é complicada e, pode de certo modo, apresentar diferenças dadas as propriedades das propostas de estimação por máxima verossimilhança simulada ou por estimativa bayesiana. De outro modo, este trabalho amplia a abordagem de Morikawa (1989) por formalizar as hipóteses necessárias a uma estimação com duas fontes de dados.

As técnicas utilizadas são intensivas em computação, portanto, podem ampliar problemas inerentes aos modelos estruturais apresentados no capítulo 2. De qualquer sorte, os métodos podem revelar informações importantes e melhorar o conhecimento do comportamento dos usuários para intervenções políticas mais acertadas, ou, no mínimo, com menor variabilidade.

CAPÍTULO 4. –DESCRIBÇÃO DA BASE DE DADOS DO TRANSPORTE DE PASSAGEIROS DO CEARÁ

4.1. INTRODUÇÃO

O transporte clandestino (irregular) foi intensificado ao final da década de 1990 e início de 2000 no Brasil decorrente da insatisfação dos usuários pelo atendimento precário por parte dos transportadores regulares destacando-se: a oferta temporal escassa, preço e/ou quantidade do serviço de transporte regular não competitivo. O estudo de Golub *et.al.* (2009), aplicado no contexto interurbano da região metropolitana do Rio de Janeiro relata este crescimento.

No caso cearense, em 2004, foram realizados cerca de 42 milhões de deslocamentos intermunicipais de passageiros e pouco menos de 40% dos deslocamentos utilizou do sistema regular por ônibus. Excetuando-se o CRAJUBAR (Crato-Juazeiro-Barbalha), o crescimento da demanda de transporte pelo sistema regular, entre 1997 e 2004, foi de apenas 2%. Por outro lado, o transporte clandestino mais que dobrou nas viagens regionais. A Tabela 4.1 apresenta a participação do transporte clandestino intermunicipal na matriz de transportes de passageiro do estado do Ceará para os anos de 1997 e 2004.

Nos deslocamentos interior-capital o transporte clandestino registrou uma redução de 15% de 1997 para 2004, conforme indicado na Figura 4.1. Algumas suposições foram levantadas nos relatórios do PDOTIP relativas a esta redução com destaque as seguintes: as ações fiscalizatórias no eixo capital-interior ou decorrente da distância elevada nos deslocamentos capital-interior em comparação com as regionais.

Tabela 4.1: Evolução das viagens no STRIP-CE (Sistema de Transporte Intermunicipal de Passageiros do Ceará) no período de 1997 a 2004.

Deslocamentos (milhões)	1997			2004		
	Regular	Clandestino	Total	Regular	Clandestino	Total
Radial (interior-capital)	4,1	3,2	7,3	4,4	1,8	6,2
Regional (*)	4,0	9,5	13,5	3,9	22,8	26,7
Total	13,5	13,1	26,6	16,3	25,9	42,2

(*) excetuando-se os deslocamentos internos na Região do CRAJUBAR. (Fonte: PDOTIP, 2005). Radiais tratam-se dos deslocamentos provenientes do interior a capital do estado (Fortaleza) e vice-versa. Regional refere-se às viagens entre cidades desconsiderando a capital.

Contrariamente, os deslocamentos regionais registraram 15% de aumento na participação do clandestino. E a hipótese mais plausível para este fenômeno seria a de maior sinergia entre as regiões decorrente do crescimento das atividades entre as cidades do interior do Ceará, destacando-se Juazeiro, Quixadá e Sobral, com a criação dos pólos industriais e/ou das universidades federais.

Segundo o PDOTIP (2005), várias vantagens fazem com que os usuários escolham o clandestino destacando-se os seguintes fatos: o acesso ao transporte regular exige maiores deslocamentos a pé dos usuários; o clandestino tem a opção de o usuário pagar a prazo; e o transporte clandestino possui uma flexibilidade maior no preço do que o regular, apresentando variação de preços inclusive durante o dia.

O tempo é um atributo importante na escolha dos usuários de transportes, pois reflete a acessibilidade do sistema. Dependendo do valor do tempo para os usuários pode ser pensado em intensificar a concorrência entre modos via mudanças operacionais de paradas e redução do tempo da viagem.

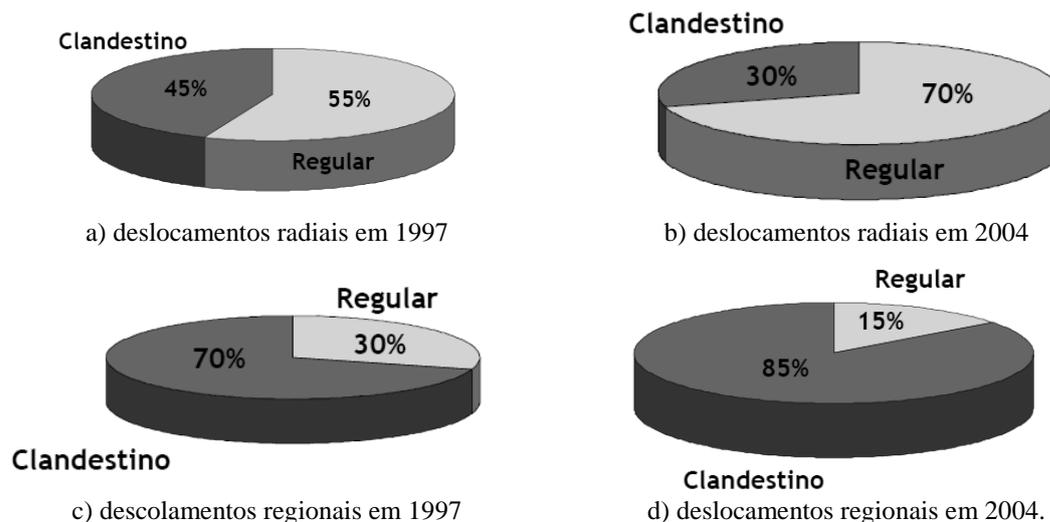


Figura 4.1: Evolução da participação do transporte clandestino nas viagens radiais 1997 – 2004 (Fonte: PDOTIP-CE 2004)

Em 2005, em razão do PDOTIP, foram feitas duas pesquisas aos usuários de transporte para embasar a reestruturação do STRIP-CE, em parte motivados pela problemática acima relatada. A primeira pesquisa era referente à satisfação dos usuários a segunda referente à sua escolha frente a mudanças no sistema ou situações hipotéticas de melhoria de qualidade do ônibus ou de inserção do modo ferroviário como alternativo para o transporte intermunicipal. Na próxima seção descrevemos os dados coletados nas pesquisas.

4.2. FONTES DE DADOS

Este item apresentara as generalidades dos dados que serão utilizados na estimação.

4.2.1. Generalidades dos dados de Preferência Revelada e Declarada

Os dados são provenientes da pesquisa de preferência declarada (PD) e de satisfação do usuário (PR) feitas para definição do Plano Diretor e Operacional do Transporte Intermunicipal de Passageiros do Estado do Ceará - PDOTIP-CE (2005). As pesquisas foram realizadas nos terminais rodoviários e embarcadas nos veículos. Os dados obtidos foram: dados socioeconômicos dos usuários, motivo e frequência de viagem, origem e destino da viagem, e a

escolha do indivíduo com relação a duas alternativas (ônibus e van, van e trem e trem e ônibus).

Para a modelagem utilizou-se a definição que trem e ônibus são modos regularmente estabelecidos enquanto que van trata-se de um modo clandestino³. Os locais de coleta dos dados foram: Fortaleza, Juazeiro e Iguatu e os entrevistados tinham residência e destino em diversos municípios do estado e em estados vizinhos, tais como Pernambuco, Piauí e Rio Grande do Norte.

Para o exercício empírico serão utilizadas as variáveis: sexo, número de carros na residência, número de moto na residência, idade do entrevistado, tempo da viagem, custo da viagem. Foram eliminadas as observações onde não existissem respostas para uma dessas variáveis.

Na pesquisa de PD um indivíduo era perguntado sobre várias hipóteses. Nesse caso adotou-se cada resposta como se fosse de um indivíduo, procedimento coerente estatisticamente conforme Ben-Akiva e Lerman (1984), por assumir uma robustez na escolha dos indivíduos, salvo quando não for cumprimento da transitividade das escolhas.

O custo da viagem usada na amostra de PR é resultado de uma pesquisa complementar executada com as cooperativas de vans e das tarifas praticadas nos ônibus, mas na pesquisa de satisfação de usuário foram feitas perguntas na escala de satisfação do usuário com relação ao custo da viagem.

Para a união das amostras adicionou-se uma variável que indica 1 se a observação pertence à PD e 0 se pertence a PR.

Os atributos coletados referentes aos dados socioeconômicos foram: sexo; idade; local de residência; grau de instrução; renda familiar; situação familiar; ocupação; disponibilidade de automóvel em casa; e disponibilidade de motocicleta em casa; os referentes ao padrão de

³ Para sustentar tal opção, foi feito antes de modelar, testes de IIA para identificar se existia tal propriedade na escolha entre ônibus e trem.

viagem foram: origem da viagem; destino da viagem; tempo de viagem; tipo de transporte utilizado; tipo de transporte mais utilizado; frequência de viagem; motivo da viagem; e adicionalmente tem-se: tempo de espera; tempo de caminhada. Os níveis escolhidos para os atributos “Tempo de Viagem” e “Custo da Tarifa”, para cada ligação, estão apresentados nas Tabelas 4.2 e 4.3, respectivamente.

Tabela 4.2: Níveis dos atributos de Tempo de Viagem

Nível 01	Nível 02	Nível 03
15min	25min	30min
20 min	30 min	40 min
40 min	50 min	60 min

Tabela 4.3: Níveis dos atributos Custo da Tarifa

Nível 01	Nível 02	Nível 03
R\$1,00	R\$1,10	R\$1,20
R\$1,00	R\$1,10	R\$1,20
R\$1,00	R\$1,10	R\$1,20

Para definição dos cartões de pesquisa foi utilizado um projeto fatorial definindo-se 06 cartões para a pesquisa, sendo exemplificados na Figura 4.2 os cartões apresentados aos usuários das ligações do Cariri.

Atributos	Alternativa 01	Alternativa 02
Modo de Viagem	 Trem	 Van
Tempo de Viagem	15min	25min
Custo da Tarifa	R\$ 1.00	R\$ 1.00

Figura 4.2: Exemplo de cartão de PD. •.

A estatística descritiva das amostras de preferência revelada e declarada é apresentada na Tabela 4.4. Observe que as variáveis têm pequenas mudanças da PD para a PR. As idades dos respondentes para ambas as amostras são parecidas, inclusive o desvio padrão destas. Na amostra de preferência declarada existem mais homens do que mulheres ocorrendo o oposto para a preferência revelada.

Os números de motos e de carros por residência e o tempo de viagem são os que apresentam maior mudança. No caso de possuir veículos, indicando a quantidade por residência, os respondentes de pesquisa revelada apresentam cinco vezes a quantidade média de veículos por residência que a declarada, isso é um fator a considerar, pois ao modelar conjuntamente os dados, possivelmente a decisão dos respondentes de declarada frente à revelada pode mudar pelo simples fato de existir a opção de veículos na residência nos respondentes da PR com uma maior intensidade.

	Pesquisa Declarada		Pesquisa Revelada	
	Média	Desvio Padrão	Média	Desvio Padrão
Sexo	0.51	0.50	0.47	0.50
Idade	34.25	14.10	34.70	14.76
Auto	0.25	0.49	1.17	0.38
Moto	0.23	0.47	1.15	0.35
Tempo_viagem1	30.36	15.88	34.96	6.37
Custo_Tarifa1	1.15	0.09	1.13	0.09
Tempo_viagem2	38.80	16.17	34.96	6.37
Custo_Tarifa2	1.03	0.05	1.09	0.03
Nicho_1	0.64	0.48	0.32	0.47
Amostra	1554.00		131.00	

Tabela 4.4: Estatística Descritiva dos Dados de PD e PR e Conjunta do STRIP-CE (Sistema de Transporte Intermunicipal de Passageiros do Ceará) para os estudos do PDOTIP-CE.

Um olhar mais focado na amostra de PD é necessário, uma vez que serão estimados modelos apenas com esta fonte de dados. Desta, pode-se observar que os tempos de viagem são pequenos com média de 38,80 minutos e desvio padrão de 16,17. As escolhas dos usuários no experimento foram significativamente maiores para o transporte regular, embora os custos fossem menores para o transporte irregular e os tempos de viagens maiores. Observa-se a predominância de (51%) de homens na amostra; a idade tem média e desvio padrão aproximados em 34 e 14 anos; a quantidade média por residência de carros e de motos é de aproximadamente 0.25 e 0.23 com desvio padrão de 0.49 e 0.47, respectivamente. Isso sinaliza para uma quantidade baixa de veículos na residência, não havendo possíveis efeitos substituição entre o transporte individual e o coletivo. Estes dados não são utilizados no

modelo, mas caracterizam variáveis de ambiente socioeconômico e que podem influenciar na percepção dos usuários conforme descrito por Jiang e Morikawa (2004).

A pesquisa ainda envolveu a avaliação de renda nos seguintes níveis: sem renda, com um salário mínimo, de um a dois salários mínimos, de dois a cinco salários mínimos, de cinco a dez salários mínimos e acima de dez salários mínimos, representando sub amostras de 204, 428, 428, 372, 132 e 16 indivíduos, respectivamente, portanto, com uma grande heterogeneidade de renda.

Fora os dados socioeconômicos, foram questionados aos pesquisados a frequência e os motivos das viagens, destes atributos: identificou-se 39,45% indivíduos que usam o transporte para trabalho, 21% para viagens comercial, 11,33% que utilizam para lazer, 8,2% para escola e 5,08% para outros. A frequência de utilização que se destacou foi o uso diário com 22,88% e na mesma proporção o uso com frequência de 2 a 5 vezes por semana, o uso mensal e semanal representou 19,14% e 16,80% respectivamente.

4.2.2. Considerações sobre a heterogeneidade na amostra da pesquisa declarada

Os dados de PD foram feitos em diferentes regiões. Existem possíveis características regionais que influenciam a decisão dos usuários. Isso ficou latente no trabalho de Freire *et. al.* (2006) que encontrou valores de correlação espacial do número total de viagens por município significativamente maiores que para variáveis socioeconômicas. Isso pode ser decorrente de, nos últimos anos, terem sido instaladas Universidades Federais, Centros de Ensino Tecnológicos e indústrias nas regiões de Sobral/ Ibiapaba e do Sertão Central e do Cariri/Centro Sul, influenciando seu potencial atrator e gerador de viagens.

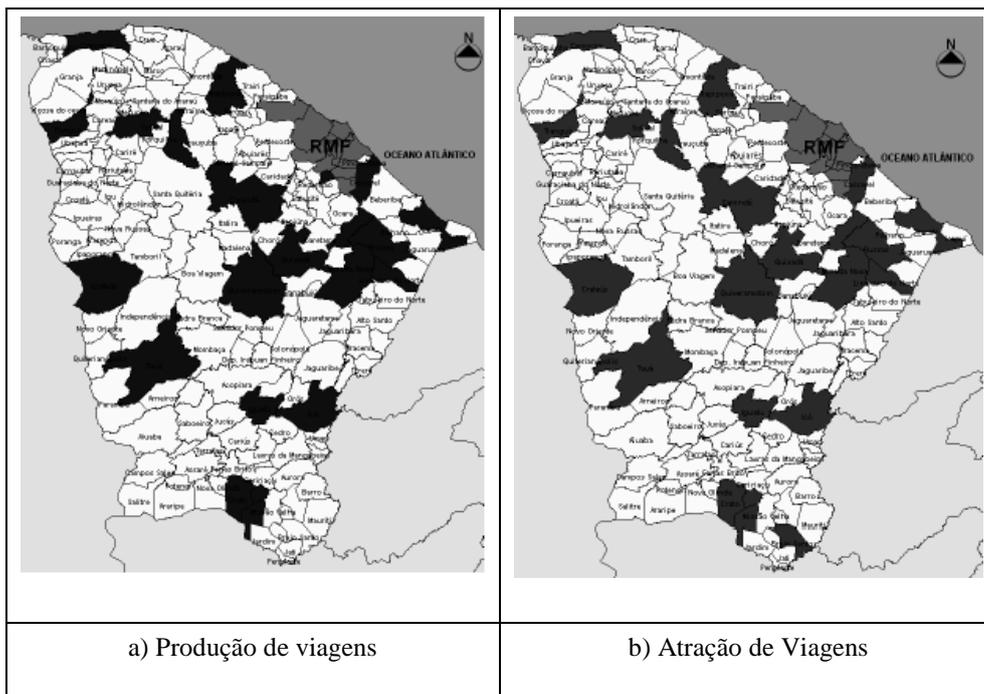


Figura 4.3 –Municípios que mais produzem e atraem viagens no Ceará. (Fonte: PDOTIP-CE, 2004)

Relativo às viagens, o relatório executivo do PDOTIP (2005) aponta problemas de mobilidade e acessibilidade dos usuários de transporte regular. Indica-se que 13% da população do Ceará, à época, não tinha acesso direto a rede de transportes e que esses necessitavam de um deslocamento médio, a pé, de 3 km para acessar a rede. No mesmo relatório indica-se que 10% da população do interior do estado, à época, dispunham de automóveis nas suas residências e que existia, ineficiência na rede, dos operadores de transportes e uma baixa qualidade dos serviços prestados.

Freire *et. al.* (2006) supõe que indivíduos de um município podem exercer influência sobre os outros indivíduos de municípios vizinhos, sendo verificado que a região do Litoral Leste do estado tem uma tendência maior de concentrar viagens.

Tabela 4.5: Residência dos respondentes por renda.

Renda	CRATO	BARBALHA	MISSAO VELHA	JUAZEIRO	OUTROS
1A2	28	52	84	92	44
1SM	40	64	68	76	44
2A5	40	40	48	56	52
5A10	4	12	28	28	0
A10	4	4	-	-	4
SR	12	16	44	52	28
Total Geral	128	188	272	304	172

Como pode ser observada na Tabela 4.5 a residência dos indivíduos está concentrada na região do CRAJUBAR e com um espectro quase uniforme em renda de um a cinco salários mensais. A variedade das rendas e localidades pode influir no parâmetro, portanto, uma estimativa com parâmetros aleatórios é justificada para o nosso estudo. Para a amostra de PR será adotado uma subamostra com valores de duração de viagem entre 30 e 45 minutos.

Do exposto, uma modelagem em que a aleatoriedade dos indivíduos seja considerada pode gerar uma informação mais robusta para fomentar intervenções públicas no mercado.

CONSIDERAÇÕES FINAIS

As fontes de dados coletadas permitem supor comportamentos diferentes entre os entrevistados, portanto, passível de uma modelagem via fatores aleatórios e, pelo fato de existirem duas fontes, modelarem tal como Morikawa (1989). Ressalte que para o caso da modelagem conjunta será utilizado uma subamostra dos dados de PR com valores compatíveis aos valores propostos na pesquisa de PD. O resultado dessas abordagens e o conhecimento relativo à valoração do tempo pelos usuários de transportes do estado do Ceará serão tratados nos capítulos que seguem.

CAPÍTULO 5. – METODOLOGIA E RESULTADOS

5.1. PROCEDIMENTO METODOLÓGICO ADOTADO

Metodologicamente serão apresentados os resultados provenientes de formas funcionais de modelos estruturados de teoria do valor de tempo com base nos trabalhos de Tuong e Hensher (1985a e 1985b) comparado com uma estimação não estruturada a exemplo de trabalhos como Train e McFadden (1978). Em concomitância, foram feitas estimativas utilizando do modelo *logit* simples, seguindo com o procedimento de enriquecimento de dados tipo Morikawa (1989) e de fatores aleatórios a exemplo de Train (1998) e Revelt e Train (1999). Feitas as estimações, foram realizados cálculos relativos ao valor do tempo das viagens, de acordo com Tuong e Hensher (1985a) para cada modelo e analisado comparando as diferenças em relação à forma funcional e ao procedimento de estimação tomando como base os resultados do modelo *logit* simples.

5.2. DESCRIÇÃO DO MODELO E RESULTADOS DAS ESTIMAÇÕES

Nesta seção serão feitas considerações sobre as funções de utilidade indireta escolhida e apresentados os resultados de estimação.

5.2.1. Estratégias de Estimação

Considerações sobre as formas funcionais

Para a estimação dos modelos serão adotadas as propostas de Tuong e Hensher (1985b) como formas funcionais. Estas estão expressas nas equações 5.1, 5.2 e 5.3. Complementarmente, será utilizado um modelo não estruturado tal como Train e McFadden (1978) sugerem com uma estrutura linear para a utilidade indireta que combina tempos de viagem com a renda dos indivíduos de uma forma multiplicativa tal como aplicado por Cherchi e Ortúzar (2002).

$$V_i = -\lambda x_i - \mu t_i \quad (5.1)$$

$$V_i = -\lambda x_i - (\mu - k_i) t_i \quad (5.2)$$

$$V_i = -\lambda x_i - (\mu - \bar{k}) t_i + (\alpha + \beta t_i^2 + \gamma x_i t_i) \quad (5.3)$$

Onde: V_i := nível da utilidade indireta associado à escolha da viagem i ;

x_i, t_i := são o custo e o tempo associado à viagem i ;

λ, μ := parâmetros a serem estimados que retratem utilidades marginais da renda e do tempo respectivamente;

k_i := multiplicador de lagrange associado ao consumo de tempo restrito a tecnologia e que retrata a utilidade marginal do tempo poupado.

Na equação 5.1 a taxa marginal do tempo é a mesma para qualquer modo de transporte, nas equações 5.2 e 5.3 estas são variáveis dependendo do modo de transportes. Por seu turno na estrutura tipo McFadden (1978) as pessoas com renda diferente valoram o tempo diferentemente, mas sem variação da taxa marginal do tempo entre os modos de transporte.

Considerações sobre as utilidades e a estimação com modelo logit simples

Na estimativa por *logit* simples a constante específica da alternativa i é o parâmetro que representa todos os efeitos de escolha da alternativa que não estão incluídos nos efeitos dos atributos observados. Os termos aleatórios serão assumidos com distribuição de probabilidade de valor extremo II. As utilidades por modo de transporte (regular e irregular) serão estimadas adotando uma forma linear com as variáveis tais como descritas nas equações 5.4 e 5.5.

$$V_{PDreg} (TVreg, Custoreg; \theta) \quad (5.4)$$

$$V_{PDIRreg} (TVirreg, Custoirreg; \theta) \quad (5.5)$$

Onde:

V_{PDreg} := Utilidade indireta para os dados de PD para o transporte regular;

$V_{PDirreg}$:= Utilidade indireta para os dados de PD para o transporte irregular;

$TVreg$:= Tempo de viagem no transporte regular (ônibus ou trem);

$Custoreg$:= Valor da tarifa no transporte regular (ônibus ou trem);

$TVirreg$:= Tempo de viagem no transporte irregular (van);

$Custoirreg$:= Valor da tarifa no transporte irregular (van);

θ := vetor de parâmetros a ser estimado.

Considerações sobre as utilidades e a estimação com junção de fontes de dados

Na proposta de junção de fontes de dados adota-se a estimação simultânea de Morikawa (1989). As funções de utilidade para os dados de PD e PR são conforme as equações 5.4, 5.5 e 5.6, esta última utilidade é a adotada para os dados de PR com as variáveis socioeconômicas.

$$V_{PR}(\text{Sexo}, \text{Idade}, \# \text{Auto}, \# \text{Moto}, \text{Custo de viagem}; \theta) \quad (5.6)$$

Assume-se que existe um parâmetro que escalona a estimação de PR compatibilizando-o com o PD, este será multiplicativo com a utilidade de PR e não aparecerá na função de PD. Ressalte-se que o atributo de “*Trade-off*” é o tempo de viagem e os de equilíbrio são o custo de transportes e as variáveis socioeconômicas. Os modelos seguiram os modelos já apresentados na seção anterior. Adicionalmente será executado o teste proposto em Hensher, Louviere e Swait (2000) após a estimação para verificar se os dados de PD e PR possuem a mesma fonte geradora de dados.

Considerações sobre a utilidade e a estimação por logit com fatores aleatórios

A forma funcional adotada para este caso foi à mesma que apresentada no *logit* simples. O parâmetro do tempo foi assumido possuir distribuição de probabilidade normal com média e variância. Além disso, por questões de facilidade para o cálculo do valor do tempo de viagem, adotou-se o parâmetro de custo da tarifa como fixo. Os parâmetros que são compostos de tempo foram considerados com distribuição normal tal como o parâmetro do tempo.

5.2.2. Resultados da estimação do logit simples

Os resultados dos diferentes modelos são apresentados na Tabela 5.1.⁴ Os parâmetros apresentaram sinais iguais ao previsto na teoria conforme equações 5.1, 5.2 e 5.3. Ou seja, valores negativos para os parâmetros do custo da viagem e do tempo, representando uma perda marginal conforme o preço e o tempo da viagem aumentam. Portanto, quanto maior o tempo de viagem, menor seria a utilidade atribuída pela viagem, valendo o mesmo para o custo da viagem.

Os parâmetros são significantes na maioria das estimações, exceto para o parâmetro do tempo (valor marginal do tempo) do modelo de Tuong e Hensher. Nas três últimas linhas são apresentados o tamanho da amostra, o pseudo R^2 de McFadden e o mínimo da função de máxima verossimilhança encontrado.

Os valores de aderência, pseudo R^2 de McFadden e o mínimo da função de máxima verossimilhança, indicam que a forma funcional de Tuong e Hensher apresenta melhor aderência, seguido pelo modelo de DeSerpa, Becker e McFadden.

⁴ O modelo logit poderia ter sido feito em estimativa multinomial ou logit agrupado, dado que o questionário foi aplicado com a decisão entre ônibus, trem e van. Na estimativa, conveniou-se chamar ônibus e trem de regular e van de irregular. Porém, antes de tal adoção, fizeram-se dois testes de acordo com o sugerido em Train (2003) sobre a existência de IIA no logit aninhado entre ônibus e trem ou no multinomial logit. Os resultados indicaram a inexistência do problema de IIA, com a rejeição da hipótese de existir IIA a 10% e a convergência do modelo logit aninhado para um logit condicional.

Os modelos de DeSerpa e Tuong e Hensher têm a hipótese de diferença do valor marginal do tempo entre diferentes atividades, no caso diferentes modos de transportes. Do resultado empírico, verifica-se que o modelo de DeSerpa tem parâmetros de valor marginal do tempo significantes estatisticamente e com sinais tal como esperado, não ocorrendo o mesmo com o modelo de Tuong e Hensher. Os valores de tempo marginal dos modelos de McFadden e Becker são significantes e com sinal conforme a teoria. Complementarmente, observando o valor do tempo variando com a renda no modelo de McFadden, verifica-se um crescimento do valor marginal do tempo com o aumento da renda, ressalvado a renda de até 10 salários mínimos que é a menor de todas e sugere um desuso do transporte por estes respondentes.

Tabela 5.1: Modelos estimados. Os valores abaixo do parâmetro são os desvios padrões e os dois asteriscos apontam o parâmetro significante a 5% de nível de significância.

	McFadden (Conceitual)	Becker (5.1)	DeSerpa (5.2)	Tuong & Hensher (5.3)
const	1.789** 0.230	1.866** 0.235	1.618** 0.242	1.839** 0.260
Temp*SR	-0.04560** 0.016	- -	- -	- -
Temp*1SL	-0.04584** 0.013	- -	- -	- -
Temp*1A2SL	-0.06156** 0.013	- -	- -	- -
Temp*2A5SL	-0.06068** 0.014	- -	- -	- -
Temp*5A10SL	-0.03997** 0.019	- -	- -	- -
Custo	-6,518** 1.174	-6,87** 1.194	-8,204** 1.245	-9,417** -1.336
Tempo	- -	-0.05693** 0.011	- -	0.008 0.039
Tempo^2	- -	- -	- -	0.000 0.000
Tempo*Custo	- -	- -	- -	-0.07848** 0.024
Tempo_Reg	- -	- -	0.07569** 0.012	- -
Tempo_Irreg	- -	- -	-0.06101** 0.011	- -
N	1564	1564	1564	1564
Pseudo R^2	0.018	0.018	0.026	0.027
lnL	-1007	-1007	-998.9	-997.9

Const=constante; Temp*SR=produto entre tempo e uma dummy para os indivíduos que não apresentaram renda fixa; Temp*1SL=produto entre tempo e uma dummy para os indivíduos que apresentaram um salário mínimo; Temp*1A2SL=produto entre tempo e uma dummy para os indivíduos que apresentaram de um a dois salários mínimos; Temp*2A5SL=produto entre tempo e uma dummy para os indivíduos que apresentaram dois a cinco salários mínimos; Temp*5A10SL=produto entre tempo e uma dummy para os indivíduos que apresentaram cinco a dez salários mínimos; Custo=custo da viagem; Tempo=tempo da viagem; Tempo^2=quadrado do tempo da viagem; Tempo*Custo=produto do tempo da viagem pelo custo; Tempo_Reg=tempo da viagem do regular; Tempo_Irreg=Tempo da viagem do irregular. A renda base para o modelo McFadden é acima de dez salários mínimos.

Os custos marginais de todos os modelos são significantes estatisticamente. Os modelos de McFadden e Becker apresentam semelhança, tal como os modelos de DeSerpa e Tuong e Hensher.

Por fim, a constante é aparentemente igual e significativa para todos os modelos, sendo o menor valor o obtido da estimação segundo DeSerpa e o maior o do modelo de Becker, este último praticamente igual ao modelo de Tuong e Hensher.

5.2.3. Resultados da Estimação do *logit* com enriquecimento de dados

O resultado do modelo com enriquecimento é apresentado na Tabela 5.2. Os parâmetros de valor marginal do tempo são significantes estatisticamente. Os parâmetros que μ se referem ao parâmetro de escalonamento da função gerada na preferência revelada e não é significativa para todos os modelos. Todas as constantes são significantes. Os pseudo R^2 apresentam valores baixos e os comparando com os obtidos na estimação simples, pode-se sugerir que a estimação com enriquecimento gera uma perda de aderência.

Os resultados replicam o obtido na primeira estimativa. O modelo de DeSerpa tem parâmetros de valor marginal do tempo significantes estatisticamente e com sinais tal como esperado, não ocorrendo o mesmo com o modelo de Tuong e Hensher. Os valores marginais de tempo dos modelos de McFadden e Becker são significantes e com sinal conforme a teoria. Complementarmente, observando o valor do tempo variando com a renda no modelo de McFadden, verifica-se um crescimento do valor marginal do tempo com o aumento da renda,

ressalvado a renda de até 10 salários mínimos que é a menor de todas e sugere um desuso do transporte por estes respondentes.

Os custos marginais de todos os modelos são significantes estatisticamente. Além disso, obteve-se uma redução do valor do custo marginal de Tuong e Hensher quando comparado à estimativa com enriquecimento frente à estimativa com *logit* simples.

Tabela 5.2: Comparação de modelos simples e com enriquecimento dos dados os valores abaixo do parâmetro representam os desvios padrões e os asteriscos indicam parâmetros significantes a 5% de nível de significância

	McFadden (Conceitual)	Becker (5.1)	DeSerpa (5.2)	Tuong & Hensher (5.3)
const	1.777 0.229	1.827 0.235	1.613 0.242	1.522 0.291
Temp*SR	-0.045 0.016	- -	- -	- -
Temp*1SL	-0.045 0.013	- -	- -	- -
Temp*1A2SL	-0.061 0.013	- -	- -	- -
Temp*2A5SL	-0.060 0.014	- -	- -	- -
Temp*5A10SL	-0.039 0.019	- -	- -	- -
Custo	-6.453 1.171	-6.661 1.196	-8.187 1.243	-8.392 1.528
Tempo	- -	-0.055 0.011	- -	0.074 0.041
Tempo^2	- -	- -	- -	-0.001 0.001
Tempo*Custo	- -	- -	- -	-0.092 0.023
Tempo_Reg	- -	- -	-0.076 0.012	- -
Tempo_Irreg	- -	- -	-0.061 0.011	- -
μ	-0.009	-0.017	-0.010	-0.025
N	1695	1695	1695	1695
Pseudo R ²	0.016	0.016	0.023	0.022
lnL	-1089.690	-1089.522	-1081.974	-1082.636
Graus de Liberdade	11	7	8	9
Teste de Hensher, Louviere & Swait (2000)	-4192.861732	-4192.525732	-4161.229732	-4160.553732

Const=constante; Temp*SR=produto entre tempo e uma *dummy* para os indivíduos que não apresentaram renda fixa; Temp*1SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram um salário mínimo; Temp*1A2SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram de um a dois salários mínimos; Temp*2A5SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram dois a cinco salários mínimos; Temp*5A10SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram cinco a dez salários mínimos; Custo=custo da viagem; Tempo=tempo da viagem; Tempo^2=quadrado do tempo da viagem; Tempo*Custo=produto do tempo da viagem pelo custo; Tempo_Reg=tempo da viagem do regular; Tempo_Irreg=Tempo da viagem do irregular. A renda base para o modelo McFadden é acima de dez salários mínimos.

Ao executar o teste de Hensher, Louviere e Swait (2000), assumindo a hipótese nula de que os dados possuem as mesmas funções geradoras, conclui-se com a rejeição da hipótese nula e, portanto, não se pode afirmar que os dados possuem a mesma função geradora dos dados, assim, pode-se concluir que os resultados não foram melhores do que os da estimativa simples. Para verificar a robustez do teste far-se-á um estudo dos valores de tempo com os modelos enriquecidos na seção onde se compara os valores de tempo. De forma complementar, fez-se uma verificação dos resultados do modelo com os dados de PR. Na Tabela 5.3. tem-se o resultado da estimação para a amostra de 131.

Tabela 5.3: Resultados do modelo com os dados de preferência revelada os asteriscos indicam parâmetros significantes a 5% de nível de significância.

Revelada		
	coeficiente	erro padrão
const	32,307	4,290
Sexo	12,500	2,357
Idade	-0,150	0,055
Auto	-10,737	0,897
Moto	-16,572	2,269
Custo	-179.519	19,493
N	131	
pseudo R ²	0,997	
Log da verossimilhança	-0,26	

Não foi possível fazer o modelo estruturado tal como apresentado nas equações 5.1, 5.2 e 5.3, por inexistência da variável tempo de viagem na amostra⁵, portanto, estes modelos, com os dados de PR, apresentam variáveis omitidas. Esta omissão de variáveis pode resultar em problemas de heterogeneidade não observada. A significância estatística só foi presente no intercepto e no parâmetro de custo. Note ainda que o pseudo R² sugere um bom ajuste do modelo.

⁵ Em verdade, existia o tempo da viagem, mas não o tempo da viagem para cada modo, mas o tempo de viagem do modo escolhido, inviabilizando a construção dos modelos estruturados. Além disso, inexistiam também as faixas de renda para desenvolver o modelo conforme McFadden.

5.2.4. Resultados da estimação do *logit* com parâmetros aleatórios

Os resultados da estimação com parâmetros aleatório são apresentados na Tabela 5.4. A leitura da Tabela 5.4 é feita pela média do parâmetro da variável e o desvio padrão deste logo abaixo. As linhas em cinza representam o desvio padrão da distribuição do parâmetro seguido pelo seu desvio padrão.

Os resultados das médias do parâmetro da variável são significantes, mas os desvios padrões da distribuição de probabilidade adotada para o parâmetro não o são na maioria dos casos da máxima verossimilhança simulada. Isso não ocorre na estimativa bayesiana.

Observa-se considerável diferença nos parâmetros estimados em máxima verossimilhança simulada para a estimativa bayesiana frente ao modelo simples. Esta diferença nos parâmetros foi em valor absoluto para mais em todos os parâmetros para todos os modelos.

A estimação por máxima verossimilhança simulada apresentou resultados muito próximos da estimação simples, sugerindo a inexistência de IIA tal como obtido nos testes sugeridos por Train (2003) e explicado na nota de rodapé três.

Os parâmetros de custo são maiores que a estimativa simples. No modelo de Tuong e Hensher, observou-se descompasso em sinal com o valor marginal do tempo.

Todos os parâmetros obtidos pela estimativa bayesiana são significantes estatisticamente. Isso não ocorreu nas interações entre tempo de viagem e renda no modelo *a la* McFadden em particular para o caso de salários entre cinco e 10 salários. Ao estimar por um procedimento bayesiano os parâmetros são praticamente o dobro da estimativa simples e da por máxima verossimilhança simulada. Infelizmente a comparação entre as estimativas não é imediata, pois o conceito de uma estimação bayesiana é diferente de uma estimativa frequentista. O fato é que a estimativa bayesiana incorpora crenças do analista e tenta adequá-las com os dados amostrais enquanto que a frequentista apenas utiliza dos dados, no momento que é utilizado o

procedimento bayesiano o técnico esta viesando os parâmetros para mais ou para menos dependendo do seu conhecimento do fenômeno.

Tabela 5.4: Comparação de modelos com parâmetros aleatórios e estimados com diferentes procedimentos. Os valores abaixo do parâmetro representam os desvios padrões e os dois asteriscos apontam o parâmetro significante a 5% de nível de significância.

	Máxima Verossimilhança Simulada				Estimativa Bayesiana			
	McFadden (Conceitual)	Becker (5.1)	DeSerpa (5.2)	Tuong & Hensher (5.3)	McFadden (Conceitual)	Becker (5.1)	DeSerpa (5.2)	Tuong & Hensher (5.3)
const	-1,796** 0.2233	-1,8745** 0.2364	-1,6332** 0.2413	-1.8417** 0.2683	-3,6607** 0.2924	-2,0951** 0.1821	-1,1223** 0.4481	-4.6236** 0.3466
Temp*SR	-0,0453** 0.016	-	-	-	-0,0552 0.0651	-	-	-
SE(Temp*SR)	0.0016 0.0187	-	-	-	0.1675 0.0344	-	-	-
Temp*1SL	-0,0464** 0.0127	-	-	-	-0,0776 0.0357	-	-	-
SE(Temp*1SL)	-0,0001 0.0136	-	-	-	0.0931 0.0171	-	-	-
Temp*1A2SL	-0,0595** 0.0126	-	-	-	-0.1261 0.0338	-	-	-
SE(Temp*1A2SL)	-0,0004 0.0105	-	-	-	0.0765 0.0121	-	-	-
Temp*2A5SL	-0,0604** 0.0142	-	-	-	-0.1205 0.0361	-	-	-
SE(Temp*2A5SL)	-0,0015 0.0134	-	-	-	0.0999 0.0172	-	-	-
Temp*5A10SL	-0,0403** 0.0192	-	-	-	0.0369 0.1113	-	-	-
SE(Temp*5A10SL)	-0,0002 0.0228	-	-	-	0.2968 0.1002	-	-	-
Custo	-6,5395** 1.1461	-6,8996** 1.2027	-8,2983** 1.2475	-9.4995** 1.3616	-15,5686** 1.4418	-7,9795** 0.891	-15,89** 1.7999	-25.2396** 2.0602
Tempo	-	-0,0564** 0.011	-	0.0117 0.0405	-	-0,0604** 0.0105	-	-0.5929 0.0852
SE(Tempo)	-	-0.001 0.0066	-	-0.0008 0.0068	-	0,0075** 0.0008	-	0.6008 0.2151
Tempo^2	-	-	-	0.0003 0.0002	-	-	-	0.0358 0.0065
SE(Tempo^2)	-	-	-	0 0.0001	-	-	-	0.0115 0.001
Tempo*Custo	-	-	-	-0.0807 0.0243	-	-	-	-0.2238 0.032
SE(Tempo*Custo)	-	-	-	0.0001 0.0064	-	-	-	0.0878 0.0138
Tempo_Reg	-	-	-0,0761** 0.012	-	-	-	-0,2027** 0.0228	-
SE(Tempo_Reg)	-	-	0.0034 0.0065	-	-	-	0,0308** 0.0045	-
Tempo_Irreg	-	-	-0,0611** 0.0111	-	-	-	-0,085** 0.0203	-
SE(Tempo_Irreg)	-	-	-0.0019 0.0039	-	-	-	0,0193** 0.0024	-
N	1564	1564	1564	1564	1564	1564	1564	1564
Pseudo R^2	0.0418	0.0419	0.0502	0.0512	-0.2448	-0.0162	-0.2351	-0.9738
RV	1.042	1.042	1.050	1.051	0.755	0.984	0.765	0.026
lnL	-982.44	-982.36	-973.84	-972.8727	-1244.8	-1041.9	-1266.4	-2023.8

Const=constante; Temp*SR=produto entre tempo e uma *dummy* para os indivíduos que não apresentaram renda fixa; Temp*1SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram um salário mínimo; Temp*1A2SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram de um a dois salários mínimos; Temp*2A5SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram dois a cinco salários mínimos; Temp*5A10SL=produto entre tempo e uma *dummy* para os indivíduos que apresentaram cinco a dez salários mínimos;

Custo=custo da viagem; Tempo=tempo da viagem; Tempo²=quadrado do tempo da viagem; Tempo*Custo=produto do tempo da viagem pelo custo; Tempo_Reg=tempo da viagem do regular; Tempo_Ireg=tempo da viagem do irregular. A renda base para o modelo McFadden é acima de dez salários mínimos.

Quando comparamos os sinais esperados frente à teoria concluímos que os modelos estimados pelo procedimento bayesiano ou por máxima verossimilhança simulada apresentam sinais coerentes com a teoria, exceto para o caso do modelo de Tuong e Hensher. Os valores de ajuste do modelo indicam a forma funcional de Becker como adequada para o caso de estimativa bayesiana enquanto que para o outro método seriam os de DeSerpa e Tuong e Hensher os de melhor resultado. Além disso, os valores de aderência para a estimativa de máxima verossimilhança simulada, pseudo R² de McFadden e o mínimo da função de máxima verossimilhança, sugerem que modelo de Tuong e Hensher apresenta melhor aderência, seguido pelo modelo de DeSerpa, Becker e McFadden. No caso dos modelos estimados por procedimento bayesiano possui valores negativos para o pseudo R². Para sanar este problema fez-se o cálculo da razão de verossimilhança e pode-se sugerir que os modelos estimados por máxima verossimilhança simulada possuem melhor aderência que o estimado por procedimento bayesiano. Referente às formas funcionais, pode-se indicar o modelo de DeSerpa como o mais aderente para o procedimento bayesiano.

Os resultados apresentam propriedades distintas e merecem um aprofundamento e teste na robustez dos resultados o que será feito na próxima seção calculando os valores de tempo de viagem.

5.3. MEDIÇÃO DOS VALORES DE TEMPO DE VIAGEM

Nesta seção apresentam-se os valores de tempo de viagem obtidos a partir das estimativas feitas na seção anterior.

5.3.1. Valor de Tempo de Viagem pela estimação do *logit* simples

Os valores de tempo de viagem para os modelos são apresentados na Tabela 5.5. O valor de tempo de viagem em McFadden foi obtido da razão pela média dos coeficientes estimados para a relação tempo renda pelo parâmetro de custo. Os outros modelos foram calculados conforme equações apresentadas na seção 2.3 com a ressalva que para o modelo de Tuong e Hensher foi utilizado o valor dos tempos e custos médios para diferenciar os valores de tempo de viagem entre modos.

Tabela 5.5: Valor de Tempo de viagem para os diferentes modelos em reais por hora.

	McFadden	Becker	DeSerpa	Tuong & Hensher
VTVReg	0.422	0.497	0.554	0.691
VTVIrreg	0.422	0.497	0.446	0.662

VTVReg= Valor do Tempo de Viagem para o transporte regular; VTVIrreg=Valor do tempo de viagem do Irregular;

No *logit* simples, os valores de tempo são iguais para os dois modos nas formas funcionais de McFadden e Becker enquanto os modelos de DeSerpa e Tuong e Hensher apresentam diferentes valores de tempo para os diferentes modos. Observa-se ainda que a diferença no valor de tempo de viagem para as diferentes formas funcionais escolhidas tem variação percentual da ordem de 15 a 30%. Retirando o modelo de Tuong e Hensher, os valores estão apresentados em termos absolutos, mas todos têm valor negativo, não gera prazer, ou seja, é um ônus aos viajantes, independentemente do modo escolhido. Vale à pena salientar que os valores são referentes ao ano de pesquisa 2005, e que neste período o salário mínimo era de 300 reais. Caso os valores fossem atualizados pelo índice geral de preços do mercado - IGP-M, os valores seriam aumentados em 38%, e caso fosse aplicado o INPC – Índice Nacional de Preços ao Consumidor, o aumento seria de 35%.

Além disso, o valor de tempo de viagem do irregular nos modelos de DeSerpa e Tuong e Hensher são menores que o regular indicando, possivelmente, uma maior aceitação do

transporte irregular frente ao regular. Adicionalmente, foi feito uma mensuração do valor do tempo para as diferentes faixas de renda utilizando o modelo de McFadden e expresso na Tabela 5.6, ocorrendo uma oscilação pequena no valor de tempo de viagem por faixa de renda. Assim, os indivíduos sem renda e com um salário mínimo apresentam valores de tempo da viagem em uma hora de 0.422 reais, indivíduos que ganham de um a dois salários com 0.597 reais, os que ganham de dois a cinco salários com 0.559 reais e os indivíduos que possuíam renda de 5 a 10 salários com valor no tempo de viagem de 0.368 reais. O último pode ser referente a pessoas de renda entre 5 a 10 salários, possivelmente, utilizarem o transporte coletivo em menor escala que os outros, portanto, valorizando pouco.

Tabela 5.6: Valor de Tempo de viagem para os diferentes níveis de renda no modelo McFadden em reais por hora.

	McFadden*SR	McFadden*1SL	McFadden*1a2SL	McFadden*2a5SL	McFadden*5a10SL
VTV	0,422	0,420	0,567	0,559	0,368

VTV= Valor do Tempo de Viagem

Não se apresenta o valor do tempo de viagem para o transporte regular e irregular, pois no caso do modelo de McFadden não existem percepções distintas pelos usuários entre o transporte regular e o irregular. Na sequência apresentamos o estudo com o valor de tempo de viagem para o modelo *logit* estimado com duas fontes de dados.

5.3.2. Valor de Tempo de Viagem pela estimação do *logit* com enriquecimento

Os valores de tempo para os modelos estimados com enriquecimento são apresentados na Tabela 5.7.

Tabela 5.7: Valor de Tempo para os diferentes modelos em reais por hora.

	McFadden	Becker	DeSerpa	Tuong & Hensher
VTVReg	0.467	0.499	0.553	0.507
VTVIrreg	0.467	0.499	0.446	0.510

VTVReg= Valor do Tempo de Viagem para o transporte regular; VTVIrreg=Valor do tempo de viagem do Irregular;

Os valores de tempo de viagem são mais elevados que os obtidos na estimativa simples. Os resultados refletem o fato que as duas fontes de dados não apresentam a mesma fonte geradora dos dados, então no máximo haveria uma distorção do valor obtido na estimativa simples. Portanto, os resultados com enriquecimento, possivelmente, não são os mais adequados para mensurar o valor do tempo de viagem.

A próxima seção apresenta o resultado dos valores de tempo de viagem para o caso de estimativas utilizando parâmetros aleatórios com diferentes formas de estimação.

5.3.3. Valor de Tempo de Viagem pela estimação do *logit* com parâmetros aleatórios

Os valores de tempo de viagem para os modelos de parâmetros aleatórios são apresentados na Tabela 5.8. e comparados ao resultado de uma estimativa simples apresentando resultados bem semelhantes para o caso da MVSL e bayesiano, mesmo que para o caso bayesiano os valores marginais das utilidades indiretas tenham sido maiores.

A exceção aos resultados é a forma funcional proposta por Tuong e Hensher que foi totalmente diferente para ambas as estimações da obtida no modelo simples. Vale salientar que o caso de Tuong e Hensher pode ser inerente a esta forma funcional possuir combinações entre variáveis e conseqüente multicolinearidade, resultando em valores de parâmetro grande com pouca ou nenhuma significância estatística.

Tabela 5.8: Valor de Tempo para os diferentes modelos em reais por hora.

		McFadden	Becker	DeSerpa	Tuong & Hensher
VTVReg	MVSL	0.462	0.490	0.550	2.769
VTVIrreg		0.462	0.490	0.442	1.052
VTVReg	Bayesiana	0.264	0.454	0.765	2.769
VTVIrreg		0.264	0.454	0.321	3.614

VTVReg= Valor do Tempo de Viagem para o transporte regular; VTVIrreg=Valor do tempo de viagem do Irregular; MVSL= Máxima Verossimilhança Simulada.

Optou-se por apresentar os resultados segundo a média do parâmetro, não incorporando o desvio padrão, uma vez que para a estimativa via máxima verossimilhança simulada resultou em o estimador dos desvios padrões dos parâmetros é não significativo o que não aconteceu na estimativa bayesiana. Destaque-se que ao se assumir o custo marginal como um parâmetro não aleatório tem-se uma medição mais fácil do valor de tempo e evita-se uma distribuição de probabilidade desconhecida.

Na seção 5.4, seguinte, faz-se uma comparação com os resultados encontrados nos modelos com diferentes formas de estimação.

5.4. COMPARAÇÃO ENTRE OS MODELOS

Os resultados encontrados nas estimações serão comparados nesta seção tomando como base o modelo *logit* simples uma vez que este foi e é largamente utilizado para mensurar valor de tempo de viagem. Foram comparadas aos pares as estimações com a estimação simples e os as formas funcionais adotadas em cada procedimento de estimação. Pode ser resumido assim:

- McFadden simples com o McFadden com enriquecimento;
- McFadden simples com o McFadden com fatores aleatórios estimado por MVSL;
- McFadden simples com o McFadden com fatores aleatórios estimado de forma bayesiana;
- Comparação entre as formas funcionais de McFadden, Becker, DeSerpa e Tuong e Hensher nas diferentes formas estimação utilizadas, *logit* simples, com enriquecimento e com fatores aleatórios para o caso de MVSL e bayesiano.

Na Tabela 5.9 os valores positivos significam que o valor de tempo de viagem estimado por *logit* simples é menor que o do processo de estimação que está sendo comparado e o negativo que é maior. A comparação entre os processos de estimação, conforme observado na Tabela 5.9

tomando como base a estimativa feita com o *logit* simples indica que o valor do tempo de viagem de um modelo com a forma funcional tal como a de Becker é estável e que indiferente ao processo de estimações ocorre apenas um escalonamento do valor marginal do tempo. Esta estabilidade não se replica aos outros modelos. No entanto, observa-se uma oscilação mais acentuada da estimativa dos valores de tempo de viagem para os modelos de McFadden e Tuong e Hensher.

Tabela 5.9: Comparação entre diferentes métodos de estimação para diferentes modelos do Valor de Tempo de viagem.

		McFadden	Becker	DeSerpa	Tuong & Hensher
VTVReg	% simples versus com enriquecimento	10%	0%	0%	-36%
VTVIrreg		10%	0%	0%	-30%
VTVReg	% simples versus MVSL	9%	-1%	-1%	75%
VTVIrreg		9%	-1%	-1%	37%
VTVReg	% simples versus bayesiana	-60%	-9%	28%	75%
VTVIrreg		-60%	-9%	-39%	82%

VTVReg= Valor do Tempo de Viagem para o transporte regular; VTVIrreg=Valor do tempo de viagem do Irregular; MVSL= Máxima Verossimilhança Simulada.

Embora os valores de tempo de viagem obtidos tenham sido muito diferentes, o modelo com enriquecimento de dados apresentou pouca variação em valor de tempo de viagem frente ao modelo *logit* simples. Nota-se ainda pouca oscilação, até mesmo menor que o modelo de enriquecimento de dados, para os valores de tempo de viagem obtidos da estimação por MVSL, exceto para o modelo de Tuong e Hensher. A estimativa bayesiana é a que apresentou maior oscilação de valor de tempo de viagem quando comparado com o modelo *logit* simples.

Em muitas das vezes a estimativa bayesiana estima valores de tempo de viagem menores que o do *logit* simples e o com enriquecimento aumenta o valor do tempo de viagem. O valor do tempo de viagem do modelo de Tuong e Hensher para os modelos de fatores aleatórios são sempre menores que o estimado por *logit* simples.

Uma comparação das formas funcionais sugere que a proposta de Tuong e Hensher tende a apresentar valores de tempo de viagem maiores que as outras formas funcionais indiferentes ao

processo de estimação. O modelo de DeSerpa tende a fazer uma compensação entre as atividades com valores de tempo de viagem distinto entre modos e complementares, indicando uma substituição entre modos.

CONSIDERAÇÕES FINAIS

Os resultados indicam que se devem utilizar as formas funcionais e os métodos de estimação com parcimônia, pois dependendo dos parâmetros obtidos, podem ser encontrados valores para tempo de viagem consideravelmente diferente. Por outro lado, as formas funcionais podem ser boas para previsão tal como discutido acima.

A forma funcional de Tuong e Hensher produz valores de tempo de viagem maiores que as outras formas funcionais e não é de mensuração adequada com qualquer procedimento para estimar o modelo com fatores aleatórios, pois distorce os valores obtidos. Entretanto, caso seja utilizado para previsão o modelo de Tuong e Hensher foi o que se mostrou mais ajustado com os dados, exceto para o caso bayesiano.

Uma escolha entre métodos de estimação sugere o enriquecimento como o que menos viesas as estimativas. Uma escolha de forma funcional indica o uso ou do modelo de Becker ou do DeSerpa por conta da sua pequena oscilação independentemente dos métodos de estimação utilizados.

Uma combinação entre o método de estimação com forma funcional sugere o modelo de Becker e DeSerpa com MVSL, tanto para obter valores de tempo de viagem como para prever eventos.

CAPÍTULO 6. – CONCLUSÕES

Os aspectos teóricos da tentativa de mensurar o insumo tempo surgiram em 1965, mas se alastra até a atualidade, ficando as seguintes respostas em aberto:

1. Como as características das pessoas influenciam no valor do tempo, níveis diferentes de renda influenciam no valor de tempo e como?
2. O local de habitação afeta na percepção de quanto vale o tempo?

A aplicação de métodos de medição mais fundamentada surgiu, somente, após a década de 1980, entretanto, sem considerar aspectos idiossincráticos dos indivíduos e apresentando apenas fatos estilizados para as perguntas em aberto na teoria.

Por outro lado, ao estudar a escolha dos usuários de transportes intermunicipais no Ceará utilizando diferentes formas de estimação da função utilidade indireta (estruturados e não estruturados) e verificando os valores dos tempos de viagem, pretendia-se entender como o tempo influi na decisão do usuário. Isso impactaria primordialmente na competição entre o transporte regular e o irregular, sendo passível avaliar os ganhos em redução de tempo do regular e quando dado este ganho se intervenções operacionais resultariam em redução da demanda pelo transporte clandestino. Isso, por si reduziria as externalidades negativas imputadas pelo transporte irregular à sociedade tais como, acidentes, sonegação de impostos e inviabilização do transporte regular.

O motivador é que em transportes, diferentemente de outros bens da economia, o tempo deve ser um atributo importante na decisão dos usuários e, portanto a concorrência entre modos pode se dar nesse nível tal como apresentado em Granou (1970). Economistas desde 1965 se debruçam em como mensurar o valor do tempo seja de forma teórica ou de forma empírica.

Este trabalho aborda empiricamente o que na teoria foi proposto particularmente para o caso dos transportes. A medição do valor de tempo uniu teoria e medição empírica com o trabalho de Tuong e Hensher (1985), entretanto, com o passar dos anos os métodos diversos de estimação por modelos de escolha discreta não foram utilizados para a medição do valor de tempo de viagem o que é feito neste trabalho com uma abordagem utilizando diferentes fontes de dados e um modelo com fatores aleatórios.

Esta tese estudou a teoria econômica do valor do tempo, as propostas de estimação com diferentes métodos e formas funcionais para avaliar quantitativamente o valor do tempo de viagem, em particular, o do transporte regular e irregular no Ceará. Conhecendo o valor do tempo de viagem para os usuários seria possível a intervenção política do Estado via redefinição da programação do transporte regular, sabendo em quanto um transporte pode interferir no outro.

Como conclusão encontrou-se que o valor do tempo é inerente aos aspectos das atividades executadas e o nível de renda dos indivíduos sinaliza como o tempo é valorizado.

Do ponto de vista empírico, conclui-se que os estudos empíricos de valor de tempo iniciaram em 1965 com Beesley e outros apresentados em McGillivray (1970), passando por Train e McFadden (1978), alcançando uma postura bem mais robusta em Tuong e Hensher (1985a e b), sendo posteriormente melhorado em Blayac e Causse (2001) e Hammadou e Jayet (2002).

Por seu turno, os métodos de estimação de modelos de escolha discreta evoluíram o método de enriquecimento de dados proposto em Morikawa (1989), pode gerar ganhos em estimação

pontual, mas dependerá da função geradora dos dados das diferentes fontes serem compatível conforme salientado em Hensher, Louviere e Swait (2000).

Já os modelos com fatores aleatório são importantes para descrever a idiosincrasia dos usuários, entretanto, a escolha do método de estimação é complicado e, pode de certo modo, apresentar diferenças dado as propriedades das propostas de estimação por máxima verossimilhança simulada ou por estimativa bayesiana, podendo levar a resultados que destoem significativamente dependendo, inclusive, do modelo estrutural adotado.

Além disso, as técnicas, por utilizarem um esforço computacional maior que os métodos convencionais, podem carrear erros numéricos que podem intensificar problemas inerentes às formas funcionais utilizadas. De qualquer sorte, os métodos podem revelar informações importantes e melhorar o conhecimento do comportamento dos usuários para intervenções políticas mais acertadas, ou, no mínimo, com menor variabilidade.

O resultado indica que se devem utilizar os métodos de estimação com parcimônia, pois dependendo dos parâmetros obtidos, podem ser encontrados valores para tempo de viagem consideravelmente diferente. Por outro lado, as formas funcionais podem ser boas para previsão tal como discutido acima. Uma escolha entre métodos de estimação sugere o enriquecimento como o que menos viesas as estimativas. Uma escolha de forma funcional indica o uso ou do modelo de Becker ou do DeSerpa por conta da sua pequena oscilação independentemente dos métodos de estimação utilizados. Já uma escolha combinada entre método de estimação e forma funcional sugere o modelo de Becker e DeSerpa com MVSL, tanto para obter valores de tempo de viagem como para prever eventos.

APÊNDICE

Apêndice 1: exemplos de estimação dos modelos com enriquecimento no R.

Modelo segundo Becker

Hipóteses de retornos marginais em tempo e custo iguais

```
logitd<-function (pars, y){
b1<-pars[1]
b2<-pars[2]
b3<-pars[3]
b4<-pars[4]
v1<-b1*y[,11]+b2*y[,4]+b3*y[,5]+b4*y[,6]
v2<-b3*y[,7]+b4*y[,8]
logl<-sum(y[,2]*log((1)/(exp(v2-v1)+1))+(1-y[,2])*log((1)/(1+exp(v1-v2))))
return(-logl)
}
```

Misto

```
logitd<-function (pars,y){
b1<-pars[1]
b2<-pars[2]
b3<-pars[3]
b4<-pars[4]
mu<-pars[5]
v1<-b1*y[,5]+b2*y[,6]+b3*y[,7]+b4*y[,8]
v2<-b3*y[,9]+b4*y[,10]
v3<-b1*y[,5]+b2*y[,6]
v4<-b3*y[,7]
FSP<-y[,4]*(y[,3]*log((1)/(exp(mu*(v2-v1)+1)))+(1-y[,3])*log((1)/(1+exp(mu*(v1-v2))))))
FRP<-(-1-y[,4])*(y[,3]*log((1)/(exp(v4-v3)+1)))+(1-y[,3])*log((1)/(1+exp(v3-v4))))
logl<-sum(FSP+FRP)
return(-logl)
}
```

Otimo

```
logitd<-function (pars, y){  
b1<-pars[1]  
b2<-pars[2]  
b3<-pars[3]  
b4<-pars[4]  
v1<-b1*(y[,9]+y[,10])+b2*y[,4]+b3*y[,5]+b4*y[,6]  
v2<-b3*y[,7]+b4*y[,8]  
logl<-sum(y[,2]*log((exp(v1))/(exp(v2)+exp(v1)))+(1-  
y[,2])*log((exp(v2))/(exp(v2)+exp(v1))))  
return(-logl)  
}
```

Modelo segundo Hensher tem a hipótese de retornos marginais de tempo diferentes, mas retornos em custo iguais

```
logitd<-function (pars, y){  
b1<-pars[1]  
b2<-pars[2]  
b3<-pars[3]  
b4<-pars[4]  
b5<-pars[5]  
v1<-b1*y[,11]+b2*y[,4]+b3*y[,5]+b4*y[,6]  
v2<-b5*y[,7]  
logl<-sum(y[,2]*log((1)/(exp(v2-v1)+1))+(1-y[,2])*log((1)/(1+exp(v1-v2))))  
return(-logl)  
}
```

Apêndice 2: exemplos do procedimento de cálculo das estatísticas dos parâmetros no R.

```
resM<-optim(rep(0.001,5),logitM,y=KKK,x=KKK, method="BFGS",hessian=T)
```

Resultados:

```

parD<-resD$par
parD<-c(parD,0)
parR<-resR$par
parR<-c(parR,0,0)
parM<-resM$par

```

```
%%Apresentação dos parâmetros
```

```
CompPar<-matrix(c(parD,parR,parM), nr=3,nc=8,byrow=T)
```

```
CompPar
```

```

      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
Declarada -10.7761577 0.031302041 0.0007218993 0.06229571 0.12507530 -0.041851330
12.059360 0.000000
Revelada -0.4206448 -0.044927428 -0.0002961624 0.09473969 -0.09972380 0.008717964
0.000000 0.000000
Mix 0.2644284 -0.004599405 -0.0001886825 -0.01619926 -0.03197993 0.004349358 -
0.425201 -10.93368

```

```
%%Apresentação das estatísticas ts dos parâmetros
```

```

HsD<-resD$hessian
HsR<-resR$hessian
HsM<-resM$hessian
OID<-solve(HsD)
OIR<-solve(HsR)
OIM<-solve(HsM)
seD<-sqrt(diag(OID))
seR<-sqrt(diag(OIR))
seM<-sqrt(diag(OIM))
seD<-c(seD,0)
seR<-c(seR,0,0)
tsD<-parD*(1/seD)

```

```

tsR<-parR*(1/seR)
tsM<-parM*(1/seM)
Compts<-matrix(c(tsD,tsR,tsM),nr=3,nc=8,byrow=T)
Compts
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
Declarda -14.257130 0.3116677 0.2025151 0.586338 1.123010 -13.499839 16.69608
NaN
Revelada -2.588195 -0.6350637 -0.1143918 1.240779 -1.406192 17.384302 NaN
NaN
Mix      4.385704 -0.5254984 -0.5928876 -1.572659 -2.650050 6.922873 -11.96669 -
6.085652

```

Apêndice 3: exemplo de formulário de questionário utilizado na pesquisa.

QUESTIONÁRIO

Macrorregião: <input type="checkbox"/>	No. Formulário: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
Município: _____ Local: _____	Data: ____/____/____
Pesquisador: _____	Hora: ____h ____min
Supervisor: _____	

N1. PESQUISA SOCIOECONÔMICA

1.1 Sexo: [] Masculino [] Feminino	1.2 Idade: ____ anos		
1.3 Reside na Cidade de: _____			
1.4 Grau de Instrução:			
[] Sem escolaridade	[] Alfabetizado	[] 1º Grau Incompleto	[] 1º Grau Completo
[] 2º Grau Incompleto	[] 2º Grau Completo	[] Superior Incompleto	[] Superior Completo
[] Pós-graduação			
1.5 Renda:			
[] Sem rendimento	[] Até 1 SM (R\$ 300)	[] De 1 a 2 SM (R\$ 600)	
[] De 2 a 5 SM (R\$ 1.500)	[] De 5 a 10 SM (R\$ 3.000)	[] Acima de 10 SM	
1.6 Situação Familiar:			
[] Chefe de Família	[] Cônjuge	[] Filho(a)	[] Outro parentesco

1.7 Ocupação:

[] Estudante	[] Dona de casa	[] Assalariado	[] Autônomo
[] Empresário	[] Aposentado	[] Desempregado	[] Outra

1.8 Disponibilidade de Automóvel em casa:

Zero 1 (um) 2 (dois) Mais de dois

1.9 Disponibilidade de Motocicleta em casa:

Zero 1 (uma) 2 (duas) Mais de duas

2. PADRÕES DE VIAGENS

2.1 Local de Origem da Viagem: _____ **Local de Destino da Viagem:** _____

2.2 Tempo estimado de viagem _____

2.3 Tempo estimado de espera pela condução _____

2.3 Tempo de caminhada até o local de espera do ônibus: _____

2.4 Tipo de Transporte Atualmente Utilizado:
 Ônibus / microônibus regular Transp. complementar (topic ou camioneta) Automóvel Particular

2.5 Tipo de Transporte Mais Utilizado:
 Ônibus / microônibus regular Transp. complementar (topic ou camioneta) Automóvel Particular

2.6 Frequência de viagem:

Diária De 2 a 5 viagens por semana Semanal Quinzenal
 Mensal Bimestral Semestral Anual 1ª Viagem

2.5 Motivos da Viagem:

Trabalho Escola Compras Saúde Lazer Outros

3. QUESTIONÁRIO DE PD

ÔNIBUS						VAN					
CARTAO	ALTERNATIVA										
1A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	1B	<input type="checkbox"/> 1	<input type="checkbox"/> 2	1A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	1B	<input type="checkbox"/> 1	<input type="checkbox"/> 2
2A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	2B	<input type="checkbox"/> 1	<input type="checkbox"/> 2	2A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	2B	<input type="checkbox"/> 1	<input type="checkbox"/> 2
3A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	3B	<input type="checkbox"/> 1	<input type="checkbox"/> 2	3A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	3B	<input type="checkbox"/> 1	<input type="checkbox"/> 2
4A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	4B	<input type="checkbox"/> 1	<input type="checkbox"/> 2	4A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	4B	<input type="checkbox"/> 1	<input type="checkbox"/> 2
5A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	5B	<input type="checkbox"/> 1	<input type="checkbox"/> 2	5A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	5B	<input type="checkbox"/> 1	<input type="checkbox"/> 2
6A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	6B	<input type="checkbox"/> 1	<input type="checkbox"/> 2	6A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	6B	<input type="checkbox"/> 1	<input type="checkbox"/> 2
7A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	7B	<input type="checkbox"/> 1	<input type="checkbox"/> 2	7A	<input type="checkbox"/> 1	<input type="checkbox"/> 2	7B	<input type="checkbox"/> 1	<input type="checkbox"/> 2

4. OBSERVAÇÃO

REFERÊNCIAS BIBLIOGRÁFICAS

- Abrantes, P. A. L., Wardman, M. R. (2011), *Meta-analysis of UK values of travel time: An update*. Transportation Research Part A 45 1–17
- Anselin, L., Raymond, J.G.M, Sergio, J. Ray, (2004), *Advances in Spatial Econometrics: Methodology, Tools, and Applications*, Springer.
- Araujo, C. E. F. ; Soares, C. H ; Silva, F. G. F. . *Análise espacial da geração de viagens no transporte público intermunicipal de passageiros do estado do Ceará*. In: pluris 2006 - Congresso Luso Brasileiro para o Planeamento Urbano, Regional, Integrado e Sustentável, 2006, Braga - Portugal, 2006.
- Arruda (1996) *Valor do Tempo de Viagem para Avaliação de Projetos de Transportes no Brasil: Um Estudo Crítico-Comparativo*. In: Confederação Nacional do Transporte. (Org.). Transporte em Transformação: Trabalhos Vencedores do Prêmio CNT. 1a ed. São Paulo: Makron Books do Brasil Editora Ltda., 1997, v. I, pp. 52-71.
- Bates, J.T., (1987) *Measuring Travel Time Values with a Discrete Choice Model: A note*. The Economic Journal, Vol. 97 No. 386 (Jun. 1987), pp. 493-498.
- Becker, G. S., (1965). *A Theory of the Allocation of Time*. The Economic Journal, Vol. 75, No. 299 (Sep., 1965), 493-517.
- Ben-Akiva, M. E., (1973) *Structure of Travel Demand Model*, Ph.D. Dissertation, Department of Civil Engineering, MIT.
- Ben-Akiva, M. E , Bierlaire, M., (1999) *Chapter 2: Discrete Choice Models with Applications to Departure Time and Route Choice*. Handbook of Transportation Science. 2nd ed. International Series in Operations Research & Management Science, Vol. 56. New York: Springer, 2003.ISBN 1-4020-7246-5
- Ben-Akiva. M. e T. Morikawa (1990) *Estimation of Switching Models from Revealed Preferences and Stated Intentions*. Transportation Research, 24A, 485-495.
- Ben-Akiva, M, Bolduc, D. e Bradley, M (1993), *Estimation of travel model choice models with randomly distributed values of time*, Transportation Research Record 1413, 88-97.
- Bhat, C.R. (2000), *Incorporating Observed and Unobserved Heterogeneity in Urban Work Mode Choice Modeling*, Transportation Science, Vol. 34, No. 2, pp. 228-238.
- Bhat, C., Guo, J., (2002), *A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modelling*, Manuscript number 122-02.
- Bhat, C. R. e Koppelman F., (1999) *A retrospective and prospective survey of time-use research*. Transportation 26: 119-139
- Blayac, T and Causse, A. (2001) *Value of travel time: a theoretical legitimization of some nonlinear representative utility in discrete choice models*. Transportation Research Part B 35 (2001) 391±400
- Brownstone, D. e K. Train (1999) *Forecasting New Product Penetration with Flexible Substitution Patterns* Journal of Econometrics, 89, 109-129.
- Cherchi E., e Ortúzar J. de D. (2002) *Mixed RP/SP models incorporating interaction effects: Modelling new suburban train services in Cagliari*. Transportation, 29, 4, 371-395
- Cherchi, E. e Ortúzar, J. de D. (2008) *Predicting best with mixed logit models: understanding some confounding effects*. En P.O. Inweldi (ed.), Transportation Research Trends. Nova Science Publishers, Nueva York.
- DeSerpa, A. C., (1971). *A Theory of the economics of time*, Economic Journal, December, 828-846.
- DeSerpa, A. C., (1972). *Microeconomic Theory and the valuation of Travel Time: Some Clarification*. Regional and Urban Economics – Vol. 2, No. 4 pp. 401-410.
- Donnea, F.X., (1971) *Consumer Behaviour, transport Mode Choice and Value of Time: Some Micro-Economic Models*. Regional and Urban Economics – Vol, 1, No. 4 355-382.
- Evans, S., 1972. *On the theory of the valuation and allocation of time*. Scottish Journal of Political Economy 19, 1–17.
- Golub, A., Balassiano, R., Araújo, A., Ferreira, E. (2009) *Regulation of the Informal Transport Sector in Rio de Janeiro, Brazil: Welfare Impacts and Policy Analysis*. Transportation, Vol. 36, pp.601-616.
- Gronau, Reuben, (1970) *The Value of Time in Passenger Transportation: The Demand for Air Travel*, National Bureau of Economic Research, New York, Occasional Paper No. 109.
- Hajivassiliou, V.A. e Ruud, P. A. (1993). *Classical Estimation Methods for LDV Models Using Simulation* Econometrics 9311002, EconWPA.
- Hammadou, H and Jayet, H, (2002), *La Valeur du temps pour les déplacements à longue distance: une evaluation sur données françaises*. Les cahiers Scientifiques du Transport. N° 42/2002- pages 3-23.
- Hensher, D. e M. Bradley (1993) *Using Stated Response Data to Enrich Revealed Preference Discrete Choice Models* Marketing Letters, 4, 39-152.
- Hensher, D., J. Louviere, e J. Swait (2000) *Stated choice methods: analysis and applications*. Cambridge University.
- Hensher, D., J. Louviere, e J. Swait (1999) *Combining Sources of Preference Data* Journal of Econometrics, 87, 97-221.

- Huq, M. (2010) *Theoretical Analysis on Valuation of Time Saving*. Mimeo.
- Jara-Díaz, S.R. e Ortúzar, J.D., (1986) *Introducing The Expenditure Rate In The Estimation of Mode Choice Models*, Journal of Transport Economics and Policy, Volume_XX111_No_3_293-308.
- Jara-Díaz, S.R. (2000) *Allocation and Valuation of Travel-Time Savings*. *Handbook of Transport Modelling*, Edited by D.A. Hensher and K.J. Button, Elsevier Science Ltda.
- Jiang, M, Morikawa, T., (2004) *Theoretical Analysis on the variation of value of travel time savings*. Transportation Research Part A 38 551-571.
- Johnson, m. B. (1966), *Travel time and the price of leisure*. Economic inquiry, 4: 135–145.
- Lancaster, K. J. (1966) *A New Approach to Consumer Theory*. Journal of Political Economy, 74, 132-157.
- LeSage (1999), *The Theory and Practice of Spatial Econometrics*, mimeo.
- Luce, R.(1959). *Individual Choice Behavior. A Theoretical Analysis*. Wiley, New York.
- Luce, R. e Suppes, P. (1965). *Preference, Utility and Subjective Probability*. In Handbook of Mathematical Psychology. Vol 3. R. Luce, R. Bush, and E. Galanter, eds. Wiley, New York.
- Klier, T. e McMillen, D. P., (2006) *Clustering of Auto Supplier Plants in the U.S.: GMM Spatial Logit for Large Samples*. Mimeo.
- MVA consultancy, Institute for transport studies at Leeds University, Transport studies at Oxford University, (1994) *Time saving: research into the value of time*. In: Layard, R., Glaister, S., (Eds.). Cost-Benefit Analysis. Cambridge, pp. 205–211.
- McFadden, D. L , (1974) *The Measurement of Urban Travel Demand*, Jornal of Public Economics.
- McFadden, D. L., (2000a) *Economic Choice – Prize Lecture*, December 8, 2000 – Economic Science.
- McFadden, D.L. (2000b) *Disaggregate Behavioral Travel Demand's RUM Side A 30-Year Retrospective*. - presentation at the International Association of Travel Behavior Analysts, Brisbane, Australia, July 2, 2000, and at RU2000, Duke University, Aug. 6, 2000.
- McGillivray, R.G., (1970) *Demand And Choice Models of Modal Split*, Journal of Transport Economics and Policy, Volume_1V_No_2_192-207.
- Mackie, Peter, S. R. Jara-Díaz and T. Fowkes (2001) *The value of travel time savings in evaluation*. Transportation Research 37 E, 91-106.
- Morikawa, T. (1989) *Incorporating Stated Preference Data in Travel Demand Analysis*, Ph.D. Dissertation, Department of Civil Engineering, MIT.
- Morikawa, T., M. Ben-Akiva, K. Yamada (1991) *Forecasting Intercity Rail Ridership Using Revealed Preference and Stated Preference Data*. Transportation Research Record, 1328, 30-35.
- Oort, C. J., (1969) *The Evaluation of Travelling Time*, Journal of Transport Economics and Policy.
- Ortúzar, J. D. e Willumsen, L. G. (1997) *Modelling Transport – Second Edition – John Wiley e sons*.
- Pinkse J. e Slade, M.E., (1998), *Contracting in Space: An application of spatial statistics to discrete-choice models*. Journal of Econometrics, 85:125-154.
- Poduje, M. A. S., (2002) *Estimacion de Valores de Disposicion a Pagar con Modelos logit Mixto*. Pontificia Universidad Catolica de Chile Escuela de Ingenhiera
- PDOTIP, (2005), Relatório Final.
- R Development Core Team (2009). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Revelt, D. and K. Train (1998), *Mixed logit with repeated choices*, Review of Economics and Statistics 80, 647–657.
- Small, K.A., (1982) *Scheduling of consumer activities: work trips*. American Economic Review 72, 467–479.
- Swait, J., Louviere, J. and Williams, M. (1994) *A Sequential Approach to Exploiting the Combined Strengths of SP and RP Data: Application to Freight Shipper Choice*. Transportation, 21:135-152.
- Thurstone, L. (1927) *A Law of Comparative Judgement*. Psychological Rev. 34: 273-286.
- Train, K. E. and McFadden, (1978), *The Goods/Leisure Tradeoff and Disaggregate Work Trip Mode Choice Models*. Transportation Research, Vol. 12, No. 5, 349-353.
- Train, K.E., McFadden, D., Ben-Akiva, M. (1987) *Recreation The demand for local telephone service: A fully discrete model of residential calling patterns and service choice*. Rand Journal of Economics 18, 109-123
- Train, K.E., (1998) *Recreation Demand Models with Taste Differences Over People*. Land Economics, Vol. 74, No. 2.
- Train, K. E., (2003), *Discrete Choice Models with Simulation*, Cambridge Press.
- Truong, P and Hensher D.A., (1985a) *Measurement of Travel Time Values and Opportunity Cost from a Discrete-Choice Model*. The Economic Journal. Vol. 95, No. 378 (Jun., 1985), pp.438-451.
- Truong, P and Hensher D.A., (1985b) *A Valuation of Travel Time Savings A Direct Experimental Approach*. Journal of Transport Economics and Policy. Vol. 19, pp.237-261.
- Vichiensan, V. (2005) *Mixed Logit Model With Structuralized Spatial Effects*. Mimeo.