

ELAINE CRISTINA SCHNEIDER DE CARVALHO

Comportamento de escolha de linha de ônibus sob a influência de  
painéis eletrônicos com previsões em tempo real sobre  
a chegada dos veículos aos pontos

São Paulo  
2013

ELAINE CRISTINA SCHNEIDER DE CARVALHO

Comportamento de escolha de linha de ônibus sob a influência de  
painéis eletrônicos com previsões em tempo real sobre  
a chegada dos veículos aos pontos

Dissertação apresentada à Escola Politécnica  
da Universidade de São Paulo para obtenção  
do Título de Mestre em Ciências

São Paulo  
2013

ELAINE CRISTINA SCHNEIDER DE CARVALHO

Comportamento de escolha de linha de ônibus sob a influência de  
painéis eletrônicos com previsões em tempo real sobre  
a chegada dos veículos aos pontos

Dissertação apresentada à Escola Politécnica  
da Universidade de São Paulo para obtenção  
do Título de Mestre em Ciências

Área de Concentração:  
Engenharia de Transportes

Orientador: Professor Titular Orlando Strambi

São Paulo  
2013

**Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.**

**São Paulo, 15 de outubro de 2013.**

**Assinatura do autor** \_\_\_\_\_

**Assinatura do orientador** \_\_\_\_\_

## **FICHA CATALOGRÁFICA**

**Carvalho, Elaine Cristina Schneider de**

**Comportamento de escolha de linha de ônibus sob a influência de painéis eletrônicos com previsões em tempo real sobre a chegada dos veículos aos pontos / E.C.S. de Carvalho. -- versão corr. -- São Paulo, 2013.**

**128 p.**

**Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Transportes.**

**1.Transporte público 2.Ônibus 3.Painéis 4.Informação 5.Demanda (Modelos) I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Transportes II.t.**

À descoberta de que o essencial nos transforma.

## **AGRADECIMENTOS**

A meu bem humorado, entusiasmado e exigente orientador, Professor Orlando Strambi, por me ensinar a dar os primeiros passos no universo da pesquisa acadêmica e por me proporcionar experiências de aprendizado ímpares.

Ao Professor Luis Ignacio Rizzi, pela intensa e valiosa colaboração com minha pesquisa, e ao colega David Palma Araneda, pela cooperação sucinta e precisa, ambos do Departamento de Engenharia de Transporte e Logística da Escola de Engenharia da Universidade Católica do Chile.

Aos Professores Claudio Barbieri da Cunha, Nicolau Dionísio Fares Gualda e José Afonso Mazzon, pelo interesse constante.

Agradeço também, pelo apoio e cooperação, a Amanda Rabelo Santos Luiz e Rosana Simone Vieira da Silva, do Serviço de Comunicação Social da Escola Politécnica da Universidade de São Paulo, às colegas pesquisadoras Ana Margarita Larrañaga e Patrícia Sauri Lavieri, à Secretaria do Departamento de Engenharia de Transportes da Escola Politécnica, à CAPES e ao CNPq.

À amiga Glaucia Guimarães Pereira, pelo carinho, cuidado e apoio imprescindíveis. A Bianca Bianchi Alves, pela amizade, acolhimento e apoio, desde o começo. A ambas, pela vontade e disponibilidade para discutir ideias e analisar criticamente meus avanços.

Aos meus primos Jussara, Paulo e Denise, e à minha tia Elisabete, por estarem ao meu lado em momentos difíceis. E aos trabalhadores e mentores do Núcleo Espírita Nosso Lar, em Florianópolis, pelo cuidadoso e amoroso processo de cura.

A Gabriel Führ, Ana Lúcia Rodrigues e colegas da Sistran Engenharia, por tornarem mais suave minha vinda para São Paulo.

Aos amigos Rosiani Soares, Cassiano Isler, Juliana Stallivieri, Andreia Chiavini, Rodrigo Serratine e Fabiana Pessôa, pela sincera amizade, que torna minha vida mais alegre.

Aos meus pais, Juarez e Marcília, por apoiarem, incentivarem e ajudarem a viabilizar minhas escolhas. Ao meu irmão Rodrigo, pelo exemplo de perseverança.

## RESUMO

Esta dissertação teve como objetivo principal investigar a influência, no comportamento de escolha de linha de passageiros de ônibus, de painéis eletrônicos em pontos de parada com informação em tempo real sobre previsões para passagem dos veículos. Para a coleta de dados, um experimento de escolha declarada com desenho eficiente foi aplicado a uma amostra de 1179 entrevistados residentes na Região Metropolitana de São Paulo (RMSP) e pertencentes à comunidade da Universidade de São Paulo. Em cada situação de escolha foram apresentadas duas linhas diferentes de ônibus que chegariam ao mesmo destino, mas não passavam pelo mesmo ponto, de modo que a escolha a ser feita era entre uma ou outra combinação de linha e ponto. Adicionalmente, apenas um dos pontos tinha painel eletrônico. As alternativas também se diferenciavam pelos seguintes atributos: intervalo de tempo programado para a passagem de veículos consecutivos da mesma linha; possível atraso em relação ao intervalo programado; tempo de viagem dentro do veículo até o destino; ocupação do veículo quando chega ao ponto; e valor da tarifa. Na amostra predominaram jovens com até 25 anos (64% da amostra), homens (60%), usuários frequentes de ônibus (80%), estudantes (81%) e entrevistados com pelo menos um automóvel no domicílio (76%). A partir das respostas ao experimento foram estimados modelos de escolha discreta *Mixed Logit Panel*, de modo a mensurar a importância relativa de cada atributo na decisão e também medir a variabilidade das preferências entre os entrevistados. Os resultados indicam que a presença do painel no ponto de ônibus tem, sim, influência sobre a escolha da linha. Os entrevistados estariam dispostos a pagar em média, pela presença de painel, R\$0,12 adicionais, equivalentes a 5 minutos de viagem. Verificou-se também que a existência de painel no ponto diminui a desutilidade marginal da espera, e isto ocorre com mais intensidade quando ela está associada ao atraso do que quando está associada ao intervalo programado entre veículos. O valor médio do tempo de viagem foi relativamente baixo: R\$1,44/hora, provavelmente devido à composição socioeconômica da amostra, com elevada proporção de estudantes. No entanto, observou-se que o comportamento de escolha de linha é bastante afetado pelas características socioeconômicas e de uso de ônibus dos entrevistados, podendo o valor do tempo chegar a R\$17,00/hora, e a disposição a pagar pelo painel a R\$0,77. Acredita-se que os resultados desta pesquisa permitem ampliar o entendimento do comportamento de escolha de linha, ao incorporar a presença de painel no ponto como elemento adicional da decisão.

Palavras-chave: Transporte público. Ônibus. Painéis. Informação. Demanda (Modelos).

## ABSTRACT

The main objective of this research is to investigate the influence on bus route choice behavior of variable message signs (VMS) displaying real time predictions of bus arrival at stops. A stated choice survey was conducted, using an efficient design experiment. Sampled individuals were asked to answer to eight choice situations, each presenting two bus routes going to the same destination but with different itineraries and boarding stops. The choice was made between two combinations of bus route and boarding stop; only one of the stops had VMS. The other attributes characterizing alternatives were: bus route headway, (possible) delay at arriving at the stop, travel time until destination, level of vehicle crowdedness when arriving at the boarding stop, and fare. Data were collected from 1179 individuals, mostly students, professors and employees of the University of São Paulo, and all of them residents of the São Paulo Metropolitan Area. The typical interviewee was 25 years old or younger (64% of the sample), male (60%), a frequent bus user (80%), student (81%) and had at least one car in his household. Mixed logit panel discrete choice models were estimated to analyze the data, capturing both the relative importance of each attribute in the decision process and systematic taste variation among individuals. Results show that VMS displaying predictions of bus arrival at stops do influence bus route choice behavior. The estimated average willingness to pay for a bus stop to have a VMS was R\$0.12, which corresponds to 5 minutes of travel time. It was also observed that the marginal disutility of waiting time decreases when there is a VMS at the stop. Disutility of waiting due to delays also decreases (more intensively) with the VMS. The average value of travel time was relatively low, compared to expectations: R\$1.44/hour, probably due to the socioeconomic profile of the sample, particularly the high proportion of students. Nevertheless, frequency of bus use and socioeconomic characteristics significantly affect route choice behavior; the value of travel time, for instance, may reach R\$17/hour, while willingness to pay for a VMS in a stop may become R\$0.77. The results indicate that incorporating the VMS as an additional component of the decision, allows for a better understanding of bus route choice behavior.

Keywords: Public transportation. Bus. Displays. Information. Demand (Models).



## LISTA DE ILUSTRAÇÕES

Figura 1 – Sequência de etapas para coleta de dados .....	52
Figura 2 – Sequência de etapas para análise de dados .....	53
Figura 3 – Exemplo de tabela com as características das alternativas .....	56
Figura 4 – Segmentação por frequência de uso de ônibus e uso de pontos com painel .....	68
Figura 5 – Segmentação por idade e sexo .....	69
Figura 6 – Segmentação por quantidade de automóveis e motocicletas no domicílio .....	71
Figura 7 – Segmentação por renda do domicílio em R\$ .....	71
Figura 8 – Exemplo de tabela com as características das alternativas do primeiro piloto .....	102
Figura 9 – Exemplo de tabela com as características das alternativas do segundo piloto.....	105
Figura 10 – Tela inicial do questionário de pesquisa .....	114

## LISTA DE TABELAS

Tabela 1 – Exemplos de situações de escolha de um desenho de experimento .....	44
Tabela 2 – Detalhamento dos atributos do experimento .....	58
Tabela 3 – Relação entre atributo <i>ocupação</i> e variáveis <i>dummy</i> .....	59
Tabela 4 – Estratégia para combinar atributos <i>intervalo</i> , <i>atraso</i> e <i>tempo de viagem</i> .....	60
Tabela 5 – Aproximações para coeficientes .....	62
Tabela 6 – Segmentação por ocupação e escolaridade.....	70
Tabela 7 – Segmentação quanto ao comportamento compensatório.....	73
Tabela 8 – Modelo base.....	75
Tabela 9 – Modelo ML <i>panel piecewise</i> .....	80
Tabela 10 – Taxas marginais de substituição (em minutos de viagem) .....	80
Tabela 11 – Taxas marginais de substituição (em minutos de intervalo, atraso ou viagem) ...	81
Tabela 12 – Modelos ML <i>panel</i> com interações entre atributos .....	83
Tabela 13 – Modelo ML <i>panel</i> com variáveis socioeconômicas e de uso do ônibus.....	89
Tabela 14 – Atributos para primeiro experimento piloto .....	103
Tabela 15 – Desenho eficiente para o primeiro experimento piloto.....	104
Tabela 16 – Modelo MNL estimado com resultados do primeiro experimento piloto .....	104
Tabela 17 – Atributos para segundo experimento piloto.....	106
Tabela 18 – Relação entre atributo <i>ocupação</i> e variáveis <i>dummy</i> .....	107
Tabela 19 – Estratégia para combinar atributos intervalo, atraso e tempo de viagem .....	107
Tabela 20 – Desenho eficiente para segundo experimento piloto.....	111

Tabela 21 – Taxas marginais de substituição a partir do segundo experimento piloto.....	112
Tabela 22 – Situações de escolha do desenho do experimento .....	113
Tabela 23 – Modelo base sem os lexicográficos quanto ao conforto.....	116
Tabela 24 – Modelo base sem os lexicográficos quanto ao custo.....	117
Tabela 25 – Modelo base sem os lexicográficos quanto ao painel.....	118
Tabela 26 – Modelo ML <i>panel</i> com interações entre atributos (1) .....	119
Tabela 27 – Modelo ML <i>panel</i> com interações entre atributos (2) .....	119
Tabela 28 – Modelo ML <i>panel</i> com interações entre atributos (3) .....	120
Tabela 29 – Modelo ML <i>panel random coefficients</i> com distribuição normal .....	121
Tabela 30 – Modelo ML <i>panel random coefficients</i> com distribuição uniforme.....	122
Tabela 31 – Modelo ML <i>panel</i> estimado a partir das respostas dos não usuários .....	123
Tabela 32 – Classificação socioeconômica e quanto ao uso de ônibus dos lexicográficos....	124
Tabela 33 – Classificação socioeconômica e de uso de ônibus por <i>escolaridade</i> .....	125
Tabela 34 – Classificação socioeconômica por <i>frequência de uso de ônibus</i> .....	126
Tabela 35 – Classificação socioeconômica e de uso de ônibus por <i>ocupação</i> .....	127
Tabela 36 – Classificação socioeconômica e de uso de ônibus por <i>idade</i> .....	128

## **LISTA DE ABREVIATURAS E SIGLAS**

AVL	Automatic Vehicle Location
IIA	Independence from Irrelevant Alternatives
ML	Mixed Logit
MNL	Multinomial Logit
PD	Preferência Declarada
PMV	Painel de Mensagem Variável
PR	Preferência Revelada
RMSP	Região Metropolitana de São Paulo
RTPI	Real Time Passenger Information
USP	Universidade de São Paulo

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>13</b>
1.1	Objetivos .....	16
1.2	Justificativas .....	17
<b>2</b>	<b>REVISÃO DA LITERATURA.....</b>	<b>18</b>
<b>2.1</b>	<b>Informações para o transporte: da identificação da necessidade ao uso.....</b>	<b>18</b>
2.1.1	<i>Classificações da informação.....</i>	<i>18</i>
2.1.2	<i>Nível de conhecimento e necessidade de informação.....</i>	<i>19</i>
2.1.3	<i>Valor da informação e sua aquisição.....</i>	<i>20</i>
2.1.4	<i>Atualização do conhecimento e decisão.....</i>	<i>22</i>
2.1.5	<i>Estratégias de decisão dos indivíduos.....</i>	<i>23</i>
<b>2.2</b>	<b>Espera no transporte público por ônibus .....</b>	<b>25</b>
2.2.1	<i>Tempo de espera objetivo .....</i>	<i>25</i>
2.2.2	<i>Tempo de espera subjetivo e a experiência da espera com caráter negativo .....</i>	<i>27</i>
2.2.3	<i>Experiências de espera com caráter positivo.....</i>	<i>28</i>
<b>2.3</b>	<b>Escolha de linha de ônibus e alocação de passageiros .....</b>	<b>30</b>
<b>2.4</b>	<b>Informação em tempo real sob outras perspectivas .....</b>	<b>31</b>
2.4.1	<i>Redução dos recursos tempo e esforço.....</i>	<i>31</i>
2.4.2	<i>Motivações e benefícios sociais e econômicos .....</i>	<i>32</i>
<b>2.5</b>	<b>Técnicas de coleta de dados e de modelagem adotadas .....</b>	<b>32</b>
2.5.1	<i>Técnicas de preferência declarada para coleta de dados.....</i>	<i>32</i>
2.5.2	<i>Modelos de escolha discreta.....</i>	<i>33</i>
2.5.3	<i>Experimentos de escolha declarada e desenhos eficientes.....</i>	<i>43</i>
<b>3</b>	<b>METODOLOGIA.....</b>	<b>51</b>
<b>3.1</b>	<b>Métodos para coleta de dados .....</b>	<b>51</b>
<b>3.2</b>	<b>Métodos para análise de dados .....</b>	<b>53</b>
<b>4</b>	<b>EXPERIMENTO COM DESENHO EFICIENTE.....</b>	<b>55</b>

<b>4.1 Criação do desenho eficiente</b> .....	<b>57</b>
4.1.1 <i>Alternativas</i> .....	57
4.1.2 <i>Atributos</i> .....	57
4.1.3 <i>Restrições ao desenho</i> .....	60
4.1.4 <i>Especificação econométrica</i> .....	61
4.1.5 <i>Aproximações para parâmetros</i> .....	61
4.1.6 <i>Situações de escolha</i> .....	62
4.1.7 <i>Desenho eficiente criado</i> .....	63
<b>4.2 Informações adicionais coletadas</b> .....	<b>63</b>
<b>4.3 Amostra e operacionalização</b> .....	<b>65</b>
<b>5 RESULTADOS E DISCUSSÃO</b> .....	<b>67</b>
<b>5.1 Caracterização da amostra</b> .....	<b>67</b>
5.1.1 <i>Caracterização socioeconômica e relativa ao uso de ônibus</i> .....	67
5.1.2 <i>Caracterização quanto ao comportamento não compensatório</i> .....	72
<b>5.2 Modelo base para o comportamento da demanda</b> .....	<b>74</b>
<b>5.3 Não linearidade das preferências</b> .....	<b>77</b>
<b>5.4 Influência das interações entre atributos</b> .....	<b>82</b>
<b>5.5 Variabilidade sistemática das preferências</b> .....	<b>86</b>
<b>6 CONCLUSÕES E RECOMENDAÇÕES</b> .....	<b>94</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>97</b>
<b>APÊNDICE A – PRIMEIRO EXPERIMENTO PILOTO</b> .....	<b>102</b>
<b>APÊNDICE B – SEGUNDO EXPERIMENTO PILOTO</b> .....	<b>105</b>
<b>APÊNDICE C – DESENHO EFICIENTE DO EXPERIMENTO</b> .....	<b>113</b>
<b>APÊNDICE D – TELA INICIAL DO QUESTIONÁRIO DE PESQUISA</b> .....	<b>114</b>
<b>APÊNDICE E – INFLUÊNCIA DO COMPORTAMENTO LEXICOGRÁFICO</b> .....	<b>115</b>
<b>APÊNDICE F – MODELOS COM INTERAÇÕES ENTRE ATRIBUTOS</b> .....	<b>119</b>
<b>APÊNDICE G – VARIABILIDADE ALEATÓRIA DAS PREFERÊNCIAS</b> .....	<b>121</b>

<b>APÊNDICE H – MODELO ML PANEL PARA NÃO USUÁRIOS .....</b>	<b>123</b>
<b>APÊNDICE I – CLASSIFICAÇÃO SOCIOECONÔMICA E DE USO DE ÔNIBUS..</b>	<b>124</b>

# 1 INTRODUÇÃO

O usuário do sistema de transportes está sujeito a incertezas de naturezas diversas enquanto se desloca pela rede. Variabilidade dos tempos de viagem no transporte individual e coletivo (comum em grandes centros urbanos sujeitos a congestionamentos), descumprimento da programação do transporte público, pouca familiaridade ou desconhecimento de rotas alternativas no transporte individual e público, desinformação em relação a níveis de congestionamento em determinadas vias e faixas horárias, bem como em relação à frequência dos serviços de transporte público, para mencionar algumas.

No que tange a variabilidade do sistema, Bonsall (2004) a diferencia da incerteza percebida pelo indivíduo: esta, embora seja influenciada pela variabilidade, envolve também aspectos psicológicos. Já Soriguera (2013), diferencia os conceitos de confiabilidade e variabilidade do sistema. Argumenta que a variabilidade está associada ao desempenho do sistema, enquanto a confiabilidade está ligada às expectativas do indivíduo em relação a este desempenho (de modo que um sistema pode ser confiável mesmo com alta variabilidade, desde que ela seja conhecida pelo indivíduo).

As estratégias adotadas pelos indivíduos para lidar com as incertezas percebidas são diversas. É possível, por exemplo: evitar modos, rotas e horários que sabidamente apresentem muitas oscilações no tempo de viagem; adotar margens de segurança no planejamento da viagem, adiantando o horário de partida; utilizar o tempo de viagem dentro do veículo para realizar outras atividades; cancelar a viagem caso o deslocamento seja custoso a ponto de superar os benefícios pretendidos com a mesma; ou ainda acessar informações adicionais sobre o sistema que auxiliem no planejamento e execução da viagem, quer sejam provenientes de opiniões de outros usuários e da mídia, ou de orientações dos gestores e operadores do sistema (BONSALL, 2004).

Para o caso específico do transporte urbano por ônibus, sua operação em ambiente apenas parcialmente controlado contribui para a irregularidade do sistema e, por consequência, afeta os tempos de viagem e espera pelos veículos nos pontos de parada. Uma forma de auxiliar os passageiros a lidar com esta irregularidade é a divulgação, em tempo real,



dos intervalos de tempo previstos para passagem das linhas, através de painéis eletrônicos instalados nos pontos de parada. Esta solução é parte do conceito RTPI (*Real Time Passenger Information*), um meio de prover informação eletrônica em tempo real sobre o transporte público a seus usuários, quer seja em pontos de parada, estações, dentro dos veículos ou via internet e telefone.

Dziekan e Kottenhoff (2007) mencionam sete importantes efeitos do uso dos painéis e ressaltam que, embora estejam relacionados uns aos outros, são uma referência útil para estudos nesta área. Tais efeitos são: (1) redução do tempo de espera percebido, (2) impactos psicológicos positivos, como redução da incerteza e maior sensação de segurança, (3) maior disposição a pagar pelo serviço, (4) ajustes no comportamento de modo a usar melhor o tempo de espera e viajar mais eficientemente, (5) efeitos na escolha do modo, (6) maior satisfação e (7) melhor imagem do serviço.

Pode-se acrescentar a esta lista de efeitos as alterações na escolha da linha de ônibus quando há mais de uma alternativa disponível, de modo que a informação usada pelos passageiros (e a forma como é usada) deve ser levada em consideração em modelos de alocação de passageiros no transporte público (LARSEN; SUNDE, 2008).

A disponibilização de informação em tempo real eleva a qualidade e a atratividade do serviço de transporte público. No entanto, Dziekan e Kottenhoff (2007) questionam se estes sistemas chegam a elevar a demanda pelos serviços de transporte, seja estimulando mais viagens dos usuários existentes ou atraindo novos usuários. Holdsworth; Enoch e Ison (2007) sugerem que tais sistemas são mais efetivos quando fazem parte de um pacote de medidas (ao invés de serem adotados isoladamente), especialmente se ações de combate ao congestionamento nas vias da cidade estiverem inclusas. Cabe também salientar que o surgimento de efeitos positivos está condicionado à confiabilidade da informação fornecida, e, por conseguinte, à confiabilidade percebida pelos usuários.

No caso da informação em tempo real sobre intervalos de tempo previstos para passagem das linhas de ônibus, através de painéis eletrônicos instalados nos pontos de parada, a compreensão das respostas dos usuários é necessária para estimar as relações de custo-benefício da implantação destes sistemas. E permite, ainda, comparar a importância dada pelo usuário e a eficiência destes sistemas frente a outras medidas aplicadas ao transporte público,

como adoção de tarifas mais baixas, redução dos tempos de viagem, melhorias no conforto dos veículos etc.

É importante salientar que o simples fato de um ponto de parada estar equipado com painel eletrônico (antes mesmo que se veja e utilize a informação por ele mostrada) representa, para o usuário, a possibilidade de que uma parte da incerteza associada à sua viagem será reduzida, dado que a existência do painel elimina a possibilidade de se esperar pelo ônibus sem saber quando ele chegará. Para o contexto de viagens rodoviárias, Razo e Gao (2013) demonstraram que o fato de uma rota estar equipada com sistema PMV (Painel de Mensagem Variável) pode torná-la mais atrativa. É pertinente, portanto, diferenciar os efeitos da informação (disponibilizada em tempo real através do painel) dos efeitos da presença do painel (o mecanismo de disponibilização da informação em tempo real), embora estejam intimamente relacionados. Enquanto a presença do painel atua na redução da incerteza, chegando a influenciar o comportamento e as decisões dos usuários (conforme será demonstrado no desenvolvimento da pesquisa), é a informação em si que dá condições ao indivíduo de avaliar com maior precisão as alternativas disponíveis e tomar decisões mais satisfatórias em sua viagem e também em sua programação de atividades.

Em concordância com as ideias do parágrafo precedente, o presente estudo se concentra nos efeitos da existência do painel no ponto de ônibus – antes mesmo que se veja e utilize a informação por ele mostrada – no comportamento de escolha de linha dos usuários de ônibus. A questão de pesquisa e suas hipóteses nula e alternativa são assim formuladas:

*A presença de painéis eletrônicos em pontos de ônibus, disponibilizando informação em tempo real sobre intervalos de tempo previstos para passagem dos veículos, influencia o comportamento de escolha de linha?*

#### **Hipótese nula**

A presença de painéis não influencia o comportamento de escolha de linha:

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

#### **Hipótese alternativa**

A presença de painéis influencia o comportamento de escolha de linha:

$$H_1: \beta_{display} \neq 0$$

sendo  $\beta_{display}$  o coeficiente do atributo *existência de painel no ponto* utilizado na análise dos resultados.

O restante do documento está assim organizado: nas próximas seções os objetivos da pesquisa e suas justificativas são formalizados; no capítulo seguinte é apresentada uma revisão da literatura pertinente; a descrição da metodologia é objeto do terceiro capítulo, enquanto o seguinte trata da elaboração do experimento para a coleta de dados desta pesquisa; finalmente, no quinto capítulo são apresentados e analisados os resultados do experimento, seguidos pelas conclusões e recomendações no sexto e último capítulo.

## 1.1 Objetivos

Em concordância com a questão de pesquisa, este estudo tem como objetivo geral:

*Investigar a influência, no comportamento de escolha de linha, da presença de painéis eletrônicos em pontos de ônibus com informação em tempo real sobre intervalos de tempo previstos para passagem dos veículos.*

Como objetivos específicos, deseja-se:

1. mensurar a influência dos seguintes atributos sobre o comportamento de escolha de linha: (1) existência de painel no ponto, (2) intervalo de tempo programado para a passagem de veículos consecutivos da mesma linha, (3) possível atraso em relação a este intervalo programado, (4) tempo de viagem dentro do veículo até o destino, (5) ocupação do veículo quando chega ao ponto e (6) valor da tarifa;
2. identificar quais interações entre estes seis atributos influenciam o comportamento de escolha de linha, e mensurar a intensidade com que isto ocorre;
3. identificar as características socioeconômicas e relativas a hábitos de uso de ônibus dos indivíduos que intensificam ou diminuem a influência dos mesmos seis atributos sobre o comportamento de escolha de linha, e mensurar a intensidade com que isto ocorre.

## 1.2 Justificativas

Embora já mencionadas na introdução deste capítulo, as justificativas desta pesquisa são sumarizadas e formalizadas abaixo:

*Ampliar o entendimento sobre as respostas dos usuários aos painéis, de modo a fornecer subsídios para a estimação das relações de custo-benefício de sua implantação e a permitir a comparação de seus benefícios com os benefícios advindos de outras medidas aplicadas ao transporte público.*

*Ampliar o entendimento sobre a decisão de escolha de linha de ônibus, de modo a contribuir para a melhoria dos modelos de alocação de passageiros às linhas do transporte público, usados em atividades de simulação e planejamento do sistema.*

## 2 REVISÃO DA LITERATURA

O conteúdo deste capítulo está dividido em cinco partes, oferecendo referencial teórico sobre: (1) a informação para os usuários do sistema de transportes, desde a identificação da necessidade até o uso para tomada de decisão; (2) a espera no transporte público por ônibus e sua relação com a informação em tempo real; (3) a escolha de linha de ônibus e sua relação com a informação em tempo real; (4) perspectivas adicionais sobre a informação em tempo real; e (5) as técnicas de coleta de dados e de modelagem utilizadas neste estudo.

É importante mencionar que, dentre os efeitos dos sistemas de informação em tempo real que fornecem os tempos previstos para passagem de veículos do transporte público, os benefícios relacionados aos tempos de espera e ao comportamento de escolha de linha são aqui enfatizados, merecendo cada um deles uma seção específica.

### 2.1 Informações para o transporte: da identificação da necessidade ao uso

Antes de descrever os mecanismos em ação desde a identificação da necessidade de informação até a incorporação desta última no processo decisório, cabe uma exposição sucinta das classificações da informação para usuários dos sistemas de transporte.

#### 2.1.1 *Classificações da informação*

Chorus; Arentze e Molin (2006) classificam a informação usada pelo usuário no planejamento e na execução de viagens em duas categorias: (1) informação para geração de alternativas (até então desconhecidas pelo usuário) e (2) para caracterização de alternativas já conhecidas. No caso do transporte público, uma nova alternativa pode ser uma linha de ônibus até então desconhecida, enquanto o tempo de viagem desta linha é um atributo que a caracteriza.

Já Grotenhuis; Wiegmans e Rietveld (2007) classificam a informação para o transporte público quanto ao momento da viagem em que é usada: antes do início, com finalidade de

planejamento; ou durante a execução, com finalidade de acompanhamento (quer a informação seja acessada em terminais e pontos de parada ou dentro dos veículos).

Soriguera (2013) propõe classificação semelhante, diferenciando a informação que ajuda o usuário a tomar decisões operacionais daquela que o ajuda a planejar sua viagem. Informações em tempo real, por exemplo, podem dar suporte a decisões operacionais, que podem ser alteradas mais facilmente, como a escolha da rota. Já informações baseadas em dados históricos, por exemplo, podem dar suporte a decisões que precisam ser planejadas com antecedência, como o horário de partida (visto que não é possível conhecer previamente o estado do sistema no momento futuro planejado para a partida).

Ettema e Timmermans (2006), por sua vez, usam as seguintes categorias para classificar a informação: retrospectiva, que se refere ao desempenho do sistema no passado; descritiva, que é relativa às características atuais do sistema; e preditiva, ou seja, relativa à projeção de estados futuros do sistema.

### ***2.1.2 Nível de conhecimento e necessidade de informação***

O nível de conhecimento do usuário do sistema de transporte pode ser definido a partir do conceito de conhecimento total ou informação perfeita. Este conceito implica na ciência de todas as alternativas e dos valores exatos de suas características, de modo a executar a decisão em um cenário de absoluta certeza. Desta forma, os desvios em relação ao conhecimento completo (quer seja por desconhecimento de alternativas ou de seus atributos) caracterizam o nível de conhecimento do indivíduo (BONSALL, 2004; CHORUS; ARENTZE; MOLIN, 2006).

Para descrever o nível de conhecimento do indivíduo, Dziekan (2008) usa o conceito de *memory representation* (representação na memória, em tradução literal para a Língua Portuguesa), que descreve o que o indivíduo sabe ou tem em mente a respeito de algo, que para o contexto do transporte pode ser uma linha de ônibus, por exemplo. De acordo com a autora, este conceito compreende os mapas cognitivos (que representam espacialmente o ambiente na mente do indivíduo) e também outros conhecimentos (tais como opções

existentes de linhas de ônibus e suas frequências), e serve como base para as escolhas dos usuários, quer sejam escolhas de destino, modo ou rota.

Para o contexto do transporte público, quanto maior a frequência de uso, mais o indivíduo aprimora seu conhecimento (*memory representation*) a respeito do serviço, dado que, ao interagir com o sistema, ele continuamente armazena informações novas, experiências e percepções, e combina as informações do ambiente com seu mapa cognitivo (DZIEKAN, 2008).

A percepção que os usuários têm do próprio conhecimento (ou desconhecimento) a respeito do sistema de transportes é uma das determinantes da necessidade de informação. Esta percepção é expressa, por exemplo, em seu grau de confiança nas próprias estimativas de tempos de espera e viagem (CHORUS et al., 2007). Para o contexto do transporte público, Grotenhuis; Wiegmans e Rietveld (2007) relacionam a necessidade de informação também com a frequência de uso (um indicador da familiaridade com o serviço), o período do dia, o dia da semana, o propósito, a duração e a distância da viagem, além de indicadores socioeconômicos.

Zhao; Yang e Zhang (2010) estudaram o comportamento de escolha de informação para serviços de transporte público. Os resultados indicaram que, dentre os tipos de informação apresentados a indivíduos com ocupação “estudante” ou com propósito de viagem “estudo”, a preferência foi por tabelas de horários com a programação do serviço, previsões de tempos de chegada dos veículos aos pontos e informação sobre atrasos. Já os entrevistados com ocupação “trabalhador assalariado” ou com propósito de viagem “trabalho”, manifestaram preferência pelos três conteúdos acima mencionados e também por mapas da cidade e do sistema de transporte público, além de informações sobre transbordos.

### ***2.1.3 Valor da informação e sua aquisição***

A informação sobre o sistema de transporte pode ser adquirida de forma passiva (através dos painéis instalados em terminais e pontos de parada do transporte público, por exemplo) ou ativa (através de buscas pela internet, por exemplo). Em qualquer uma das situações, a aquisição de informação é descrita como uma decisão de custo-benefício, e

somente será realizada quando o usuário sentir que os benefícios percebidos da informação superam o custo percebido de sua busca e aquisição, que pode incluir custos monetários, investimentos de tempo, esforço e atenção (CHORUS; ARENTZE; MOLIN, 2006). Assim, a possibilidade de não aquisição de informação e de consequente permanência em um estado parcial de ignorância existe e pode ser uma escolha racional (LARSEN; SUNDE, 2008).

No que se refere ao valor da informação, Soriguera (2013) o calcula como a diferença entre os custos de viagem com e sem informação. E, para a segunda situação, contabiliza, além do possível tempo adicional de viagem, os custos de uma possível reprogramação de atividades e o estresse inerente a esta situação.

Já Chorus; Molin e Van Wee (2006) usam a perspectiva do arrependimento esperado para calcular o valor da informação. Assume-se que, para dada situação de escolha, o indivíduo seleciona a alternativa que, em sua percepção, irá gerar o menor arrependimento (que é mensurado numericamente). A diferença entre os menores arrependimentos associados a situações de escolha com e sem informação é o valor percebido da informação.

Ao estudar o valor da informação sobre tempos de viagem em automóvel, Soriguera (2013) encontrou uma curva assimétrica com cauda alongada à direita para a distribuição de frequências do valor da informação. Uma das características deste tipo de curva é que a moda é inferior à média. No contexto da pesquisa do autor, tal característica pode refletir o fato de que, na maior parte das vezes em que os viajantes usam a informação sobre os tempos de viagem, obtêm benefícios inferiores ao benefício médio que pode ser conseguido com o uso da informação. E o formato da curva pode, segundo o autor, explicar a baixa disposição dos usuários a pagar pela informação: considerando que os benefícios que se repetem mais frequentemente são lembrados mais claramente pelos indivíduos, a disposição a pagar pela informação estaria mais relacionada à moda do que à média da curva.

Fazendo referência ao exposto no primeiro parágrafo, no experimento realizado para esta pesquisa, ao selecionar uma alternativa com painel eletrônico em detrimento de uma alternativa sem painel, o participante está em algum grau valorizando os benefícios que a informação poderá trazer e absorvendo os custos desta escolha (relacionados a possíveis características negativas da alternativa selecionada). Por outro lado, quando escolhe a alternativa sem painel, está optando por permanecer em estado momentâneo de ignorância



quanto ao momento previsto da passagem da linha pelo ponto (muito embora possa ter expectativas a respeito do tempo de espera, alimentadas, por exemplo, por sua experiência acumulada).

#### ***2.1.4 Atualização do conhecimento e decisão***

Uma vez adquiridas, as informações adicionais são agregadas ao conjunto de conhecimentos pré-existentes (derivados da própria experiência e de informações adquiridas anteriormente), de modo a atualizá-lo, e as decisões tomadas a partir de então dependerão do estilo (ou estratégia) de decisão do usuário (CHORUS; MOLIN; VAN WEE, 2006), bem como de sua atitude de propensão ou aversão ao risco (BONSALL, 2004).

Os mecanismos de atualização do conjunto de conhecimentos podem ser explicados sob as perspectivas da introdução de novos conhecimentos e da atualização dos já existentes. Assim, em concordância com a classificação de Chorus; Arentze e Molin (2006) apresentada na Seção 2.1.1, a informação introduz novos conhecimentos quando revela ao usuário (1) alternativas de viagem até então desconhecidas e (2) atributos até então desconhecidos de alternativas de viagem já conhecidas; e atualiza os conhecimentos já existentes quando apresenta novos valores para atributos já conhecidos de alternativas de viagem igualmente já conhecidas. Através destes mecanismos, a informação pode reformatar o contexto e a complexidade da escolha aos olhos do usuário, sendo que a intensidade deste processo é maior na medida em que ele percebe seus conhecimentos iniciais (prévios à aquisição de informação) como não confiáveis, e as novas informações como confiáveis (CHORUS; MOLIN; VAN WEE, 2006).

No que tange à confiabilidade da informação, ressalta-se que ela afeta não só o processo de atualização do conhecimento, mas também a decisão de aquisição. Ben-Elia et al. (2013) definem acurácia da informação como a habilidade do sistema gerador da mesma em reduzir a discrepância entre suas estimativas e os tempos reais experimentados pelos usuários. Acrescentam que a informação imprecisa pode ser entendida pelos indivíduos como um elemento agravador do risco. Em um contexto de informação preditiva sobre tempos de viagem, Ettema e Timmermans (2006) diferenciam o tempo de viagem previsto da percepção

que o indivíduo tem da previsão. Argumentam que, através da experiência, o usuário passa a conhecer e avaliar a confiabilidade da previsão recebida.

Ben-Elia et al. (2013) demonstraram que o decréscimo na acurácia da informação fornecida induz a menores taxas de uso da informação e a comportamentos de aversão ao risco: uma vez que não se pode confiar na informação recebida, opta-se pela alternativa que pareça mais segura. No entanto, quando se trata de informação de natureza prescritiva (ou seja: quando o sistema de informação indica determinada alternativa como a melhor opção), a pesquisa conduzida pelos autores revelou que o nível de acurácia da informação não influencia sua taxa de aceitação. Ou seja: mesmo quando a prescrição não é acurada os indivíduos continuam tomando suas decisões com base nela, comportando-se segundo a heurística da ancoragem.

Um resultado interessante encontrado por Soriguera (2013), é que o valor da informação só é afetado por sua acurácia até um limite tido como nível mínimo necessário de acurácia. A partir deste limite, melhorias na acurácia resultam em pequenos benefícios marginais.

### ***2.1.5 Estratégias de decisão dos indivíduos***

Nesta seção serão descritas estratégias de decisão empregadas pelos indivíduos, muito embora não se almeje fazer uma revisão exaustiva do assunto. Para melhor compreender o conteúdo dos parágrafos seguintes, recomenda-se considerar o contexto de um indivíduo tomador de decisão em face de um conjunto de alternativas disponíveis (caracterizadas cada uma por seus atributos), dentre as quais uma deverá ser escolhida.

Ben-Akiva e Lerman (1985) referem-se às estratégias como regras de decisão, e as classificam em regras compensatórias e não compensatórias. Ainda lembram que, no caso das não compensatórias, pode-se chegar a uma decisão através da aplicação isolada de uma regra ou da combinação de duas ou mais. Abaixo é reproduzida a lista de regras mencionadas pelos autores:

- Regra da dominância (não compensatória): é usada quando uma das alternativas disponíveis é superior às demais em pelo menos um atributo, e não é inferior às demais nos atributos restantes. Tal alternativa é a escolhida;
- Regra da satisfação (não compensatória): pressupõe que para cada atributo o indivíduo tenha um nível de exigência mínimo. Desta forma, uma alternativa que tenha pelo menos um atributo que não satisfaça este nível de exigência será excluída do processo de decisão;
- Regra lexicográfica (não compensatória): pressupõe que o tomador de decisão ordene os atributos em ordem decrescente de importância. Assim, a alternativa que for mais atraente para o atributo mais importante será escolhida;
- Regra da maximização da utilidade (compensatória): pressupõe que o grau de atratividade de uma alternativa seja representado por uma função matemática de seus atributos (a função utilidade), que, quando calculada, resulta em um escalar (a utilidade). O tomador de decisão escolhe a alternativa com a maior utilidade, otimizando sua satisfação. A maximização da utilidade é dita compensatória porque o indivíduo usa regras de compensação entre os atributos que considera negativos e aqueles que considera positivos de cada alternativa, podendo escolher uma opção que tenha características negativas, desde que sejam compensadas pelas positivas. Ao escolher determinada linha de ônibus para se deslocar, por exemplo, a alta ocupação do veículo pode ser compensada por um tempo de viagem mais curto. Por ser elemento fundamental das técnicas econométricas usadas para analisar os resultados desta pesquisa, esta estratégia de decisão será explorada com mais detalhes na Seção 2.5.2 deste mesmo capítulo.

Outra estratégia de decisão, baseada no arrependimento (e não na utilidade ou satisfação), é a minimização do arrependimento esperado. Segundo esta estratégia, o indivíduo faz sua escolha de modo a evitar que qualquer alternativa não escolhida se torne mais atrativa do que aquela que foi selecionada, pois isto causaria arrependimento. Assim, o tomador de decisão prevê o arrependimento associado à escolha de cada alternativa (o que, para fins de modelagem, é traduzido por uma função matemática), e elege aquela que tem o

menor arrependimento esperado. Chorus; Arentze e Timmermans (2008) enquadram esta estratégia no grupo das regras semi-compensatórias e não compensatórias, visto que o mau desempenho em um atributo (a alta ocupação de um ônibus, por exemplo) não é necessariamente compensado pela boa performance em outro (o tempo de viagem baixo, por exemplo).

Um estilo adicional de tomada de decisão é o comportamento habitual, em que não há avaliação deliberada das alternativas (VERPLANKEN; AARTS; VAN KNIPPENBERG, 1997) e a aquisição de novas informações é extremamente limitada, pois os esforços cognitivos tendem a ser reduzidos a um nível mínimo (AARTS; VERPLANKEN; VAN KNIPPENBERG, 1997).

Cabe ainda comentar que, para um mesmo contexto, pessoas diferentes podem aplicar regras distintas, e que o mesmo indivíduo pode variar suas estratégias de acordo com a situação, ou mesmo alterá-las ao passar do tempo.

## **2.2 Espera no transporte público por ônibus**

Esta seção diferencia o tempo de espera objetivo do subjetivo: o primeiro é aquele que pode ser medido pelo relógio, enquanto o segundo é a percepção individual do tempo. Além de descrever o tempo de espera sob ambas as perspectivas, esta seção analisa a experiência da espera (que pode adquirir caráter negativo ou positivo) e discute o papel da informação em tempo real na redução dos tempos de espera objetivo e subjetivo, e na criação de experiências mais positivas.

### ***2.2.1 Tempo de espera objetivo***

No transporte público por ônibus, o tempo de espera objetivo é função tanto dos horários de passagem dos ônibus nos pontos de parada, quanto dos horários de chegada dos passageiros a estes locais. Assim, ambos devem ser considerados quando se trata de estratégias para redução do tempo de espera.

Os horários de passagem dos veículos dependem tanto da frequência e regularidade do serviço quanto dos congestionamentos na rede viária. Desta forma, ações de combate aos congestionamentos e para aumento da frequência e regularidade do serviço colaboram para a redução dos tempos de espera.

No que se refere à regularidade, políticas de gerenciamento da frota podem ser empregadas para aumentar a aderência à programação. Um exemplo é a alocação de intervalos de tempo adicionais ao longo dos itinerários, de modo a absorver os atrasos na operação (CEDER, 2007). Assim, veículos à frente ou de acordo com a programação esperam parados nos pontos até que o horário de prosseguir sua viagem seja atingido, ao contrário dos veículos atrasados. Cabe acrescentar que o gerenciamento da frota pode ser auxiliado por práticas de rastreamento de veículos através de sistemas AVL (*Automatic Vehicle Location*), demonstrando que esta tecnologia pode auxiliar na redução dos tempos objetivos de espera.

Já a chegada de passageiros aos pontos pode se dar de forma aleatória ou em sincronia com os horários de passagem dos veículos. A primeira ocorre quando os usuários não estão cientes da programação das linhas ou quando a irregularidade do serviço ou sua frequência são altas. No caso da sincronização, os picos de chegada dos passageiros aos pontos ocorrem poucos minutos antes do horário programado para a passagem das linhas, e requer que os usuários estejam cientes da programação do serviço. Geralmente está associada a serviços regulares e de baixa frequência. Ressalta-se que a sincronização das chegadas com os horários de passagem das linhas é, em si, uma estratégia do usuário para minimizar o tempo de espera e sua variabilidade, elevando a importância da divulgação das tabelas de horários e da sua elaboração de modo a permitirem fácil compreensão (HALL, 2001).

Sob o ponto de vista da demanda, além do ponto de vista da oferta, os sistemas AVL também podem proporcionar redução do tempo de espera, através da utilização de seus dados em sistemas de informação em tempo real para passageiros do transporte público, quer estes dados sejam disponibilizados durante a viagem (nos terminais, pontos de parada ou dentro dos veículos), ou antes de seu início, através da internet, por exemplo (HALL, 2001).

### ***2.2.2 Tempo de espera subjetivo e a experiência da espera com caráter negativo***

O ato de esperar é um fenômeno que geralmente provoca reações emocionais nos indivíduos. Dependendo do contexto, pode ser uma experiência prazerosa ou uma imposição desagradável, sendo na maior parte das vezes negativa em termos psicológicos, pois implica em renunciar a formas mais produtivas e recompensadoras de usar o tempo e porque aumenta o investimento necessário para alcançar o objetivo (DUBÉ-RIOUX; SCHMITT; LECLERC, 1989).

Assim, dependendo dos acontecimentos que preenchem um intervalo de tempo, bem como da forma como são processados e armazenados pelo indivíduo e das emoções associadas, dentre outros fatores, o intervalo percebido pelo indivíduo pode sofrer distorções em relação ao tempo objetivo (DROIT-VOLET; MECK, 2007; MANTEL; KELLARIS, 2003). Este intervalo percebido, quer seja mais longo ou mais curto do que o tempo objetivo, é chamado de tempo subjetivo ou psicológico, definido por Meck (2005) como a experiência interna e individual da passagem do tempo. De acordo com o mesmo autor, sua estimação requer o uso de uma espécie de contador interno para medir o tempo, porém sem recurso a relógios ou marcadores de tempo externos. É importante lembrar que a duração subjetiva de um evento pode influenciar as avaliações e a satisfação dos indivíduos em relação a um serviço e, conseqüentemente, seu comportamento (DUBÉ-RIOUX; SCHMITT; LECLERC, 1989).

Maister (1985) propõe diversas situações em que o tempo de espera percebido é distorcido e se torna superior ao tempo objetivo, dentre elas: (1) o tempo desocupado parece mais longo do que o tempo ocupado com alguma atividade; (2) esperas nas etapas que precedem uma atividade parecem mais longas do que esperas durante a execução da atividade; (3) a ansiedade faz com que a espera pareça mais longa; (4) esperas com durações desconhecidas parecem mais longas do que esperas com durações finitas e conhecidas; (5) esperas não explicadas parecem mais longas do que esperas justificadas; e (6) esperar sozinho parece mais demorado do que esperar acompanhado.

O segundo item desta lista, se interpretado no contexto do transporte público, está em conformidade com o fato de que o tempo de espera para embarque (a etapa que precede a

atividade de viajar no ônibus) é mais desconfortável e tem menor utilidade do que o tempo de viagem dentro do ônibus. Abrantes e Wardman (2011), em meta-análise de 90 estudos sobre valoração do tempo de espera, concluíram que o valor médio do tempo de espera é igual a 1,7 vezes o valor do tempo de viagem dentro do veículo, variando de 1,43 em estudos de preferência revelada até 2,32 em estudos de preferência declarada. Não há alterações expressivas destes números para diferentes propósitos de viagem ou modos.

Já o terceiro e o quarto itens estão de acordo com Psarros et al. (2011) e Dziekan e Kottenhoff (2007), que ressaltam que as distorções sofridas pelo tempo de espera subjetivo em relação ao tempo objetivo tornam-se maiores quando os passageiros estão desprovidos de informações a respeito dos horários de passagem das linhas (situação que, por si só, é geradora de ansiedade).

Além das situações listadas por Maister (1985), Dubé-Rioux; Schmitt e Leclerc (1989) argumentam que a espera e os atrasos são mais desagradáveis na medida em que a necessidade do indivíduo é maior. Embora os autores tenham feito esta proposição para o contexto de idas a restaurantes (onde a fome mede o grau da necessidade), é razoável afirmar que, para o transporte público, quanto menor o tempo de que o usuário dispõe para alcançar seu destino (ou quanto maior sua pressa), mais desagradável se torna a espera pelo ônibus.

Vale notar que o usuário de ônibus, enquanto espera nos terminais e pontos de parada, pode estar exposto a condições climáticas adversas, grandes concentrações de pessoas, ruído e poluição, elementos que também influenciam os tempos de espera percebidos, geralmente tornando-os superiores aos reais (PSARROS et al., 2011). Para o contexto de determinadas localidades, pode-se acrescentar a estes fatores a violência urbana, cuja influência sobre a percepção do tempo de espera pode ser acentuada, particularmente, quando as viagens são realizadas à noite.

### ***2.2.3 Experiências de espera com caráter positivo***

Em estudo relacionando experiência de espera e grau de satisfação com o serviço de transporte público por ônibus, Friman (2010) demonstrou que reações afetivas positivas podem ocorrer durante a etapa de espera que precede a viagem de ônibus. Como exemplo,

menciona a situação em que o usuário espera pela partida dentro do veículo, sugerindo que ele experimentaria a sensação de “ter entrado” no sistema e de que o serviço já teria começado. A autora também argumenta que intervenções no sentido de melhorar a experiência da espera podem aumentar significativamente a qualidade percebida e a atratividade do transporte público.

Pode-se afirmar que a instalação de painéis eletrônicos nos pontos de parada e terminais é um caso de intervenção positiva. Dziekan e Vermeulen (2006), por exemplo, estudaram as percepções de tempo de espera no transporte público antes e depois da implantação de painéis eletrônicos disponibilizando informação em tempo real sobre a passagem de uma *tramline* na Holanda. A pesquisa, feita em três momentos com a mesma amostra de usuários, revelou uma queda de aproximadamente 20% no tempo médio de espera percebido entre a primeira e a segunda medição (feitas, respectivamente, um mês antes e três meses após a implantação). A terceira medição (que ocorreu 16 meses após a implantação) revelou uma queda discreta em relação à segunda. Para o caso específico desta *tramline*, e com base em estimativas de custos para disponibilizar informação em tempo real em oito pontos, e de custos incorridos em ações para elevar a frequência do serviço, Dziekan e Kottenhoff (2007) sugerem que é cinco vezes mais caro obter uma redução do tempo de espera objetivo através do aumento da frequência, do que obter uma redução equivalente do tempo de espera subjetivo através da informação em tempo real.

Friman (2010) comparou os níveis de satisfação com o transporte por ônibus de dois cenários de espera distintos: um dito positivo e outro dito neutro. No primeiro, o usuário chega ao ponto imediatamente antes do horário programado para passagem da linha e o veículo já está no local, permitindo que embarque e aguarde alguns minutos para a partida, embora não receba informações sobre o horário exato em que a partida ocorrerá. Já no segundo cenário, o usuário chega ao ponto imediatamente antes do horário programado para passagem da linha, mas a mesma está atrasada. Recebe, então, de maneira contínua, informação a respeito do horário previsto de chegada do veículo ao ponto. O nível médio de satisfação do cenário neutro foi equivalente a 75% do nível médio do cenário positivo, apontando para o potencial que tem a informação em tornar a experiência de espera mais positiva.



### **2.3 Escolha de linha de ônibus e alocação de passageiros**

Os modelos de escolha de linha são um dos elementos fundamentais dos modelos de alocação de passageiros a linhas do transporte público (LIU; BUNKER; FERREIRA, 2010), e o pressuposto clássico destes modelos é de que o usuário embarca no primeiro veículo que passa no ponto de parada, desde que tal veículo pertença a seu conjunto de linhas atrativas (GENTILE; NGUYEN; PALLOTTINO, 2005).

Larsen e Sunde (2008), no entanto, argumentam que embarcar no primeiro veículo pode não ser uma estratégia racional, a menos que as linhas tenham tempos de viagem idênticos. Caso contrário, pode ser vantajoso deixar passar a linha de maior tempo de viagem e esperar pela chegada da linha cujo tempo é menor. Contudo, a disposição para fazê-lo depende das expectativas que o usuário tem quanto à espera adicional no ponto, e quanto à economia do tempo dentro do veículo.

Quando informação em tempo real é disponibilizada nos terminais e pontos de parada, o usuário passa a ter expectativas mais realistas sobre os tempos de espera associados a cada linha, tendo assim mais elementos para decidir em qual embarcar. A escolha pode ser feita pela linha que oferecer a melhor combinação de tempo de espera exibido pelo sistema e tempo estimado de viagem até o destino. Neste contexto, o pressuposto clássico para alocação de passageiros não se mantém, e a presença de informação em tempo real nos terminais e pontos de parada deve ser considerada nas simulações e dimensionamentos da rede de transportes para propósitos de planejamento (GENTILE; NGUYEN; PALLOTTINO, 2005).

Cabe salientar que a simples presença do painel no ponto de parada (situação explorada no experimento desta pesquisa), também tem impactos na escolha da linha, uma vez que é indicativa de redução da incerteza da viagem.

Sob uma perspectiva mais ampla, considerando inclusive a informação em tempo real disponibilizada antes do início da viagem (acessada através da internet, por exemplo), os impactos vão além da escolha da linha e estendem-se à escolha do horário de início da viagem (quando o usuário sai de sua origem), alterando assim tanto os padrões espaciais quanto os

padrões temporais dos fluxos alocados na rede de transportes (LIU; BUNKER; FERREIRA, 2010).

## **2.4 Informação em tempo real sob outras perspectivas**

Além dos efeitos da informação em tempo real e dos mecanismos que envolvem seu uso (apresentados nos itens anteriores), é pertinente descrever outras perspectivas encontradas na literatura, que são expostas a seguir.

### ***2.4.1 Redução dos recursos tempo e esforço***

De acordo com Stradling (2002), quando o transporte público tem caráter não confiável, as viagens geram esforços afetivo, cognitivo e físico: incerteza, preocupação e estresse consomem esforço afetivo; a necessidade de adaptar o planejamento feito inicialmente consome esforço cognitivo; e executar as adaptações consome esforço físico. Adicionalmente aos três tipos de esforços, mais tempo e dinheiro podem ser requeridos.

A informação para os usuários do transporte público possibilita economia dos recursos tempo e esforço despendidos nas viagens, reduzindo seu custo. Para Grotenhuis; Wiegman e Rietveld (2007), as economias de tempo referem-se tanto às reduções de tempo de viagem decorrentes do uso da informação, quanto às reduções de tempo de processamento da informação usada durante o planejamento e a execução da viagem. Já as economias de esforço referem-se aos esforços físico, cognitivo e afetivo relacionados à viagem, conforme descritos no parágrafo precedente. Stradling (2002) sugere que cada usuário tem uma quota pessoal de recursos de tempo e esforço, que são consumidos pelas viagens.

No que se refere aos esforços afetivos, Zhang; Shen e Clifton (2008) examinaram os efeitos comportamentais e psicológicos da informação em tempo real sobre intervalos de tempo previstos para chegadas de ônibus em usuários do serviço de *shuttle* da Universidade de Maryland, nos Estados Unidos da América. Os resultados indicaram que o uso do sistema aumentou significativamente o sentimento de segurança nas viagens noturnas.

### ***2.4.2 Motivações e benefícios sociais e econômicos***

Zhang; Shen e Clifton (2008) citam razões sociais e econômicas como motivações para o desenvolvimento e implementação de sistemas de informação em tempo real. Dentre as primeiras estão a promoção de inclusão social através de melhorias no transporte público, (balanceando as vantagens e desvantagens entre o transporte público e o privado), especialmente para a população que depende fortemente dele. Já as motivações econômicas são justificadas tanto pela expectativa de aumento da demanda do transporte público (o que pode ser benéfico também para o meio ambiente, caso o sistema incentive a migração do transporte privado para o público), quanto pelas economias de tempo que podem ser obtidas pela população, através das mudanças em suas escolhas de viagem e também devido ao potencial que estes sistemas têm para redução do congestionamento.

## **2.5 Técnicas de coleta de dados e de modelagem adotadas**

Nesta seção serão expostos os fundamentos das técnicas de coleta de dados e de modelagem econométrica adotadas neste trabalho, na seguinte ordem: (1) técnicas de preferência declarada, (2) modelos de escolha discreta e (3) experimentos de escolha declarada e desenhos eficientes. Embora a sequência de aplicação das técnicas tenha sido diferente (com o desenho do experimento precedendo a modelagem), é necessário expor os princípios dos modelos de escolha discreta antes, visto que são parte da fundamentação da teoria dos desenhos eficientes.

### ***2.5.1 Técnicas de preferência declarada para coleta de dados***

Ao contrário das técnicas de Preferência Revelada (PR), em que são obtidos dados a respeito das escolhas de fato realizadas pelos indivíduos em contextos reais, as técnicas de Preferência Declarada (PD) obtêm dados sobre as intenções de escolha dos indivíduos em situações hipotéticas. E, por este motivo, possibilitam maior controle das variáveis que afetam o comportamento da demanda, sendo que a indicação de seu uso é reforçada na medida em que não têm as limitações das técnicas de PR, mencionadas por Ortúzar e Willumsen (2011):

(a) escolhas reais nem sempre envolvem variabilidade suficiente entre as alternativas para permitir a construção de bons modelos econométricos; (b) quando poucos fatores conseguem explicar quase todo o comportamento observado, torna-se difícil detectar a importância relativa das outras variáveis de interesse, como é o caso de variáveis qualitativas, tais como a presença de painéis eletrônicos instalados em pontos de ônibus fornecendo informações em tempo real sobre o serviço; (c) inadequação para coletar respostas para situações totalmente novas, como uma nova tecnologia para transporte de pessoas, por exemplo.

Não obstante as vantagens das técnicas de PD, não há garantias de que a escolha declarada pelo indivíduo seja de fato realizada caso o contexto hipotético se torne real. Ortúzar e Willumsen (2011) enfatizam que as situações hipotéticas apresentadas aos entrevistados devem ser compreensíveis, parecer plausíveis e realistas e também estar relacionadas com as experiências já vivenciadas por eles.

Os mesmos autores observam que as três modalidades mais comuns entre as técnicas de PD são a Análise de Contingente (*Contingent Valuation*, na Língua Inglesa), a Análise Conjunta (*Conjoint Analysis*, na Língua Inglesa) e a Escolha Declarada (*Stated Choice*, na Língua Inglesa), sendo a última a dominante na área de Transportes e a aplicada neste estudo.

Na modalidade de Escolha Declarada é apresentado ao entrevistado um conjunto de alternativas, dentre as quais ele deve escolher sua preferida. Cada alternativa é caracterizada por um conjunto de atributos, como tempo de viagem e de espera, preço e conforto, por exemplo. Uma característica positiva da modalidade de Escolha Declarada é que o ato de selecionar a alternativa preferida é uma tarefa comum nas situações reais de escolha (ORTÚZAR; WILLUMSEN, 2011).

### **2.5.2 Modelos de escolha discreta**

Os modelos de escolha discreta são ferramentas econométricas largamente usadas na área de Transportes para previsão de demanda, e seus fundamentos são aqui expostos com base em Ben-Akiva e Lerman (1985), Ortúzar e Willumsen (2011) e Train (2009).

Tais modelos estimam a probabilidade de que um indivíduo, quando confrontado com um conjunto finito de alternativas mutuamente excludentes, escolha aquela que considere mais atrativa, conforme será mostrado nos parágrafos seguintes desta mesma seção. A atratividade, também denominada utilidade, é função tanto dos atributos das alternativas quanto das características (em geral socioeconômicas) do tomador de decisão. Assim, a utilidade da alternativa  $i$  para o indivíduo  $n$  terá a seguinte especificação:

$$V_{in} = \beta_1 \cdot v_{in1} + \beta_2 \cdot v_{in2} + \beta_3 \cdot v_{in3} + \dots + \beta_K \cdot v_{inK} \quad (1)$$

sendo que  $v_{ink}$  são as variáveis explicativas observadas pelo modelador, e representam tanto os atributos da alternativa  $i$  quanto as características do indivíduo  $n$ . Já os coeficientes  $\beta_k$ , também chamados de utilidades marginais, são constantes que indicam a importância de cada  $v_{ink}$  na composição da utilidade  $V_{in}$ . No caso da eq. (1), os coeficientes  $\beta_k$  são os mesmos para todas as alternativas (ou seja, no caso de haver as alternativas  $i$  e  $j$ , por exemplo,  $\beta_{ik} = \beta_{jk} = \beta_k$ ) e, por este motivo, são chamados de coeficientes genéricos. Trata-se, portanto, de um caso específico (restrito) de uma formulação mais geral, utilizada aqui como exemplo para apresentação dos princípios dos modelos de escolha discreta. Cabe ainda comentar que a forma aditiva e linear quanto aos coeficientes é a mais comumente adotada para a função utilidade.

Tem-se que cada indivíduo  $n$ , ao decidir pela alternativa que lhe pareça mais atrativa (ou seja, que tenha o maior  $V_{in}$ ) dentro de seu conjunto de alternativas disponíveis  $C_n$ , está maximizando a utilidade de sua escolha (conforme introduzido na Seção 2.1.5 deste mesmo capítulo) e valendo-se do comportamento racional. Para Ben-Akiva e Lerman (1985), o termo “comportamento racional” está relacionado às crenças de um observador externo sobre qual deveria ser o resultado de uma decisão tomada por um indivíduo. A definição por eles adotada caracteriza o comportamento racional como consistente e transitivo. A consistência está relacionada ao fato de que o indivíduo sempre toma a mesma decisão sob circunstâncias idênticas. E a transitividade significa que as escolhas do indivíduo obedecem à seguinte lógica: se a alternativa  $a$  é preferida em relação à  $b$ , e esta é preferida em relação à  $c$ , então a alternativa  $a$  é preferida em relação à  $c$ .

No entanto, o comportamento real dos indivíduos nem sempre se mostra, aos olhos do observador, consistente e transitivo, de modo que não necessariamente a alternativa com maior  $V_{in}$  é escolhida, indicando que o mecanismo de escolha pode ser entendido como probabilístico (e não determinístico).

Muito embora este comportamento não racional possa ser causado pelo tomador de decisão (quando ele, por exemplo, falha ao armazenar e processar informações sobre as alternativas e, como consequência, não escolhe a alternativa ótima), a *teoria da utilidade aleatória* assume que o indivíduo sempre escolhe a alternativa mais atrativa para si, e toda a aleatoriedade verificada no seu comportamento é atribuída a limitações do observador, que pode não conhecer ou não conseguir medir todos os elementos que influenciam o processo decisório.

Portanto, a utilidade passa a ser tratada como variável aleatória, composta pela soma de um componente determinístico, também chamado de sistemático e representado por  $V_{in}$ , (definido na eq. 1) e de um componente estocástico, também chamado de componente aleatório, não observado ou termo de erro e representado por  $\varepsilon_{in}$ . Assim, a utilidade da alternativa  $i$  para o indivíduo  $n$  passa a ser:

$$U_{in} = V_{in} + \varepsilon_{in} \quad (2)$$

De modo que o indivíduo  $n$  escolherá a alternativa  $i$  se a sua utilidade ( $U_{in}$ ), agora composta por um componente determinístico e um estocástico, for superior à utilidade de qualquer outra alternativa  $j$  disponível em seu conjunto de escolha  $C_n$ :

$$U_{in} > U_{jn}, \forall j \neq i \quad (3)$$

$$V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn}, \forall j \neq i$$

$$V_{in} - V_{jn} > \varepsilon_{jn} - \varepsilon_{in}, \forall j \neq i$$

No entanto, sendo o termo  $\varepsilon_{jn} - \varepsilon_{in}$  desconhecido pelo observador externo, não é possível afirmar que a desigualdade da eq. (3) seja verdadeira e que a alternativa  $i$  seja de fato escolhida pelo indivíduo  $n$ . Deste modo, a alternativa  $i$  será escolhida segundo a probabilidade:

$$P_n(i) = P[V_{in} - V_{jn} > \varepsilon_{jn} - \varepsilon_{in}, \forall j \neq i] \quad (4)$$

Assim, têm-se as probabilidades de escolha de cada alternativa presente em  $C_n$ , que são os resultados dos modelos de escolha discreta. No entanto, a derivação de um modelo de escolha discreta específico requer que se especifique uma distribuição de probabilidade para os componentes aleatórios da utilidade ( $\varepsilon_{in}$  e  $\varepsilon_{jn}$ , na eq. 4), conforme será descrito nos próximos dois itens, que são sucedidos por explicações sobre o processo de estimação dos modelos e sobre alguns procedimentos para avaliar os modelos estimados.

#### (a) O modelo logit

Quando se assume que os termos aleatórios da utilidade têm distribuições Gumbel idênticas e independentes, obtém-se o modelo logit multinomial (cuja sigla na Língua Inglesa é MNL, de *multinomial logit*). Sua equação para a probabilidade de escolha da alternativa  $i$  pelo indivíduo  $n$  tem forma fechada e é dada por:

$$P_n(i) = \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}}, \forall j \in C_n \quad (5)$$

Uma propriedade deste modelo é a independência de alternativas irrelevantes (cuja sigla na Língua Inglesa é IIA, de *Independence from Irrelevant Alternatives*). Segundo esta propriedade, a razão entre as probabilidades de duas alternativas que fazem parte de um conjunto de escolha  $C_n$  é constante e independente da existência de outras alternativas no conjunto. Ou seja, a taxa de substituição entre alternativas é constante:

$$\frac{P_n(i)}{P_n(q)} = \frac{e^{V_{in}} / \sum_j e^{V_{jn}}}{e^{V_{qn}} / \sum_j e^{V_{jn}}} = \frac{e^{V_{in}}}{e^{V_{qn}}} = e^{V_{in} - V_{qn}} \quad (6)$$

Assim, o modelo MNL é apropriado para situações em que as alternativas apresentem esta característica. Caso contrário, ele poderá gerar resultados falsos, superestimando ou subestimando a probabilidade de escolha das alternativas, como ocorre com alternativas que têm correlações não observadas entre si. A este respeito, Train (2009) argumenta que, quando o componente sistemático da utilidade é realmente bem especificado, explicitando todas as variáveis explicativas do comportamento e fazendo com que o termo de erro torne-se apenas ruído branco, a IIA é sempre respeitada.

Outra característica do modelo MNL é que ele não captura variações puramente aleatórias das preferências, muito embora seja possível representar as variações sistemáticas através de melhorias na especificação do componente determinístico da função utilidade (incluindo interações entre variáveis, por exemplo).

Mais uma particularidade deste modelo é que, sendo os termos de erro independentes uns dos outros, não é possível capturar correlações não observadas entre as várias respostas de um mesmo indivíduo.

### **(b) Os modelos logit mistos**

Os modelos logit mistos (referenciados pela sigla ML de *mixed logit*, na Língua Inglesa) são bastante flexíveis, visto que não apresentam as três limitações do modelo MNL mencionadas anteriormente, permitindo assim a inclusão de alternativas que apresentem correlações não observadas, capturando variações puramente aleatórias das preferências e também correlações não observadas entre as respostas de um mesmo indivíduo. No entanto, de forma similar ao modelo MNL, os termos de erro dos modelos ML também têm distribuições Gumbel idênticas e independentes.

Os modelos ML podem ser derivados sob diferentes especificações, no entanto apenas duas são de interesse desta pesquisa: (1) os modelos ML que capturam as variações aleatórias das preferências dos indivíduos e (2) os modelos ML que capturam as correlações não observadas entre as respostas do mesmo indivíduo. Cabe observar que estas duas especificações podem ser usadas em conjunto.

O primeiro dos modelos mencionados é denominado, na Língua Inglesa, ML *random coefficients* (coeficientes aleatórios), pois assume que os coeficientes  $\beta_k$  da parte determinística da função utilidade ( $V_{in}$ ) podem ser, ao invés de constantes (como no modelo MNL), variáveis aleatórias, assumindo valores diferentes entre os tomadores de decisão de acordo com uma função densidade de probabilidade (que pode ser contínua ou discreta). Assim, a probabilidade de que um indivíduo  $n$  escolha a alternativa  $i$  passa a ser:



$$P_n(i) = \int_{\beta} \left( \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}} \right) f(\beta) d\beta, \forall j \neq i \quad (7)$$

Ou seja, a probabilidade de escolha calculada segundo o modelo MNL ( $e^{V_{in}}/\sum_j e^{V_{jn}}$ ) é integrada sobre  $f(\beta)$ , a curva de distribuição de  $\beta$  (que aqui representa o vetor dos coeficientes  $\beta_k$ ).

Já o segundo modelo é denominado, na Língua Inglesa, de ML *panel*, em referência às pesquisas em que são coletadas várias repostas do mesmo indivíduo, simultaneamente ou ao longo do tempo (sendo que cada resposta está associada a um momento  $t$ ). Esta especificação captura as correlações não observadas entre as várias repostas do mesmo tomador de decisão. Para tanto, é acrescentada à função utilidade a variável  $\alpha_n$  (um termo aleatório de erro associado ao indivíduo  $n$ ), e a equação da utilidade para o indivíduo  $n$ , escolhendo a alternativa  $i$  no momento  $t$  passa a ser:

$$U_{int} = V_{int} + \alpha_n + \varepsilon_{int} \quad (8)$$

Vale lembrar que, para esta especificação, os termos de erro  $\varepsilon_{int}$  permanecem independentes entre os indivíduos, entre as repostas do mesmo indivíduo e entre as alternativas.

Tem-se, portanto, que a probabilidade de que o indivíduo  $n$  faça determinada sequência de escolhas (cada uma associada a um momento  $t$ ) é:

$$P_n(i_1, i_2, \dots, i_T) = \int_{\alpha} \prod_t \left[ \frac{e^{V_{int} + \alpha_n}}{\sum_j e^{V_{jnt} + \alpha_n}} \right] f(\alpha) d\alpha, \forall j \neq i \quad (9)$$

Ou seja, o produtório em  $t$  da probabilidade de que o indivíduo  $n$  escolha a alternativa  $i$  no momento  $t$  ( $e^{V_{int} + \alpha_n}/\sum_j e^{V_{jnt} + \alpha_n}$ ) é integrado sobre a curva de distribuição de  $\alpha$ , que aqui representa o vetor de todos os  $\alpha_n$ .

### (c) Processo de estimação dos coeficientes

O processo de estimação de um modelo de escolha discreta parte das observações de uma amostra de indivíduos e suas escolhas, e consiste em encontrar valores para os

coeficientes da função utilidade que melhor retratem o comportamento observado da amostra (os coeficientes  $\beta_k$ , no caso do modelo MNL; os parâmetros que descrevem a distribuição  $f(\beta)$ , no caso do modelo ML *random coefficients*; e também os parâmetros que descrevem a distribuição  $f(\alpha)$ , no caso do modelo ML *panel*).

O procedimento clássico para estimar os modelos já discutidos é a máxima verossimilhança, e os elementos essenciais desta técnica serão apresentados a seguir, com foco no modelo MNL. Detalhamentos específicos para a estimação de modelos ML (por envolverem etapas de simulação), não fazem parte do escopo desta seção.

A máxima verossimilhança  $\mathcal{L}^*(\beta)$  de uma amostra de observações é calculada como o produtório em  $n$  do produto das probabilidades de escolha de um indivíduo  $n$  para cada alternativa  $j$  pertencente a  $C_n$ :

$$\mathcal{L}^*(\beta) = \prod_n \prod_j P_n(j)^{y_{jn}} \quad (10)$$

onde  $P_n(j)$  é definido conforme a eq. (5), sendo portanto uma função do vetor  $\beta$ . Já  $y_{jn}$  é obtido a partir das observações da amostra e revela se a alternativa  $j$  foi ou não selecionada pelo indivíduo  $n$  (assumindo o valor 1 no caso de ter sido escolhida, e 0 caso contrário). Desta forma,  $\mathcal{L}^*(\beta)$  fornece a probabilidade de ocorrência das escolhas observadas na amostra quando elas são calculadas a partir do vetor  $\beta$  a ser estimado. Assim, tem-se que, quanto maior  $\mathcal{L}^*(\beta)$ , mais próximos da realidade estão os resultados que o modelo produz.

Por conveniência, costuma-se trabalhar com o logaritmo da máxima verossimilhança,  $\mathcal{L}(\beta)$  – de modo que  $\ln \mathcal{L}^*(\beta) = \mathcal{L}(\beta)$  – que é maximizado a fim de encontrar o vetor  $\beta$ :

$$\max_{\beta} \mathcal{L}(\beta) \quad (11)$$

$$\frac{\partial \mathcal{L}(\beta)}{\partial \beta_k} = 0$$

Tal procedimento fornece o valor médio de cada um dos coeficientes do vetor  $\beta$ . Já as variâncias dos coeficientes são obtidas através da matriz de variância e covariância assintótica do modelo. Para seu cálculo, é necessário obter primeiro a matriz  $K \times K$  com as derivadas de

segunda ordem de  $\mathcal{L}(\beta)$  em relação aos coeficientes já estimados, denominada matriz Hessiana e representada por  $\nabla^2 \mathcal{L}(\beta)$ . A partir dela obtém-se a matriz de variância e covariância assintótica:

$$\Omega_N = -\mathcal{E} [\nabla^2 \mathcal{L}(\beta)]^{-1} \quad (12)$$

Os elementos da diagonal de  $\Omega_N$  são as variâncias estimadas dos coeficientes  $\beta_k$ , cujas raízes quadradas são os erros padrão.

#### (d) Elementos para análise e comparação dos modelos estimados

A escolha de um bom modelo evolui a criação de especificações alternativas para a função utilidade, bem como métodos para analisar e comparar os modelos. Os métodos mais comumente usados (e os empregados neste estudo) serão apresentados a seguir, e se dividem em métodos para avaliar os coeficientes estimados e métodos para avaliar os modelos como um todo.

Fazem parte da avaliação dos coeficientes:

- a comparação do sinal do coeficiente estimado com o sinal esperado para ele;
- a aplicação do *teste estatístico t* (fruto da divisão do valor estimado para o coeficiente por seu erro padrão), que testa a hipótese nula de que o coeficiente é igual a zero ( $H_0: \beta_k = 0$ ) para determinado nível de significância (geralmente 5%);
- o cálculo das *taxas marginais de substituição* ou equivalências entre as variáveis independentes do modelo (indicando quais quantidades de cada variável geram a mesma utilidade), seguido da comparação com as taxas encontradas em outros estudos.

O detalhamento do conceito de taxa marginal de substituição requer um exemplo ilustrativo. Suponha-se que a parte determinística da utilidade da alternativa  $i$  para o indivíduo  $n$  seja função dos atributos preço ( $p_{in}$ ), tempo de viagem ( $t_{in}$ ) e conforto ( $c_{in}$ ):

$$V_{in} = \beta_{preço} \cdot p_{in} + \beta_{tempo} \cdot t_{in} + \beta_{conforto} \cdot c_{in}$$

Caso se deseje reduzir o tempo de viagem em uma unidade, sem que  $V_{in}$  se altere, será necessário que alguma outra variável se modifique (o preço, digamos). Para descobrir qual a variação  $\delta$  no preço capaz de manter  $V_{in}$  constante quando uma unidade é subtraída do tempo, deve-se reescrever a equação da parte determinística da utilidade:

$$V'_{in} = \beta_{preço} \cdot (p_{in} + \delta) + \beta_{tempo} \cdot (t_{in} - 1) + \beta_{conforto} \cdot c_{in}$$

Igualando  $V_{in}$  e  $V'_{in}$  e eliminando o termo  $\beta_{conforto} \cdot c_{in}$  (comum a ambas as equações) tem-se:

$$\beta_{preço} \cdot p_{in} + \beta_{tempo} \cdot t_{in} = \beta_{preço} \cdot (p_{in} + \delta) + \beta_{tempo} \cdot (t_{in} - 1)$$

$$\beta_{preço} \cdot p_{in} + \beta_{tempo} \cdot t_{in} = \beta_{preço} \cdot p_{in} + \beta_{preço} \cdot \delta + \beta_{tempo} \cdot t_{in} - \beta_{tempo}$$

$$0 = \beta_{preço} \cdot \delta - \beta_{tempo}$$

$$\delta = \frac{\beta_{tempo}}{\beta_{preço}}$$

Assim, tem-se que, ao reduzir uma unidade do tempo de viagem, é necessário aumentar o preço em  $\delta$  unidades para que a utilidade não se altere (para que  $V_{in} = V'_{in}$ ). E o valor de  $\delta$  é igual à razão  $\beta_{tempo}/\beta_{preço}$ , também chamada de taxa marginal de substituição entre os atributos tempo e preço. Outras taxas marginais de substituição (entre outros pares de atributos) podem ser calculadas, porém o exemplo dado corresponde a uma taxa comumente usada e também denominada *valor da economia do tempo de viagem* (ou simplesmente *valor do tempo*), que costuma variar conforme o propósito da viagem e a renda do indivíduo ou de seu domicílio.

Quando o denominador das taxas marginais de substituição é o coeficiente do preço (ou do custo), elas representam a *disposição a pagar* por alterações no atributo do numerador. A disposição a pagar pode ser interpretada como a quantidade de dinheiro que o indivíduo está disposto a desembolsar pelo acréscimo ou decréscimo (sempre aquele que elevar a utilidade) no valor de outro atributo.

No que se refere a avaliação e comparação de modelos, duas medidas de ajuste comumente usadas são o *rho quadrado* ( $\rho^2$ ) e o *rho quadrado ajustado* ( $\bar{\rho}^2$ ), que permitem comparar modelos estimados a partir da mesma amostra de observações:

$$\rho^2 = 1 - \frac{\mathcal{L}(\beta)}{\mathcal{L}(0)} = \frac{\mathcal{L}(0) - \mathcal{L}(\beta)}{\mathcal{L}(0)} \quad (13)$$

$$\bar{\rho}^2 = 1 - \frac{\mathcal{L}(\beta) - K}{\mathcal{L}(0)} = \frac{\mathcal{L}(0) - \mathcal{L}(\beta) + K}{\mathcal{L}(0)} \quad (14)$$

sendo que  $\mathcal{L}(\beta)$  é o logaritmo da máxima verossimilhança calculada a partir dos coeficientes estimados, e  $\mathcal{L}(0)$  é o logaritmo da máxima verossimilhança obtida quando todos os coeficientes são iguais a zero (e portanto a probabilidade de escolha de cada alternativa é igual a  $1/J$ , sendo  $J$  a quantidade de alternativas disponíveis), o que equivale a não haver modelo algum. O rho quadrado mede o quanto o modelo estimado é superior, quanto à sua capacidade de prever o comportamento da amostra, em relação a um modelo em que todas as alternativas tenham a mesma probabilidade de escolha ( $1/J$ ). Esta medida de ajuste varia de 0, quando o modelo estimado não consegue explicar o comportamento da amostra (ou seja,  $\mathcal{L}(\beta) = \mathcal{L}(0)$ ) a 1, quando o modelo estimado prediz com perfeição o comportamento da amostra ( $\mathcal{L}^*(\beta) = 1$  e, por consequência,  $\mathcal{L}(\beta) = 0$ ).

O rho quadrado ajustado é uma modificação do rho quadrado (conforme mostra a eq. 14), que permite comparar modelos estimados a partir da mesma amostra de observações, mas que tenham quantidades diferentes de parâmetros (onde  $K$  é a quantidade de parâmetros estimados). Pode-se dizer que, para modelos estimados a partir da mesma amostra, aquele que apresentar maior rho quadrado ajustado, independente da quantidade de coeficientes estimados, tem melhor ajuste.

Outra maneira de comparar diferentes especificações, desde que a estimação seja feita a partir da mesma amostra, é o teste da razão da máxima verossimilhança. Supondo que a um modelo sejam impostas restrições (como igualar todos ou alguns coeficientes a zero, ou então igualar dois ou mais coeficientes entre si), é possível testar a hipótese de que as restrições são verdadeiras com o teste da razão da máxima verossimilhança. Seja  $H_0$  a hipótese nula (de que dois coeficientes são iguais a zero, por exemplo) que se deseja testar:

$$H_0: \beta_1 = \beta_2 = 0$$

E seja  $\mathcal{L}^*(\beta_{SR})$  a máxima verossimilhança do modelo estimado sem a restrição (ou seja, a partir de uma especificação de  $V_{in}$  que contenha  $\beta_1$  e  $\beta_2$ ), e  $\mathcal{L}^*(\beta_{CR})$  a máxima verossimilhança do modelo estimado com a restrição imposta por  $H_0$  (ou seja, a partir de uma especificação de  $V_{in}$  que não contenha  $\beta_1$  e  $\beta_2$ , ou que fixe seus valores em zero). O teste da razão da máxima verossimilhança é assim definido:

$$-2 \cdot \ln \left[ \frac{\mathcal{L}^*(\beta_{CR})}{\mathcal{L}^*(\beta_{SR})} \right] = -2 \cdot [\mathcal{L}(\beta_{CR}) - \mathcal{L}(\beta_{SR})] \quad (15)$$

Esta estatística tem distribuição chi quadrado com número de graus de liberdade igual à quantidade de restrições impostas ao modelo (duas, para o exemplo usado em que  $H_0: \beta_1 = \beta_2 = 0$ ). Se o valor da estatística exceder o valor crítico do chi quadrado com o número apropriado de graus de liberdade, a hipótese nula pode ser rejeitada. Nas aplicações práticas, este teste geralmente é usado para comparar o modelo estimado (irrestrito) com aquele em que todos os coeficientes são iguais a zero (restrito).

### 2.5.3 Experimentos de escolha declarada e desenhos eficientes

A criação de um experimento de escolha declarada, para Bliemer e Rose (2009), é composta por três etapas, que são descritas nos parágrafos seguintes.

A **primeira** delas é a especificação do modelo de escolha discreta que se pretende estimar a partir dos dados coletados com o experimento. Para tanto, são necessárias as definições listadas abaixo, acompanhadas de comentários e exemplos:

#### 1. Alternativas que farão parte do experimento

No caso de um experimento sobre escolha de rota com duas alternativas, por exemplo, elas podem ser *rota a* e *rota b*. Já para um experimento de escolha de modo com três alternativas, elas podem ser *ônibus*, *automóvel* e *trem*;

## 2. Atributos que caracterizarão cada alternativa

Cabe considerar que há atributos comuns a algumas ou a todas as alternativas, e também há atributos exclusivos para uma alternativa. No caso do experimento sobre escolha de modo citado anteriormente, o tempo de viagem é um atributo comum às alternativas ônibus, automóvel e trem. Já o atributo quantidade de transbordos só caracteriza as alternativas ônibus e trem;

## 3. Tipo de modelo econométrico a ser estimado

MNL, ML *panel*, ML *random coefficients* etc.;

## 4. Especificação do componente determinístico da função utilidade para todas as alternativas

Este componente pode incluir, além dos efeitos principais, efeitos de interações e também efeitos não lineares. A título de exemplo, para o caso do experimento de escolha de modo seria necessário criar o componente determinístico da função utilidade para os três modos.

A **segunda** etapa é a geração do desenho do experimento, que nada mais é do que um grupo de *situações de escolha*. Uma situação de escolha é uma combinação de valores de atributos para as alternativas. A Tabela 1 (abaixo) mostra um exemplo de desenho em que são exibidas três situações de escolha criadas para o experimento de escolha de rota mencionado anteriormente:

Situação de escolha	Rota a		Rota b	
	Custo	Tempo de viagem	Custo	Tempo de viagem
1	15 \$	35 min	20 \$	25 min
2	20 \$	25 min	15 \$	30 min
3	20 \$	30 min	25 \$	25 min
...	...	...	...	...

**Tabela 1 – Exemplos de situações de escolha de um desenho de experimento**

A criação do desenho também requer definições prévias, que são listadas abaixo, acompanhadas de comentários e exemplos (a numeração desta lista é contínua à numeração da lista das definições da primeira etapa):

5. Forma de identificação das alternativas: com rótulos ou não

Geralmente experimentos de escolha de modo têm suas alternativas rotuladas: ônibus, automóvel, trem etc. Já experimentos de escolha de rota identificam as alternativas de maneira genérica: rota a, rota b, etc.;

6. Balancear níveis dos atributos no desenho: sim ou não

Balancear os níveis dos atributos significa que cada nível de um atributo deve aparecer no desenho a mesma quantidade de vezes, a fim de que todos os pontos do intervalo de variação do atributo sejam igualmente representados. Esta recomendação pode elevar a quantidade mínima de situações de escolha a serem geradas, que passa a ser o mínimo múltiplo comum dos níveis dos diferentes atributos (por exemplo: se o experimento tem 3 atributos, o primeiro com 3 níveis, o segundo com 4 e o terceiro com 2, então deverão ser criadas 12 situações de escolha para garantir o equilíbrio entre os níveis de cada atributo). Embora haja circunstâncias em que esta característica não seja desejada, ela geralmente é adotada em desenhos de experimento;

7. Intervalo de variação de cada atributo

É o intervalo entre seu menor e seu maior nível, também chamado de amplitude do atributo. Novamente no exemplo da Tabela 1, ambos os atributos têm intervalo de variação igual a 10 unidades (\$ ou minutos).

Não obstante os fatos de que amplitudes maiores são preferíveis sob o ponto de vista estatístico, e de que os modelos são aplicáveis apenas para os intervalos dos atributos a partir dos quais foram estimados, a definição do intervalo de variação deve ser balizada por aspectos práticos, tal como o realismo do contexto do experimento;

8. Níveis de cada atributo

Uma vez definidos os intervalos de variação dos atributos, dá-se a definição de seus níveis, que depende de quantos se deseja criar e dos intervalos entre níveis consecutivos. No exemplo da Tabela 1, o atributo custo tem 3 níveis equidistantes (15 \$, 20 \$ e 25 \$), assim como o atributo tempo de viagem (25, 30 e 35 minutos);

9. Quantidade  $S$  de situações de escolha a serem criadas

Recomenda-se que o número de diferentes situações de escolha a serem criadas seja, no mínimo, igual ao número de parâmetros do modelo de escolha discreta que se



pretende estimar a partir dos resultados do experimento. Cabe comentar que  $S$  depende também dos níveis dos atributos, caso se deseje manter seu balanceamento;

#### 10. Tipo de desenho de experimento a ser adotado

Dois categorias de desenho podem ser citadas: o desenho fatorial completo (Full Factorial Design, na Língua Inglesa) e os fatoriais fracionados (Fractorial Factorial Designs, na Língua Inglesa). O primeiro consiste em gerar todas as situações de escolha possíveis. No exemplo da Tabela 1, tanto para a rota  $a$  quanto a rota  $b$  há 9 combinações possíveis, que, quando cruzadas, resultam em 81 ( $9^2$ ) diferentes situações de escolha. Já os desenhos fatoriais fracionados são uma seleção de situações de escolha a partir do desenho fatorial completo. Um tipo de desenho fatorial fracionado bastante conhecido é o desenho ortogonal, em que a correlação entre os níveis dos atributos do desenho é igual a zero. Outro tipo é o desenho eficiente (o escolhido para esta pesquisa), que minimiza as variâncias previstas para os parâmetros que serão estimados pelos modelos de escolha discreta. Estes desenhos serão detalhados ainda nesta seção.

Finalmente, a **terceira** etapa para a criação de um experimento de escolha declarada consiste na construção do questionário que será apresentado aos entrevistados, quer seja em papel, através do computador, presencialmente ou pela internet etc. Um elemento importante que diz respeito a esta etapa é a aleatorização das situações de escolha: não devem ser apresentadas sempre na mesma ordem aos entrevistados. É possível também aleatorizar a sequência em que os atributos aparecem dentro de cada situação de escolha.

Os parágrafos seguintes desta seção são dedicados à descrição dos desenhos eficientes (seu embasamento teórico e requerimentos à sua aplicação) e as informações expostas se baseiam fortemente em Bliemer e Rose (2009), Ortúzar e Willumsen (2011) e no manual do software Ngene 1.1 (CHOICE METRICS, 2011).

#### **Desenhos eficientes**

Esta modalidade de desenho gera um grupo de situações de escolha tal que, quando transformado em questionário de pesquisa e aplicado a uma amostra, permite estimar modelos de escolha discreta cujos parâmetros tenham as menores (co)variâncias possíveis (sendo, portanto, mais confiáveis e qualificando o desenho como um desenho eficiente), sem com isto

requerer amostras maiores. Alternativamente, os desenhos eficientes permitem atingir (co)variâncias para os parâmetros iguais às obtidas através de outros tipos de desenhos, porém usando amostras menores.

Idealmente, o procedimento para encontrar o melhor desenho eficiente exigiria que, para todas as combinações de  $S$  situações de escolha que poderiam ser formadas a partir do desenho fatorial completo, fossem calculadas medidas de ineficiência. E a combinação das  $S$  situações de escolha com a menor ineficiência seria considerada o desenho ótimo. No entanto, em geral a quantidade de combinações possíveis é tão grande que torna inviável este procedimento. Portanto, são usados algoritmos para selecionar desenhos e medir sua ineficiência, em um processo iterativo até que um desenho satisfatoriamente eficiente seja encontrado (ou seja, com a medida de ineficiência suficientemente baixa), muito embora a descrição de tais algoritmos esteja fora do escopo deste trabalho.

Medir a ineficiência de um desenho é medir o “quão pequena” está a (co)variância dos parâmetros associados a ele, através do cálculo da matriz de variância e covariância assintótica  $\Omega_N$  (definida na eq. 12). No entanto, os valores estimados para os parâmetros, necessários para o cálculo de  $\Omega_N$ , não estão disponíveis, pois somente serão obtidos após a aplicação do desenho a uma amostra, seguida da estimação dos modelos. Para contornar esta dificuldade, são usadas aproximações para os parâmetros: valores que se acredita serem próximos aos que serão obtidos quando o modelo for estimado. Resultados de outras pesquisas ou mesmo experimentos piloto podem ser usados para definir as aproximações para cada parâmetro, que podem ser valores fixos ou aleatórios (com distribuições de probabilidade geralmente normais ou uniformes). Caso não se tenha conhecimento algum sobre o valor dos parâmetros, o valor zero é usado no lugar das aproximações. Detalhamentos do cálculo de  $\Omega_N$ , que pode tanto utilizar métodos de simulação quanto ser feito analiticamente, estão fora do escopo deste trabalho.

Um indicador de ineficiência bastante usado (e o adotado nesta pesquisa) é o  $D_{error}$ , calculado como o determinante da matriz  $\Omega_1$ , que é a matriz de variância e covariância para apenas um indivíduo:

$$D_{error} = \det(\Omega_1)^{1/L} \quad (16)$$

Quando são usadas distribuições de probabilidades para as aproximações dos parâmetros, tem-se o *Bayesian D\_error* ( $D_{b\_error}$ ) dado por:

$$D_{b\_error} = \int_{\beta} \det(\Omega_1)^{1/L} f(\beta) d\beta \quad (17)$$

Para que ambos os indicadores se tornem independentes do tamanho do experimento (ou seja, da quantidade L de parâmetros estimados), a matriz  $\Omega_1$  é normalizada pela potência  $1/L$ .

Outro indicador interessante relacionado aos desenhos eficientes é o  $S\_estimate$ , que fornece uma indicação do tamanho mínimo da amostra para que se estimem parâmetros significativamente diferentes de zero. Os tamanhos mínimos de amostra são calculados para cada um dos parâmetros, de modo que o maior tamanho de amostra calculado corresponde à indicação da amostra mínima para o desenho. No entanto, a realização do experimento com a amostra do tamanho indicado não garante parâmetros estatisticamente significativos, de modo que o uso deste indicador é mais apropriado para comparar desenhos diferentes do que para determinar os tamanhos das amostras.

Vale mencionar também o  $B\_estimate$ , uma medida do grau de balanceamento das utilidades das alternativas que fazem parte de cada situação de escolha do desenho. Este indicador está relacionado à quantidade de informação (para fins de modelagem) que pode ser obtida a partir de cada situação. Para maximizar esta quantidade de informação, o desenho eficiente é gerado de modo que as utilidades das alternativas de uma mesma situação de escolha não sejam nem muito próximas (pelo risco de que o indivíduo selecione uma alternativa aleatoriamente), nem muito distantes (a ponto de tornar uma alternativa muito superior à outra e dispensar o uso de regras compensatórias de decisão, como ocorre quando há alternativas dominantes). Não há um valor ideal para o  $B\_estimate$ , mas ele costuma variar entre 70% e 90% para os desenhos eficientes.

Além das definições necessárias à primeira e segunda etapas da criação de um experimento, cabe formalizar mais quatro definições prévias à criação do desenho

(acompanhadas de comentários e exemplos), bem como uma observação pertinente a respeito do modelo econométrico e sua especificação. A numeração desta lista é contínua à da lista das definições da segunda etapa:

11. Aproximações para os parâmetros do componente determinístico da função utilidade

Conforme mencionado anteriormente, podem ser valores fixos ou aleatórios. Dado que o desenho eficiente depende fortemente da acurácia das aproximações fornecidas, igualar as aproximações a zero só é recomendado em caso de total desconhecimento dos valores dos parâmetros;

12. Indicador de ineficiência a ser usado pelos algoritmos

É necessário definir qual indicador se deseja que o algoritmo utilize (para esta pesquisa foi escolhido o  $D_b\_error$ );

13. Restrições ao desenho

É possível impor restrições ao desenho, caso se deseje evitar determinadas combinações de níveis de atributos. No exemplo da Tabela 1, por exemplo, é possível evitar situações de escolha em que (pelo menos) uma das alternativas tenha custo muito alto e tempo de viagem igualmente alto, por se considerar que tal combinação não seria realista;

14. Quantidade de blocos

É possível criar blocos de situações de escolha, de modo que o mesmo entrevistado não precise responder a todas as situações do desenho. Um desenho com 24 situações de escolha, por exemplo, pode ter 2 blocos de 12 situações ou 3 de 8.

Cabe comentar que a quantidade de situações de escolha à qual cada indivíduo é submetido não pode ser grande a ponto de causar fadiga (visto que ela influencia os mecanismos de escolha);

15. Sobre o tipo de modelo econométrico e sobre a especificação do componente determinístico da função utilidade

É importante ressaltar que as matrizes  $\Omega_N$  e  $\Omega_1$  dependem tanto do modelo econométrico adotado quanto da especificação da parte determinística da utilidade. Portanto, diferentes modelos (no que diz respeito ao tipo e à especificação) produzem

diferentes matrizes  $\Omega_N$  e  $\Omega_1$ , de modo que um desenho gerado para determinado modelo não necessariamente produzirá resultados igualmente eficientes para outro. Apesar disto, é comum adotar especificações MNL mesmo quando se pretende estimar outros tipos de modelos após a coleta dos dados (como modelos ML, por exemplo), pelo fato de que o tempo computacional para gerar desenhos eficientes a partir de especificações MNL é pequeno.

### 3 METODOLOGIA

A natureza da pesquisa desenvolvida é descritiva (no que se refere aos seus objetivos) e experimental (no que se refere aos procedimentos técnicos utilizados).

De acordo com Gil (2002, p. 42), “as pesquisas descritivas têm como objetivo primordial a descrição das características de determinada população ou fenômeno ou, então, o estabelecimento de relações entre variáveis”. Já as pesquisas experimentais, também de acordo com Gil (2002, p. 47), consistem essencialmente em “determinar um objeto de estudo, selecionar as variáveis que seriam capazes de influenciá-lo, definir as formas de controle e de observação dos efeitos que a variável produz no objeto”.

Neste trabalho, o fenômeno estudado (através de técnica experimental) é o comportamento de escolha de linha de ônibus, que é função dos seis atributos já descritos nos objetivos do estudo, na Seção 1.1: (1) existência de painel no ponto, (2) intervalo de tempo programado para a passagem de veículos consecutivos da mesma linha, (3) possível atraso em relação a este intervalo programado, (4) tempo de viagem dentro do veículo até o destino, (5) ocupação do veículo quando chega ao ponto e (6) valor da tarifa. Usando a terminologia de Gil (2002), entende-se que os atributos são as variáveis que influenciam o fenômeno estudado.

A seguir serão apresentados os métodos de coleta e de análise de dados escolhidos para o desenvolvimento da pesquisa.

#### 3.1 Métodos para coleta de dados

Para a coleta de dados foi realizado um experimento de escolha declarada com desenho eficiente aplicado à comunidade da Universidade de São Paulo (USP), que foi precedido por dois experimentos piloto, também de escolha declarada e com desenho eficiente, conforme mostra a Figura 1. Os resultados do primeiro experimento piloto (taxas marginais de substituição entre atributos) foram utilizados para criar o desenho eficiente do segundo experimento piloto. E os resultados deste último (intervalos de variação dos

coeficientes dos atributos e taxas marginais de substituição entre atributos), por sua vez, foram utilizados para criar o desenho eficiente do experimento final. A elaboração e os resultados do primeiro e do segundo experimento piloto são apresentados, respectivamente, nos Apêndices A e B, enquanto a elaboração do experimento final é objeto do Capítulo 4 e seus resultados são descritos no Capítulo 5.

A Figura 1 também mostra as definições que precederam a realização do experimento final, com a mesma nomenclatura utilizada no Capítulo 4.

Cabe comentar que a ferramenta Ngene 1.0 ([www.choice-metrics.com](http://www.choice-metrics.com)) foi utilizada para criar os desenhos eficientes dos dois experimentos piloto e do experimento final.

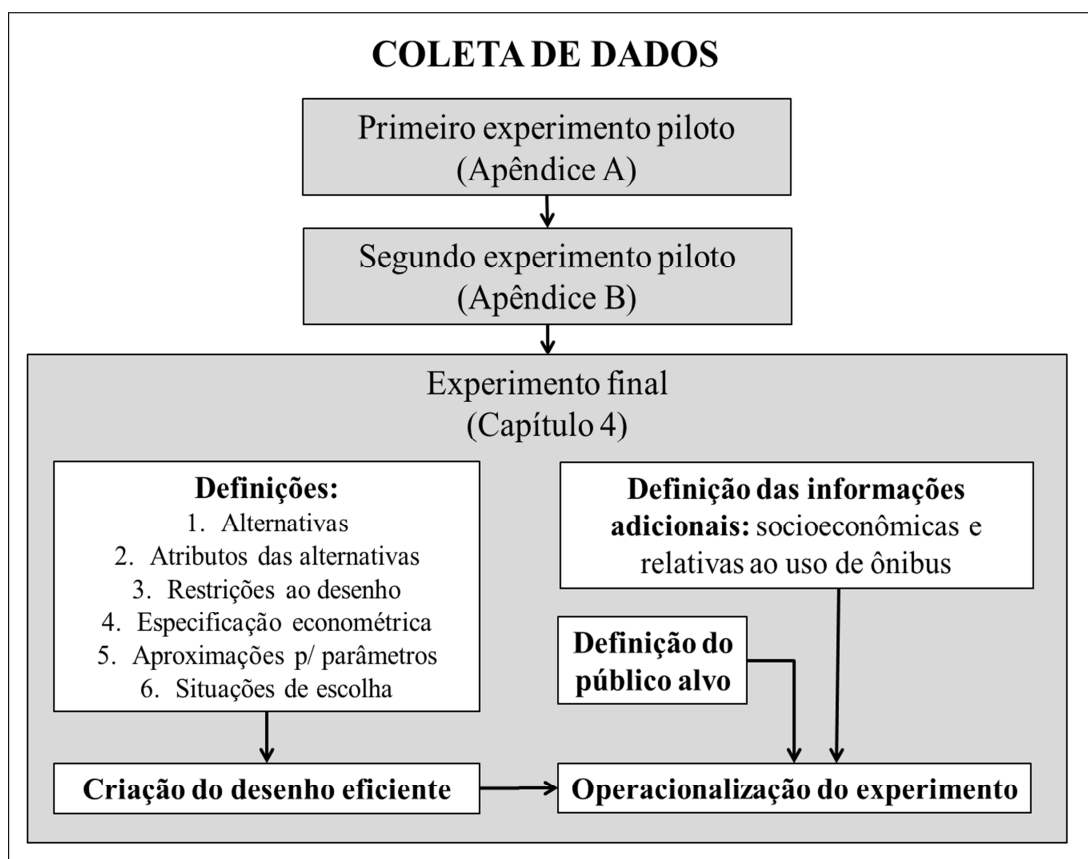


Figura 1 – Sequência de etapas para coleta de dados

### 3.2 Métodos para análise de dados

As etapas da análise de dados (ou seja, das repostas obtidas no experimento final) são exibidas na Figura 2, e desenvolvidas no Capítulo 5.

A primeira etapa, de caracterização da amostra, refere-se à classificação socioeconômica e relativa a hábitos de uso de ônibus dos entrevistados. Todas as etapas consecutivas a ela referem-se à estimação de modelos de escolha discreta *ML panel* a partir das respostas de todos os entrevistados da amostra.

O *modelo base* contém apenas os efeitos principais dos atributos das alternativas, enquanto os demais visam estudar a não linearidade das preferências, as interações entre atributos e a variabilidade sistemática das preferências dos entrevistados, esta última explorada através de um modelo com interações entre os atributos e as variáveis socioeconômicas e de uso de ônibus dos entrevistados.

Todos os modelos estimados respondem à questão de pesquisa apresentada no Capítulo 1, ao objetivo geral e ao primeiro objetivo específico da pesquisa (descritos na Seção 1.1). Os *modelos com interações entre atributos*, especificamente, também respondem ao segundo objetivo específico, enquanto o *modelo com variabilidade sistemática das preferências* responde também ao terceiro objetivo específico.

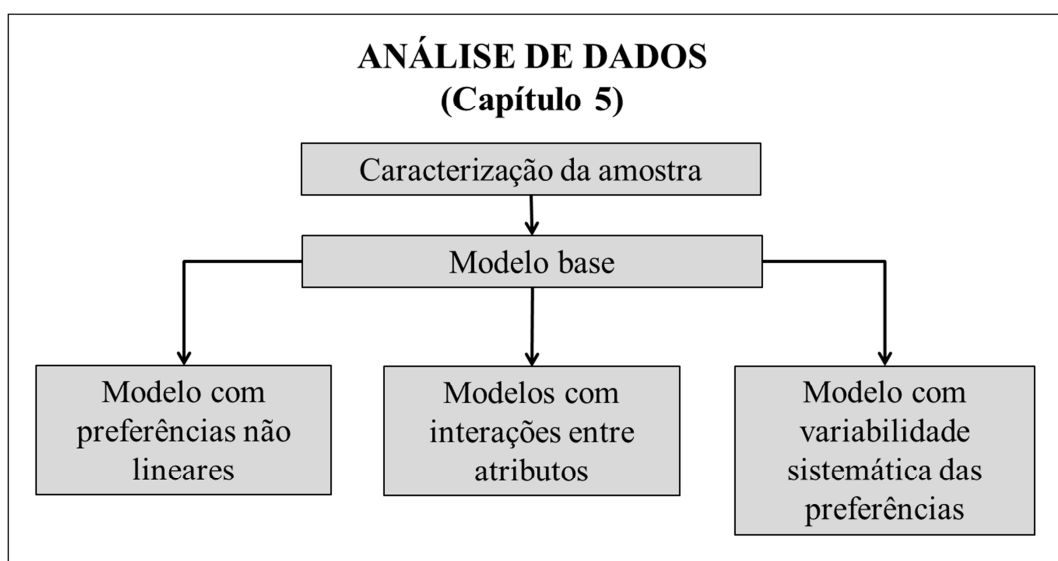


Figura 2 – Sequência de etapas para análise de dados



Encontrar os melhores modelos, que são os apresentados neste trabalho, envolveu um processo iterativo de especificação, estimação e análise de modelos. Ao todo, aproximadamente mil modelos foram estimados, todos usando o software Biogeme (BIERLAIRE, 2003). Especificamente para os modelos que envolviam interações, a impossibilidade de estimar todas as especificações possíveis (que no caso do *modelo com variabilidade sistemática das preferências*, chegava à ordem de  $10^{50}$ ), exigiu a utilização de heurísticas para que se chegasse a bons modelos dentro do período de tempo disponível para a realização das análises. A principal delas consistiu em incluir na especificação todas as interações de interesse e, após estimar o modelo, a interação cujo coeficiente tivesse o menor valor do teste  $t$  (desde que fosse estatisticamente não significativo) era retirada da especificação. O modelo era novamente estimado e o procedimento era repetido até que restassem apenas coeficientes estatisticamente significativos no modelo.

Cabe salientar que os modelos estimados usando as heurísticas tinham todos especificação MNL. Somente após um bom modelo ser encontrado, sua especificação era transformada em ML *panel*, em geral sem grandes alterações nos resultados. Este procedimento teve grande impacto no tempo total para a estimação das centenas de modelos criados durante a aplicação das heurísticas, visto que o tempo de processamento computacional para estimar um MNL é bastante inferior ao tempo para estimar um ML *panel*.

## 4 EXPERIMENTO COM DESENHO EFICIENTE

É objeto deste capítulo a elaboração do experimento para coleta de dados usado nesta pesquisa, que faz uso de técnicas de escolha declarada e foi construído a partir de um desenho eficiente (sendo que as atividades relativas à elaboração e envio do questionário à amostra já foram descritas no Capítulo 3). No entanto, antes da descrição das etapas de sua elaboração, faz-se necessária uma exposição prévia do contexto do experimento.

Em cada situação de escolha foram apresentadas ao entrevistado duas linhas diferentes de ônibus que podem levá-lo até um destino determinado para que compareça a um compromisso. As linhas, no entanto, têm itinerários diferentes e não passam pelo mesmo ponto, de modo que a escolha a ser feita é entre uma ou outra combinação de linha e ponto. Adicionalmente, apenas um dos pontos tem painel eletrônico exibindo os intervalos de tempo **previstos** para passagem das linhas.

Além da **existência de painel no ponto**, as alternativas se diferenciam pelos seguintes atributos:

- (1) intervalo de tempo **programado** para a passagem de veículos consecutivos da mesma linha;
- (2) possível atraso em relação a este intervalo programado;
- (3) tempo de viagem dentro do veículo até o destino;
- (4) ocupação do veículo quando chega ao ponto;
- (5) custo da viagem (representado pelo valor da tarifa).

Abaixo é apresentada a instrução comum a todas as situações de escolha, seguida de um exemplo de tabela (na Figura 3) com as características de cada alternativa de uma situação de escolha:

*“Suponha que você irá de ônibus a um compromisso e que tanto a LINHA 1 quanto a LINHA 2 servem para levá-lo(a) até o local desejado. Porém, estas linhas não passam pelo mesmo ponto. Você tem, então, duas alternativas: **andar em uma direção até o PONTO 1 e pegar a LINHA 1** ou **andar em outra direção até o PONTO 2 e pegar a LINHA 2**.*

Considere que os tempos de caminhada até qualquer um dos pontos são pequenos e parecidos. Os pontos podem ou não ter **painel eletrônico** mostrando quanto tempo falta para o ônibus passar. Levando em conta esta e outras características de cada alternativa (apresentadas na tabela abaixo), **qual você escolhe?**”

	PONTO 1 + LINHA 1	PONTO 2 + LINHA 2
<b>A linha está programada para passar a cada:</b>	20 minutos	15 minutos
<b>Mas pode demorar até:</b>	39 minutos	29 minutos
<b>O ponto tem painel eletrônico?</b>	Sim	Não
<b>O tempo de viagem (dentro do ônibus) é:</b>	50 minutos	50 minutos
<b>A passagem custa:</b>	R\$ 3,00	R\$ 3,30
<b>O ônibus costuma estar:</b>	Quase vazio. Você poderá viajar sentado.	Nem vazio, nem lotado. Todos os bancos estão ocupados e você terá que viajar em pé (com poucas outras pessoas em pé).

Figura 3 – Exemplo de tabela com as características das alternativas

Cabe destacar as seguintes características do contexto do experimento:

- os pontos ficam em direções opostas, de modo que não é possível passar por um deles enquanto se dirige ao outro;
- o tempo de caminhada para acessar o ponto é considerado pequeno e semelhante para ambas as alternativas. Ou seja, trata-se de uma característica que não as diferencia (não torna uma superior à outra);
- o possível atraso da linha não é exibido diretamente nas situações de escolha, e sim indiretamente: é somado ao intervalo **programado** e esta soma é apresentada ao entrevistado como um tempo máximo de espera pelo ônibus;
- a razão de haver dois pontos diferentes é permitir que o painel esteja presente para apenas uma alternativa, o que não seria possível caso ambas as linhas de ônibus passassem pelo mesmo ponto de parada (supondo que o painel exiba informações sobre todas as linhas que passem pelo ponto);
- a decisão do entrevistado é tomada sem que ele conheça o intervalo de tempo **previsto** para a passagem da linha pelo ponto (ele sabe apenas que um ponto tem painel e outro não, mas não sabe qual a informação mostrada pelo painel).

As etapas para criação do desenho eficiente são descritas na próxima seção, que é seguida da apresentação de informações adicionais que foram coletadas dos entrevistados (dados socioeconômicos e referentes a hábitos de uso de ônibus) e de comentários a respeito da obtenção da amostra e da operacionalização do experimento.

## **4.1 Criação do desenho eficiente**

São objeto desta seção as definições prévias à criação do experimento com desenho eficiente, bem como uma breve descrição do desenho obtido.

Salienta-se que desenhos diferentes foram criados e avaliados, sob um conjunto de definições diferentes das apresentadas a seguir. No entanto, o conteúdo das seções seguintes refere-se apenas ao desenho escolhido para realização do experimento.

### ***4.1.1 Alternativas***

Conforme mencionado na introdução deste capítulo, são duas as alternativas deste experimento, e foram apresentadas sem rótulos: *Ponto 1 + Linha 1* e *Ponto 2 + Linha 2*.

As definições acima correspondem às listadas na Seção 2.5.3: *(1) Alternativas que farão parte do experimento* e *(5) Forma de identificação das alternativas: com rótulos ou não*.

### ***4.1.2 Atributos***

São seis os atributos escolhidos para o experimento (sendo que todos são comuns a ambas as alternativas) e estão detalhados na Tabela 2 (que traz também a notação de seus coeficientes no componente determinístico da função utilidade):

<b>Atributos</b>	<b>Níveis</b>	<b>Coefficientes na função utilidade</b>
Existência de <b>painel</b> no ponto	1 = Sim / 0 = Não	$\beta_{display}$
<b>Intervalo</b> de tempo programado para a passagem de veículos consecutivos da mesma linha	5 / 10 / 15 / 20 minutos	$\beta_{interv}$
Possível <b>atraso</b> em relação ao intervalo programado	4 / 9 / 14 / 19 minutos	$\beta_{atraso}$
<b>Tempo de viagem</b> dentro do veículo	20 / 30 / 40 / 50 minutos	$\beta_{tviagem}$
<b>Ocupação</b> do veículo quando chega ao ponto	Mínima ( <i>dummy</i> : ocup0) “ <i>Quase vazio. Você poderá viajar sentado</i> ”.	$\beta_{ocup0}$ e $\beta_{ocup1}$
	Intermediária ( <i>dummy</i> : ocup1) “ <i>Nem vazio, nem lotado. Todos os bancos estão ocupados e você terá que viajar em pé (com poucas outras pessoas em pé)</i> ”.	
	Máxima “ <i>Lotado. Todos os bancos estão ocupados e você terá que viajar em pé (com muitas outras pessoas em pé)</i> ”.	
<b>Custo</b> da viagem	3,00 / 3,30 R\$	$\beta_{custo}$

**Tabela 2 – Detalhamento dos atributos do experimento**

Para os atributos *intervalo*, *atraso* e *tempo de viagem*, a quantidade de níveis escolhida (quatro) permitiu representá-los com amplitude suficiente e, ao mesmo tempo, utilizar intervalos pequenos entre níveis consecutivos. O uso de apenas dois ou três níveis iria requerer que a amplitude fosse reduzida ou que fossem adotados intervalos maiores entre níveis consecutivos. E a adoção de intervalos maiores poderia gerar situações de escolha nas quais a combinação dos valores dos três atributos fosse muito mais atrativa para uma alternativa do que para outra, havendo o risco de que nem mesmo se os demais atributos da outra alternativa fossem bastante positivos (como presença do painel, tarifa menor ou baixa ocupação do veículo, por exemplo), eles seriam suficientes para compensar a desvantagem de

sua combinação de *intervalo*, *atraso* e *tempo de viagem*, caracterizando uma situação de escolha bastante desbalanceada.

Cabe comentar que, para o atributo *tempo de viagem*, mais especificamente, os níveis selecionados estão dentro do intervalo de duração das viagens com transporte público registradas pela Pesquisa Origem e Destino 2007 da RMSP (COMPANHIA DO METROPOLITANO DE SÃO PAULO, 2008).

Para o atributo *painel*, a definição de dois níveis é óbvia, enquanto que para o atributo *ocupação*, o fato de ser uma variável categórica motivou a escolha de três níveis (cujas descrições utilizadas no experimento são também apresentadas na Tabela 2) e o uso de duas variáveis *dummy* (para representar o atributo no componente determinístico da função utilidade). A primeira das variáveis *dummy* (*ocup0*), cujo coeficiente é  $\beta_{ocup0}$ , assume o valor 1 quando o nível de ocupação é mínimo, e 0 caso contrário. Já a outra variável (*ocup1*), cujo coeficiente é  $\beta_{ocup1}$ , assume o valor 1 quando o nível de ocupação é intermediário, e 0 caso contrário. Deste modo, quando a ocupação do veículo é máxima, tanto *ocup0* quanto *ocup1* assumem o valor 0. O nível máximo do atributo *ocupação* é dito nível base desta codificação *dummy*. A Tabela 3 a seguir sumariza a relação entre os níveis do atributo *ocupação* e as duas variáveis *dummy*:

Ocupação do veículo	ocup0	ocup1
Mínima	1	0
Intermediária	0	1
Máxima	0	0

**Tabela 3 – Relação entre atributo *ocupação* e variáveis *dummy***

No que se refere ao atributo custo, seu primeiro nível (R\$3,00) corresponde ao valor da tarifa municipal de ônibus no período da realização do experimento. Optou-se por limitar o nível máximo deste atributo a 10% acima do nível mínimo, de modo a manter o realismo do contexto do experimento. O nível máximo foi fixado, portanto, em R\$3,30. Nenhum nível intermediário (como R\$3,15, por exemplo) foi adicionado pelo risco de que diferenças menores entre as tarifas de duas alternativas (como 10 ou 15 centavos, por exemplo) não fossem suficientes para diferenciá-las quanto ao custo, implicando em situações de escolha com menores chances de coletar informações sobre o comportamento dos entrevistados.

O conteúdo desta seção corresponde às seguintes definições listadas na Seção 2.5.3: (2) *Atributos que caracterizarão cada alternativa*; (7) *Intervalo de variação de cada atributo* e (8) *Níveis de cada atributo*.

#### 4.1.3 Restrições ao desenho

Os níveis dos atributos *intervalo*, *atraso* e *tempo de viagem* foram segmentados em dois grupos: “valores menores” e “valores maiores” (conforme Tabela 4 abaixo), de modo que duas alternativas concorrentes tivessem valores para o mesmo atributo que pertencessem ao mesmo grupo (o grupo dos “valores menores” ou o grupo dos “valores maiores”). Por exemplo: se a alternativa 1 tivesse intervalo de 5 minutos, atraso de 14 minutos e tempo de viagem de 30 minutos, a alternativa 2 deveria ter intervalo de 5 ou 10 minutos (mas nunca de 15 ou 20 minutos), atraso de 14 ou 19 minutos (mas nunca de 4 ou 9 minutos) e tempo de viagem de 20 ou 30 minutos (mas nunca de 40 ou 50 minutos).

Atributo	Valores menores		Valores maiores	
Intervalo	5	10	15	20
Atraso	4	9	14	19
Tempo de viagem	20	30	40	50

Tabela 4 – Estratégia para combinar atributos *intervalo*, *atraso* e *tempo de viagem*

Esta estratégia foi adotada com a finalidade de evitar que a combinação dos atributos *intervalo*, *atraso* e *tempo de viagem* tornasse uma alternativa muito mais atrativa do que a outra, ocasionando os mesmos riscos (descritos na seção precedente) da utilização de menor quantidade de níveis para estes três atributos.

O conteúdo desta seção corresponde à seguinte definição listada na Seção 2.5.3: (13) *Restrições ao desenho*.

#### 4.1.4 Especificação econométrica

O modelo MNL foi adotado, e o componente determinístico da utilidade ( $V_{in}$ ) para o indivíduo  $n$  e a alternativa  $i$ , foi especificado apenas com os efeitos principais, conforme mostra a equação abaixo:

$$\begin{aligned}
 V_{in} = & \beta_{display} \cdot display_{in} + \beta_{interv} \cdot interv_{in} + \\
 & \beta_{atraso} \cdot atraso_{in} + \beta_{tviagem} \cdot tviagem_{in} + \\
 & \beta_{ocup0} \cdot ocup0_{in} + \beta_{ocup1} \cdot ocup1_{in} + \\
 & \beta_{custo} \cdot custo_{in}
 \end{aligned} \tag{18}$$

O conteúdo desta seção corresponde às seguintes definições listadas na Seção 2.5.3: (3) *Tipo de modelo econométrico a ser estimado* e (4) *Especificação do componente determinístico da função utilidade para todas as alternativas*.

#### 4.1.5 Aproximações para parâmetros

Foram adotadas distribuições uniformes para as aproximações dos coeficientes, deduzidas a partir do intervalo de variação do coeficiente do tempo de viagem ( $\beta_{tviagem}$ ) e das taxas marginais de substituição entre variáveis, ambos obtidos a partir dos modelos de escolha discreta estimados com os resultados do segundo experimento piloto (conforme a Tabela 21 do Apêndice B).

A partir da distribuição de  $\beta_{tviagem} \sim U(-0,125, -0,081)$ , as distribuições para os demais coeficientes foram derivadas. Por exemplo: o valor médio para a razão entre os coeficientes do custo e do tempo de viagem é de 20,45 (conforme a Tabela 21 do Apêndice B). Assim, a distribuição para  $\beta_{tviagem}$  é multiplicada por 20,45 para que se obtenha a distribuição para  $\beta_{custo}$ :  $U(-2,556, -1,656)$ .

As aproximações das distribuições uniformes obtidas para os coeficientes que fazem parte da eq. (18) estão sumarizadas na Tabela 5 a seguir:



<b>Coefficiente</b>	<b>Aproximação para distribuição uniforme</b>
$\beta_{display}$	(0,350 , 0,229)
$\beta_{interv}$	(-0,177 , -0,115)
$\beta_{atraso}$	(-0,160 , -0,104)
$\beta_{tviagem}$	(-0,125 , -0,081)
$\beta_{ocup0}$	(1,783 , 2,720)
$\beta_{ocup1}$	(1,135 , 1,731)
$\beta_{custo}$	(-2,544 , -1,668)

**Tabela 5 – Aproximações para coeficientes**

O conteúdo desta seção corresponde à seguinte definição listada na Seção 2.5.3: (11) *Aproximações para os parâmetros do componente determinístico da função utilidade.*

#### **4.1.6 Situações de escolha**

Com base em outros estudos da literatura da área de Transportes (como o de Navarrete, 2010) e na experiência do segundo experimento piloto (em que cada entrevistado respondeu a 6 situações de escolha), optou-se pela criação de 24 diferentes situações de escolha agrupadas em 3 blocos, de modo que cada entrevistado desse 8 respostas.

Acredita-se que esta quantidade não tenha causado fadiga, dado que a forma de comunicação das situações de escolha aos entrevistados era de fácil compreensão e o número de atributos para análise foi limitado a seis. Além disso, 80% dos entrevistados levaram no máximo 12 minutos para o preenchimento do questionário.

Cabe comentar que o número de situações de escolha geradas respeitou as recomendações quanto ao balanceamento dos níveis dos atributos (visto que há atributos com 2, 3 e 4 níveis, sendo seu mínimo múltiplo comum igual a 12) e quanto à quantidade de coeficientes a serem estimados nos modelos de escolha discreta.

Conforme mencionado anteriormente, a medida de ineficiência escolhida para criar o desenho foi o  $D_b\_error$ .

O conteúdo desta seção corresponde às seguintes definições listadas na Seção 2.5.3: (6) *Balancear níveis dos atributos no desenho: sim ou não*; (9) *Quantidade S de situações de escolha a serem criadas*; (12) *Indicador de ineficiência a ser usado pelos algoritmos*; (14) *Quantidade de blocos*.

#### **4.1.7 Desenho eficiente criado**

O desenho criado conforme as definições precedentes é apresentado na Tabela 22 do Apêndice C. A título de informação, o valor do  $D_b\_error$  do desenho foi igual a 0,1061, e para o indicador  $B\_estimate$  o valor foi de 81,71 (encontrando-se dentro do intervalo entre 70% e 90%, comum para desenhos eficientes). Já seu  $S\_estimate$  foi de 10,71. Ou seja: recomenda-se que o desenho (todas as 24 situações de escolha) seja respondido pelo menos 11 vezes, o que equivale a 33 entrevistados respondendo 1 bloco com 8 situações de escolha cada um. Conforme será descrito no próximo capítulo, a quantidade de respondentes obtida no experimento (1179) foi bastante superior ao limite mínimo indicado pelo  $S\_estimate$ .

## **4.2 Informações adicionais coletadas**

A seguir são apresentadas outras variáveis coletadas no experimento: variáveis socioeconômicas dos entrevistados e quanto aos hábitos de uso do ônibus. As opções de resposta (ou seja, os valores que estas variáveis podem assumir) também foram listadas:

- Local de residência:
  - RMSP;
  - Fora da RMSP (entrevistado é solicitado a indicar estado e município);
- Frequência de uso de ônibus urbano ou metropolitano:
  - Nunca;
  - 1 a 2 vezes por mês;
  - 1 a 2 vezes por semana;

- 3 ou mais vezes por semana;
- Presença de painéis eletrônicos nos pontos que costuma usar:
  - Sim;
  - Não;
- Sexo:
  - Feminino;
  - Masculino;
- Faixa de idade:
  - Até 17 anos;
  - Entre 18 e 25 anos;
  - Entre 26 e 30 anos;
  - Entre 31 e 35 anos;
  - Entre 36 e 40 anos;
  - Entre 41 e 45 anos;
  - Entre 46 e 50 anos;
  - Entre 51 e 55 anos;
  - Entre 56 e 60 anos;
  - Entre 61 e 65 anos;
  - Entre 66 e 70 anos;
  - 71 anos ou mais;
- Ocupação:
  - Estudo;
  - Estudo e trabalho / estágio;
  - Trabalho / estágio;
  - Outra;

*Observação: após a coleta dos dados, concluiu-se que a categoria “Trabalho / estágio” é inadequada, visto que só pode fazer estágio quem é estudante. Assim, os entrevistados que são estudantes fazendo estágio podem ter selecionado tanto a ocupação “Estudo e trabalho / estágio” quanto a ocupação “Trabalho / estágio”, de modo que não se sabe quantos entrevistados, dentro da categoria “Trabalho / estágio”, são estudantes.*

- Grau de escolaridade:
  - Ensino Fundamental (mesmo que incompleto);
  - Ensino Médio (mesmo que incompleto);

- Superior incompleto;
- Superior completo;
- Pós-graduação (mesmo que incompleto);
- Quantidade de automóveis, camionetes ou peruas no domicílio:
  - Nenhum;
  - 1;
  - 2;
  - 3;
  - 4 ou mais;
- Quantidade de motocicletas no domicílio:
  - Nenhum;
  - 1;
  - 2;
  - 3;
  - 4 ou mais;
- Renda bruta mensal aproximada do domicílio:
  - Até R\$ 1.400,00;
  - Mais de R\$ 1.400,00 e até R\$ 2.800,00;
  - Mais de R\$ 2.800,00 e até R\$ 4.900,00;
  - Mais de R\$ 4.900,00 e até R\$ 7.000,00;
  - Mais de R\$ 7.000,00 e até R\$ 10.500,00;
  - Mais de R\$ 10.500,00 e até R\$ 14.000,00;
  - Mais de R\$ 14.000,00.

Cabe comentar que os limites inferior e superior de cada faixa de renda são múltiplos de R\$700,00, que é o valor aproximado do salário mínimo vigente no Estado de São Paulo na época da elaboração do experimento.

### **4.3 Amostra e operacionalização**

O público alvo desta pesquisa foi determinado por conveniência. Pela facilidade de envio do questionário de pesquisa, foi adotada a comunidade da USP, dos *campi* localizados no município de São Paulo, abrangendo alunos (predominantemente), professores e funcionários.

A amostra, portanto, foi determinada pela quantidade de respostas obtidas a partir do público alvo. Ressalta-se que nem o público alvo, nem a amostra, são representativos da população de usuários de ônibus da RMSP. No entanto, a interpretação das respostas do experimento à luz das características socioeconômicas dos entrevistados torna os resultados obtidos a partir da amostra mais próximos do comportamento dos passageiros da RMSP.

O questionário da pesquisa foi construído com a ferramenta Qualtrics ([www.qualtrics.com](http://www.qualtrics.com)) e hospedado numa página da internet. Uma cópia da tela inicial do questionário é exibida no Apêndice D.

O envio do questionário à comunidade da USP ficou a cargo do Serviço de Comunicação Social da Escola Politécnica, e ocorreu por meio de correio eletrônico, além de ter sido divulgado no *site* da Escola Politécnica e em seu perfil numa rede social virtual.

Cabe salientar que, no questionário, as situações de escolha foram apresentadas aos entrevistados individualmente, de modo que uma nova situação só aparecesse na tela quando o participante tivesse feito a escolha para a situação anterior, não sendo permitido voltar à situação anterior para alterá-la ou consultá-la.

As oito situações de escolha apresentadas a cada entrevistado, o foram em sequência aleatória. A localização das alternativas 1 e 2 na tabela com suas características também mudava: ora a alternativa 1 estava à direita (e a 2 à esquerda), ora o contrário. A ordem de apresentação dos atributos nesta tabela também se alterava: ora o atributo *painel* era o primeiro da lista, ora o terceiro.

## 5 RESULTADOS E DISCUSSÃO

O experimento obteve 1179 respostas completas de indivíduos residentes na RMSP, a partir das quais foram realizadas análises que se dividem em cinco categorias (conforme a Figura 2 da Seção 3.2):

- (1) caracterização da amostra;
- (2) estimação do primeiro modelo para o comportamento da amostra (chamado de modelo base);
- (3) estudo da não linearidade das preferências;
- (4) estudo dos efeitos das interações entre atributos sobre as preferências; e
- (5) estudo da variabilidade sistemática das preferências entre os entrevistados;

As cinco categorias de análise são descritas a seguir e correspondem, na mesma sequência apresentada no parágrafo precedente, às cinco seções deste capítulo.

### 5.1 Caracterização da amostra

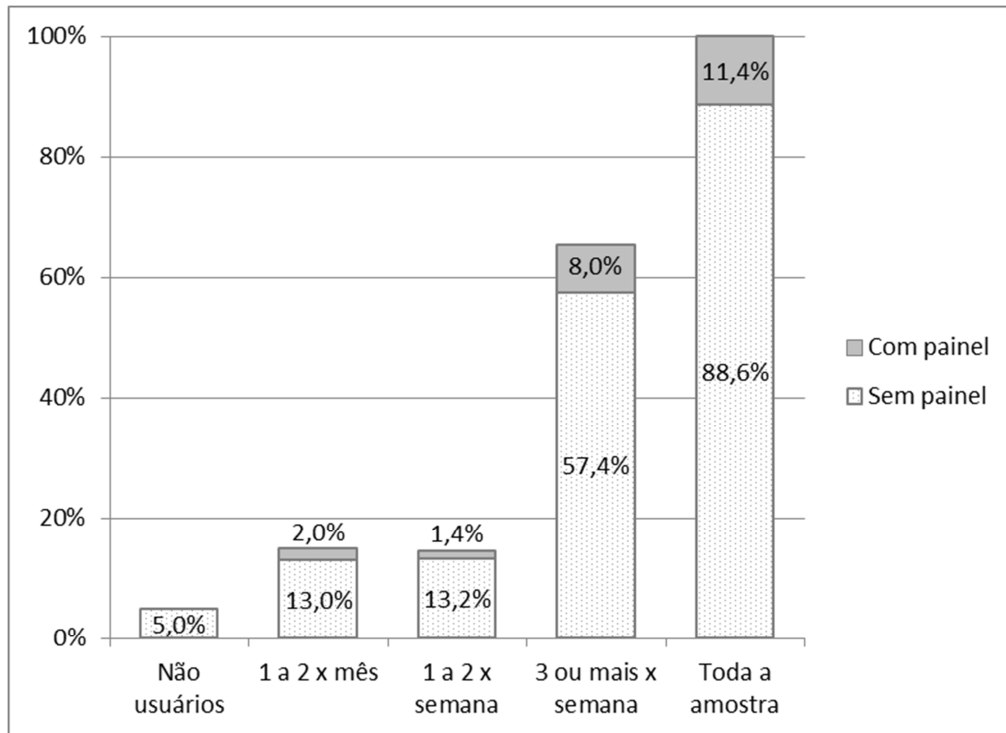
A seguir são descritas as características da amostra sob a perspectiva socioeconômica e relativa a hábitos de uso de ônibus e também sob a perspectiva do comportamento não compensatório.

#### 5.1.1 *Caracterização socioeconômica e relativa ao uso de ônibus*

Na amostra coletada predominam usuários frequentes de ônibus, conforme se verifica na Figura 4: 80% usam semanalmente este modo de transporte, sendo que 65,4% do total de entrevistados usam 3 ou mais vezes por semana. Apenas 5% da amostra não é usuária de ônibus e os restantes 15% usam esporadicamente.

No entanto, mesmo com a alta proporção de usuários de ônibus na amostra, apenas 11,4% do total de entrevistados declararam que há painéis eletrônicos nos pontos de ônibus

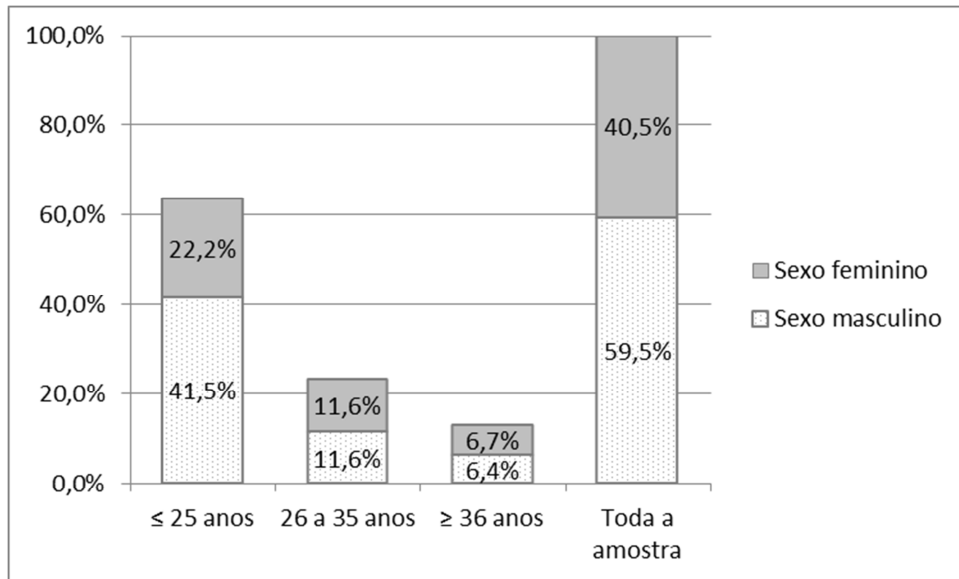
que costumam utilizar (o que talvez seja um reflexo da quantidade de painéis na RMSP, proporcionalmente à quantidade de pontos), conforme mostra o mesmo gráfico.



**Figura 4 – Segmentação por frequência de uso de ônibus e uso de pontos com painel (% sobre o total da amostra)**

No que se refere a sexo e idade, 59,5% dos respondentes são do sexo masculino e 63,7% tem até 25 anos (conforme a Figura 5), o que é reflexo das características do público para o qual o questionário foi enviado: a comunidade dos *campi* da USP situados na cidade de São Paulo, com predominância de alunos de graduação (sabe-se que houve intensa divulgação à comunidade da Escola Politécnica<sup>1</sup>). Destaca-se que o desequilíbrio entre o sexo masculino e o feminino ocorre apenas para a primeira faixa de idade (até 25 anos): aproximadamente 35% dos entrevistados desta faixa são do sexo feminino e 65% do masculino. Para as duas outras faixas de idade (entre 26 e 35 anos e a partir de 36 anos), a quantidade de homens e mulheres na amostra é igual ou muito semelhante.

<sup>1</sup> O envio do questionário à comunidade da USP ficou a cargo do Serviço de Comunicação Social da Escola Politécnica, e não se teve acesso à lista completa das unidades da universidade para cujos alunos, professores e funcionários o questionário foi enviado.



**Figura 5 – Segmentação por idade e sexo (% sobre o total da amostra)**

Como se verifica na Tabela 6, e novamente em conformidade com as características do público para o qual o questionário foi enviado, 81,1% dos entrevistados se declararam estudantes (e pode-se inferir que cursam essencialmente a graduação ou a pós-graduação), sendo que aproximadamente a metade (42,5% do total da amostra) não exerce atividade remunerada simultaneamente, ao contrário dos restantes (38,6% do total da amostra), que também trabalham ou fazem estágio. Quanto aos 14,4% que declarou o trabalho ou o estágio como sua atividade, cabe lembrar que os entrevistados que são estudantes fazendo estágio podem ter selecionado tanto a ocupação “Estudo e trabalho / estágio” quanto a ocupação “Trabalho / estágio”, de modo que não se sabe quantos entrevistados, dentro da categoria “Trabalho / estágio”, são estudantes (conforme mencionado na Seção 4.2).

O nível de escolaridade, conforme esperado, é alto: 54,7% já iniciaram o curso superior (porém ainda não o completaram), 11,1% já terminaram o curso superior (porém não iniciaram uma pós-graduação), 30% já iniciaram uma pós-graduação (quer a tenham completado ou não) e os 4,2% restantes têm o Ensino Fundamental ou Médio iniciado (quer os tenham concluído ou não).

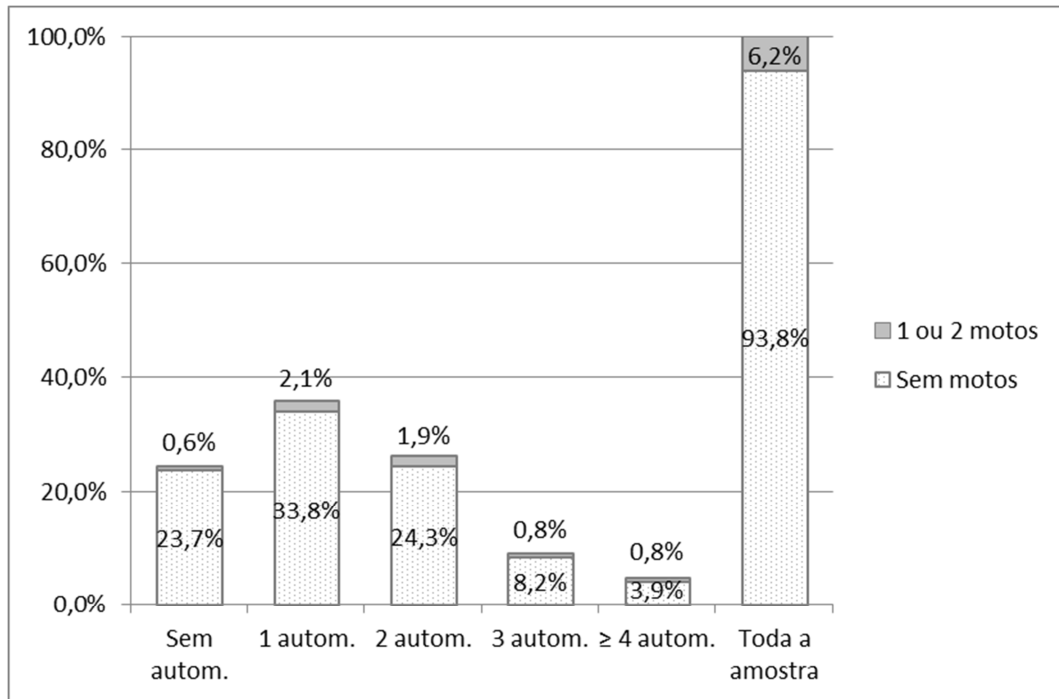


<b>Escolaridade / Ocupação</b>	<b>Fund. / Médio *</b>	<b>Superior incompleto</b>	<b>Superior completo</b>	<b>Pós- graduação *</b>	<b>Total</b>
<i>Estudo</i>	3,1%	31,6%	0,6%	7,2%	42,5%
<i>Estudo e trabalho / estágio</i>	0,3%	21,8%	3,7%	12,7%	38,6%
<i>Trabalho / estágio</i>	0,4%	1,3%	5,4%	7,3%	14,4%
<i>Outra</i>	0,3%	0,1%	1,4%	2,8%	4,5%
<b>Total</b>	<b>4,2%</b>	<b>54,7%</b>	<b>11,1%</b>	<b>30,0%</b>	<b>100%</b>

**Tabela 6 – Segmentação por ocupação e escolaridade (% sobre o total da amostra, \* mesmo que incompleto)**

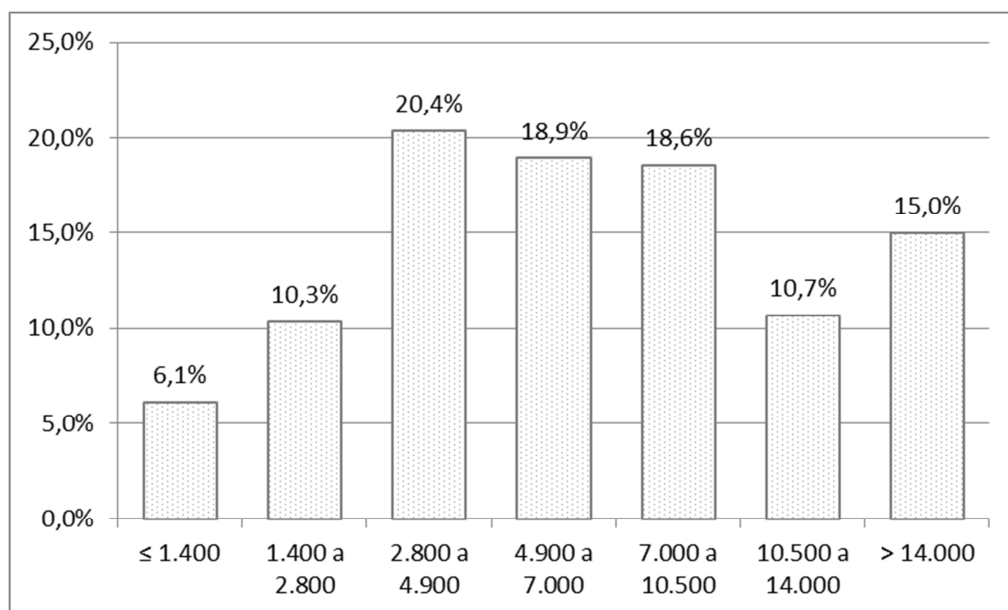
Quanto à posse de automóveis no domicílio (aqui se incluem também as camionetes e peruas), conforme se verifica na Figura 6, a amostra se concentra em três categorias: indivíduos que vivem em domicílios sem automóveis (24,3% dos entrevistados), com 1 automóvel (35,9%) e com 2 automóveis (26,2%). Os 13,7% restantes correspondem a entrevistados em cujos domicílios há 3 ou mais automóveis.

Ao analisar a posse de motocicletas, verifica-se que pequena parcela da amostra possui motocicleta no domicílio (apenas 6,2%), sendo que os entrevistados com 4 ou mais automóveis no domicílio são aqueles que, proporcionalmente, possuem mais motos (aproximadamente 17% deles possuem motos).



**Figura 6 – Segmentação por quantidade de automóveis e motocicletas no domicílio (% sobre o total da amostra)**

A amostra é distribuída por todas as faixas de renda domiciliar, conforme pode ser verificado na Figura 7, com predominância das três faixas centrais: de R\$2.800,00 a R\$4.900,00 (20,4% dos entrevistados), de R\$4.900,00 a R\$7.000,00 (18,9%) e de R\$7.000,00 a R\$10.500,00 (18,6%).



**Figura 7 – Segmentação por renda do domicílio em R\$ (% sobre o total da amostra)**

### ***5.1.2 Caracterização quanto ao comportamento não compensatório***

A análise das respostas dadas por cada entrevistado permitiu identificar aqueles que apresentaram comportamento não compensatório. Foram encontrados quatro tipos de comportamentos não compensatórios:

1. quando, para todas as situações de escolha, o entrevistado selecionou a alternativa que estava na mesma posição (à direita ou à esquerda da tabela que mostrava os atributos das alternativas);
2. quando, para todas as situações de escolha, o entrevistado selecionou a alternativa de menor tarifa, caracterizando o comportamento lexicográfico quanto ao custo. Destaca-se que, em nenhuma situação de escolha do experimento, as alternativas tinham tarifas iguais;
3. quando, para todas as situações de escolha, o entrevistado selecionou a alternativa com painel eletrônico, caracterizando o comportamento lexicográfico quanto ao painel. Destaca-se que, em todas as situações de escolha do experimento, apenas uma das alternativas tinha painel;
4. quando, para todas as situações de escolha, o entrevistado selecionou a alternativa com maior nível de conforto, caracterizando o comportamento lexicográfico quanto ao conforto. Destaca-se que, em apenas 1 das 24 situações de escolha do experimento, os níveis de conforto das alternativas eram iguais.

Pela Tabela 7 verifica-se que quase 25% dos entrevistados apresentaram algum tipo de comportamento não compensatório, sendo que os mais frequentes foram quanto ao custo (5,8% da amostra) e principalmente quanto ao conforto (17,3% da amostra).

<b>Comportamento</b>	<b>Quant.</b>	<b>%</b>
Não compensatório	290	24,6%
1. Sempre escolhe alternativa da direita/esquerda	4	0,3%
2. Lexicográfico quanto ao custo	68	5,8%
3. Lexicográfico quanto ao painel	14	1,2%
4. Lexicográfico quanto ao conforto	204	17,3%
Compensatório	889	75,4%
<b>Total</b>	<b>1179</b>	

**Tabela 7 – Segmentação quanto ao comportamento compensatório**

Ressalta-se que o comportamento supostamente não compensatório apresentado no experimento não significa que o entrevistado:

- (a) tenha de fato usado regras não compensatórias para suas escolhas. Ou seja: pode ter usado regras compensatórias cujo resultado foi idêntico, para o conjunto de situações de escolha que lhe foi apresentado, ao resultado das escolhas feitas por um indivíduo que usasse regras realmente não compensatórias. Por este motivo, os 290 indivíduos que apresentaram comportamento não compensatório serão mantidos na amostra e incluídos nas análises das seções seguintes; ou,
- (b) sempre use regras não compensatórias para suas decisões. Ou seja: pode tê-las usado para as situações de escolha que lhe foram apresentadas, mas não necessariamente o faria para outras combinações de níveis dos atributos, ou mesmo para níveis dos atributos fora dos intervalos de variação adotados no experimento.

O efeito do comportamento não compensatório sobre os modelos de escolha discreta estimados é bastante coerente. Em comparação a um modelo ML *panel* estimado apenas com os efeitos principais e para a amostra completa (denominado modelo base e explorado na próxima seção), quando dela são retirados os 204 lexicográficos quanto ao conforto e o modelo é reestimado, os coeficientes do atributo conforto diminuem em valor absoluto, enquanto os coeficientes dos demais atributos sofrem aumento em seus valores absolutos. Já quando são retirados da amostra os 68 lexicográficos quanto ao custo (ao invés dos lexicográficos quanto ao conforto), é o coeficiente do custo que cai em valor absoluto, ao

mesmo tempo em que aumentam, também em valor absoluto, os demais. Ao se retirar da amostra os 14 lexicográficos quanto ao painel (ao invés dos outros grupos de lexicográficos), as alterações no modelo também são coerentes embora muito discretas, dado que este grupo é bastante pequeno. Para o último grupo com comportamento não compensatório (os que sempre escolheram a alternativa na mesma posição), não foi estimado nenhum modelo, tal como feito para os demais grupos, por se tratar de apenas 4 entrevistados. A apresentação e a discussão destes modelos está no Apêndice E. No entanto, recomenda-se sua leitura após a Seção 5.2 a seguir.

## **5.2 Modelo base para o comportamento da demanda**

Nesta seção é apresentado o modelo base, o mais simples dentre os estimados neste estudo, por ser especificado apenas com os efeitos principais dos seis atributos e sem levar em consideração que há variabilidade de preferências entre os entrevistados. Trata-se de um modelo ML *panel* estimado a partir das 8 respostas de cada um dos 1179 entrevistados, totalizando 9432 observações. Os resultados são mostrados na Tabela 8, à qual foram adicionadas duas colunas à direita: a primeira indicando a disposição a pagar (em R\$) pelo acréscimo ou decréscimo (sempre aquele que aumentar a utilidade) de uma unidade de cada atributo, e a segunda indicando a taxa marginal de substituição de uma unidade de cada atributo por minutos de viagem.

Antes de iniciar a interpretação do modelo, cabe fazer uma observação, válida para todos os modelos apresentados neste capítulo e também nos apêndices. Quando a utilidade marginal de um atributo tem sinal positivo, quanto maior seu valor absoluto, maior é sua contribuição para a satisfação do indivíduo e para o aumento da utilidade. Já quando a utilidade marginal tem sinal negativo, quanto maior seu valor absoluto, maior é sua contribuição para a insatisfação do indivíduo e para o aumento da desutilidade.

<b>Parâmetro</b>	<b>Valor</b>	<b>Erro padrão</b>	<b>Teste t</b>	<b>Disposição a pagar</b>	<b>Taxa substit.</b>
$\beta_{display}$	0,328	0,026	12,870	-0,119	-5,0
$\beta_{interv}$	-0,165	0,008	-21,950	0,060	2,5
$\beta_{atraso}$	-0,068	0,007	-9,250	0,025	1,0
$\beta_{tviagem}$	-0,066	0,004	-15,570	0,024	--
$\beta_{ocup0}$	2,750	0,072	38,240	-1,000	-41,7
$\beta_{ocup1}$	1,720	0,050	34,340	-0,625	-26,1
$\beta_{custo}$	-2,750	0,106	-25,830	--	41,7
$\sigma_{panel}$	-0,302	0,057	-5,320	--	--

Número de observações: 9432

Número de indivíduos: 1179

$\mathcal{L}(0) = -6537,76$

$\mathcal{L}(\beta) = -5219,88$

$-2 \cdot [\mathcal{L}(0) - \mathcal{L}(\beta)] = 2635,77$

$\rho^2 = 0,202$

$\rho^2$  ajustado = 0,200

**Tabela 8 – Modelo base**

Verifica-se que todos os parâmetros têm os sinais esperados (os coeficientes para custo e tempo têm sinal negativo, enquanto aqueles que indicam mais conforto e presença de painel no ponto têm sinal positivo) e são diferentes de zero ao nível de significância de 5%, o que, no caso do parâmetro  $\sigma_{panel}$ , indica que o modelo está capturando as correlações não observadas entre as respostas dadas por um mesmo entrevistado.

Quanto ao teste da razão da máxima verossimilhança e ao rho quadrado ajustado, o resultado do primeiro rejeita a hipótese de que os parâmetros sejam iguais a zero, e o segundo (que tem valor igual a 0,200) representa o poder explicativo do modelo e será um importante elemento para comparação com os modelos apresentados nas seções seguintes.

A análise dos coeficientes do custo e do tempo de viagem sugere que o valor do tempo é baixo para a amostra entrevistada, visto que o indivíduo estaria disposto a pagar apenas 2,4 centavos para reduzir 1 minuto de viagem (ou R\$1,44 / hora). Esta taxa é inferior à obtida no segundo experimento piloto, que foi de 4,9 centavos por minuto de viagem (ou R\$2,94 / hora), conforme se verifica na Tabela 21 do Apêndice B.

A comparação entre os coeficientes do intervalo programado entre veículos e do atraso mostra que 1 minuto de intervalo exerce 2,43 vezes mais influência na decisão do que 1 minuto de atraso. Esta diferença entre os valores dos coeficientes também se reflete na taxa marginal de substituição de 1 minuto de intervalo e de 1 minuto de atraso por minutos de viagem: 2,5 e 1,0 minutos de viagem, respectivamente, sendo que a primeira das taxas está de acordo com as relações costumeiramente encontradas entre tempo de espera e de viagem nos estudos da área de Transportes, a exemplo do que concluiu Wardman (2004). Pelo fato de o atraso ser uma espera com duração desconhecida e não justificada, conforme a proposição de Maister (1985) reproduzida na Seção 2.2.2, se esperava que seu coeficiente fosse superior (ou, no mínimo, semelhante) ao coeficiente do intervalo programado entre veículos. Possíveis razões para a menor importância do atraso (e o baixo valor relativo de seu coeficiente) nos resultados do experimento são:

- (a) situações de escolha com atrasos grandes (de 14 ou 19 minutos, por exemplo) e intervalos pequenos (como 5 minutos), que podem ter sido interpretadas como pouco realistas pelos entrevistados;
- (b) presença de entrevistados, na amostra, que não usam ônibus ou que usam esporadicamente. Eles podem não saber avaliar o atraso e suas implicações, conforme será mostrado na Seção 5.5 e no Apêndice H;
- (c) forma adotada para comunicar o atraso no experimento, que pode não ter enfatizado adequadamente o atributo. O atraso foi comunicado indiretamente através do tempo máximo de espera (que é soma do intervalo com o atraso), e como uma possibilidade que poderia ou não ocorrer e teria duração variável (sem a indicação, contudo, de probabilidades de ocorrência);
- (d) ausência, no experimento, do motivo específico da viagem (ex.: trabalho, educação etc.) e de um limite de tempo para chegar ao destino, correspondente ao horário de um compromisso e representando pressão de tempo.

No que se refere ao painel, sua presença no ponto de ônibus gera a mesma utilidade que a redução de 5 minutos de viagem (ou 2 minutos de intervalo, ou 4,8 minutos de atraso), e o entrevistado estaria disposto a pagar cerca de 12 centavos a mais na tarifa se o ponto estivesse equipado com painel.

Quanto à ocupação do veículo, verifica-se que a utilidade marginal de passar do nível máximo para o intermediário (igual a 1,720) é superior à utilidade marginal de passar do nível intermediário para o mínimo (igual a 1,030). Em outras palavras, descobrir que o ônibus não está lotado (embora se tenha que viajar em pé), gera mais satisfação do que, já sabendo que o ônibus não está lotado, descobrir que será possível viajar sentado. Isto também se reflete na disposição a pagar para viajar com mais conforto, que é de R\$0,63 quando se passa do nível máximo da ocupação para o intermediário, e de R\$0,37 quando se passa do intermediário para o mínimo. Semelhante situação ocorre com as taxas marginais de substituição entre a ocupação do veículo e o tempo de viagem, muito superiores às encontradas nos resultados do segundo experimento piloto, conforme a Tabela 21 do Apêndice B, e aos resultados de Kim; Lee e Oh (2009).

A comparação do coeficiente da presença do painel com os dois coeficientes de ocupação do veículo mostra que a utilidade marginal gerada pelo nível mínimo da ocupação é 8,4 vezes maior do que a utilidade marginal gerada quando há painel no ponto. Já a razão entre as utilidades marginais do nível intermediário da ocupação e da presença do painel é 5,2. Uma possível explicação para a maior importância da ocupação do veículo na decisão dos entrevistados é que, quando se trata do tempo de espera, o painel é o mecanismo que revelará a previsão do tempo de espera pelo ônibus (sendo que a previsão em si não é fornecida ao entrevistado no experimento), enquanto que, em se tratando do conforto durante a viagem, a própria previsão do nível de ocupação a ser encontrado é comunicada ao entrevistado.

### **5.3 Não linearidade das preferências**

O objetivo da especificação desenvolvida nesta seção é explorar a possível existência de não linearidade para os coeficientes do intervalo programado entre veículos, do atraso e do tempo de viagem. Assume-se que os primeiros minutos de espera, quer seja devida ao intervalo programado quer ao atraso, ou de viagem influenciam a decisão do usuário com intensidade diferente dos últimos minutos de espera ou viagem. Para tanto, foi estimado um modelo ML *panel piecewise*, cuja especificação é brevemente descrita a seguir.



O que diferencia o modelo ML *panel piecewise* e o modelo base apresentado anteriormente, é que os atributos intervalo programado, atraso e tempo de viagem foram separados em duas partes, conforme as equações abaixo:

$$interv = interv\_a + interv\_b, \text{ dado que:}$$

$$interv\_a = \min(interv, 10)$$

$$interv\_b = \max(0, interv - 10)$$

$$atraso = atraso\_a + atraso\_b, \text{ dado que:}$$

$$atraso\_a = \min(atraso, 9)$$

$$atraso\_b = \max(0, atraso - 9)$$

$$tviagem = tviagem\_a + tviagem\_b, \text{ dado que:}$$

$$tviagem\_a = \min(tviagem, 30)$$

$$tviagem\_b = \max(0, tviagem - 30)$$

De modo que a especificação do componente determinístico da utilidade para o indivíduo  $n$  e a alternativa  $i$  se torna:

$$\begin{aligned} V_{in} = & \beta_{display} \cdot display_{in} + \\ & \beta_{interv\_a} \cdot interv\_a_{in} + \beta_{interv\_b} \cdot interv\_b_{in} + \\ & \beta_{atraso\_a} \cdot atraso\_a_{in} + \beta_{atraso\_b} \cdot atraso\_b_{in} + \\ & \beta_{tviagem\_a} \cdot tviagem\_a_{in} + \beta_{tviagem\_b} \cdot tviagem\_b_{in} + \\ & \beta_{ocup0} \cdot ocup0_{in} + \beta_{ocup1} \cdot ocup1_{in} + \\ & \beta_{custo} \cdot custo_{in} \end{aligned} \quad (18)$$

sendo que, para apenas uma das duas alternativas é acrescentado à  $V_{in}$  o termo  $\sigma_{panel}$ , que captura as correlações não observadas entre as repostas do mesmo entrevistado.

Devido à separação dos atributos em duas partes, quando, por exemplo, o atraso associado à determinada alternativa for de 14 minutos, a variável  $atraso\_a$  será igual a 9, e a variável  $atraso\_b$  igual a 5 (sendo 14 a soma de ambas). Cada uma das novas variáveis é multiplicada por um coeficiente diferente, o que significa que cada um dos primeiros 9 minutos de atraso terá uma contribuição diferente para a utilidade do que cada um dos 5 últimos minutos de atraso.

Os resultados do modelo ML *panel piecewise* são apresentados na Tabela 9, à qual foi acrescentada uma coluna à direita com os valores dos parâmetros no modelo base. Observa-se que todos os parâmetros do modelo em discussão têm os sinais esperados e são diferentes de zero ao nível de significância de 5%.

Verifica-se que o coeficiente do intervalo programado entre veículos é maior para os primeiros 10 minutos de espera ( $\beta_{\text{interv\_a}} = -0,189$ ), decrescendo 18,5% em valor absoluto a partir do décimo primeiro minuto ( $\beta_{\text{interv\_b}} = -0,154$ ). Ou seja: quanto mais próximo o momento da chegada do ônibus ao ponto, menor a desutilidade marginal da espera associada ao intervalo programado.

Já o coeficiente do atraso apresenta comportamento oposto ao do intervalo programado entre veículos: a utilidade marginal dos 9 primeiros minutos de atraso ( $\beta_{\text{atraso\_a}} = -0,047$ ) aumenta quase 2,5 vezes para os minutos seguintes ( $\beta_{\text{atraso\_b}} = -0,117$ ). Uma possível explicação para a diferença entre os coeficientes do intervalo e do atraso é que, ao contrário das esperas finitas advindas do intervalo, as esperas advindas do atraso não são programadas (de modo que não se sabe quando o ônibus chegará ao ponto), tornando a desutilidade marginal do atraso maior na medida em que ele cresce.

O coeficiente do tempo de viagem tem comportamento semelhante ao do intervalo programado entre veículos, visto que é maior para os primeiros 30 minutos dentro do veículo ( $\beta_{\text{tviagem\_a}} = -0,097$ ), e sofre queda de 62,9% em valor absoluto para os seguintes ( $\beta_{\text{tviagem\_b}} = -0,036$ ). Aplicando raciocínio semelhante ao usado para interpretar o coeficiente do intervalo, pode-se deduzir que, quanto mais próximo o momento do desembarque, menor a desutilidade marginal do tempo de viagem.

Parâmetro	Valor	Erro padrão	Teste t	Modelo Base
$\beta_{display}$	0,302	0,028	10,770	0,328
$\beta_{interv\_a}$	-0,189	0,011	-17,180	-0,165
$\beta_{interv\_b}$	-0,154	0,011	-14,190	
$\beta_{atraso\_a}$	-0,047	0,011	-4,450	-0,068
$\beta_{atraso\_b}$	-0,117	0,013	-8,980	
$\beta_{tviagem\_a}$	-0,097	0,006	-15,790	-0,066
$\beta_{tviagem\_b}$	-0,036	0,006	-6,200	
$\beta_{ocup0}$	2,890	0,079	36,800	2,750
$\beta_{ocup1}$	1,760	0,052	33,860	1,720
$\beta_{custo}$	-3,130	0,126	-24,900	-2,750
$\sigma_{panel}$	-0,308	0,056	-5,470	-0,302
Número de observações: 9432				9432
Número de indivíduos: 1179				1179
$\log \mathcal{L}(0) = -6537,76$				-6537,76
$\log \mathcal{L}(\beta) = -5191,33$				-5219,88
$-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] = 2692,86$				2635,77
$\rho^2 = 0,206$				0,202
$\rho^2$ ajustado = 0,204				0,200

Tabela 9 – Modelo ML *panel piecewise*

As taxas marginais de substituição entre os dois coeficientes do intervalo ( $\beta_{interv\_a}$  e  $\beta_{interv\_b}$ ) e os dois coeficientes do tempo de viagem ( $\beta_{tviagem\_a}$  e  $\beta_{tviagem\_b}$ ) são apresentados na Tabela 10 (abaixo), bem como as taxas marginais de substituição entre os dois coeficientes do atraso ( $\beta_{atraso\_a}$  e  $\beta_{atraso\_b}$ ) e os dois coeficientes do tempo de viagem.

	$\beta_{interv\_a}$ (até o 10º minuto)	$\beta_{interv\_b}$ (a partir do 11º minuto)	$\beta_{atraso\_a}$ (até o 9º minuto)	$\beta_{atraso\_b}$ (a partir do 10º minuto)
$\beta_{tviagem\_a}$ (até o 30º minuto)	1,95	1,59	0,48	1,21
$\beta_{tviagem\_b}$ (a partir do 31º minuto)	5,25	4,28	1,31	3,25

Tabela 10 – Taxas marginais de substituição (em minutos de viagem)

Verifica-se, pela análise da Tabela 10, que tanto a taxa marginal de substituição entre o intervalo e o tempo de viagem, quanto entre o atraso e o tempo de viagem apresentam variação. Destaca-se a primeira, que varia entre 1,59 e 5,25. Entende-se, portanto, que a redução de 1 minuto no final da espera (quando esta é decorrente de intervalo superior a 10 minutos) pode ser trocada por 1,59 minutos adicionais no início da viagem. Similarmente, a redução de 1 minuto no início da espera (quando esta é decorrente de intervalo) pode ser trocada por 5,25 minutos adicionais no final da viagem (para viagens com duração superior a 30 minutos).

A Tabela 11 (abaixo) traz as taxas marginais de substituição entre a presença do painel e o intervalo programado entre veículos, o atraso e o tempo de viagem.

	$\beta_{\text{display}}$
$\beta_{\text{interv\_a}}$ (até o 10º minuto)	-1,60
$\beta_{\text{interv\_b}}$ (a partir do 11º minuto)	-1,96
$\beta_{\text{atraso\_a}}$ (até o 9º minuto)	-6,43
$\beta_{\text{atraso\_b}}$ (a partir do 10º minuto)	-2,58
$\beta_{\text{tviagem\_a}}$ (até o 30º minuto)	-3,11
$\beta_{\text{tviagem\_b}}$ (a partir do 31º minuto)	-8,39

**Tabela 11 – Taxas marginais de substituição (em minutos de intervalo, atraso ou viagem)**

Assim como se verificou na Tabela 10, as taxas marginais de substituição da Tabela 11 também apresentam variação, destacando-se: a taxa de -6,43 entre a presença do painel e o atraso, o que significa que a presença do painel pode ser trocada por 6,43 minutos adicionais no início do atraso, e a taxa de -8,39 entre a presença do painel e o tempo de viagem, o que significa que a presença do painel pode ser trocada por 8,39 minutos adicionais no final da viagem (para viagens com duração acima de 30 minutos).

Pode-se dizer que a inclusão da não linearidade das preferências na especificação leva a um modelo que melhor explica o comportamento da amostra, visto que o rho quadrado ajustado do modelo em discussão é de 0,204, enquanto o do modelo base é de 0,200.

#### **5.4 Influência das interações entre atributos**

As especificações apresentadas nesta seção visam explorar a influência das interações entre atributos no comportamento da amostra.

Cabe lembrar, como exposto na Seção 3.2, que os modelos aqui apresentados são o resultado de um processo iterativo de especificação, estimação e análise de aproximadamente mil outros modelos. Desta forma, quando determinada interação não faz parte de um modelo apresentado, entende-se que, durante o processo iterativo, ela não se mostrou estatisticamente significativa ou, mesmo tendo se mostrado significativa, apresentava sinal contrário ao esperado, ou ainda fazia parte de um modelo com poder explicativo inferior ou incoerente com o comportamento de escolha de linha. Esta observação também é válida para o modelo discutido na próxima seção.

Três modelos ML *panel* contendo interações entre atributos são apresentados na Tabela 12, à qual foi acrescentada, à direita, uma coluna com os valores dos parâmetros no modelo base. Todos os coeficientes dos três novos modelos têm os sinais esperados e são diferentes de zero ao nível de significância de 5%. Os valores do erro padrão e do teste *t* para os parâmetros dos três modelos podem ser encontrados no Apêndice F.

Parâmetro	Modelo Int.1	Modelo Int.2	Modelo Int.3	Modelo Base
$\beta_{display}$	--	--	0,304	0,328
$\beta_{interv}$	-0,158	-0,166	-0,171	-0,165
$\beta_{interv\_display}$	--	0,022	--	--
$\beta_{atraso}$	-0,071	-0,069	--	-0,068
$\beta_{atraso\_interv\{5,10\}}$	--	--	-0,044	--
$\beta_{atraso\_interv\{15,20\}}$	--	--	-0,097	--
$\beta_{atraso\_display}$	0,017	--	--	--
$\beta_{tviagem}$	-0,081	-0,063	-0,070	-0,066
$\beta_{tviagem\_ocup0}$	0,025	0,010	0,011	--
$\beta_{tviagem\_ocup1}$	0,014	--	--	--
$\beta_{ocup0}$	1,890	2,390	2,380	2,750
$\beta_{ocup1}$	1,230	1,700	1,700	1,720
$\beta_{custo}$	-2,730	-2,760	-2,880	-2,750
$\sigma_{panel}$	-0,292	-0,305	-0,303	-0,302
Número de observações:	9432	9432	9432	9432
Número de indivíduos:	1179	1179	1179	1179
$\log \mathcal{L}(0) =$	-6537,76	-6537,76	-6537,76	-6537,76
$\log \mathcal{L}(\beta) =$	-5230,17	-5212,36	-5205,60	-5219,88
$-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] =$	2615,18	2650,82	2664,32	2635,77
$\rho^2 =$	0,200	0,203	0,204	0,202
$\rho^2$ ajustado =	0,198	0,201	0,202	0,200

**Tabela 12 – Modelos ML *panel* com interações entre atributos**

A interação entre o tempo de viagem e a ocupação do veículo é comum aos três modelos, e age de maneira a diminuir a desutilidade marginal do tempo de viagem, revelando a importância do conforto para a redução da insatisfação gerada durante o deslocamento. O efeito desta interação sobre os três modelos pode ser assim sumarizado:

- o modelo Int.1 é o único que interage tanto com o nível mínimo quanto com o nível intermediário da ocupação. É, também, o que apresenta maior intervalo de variação para a utilidade marginal do tempo de viagem: o coeficiente deste atributo é igual a -0,081 ( $\beta_{tviagem}$ ) quando o veículo está lotado, -0,067 ( $\beta_{tviagem} + \beta_{tviagem\_ocup1}$ ) quando a ocupação é intermediária, e -0,056 ( $\beta_{tviagem} + \beta_{tviagem\_ocup0}$ ) quando o veículo está quase vazio;

- no modelo Int.2, o coeficiente do tempo de viagem é igual a  $-0,063$  ( $\beta_{tviagem}$ ) quando o veículo está lotado ou com ocupação intermediária, e igual a  $-0,053$  ( $\beta_{tviagem} + \beta_{tviagem\_ocup0}$ ) quando o veículo tem ocupação mínima;
- no modelo Int.3, a utilidade marginal do tempo de viagem é de  $-0,070$  ( $\beta_{tviagem}$ ) quando o ônibus está lotado ou com ocupação intermediária, e de  $-0,059$  ( $\beta_{tviagem} + \beta_{tviagem\_ocup0}$ ) quando o veículo está quase vazio.

A seguir, aspectos específicos de cada modelo serão apresentados.

#### (a) Modelo ML *panel* com interações 1

O modelo Int.1 também apresenta interação entre o atraso e a presença de painel no ponto, substituindo o efeito principal deste último atributo. Verifica-se que a utilidade marginal do atraso cai 23,9% em valor absoluto quando há painel no ponto: passa de  $-0,071$  ( $\beta_{atraso}$ ) na situação sem painel, para  $-0,054$  ( $\beta_{atraso} + \beta_{atraso\_display}$ ) na situação com painel, o que confirma a eficácia do painel em tornar menos custosa a espera pelo ônibus devida ao atraso.

Quando o coeficiente do atraso atinge seu valor absoluto máximo (ou seja,  $-0,071$ , quando não há painel no ponto) e o coeficiente do tempo de viagem, o seu valor absoluto mínimo (ou seja,  $-0,056$ , quando o veículo tem ocupação mínima e se pode viajar sentado), tem-se a maior taxa marginal de substituição entre atraso e tempo de viagem obtida a partir do modelo em discussão: 1,27 minutos de viagem por minuto de atraso. Já quando o coeficiente do atraso alcança seu valor absoluto mínimo (ou seja,  $-0,054$ , quando há painel no ponto) e o coeficiente do tempo de viagem, o seu valor absoluto máximo (ou seja,  $-0,081$ , quando é necessário viajar em pé e o ônibus está lotado), tem-se a menor taxa marginal de substituição entre os dois atributos: 0,67 minutos de viagem por minuto de atraso.

Destaca-se que as interações incluídas e a ausência do efeito principal da presença de painel não tornam este modelo superior ao modelo base em sua capacidade de explicar o comportamento de escolha dos entrevistados, visto que o rho quadrado ajustado cai de 0,200 no modelo base para 0,198 no modelo em discussão.

### (b) Modelo ML *panel* com interações 2

Este modelo, ao invés da interação entre a presença de painel e o atraso (como ocorre em Int.1), apresenta a interação entre o primeiro atributo e o intervalo programado entre veículos. Neste modelo, o efeito principal da presença do painel também é substituído pela interação.

Observa-se que a presença do painel também diminui o custo da espera associada ao intervalo programado entre veículos: a utilidade marginal do intervalo cai 13,3% em valor absoluto quando há painel no ponto, passando de -0,166 ( $\beta_{\text{interv}}$ ) na situação sem painel, para -0,144 ( $\beta_{\text{interv}} + \beta_{\text{interv\_display}}$ ) na situação com painel.

As taxas marginais de substituição entre intervalo e tempo de viagem também variam para o modelo em discussão, tal como ocorre entre atraso e tempo de viagem no modelo Int.1. A maior taxa, de 3,13 minutos de viagem por minuto de intervalo, ocorre quando não há painel no ponto durante a espera e o ônibus está quase vazio, sendo possível viajar sentado. Já a menor taxa, de 2,29 minutos de viagem por minuto de intervalo, está associada à presença de painel no ponto e aos níveis intermediário e máximo da ocupação do ônibus.

Verifica-se que a especificação deste modelo, com a inclusão das interações e a ausência do efeito principal da presença de painel, torna seu poder explicativo, que é de 0,201, discretamente superior ao do modelo base, que é igual a 0,200.

### (c) Modelo ML *panel* com interações 3

No último dos três modelos (Int.3), além da interação entre o tempo de viagem e a ocupação do veículo, o atraso interage com duas variáveis *dummy*: a primeira indica se o intervalo programado entre veículos é igual a 5 ou 10 minutos, e a segunda indica se o intervalo é igual a 15 ou 20 minutos.

Os resultados mostram que a utilidade marginal do atraso, quando é precedido de um intervalo programado menor (ou seja, de 5 ou 10 minutos) é igual a -0,044, e aumenta 2,2 vezes em valor absoluto, tornando-se igual a -0,097 quando o atraso é precedido de um intervalo maior (ou seja, de 15 ou 20 minutos). Pode-se concluir, então, que o atraso gera



mais insatisfação quando ocorre após um período maior de espera decorrente do intervalo programado.

A taxa marginal de substituição entre atraso e tempo de viagem também varia para este modelo: de 0,63 minutos de viagem por minuto de atraso, quando o atraso é precedido de intervalo programado menor e o ônibus tem ocupação intermediária ou máxima, a 1,64 minutos de viagem por minuto de atraso, quando o intervalo que precede o atraso é maior e o ônibus está quase vazio.

Assim como no modelo Int.2, o rho quadrado ajustado do modelo em discussão também é discretamente superior ao do modelo base, além de ser o maior dentre os três modelos apresentados na Seção 5.4.

## **5.5 Variabilidade sistemática das preferências**

No modelo apresentado nesta seção, a variabilidade sistemática das preferências entre os entrevistados é explorada. Para tanto, variáveis socioeconômicas e de uso de ônibus foram incluídas na especificação do componente determinístico da utilidade, através da sua interação com os atributos das alternativas.

Cabe comentar que a variabilidade puramente aleatória das preferências também foi explorada, através da estimação de modelos ML *panel random coefficients*, sob duas especificações diferentes: a primeira assumindo que os coeficientes dos atributos têm distribuição normal de probabilidades e a segunda admitindo que os coeficientes variam de acordo com uma distribuição uniforme. No primeiro caso, apenas os coeficientes do tempo de viagem e do nível mínimo da ocupação do veículo apresentaram variação aleatória. Já para o segundo modelo, foram os coeficientes do tempo de viagem e do intervalo programado entre veículos que exibiram variação aleatória. Os resultados de ambos os modelos são apresentados e brevemente discutidos no Apêndice G, visto que suas contribuições para a compreensão do comportamento de escolha de linha são menores.

Os resultados do modelo em discussão nesta seção (que é *ML panel*) são exibidos na Tabela 13, à qual foi acrescentada uma coluna à direita com os valores dos parâmetros no modelo base.

Verifica-se expressivo aumento do rho quadrado ajustado no modelo em discussão: passa de 0,200 (valor no modelo base) para 0,217, indicando que o acréscimo das interações na especificação torna o modelo superior quanto à capacidade de explicar o comportamento de escolha dos entrevistados.

Antes de prosseguir com a análise mais detalhada do modelo, cabe apresentar as variáveis *dummy* utilizadas na especificação:

- *autos4* - assume valor 1 quando, no domicílio do entrevistado, há quatro ou mais automóveis, e 0 em caso contrário;
- *estudante* - assume valor 1 quando o entrevistado é estudante e não trabalha ou faz estágio, e 0 em caso contrário;
- *freqmensal* - assume valor 1 quando o entrevistado usa ônibus de 1 a 2 vezes por mês, e 0 em caso contrário;
- *freqsemanal* - assume valor 1 quando o entrevistado usa ônibus semanalmente, e 0 em caso contrário;
- *freqsemanal3* - assume valor 1 quando o entrevistado usa ônibus 3 ou mais vezes por semana, e 0 em caso contrário;
- *idade36* - assume valor 1 quando o entrevistado tem 36 anos ou mais, e 0 em caso contrário;
- *posgrad* - assume valor 1 quando o entrevistado já iniciou a pós-graduação (quer a tenha concluído ou não), e 0 em caso contrário;
- *renda* - assume valor 0 quando o entrevistado tem renda domiciliar até R\$1.400 ou entre R\$4.900 e R\$7.000, e 1 quando o entrevistado tem qualquer outra renda domiciliar;
- *renda14mil* - assume valor 1 quando a renda domiciliar do entrevistado é acima de R\$14.000, e 0 em caso contrário;
- *superiorcomp* - assume valor 1 quando o entrevistado já concluiu o Ensino Superior, porém não ingressou na pós-graduação, e 0 em caso contrário.

Dentre as variáveis socioeconômicas e de uso de ônibus, apenas três não fazem parte deste modelo e, portanto, parecem não afetar (ao menos diretamente) o comportamento de escolha de linha. São elas: uso de pontos de ônibus equipados com painel eletrônico, sexo e quantidade de motocicletas no domicílio.

Observa-se que, dependendo das características socioeconômicas e de uso de ônibus dos entrevistados, as utilidades marginais dos atributos e, por consequência, as taxas marginais de substituição entre atributos se alteram fortemente. Dois exemplos importantes são o valor do tempo e a disposição a pagar pelo painel. Um indivíduo que (simultaneamente) tenha 4 ou mais automóveis no domicílio, tenha iniciado a pós-graduação (quer a tenha concluído ou não), tenha 36 anos ou mais, tenha renda domiciliar acima de R\$14.000 e não use ônibus com frequência igual ou superior a 3 vezes por semana, terá valor do tempo de 28,3 centavos de real por minuto de viagem (aproximadamente R\$17,00/hora), bastante acima do valor do tempo no modelo base, que é de 2,4 centavos por minuto de viagem (R\$1,44/hora). Outro indivíduo que tenha as mesmas características, com exceção do fato de que pode ter qualquer quantidade de automóveis (ou mesmo não ter automóveis), estará disposto a pagar 76,6 centavos pela disponibilidade de painel no ponto, sendo esta taxa no modelo base bastante inferior (aproximadamente 12 centavos).

Salienta-se que o efeito das variáveis socioeconômicas e de uso de ônibus sobre os atributos pode se dar tanto diretamente (como é o caso, por exemplo, do efeito da renda sobre o custo), quanto indiretamente. Neste segundo caso, acredita-se que a variável represente determinado grupo de entrevistados e sintetize seu conjunto de características, sendo possível admitir que é este conjunto que determina a influência do atributo sobre o comportamento.

Parâmetro	Valor	Erro padrão	Teste t	Modelo Base
$\beta_{display}$	0,341	0,026	13,190	0,328
$\beta_{interv}$	-0,161	0,008	-19,940	-0,165
$\beta_{interv\_autos4}$	-0,094	0,039	-2,450	--
$\beta_{interv\_superiorcomp}$	-0,048	0,021	-2,300	--
$\beta_{atraso}$	--	--	--	-0,068
$\beta_{atraso\_freqmensal}$	-0,066	0,017	-3,850	--
$\beta_{atraso\_freqsemanal}$	-0,080	0,008	-9,860	--
$\beta_{tviagem}$	-0,066	0,004	-15,100	-0,066
$\beta_{tviagem\_autos4}$	-0,060	0,021	-2,900	--
$\beta_{ocup0}$	2,410	0,098	24,540	2,750
$\beta_{ocup0\_autos4}$	1,140	0,327	3,500	--
$\beta_{ocup0\_renda}$	0,658	0,088	7,470	--
$\beta_{ocup0\_estudante}$	-0,260	0,065	-4,030	--
$\beta_{ocup1}$	1,350	0,076	17,780	1,720
$\beta_{ocup1\_autos4}$	0,765	0,258	2,970	--
$\beta_{ocup1\_renda}$	0,463	0,081	5,730	--
$\beta_{ocup1\_superiorcomp}$	0,372	0,093	4,000	--
$\beta_{custo}$	-2,690	0,187	-14,370	-2,750
$\beta_{custo\_freqsemanal3}$	-0,799	0,183	-4,360	--
$\beta_{custo\_renda14mil}$	0,518	0,249	2,080	--
$\beta_{custo\_posgrad}$	0,617	0,194	3,180	--
$\beta_{custo\_idade36}$	1,110	0,271	4,090	--
$\sigma_{panel}$	-0,287	0,061	-4,710	-0,302
Número de observações:	9432			9432
Número de indivíduos:	1179			1179
$\log \mathcal{L}(0) =$	-6537,76			-6537,76
$\log \mathcal{L}(\beta) =$	-5096,82			-5219,88
$-2 \cdot [ \log \mathcal{L}(0) - \log \mathcal{L}(\beta) ] =$	2881,90			2635,77
$\rho^2 =$	0,220			0,202
$\rho^2$ ajustado =	0,217			0,200

**Tabela 13 – Modelo ML *panel* com variáveis socioeconômicas e de uso do ônibus**

A seguir as interações para cada um dos atributos serão analisadas, com exceção do atributo *presença de painel no ponto*, para o qual nenhuma interação com variáveis socioeconômicas ou de uso de ônibus foi significativa.

**(a) Atributo *intervalo programado entre veículos***

O efeito do intervalo programado entre veículos no comportamento de escolha de linha é reforçado pelas interações com as variáveis *dummy autos4* e *superiorcomp*. Devido a estas interações, o coeficiente do intervalo passa a variar entre -0,161 e -0,303, ou seja: 2% abaixo e 84% acima (ambos em valor absoluto) do seu valor no modelo base.

Quanto à interação com *autos4*, infere-se que indivíduos com maior disponibilidade de automóveis se incomodem mais com as esperas do transporte público.

Já a influência da variável *superiorcomp* sobre o intervalo é, provavelmente, indireta. Esta variável representa um grupo com baixa presença de entrevistados que usam ônibus semanalmente e maior parcela de indivíduos que exercem atividade profissional (trabalho ou estágio), comparativamente às demais categorias da variável socioeconômica *escolaridade* (conforme se verifica no Apêndice I); estas características, conjuntamente, podem influenciar o efeito do intervalo programado sobre o comportamento.

**(b) Atributo *atraso***

No modelo em análise, o atraso interage com as variáveis *dummy freqmensal* e *freqsemanal*, fazendo com que seu coeficiente passe a variar entre -0,066 e -0,080, ou seja: 3% abaixo e 18% acima (ambos em valor absoluto) do seu valor no modelo base.

Verifica-se que usuários com frequência semanal são um pouco mais intolerantes ao atraso do que usuários com frequência mensal. Uma possível explicação é que os usuários semanais sofrem as consequências negativas do atraso mais vezes do que os usuários esporádicos.

No entanto, o coeficiente  $\beta_{\text{atraso}}$  (que, no modelo em discussão, representa a influência do atraso na decisão dos não usuários de ônibus), é igual a zero<sup>2</sup>, o que significa que os entrevistados que não usam ônibus não levam em consideração a possibilidade de atra-

---

<sup>2</sup> Um modelo com especificação similar à exibida na Tabela 13, porém incluindo  $\beta_{\text{atraso}}$ , foi estimado e os resultados mostraram que  $\beta_{\text{atraso}}$  era estatisticamente igual a zero. Por este motivo, este coeficiente foi retirado da especificação final, cujos resultados são os apresentados na Tabela 13.

so em sua decisão. Infere-se que este grupo não tenha experiência de uso de ônibus e, portanto, não saiba avaliar este atributo. Cabe comentar que um modelo ML *panel* especificado apenas com os efeitos principais dos atributos e estimado a partir das respostas dos 59 entrevistados que não usam ônibus, resultou em  $\beta_{\text{atraso}}$  estatisticamente igual a zero (conforme se verifica no Apêndice H).

**(c) Atributo *tempo de viagem***

O efeito do tempo de viagem no comportamento de escolha de linha é reforçado pela interação com a variável *dummy autos4*, de modo que seu coeficiente passa a variar entre -0,066, valor igual ao observado no modelo base, e -0,126, que está 91% acima (em valor absoluto) do observado no modelo base. Entende-se que entrevistados com maior disponibilidade de automóveis se incomodem mais com os tempos de viagem do transporte público<sup>3</sup>.

**(d) Atributo *ocupação do veículo (nível mínimo da ocupação)***

As variáveis *dummy autos4*, *renda* e *estudante* interagem com o nível mínimo da ocupação do veículo (sendo que apenas a última enfraquece o efeito do atributo), de tal forma que seu coeficiente passa a variar entre 2,150 e 4,208, ou seja: 22% abaixo e 53% acima do seu valor no modelo base.

Quanto à interação com *autos4*, infere-se que a maior disponibilidade (e conseqüente uso) de automóveis habitue os indivíduos a condições de conforto superiores às encontradas no ônibus.

Já a interação com a variável *renda* mostra que os indivíduos com renda domiciliar até R\$1.400 e entre R\$4.900 e R\$7.000 valorizam menos o conforto do que os demais. Sugere-se que os indivíduos com a mais baixa categoria de renda tenham maior dependência do ônibus (de fato, 93% deles usam ônibus semanalmente, e 62,5% não possuem automóvel no domicí-

---

<sup>3</sup> De fato, o tempo médio das viagens com transporte coletivo na RMSP é de aproximadamente 67 minutos, contra 31 minutos para as viagens com transporte individual, segundo a Pesquisa Origem e Destino 2007 para a RMSP (COMPANHIA DO METROPOLITANO DE SÃO PAULO, 2008).

lio), e por este motivo valorizam o conforto com menor intensidade do que valorizam outros atributos (como o custo, por exemplo).

Quanto aos indivíduos da faixa de renda domiciliar entre R\$4.900 e R\$7.000, sugere-se que a menor importância dada ao conforto na decisão se deva ao fato de que este grupo é o que mais concentra lexicográficos quanto ao custo: 24 indivíduos, o que corresponde a 35% de todos os lexicográficos quanto ao custo da amostra e 11% dos entrevistados desta faixa de renda. Lexicográficos quanto ao custo fazem suas escolhas com base somente no valor da tarifa, em detrimento dos demais atributos.

No que se refere à interação com a variável *estudante*, acredita-se que sua influência seja indireta e esteja relacionada a outras características do grupo de entrevistados que são estudantes, como a maior presença de usuários semanais de ônibus e a maior parcela de entrevistados do sexo masculino ou com até 25 anos, comparativamente às demais categorias da variável socioeconômica *ocupação*, conforme se verifica no Apêndice I.

**(e) Atributo *ocupação do veículo* (nível intermediário da *ocupação*)**

Três variáveis *dummy* interagem com o nível intermediário da ocupação do veículo, todas reforçando seu efeito sobre o comportamento: *autos4*, *renda* e *superiorcomp*. Destaca-se que as duas primeiras também interagem com o nível mínimo da ocupação (também reforçando seu efeito), de modo que as explicações para estas interações são as mesmas expostas anteriormente.

Devido às interações, o coeficiente do nível intermediário da ocupação do veículo passa a variar entre 1,350 e 2,950, ou seja: 22% abaixo e 72% acima do seu valor no modelo base.

No que se refere à variável *superiorcomp*, acredita-se que sua influência seja indireta. Características do grupo de entrevistados que já concluíram o Ensino Superior, tais como a menor presença de indivíduos que usam ônibus semanalmente e a maior parcela de entrevistados do sexo feminino, comparativamente às demais categorias da variável socioeconômica *escolaridade*, (conforme se verifica no Apêndice I), podem influenciar o efeito do nível intermediário da ocupação sobre o comportamento de escolha de linha.

**(f) Atributo *custo***

O atributo *custo* interage com as variáveis *dummy freqsemanal3*, *renda14mil*, *posgrad* e *idade36*, sendo *freqsemanal3* a única que intensifica o efeito do custo sobre o comportamento de escolha de linha. Devido a estas interações, o coeficiente do custo passa a variar entre -0,445 e -3,489, ou seja: 84% abaixo e 27% acima (ambos em valor absoluto) do seu valor no modelo base.

No que se refere à interação com a variável *freqsemanal3*, uma explicação plausível é que, ao andar mais de ônibus, cresce a participação deste modo para as despesas individuais. Além disto, este grupo tem, comparativamente às demais categorias da variável *frequência de uso de ônibus*, pequena proporção de entrevistados que exercem atividade remunerada (trabalho ou estágio) e maior concentração de entrevistados nas faixas de renda domiciliar mais baixas (conforme se verifica no Apêndice I).

A influência da variável *renda14mil* sobre o efeito do custo é direta e, conforme o esperado, enfraquece seu efeito.

Acredita-se que o efeito das variáveis *posgrad* e *idade36* sobre o custo seja indireto. Em ambos os casos, características tais como a menor proporção de entrevistados que usam ônibus semanalmente e a maior proporção de indivíduos nas faixas de renda domiciliar acima de R\$7.000 (comparativamente às demais categorias das variáveis socioeconômicas *escolaridade* e *idade*, conforme se verifica no Apêndice I), podem justificar a interação.



## 6 CONCLUSÕES E RECOMENDAÇÕES

O objetivo geral desta pesquisa foi medir a influência, na escolha de linha, de painéis eletrônicos nos pontos de ônibus que informam os intervalos de tempo previstos para passagem das linhas. Um experimento de escolha declarada com desenho eficiente foi aplicado a uma amostra de 1179 entrevistados pertencentes à comunidade da USP.

Os entrevistados são predominantemente jovens com até 25 anos (63,7% da amostra), do sexo masculino (59,5%), usuários frequentes de ônibus (80%), estudantes (81,1%) e com pelo menos um automóvel no domicílio (75,5%).

Para analisar os resultados, foram estimados os seguintes tipos de modelos *ML panel*: apenas com efeitos principais, a partir dos quais se obtiveram as taxas médias de substituição entre atributos; considerando a não linearidade das preferências; incorporando as interações entre atributos; e considerando a variabilidade sistemática das preferências entre indivíduos.

Todos os modelos respondem à questão de pesquisa formalizada no Capítulo 1, confirmando que a presença do painel no ponto de ônibus tem, sim, influência sobre a escolha da linha. Além desta, as demais conclusões da pesquisa foram:

- a presença de painel no ponto produz a mesma utilidade que a redução de 5 minutos de viagem, e os entrevistados estariam dispostos a pagar, em média, cerca de R\$0,12 adicionais para que o ponto tivesse painel;
- a existência de painel no ponto diminui a desutilidade marginal da espera, e isto ocorre com mais intensidade quando a espera está associada ao *atraso* do que quando está associada ao *intervalo programado entre veículos*. No caso do atraso, a desutilidade marginal da espera cai 24% ao passar da condição “sem painel” para a condição “com painel”. Já no caso do intervalo, a queda é de 13%;
- a taxa marginal de substituição entre tempo de espera associado ao *intervalo programado* e tempo de viagem está, em média, dentro das expectativas: 1 minuto de espera equivale a 2,5 minutos de viagem. Já a taxa entre o tempo de espera

associado ao *atraso* e o tempo de viagem é, em média, bastante inferior ao que se esperava: 1 minuto de espera por atraso equivale a 1 minuto de viagem;

- a desutilidade marginal do atraso é 2,2 vezes maior quando ele é precedido por intervalos programados maiores (15 ou 20 minutos), do que quando precedido por intervalos menores (5 ou 10 minutos);
- os efeitos do *intervalo programado*, do *tempo de viagem* e do *atraso* sobre o comportamento não são lineares: a desutilidade marginal do intervalo é maior para os primeiros 10 minutos de espera, e decresce 19% para os minutos seguintes; já a desutilidade marginal do tempo de viagem é maior para os primeiros 30 minutos de viagem, e decresce 63% para os minutos seguintes; e a desutilidade marginal do atraso é menor para os primeiros 9 minutos de espera, aumentando em 149% para os minutos seguintes;
- o valor médio do tempo de viagem, decorrente das características socioeconômicas da amostra, é considerado baixo: R\$1,44 / hora;
- o maior conforto durante a viagem diminui a desutilidade marginal do tempo de viagem: a possibilidade de viajar sentado em um veículo quase vazio (ao invés de viajar em pé em um veículo lotado) reduz a desutilidade marginal do tempo de viagem entre 16% e 31%, dependendo da especificação adotada para o componente determinístico da utilidade. Já a possibilidade de viajar em pé em um veículo não lotado reduz a desutilidade marginal do tempo de viagem em 17%;
- viajar sentado em um veículo quase vazio (ao invés de viajar em pé em um veículo lotado) produz, em média, a mesma utilidade que reduzir em cerca de 42 minutos o tempo de viagem. Já a possibilidade de viajar em pé em um veículo não lotado produz, em média, a mesma utilidade que reduzir em cerca de 26 minutos o tempo de viagem;
- há bastante variabilidade sistemática das preferências entre os entrevistados, de modo que o comportamento de escolha de linha é significativamente afetado pelas características socioeconômicas e de uso de ônibus dos indivíduos. Verifica-se que o valor do tempo, para determinado grupo de entrevistados, pode chegar a R\$17,00/hora, enquanto a disposição a pagar pelo painel pode chegar a R\$0,77.

Acredita-se que, dada a baixa importância relativa do coeficiente do atraso nos resultados dos modelos, a forma de comunicação deste atributo no experimento não foi suficiente para capturar a complexidade do comportamento dos entrevistados em relação ao atraso e à sua interação com o painel. Sugere-se, para experimentos futuros, comunicar as probabilidades de ocorrência de diversos valores de atraso, por exemplo: 50% de probabilidade de não haver atraso, 30% de probabilidade de ocorrer atraso de 5 minutos e 20% de probabilidade de ocorrer atraso de 8 minutos.

Em estudos futuros, sugere-se, ainda, estudar a influência do painel em contextos com diferentes motivos de viagem (trabalho, estudo, lazer etc.) e condições de pressão de tempo, e também em função da confiabilidade da informação fornecida pelos painéis, além da investigação da incorporação e do uso da informação em tempo real no processo decisório dos passageiros.

Cabe ressaltar que os resultados obtidos são específicos para a amostra coletada e não representam as preferências de um usuário de ônibus médio da RMSP. No entanto, a metodologia adotada pode ser estendida a outras amostras, com a finalidade de representar mais fielmente o comportamento de tais usuários.

Acredita-se que esta pesquisa, ao incorporar a presença de painel no ponto como elemento adicional da decisão, permita ampliar o entendimento da decisão de escolha de linha, revelando a importância relativa da expectativa de redução da incerteza durante a etapa de espera pelo ônibus (e, portanto, da expectativa de uma experiência de espera mais positiva), frente a outros atributos do transporte público por ônibus.

## REFERÊNCIAS BIBLIOGRÁFICAS

AARTS, H.; VERPLANKEN, B.; VAN KNIPPENBERG, A. **Habit and information use in travel mode choices.** Acta Psychologica, v. 96, n. 1-2, p. 1-14, 1997.

ABRANTES, P. A. L.; WARDMAN, M. R. **Meta-analysis of UK values of travel time: an update.** Transportation Research Part A, v. 45, n. 1, p. 1-17, 2011.

BEN-AKIVA, M.; LERMAN, S. R. **Discrete Choice Analysis: Theory and Application to Travel Demand.** 1. ed. Cambridge: The MIT Press, 1985. 390 p.

BEN-ELIA, E.; DI PACE, R.; BIFULCO, G. N.; SHIFTAN, Y. **The impact of travel information's accuracy on route-choice.** Transportation Research Part C, v. 26, p. 146-159, 2013.

BIERLAIRE, M. Biogeme: a free package for the estimation of discrete choice models. In: 3<sup>rd</sup> SWISS TRANSPORTATION RESEARCH CONFERENCE. **Conference Paper.** Ascona, 2003.

BLIEMER, M. C. J.; ROSE, J. M. **Designing stated choice experiments: state of the art.** In: Kitamura, R., Yoshii, T., Yamamoto, T. The Expanding Sphere of Travel Behaviour Research: Selected Papers from the 11<sup>th</sup> International Conference on Travel Behaviour Research. United Kingdom: Emerald, 2009. 924 p.

BONSALL, P. **Traveller behavior: decision-making in an unpredictable world.** Journal of Intelligent Transportation Systems, v. 8, n. 1, p. 45-60, 2004.

CARVALHO, E. C. S.; STRAMBI, O. Comportamento de usuários de ônibus urbanos frente à disponibilização de informações em tempo real sobre tempos de partida das linhas. In: XXVI CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES. Joinville, 2012.

CEDER, A. **Public Transit Planning and Operation: Theory, Modeling and Practice.** 1. ed. Oxford: Elsevier, 2007. 626 p.

CHOICE METRICS. **Ngene 1.1 User Manual and Reference Guide.** Choice Metrics, 2011. 247 p.

CHORUS, C. G.; ARENTZE, T. A.; MOLIN, E. J. E. **The value of travel information: decision strategy-specific conceptualizations and numerical examples.** *Transportation Research Part B*, v. 40, n. 6, p. 504-519, 2006.

CHORUS, C. G.; ARENTZE, T. A.; TIMMERMANS, H. J. P. **A random regret-minimization model of travel choice.** *Transportation Research Part B*, v. 42, n. 1, p. 1-18, 2008.

CHORUS, C. G.; ARENTZE, T. A.; TIMMERMANS, H. J. P.; MOLIN, E. J. E.; VAN WEE, B. **Travelers' need for information in traffic and transit: results from a web survey.** *Journal of Intelligent Transportation Systems*, v. 11, n. 2, p. 57-67, 2007.

CHORUS, C. G.; MOLIN, E. J. E.; VAN WEE, B. **Use and effects of advanced traveller information services (ATIS): a review of the literature.** *Transport Reviews*, v. 26, n. 2, p. 127-149, 2006.

COMPANHIA DO METROPOLITANO DE SÃO PAULO. **Pesquisa Origem e Destino 2007 da Região Metropolitana de São Paulo: Síntese das Informações da Pesquisa Domiciliar.** São Paulo: 2008.

DROIT-VOLET, S.; MECK, W. H. **How emotions colour our perception of time.** *Trends in Cognitive Sciences*, v. 11, n. 12, p. 504-513, 2007.

DUBÉ-RIOUX, L.; SCHMITT, B. H.; LECLERC, F. **Consumers' reactions to waiting: when delays affect the perception of service quality.** *Advances in Consumer Research*, v. 16, p. 59-63, 1989.

DZIEKAN, K. **Ease-of-use in public transportation: a user perspective on information and orientation aspects.** 2008. 63 p. Tese (Doutorado) - Royal Institute of Technology, Estocolmo, 2008.

DZIEKAN, K.; KOTTENHOFF, K. **Dynamic at-stop real-time information displays for public transport: effects on customers.** *Transportation Research Part A*, v. 41, n. 6, p. 489-501, 2007.

DZIEKAN, K.; VERMEULEN, A. **Psychological effects of and design preferences for real-time information displays.** *Journal of Public Transportation*, v. 9, n. 1, p. 71-89, 2006.

ETTEMA, D.; TIMMERMANS, H. **Costs of travel time uncertainty and benefits of travel time information: conceptual model and numerical examples.** *Transportation Research Part C*, v. 14, n. 5, p. 335-350, 2006.

FRIMAN, M. **Affective dimensions of the waiting experience.** *Transportation Research Part F*, v. 13, n. 3, p. 197-205, 2010.

GENTILE, G.; NGUYEN, S.; PALLOTTINO, S. **Route choice on transit networks with online information at stops.** *Transportation Science*, v. 39, n. 3, p. 289-297, 2005.

GIL, A. C. **Como Elaborar Projetos de Pesquisa.** 4. ed. São Paulo: Atlas, 2002. 175 p.

GROTENHUIS, J. W.; WIEGMANS, B. W.; RIETVELD, P. **The desired quality of integrated multimodal travel information in public transport: customer needs for time and effort savings.** *Transport Policy*, v. 14, n. 1, p. 27-38, 2007.

HALL, R. W. **Passenger waiting time and information acquisition using automatic vehicle location for verification.** *Transportation Planning and Technology*, v. 24, n. 3, p. 249-269, 2001.

HOLDSWORTH, N.; ENOCH, M. P.; ISON, S. G. **Examining the political and practical reality of bus-based real time passenger information.** *Transportation Planning and Technology*, v. 30, n. 2-3, p. 183-204, 2007.

KIM, J.; LEE, B.; OH, S. **Passenger choice models for analysis of impacts of real-time bus information on crowdedness.** *Transportation Research Record*, n. 2112, p. 119-126, 2009.

LARSEN, O. I.; SUNDE, Ø. **Waiting time and the role and value of information in scheduled transport.** *Research in Transportation Economics*, v. 23, n. 1, p. 41-52, 2008.

LIU, Y.; BUNKER, J.; FERREIRA, L. **Transit users' route-choice modelling in transit assignment: a review.** *Transport Reviews*, v. 30, n. 6, p. 753-769, 2010.

MAISTER, D. H. **The psychology of waiting lines.** In: Czepiel, J. A., Solomon, M. R., Surprenant, C. F. *The service encounter: managing employee/customer interaction in service businesses.* 1. ed. Lexington: Lexington Books, 1985. 338 p.

MANTEL, S. P.; KELLARIS, J. J. **Cognitive determinants of consumers' time perceptions: the impact of resources required and available.** Journal of Consumer Research, v. 29, n. 4, p. 531-538, 2003.

MECK, W. H. **Neuropsychology of timing and time perception.** Brain and Cognition, v. 58, n. 1, p. 1-8, 2005.

NAVARRETE, F. J. **Valoración subjetiva de los tipos de transbordo en transporte público: caso Santiago de Chile.** 2010. 149 p. Dissertação (Mestrado) - Pontificia Universidad Católica de Chile, Santiago de Chile, 2010.

ORTÚZAR, J. D.; WILLUMSEN, L. G. **Modelling Transport.** 4. ed. Great Britain: John Wiley & Sons, 2011. 586 p.

PSARROS, I., KEPAPTSOGLU, K.; KARLAFTIS, M. G.; EFSTATHIADIS, S. Wait time at bus stops: an empirical investigation of passenger perceptions. In: 90<sup>th</sup> TRB ANNUAL MEETING. **Conference Paper.** Washington, 2011.

RAZO, M.; GAO, S. **A rank-dependent expected utility model for strategic route choice with stated preference data.** Transportation Research Part C, v. 27, Special Issue, p. 117-130, 2013.

SORIGUERA, F. Value of freeway travel time information. In: 92<sup>nd</sup> TRB ANNUAL MEETING. **Conference Paper.** Washington, 2013.

STRADLING, S. G. **Transport user needs and marketing public transport.** Institution of Civil Engineers – Municipal Engineer, v. 151, n. 1, p. 23-28, 2002.

TRAIN, K. E. **Discrete Choice Methods with Simulation.** 2. ed. New York: Cambridge University Press, 2009. 388 p.

VERPLANKEN, B.; AARTS, H.; VAN KNIPPENBERG, A. **Habit, information acquisition, and the process of making travel mode choices.** European Journal of Social Psychology, v. 27, n. 5, p. 539-560, 1997.

WARDMAN, M. **Public transport values of time.** Transport Policy, v. 11, n. 4, p. 363-377, 2004.

WARDMAN, M.; WHELAN, G. **Twenty years of rail crowding valuation studies: evidence and lessons from British experience.** *Transport Reviews*, n. 31, v. 3, p. 379-398, 2011.

ZHANG, F.; SHEN, Q.; CLIFTON, K. J. **Examination of traveler responses to real-time information about bus arrivals using panel data.** *Transportation Research Record*, n. 2082, p. 107-115, 2008.

ZHAO, S.; YANG, Z.; ZHANG, X. Traveller's information choice behaviour in response to transit information services. In: **12<sup>th</sup> WORLD CONFERENCE IN TRANSPORTATION RESEARCH. Conference Paper.** Lisboa, 2010.



## APÊNDICE A – PRIMEIRO EXPERIMENTO PILOTO

### (a) Contexto do experimento

Com a finalidade de descrever o contexto do experimento, é apresentada a seguir a instrução dada em todas as situações de escolha, seguida de um exemplo de tabela (na Figura 8) com a caracterização das alternativas:

*“Suponha que você tem um compromisso, e tanto a LINHA 1 quanto a LINHA 2 podem deixá-lo(a) no local do compromisso. O tempo de viagem (dentro do ônibus) é igual para as duas linhas. No entanto, estas linhas não passam pelo mesmo ponto. Você tem, então, duas alternativas: **andar até o PONTO 1 e pegar a LINHA 1** ou **andar até o PONTO 2 e pegar a LINHA 2**. Considere que os tempos de caminhada até os pontos são pequenos, e que apenas um dos pontos tem painel eletrônico mostrando quanto tempo falta para o ônibus passar. Levando em conta as características de cada alternativa (apresentadas na tabela abaixo), **qual você escolhe?**”*

	PONTO 1 + LINHA 1	PONTO 2 + LINHA 2
<b>A linha está programada para passar a cada:</b>	18 minutos	12 minutos
<b>Mas pode demorar até:</b>	26 minutos	24 minutos
<b>O ponto tem painel eletrônico?</b>	Sim	Não

Figura 8 – Exemplo de tabela com as características das alternativas do primeiro piloto

### (b) Principais definições prévias à criação do experimento

As duas alternativas do experimento (*Ponto 1 + Linha 1* e *Ponto 2 + Linha 2*) foram caracterizadas pelos atributos constantes na Tabela 14 a seguir:

<b>Atributos</b>	<b>Níveis</b>	<b>Coefficientes na função utilidade</b>
Existência de <b>painel</b> no ponto	1 = Sim / 0 = Não	$\beta_{display}$
<b>Intervalo</b> de tempo programado para a passagem de veículos consecutivos da mesma linha	6 / 12 / 18 minutos	$\beta_{interv}$
Possível <b>atraso</b> em relação ao intervalo programado	4 / 8 / 12 minutos	$\beta_{atraso}$

**Tabela 14 – Atributos para primeiro experimento piloto**

O modelo econométrico escolhido foi o MNL, e o componente determinístico da função utilidade ( $V_{in}$ , onde  $n$  refere-se ao indivíduo, e  $i$  à alternativa), idêntico para ambas as alternativas, foi especificado com efeitos principais e interações:

$$\begin{aligned}
 V_{in} = & \beta_{display} \cdot display_{in} + \beta_{interv} \cdot interv_{in} + \\
 & \beta_{atraso} \cdot atraso_{in} + \\
 & \beta_{interv\_display} \cdot (interv_{in} \cdot display_{in}) + \\
 & \beta_{atraso\_display} \cdot (atraso_{in} \cdot display_{in})
 \end{aligned}
 \tag{19}$$

Não foram adotadas aproximações para os coeficientes (a fim de criar o desenho eficiente). Ao invés disto, todos os coeficientes da eq. (19) foram definidos como iguais a zero.

Optou-se pela criação de 12 situações de escolha, agrupadas em 6 blocos (de modo que cada entrevistado respondesse a 2 situações de escolha). O indicador de ineficiência escolhido para ser usado pelo algoritmo (para criação do desenho eficiente) foi o  $D_{error}$ .

### (c) Desenho eficiente criado

O desenho criado, mostrado na Tabela 15, é balanceado para os níveis dos atributos *painel* e *intervalo*. No entanto, não é para os níveis do atributo *atraso*: o nível *4 minutos* apareceu 10 vezes no desenho, enquanto o *8 minutos* apareceu 6, e o *12 minutos* apareceu 8 vezes.

Bloco	Sit. escolha	Ponto 1 + Linha 1			Ponto 2 + Linha 2		
		Painel	Intervalo	Atraso	Painel	Intervalo	Atraso
1	1	1	6	8	0	6	12
1	2	1	18	4	0	18	8
2	3	0	18	8	1	12	12
2	4	0	12	12	1	6	4
3	5	1	12	8	0	12	12
3	6	1	12	4	0	18	4
4	7	1	6	4	0	6	8
4	8	0	18	4	1	12	8
5	9	0	12	4	1	6	4
5	10	0	18	12	1	6	4
6	11	1	18	4	0	12	12
6	12	1	6	12	0	18	12

Tabela 15 – Desenho eficiente para o primeiro experimento piloto

**(d) Resultados da coleta de dados**

A partir das 118 respostas obtidas (2 situações de escolha respondidas por cada um dos 59 entrevistados) foi estimado o modelo MNL mostrado na Tabela 16 abaixo:

Coeficiente	Valor	Erro padrão	Teste t
$\beta_{display}$	1,760	0,425	4,140
$\beta_{interv}$	-0,347	0,084	-4,150
$\beta_{atraso}$	-0,255	0,088	-2,890

Número de observações: 118  
Número de indivíduos: 118  
 $\mathcal{L}(0) = -81,79$   
 $\mathcal{L}(\beta) = -66,26$   
 $-2 \cdot [\mathcal{L}(0) - \mathcal{L}(\beta)] = 31,06$   
 $\rho^2 = 0,190$   
 $\rho^2 \text{ ajustado} = 0,153$

Tabela 16 – Modelo MNL estimado com resultados do primeiro experimento piloto

## APÊNDICE B – SEGUNDO EXPERIMENTO PILOTO

### (a) Contexto do experimento

Com a finalidade de descrever o contexto do experimento, é apresentada a seguir a instrução dada em todas as situações de escolha, seguida de um exemplo de tabela (na Figura 9) com a caracterização das alternativas:

*“Suponha que você irá de ônibus a um compromisso e que tanto a LINHA 1 quanto a LINHA 2 servem para levá-lo(a) até o local desejado. Porém, estas linhas não passam pelo mesmo ponto. Você tem, então, duas alternativas: **andar em uma direção até o PONTO 1 e pegar a LINHA 1** ou **andar em outra direção até o PONTO 2 e pegar a LINHA 2**. Considere que os tempos de caminhada até os pontos são pequenos. Os pontos podem ou não ter **painel eletrônico** mostrando quanto tempo falta para o ônibus passar. Levando em conta esta e outras características de cada alternativa (apresentadas na tabela abaixo), **qual você escolhe?**”*

	PONTO 1 + LINHA 1	PONTO 2 + LINHA 2
<b>A linha está programada para passar a cada:</b>	20 minutos	15 minutos
<b>Mas pode demorar até:</b>	39 minutos	34 minutos
<b>O ponto tem painel eletrônico?</b>	Não	Sim
<b>O tempo de viagem (dentro do ônibus) é:</b>	50 minutos	50 minutos
<b>A passagem custa:</b>	R\$ 3,00	R\$ 3,30
<b>O ônibus costuma estar:</b>	Quase vazio. Você poderá viajar sentado.	Lotado. Todos os bancos estão ocupados e você terá que viajar em pé (com muitas outras pessoas em pé).

Figura 9 – Exemplo de tabela com as características das alternativas do segundo piloto

### (b) Principais definições prévias à criação do experimento

#### *Alternativas e atributos*

As duas alternativas do experimento (*Ponto 1 + Linha 1* e *Ponto 2 + Linha 2*) foram caracterizadas pelos atributos constantes na Tabela 17 a seguir:

<b>Atributos</b>	<b>Níveis</b>	<b>Coefficientes na função utilidade</b>
Existência de <b>painel</b> no ponto	1 = Sim / 0 = Não	$\beta_{display}$
<b>Intervalo</b> de tempo programado para a passagem de veículos consecutivos da mesma linha	5 / 10 / 15 / 20 minutos	$\beta_{interv}$
Possível <b>atraso</b> em relação ao intervalo programado	4 / 9 / 14 / 19 minutos	$\beta_{atraso}$
<b>Tempo de viagem</b> dentro do veículo	20 / 30 / 40 / 50 minutos	$\beta_{tviagem}$
<b>Ocupação</b> do veículo quando chega ao ponto	Mínima ( <i>dummy</i> : ocup0) “Quase vazio. Você poderá viajar sentado”.	$\beta_{ocup0}$ e $\beta_{ocup1}$
	Intermediária ( <i>dummy</i> : ocup1) “Nem vazio, nem lotado. Todos os bancos estão ocupados e você terá que viajar em pé (com poucas outras pessoas em pé)”.	
	Máxima “Lotado. Todos os bancos estão ocupados e você terá que viajar em pé (com muitas outras pessoas em pé)”.	
<b>Custo</b> da viagem	3,00 / 3,30 R\$	$\beta_{custo}$

**Tabela 17 – Atributos para segundo experimento piloto**

Ressalta-se que estes atributos e seus níveis são exatamente os mesmos utilizados no experimento para coleta de dados desta pesquisa (descrito no Capítulo 4).

Foram criadas duas variáveis *dummy* (*ocup0* e *ocup1*) para representar efeitos não lineares do atributo *ocupação*. A Tabela 18 a seguir sumariza a relação entre o atributo e as variáveis *dummy*:

Ocupação do veículo	ocup0	ocup1
Mínima	1	0
Intermediária	0	1
Máxima	0	0

**Tabela 18 – Relação entre atributo *ocupação* e variáveis *dummy***

### ***Restrições ao desenho***

Os níveis dos atributos *intervalo*, *atraso* e *tempo de viagem* foram segmentados em dois grupos: “valores menores” e “valores maiores” (conforme Tabela 19 abaixo), de modo que duas alternativas concorrentes tivessem valores para o mesmo atributo que pertencessem ao mesmo grupo (o grupo dos “valores menores” ou o grupo dos “valores maiores”).

Atributo	Valores menores		Valores maiores	
Intervalo	5	10	15	20
Atraso	4	9	14	19
Tempo de viagem	20	30	40	50

**Tabela 19 – Estratégia para combinar atributos *intervalo*, *atraso* e *tempo de viagem***

Esta restrição é exatamente a mesma aplicada ao experimento para coleta de dados desta pesquisa (descrito no Capítulo 4).

### ***Especificação econométrica***

O modelo econométrico escolhido foi o MNL, e o componente determinístico da função utilidade ( $V_{in}$ , onde  $n$  refere-se ao indivíduo, e  $i$  à alternativa), idêntico para ambas as alternativas, foi especificado com efeitos principais e interações:

$$\begin{aligned}
V_{in} = & \beta_{display} \cdot display_{in} + \beta_{interv} \cdot interv_{in} + \\
& \beta_{atraso} \cdot atraso_{in} + \beta_{tviagem} \cdot tviagem_{in} + \\
& \beta_{ocup0} \cdot ocup0_{in} + \beta_{ocup1} \cdot ocup1_{in} + \beta_{custo} \cdot custo_{in} + \\
& \beta_{interv\_display} \cdot (interv_{in} \cdot display_{in}) + \\
& \beta_{atraso\_display} \cdot (atraso_{in} \cdot display_{in}) + \\
& \beta_{tviagem\_ocup0} \cdot (tviagem_{in} \cdot ocup0_{in}) + \\
& \beta_{tviagem\_ocup1} \cdot (tviagem_{in} \cdot ocup1_{in}) +
\end{aligned} \tag{20}$$

### *Aproximações para parâmetros*

Foram adotadas distribuições uniformes aproximadas para os coeficientes (a fim de criar o desenho eficiente), derivadas de referências da literatura da área de Transportes. As aproximações para todos os coeficientes da eq. (20) e as referências utilizadas estão listadas abaixo:

- $\beta_{tviagem} \sim U(-0,08, -0,02)$

Adotaram-se valores próximos aos encontrados por Carvalho e Strambi (2012) e Navarrete (2010). Ambas as pesquisas estimaram coeficientes para o tempo de viagem no transporte público para usuários de São Paulo e Santiago do Chile, respectivamente.

- $\beta_{interv} \sim U(-0,24, -0,03)$

Adotou-se o intervalo encontrado por Wardman (2004) para a taxa marginal de substituição entre intervalo e tempo de viagem:  $(\beta_{intervalo} / \beta_{tviagem}) \in [1,5, 3,0]$ . Então, fez-se:  $\beta_{interv} \sim U(-0,08 \cdot 3, -0,02 \cdot 1,5) = U(-0,24, -0,03)$ .

- $\beta_{atraso} \sim U(-0,24, -0,03)$

Adotou-se o mesmo raciocínio aplicado a  $\beta_{interv}$ .

- $\beta_{display} \sim U(0,18, 1,44)$

A partir das taxas marginais de substituição entre painel e intervalo ( $\beta_{display} / \beta_{intervalo} = -5,07$ ) e entre painel e atraso ( $\beta_{display} / \beta_{atraso} = -6,90$ ) obtidas dos resultados do primeiro experimento piloto (conforme Tabela 16 do Apêndice A), foi adotada a aproximação:  $\beta_{display} / \beta_{interv} \sim \beta_{display} / \beta_{atraso} \sim -6,00$ . E então, fez-se:  $\beta_{display} \sim U(-0,03 \cdot -6, -0,24 \cdot -6) = U(0,18, 1,44)$ .

- $\beta_{\text{ocup0}} \sim U(0,40, 1,60)$

Adotou-se o resultado encontrado por Kim; Lee e Oh (2009), segundo o qual o usuário de ônibus aceitaria que o tempo de viagem aumentasse em 20 minutos se pudesse viajar sentado ao invés de viajar em pé em um veículo com alta ocupação. Portanto, foi adotada a taxa:  $\beta_{\text{ocup0}} / \beta_{\text{tviagem}} = -20,00$ .

E então, fez-se:  $\beta_{\text{ocup0}} \sim U(-0,02 \cdot -20, -0,08 \cdot -20) = U(0,40, 1,60)$ .

- $\beta_{\text{ocup1}} \sim U(0,20, 0,80)$

Adotou-se o resultado encontrado por Kim; Lee e Oh (2009), segundo o qual o usuário de ônibus aceitaria que o tempo de viagem aumentasse em 10 minutos se pudesse viajar em pé (porém com a possibilidade de se movimentar com facilidade dentro do veículo) ao invés de viajar em pé em um veículo com alta ocupação. Portanto, foi adotada a taxa:  $\beta_{\text{ocup1}} / \beta_{\text{tviagem}} = -10,00$ .

E então, fez-se:  $\beta_{\text{ocup1}} \sim U(-0,02 \cdot -10, -0,08 \cdot -10) = U(0,20, 0,80)$ .

- $\beta_{\text{custo}} \sim U(-1,60, -0,24)$

Adotou-se uma faixa de variação como referência para o valor do tempo: de R\$3,00 por hora a R\$5,00 por hora. Segue breve descrição da dedução desta referência: a renda média familiar mensal para a RMSP em outubro de 2007 (divulgada pela pesquisa Origem e Destino 2007 da RMSP, igual a R\$2211,00) foi corrigida pela cesta básica e pelo salário mínimo para valores de dezembro de 2012 (período da realização deste experimento piloto) e arredondada para R\$3500,00. Para um mês com 180 horas de trabalho, esta renda corresponde a aproximadamente R\$20,00 por hora trabalhada. Supondo que a renda de um indivíduo da família (e lembrando que a maior parte da amostra que seria pesquisada era composta por estudantes) varie entre 30% e 50% da renda familiar, tem-se que a renda média individual varia de R\$6,00 a R\$10,00 por hora. Supondo que o valor comportamental do tempo de viagem (para uma viagem genérica, não necessariamente a trabalho) seja a metade disto, tem-se que:  $\beta_{\text{custo}} / \beta_{\text{tviagem}} \in [3, 5]$  em R\$/h. Portanto foram adotadas as taxas:  $\beta_{\text{tviagem}} / \beta_{\text{custo}} = 5/60$  e  $\beta_{\text{tviagem}} / \beta_{\text{custo}} = 3/60$ .

E então, fez-se:  $\beta_{\text{custo}} \sim U(-0,08 \cdot 60/3, -0,02 \cdot 60/5) = U(-1,60, -0,24)$ .

- $\beta_{\text{interv\_display}} \sim U(0,006, 0,048)$

Adotou-se a relação encontrada por Dziekann e Kottenhoff (2007), segundo a qual os painéis têm potencial para reduzir o tempo percebido de espera em até 20%. Muito embora, neste experimento, a decisão do entrevistado seja tomada sem que ele “veja” a informação mostrada pelo painel, admitiu-se que 1 minuto de espera com painel produz 80% da utilidade de 1 minuto de espera sem painel (ou seja: 20% a menos). Portanto, foi adotada a taxa:  $(\beta_{\text{interv}} + \beta_{\text{interv\_display}}) / \beta_{\text{interv}} = 0,80$ , que equivale a  $\beta_{\text{interv\_display}} = -0,20 \cdot \beta_{\text{interv}}$ .

E então, fez-se:  $\beta_{\text{interv\_display}} \sim U(-0,03 \cdot -0,20, -0,24 \cdot -0,20) = U(0,006, 0,048)$ .



- $\beta_{\text{atraso\_display}} \sim U(0,006, 0,048)$

Adotou-se o mesmo raciocínio aplicado a  $\beta_{\text{interv\_display}}$ .

- $\beta_{\text{tviagem\_ocup0}} \sim U(0,012, 0,048)$

Adotou-se a relação encontrada por Wardman e Whelan (2011), segundo a qual 1 minuto de viagem em pé em veículo de alta ocupação (lotado) produz a mesma utilidade que 2,5 minutos de viagem em uma situação confortável (viajar sentado). Portanto, foi adotada a taxa:  $\beta_{\text{tviagem}} / (\beta_{\text{tviagem}} + \beta_{\text{tviagem\_ocup0}}) = 2,5$ , que equivale a:  $\beta_{\text{tviagem\_ocup0}} = -0,60 \cdot \beta_{\text{tviagem}}$ .

E então, fez-se:  $\beta_{\text{tviagem\_ocup0}} \sim U(-0,02 \cdot -0,60, -0,08 \cdot -0,60) = U(0,012, 0,048)$ .

- $\beta_{\text{tviagem\_ocup1}} \sim U(0,008, 0,032)$

Adotou-se a relação encontrada por Wardman e Whelan (2011), segundo a qual 1 minuto de viagem em pé em veículo de alta ocupação (lotado) produz a mesma utilidade que 1,67 minutos de viagem também em pé, porém com possibilidade de se movimentar com facilidade no interior do veículo. Portanto, foi adotada a taxa:  $\beta_{\text{tviagem}} / (\beta_{\text{tviagem}} + \beta_{\text{tviagem\_ocup1}}) = 1,67$ , que equivale a:  $\beta_{\text{tviagem\_ocup1}} = -0,40 \cdot \beta_{\text{tviagem}}$ .

E então, fez-se:  $\beta_{\text{tviagem\_ocup1}} \sim U(-0,02 \cdot -0,40, -0,08 \cdot -0,40) = U(0,008, 0,032)$ .

### ***Situações de escolha***

Optou-se pela criação de 24 situações de escolha, agrupadas em 4 blocos (de modo que cada entrevistado respondesse a 6 situações de escolha), respeitando o balanceamento entre níveis do mesmo atributo. O indicador de ineficiência escolhido para ser usado pelo algoritmo (para criação do desenho eficiente) foi o  $D_{b\_error}$ .

### **(c) Desenho eficiente criado**

O desenho criado é mostrado na Tabela 20, a seguir:

Bloco	Situação de escolha	Ponto 1 + Linha 1						Ponto 2 + Linha 2					
		Painel	Intervalo	Atraso	Tempo de viagem	Ocupação	Custo	Painel	Intervalo	Atraso	Tempo de viagem	Ocupação	Custo
1	1	1	10	9	40	0	3,3	1	10	9	40	2	3
1	2	1	15	4	30	1	3	0	15	9	30	0	3,3
1	3	0	5	14	40	0	3,3	1	10	14	50	1	3
1	4	1	20	4	50	2	3,3	0	20	4	40	1	3
1	5	1	5	9	20	2	3	0	5	4	30	1	3,3
1	6	0	15	14	20	1	3,3	0	20	14	20	0	3
2	7	0	20	14	30	1	3,3	1	20	19	20	2	3
2	8	1	20	19	30	2	3,3	0	20	14	30	1	3
2	9	0	5	19	20	0	3	1	5	19	20	2	3,3
2	10	0	20	19	50	0	3	1	15	19	50	2	3,3
2	11	0	5	14	40	1	3	1	10	14	40	0	3,3
2	12	0	10	4	50	0	3	1	5	9	40	1	3,3
3	13	1	20	9	20	1	3	0	15	4	30	0	3,3
3	14	0	20	9	20	1	3,3	0	15	9	30	2	3
3	15	1	15	19	30	1	3	0	15	14	20	0	3,3
3	16	0	5	4	40	2	3	1	10	9	50	1	3,3
3	17	1	5	4	50	2	3	0	5	4	50	1	3,3
3	18	1	10	19	50	2	3,3	0	5	19	50	1	3
4	19	1	15	14	20	2	3,3	1	20	19	20	0	3
4	20	1	15	4	50	1	3,3	0	15	4	40	0	3
4	21	0	15	9	30	0	3	1	20	4	20	2	3,3
4	22	1	10	9	30	0	3,3	1	10	9	30	2	3
4	23	0	10	14	40	2	3	0	10	19	50	0	3,3
4	24	0	10	19	40	0	3,3	1	5	14	40	2	3

Tabela 20 – Desenho eficiente para segundo experimento piloto

**(d) Resultados da coleta de dados**

A partir das 210 respostas obtidas (6 situações de escolha respondidas por cada um dos 35 entrevistados) foram estimadas dezenas de modelos diferentes. Para aqueles cujos coeficientes tiveram os sinais esperados e foram estatisticamente significativos, foram calculadas as taxas marginais de substituição entre cada atributo e o *tempo de viagem*. Os valores médio, mínimo e máximo para cada taxa são exibidos na Tabela 21 abaixo:

<b>Relações</b>	<b>Média</b>	<b>Mínimo</b>	<b>Máximo</b>
$\beta_{\text{display}} / \beta_{\text{tviagem}}$	-2,81	-1,38	-3,96
$\beta_{\text{interv}} / \beta_{\text{tviagem}}$	1,42	1,19	1,67
$\beta_{\text{atraso}} / \beta_{\text{tviagem}}$	1,28	1,08	1,50
$\beta_{\text{ocup0}} / \beta_{\text{tviagem}}$	-21,86	-24,41	-17,65
$\beta_{\text{ocup1}} / \beta_{\text{tviagem}}$	-13,91	-15,75	-12,44
$\beta_{\text{custo}} / \beta_{\text{tviagem}}$	20,45	17,65	22,63

**Tabela 21 – Taxas marginais de substituição a partir do segundo experimento piloto**


Ainda para estes mesmos modelos, o valor do coeficiente do *tempo de viagem* variou entre -0,125 e -0,081.

## APÊNDICE C – DESENHO EFICIENTE DO EXPERIMENTO

Bloco	Situação de escolha	Ponto 1 + Linha 1						Ponto 2 + Linha 2					
		Painel	Intervalo	Atraso	Tempo de viagem	Ocupação	Custo	Painel	Intervalo	Atraso	Tempo de viagem	Ocupação	Custo
1	1	0	10	19	20	0	3,3	1	5	14	30	1	3
1	2	1	5	14	30	2	3	0	10	19	20	1	3,3
1	3	1	10	4	30	2	3	0	5	9	30	1	3,3
1	4	1	20	14	20	1	3,3	0	15	19	30	0	3
1	5	0	5	19	50	1	3,3	1	10	14	40	2	3
1	6	0	15	14	40	0	3,3	1	15	14	40	2	3
1	7	1	15	9	30	0	3,3	0	20	4	20	1	3
1	8	0	20	4	50	2	3	1	20	4	50	0	3,3
2	9	1	15	19	20	1	3	0	20	19	20	0	3,3
2	10	0	5	9	50	0	3	1	5	9	40	2	3,3
2	11	0	10	14	30	2	3	1	10	14	30	0	3,3
2	12	0	20	4	30	0	3,3	1	15	9	20	2	3
2	13	0	15	9	40	1	3	1	15	4	50	0	3,3
2	14	1	10	9	50	1	3,3	0	5	9	40	2	3
2	15	1	10	9	50	2	3	0	10	4	50	2	3,3
2	16	1	20	9	40	1	3	0	20	9	50	0	3,3
3	17	0	20	4	20	0	3,3	1	20	4	20	2	3
3	18	0	15	4	40	2	3	1	15	4	40	1	3,3
3	19	1	10	19	20	1	3,3	0	10	19	20	0	3
3	20	1	20	19	50	0	3	0	15	14	50	1	3,3
3	21	1	15	19	20	2	3,3	0	20	19	30	1	3
3	22	0	5	14	40	2	3,3	1	10	19	50	1	3
3	23	1	5	14	30	0	3,3	0	5	14	30	2	3
3	24	0	5	4	40	1	3	1	5	9	40	0	3,3

**Tabela 22 – Situações de escolha do desenho do experimento**

## APÊNDICE D – TELA INICIAL DO QUESTIONÁRIO DE PESQUISA



Escola Politécnica da Universidade de São Paulo  
Departamento de Engenharia de Transportes

O Departamento de Engenharia de Transportes da Escola Politécnica da Universidade de São Paulo agradece sua participação na pesquisa sobre os painéis eletrônicos em pontos de ônibus.

A pesquisa é composta por 8 questões, seguidas de algumas perguntas socioeconômicas. Não levará mais do que 10 minutos para ser preenchida.

As respostas são anônimas e serão utilizadas apenas para esta pesquisa.

Se desejar saber mais sobre este trabalho ou tiver dúvidas e comentários, escreva para [elaine.carvalho@usp.br](mailto:elaine.carvalho@usp.br).

**Figura 10 – Tela inicial do questionário de pesquisa**

## APÊNDICE E – INFLUÊNCIA DO COMPORTAMENTO LEXICOGRÁFICO

### (a) Modelo base sem os 204 lexicográficos quanto ao conforto

Conforme se verifica na Tabela 23, todos os parâmetros deste modelo são diferentes de zero ao nível de significância de 5% e têm os sinais esperados. Em comparação com o modelo base (cujos resultados estão na coluna à direita da tabela), observa-se que houve queda de  $\beta_{ocup0}$  e discreta diminuição de  $\beta_{ocup1}$ . Pode-se supor que os lexicográficos quanto ao conforto sempre escolham a alternativa que lhes permita viajar sentados, porém utilizam regras compensatórias quando nenhuma das alternativas tem veículo com ocupação mínima. Os coeficientes dos demais atributos apresentam aumento do valor absoluto. Estas variações (em relação ao modelo base) são coerentes com o fato de se retirar da amostra os indivíduos que aparentemente valorizam apenas o conforto em sua decisão em detrimento dos demais atributos.

Conforme esperado, houve queda da disposição a pagar para viajar com mais conforto (ou seja, em veículo com menor ocupação), explicada conjuntamente pela redução dos coeficientes  $\beta_{ocup0}$  e  $\beta_{ocup1}$  e pelo aumento em valor absoluto do coeficiente  $\beta_{custo}$ .

Acredita-se que as seguintes características socioeconômicas e de uso de ônibus deste grupo de usuários possam estar relacionadas ao seu comportamento lexicográfico quanto ao conforto:

- maior concentração de não usuários de ônibus (10% contra 5% para a amostra completa);
- maior proporção de entrevistados com idade igual ou superior a 36 anos (23% contra 13% para a amostra completa);
- maior concentração de entrevistados do sexo feminino (46% contra 41% para a amostra completa);
- maior proporção de entrevistados nas faixas de renda domiciliar acima de R\$7.000 (50% contra 44% para a amostra completa).

Na Tabela 32 do Apêndice I pode ser encontrada a classificação socioeconômica e relativa ao uso do ônibus para os três grupos de lexicográficos.

Parâmetro	Valor	Erro padrão	Teste t	Modelo Base
$\beta_{display}$	0,365	0,028	13,290	0,328
$\beta_{interv}$	-0,200	0,008	-24,360	-0,165
$\beta_{atraso}$	-0,109	0,008	-13,780	-0,068
$\beta_{tviagem}$	-0,078	0,005	-17,180	-0,066
$\beta_{ocup0}$	2,550	0,074	34,390	2,750
$\beta_{ocup1}$	1,700	0,053	31,810	1,720
$\beta_{custo}$	-3,110	0,115	-27,020	-2,750
$\sigma_{panel}$	0,215	0,079	2,730	-0,302
Número de observações: 7800				9432
Número de indivíduos: 975				1179
$\log \mathcal{L}(0) = -5406,55$				-6537,76
$\log \mathcal{L}(\beta) = -4430,16$				-5219,88
$-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] = 1952,78$				2635,77
$\rho^2 = 0,181$				0,202
$\rho^2 \text{ ajustado} = 0,179$				0,200

**Tabela 23 – Modelo base sem os lexicográficos quanto ao conforto**

#### (b) Modelo base sem os 68 lexicográficos quanto ao custo

Conforme mostra a Tabela 24, todos os parâmetros deste modelo são diferentes de zero ao nível de significância de 5% e têm os sinais esperados. Em comparação com o modelo base (cujos resultados estão na coluna à direita da tabela) observa-se que houve queda do valor absoluto do coeficiente do custo ( $\beta_{custo}$ ) e aumento do valor absoluto dos demais coeficientes, o que é coerente com o fato de se retirar da amostra os indivíduos que aparentemente usam apenas o valor da tarifa em sua decisão em detrimento dos demais atributos. Como consequência, a disposição a pagar aumentou para todos os atributos.

Sugere-se que as seguintes características socioeconômicas e de uso de ônibus deste grupo estejam relacionadas ao seu comportamento lexicográfico quanto ao custo:

- maior concentração de usuários com frequência semanal de uso de ônibus (97% contra 80% para a amostra completa);

- maior proporção de entrevistados que se dedicam exclusivamente ao estudo (62% contra 42% para a amostra completa);
- maior proporção de entrevistados sem automóvel no domicílio (35% contra 24% para a amostra completa);
- maior proporção de entrevistados nas faixas de renda domiciliar abaixo de R\$7.000 (78% contra 55% para a amostra completa).

Parâmetro	Valor	Erro padrão	Teste t	Modelo Base
$\beta_{\text{display}}$	0,385	0,027	14,130	0,328
$\beta_{\text{interv}}$	-0,178	0,008	-22,370	-0,165
$\beta_{\text{atraso}}$	-0,089	0,008	-11,510	-0,068
$\beta_{\text{tviagem}}$	-0,074	0,005	-16,240	-0,066
$\beta_{\text{ocup0}}$	3,080	0,081	37,840	2,750
$\beta_{\text{ocup1}}$	1,910	0,055	34,390	1,720
$\beta_{\text{custo}}$	-2,490	0,109	-22,790	-2,750
$\sigma_{\text{panel}}$	0,226	0,076	2,970	-0,302
Número de observações: 8888				9432
Número de indivíduos: 1111				1179
$\log \mathcal{L}(0) = -6160,69$				-6537,76
$\log \mathcal{L}(\beta) = -4725,88$				-5219,88
$-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] = 2869,63$				2635,77
$\rho^2 = 0,233$				0,202
$\rho^2 \text{ ajustado} = 0,232$				0,200

**Tabela 24 – Modelo base sem os lexicográficos quanto ao custo**

### (c) Modelo base sem os 14 lexicográficos quanto ao painel

Para este modelo, assim como para os anteriores, todos os parâmetros são diferentes de zero ao nível de significância de 5% e têm os sinais esperados (conforme mostra a Tabela 25). Em comparação com o modelo base, observa-se que houve queda do coeficiente do painel ( $\beta_{\text{display}}$ ), o que é condizente com uma amostra que não contenha indivíduos que, aparentemente, sempre escolhem a alternativa com painel. Os demais coeficientes oscilaram discretamente, geralmente com acréscimo do valor absoluto.



No que se refere às características socioeconômicas e de uso de ônibus deste grupo, não é possível relacioná-las ao comportamento lexicográfico dado o pequeno tamanho do grupo (apenas 14 entrevistados).

<b>Parâmetro</b>	<b>Valor</b>	<b>Erro padrão</b>	<b>Teste t</b>	<b>Modelo Base</b>
$\beta_{\text{display}}$	0,300	0,026	11,680	0,328
$\beta_{\text{interv}}$	-0,170	0,008	-22,210	-0,165
$\beta_{\text{atraso}}$	-0,067	0,007	-9,060	-0,068
$\beta_{\text{tviagem}}$	-0,067	0,004	-15,690	-0,066
$\beta_{\text{ocup0}}$	2,810	0,074	38,260	2,750
$\beta_{\text{ocup1}}$	1,760	0,051	34,500	1,720
$\beta_{\text{custo}}$	-2,800	0,108	-25,860	-2,750
$\sigma_{\text{panel}}$	-0,321	0,056	-5,700	-0,302
Número de observações: 9320				9432
Número de indivíduos: 1165				1179
$\log \mathcal{L}(0) = -6460,13$				-6537,76
$\log \mathcal{L}(\beta) = -5122,27$				-5219,88
$-2 \cdot [ \log \mathcal{L}(0) - \log \mathcal{L}(\beta) ] = 2675,73$				2635,77
$\rho^2 = 0,207$				0,202
$\rho^2 \text{ ajustado} = 0,206$				0,200

**Tabela 25 – Modelo base sem os lexicográficos quanto ao painel**

## APÊNDICE F – MODELOS COM INTERAÇÕES ENTRE ATRIBUTOS

Parâmetro	Valor	Erro padrão	Teste t
$\beta_{\text{interv}}$	-0,158	0,008	-20,710
$\beta_{\text{atraso}}$	-0,071	0,008	-9,200
$\beta_{\text{atraso\_display}}$	0,017	0,002	7,760
$\beta_{\text{tviagem}}$	-0,081	0,006	-13,100
$\beta_{\text{tviagem\_ocup0}}$	0,025	0,005	5,390
$\beta_{\text{tviagem\_ocup1}}$	0,014	0,004	3,170
$\beta_{\text{ocup0}}$	1,890	0,160	11,790
$\beta_{\text{ocup1}}$	1,230	0,149	8,220
$\beta_{\text{custo}}$	-2,730	0,109	-25,110
$\sigma_{\text{panel}}$	-0,292	0,058	-5,070

Número de observações: 9432

Número de indivíduos: 1179

$\log \mathcal{L}(0) = -6537,76$

$\log \mathcal{L}(\beta) = -5230,17$

$-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] = 2615,18$

$\rho^2 = 0,200$

$\rho^2$  ajustado = 0,198

**Tabela 26 – Modelo ML *panel* com interações entre atributos (1)**

Parâmetro	Valor	Erro padrão	Teste t
$\beta_{\text{interv}}$	-0,166	0,008	-21,910
$\beta_{\text{interv\_display}}$	0,022	0,002	10,710
$\beta_{\text{atraso}}$	-0,069	0,007	-9,240
$\beta_{\text{tviagem}}$	-0,063	0,004	-14,620
$\beta_{\text{tviagem\_ocup0}}$	0,010	0,003	3,330
$\beta_{\text{ocup0}}$	2,390	0,124	19,290
$\beta_{\text{ocup1}}$	1,700	0,050	34,350
$\beta_{\text{custo}}$	-2,760	0,108	-25,470
$\sigma_{\text{panel}}$	-0,305	0,057	-5,400

Número de observações: 9432

Número de indivíduos: 1179

$\log \mathcal{L}(0) = -6537,76$

$\log \mathcal{L}(\beta) = -5212,36$

$-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] = 2650,82$

$\rho^2 = 0,203$

$\rho^2$  ajustado = 0,201

**Tabela 27 – Modelo ML *panel* com interações entre atributos (2)**

Parâmetro	Valor	Erro padrão	Teste t
$\beta_{display}$	0,304	0,027	11,130
$\beta_{interv}$	-0,171	0,008	-21,860
$\beta_{atraso\_interv\{5,10\}}$	-0,044	0,010	-4,480
$\beta_{atraso\_interv\{15,20\}}$	-0,097	0,012	-8,240
$\beta_{tviagem}$	-0,070	0,004	-16,020
$\beta_{tviagem\_ocup0}$	0,011	0,003	3,680
$\beta_{ocup0}$	2,380	0,123	19,380
$\beta_{ocup1}$	1,700	0,051	33,420
$\beta_{custo}$	-2,880	0,113	-25,350
$\sigma_{panel}$	-0,303	0,057	-5,310

Número de observações: 9432  
Número de indivíduos: 1179  
 $\log \mathcal{L}(0) = -6537,76$   
 $\log \mathcal{L}(\beta) = -5205,60$   
 $-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] = 2664,32$   
 $\rho^2 = 0,204$   
 $\rho^2 \text{ ajustado} = 0,202$

**Tabela 28 – Modelo ML *panel* com interações entre atributos (3)**

## APÊNDICE G – VARIÁVEL ALEATÓRIA DAS PREFERÊNCIAS

### (a) Modelo ML *panel random coefficients* com distribuição normal de probabilidades

O modelo estimado é apresentado na Tabela 29 e todos os seus parâmetros são diferentes de zero ao nível de significância de 5% e têm os sinais esperados. Verifica-se que  $\beta_{tviagem}$  tem média ( $\mu$ ) de -0,062 (valor semelhante ao encontrado para o coeficiente no modelo base) e desvio padrão ( $\sigma$ ) igual a 0,138, o que significa que o coeficiente do tempo de viagem será negativo com probabilidade de 67,36%. Já  $\beta_{ocup0}$  assume média de 3,070 (superior ao valor do coeficiente no modelo base) e desvio padrão de -0,568, o que significa que este coeficiente é sempre positivo.

Em comparação com o modelo base, houve aumento nos valores absolutos dos demais coeficientes e o rho quadrado ajustado sofreu elevação discreta, indicando que o modelo em discussão é um pouco superior ao modelo base quanto à capacidade de reproduzir o comportamento de escolha dos entrevistados.

Parâmetro	Valor	Erro padrão	Teste t	Modelo Base
$\beta_{display}$	0,330	0,032	10,240	0,328
$\beta_{interv}$	-0,185	0,012	-15,850	-0,165
$\beta_{atraso}$	-0,070	0,010	-6,840	-0,068
$\beta_{tviagem} (\mu)$	-0,062	0,007	-9,270	-0,066
$\beta_{tviagem} (\sigma)$	0,138	0,022	6,150	
$\beta_{ocup0} (\mu)$	3,070	0,147	20,830	2,750
$\beta_{ocup0} (\sigma)$	-0,568	0,234	-2,430	
$\beta_{ocup1}$	1,930	0,091	21,190	1,720
$\beta_{custo}$	-3,330	0,210	-15,870	-2,750
$\sigma_{panel}$	-0,353	0,069	-5,080	-0,302
Número de observações:	9432			9432
Número de indivíduos:	1179			1179
$\log \mathcal{L} (0)$	= -6537,76			-6537,76
$\log \mathcal{L} (\beta)$	= -5210,58			-5219,88
$-2 \cdot [ \log \mathcal{L} (0) - \log \mathcal{L} (\beta) ]$	= 2654,36			2635,77
$\rho^2$	= 0,203			0,202
$\rho^2$ ajustado	= 0,201			0,200

Tabela 29 – Modelo ML *panel random coefficients* com distribuição normal

**(b) Modelo ML *panel random coefficients* com distribuição uniforme de probabilidades**

Todos os parâmetros do modelo estimado (apresentado na Tabela 30) são diferentes de zero ao nível de significância de 5% e têm os sinais esperados. Verifica-se que  $\beta_{\text{interv}}$  tem média ( $\mu$ ) igual -0,183 (superior ao valor do coeficiente no modelo base) e varia com distribuição uniforme dentro do intervalo de  $\mu - s$  (-0,483) a  $\mu + s$  (0,117), o que significa que a probabilidade deste coeficiente ter sinal negativo, conforme se espera, é de 80,5%. Já  $\beta_{\text{tviagem}}$  tem média de -0,055 (inferior ao valor do coeficiente no modelo base) e varia com distribuição uniforme dentro do intervalo de -0,282 a 0,172, o que significa que há uma probabilidade de 62,1% (não muito alta, porém razoável) de que este coeficiente tenha sinal negativo, conforme se espera.

Em comparação com o modelo base, houve alteração nos valores dos coeficientes dos demais atributos, e o rho quadrado ajustado teve discreto aumento, indicando que a especificação do modelo em discussão é superior à do modelo base (e também à do primeiro modelo discutido neste apêndice) quanto à capacidade de reproduzir o comportamento de escolha dos entrevistados.

Parâmetro	Valor	Erro padrão	Teste t	Modelo Base
$\beta_{\text{display}}$	0,326	0,031	10,480	0,328
$\beta_{\text{interv}} (\mu)$	-0,183	0,011	-17,440	-0,165
$\beta_{\text{interv}} (s)$	-0,300	0,100	-3,010	
$\beta_{\text{atraso}}$	-0,058	0,010	-5,930	-0,068
$\beta_{\text{tviagem}} (\mu)$	-0,055	0,006	-8,960	-0,066
$\beta_{\text{tviagem}} (s)$	0,227	0,039	5,860	
$\beta_{\text{ocup0}}$	2,990	0,094	31,890	2,750
$\beta_{\text{ocup1}}$	1,900	0,070	27,240	1,720
$\beta_{\text{custo}}$	-3,330	0,185	-18,030	-2,750
$\sigma_{\text{panel}}$	-0,381	0,068	-5,630	-0,302
Número de observações:	9432			9432
Número de indivíduos:	1179			1179
$\log \mathcal{L}(0)$	-6537,76			-6537,76
$\log \mathcal{L}(\beta)$	-5206,35			-5219,88
$-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)]$	2662,84			2635,77
$\rho^2$	0,204			0,202
$\rho^2$ ajustado	0,202			0,200

**Tabela 30 – Modelo ML *panel random coefficients* com distribuição uniforme**

## APÊNDICE H – MODELO ML PANEL PARA NÃO USUÁRIOS

<b>Coefficiente</b>	<b>Valor</b>	<b>Erro padrão</b>	<b>Teste t</b>
$\beta_{\text{display}}$	0,349	0,133	2,630
$\beta_{\text{interv}}$	-0,198	0,039	-5,140
$\beta_{\text{atraso}}$	-0,006	0,037	-0,150
$\beta_{\text{tviagem}}$	-0,093	0,023	-4,100
$\beta_{\text{ocup0}}$	3,480	0,419	8,320
$\beta_{\text{ocup1}}$	2,160	0,280	7,710
$\beta_{\text{custo}}$	-1,950	0,526	-3,700
$\sigma_{\text{panel}}$	0,302	0,301	1,000

Número de observações: 472  
Número de indivíduos: 59  
 $\log \mathcal{L}(0) = -327,16$   
 $\log \mathcal{L}(\beta) = -220,48$   
 $-2 \cdot [\log \mathcal{L}(0) - \log \mathcal{L}(\beta)] = 213,67$   
 $\rho^2 = 0,326$   
 $\rho^2$  ajustado = 0,302

**Tabela 31 – Modelo ML *panel* estimado a partir das respostas dos não usuários**

## APÊNDICE I – CLASSIFICAÇÃO SOCIOECONÔMICA E DE USO DE ÔNIBUS

	Toda a amostra	Lexic. quanto ao conforto	Lexic. quanto ao custo	Lexic. quanto ao painel
<b>Frequência de uso do ônibus</b>				
Nunca	5%	10%	3%	14%
1 a 2 x mês	15%	15%	0%	21%
1 a 2 x semana	15%	15%	16%	0%
3 ou mais x semana	65%	60%	81%	64%
<b>Sexo</b>				
Feminino	41%	46%	32%	36%
Masculino	59%	54%	68%	64%
<b>Idade</b>				
Até 25 anos	64%	50%	72%	50%
Entre 26 e 35 anos	23%	27%	22%	36%
A partir de 36 anos	13%	23%	6%	14%
<b>Ocupação</b>				
Outra	4%	6%	0%	0%
Estudo	42%	31%	62%	57%
Trabalho / estágio	14%	19%	6%	21%
Estudo e trabalho / estágio	39%	44%	32%	21%
<b>Escolaridade</b>				
Fundamental / Médio *	4%	4%	3%	0%
Superior incompleto	55%	46%	68%	50%
Superior completo	11%	11%	7%	0%
Pós-graduação *	30%	39%	22%	50%
<i>* mesmo que incompleto</i>				
<b>Automóveis no domicílio</b>				
Nenhum	24%	23%	35%	43%
1	36%	41%	34%	14%
2	26%	23%	24%	29%
3	9%	6%	7%	14%
4 ou mais	5%	7%	0%	0%
<b>Renda do domicílio</b>				
Até R\$ 1.400	6%	5%	16%	14%
De R\$ 1.401 a R\$ 2.800	10%	9%	12%	14%
De R\$ 2.801 a R\$ 4.900	20%	25%	15%	21%
De R\$ 4.901 a R\$ 7.000	19%	12%	35%	14%
De R\$ 7.001 a R\$ 10.500	19%	21%	9%	14%
De R\$ 10.501 a R\$ 14.000	11%	12%	7%	7%
Mais de R\$ 14.000	15%	17%	6%	14%

**Tabela 32 – Classificação socioeconômica e quanto ao uso de ônibus dos lexicógrafos**

	Toda a amostra	Fund. / Médio*	Superior incomp.	Superior comp.	Pós-grad.*
<b>Frequência de uso do ônibus</b>					
Nunca	5,0%	0,0%	3,4%	7,6%	7,6%
1 a 2 x mês	15,0%	6,1%	11,2%	19,1%	21,8%
1 a 2 x semana	14,6%	10,2%	12,7%	11,5%	19,8%
3 ou mais x semana	65,4%	83,7%	72,7%	61,8%	50,8%
<b>Presença de painel nos pontos</b>					
Não se aplica	5,0%	0,0%	3,4%	7,6%	7,6%
Não	83,6%	89,8%	87,3%	76,3%	78,8%
Sim	11,4%	10,2%	9,3%	16,0%	13,6%
<b>Sexo</b>					
Feminino	40,5%	34,7%	34,6%	47,3%	49,7%
Masculino	59,5%	65,3%	65,4%	52,7%	50,3%
<b>Idade</b>					
Até 25 anos	63,7%	81,6%	92,6%	40,5%	17,2%
Entre 26 e 35 anos	23,2%	4,1%	6,5%	35,1%	52,0%
A partir de 36 anos	13,1%	14,3%	0,9%	24,4%	30,8%
<b>Ocupação</b>					
Outra	4,5%	6,1%	0,2%	12,2%	9,3%
Estudo	42,5%	75,5%	57,7%	5,3%	24,0%
Trabalho / estágio	14,4%	10,2%	2,3%	48,9%	24,3%
Estudo e trabalho / estágio	38,6%	8,2%	39,8%	33,6%	42,4%
<b>Automóveis no domicílio</b>					
Nenhum	24,3%	20,4%	21,9%	26,7%	28,2%
1	35,9%	44,9%	31,9%	42,0%	39,5%
2	26,1%	22,4%	29,1%	20,6%	23,2%
3	9,0%	8,2%	11,9%	3,8%	5,6%
4 ou mais	4,7%	4,1%	5,1%	6,9%	3,4%
<b>Motos no domicílio</b>					
Nenhuma	93,8%	95,9%	93,8%	91,6%	94,4%
1	5,5%	4,1%	5,3%	8,4%	5,1%
2	0,7%	0,0%	0,9%	0,0%	0,6%
<b>Renda do domicílio</b>					
Até R\$ 1.400	6,1%	10,2%	7,3%	3,1%	4,5%
De R\$ 1.401 a R\$ 2.800	10,3%	8,2%	12,4%	5,3%	8,8%
De R\$ 2.801 a R\$ 4.900	20,4%	28,6%	19,8%	28,2%	17,2%
De R\$ 4.901 a R\$ 7.000	18,9%	18,4%	19,1%	22,1%	17,5%
De R\$ 7.001 a R\$ 10.500	18,6%	12,2%	16,4%	16,8%	24,0%
De R\$ 10.501 a R\$ 14.000	10,7%	4,1%	11,2%	9,9%	11,0%
Mais de R\$ 14.000	15,0%	18,4%	13,8%	14,5%	16,9%

Tabela 33 – Classificação socioeconômica e de uso de ônibus por *escolaridade* (\*mesmo que incompleto)



	<b>Toda a amostra</b>	<b>Não usuários</b>	<b>1 a 2 x mês</b>	<b>1 a 2 x semana</b>	<b>3 ou mais x semana</b>
<b>Painel nos pontos</b>					
Não se aplica	5,0%	100,0%	0,0%	0,0%	0,0%
Não	83,6%	0,0%	86,4%	90,7%	87,8%
Sim	11,4%	0,0%	13,6%	9,3%	12,2%
<b>Sexo</b>					
Feminino	40,5%	47,5%	39,5%	38,4%	40,7%
Masculino	59,5%	52,5%	60,5%	61,6%	59,3%
<b>Idade</b>					
Até 25 anos	63,7%	42,4%	49,2%	55,2%	62,2%
Entre 26 e 35 anos	23,2%	27,1%	25,4%	25,0%	31,3%
A partir de 36 anos	13,1%	30,5%	25,4%	19,8%	6,5%
<b>Ocupação</b>					
Outra	4,5%	5,1%	7,3%	8,1%	3,0%
Estudo	42,5%	23,7%	33,9%	43,6%	45,7%
Trabalho / estágio	14,4%	35,6%	23,7%	15,1%	10,5%
Estudo e trabalho / estágio	38,6%	35,6%	35,0%	33,1%	40,9%
<b>Escolaridade</b>					
Fundamental / Médio *	4,2%	0,0%	1,7%	2,9%	5,3%
Superior incompleto	54,7%	37,3%	40,7%	47,7%	60,8%
Superior completo	11,1%	16,9%	14,1%	8,7%	10,5%
Pós-graduação *	30,0%	45,8%	43,5%	40,7%	23,3%
<i>* mesmo que incompleto</i>					
<b>Automóveis no domicílio</b>					
Nenhum	24,3%	3,4%	5,1%	21,5%	30,9%
1	35,9%	16,9%	31,1%	40,1%	37,5%
2	26,1%	39,0%	32,8%	24,4%	24,0%
3	9,0%	22,0%	20,3%	9,3%	5,3%
4 ou mais	4,7%	18,6%	10,7%	4,7%	2,3%
<b>Motos no domicílio</b>					
Nenhuma	93,8%	93,2%	91,0%	94,2%	94,4%
1	5,5%	6,8%	7,3%	4,7%	5,2%
2	0,7%	0,0%	1,7%	1,2%	0,4%
<b>Renda do domicílio</b>					
Até R\$ 1.400	6,1%	0,0%	2,8%	5,8%	7,4%
De R\$ 1.401 a R\$ 2.800	10,3%	5,1%	5,1%	9,3%	12,2%
De R\$ 2.801 a R\$ 4.900	20,4%	10,2%	13,0%	15,7%	23,9%
De R\$ 4.901 a R\$ 7.000	18,9%	6,8%	15,3%	18,6%	20,8%
De R\$ 7.001 a R\$ 10.500	18,6%	11,9%	22,0%	21,5%	17,6%
De R\$ 10.501 a R\$ 14.000	10,7%	22,0%	10,2%	14,5%	9,1%
Mais de R\$ 14.000	15,0%	44,1%	31,6%	14,5%	9,1%

Tabela 34 – Classificação socioeconômica por frequência de uso de ônibus

	Todos	Outra	Estudo	Trabalho / estágio	Estudo e trabalho / estágio
<b>Freq. de uso do ônibus</b>					
Nunca	5,0%	5,7%	2,8%	12,4%	4,6%
1 a 2 x mês	15,0%	24,5%	12,0%	24,7%	13,6%
1 a 2 x semana	14,6%	26,4%	15,0%	15,3%	12,5%
3 ou mais x semana	65,4%	43,4%	70,3%	47,6%	69,2%
<b>Painel nos pontos</b>					
Não se aplica	5,0%	5,7%	2,8%	12,4%	4,6%
Não	83,6%	84,9%	88,2%	73,5%	82,2%
Sim	11,4%	9,4%	9,0%	14,1%	13,2%
<b>Sexo</b>					
Feminino	40,5%	60,4%	32,7%	52,9%	42,2%
Masculino	59,5%	39,6%	67,3%	47,1%	57,8%
<b>Idade</b>					
Até 25 anos	63,7%	9,4%	86,2%	18,8%	62,0%
Entre 26 e 35 anos	23,2%	22,6%	12,2%	40,6%	29,0%
A partir de 36 anos	13,1%	67,9%	1,6%	40,6%	9,0%
<b>Escolaridade</b>					
Fundamental / Médio *	4,2%	5,7%	7,4%	2,9%	0,9%
Superior incompleto	54,7%	1,9%	74,3%	8,8%	56,5%
Superior completo	11,1%	30,2%	1,4%	37,6%	9,7%
Pós-graduação *	30,0%	62,3%	17,0%	50,6%	33,0%
<i>* mesmo que incompleto</i>					
<b>Automóveis no domicílio</b>					
Nenhum	24,3%	13,2%	22,8%	24,7%	27,0%
1	35,9%	49,1%	32,7%	38,8%	36,7%
2	26,1%	32,1%	29,3%	25,9%	22,0%
3	9,0%	3,8%	10,4%	4,7%	9,7%
4 ou mais	4,7%	1,9%	4,8%	5,9%	4,6%
<b>Motos no domicílio</b>					
Nenhuma	93,8%	96,2%	94,8%	92,4%	93,0%
1	5,5%	3,8%	4,2%	6,5%	6,8%
2	0,7%	0,0%	1,0%	1,2%	0,2%
<b>Renda do domicílio</b>					
Até R\$ 1.400	6,1%	5,7%	9,6%	0,0%	4,6%
De R\$ 1.401 a R\$ 2.800	10,3%	5,7%	11,0%	2,9%	13,0%
De R\$ 2.801 a R\$ 4.900	20,4%	15,1%	20,6%	20,6%	20,7%
De R\$ 4.901 a R\$ 7.000	18,9%	18,9%	18,0%	14,7%	21,5%
De R\$ 7.001 a R\$ 10.500	18,6%	15,1%	15,4%	24,7%	20,2%
De R\$ 10.501 a R\$ 14.000	10,7%	18,9%	10,4%	12,9%	9,2%
Mais de R\$ 14.000	15,0%	20,8%	15,2%	24,1%	10,8%

Tabela 35 – Classificação socioeconômica e de uso de ônibus por ocupação

	<b>Toda a amostra</b>	<b>Até 25 anos</b>	<b>Entre 26 e 35 anos</b>	<b>A partir de 36 anos</b>
<b>Freq. de uso do ônibus</b>				
Nunca	5,0%	3,3%	5,8%	11,7%
1 a 2 x mês	15,0%	11,6%	16,4%	29,2%
1 a 2 x semana	14,6%	12,6%	15,7%	22,1%
3 ou mais x semana	65,4%	72,4%	62,0%	37,0%
<b>Painel nos pontos</b>				
Não se aplica	5,0%	3,3%	5,8%	11,7%
Não	83,6%	86,6%	80,7%	74,7%
Sim	11,4%	10,1%	13,5%	13,6%
<b>Sexo</b>				
Feminino	40,5%	34,9%	50,0%	51,3%
Masculino	59,5%	65,1%	50,0%	48,7%
<b>Ocupação</b>				
Outra	4,5%	0,7%	4,4%	23,4%
Estudo	42,5%	57,5%	22,3%	5,2%
Trabalho / estágio	14,4%	4,3%	25,2%	44,8%
Estudo e trabalho / estágio	38,6%	37,5%	48,2%	26,6%
<b>Escolaridade</b>				
Fundamental / Médio *	4,2%	5,3%	0,7%	4,5%
Superior incompleto	54,7%	79,5%	15,3%	3,9%
Superior completo	11,1%	7,1%	16,8%	20,8%
Pós-graduação *	30,0%	8,1%	67,2%	70,8%
<i>* mesmo que incompleto</i>				
<b>Automóveis no domicílio</b>				
Nenhum	24,3%	20,8%	38,3%	16,2%
1	35,9%	33,3%	37,2%	46,1%
2	26,1%	28,5%	16,8%	31,2%
3	9,0%	12,0%	4,0%	3,2%
4 ou mais	4,7%	5,5%	3,6%	3,2%
<b>Motos no domicílio</b>				
Nenhuma	93,8%	94,3%	93,1%	92,9%
1	5,5%	4,9%	6,6%	6,5%
2	0,7%	0,8%	0,4%	0,6%
<b>Renda do domicílio</b>				
Até R\$ 1.400	6,1%	6,8%	6,6%	1,9%
De R\$ 1.401 a R\$ 2.800	10,3%	11,5%	9,9%	5,8%
De R\$ 2.801 a R\$ 4.900	20,4%	21,7%	20,8%	13,0%
De R\$ 4.901 a R\$ 7.000	18,9%	19,3%	21,2%	13,0%
De R\$ 7.001 a R\$ 10.500	18,6%	15,4%	22,3%	27,3%
De R\$ 10.501 a R\$ 14.000	10,7%	10,8%	9,9%	11,7%
Mais de R\$ 14.000	15,0%	14,5%	9,5%	27,3%

Tabela 36 – Classificação socioeconômica e de uso de ônibus por idade