GABRIEL PEREIRA CALDEIRA

DEPENDÊNCIA ESPACIAL SOB A ABORDAGEM DE MINIMIZAÇÃO DO
ARREPENDIMENTO EM MODELOS DE ESCOLHA DE MODO DE TRANSPORTE

SÃO PAULO

2022

GABRIEL PEREIRA CALDEIRA

DEPENDÊNCIA ESPACIAL SOB A ABORDAGEM DE MINIMIZAÇÃO DO
ARREPENDIMENTO EM MODELOS DE ESCOLHA DE MODO DE TRANSPORTE

Versão Corrigida

(Versão original encontra-se na unidade que aloja o Programa de Pós-graduação)

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

Área de concentração: Engenharia de Transportes

Orientador: Prof. Dr. Cassiano Augusto Isler

SÃO PAULO

2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

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, 03 de Fevereiro de 2022

Assinatura do autor: Gabriel P. Caldeira

Assinatura do orientador:

Catálogo-na-publicação

Caldeira, Gabriel Pereira

Dependência espacial sob a abordagem de minimização do arrependimento em modelos de escolha de modo de transporte / G. P. Caldeira – versão corr. – São Paulo, 2022.
143 p.

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

1.Minimização do Arrependimento Aleatório 2.Maximização da Utilidade Aleatória 3.Modelos de escolha discreta 4.Econometria Espacial
5. Dependência Espacial I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Transportes II.t.

À ciência brasileira que é vítima constante do obscurantismo que por hora nos assola

AGRADECIMENTOS

Talvez seja egoísmo, talvez seja orgulho, não sei, mas gostaria de começar a agradecer a mim mesmo por ter chegado até esse momento. Obviamente não cheguei aqui sozinho, e por isso quero agradecer a minha família que sempre torceu pelo meu sucesso, em especial às pessoas mais importantes da minha vida, Mami, Dri e Lu. Amo vocês! E só estou aqui porque tenho vocês comigo.

Obrigado ao meu orientador Cassiano por ser o melhor orientador que eu poderia ter e por me aceitar quando achei que não teria mais lugar nessa Universidade. Obrigado por me incentivar a sempre melhorar e por sempre me puxar para frente. Obrigado por apoiar os meus anseios de pesquisa, mesmo não sabendo onde chegaríamos com as ideias malucas.

Gostaria de agradecer ao Prof. Cláudio Marte por ser a primeira pessoa a me receber na Poli e abrir as portas da universidade para mim. Obrigado por ter um coração tão bom e gentil. Serei sempre grato pelo carinho e entendimento. Aos demais professores da Poli-USP que me inspiram constantemente: Orlando Strambi, Mariana Giannotti e Cláudio Barbieri, obrigado! Foi ótimo poder aprender com vocês e levarei esses aprendizados para sempre. Obrigado aos professores da FEA-USP, em especial ao Prof Fávero que me fez amar estudar econometria.

Obrigado aos meus amigos Lucas e Renata, que dividiram tantos almoços, conversas e cookies (a pandemia nos levou até esses momentos simples). Obrigado por escreverem comigo e me fazerem sorrir presencialmente ou online. Obrigado por me escutarem no caminho do RU falando sobre como eu não conseguiria terminar esse trabalho e depois rir das fofocas da universidade. Obrigado também às pessoas maravilhosas que a Poli me trouxe: Valéria e Raíssa, foi ótimo conviver com vocês, mesmo separados por algumas paredes.

Obrigado às princesinhas do LabGeo Tainá Bitterncourt, Bruna Pizzol e Beatriz Moura. Vocês são mulheres inspiradoras e me fazem acreditar em cidades mais justas e menos desiguais. Um obrigado especial para a Bru por me ajudar em um dos momentos mais difíceis na universidade e por me receber tão bem no Banco Mundial. Vou guardar para sempre nossos cafés da tarde!

Obrigado aos meus amigos de trabalho Beatriz Moura (de novo sim!), Carlos Bellas, Luca di Biase, Maria Inês Lippe e Tais Fonseca. Trabalhar em um dos maiores bancos de desenvolvimento do planeta não teria sido uma experiência tão prazerosa se não fosse por vocês. Obrigado pela paciência e por me ensinarem todo dia algo novo.

Obrigado aos meus amigos que estão comigo há anos, em especial ao meu trio favorito do mundo: Gabi, Ma e Doug. Amo vocês mesmo sendo os piores amigos do mundo. Obrigado Ma por me receber em Floripa no meu último mês do mestrado. Amiga, se não fosse você, não sei se teria conseguido. Te amo infinito e passei a amar o novo agregado também!

Obrigado aos meus amigos que mesmo longe ainda estão no meu coração: Ana, Lucas, João, Gi, Le, Deko e Bruno. Não importa a cidade ou o tempo sem se falar, continuo com vocês no coração. Obrigado também ao meu mozinho (Amanda para quem não conhece), que está sempre comigo e me alimentando bem. Te amo mozi! Obrigado aos meus amigos da EESC: Alceu, Samuel, Marina e Renan por dividirem comigo os melhores momentos na ANPET (saudades da vida presencial). Tenho um carinho e admiração enorme por vocês.

Obrigado às professoras Cira Pitombo e Ana Larrañaga por contribuírem enormemente na minha banca de qualificação para a melhora desta dissertação. Vocês são mulheres inspiradoras dentro da academia e admiro muito o trabalho de vocês.

Obrigado a todos aqueles que não nomeei aqui, mas que me ajudaram a chegar nesse momento, seja no desenvolvimento direto da pesquisa, seja na vida através de conversas e/ou experiências. Obrigado, obrigado, obrigado!

Obrigado ao CNPq pelo financiamento desta pesquisa. Obrigado a universidade pública pela oportunidade de fazer ciência num país tão solapado por forças retrógradas. Obrigado a ciência e aos cientistas pelas vacinas e por salvarem todos ao meu redor num período tão dramático.

Obrigado às forças e energias que me trouxeram até aqui e que me ajudaram a sobreviver a esse momento tão difícil em que vivemos.

Por fim, mas não menos importante, obrigado à minha bebezinha pet Zoe, te amo filha.

“Zones don’t travel, people travel!”

Travel Demand Forecasting Project (1972) apud Jang (2018)

RESUMO

Diferentes abordagens de análise de comportamentos relacionados a viagens têm se apresentado como alternativa à usual maximização da utilidade aleatória (*Random Utility Maximization – RUM*). A abordagem de minimização do arrependimento aleatório (*Random Regret Minimization – RRM*) postula que os indivíduos buscam minimizar seu arrependimento na escolha de uma alternativa quanto às demais disponíveis. Além disso, análises que tomam o espaço geográfico como parte explicadora dos fenômenos sociais têm crescido no âmbito da Engenharia de Transportes. Assim, para além de uma característica dada, o espaço se torna elemento estruturante de padrões observáveis no espaço social e geográfico. Neste contexto, o objetivo principal desta pesquisa é analisar efeitos de dependência espacial em modelos de escolha discreta sob a abordagem *RRM* no contexto de escolhas de modos de transportes em viagens por motivo trabalho na cidade São Paulo, Brasil. Dados de Preferência Revelada oriundos da pesquisa OD2017 do Metrô-SP foram utilizados na modelagem considerando características sociodemográficas dos indivíduos e atributos de nível de serviço. Os modelos foram comparados sob as abordagens *RUM* e *RRM* apenas com atributos de nível de serviço (*LOS*) e com atributos de nível de serviço e sociodemográficos (*LOS + SOCIO*). Além disso, foram estimados modelos com e sem a presença de efeitos de interações espaciais nas variáveis de nível de serviço (efeitos exógenos) pela especificação de dois tipos de matrizes de vizinhança, pelo critério de distância e de *k*-vizinhos mais próximos. Os valores limites de distância foram variados entre 1.000 metros e 2.000 metros em intervalos de 100 metros e os de número de vizinhos entre 10 e 100 em intervalos de 10 vizinhos. Os modelos *LOS* e *LOS + SOCIO* baseados na abordagem *RRM* tiveram melhor desempenho em termos de log-verossimilhança, inclusive com a presença de efeitos espaciais. O modelo com o melhor resultado foi o híbrido *RUM-RRM* com efeitos de interação espacial sob a abordagem *RRM*. Análises de Valor de Tempo de Viagem (Value of Travel Time-VTT) indicaram que os modelos sem efeitos de dependência espacial superestimaram-no para a alternativa automóvel em comparação àqueles que consideram tais efeitos. No caso da abordagem *RUM* esse valor foi superestimado em até 33%, enquanto na abordagem *RRM* a diferença máxima foi de 8%. Os resultados das estimativas de elasticidades mostraram que os atributos relacionados ao transporte público são elásticos, indicando que políticas públicas de transporte com enfoque nesses atributos têm maior potencial de atrair ou afastar a demanda por essas alternativas.

Palavras-chave: Minimização do Arrependimento Aleatório, Maximização da Utilidade Aleatória, Modelos de escolha discreta, Econometria Espacial, Dependência Espacial.

ABSTRACT

Different approaches of travel behaviour analysis have been proposed as alternative to the well-known Random Utility Maximisation approach (RUM). The Random Regret Minimisation (RRM) approach states that individuals try to minimize their regret when choosing among a set of available alternatives. Furthermore, analyses that consider the geographic space as an explanatory part of social phenomena have grown in Transportation Engineering in particular. Thus, beyond a given characteristic, space becomes a structuring element of observable patterns in social and geographical space. In this context, the main goal of this research is to assess spatial dependency effects in a discrete choice model under the *RRM* approach for commuting mode choice in the city of São Paulo, Brazil. Revealed Preference data from the OD 2017 of the Metrô-SP survey were used including the individual sociodemographic characteristics and the level-of-service attributes. Models under the *RUM* and *RRM* approaches were compared solely with level-of-service attributes (*LOS*) and with level-of-service and sociodemographic attributes (*LOS + SOCIO*). In addition, the models were estimated with and without the presence of spatial dependency effect on the *LOS* variables (exogenous interaction effects) given two specifications of the spatial weight matrix based on distance and *k* nearest neighbors criteria. The thresholds for the distance specification varied from 1,000 meters to 2,000 meters by 100 meters, and the number of neighbors varied from 10 to 100 by 10 neighbors. The results indicate that models based on the *RRM* approach outperformed in terms of log-likelihood the *LOS* and *LOS + SOCIO* models, and also those including spatial dependency. The best model was the hybrid *RUM-RRM* model with the presence of spatial exogeneous interaction effects under the *RRM* approach. Results of time value of time (VTT) estimations indicated that the models without spatial dependency effects overestimated the value of this measure for the Car alternative compared to those that controlled spatial dependency in the model. In the case of the *RUM* approach the VTT was overestimated by up to 33%, while in the *RRM* approach this difference was up to 8%. The analysis of elasticities showed that the attributes related to public transportation are elastic, indicating that transport policies that focus on these attributes have greater potential to attract or drive away demand for these alternatives.

Keywords: Random Regret Minimization; Random Utility Maximization; Discrete Choice models; Spatial Econometrics; Spatial Dependency.

LISTA DE FIGURAS

Figura 1 – Parcela do arrependimento do modelo de 2008 e 2010.	30
Figura 2 – Mapa de corredores e faixas exclusivas de ônibus de São Paulo.....	54
Figura 3 – Mapa de linhas de trens suburbanos e metrô de São Paulo.....	56
Figura 4 – Zonas de tráfego da Pesquisa OD 2017.	57
Figura 5 – Processo de tratamento dos dados da OD 2017.	58
Figura 6 – Fluxograma de caracterização de atributos de nível de serviço.....	62
Figura 7 – Escolha das alternativas da amostra de modelagem.	70
Figura 8 – Efeito de dependência espacial do tempo de viagem por ônibus (RUM).	76
Figura 9 – Efeito de dependência espacial do tempo de viagem por ônibus e metrô (RRM). .	78
Figura 10 – Log-Verossimilhança dos modelos <i>LOS</i> para matriz baseada em distância.	84
Figura 11 – Log-Verossimilhança dos modelos <i>LOS</i> para matriz baseada em <i>knn</i>	85
Figura 12 – AIC dos modelos <i>LOS</i> para matriz baseada em distância.	88
Figura 13 – AIC dos modelos <i>LOS</i> para matriz baseada em <i>k</i> vizinhos mais próximos.	89
Figura 14 – Log-Verossimilhança dos modelos <i>LOS + SOCIO</i> para matriz de distância.....	90
Figura 15 – Log-Verossimilhança dos modelos <i>LOS + SOCIO</i> para matriz baseada em <i>knn</i> .	91
Figura 16 – AIC dos modelos <i>LOS + SOCIO</i> para matriz baseada em distância.....	94
Figura 17 – Log-Verossimilhança dos modelos <i>LOS + SOCIO</i> para matriz baseada em <i>knn</i> .	94

LISTA DE TABELAS

Tabela 1 – Estatística descritiva dos atributos utilizados para estimação.	69
Tabela 2 – Resultados gerais dos modelos <i>LOS</i>	89
Tabela 3 – Resultados gerais dos modelos <i>LOS + SOCIO</i>	95
Tabela 4 – Parâmetros estimados dos modelos <i>LOS</i>	95
Tabela 5 – Parâmetros estimados dos modelos <i>LOS + SOCIO</i>	96
Tabela 6 – VTT dos modelos <i>LOS</i> e <i>LOS + SOCIO</i>	100
Tabela 7 – Elasticidade agregada direta no ponto dos modelos <i>LOS</i> e <i>LOS + SOCIO</i>	103

LISTA DE ABREVIATURAS E SIGLAS

AIC	Critério de Informação de <i>Akaike</i>
API	Interface de Programação de Aplicação
BIC	Critério de Informação Bayesiano
BU	Bilhete Único
B&S	Teste de Ben Akiva e Swait
CPTM	Companhia Paulista de Trens Metropolitanos
CRRM	Modelo de Minimização do Arrependimento Aleatório Clássico
EM	<i>Expection-Maximization</i>
GEV	Modelo de Valor Extremo Generalizado
GHK	<i>Geweke-Hajivassiliou-Keene</i>
GMM	<i>Generalized Method of Moments</i>
GNNM	<i>General Nonlinear Nested Model</i>
GTFS	<i>General Transit Feed Specification</i>
IIA	Independência de Alternativas Irrelevantes
IID	Independente e Identicamente Distribuído
IPCA	Índice de Preço ao Consumidor Amplo
<i>knn</i>	<i>k-nearest neighbor</i>
LOS	<i>Level of service</i>
MACML	<i>Maximum Approximate Composite Marginal Likelihood</i>
MCMC	<i>Markov Chain Monte Carlo</i>
METRO	Companhia do Metropolitano de São Paulo
MV	Maximização de Verossimilhança
OD 2017	Pesquisa Origem e Destino 2017
OTP	<i>Open Trip Planner</i>
PD	Preferência Declarada
PIB	Produto Interno Bruto
PR	Preferência Revelada
P-RRM	Modelo de Minimização do Arrependimento Aleatório Puro
RAM	<i>Relative Advantage Maximising</i>
RDM	<i>Random Disutility Minimization</i>
RIS	<i>Recursive Importance Sampling</i>
RMSP	Região Metropolitana de São Paulo

RMU	<i>Random Modified Utility</i>
RRM	<i>Random Regret Minimisation</i>
RUM	<i>Random Utility Maximisation</i>
SAR	<i>Spatial Autoregressive model</i>
SARAR	<i>Spatial Autoregressive Autoregressive model</i>
SDM	<i>Spatial Durbin Model</i>
SDEM	<i>Spatial Durbin Error Model</i>
SEM	<i>Spatial Error Model</i>
SLX	<i>Spatial Lag of X model</i>
SMA	<i>Spatial Moving Average model</i>
TA	Teoria do Arrependimento
VTT	Valor do Tempo de Viagem

SUMÁRIO

1 INTRODUÇÃO	16
1.1 OBJETIVOS	19
1.2 JUSTIFICATIVA	19
1.3 ESTRUTURA DO TEXTO.....	21
2 ABORDAGENS COMPORTAMENTAIS EM MODELOS DE ESCOLHA	
DISCRETA	22
2.1 MAXIMIZAÇÃO DA UTILIDADE ALEATÓRIA – <i>RUM</i>	22
2.1.1 Tipos de dados	24
2.1.2 Estimador.....	24
2.1.3 Medidas para avaliação de políticas de transporte	26
2.2 MINIMIZAÇÃO DO ARREPENDIMENTO ALEATÓRIO - <i>RRM</i>	26
2.2.1 Especificações da função de arrependimento	31
2.2.2 Heterogeneidade de estratégias de decisões no processamento de atributos.....	31
2.2.3 Propriedades	32
2.2.4 Tipos de dados	33
2.2.5 Interpretação dos parâmetros.....	34
2.2.6 Medidas para avaliação de políticas de transporte	34
2.2.7 Aplicações	36
3 DEPENDÊNCIA ESPACIAL EM MODELOS DE ESCOLHA DISCRETA.....	39
3.1 ESPECIFICAÇÃO DE PROCESSOS ESPACIAIS	39
3.2 ESPECIFICAÇÃO DE MODELOS DE ESCOLHA DISCRETA ESPACIAIS	42
3.3 MATRIZ DE VIZINHANÇA E PESOS ESPACIAIS.....	46
3.3.1 Critérios de vizinhança	46
3.3.2 Propriedades da matriz de vizinhança	47
3.4 ESTIMADORES	48
3.5 MEDIDAS PARA AVALIAÇÃO DE POLÍTICAS DE TRANSPORTES	49
3.6 APLICAÇÕES	50
4 ÁREA DE ESTUDO, DADOS E MÉTODO	53
4.1 ÁREA DE ESTUDO	53
4.2 BASE DE DADOS	56
4.2.1 Motivo de viagem.....	58
4.2.2 Conjunto de alternativas	59

4.2.3 Viagens no município de São Paulo	60
4.2.4 Atributos sociodemográficos	60
4.3 CARACTERIZAÇÃO DE ATRIBUTOS DE NÍVEL DE SERVIÇO	61
4.3.1 Nível dos atributos	62
4.3.2 Distâncias de acesso, egresso e total	63
4.3.3 Tempo e custo de viagem <i>BUS</i>	64
4.3.4 Tempo e custo de viagem <i>RAIL</i>	65
4.3.5 Tempo e custo de viagem <i>CAR</i>	67
4.4 ATRIBUTOS PARA MODELAGEM	68
4.5 MÉTODO	71
4.5.1 Especificação da matriz de vizinhança e pesos espaciais	71
4.6 ESPECIFICAÇÃO DOS MODELOS	72
4.6.1 Modelos <i>RUM</i> não espaciais	72
4.6.2 Modelos <i>RRM</i> e <i>RMU</i> não espaciais	73
4.6.3 Modelos <i>RUM</i> espaciais	75
4.6.4 Modelos <i>RRM</i> e <i>RMU</i> espaciais	78
4.6.5 VTT e elasticidades	82
4.7 SOFTWARE	83
5 RESULTADOS	84
5.1 RESULTADOS GERAIS	84
5.1.1 Modelos com atributos de nível de serviço	84
5.1.2 Modelos com atributos de nível de serviço e sociodemográficos	90
5.2 RESULTADOS DOS PARÂMETROS ESTIMADOS	95
5.2.1 Parâmetros dos atributos de nível de serviço	96
5.2.2 Parâmetros de atributos sociodemográficos	98
5.2.3 Parâmetros espaciais	99
5.3 VALORES DE TEMPO DE VIAGEM E ELASTICIDADES	100
5.4 IMPLICAÇÃO DAS ABORDAGENS PARA FORMULAÇÃO DE POLÍTICAS PÚBLICAS DE TRANSPORTE	104
6 CONSIDERAÇÕES FINAIS	106
6.1 LIMITAÇÕES E RECOMENDAÇÕES PARA TRABALHOS FUTUROS	109
REFERÊNCIAS	111
APÊNDICE 1 – RESULTADOS DOS MODELOS ESPACIAIS	128

1 INTRODUÇÃO

A evolução das relações e interações nas cidades exige que os sistemas de transportes e as formas de se deslocar atendam a demanda pelas atividades criadas por essas interações. O crescimento demográfico experimentado nas cidades, principalmente do Sul Global (DAVIS, 2006), associado a um planejamento focado em estimular o uso do automóvel trouxe desafios que municípios, no caso do Brasil, ainda não conseguiram enfrentar. Consequentemente, as externalidades negativas advindas de um processo intenso de motorização e incentivo ao uso de modos individuais motorizados são sentidas por toda a sociedade, principalmente pelas parcelas mais vulneráveis (VASCONCELLOS, 2018).

No campo do planejamento de transportes, a mudança da perspectiva de análise da demanda tem contribuído para enfrentar esses novos desafios. A evolução da análise focada em viagens presente na técnica sequencial, ou 4 etapas, para uma abordagem de análise de atividades ajuda a compreender melhor os processos comportamentais e a maneira como as pessoas escolhem se deslocar (ou não) no espaço urbano (PINJARI; BHAT, 2011; RASOULI; TIMMERMANS, 2014).

Modelos de escolha discreta são utilizados como ferramenta para análise do comportamento relacionado a viagens, com o objetivo de entender os fatores e processos que afetam o comportamento de escolha dos indivíduos, seja na técnica sequencial ou na abordagem baseada em atividades. Além disso, tais modelos são base para a derivação de medidas econômicas utilizadas em avaliação de viabilidade de projetos de transporte, como o valor do tempo de viagem e análise de bem-estar social (HAGHANI; BLIEMER; HENSHER, 2021), além de servirem de ferramenta para análise de divisão modal na técnica sequencial de 4 etapas.

Modelos de escolha discreta têm sido tradicionalmente formulados com base na abordagem de maximização da utilidade aleatória (*Random Utility Maximization – RUM* daqui em diante), baseada nos axiomas econômicos de escolha de Luce (1959) e formalizados por Mcfadden (1974). Tal abordagem tem como premissa que indivíduos escolhem a alternativa com a maior utilidade dentre aquelas disponíveis, caracterizada de forma geral pelos atributos específicos das alternativas e de características sociodemográficas dos indivíduos, sendo utilizada em diversas áreas, como transportes, economia da saúde e de recursos naturais (HAGHANI; BLIEMER; HENSHER, 2021).

Mais recentemente, outras abordagens de análises de comportamento têm se apresentado como alternativa à *RUM*, que desde seu desenvolvimento tem sido a de maior disseminação nas análises comportamentais na área de transportes (CHORUS, 2012a). A

minimização do arrependimento aleatório (*Random Regret Minimization – RRM* daqui em diante) é um exemplo, preconizando que os indivíduos buscam minimizar seu arrependimento na escolha de uma alternativa frente as demais alternativas disponíveis (CHORUS; ARENTZE; TIMMERMANS, 2008).

Essa abordagem é baseada na teoria econômica do arrependimento (BELL, 1982; LOOMES; SUGDEN, 1982) que postula que o arrependimento é uma emoção determinante de escolhas em situações de risco (como apostas monetárias, por exemplo), além de ser considerada em outras áreas como fator determinante no processo de tomada de decisão, como a psicologia (ZEELLENBERG, 1999; ZEELLENBERG; PIETERS, 2007) e marketing (INMAN; DYER; JIANMIN, 1997; SIMONSON, 1992; TAYLOR, 1997).

Dessa forma, nos últimos dez anos, o número de trabalhos que fazem uso da abordagem *RRM* aumentou gradativamente, tanto na área de transportes onde foi inicialmente proposta, quanto em outras áreas do conhecimento, como saúde (PAUL *et al.*, 2018) e economia de recursos naturais (THIENE; BOERI; CHORUS, 2012). Entre as análises desenvolvidas, a maioria dos estudos buscou comparar a modelagem *RUM* com sua contraparte, *RRM*, a fim de analisar seu desempenho. Em menor escala, outros avanços foram feitos para compreender as propriedades da abordagem *RRM* (CHORUS, 2014; DEKKER, 2014; JANG; RASOULI; TIMMERMANS, 2017; VAN CRANENBURGH; GUEVARA; CHORUS, 2015).

Apesar de diversas pesquisas analisarem comparativamente o desempenho da abordagem *RRM* com sua contraparte *RUM*, esta não se propõe como alternativa superior em termos de desempenho ou em representação do comportamento humano. Mas sim como alternativa de modelagem comportamental para representação e entendimento do comportamento individual trazendo novos insights e implicações para políticas públicas através de novas propriedades como a presença de efeitos semi-compensatórios, por exemplo (CHORUS, 2012d). Dessa forma, a exploração da abordagem *RRM* é importante pois traz novas compreensões acerca do comportamento humano para a proposição de políticas de transporte eficazes, e não apenas para comparação de desempenho com sua contraparte *RUM*.

Simultâneo a esse processo, cresceram nas ciências sociais análises que tomam o espaço geográfico e social como parte explicadora dos fenômenos físicos e sociais. Assim, para além de uma característica dada, o espaço se torna elemento estruturante de padrões observáveis no espaço social e geográfico (ANSELIN, 1988). A primeira lei da Geografia, ou primeira lei de Tobler, preconiza que “todas as coisas estão relacionadas, mas coisas mais próximas estão mais relacionadas entre si que as mais distantes” (TOBLER, 1970), tendo inspirado diversas áreas a refletir sobre a importância de incorporar o espaço como elemento de análise. Assim,

mudanças significativas foram feitas em relação aos modelos que analisam eventos naturais com o surgimento da geoestatística, ou fenômenos sociais com o desenvolvimento da econometria espacial (ANSELIN, 2003).

O desenvolvimento da econometria espacial se deu principalmente pela necessidade de reflexão acerca dos pressupostos dos modelos da econometria que consideram que as observações, indivíduos ou regiões, são independentes. Quando considerada a posição que esses agentes tomam, padrões espaciais e efeitos de interações espaciais e sociais podem existir, violando a premissa de independência. Além disso, a forma de representação do fenômeno analisado no espaço pode causar efeitos de dependência entre as unidades sob análise. O objetivo da econometria espacial, então, tem sido analisar a presença desses efeitos que podem ser definidos como efeitos de dependência espacial e heterogeneidade espacial (ANSELIN, 1988).

Na proposição de modelos econométricos que incorporem efeitos de dependência espacial, diversos avanços foram feitos em modelos de variáveis contínuas e transversais, com um corpo teórico consolidado (ANSELIN, 2010). Além disso, há uma grande quantidade de trabalhos empíricos aplicados nas ciências regionais e na área de transportes.

A incorporação de efeitos espaciais em modelos não lineares, particularmente nos modelos de escolha discreta, no entanto, é menos explorada, dado que a inclusão desses efeitos apresenta um grau de complexidade maior em relação aos modelos lineares (CALABRESE; ELKINK, 2014). Na literatura de transportes, modelos de escolha discreta que incorporam efeitos de dependência espacial são explorados em menor quantidade e em poucos contextos de aplicação (GOETZKE, 2008; GOETZKE e WEINBERGER, 2012; PIKE e LUBELL, 2016; SHENG e SHARP, 2019; SIDHARTHAN *et al.*, 2011; WANG; AKAR; GULDMANN, 2015).

Assim, observa-se que as complexidades associadas aos fenômenos sociais alavancaram o desenvolvimento de técnicas que buscam compreender tanto as diferentes estratégias utilizadas no processo de tomada de decisão, assim como tais comportamentos também são condicionados pelo espaço geográfico e social de indivíduos ou regiões e suas interações.

No âmbito da abordagem *RRM* observa-se, no entanto, que a incorporação de efeitos de dependência espacial ainda não foi discutida na literatura até o presente momento. Rasouli e Timmermans (2016) mencionam a incorporação de efeitos de autocorrelação espacial em modelos de escolha discreta sob a abordagem *RRM* no contexto de escolhas de rotas, sendo o único trabalho encontrado até o momento a abordar o assunto, mesmo que de maneira superficial. Assim, aqueles autores inicialmente trouxeram a discussão sobre a existência de

tais efeitos e sua importância na abordagem *RRM* no contexto de comportamentos relacionados a viagens.

1.1 OBJETIVOS

O objetivo principal desta pesquisa é analisar um modelo de escolha discreta sob a abordagem de Minimização do Arrependimento Aleatório que incorpore efeitos de dependência espacial exógenos através de um estudo de caso no contexto de escolha de modos de transportes em viagens urbanas por motivo trabalho na cidade São Paulo, Brasil.

Para atingir o objetivo principal são estabelecidos os seguintes objetivos específicos:

- Sistematizar o procedimento de caracterização de atributos de nível de serviço (*Level of service* - LOS) para modelagem de comportamento relacionado a viagens usando modelos de escolha discreta.
- Analisar comparativamente as abordagens *RUM* e *RRM* no contexto de escolha de modo de transportes em viagens a trabalho sem efeitos de dependência espacial.
- Analisar comparativamente modelos sob as abordagens *RUM* e *RRM* com a presença de efeitos de dependência espacial no contexto de escolha de modos de transporte em viagens por motivo trabalho.
- Analisar as vantagens e desvantagens associadas às abordagens quanto à modelagem de escolha de modos de transportes e possíveis contribuições para análise de políticas públicas de transporte.

1.2 JUSTIFICATIVA

Como contextualizado anteriormente, é importante entender os processos que influenciam o comportamento relacionado a viagens para garantir políticas de transporte que melhorem a qualidade de vida da população de forma geral, e das parcelas mais vulneráveis particularmente. Atualmente, diferentes modelos buscam explicitar e explicar esses processos, incorporando, por exemplo, heterogeneidades nas estratégias de decisão (HESS *et al.*, 2012), no processamento de atributos (CHORUS; ROSE; HENSHER, 2013) e heterogeneidades inerentes ou não observadas (pelo analista) nos tomadores de decisão (GREENE; HENSHER, 2003; HENSHER; GREENE, 2003).

O crescimento de pesquisas que buscam analisar o comportamento relacionado a viagens sob perspectivas diferentes da abordagem *RUM* também evidencia como tais abordagens contribuem para a explicação das complexidades que se apresentam no processo de tomada de decisão (CHORUS, 2012a; HANCOCK *et al.*, 2020a; HANCOCK; HESS; CHOUDHURY, 2018). No entanto, observam-se duas lacunas na literatura de transportes quanto à aplicação de modelos que incorporam efeitos de dependência espacial e modelos sob a abordagem *RRM*.

A primeira lacuna, como será discutido detalhadamente mais adiante, refere-se às análises de comportamento relacionado a viagens que consideram efeitos de dependência espacial. A maior parte das pesquisas até o momento assumem a premissa de existência de efeitos espaciais autorregressivos (escolhas dos indivíduos são interdependentes no espaço geográfico/social), refletindo o estado da arte da econometria espacial (ELHORST, 2010). Assim, modelos de escolha discreta do tipo SAR (*Spatial Autoregressive*) e SEM (*Spatial Autoregressive Error*) têm sido os mais utilizados em comparação a outros modelos de dependência espacial, como o SLX (*Spatial Lag of X*), SDM (*Spatial Durbin*) e SDEM (*Spatial Durbin Error*) (CALDEIRA; ISLER; GIANNOTTI, 2021).

Este enfoque nos modelos autorregressivos reflete uma postura advinda da literatura da Econometria Espacial e das Ciências Regionais, em que o modelo SLX é menos explorado por teóricos dada sua simplicidade em relação aos modelos autorregressivos, e o maior interesse em analisar as condições de identificação e regularidade daqueles modelos (VEGA; ELHORST, 2015).

Além disso, acredita-se que essa prática tem relação com uma visão de contexto empírico acerca do comportamento individual relacionado a viagens, que na maioria dos casos é feita no Norte Global, onde a distribuição e disponibilidade dos sistemas de transportes é menos desigual. Nos países do Sul, vários estudos que utilizaram métricas de acessibilidade levando em conta a componente do sistema de transportes (GEURS; VAN WEE, 2004), evidenciam o padrão desigual e espacialmente concentrado dos sistemas de transporte que geram acessos desiguais de regiões a empregos por diferentes modos de transporte (LESSA; LOBO; CARDOSO, 2019; PEREIRA, 2018; PEREIRA *et al.*, 2019; PRITCHARD *et al.*, 2019; SLOVIC *et al.*, 2019).

Assim, encontra-se não apenas no campo da Engenharia de Transportes, mas nas Ciências Regionais e Econometria Espacial de forma geral, poucas aplicações de modelos de escolha discreta que consideram efeitos de dependência espacial exógena (modelo SLX)

(ALAMÁ-SABATER; ARTAL-TUR; NAVARRO-AZORÍN, 2011; AUTANT-BERNARD, 2007).

A segunda lacuna diz respeito à análise da abordagem *RRM* em contextos empíricos no Sul Global e à incorporação de efeitos de dependência espacial. Jing *et al.* (2018) apontam que aplicações de modelos de escolha de modos de transportes no contexto da abordagem *RRM* em países do Sul Global ainda é pouco explorada. Rasouli e Timmermans (2016) salientam a necessidade de incorporação de autocorrelação espacial (dependência espacial) em modelos sob a abordagem *RRM*.

Assim, esta pesquisa contribui para a literatura que trata do comportamento relacionado a viagens sob a perspectiva da econometria espacial pela análise de efeitos espaciais exógenos no contexto de escolhas de modos de transportes, e também quanto à consideração de efeitos de dependência espacial e suas implicações em modelos sob a abordagem *RRM* e suas aplicações no Sul Global.

1.3 ESTRUTURA DO TEXTO

Este texto é composto por 6 capítulos. Após esta introdução, o capítulo 2 discute as abordagens *RUM* e *RRM*, apresentando seus principais aspectos relacionados à especificação, estimação, interpretação, suas propriedades e limitações. O capítulo 3 discute a questão de dependência espacial em geral e aquelas específicas a modelos de escolha discreta, além de suas aplicações na análise de comportamento relacionado a viagens. O capítulo 4 apresenta a área de estudo, o processo de construção do banco de dados e a especificação dos modelos propostos. O capítulo 5 descreve os resultados obtidos e as vantagens e desvantagens das abordagens, assim como seus impactos na avaliação de políticas de transportes. Por fim, o capítulo 6 conclui o trabalho, discute suas limitações e aponta direções futuras de pesquisa.

2 ABORDAGENS COMPORTAMENTAIS EM MODELOS DE ESCOLHA DISCRETA

Este capítulo discute as abordagens *RUM* e *RRM* quanto a seus aspectos de especificação, estimação, necessidade de dados e indicadores para análise de políticas de transporte.

2.1 MAXIMIZAÇÃO DA UTILIDADE ALEATÓRIA – *RUM*

O uso de modelos probabilísticos possibilitou que diversos fenômenos mensurados qualitativamente pudessem ser analisados de forma estatística e inferencial. Dentre essas técnicas de análise destaca-se a modelagem de escolha discreta, também conhecida por regressão logística em algumas áreas em tarefas de classificação.

A abordagem mais utilizada para derivação de modelos de escolha discreta é a de maximização da utilidade aleatória proposta na década de 1970 (MCFADDEN, 1974). A principal premissa dessa abordagem é que mediante um conjunto de alternativas disponíveis para o indivíduo, este optará pela alternativa que oferece a maior utilidade dentre as demais. A utilidade é definida de forma geral como função de características sociodemográficas, atributos de serviço das alternativas, e assume-se que existem determinantes da escolha não observados ou não observáveis pelo analista.

Assim, a utilidade de uma alternativa é medida por uma parcela mensurável (determinística) e uma aleatória (erro aleatório). Considera-se que o termo de erro aleatório não pode ser medido, ou não é conhecido pelo analista, e pode representar idiosincrasias dos indivíduos quanto às suas escolhas. Essa relação pode ser observada na equação (1).

$$U_{in} = V_{in} + \varepsilon_{in} = \sum_{k=1}^K \beta_{ki} x_{ikn} + \varepsilon_{in} \quad (1)$$

em que U_{in} é a utilidade da alternativa i para indivíduo n ;

V_{in} é a parcela determinística da utilidade da alternativa i para o indivíduo n ;

ε_{in} é o termo de erro aleatório da alternativa i para o indivíduo n ;

x_{ikn} são os k atributos que influenciam a utilidade do indivíduo n quanto à alternativa i ; e

β_{ki} são os parâmetros associados aos atributos k e à alternativa i .

Assumindo uma distribuição de Valor Extremo do Tipo I, o modelo probabilístico resultante é conhecido por Logit, e no caso da distribuição Normal o modelo resultante é denominado Probit. Para o caso do modelo Logit, em que assume-se que os termos de erro são Independentes e Identicamente Distribuídos (IID), a probabilidade é calculada conforme a equação (2) (BEN-AKIVA; LERMAN, 1985; TRAIN, 2009).

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \quad (2)$$

A expressão acima representa a probabilidade de um indivíduo n escolher uma alternativa em um conjunto de alternativas $j \geq 2$. Embora menos complexo de estimar, o modelo Logit apresenta limitações em suas propriedades que por vezes pode não representar da melhor maneira a realidade. Uma das limitações mais relevantes é conhecida por Independência de Alternativas Irrelevantes (*Independence of Irrelevant Alternatives* - IIA), em que a relação entre as probabilidades de escolha entre alternativas não se altera mesmo pela inserção de uma nova alternativa ao conjunto daquelas disponíveis.

Além disso, a premissa de independência dos termos de erro pode não representar de forma fidedigna a realidade, pois existem situações em que tanto as alternativas quanto os indivíduos podem ser correlacionados. Este trabalho não apresenta detalhadamente todas as características e propriedades dos modelos de escolha discreta sob a abordagem *RUM*. No entanto, há na literatura uma vasta quantidade de pesquisas, manuais e livros em que modelos de escolha discreta sob a abordagem *RUM* são apresentados de forma simples e didática, assim como suas propriedades e limitações são discutidas em maior detalhe (BEN-AKIVA; LERMAN, 1985; GREENE, 2012; KOPPELMAN; BHAT, 2006; TRAIN, 2009).

Além dos modelos de escolha discreta já discutidos, diversos avanços foram feitos a fim de lidar com as limitações dos modelos Logit e Probit. Por exemplo, modelos que consideram a correlação entre alternativas já são amplamente utilizados, conhecidos por aqueles da família de distribuição de valor extremo generalizado, *Generalized Extreme Value Models* - GEV. Os modelos mais conhecidos são o aninhado (*Nested Logit*) e aninhado cruzado (*Cross Nested Logit*) (TRAIN, 2009). Esses modelos têm como principais vantagens a forma fechada da função densidade acumulada de probabilidade, além de controlar a existência de correlação entre as alternativas.

Outros modelos mais realistas e que nas últimas duas décadas vêm sendo bastante utilizados na literatura são conhecidos por modelos mistos (BROWNSTONE; BUNCH; TRAIN, 2000), híbridos (BEN-AKIVA *et al.*, 2002) e discretos-contínuos múltiplos de Valor Extremo (BHAT, 2008).

2.1.1 Tipos de dados

Diversas fontes de dados podem ser utilizadas para a estimação de modelos de escolha discreta. As fontes mais frequentes são conhecidas por dados de Preferência Revelada (PR) e de Preferência Declarada (PD), coletados por meio de entrevistas em campo ou *online*. Ambas as fontes apresentam vantagens e desvantagens em termos de modelagem e dos resultados obtidos. Dessa forma, a utilização de um ou de outro tipo de dado dependerá fundamentalmente do propósito e da disponibilidade de recursos para realização do estudo (TRAIN, 2009).

Os dados de PR foram e são muito utilizados pois são, de forma geral, publicamente disponibilizados por agências de planejamento de transportes, como é o caso da cidade de São Paulo com a pesquisa Origem e Destino, por exemplo (METRÔ-SP, 2017a). A principal vantagem desse tipo de dado é que fornece o comportamento real de escolha dos indivíduos. No entanto, dados de PR apresentam algumas desvantagens, como a incapacidade de considerar alternativas que não existem no momento de realização da pesquisa (TRAIN, 2009). Essa limitação impede a análise do comportamento de indivíduos pela inserção de uma alternativa inexistente, como um novo sistema de transportes, por exemplo.

Além disso, dados de PR podem conter pouca variabilidade nos atributos das alternativas devido a um fenômeno de equilíbrio de mercados (TRAIN, 2009), que pode prejudicar a correta estimação do seu efeito sobre o comportamento do indivíduo quando um atributo sofre alguma variação importante.

Os dados de PD, por sua vez, têm como característica a declaração do indivíduo sobre sua preferência por alguma alternativa em um cenário hipotético. Assim, a principal vantagem de utilização desse tipo de dado está na possibilidade de considerar alternativas não existentes no momento da pesquisa (TRAIN, 2009). Outra característica importante de dados de PD é que a variabilidade dos atributos das alternativas é definida pelo pesquisador. Uma das limitações de dados de PD é a disponibilidade de recursos para realização da pesquisa, apesar do uso de plataformas baseadas na internet facilitarem sua aplicação. Morikawa (1994) discute mais detalhadamente o uso de dados de PR e PD.

2.1.2 Estimador

A estimação dos parâmetros de modelos de escolha discreta é feita em sua grande maioria via método de máxima verossimilhança (GREENE, 2012), cujo objetivo é encontrar o conjunto de parâmetros β que maximiza a verossimilhança (ou probabilidade) da amostra

observada se repetir. Assim, assumindo que a variável de decisão segue uma distribuição de Bernoulli, a função de verossimilhança para uma amostra com n observações pode ser escrita conforme a equação (3) (GREENE, 2012).

$$L(\beta|dados) = \prod_{i=1}^n [F(\mathbf{x}'\beta)^{y_i}][F(\mathbf{x}'\beta)]^{1-y_i} \quad (3)$$

em que L é a função de verossimilhança;

$F(\cdot)$ é uma função de densidade de probabilidade acumulada;

$\mathbf{x}'$ é o vetor de variáveis explicativas do modelo;

y_i assume valor 1 se a alternativa é escolhida e 0 caso contrário; e

β é o vetor de parâmetros a serem estimados.

O procedimento de estimação é feito iterativamente, dado que a verossimilhança é não linear. Além disso, aplica-se o operador logaritmo na função de verossimilhança conforme a equação (4) para facilitar o seu cálculo computacionalmente.

$$\ln L = \sum_{i=1}^n \{y_i \ln F(\mathbf{x}'\beta) + (1 - y_i) \ln[1 - F(\mathbf{x}'\beta)]\} \quad (4)$$

Como a função logarítmica é monotônica, a função de verossimilhança mantém todas as propriedades necessárias para determinação do seu ponto de máximo. Ainda segundo Greene (2012), as condições de primeira e segunda ordem para o caso em que os termos de erro assumem distribuição Valor Extremo do Tipo I podem ser escritas, respectivamente, conforme as equações (5) e (6).

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^n (y_i - F_i) \mathbf{x}_i = \mathbf{0} \quad (5)$$

$$\mathbf{H} = \frac{\partial^2 \ln L}{\partial \beta \partial \beta'} = - \sum_{i=1}^n F_i (y_i - F_i) \mathbf{x}_i \mathbf{x}_i' = \mathbf{0} \quad (6)$$

em que $\mathbf{H}$ é a matriz de derivadas de segunda ordem, também denominada matriz de informação, matriz de Fischer ou matriz Hessiana.

A definição da matriz Hessiana é importante pois representa a variância dos parâmetros estimados, permitindo que testes estatísticos possam ser realizados a fim de avaliar o modelo estimado. As propriedades do estimador de máxima verossimilhança como consistência, eficiência e propriedades assintóticas podem ser encontradas em Greene (2012).

Para avaliar o desempenho e a significância dos modelos, uma série de testes estatísticos pode ser utilizada, como por exemplo: teste t de *Student* para avaliar a significância dos parâmetros estimados; teste de razão de verossimilhança; critério de informação de *Akaike* (AIC); critério de informação Bayesiano (BIC) e medida ρ^2 . O leitor obtém detalhes sobre a expressão matemática e as propriedades de cada teste e medida em Greene (2012). Para analisar a diferença de desempenho entre modelos de escolha discreta não aninhados (modelos que não são um caso especial do outro) pode-se utilizar o *Akaike Likelihood Ratio Index* proposto por Ben-Akiva e Swait (1986), doravante teste de B&S. Esse teste avalia se a diferença de desempenho em termos de log-verossimilhança de dois modelos não aninhados é estatisticamente significativa.

2.1.3 Medidas para avaliação de políticas de transporte

Diferente de modelos lineares, os parâmetros de modelos de escolha discreta não têm uma interpretação direta da variável de interesse em relação à variável dependente. Assim, para avaliação de diferentes políticas de transporte são utilizadas frequentemente medidas de disposição a pagar (*Willingness-to-Pay* – WTP), ou ainda elasticidades (agregadas/desagregadas ou pontuais/arco) das variáveis de interesse (BEN-AKIVA; LERMAN, 1985; KOPPELMAN; BHAT, 2006).

A medida mais utilizada para análise de viabilidade de projetos de transporte é o valor do tempo (*Value of Travel Time*, VTT), sendo caracterizada como a taxa de substituição do tempo pelo custo, ou seja, o quanto os indivíduos estão dispostos a pagar para que seu tempo seja reduzido.

As elasticidades são úteis para análise de cenários de políticas públicas, de maneira que indica a parcela de mercado das alternativas analisadas assumindo alguma mudança na variável da política de interesse, como é o caso do tempo de viagem ou custo monetário. Não são exploradas aqui suas formulações matemáticas no contexto da abordagem *RUM*, sendo o leitor direcionado a textos que apresentam tais aspectos em maior detalhe (BEN-AKIVA; LERMAN, 1985; HESS; BIERLAIRE; POLAK, 2005; KOPPELMAN; BHAT, 2006).

2.2 MINIMIZAÇÃO DO ARREPENDIMENTO ALEATÓRIO - *RRM*

A modelagem baseada na abordagem *RUM*, apesar de disseminada no âmbito acadêmico e aplicações práticas, nunca foi exclusiva para análise de processos de tomada de

decisão. Nos últimos 20 anos, no entanto, maior atenção tem sido dada a estruturas e estratégias de decisão que buscam superar as premissas associadas à maximização da utilidade aleatória (CHORUS, 2012a; HANCOCK *et al.*, 2020b; HANCOCK; HESS; CHOUDHURY, 2018), como por exemplo, a de racionalidade ilimitada (RASOULI; TIMMERMANS, 2015) e de efeitos compensatórios (THIENE; BOERI; CHORUS, 2012).

À luz desses novos paradigmas, a abordagem de minimização do arrependimento aleatório (*Random Regret Minimisation, RRM*) tem ganhado destaque nos estudos de comportamento relacionado a viagens. Proposta por Chorus, Arentze e Timmermans (2008), a abordagem *RRM* tem como premissa que o indivíduo baseia sua escolha entre alternativas a fim de evitar que uma alternativa não escolhida torne-se mais atrativa do que aquela escolhida, causando desse modo, arrependimento.

Baseada na Teoria do Arrependimento (TA) (BELL, 1982; LOOMES; SUGDEN, 1982), a abordagem *RRM* também se fundamenta em evidências de outras áreas em que a noção de arrependimento como determinante da escolha também é utilizada, como a psicologia (ZEELLENBERG, 1999), marketing (INMAN; DYER; JIANMIN, 1997; SIMONSON, 1992) e administração (BELL, 1982). A proposta de Chorus, Arentze e Timmermans (2008) foi introduzir a noção de arrependimento na análise do comportamento relacionado a viagens por meio do ferramental econométrico já consolidado na modelagem de escolha discreta.

Apesar da noção de arrependimento derivada da TA, ambas as abordagens diferem em alguns aspectos. Como destacado em Thiene, Boeri e Chorus (2012), a TA concentra-se em escolhas que apresentam risco e tem como objetivo capturar anomalias de escolha que não são bem caracterizadas na teoria de maximização da utilidade esperada. Em contraste, a abordagem *RRM* considera situações de risco e risco inexistente, objetivando principalmente capturar efeitos de comportamentos semi-compensatórios e efeitos de composição do conjunto de escolha (THIENE; BOERI; CHORUS, 2012).

Além disso, a TA tem enfoque na análise de escolha de alternativas com atributos únicos, diferente da abordagem *RRM* que permite a análise de escolhas com mais de um atributo. Por fim, na TA tem-se um modelo determinístico, enquanto a abordagem *RRM* é desenvolvida com base na modelagem de escolha discreta, que inclui heterogeneidades aleatórias por meio de um termo de erro aleatório (THIENE; BOERI; CHORUS, 2012). Neste trabalho não será demonstrada a derivação matemática da abordagem *RRM* a partir da TA. O leitor mais interessado pode encontrá-la em Chorus, Arentze e Timmermans (2008).

A parcela determinística do arrependimento foi primeiramente definida como uma função de comparações binárias entre os atributos das alternativas disponíveis, como mostrado nas equações (7), (8), (9) e (10).

$$R_{ij} = \varphi_x(x_i, x_j) + \varphi_y(y_i, y_j) + \varphi_z(z_i, z_j) \quad (7)$$

$$\varphi_x(x_i, x_j) = \max \{0, \beta_x(x_j - x_i)\} \quad (8)$$

$$\varphi_y(y_i, y_j) = \max \{0, \beta_y(y_j - y_i)\} \quad (9)$$

$$\varphi_z(z_i, z_j) = \max \{0, \beta_z(z_j - z_i)\} \quad (10)$$

em que R_{ij} é o arrependimento da alternativa i em relação à alternativa j em termos dos atributos x , y e z ;

$\varphi_x(x_i, x_j)$ é a parcela de arrependimento do atributo x da alternativa i comparado com o atributo x da alternativa j com melhor desempenho entre as demais alternativas; e $\beta_x, \beta_y, \beta_z$ são os parâmetros a serem estimados.

Chorus, Arentze e Timmermans (2008) salientam que os sinais dos parâmetros são estimados conjuntamente e que apresentam a direção esperada (teórica) para o atributo considerado analogamente à abordagem *RUM*. Assim, quando o desempenho da alternativa i é pior do que a alternativa comparada, o indivíduo experimenta o arrependimento. Vale salientar que o operador “max” limita o arrependimento de tal forma que nenhum arrependimento é produzido se o atributo da alternativa i tem um desempenho melhor que o atributo da alternativa j , ou seja, $\varphi_{ij} = 0$.

A forma funcional do arrependimento de um indivíduo escolher uma alternativa i em um conjunto de alternativas pode ser definida em termos de uma parcela determinística e uma parcela aleatória conforme a equação (11).

$$RR_i = R_{ij} + \varepsilon_i \forall i \neq j \quad (11)$$

em que RR_i é o arrependimento em escolher a alternativa i ;

R_{ij} é a soma das combinações de arrependimento binário para cada atributo; e

ε_i é um termo de erro de arrependimento aleatório.

Assumindo que o termo de erro seja representado por uma distribuição Valor Extremo do tipo I, a probabilidade do indivíduo escolher uma alternativa assume a mesma equação do modelo Logit. No entanto, ao invés de apresentar uma função aditiva e linear nos parâmetros, a equação do arrependimento apresenta uma descontinuidade em zero, não sendo facilmente

derivável, e dessa forma o modelo proposto não é facilmente estimado usando-se técnicas existentes implementadas em softwares disponíveis. Mesmo sendo estimado pelo método de máxima verossimilhança, o modelo exige que a codificação seja feita para cada caso específico, dificultando a disseminação de aplicações práticas e empíricas da abordagem *RRM* (CHORUS, 2010; CHORUS; ARENTZE; TIMMERMANS, 2008).

Outra limitação importante do modelo de Chorus, Arentze e Timmermans (2008) é que a comparação entre atributos das alternativas é feita apenas entre a de melhor desempenho e a alternativa considerada. Chorus (2010) discute que o arrependimento também poderia ser experimentado pela comparação da alternativa considerada com as demais alternativas que tenham melhor desempenho.

Chorus (2010) propôs, então, uma nova formulação da função de arrependimento que supera tais limitantes. Assim, Chorus (2010) propôs uma formulação contínua para o arrependimento, e desse modo, facilmente derivável. Além disso, a nova função expande a comparação entre a alternativa considerada com aquela de melhor desempenho para a comparação entre todas as alternativas.

Para derivação do modelo econométrico proposto, Chorus (2010) considera que o analista não dispõe ou não conhece todas as variáveis que influenciam a escolha dos indivíduos e as descreve em termos de um termo de arrependimento aleatório, considerando que os níveis do arrependimento não são determinísticos. A forma funcional do arrependimento binário pode ser escrita em termos de dois componentes de erro aleatório conforme a equação (12) (CHORUS, 2010).

$$R_{ij}^k = \max \{0 + v_{0k}, \beta_k(x_{jk} - x_{ik}) + v_{xk}\} \quad (12)$$

em que R_{ij}^k é o arrependimento binário da alternativa i em relação a alternativa j para o atributo k ; e

v_{0k} e v_{xk} são termos de erro com distribuição Valor Extremo do Tipo I e IID.

Para calcular as probabilidades, os termos aleatórios devem ser integrados. Chorus (2010) apresenta o termo de arrependimento binário em uma formulação logsum baseada em Ben Akiva (1985) conforme equação (13).

$$\int_v [R_{ij}^k \cdot f(v) dv] = \ln(1 + \exp[\beta_k \cdot (x_{jk} - x_{ik})]) \quad (13)$$

Assim, o arrependimento experimentado pela escolha de uma alternativa i em termos de k atributos comparados com $j \geq 2$ alternativas pode ser escrito conforme a equação (14).

$$R_i = \sum_{j \neq i} \sum_{k=1 \dots K} \ln(1 + \exp[\beta_k \cdot (x_{jk} - x_{ik})]) \quad (14)$$

O termo R_i representa a soma dos arrependimentos binários para os K atributos com as $j \geq 2$ alternativas. A diferença entre os termos da parcela do arrependimento do modelo de 2008 e do modelo de 2010 podem ser observados na Figura 1. Observa-se que as funções se diferem no intervalo próximo de zero com valor máximo de 0,693147 nesse ponto.

Figura 1 – Parcela do arrependimento do modelo de 2008 e 2010.

Fonte: Adaptado de Chorus (2010).

Por fim, escrevendo o arrependimento em função da combinação da parcela observável e do termo de erro aleatório com distribuição Valor Extremo do Tipo I e IID, a probabilidade de escolha de uma alternativa baseada na abordagem *RRM* pode ser definida conforme a equação (15).

$$P_i = \frac{\exp(-R_i)}{\sum_{j=1 \dots J} \exp(-R_j)} \quad (15)$$

Nessa expressão, o sinal negativo é inserido pois minimizar o arrependimento é equivalente matematicamente a maximizar seu valor negativo. O modelo proposto por Chorus (2010) então se reduz a uma função contínua, derivável e que pode ser calculada pelas ferramentas de estimativa disponíveis, sem a necessidade de codificação específica das equações de verossimilhança. Destaca-se que Chorus (2010) explicita o caso específico com apenas duas alternativas, cujos parâmetros estimados equivalem àqueles obtidos na abordagem *RUM*.

2.2.1 Especificações da função de arrependimento

Mais recentemente, novas especificações da função de arrependimento foram propostas na literatura como resultado de generalizações baseadas no modelo proposto por Chorus (2010). Entre eles pode-se citar Chorus (2014), que propôs um modelo de minimização do arrependimento aleatório generalizado (*Generalized Random Regret Minimisation model - GRRM*), cuja constante unitária da parcela determinística do arrependimento é substituída por um parâmetro a ser estimado no intervalo entre zero e um. Quando o parâmetro assume valor zero, o componente de arrependimento desaparece e o atributo é estimado sob a abordagem *RUM*. Se o parâmetro assume valor diferente de zero o modelo é estimado sob a abordagem *RRM*, resultando no modelo clássico (CHORUS, 2010) quando o parâmetro é igual um.

Os modelos *RRM* mais recentes encontrados na literatura são oriundos do trabalho de van Cranenburgh, Guevara e Chorus (2015). Os autores propõem um novo conceito na abordagem *RRM* denominada de profundidade do arrependimento, que se refere à medida com que modelos baseados na abordagem *RRM* impõe o comportamento de minimização do arrependimento. Cranenburgh, Guevara e Chorus (2015) propuseram duas novas especificações, *Pure-RRM (P-RRM)* e μ *RRM*, em que a primeira assume que indivíduos fazem a escolha apenas no domínio do arrependimento, e a segunda é uma generalização dos modelos *CRRM* e *P-RRM* por meio da identificação do parâmetro de profundidade do arrependimento em conjunto com o parâmetro de escala do modelo Logit.

2.2.2 Heterogeneidade de estratégias de decisões no processamento de atributos

A abordagem *RRM* preconiza que o arrependimento é experimentado quando uma alternativa escolhida, em termos de um ou mais atributos, tem desempenho pior comparado ao mesmo atributo nas demais alternativas. No entanto, a modelagem baseada na abordagem *RRM* não se limita à análise de atributos gerais para todas as alternativas. Assim, atributos específicos da alternativa, variáveis sociodemográficas e variáveis *dummy* podem ser utilizadas nas funções de arrependimento (CHORUS, 2012b, 2012d).

Chorus, Rose e Hensher (2013) propuseram que diferentes atributos podem ser processados por diferentes estratégias de decisão, ou seja, indivíduos podem maximizar a utilidade de algum atributo enquanto tentam diminuir o arrependimento em relação a outro. Chorus, Rose e Hensher (2013) afirmam que heterogeneidades no processamento de atributos por diferentes estratégias (*RUM* e *RRM*) podem melhorar a análise do comportamento

relacionado a viagens. A utilidade de uma alternativa em que se considera ambas as abordagens, denominada utilidade modificada aleatória (*Random Modified Utility*, *RMU*), pode ser observada na equação (16).

$$RMU_i = \sum_{k=1 \dots q} \beta_{ik} x_{ik} - \sum_{k=q+1 \dots K} \sum_{i \neq j} \ln (1 + \exp (\beta_k (x_{jk} - x_{ik}))) + \varepsilon_i \quad (16)$$

em que RMU_i é a utilidade modificada para a alternativa i ;

x_{ik} e x_{jk} são os q atributos da alternativa i sob a premissa da abordagem *RUM* e os $K - q$ atributos sob a premissa da abordagem *RRM*, respectivamente; e ε_i é o termo de erro aleatório.

Chorus, Rose e Hensher (2013) analisaram três modelos, *RUM*, *RRM* e *RMU*, utilizando dois bancos de dados. O primeiro com dados de PR acerca de escolhas de destino e o segundo de dados de PD sobre escolhas de rotas. Para os dois bancos de dados o modelo híbrido *RMU* teve melhor desempenho, indicando a presença de heterogeneidade no processamento dos atributos. Os autores também discutem as diferenças nas medidas de disposição a pagar (WTP) dos modelos estimados.

2.2.3 Propriedades

Além de trazer novos recursos para a análise do comportamento relacionado a viagens, a abordagem *RRM* também apresenta propriedades que não podem ser observadas nos modelos Logit baseados na abordagem *RUM* (THIENE; BOERI; CHORUS, 2012), destacando-se a sensibilidade à composição do conjunto de escolhas (ZHANG *et al.*, 2004) e o efeito de tomada de decisão semi-compensatório (ARENTZE; TIMMERMANS, 2007; SWAIT, 2001).

A primeira refere-se à propriedade IIA do modelo Logit, que preconiza que a inserção de uma nova alternativa ao conjunto de escolhas não afeta a razão de probabilidade entre duas alternativas. A segunda propriedade, conhecida por efeito semi-compensatório, é diretamente resultante da forma convexa da função do arrependimento (Chorus, 2012c).

Assim, na abordagem *RRM* a propriedade IIA não é válida para o modelo Logit visto que a probabilidade de uma alternativa depende da comparação dos atributos dessa alternativa com as demais disponíveis (CHORUS, 2012d). Naquele mesmo trabalho o autor salienta que a abordagem *RRM* não viola a propriedade IID, mas afirma que essas limitações podem ser

superadas na abordagem *RRM* da mesma forma que na abordagem *RUM* por meio de modelos como os aninhados e mistos.

A propriedade de efeito semi-compensatório resulta da convexidade da função do arrependimento, tal que a melhora do desempenho de uma alternativa em termos de um atributo com bom desempenho resulta em pequeno decréscimo do arrependimento, enquanto a deterioração de outro atributo igualmente importante em uma alternativa com baixo desempenho resulta em aumento substancial do arrependimento (CHORUS, 2012d).

Assim, uma implicação importante dessa propriedade é que na abordagem *RRM* as alternativas que têm atributos intermediários (*in-between*) em relação às demais recebem um “bônus de parcela de mercado”. Chorus (2012d) apresenta um exemplo numérico para ilustrar essa propriedade e destaca que essa condição está de acordo com evidências empíricas encontradas em outras áreas, cujo fenômeno é conhecido por *compromise effect*.

2.2.4 Tipos de dados

A discussão feita na seção 2.1.1 sobre dados para modelagem sob a abordagem *RUM* pode ser estendida para a abordagem *RRM*. No entanto, o uso de dados de PR tem implicações mais significativas que na abordagem *RUM*. No contexto da abordagem *RRM*, Wong *et al.* (2020) discutem as dificuldades associadas ao uso de dados de PR. Os autores destacam que a ausência de variabilidade sistemática no nível dos atributos e o desconhecimento do conjunto real de alternativas considerado pelo indivíduo podem ser fatores limitantes. Os autores também destacam que a maioria dos trabalhos empíricos têm utilizado dados de PD na abordagem *RRM* pela sua facilidade em determinar o conjunto de escolhas e pela possibilidade de controlar a variabilidade no nível dos atributos.

Além disso, Wong *et al.* (2020) elencam algumas recomendações para realização de análises com dados de PR na abordagem *RRM*:

- Pelo menos duas alternativas devem ser consideradas em adição à alternativa revelada, visto que modelos binários produzem os mesmos resultados para a abordagem *RUM* e *RRM*;
- Devem ser estabelecidos atributos numéricos para as alternativas para que comparações no nível do atributo possam ser realizadas;
- Deve haver variação suficiente nos níveis dos atributos e nas diferenças entre os respectivos níveis.

Essas recomendações não são mandatórias, principalmente no que diz respeito à variação nos atributos, mas os autores salientam que a falta de variação sistemática nos níveis dos atributos pode impedir a identificação do arrependimento mesmo quando presente no processo de tomada de decisão. Esse problema ocorre pois funções não lineares (arrependimento) podem ser aproximadas por funções lineares (utilidade) para pequenos valores (diferenças nos níveis dos atributos), analogamente ao apresentado no trabalho de Rasouli e Timmermans (2016). Além disso, Wong *et al.* (2020) destacam que modelos *RRM* que usaram dados de PR apresentaram melhor desempenho, de maneira geral, que a abordagem *RUM* em comparação ao uso de dados de PD.

2.2.5 Interpretação dos parâmetros

A interpretação dos parâmetros sob a abordagem *RRM* deve ser diferente em comparação aos resultados dos modelos estimados sob a abordagem *RUM*. Chorus (2012c) destaca que sob esta abordagem o parâmetro estimado indica um acréscimo (decréscimo) na utilidade de uma alternativa causada pelo acréscimo (decréscimo) em uma unidade no valor do atributo, enquanto naquela o parâmetro estimado refere-se ao potencial acréscimo (decréscimo) no arrependimento causado pela comparação da alternativa considerada em relação às demais alternativas devido ao acréscimo (decréscimo) no valor do atributo dessa alternativa.

Chorus (2012c) destaca que a comparação das duas abordagens em termos das magnitudes dos parâmetros não tem significado prático, argumentando que a magnitude do parâmetro de modelos sob a abordagem *RRM* é inversamente proporcional ao tamanho do conjunto de alternativas. Assim, quanto maior o número de alternativas, menor é a magnitude do parâmetro estimado. Além disso, maiores (menores) valores para os parâmetros dos modelos sob a abordagem *RRM* não implicam em afirmar que um parâmetro é mais importante que outro entre as abordagens *RUM* e *RRM*, nem tão pouco que há mais (menos) heterogeneidades não observadas em modelos *RRM* por sua maior (menor) magnitude (CHORUS, 2012c).

2.2.6 Medidas para avaliação de políticas de transporte

No caso da abordagem *RRM*, a derivação de medidas para avaliação de políticas de transporte se difere da abordagem *RUM*. O trabalho de Chorus (2012d) foi o primeiro a discutir

a formulação da medida de VTT para modelos de escolha discreta baseados na abordagem *RRM*¹, cuja formulação para o modelo *CRRM* pode ser verificada na equação (17).

$$VTT_{RRM} = \frac{\partial R_i / \partial TT_i}{\partial R_i / \partial TC_i} = \frac{\sum_{j \neq i} -\beta_{TT} / (1 + 1/\exp [\beta_{TT} \cdot (TT_j - TT_i)])}{\sum_{j \neq i} -\beta_{TC} / (1 + 1/\exp [\beta_{TC} \cdot (TC_j - TC_i)])} \quad (17)$$

em que VTT_{RRM} é o valor do tempo do modelo *CRRM*;

R_i é a parcela determinística do arrependimento em relação a alternativa i ;

TT_i e TC_i são os atributos de tempo e custo da alternativa i , respectivamente; e

β_{TT} e β_{TC} são o parâmetro de tempo e custo estimados, respectivamente.

Como se observa, diferentemente da métrica sob a abordagem *RUM*, o VTT_{RRM} é específico para o conjunto de alternativas considerado pelos indivíduos, além do valor do tempo de uma alternativa ser dependente dos atributos das demais. Dessa forma, o VTT_{RRM} muda conforme se altera o desempenho relativo das alternativas (CHORUS, 2012d). Dekker (2014) faz uma discussão profunda sobre a derivação do VTT no contexto da abordagem *RUM* e *RRM*, destacando suas principais diferenças e limitações.

Dekker (2014) aponta que a discussão feita por Chorus (2012d) não apresenta todas as implicações da medida VTT para análise de políticas de transporte. Dekker (2014) discute que tal medida é baseada na função de arrependimento indireto condicional e na noção econômica de indiferença, de maneira que o VTT nesta abordagem representa a disposição a pagar por reduções no tempo de viagem dado que a alternativa i é escolhida. Assim, essa medida informa o analista como o tempo e o custo da alternativa i pode ser ajustado a fim de manter os usuários dessa alternativa indiferentes em relação às demais.

Além disso, a propriedade de dependência (sensibilidade) do conjunto de escolha implica que usuários das demais alternativas também são afetados pela mudança nos níveis de tempo e custo da alternativa i . Ou seja, a taxa (VTT) na qual cada usuário precisa ser compensado difere entre as alternativas.

Chorus, Rose e Hensher (2013) foram os primeiros a apresentarem a formulação das equações de elasticidade do modelo *CRRM*. Belgiawan *et al.* (2019) e van Cranenburgh, Guevara e Chorus (2015) discutiram tais medidas no contexto dos modelos *P-RRM* e μ *RRM*. A equação (18) apresenta a equação para o cálculo das elasticidades sob a abordagem *RRM*.

¹ Alguns autores discutem sobre o uso correto do termo *Value of Travel Time* (VTT) no caso de modelos sob a abordagem *RRM* (FORUM APOLLO, 2021), dado que não apresentam consistência com a abordagem de maximização da utilidade (HESS; DALY; BATLEY, 2018). No entanto, esse termo é utilizado nessa dissertação baseado nas discussões feitas por Dekker (2014) e seu uso em outros trabalhos (BELGIAWAN *et al.*, 2019).

$$E_{X_{kq}}^{RRM} = \left(-\frac{\partial R_{iq}^{RRM}}{\partial X_{kq}} + \sum_{\substack{j \in J \\ j \neq i \\ j=1}}^J P_{jq} \cdot \frac{\partial R_{jq}^{RRM}}{\partial X_{kq}} \right) \cdot X_{kq} \quad (18)$$

em que $E_{X_{kq}}^{RRM}$ é a elasticidade desagregada relativa ao atributo X_{kq} da alternativa i referente ao indivíduo q ;

R_{iq}^{RRM} é a parcela determinística do arrependimento referente a alternativa i e indivíduo q ;

R_{jq}^{RRM} é a parcela determinística do arrependimento referente a alternativa j e indivíduo q ;

P_{jq} é a probabilidade de escolha da alternativa j e indivíduo q ; e

X_{kq} é o nível do atributo k da alternativa i referente ao indivíduo q .

2.2.7 Aplicações

Apesar de relativamente recentes na literatura em comparação a abordagem *RUM*, os modelos baseados na abordagem *RRM* tem ganhado força principalmente na área de transportes onde foi originalmente proposto. No entanto, outras áreas também têm feito uso dessa nova abordagem. Por exemplo, em estudos na área de economia da saúde (DE BEKKER-GROB; CHORUS, 2013; PAUL *et al.*, 2018), economia de recursos naturais (BOERI; LONGO, 2017; THIENE; BOERI; CHORUS, 2012), excedente do consumidor (DEKKER; CHORUS, 2018), segurança alimentar (DENNIS *et al.*, 2020) e alimentação (BIONDI *et al.*, 2019).

Na área de transportes, os modelos baseados na abordagem *RRM* têm sido usados em diferentes contextos, como segurança viária (KAPLAN; PRATO, 2012), transporte de cargas (BOERI; MASIERO, 2014; KEYA; ANOWAR; ELURU, 2018), equilíbrio de tráfego e escolha de rotas (LI *et al.*, 2018; LI; HUANG, 2017; VAN CRANENBURGH; PRATO, 2016; XU; ZHOU; XU, 2020), demanda por locais de recreação (BOERI *et al.*, 2012) e preferências de tomadores de decisão acerca de políticas de transporte (CHORUS *et al.*, 2011).

No que se refere ao uso dados de PR, os trabalhos de Chorus, Cranenburgh, e Dekker (2014) e Wong *et al.* (2020) apontam que o uso tem sido bastante limitado. As 42 comparações entre *RUM* e *RRM* analisadas por Chorus, Cranenburgh, e Dekker (2014) indicaram que 37 usaram dados de PD, enquanto apenas 5 usaram dados de PR. Uma revisão da literatura sobre modelos de escolha de modos de transporte e de escolha de rotas sob a abordagem *RRM* pode

ser consultada em Jing *et al.* (2018), em que os autores destacam as poucas pesquisas sobre escolhas de modo de transporte e a reduzida exploração dessa abordagem em países do Sul Global.

O trabalho de Leong e Hensher (2015) compararam a aplicação de modelos baseados na abordagem *RUM*, *RRM* e *RAM* (*Relative Advantage Maximising*) com objetivo de analisar o desempenho de modelos que apresentem a propriedade de efeito semi-compensatório existente nos modelos *RRM* e *RAM* com a abordagem *RUM*. O trabalho utilizou oito bancos de dados de PD, sendo os sete primeiros relacionados a escolha de rotas em cenários de aplicação de cobrança tarifária (*toll-road*), e o último relacionado a escolha de modos de transporte a fim de analisar a implantação de uma linha de metrô. Os autores também estimaram um modelo híbrido *RUM-RRM*. O modelo *RAM* teve melhor desempenho em termos de log-verossimilhança em sete dos oito bancos de dados. O modelo *RRM* teve melhor desempenho em apenas um banco de dados, mas as diferenças nos ajustes entre os modelos foram pequenas.

Parthan e Srinivasan (2013) analisaram a presença de efeitos semi-compensatórios na escolha de modos de transporte em viagens a trabalho na cidade de Channai, China. Os autores estimaram modelos com diferentes estratégias de decisão (*RUM*, *RRM* e *Random Disutility Minimization - RDM*) incluindo também nas funções utilidade fatores subjetivos acerca dos sistemas de transportes. O conjunto de escolhas era composto por seis alternativas: veículo motorizado sobre duas rodas, veículo motorizado sobre quatro rodas, ônibus, trem, paratransito e modo não motorizado. Os resultados indicaram que os modelos com efeitos semi-compensatórios (*RDM* e *RRM*) tiveram desempenho melhor que o modelo baseado na abordagem *RUM*.

O trabalho de Caldeira e Isler (2021) apresenta uma revisão da literatura dos diferentes tipos de heterogeneidades exploradas no contexto da abordagem *RRM*, aplicando-as no contexto de escolha de modos de transporte na cidade de São Paulo, explorando heterogeneidades de estratégias de decisão, no processamento de atributos e na escala do arrependimento para os atributos custo e tempo. Apesar de modelos mais complexos, com maior heterogeneidade, apresentarem melhores resultados em termos de log-verossimilhança, as medidas de análise de políticas como VTT e elasticidades não apresentaram resultados razoáveis em comparação aos modelos mais simples.

Por fim, entre as pesquisas desenvolvidas até o momento que utilizaram a abordagem *RRM*, observa-se uma vasta variedade de aplicações além da área de transportes e análises de comportamento relacionado a viagens. No entanto, a quantidade de estudos que analisaram escolhas de modo de transporte sob a abordagem *RRM* ainda é pequena, principalmente no Sul

Global. Assim, acredita-se que a presente pesquisa contribui para o estado da arte da abordagem *RRM* ao analisar o comportamento de escolha de modos de transportes em contexto de aplicação do Sul Global.

3 DEPENDÊNCIA ESPACIAL EM MODELOS DE ESCOLHA DISCRETA

Neste capítulo são discutidos a especificação de dependência espacial em modelos lineares e de escolha discreta, estimadores de modelos de escolha discreta espacial, especificação de matrizes de vizinhança e pesos espaciais e aplicações empíricas de modelos de escolha discreta espaciais na literatura de análise de comportamento relacionado a viagens.

3.1 ESPECIFICAÇÃO DE PROCESSOS ESPACIAIS

O objetivo da econometria espacial, de forma geral, é identificar a presença de processos espaciais nos fenômenos das ciências regionais e da geografia humana. Dessa forma, busca-se tratar dos problemas decorrentes dos pressupostos de modelos econométricos utilizados nas ciências sociais que assumem que regiões ou indivíduos são independentes. Quando o espaço é incorporado nas análises, dificilmente essa premissa é válida e se a dependência espacial não é controlada os parâmetros estimados tendem a ser enviesados, além de negligenciar os processos espaciais subjacentes ao fenômeno sob análise.

Assim, a dependência espacial pode ser definida, em termos gerais, como “a existência do relacionamento funcional entre o que acontece em um ponto no espaço e o que acontece em qualquer outro” (ANSELIN, 1988). Além disso, segundo Anselin (1988), a dependência espacial pode ser o resultado de problemas de mensuração, como por exemplo a delimitação arbitrária ou problemas de agregação de unidades espaciais.

Ainda mais importante que os erros de medida, a dependência espacial pode ser também o resultado de uma variedade de fenômenos de interações espaciais, ou ainda efeitos de *spillovers*, ou seja, o transbordamento do efeito de uma variável no espaço, afetando direta e indiretamente as observações vizinhas. Este último fator é mais relevante e segue da importância do espaço como elemento na estruturação de explicações do comportamento humano (ANSELIN, 1988).

Assim, a noção de dependência espacial implica a necessidade de definir quais outras unidades no sistema espacial têm influência sobre a unidade considerada. Anselin (1988) aponta que isso pode ser expresso em noções topológicas de vizinhança ou de vizinho mais próximo. Definido o conjunto de unidades vizinhas de uma observação, essa relação e esse conjunto podem ser representados por meio de uma matriz de conectividade, ou matriz de vizinhança.

O resultado da dependência espacial nos modelos econométricos é diretamente observado na matriz de variância-covariância dos termos de erro dos modelos, onde se observa

que a covariância das observações é diferente de zero, além dos termos de variância não serem constantes nos modelos autorregressivos (ANSELIN, 2003).

Desse modo, a dependência espacial pode ser controlada diretamente através da matriz de variância-covariância definindo-se uma função de decaimento de distância, ou ainda representando esse decaimento por uma função de passo sob a forma não paramétrica, como é o caso na Geoestatística (ANSELIN, 2003).

A estrutura espacial também pode ser capturada indiretamente, em que um processo estocástico espacial é especificado de tal forma que relaciona o valor de uma variável aleatória em uma localização aos valores da variável aleatória em suas localizações vizinhas. Assim, ao invés de ligar todos os pares de observações através de uma função de decaimento, os vizinhos de cada observação são definidos pela matriz de pesos espaciais usada na econometria espacial.

Anselin (2003) define dois tipos de efeitos espaciais que podem surgir a partir da especificação de um modelo espacial em termos de seus impactos: global e local. Os efeitos globais surgem da especificação de um processo espacial em que o efeito de interação entre as observações se distribui por todo o espaço das observações enquanto os efeitos locais impactam apenas a vizinhança direta de uma observação.

Considere o modelo linear apresentado nas equações (19) e (20).

$$y = \rho W_1 y + W_3 X \theta + X \beta + u \quad (19)$$

$$u = \lambda W_3 u + \varepsilon \quad (20)$$

em que y é o vetor ($n \times 1$) da variável dependente;

W_1, W_2, W_3 são matrizes de pesos espaciais de dimensão ($n \times n$);

X é a matriz ($n \times k$) de variáveis independentes com k variáveis;

ρ, λ são parâmetros escalares a serem estimados e $-1 < \rho < 1$ e $-1 < \lambda < 1$;

θ, β são vetores ($k \times 1$) de parâmetros a serem estimados;

u é um vetor ($n \times 1$) de erros correlacionados espacialmente; e

ε é um vetor ($n \times 1$) de erros aleatórios IID que segue uma distribuição Normal.

A partir desse modelo é possível definir uma variedade de modelos lineares que controlam a dependência espacial presente nos dados. O modelo com $\theta = 0$ e $\lambda = 0$ é conhecido na literatura por *Spatial Autoregressive model* (SAR) e controla a dependência espacial na variável dependente. Quando $\rho = 0$ e $\lambda = 0$, obtém-se o modelo *Spatial Lag of X* (SLX) e a dependência espacial presente nas variáveis independentes é controlada. O modelo em que $\rho = 0$ e $\theta = 0$ é conhecido por *Spatial Error model* (SEM) e controla a dependência

espacial nos erros. Uma variedade de outros modelos é alcançada a partir do modelo da equação (19) pela especificação dos parâmetros, mas os modelos acima discutidos são os mais conhecidos. Além disso, os modelos SAR e SEM são de efeitos globais, enquanto o modelo SLX é de efeito local.

O efeito global pode ser observado em modelos lineares em que há dependência espacial nos termos de erro (SEM) por exemplo, conforme mostrado na equação (21).

$$u = \lambda W u + \varepsilon \quad (21)$$

Assim, resolvendo a equação para o termo de erro espacialmente correlacionado, obtém-se conforme equação (22):

$$u = [I - \lambda W]^{-1} \varepsilon \quad (22)$$

Anselin (2003) discute que a matriz inversa pode ser expressa através de uma série de expansão $[I - \lambda W]^{-1} = I + \lambda W + \lambda^2 W^2 + \dots$. As potências de W são matrizes de vizinhança de ordem superior, tal que a matriz de vizinhança de ordem dois (W^2) considera que são vizinhos os vizinhos dos vizinhos, e assim por diante. A partir dessa formulação é possível observar que os efeitos são propagados de uma observação para outra de tal maneira que o efeito é maior nas proximidades de uma observação, e vai diminuindo para as demais unidades espaciais (visto que $-1 \leq \lambda \leq +1$). Essa formulação segue a primeira lei da geografia que preconiza que “todas as coisas estão relacionadas, mas coisas mais próximas estão mais relacionadas entre si que as mais distantes” (TOBLER, 1970).

Anselin (2003) ainda destaca que as variâncias, termos da diagonal principal da matriz de variância-covariância, dependem dos elementos das diagonais das matrizes da expansão do termo $[I - \lambda W]^{-1}$ e assim, no caso em que a estrutura de vizinhança não é constante para todas as observações (maioria das situações), os termos de erro serão heterocedásticos mesmo que o processo inicial não o seja.

Os efeitos locais são mais simples de observar dado que seu impacto é limitado às observações diretamente vizinhas, em que um processo de especificação de média móvel (*moving average specification* - SMA), por exemplo, é adotado, como apresentado na equação (23) (ANSELIN, 2003).

$$u = \gamma W \varepsilon + \varepsilon \quad (23)$$

em que γ é o parâmetro espacial.

Como essa já é forma reduzida da equação, ela não implica em inversão da matriz de vizinhança. Assim, a matriz de variância-covariância pode ser expressa pela equação (24).

$$E[\varepsilon\varepsilon'] = \sigma^2[(I + \gamma W)(I + \gamma W)' = \sigma^2[I + \gamma(W + W') + \gamma^2 WW']] \quad (24)$$

Nessa equação, os elementos diferentes de zero fora da diagonal principal da matriz de variância-covariância são apenas os elementos diferentes de zero na matriz de vizinhança W (de forma equivalente W') e WW' . Assim, o efeito espacial se propaga até os vizinhos de segunda ordem de uma observação (WW'), não se estendendo dessa forma para as demais observações distribuídas na área de estudo, como acontece para o efeito global (ANSELIN, 2003).

O campo da econometria espacial, então, tem se desenvolvido de forma que diferentes especificações são encontradas na literatura, assim como análises empíricas têm sido realizadas em diversas áreas das ciências sociais. Especificamente na Engenharia de Transportes, Goetzke (2003) foi um dos primeiros autores a discutir a estimação enviesada em modelos de escolha de modos de transportes que não consideram a presença de dependência espacial.

3.2 ESPECIFICAÇÃO DE MODELOS DE ESCOLHA DISCRETA ESPACIAIS

A especificação de modelos de escolha discreta espaciais não é uma tarefa simples, dado que na literatura existem poucos testes para determinar a presença de dependência espacial em modelos não lineares, além de sua aplicação não ser direta como no caso dos modelos não espaciais. Como o uso desses modelos foi primeiramente desenvolvido na área das ciências regionais, onde se analisam problemas de escolhas de uso/parcelamento do solo, análise de difusão de inovações e localização de firmas, a utilização de modelos do tipo SAR e SEM sempre tiveram maior apelo teórico. Na área de transportes o uso de modelos de escolha discreta espaciais, quando foram aplicados, também focaram em grande medida na especificação do modelo do tipo SAR (CALDEIRA; ISLER; GIANNOTTI, 2021).

De acordo com Billé e Arbia (2019) uma especificação geral para a função utilidade para um modelo binário que considera efeitos espaciais, denominado modelo aninhado não linear geral (GNNM – *General nonlinear nested model*), pode ser escrita pelas equações (25), (26) e (27).

$$\mathbf{y}^* = \rho W_{1,n} \mathbf{y}_n^* + \mathbf{X}_n \boldsymbol{\beta} + W_{2,n} \tilde{\mathbf{X}}_n \boldsymbol{\theta} + \mathbf{u}_n \quad (25)$$

$$\mathbf{u}_n = \lambda \mathbf{M}_n \mathbf{u}_n + \varepsilon_n, \quad \varepsilon_n \sim \varepsilon_n(\mathbf{0}_n, \Sigma_\varepsilon) \quad (26)$$

$$\mathbf{y}_n = \mathbb{I}_n(\mathbf{y}_n^* > \mathbf{0}_n) \quad (27)$$

em que $\mathbf{y}^*$ é o vetor coluna de variáveis dependentes contínuas e latentes de dimensão n ;

$\mathbf{y}_n$, é o vetor coluna de dimensão n de variáveis dependentes binárias definidas pela função indicadora $\mathbb{I}_n(\mathbf{y}_n^* > \mathbf{0}_n) = I(y_{n1}^* > 0), \dots, I(y_{nn}^* > 0)'$ de dimensão n ;

$\mathbf{X}_n$ é a matriz $(n \times k)$ de variáveis exógenas, incluindo o termo constante; $\tilde{\mathbf{X}}_n$ é a mesma matriz de variáveis exógenas, excluindo-se apenas o termo constante;

$\mathbf{W}_{1,n}$ e $\mathbf{W}_{2,n}$ e $\mathbf{M}_n$ são matrizes espaciais $(n \times n)$ de constantes conhecidas;

$\boldsymbol{\theta}$ é um vetor coluna de dimensão k que capta os efeitos de correlação espacial local;

$\boldsymbol{\gamma} = (\boldsymbol{\beta}', \rho, \lambda)'$ é um vetor de dimensão $k + 2$ de parâmetros com coeficientes autorregressivos ρ e λ que capturam efeitos de *spillovers* globais e com $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)'$;

ε_n é um vetor normal multivariado de inovações com média zero e variância $\sigma_\varepsilon^2 < \infty$, tal que $\Sigma_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_n$.

Visto que se assume uma distribuição normal para o termo de erro, o modelo probabilístico resultante é o Probit. Esse é um ponto importante de se destacar, pois na literatura de econometria espacial o uso do modelo Logit é bastante restrito em aplicações empíricas (CALDEIRA; ISLER; GIANNOTTI, 2021), principalmente pela evidencia apresentada por Anselin (2002) de que a derivação do modelo Logit para o caso dos modelos de escolha discreta espaciais autorregressivos é mais complicada.

Ainda, como discutido em Billé e Arbia (2019), diferentes modelos espaciais podem ser derivados a partir do modelo GNNM. Por exemplo, restringindo-se os parâmetros espaciais $\boldsymbol{\theta} = 0$, obtêm-se o modelo *Spatial Autoregressive Autoregressive* (SARAR) Probit. Cabe destacar que os autores afirmam que para o vetor $\boldsymbol{\theta}$, todos os parâmetros podem ser utilizados, como também apenas um subconjunto deles. Interessante notar que apesar dos autores derivarem os modelos SAR ($\boldsymbol{\theta} = 0$ e $\lambda = 0$) e SEM ($\boldsymbol{\theta} = 0$ e $\rho = 0$), não fazem menção ao modelo SLX em que $\lambda = 0$ e $\rho = 0$, refletindo o negligenciamento desse modelo nas análises empíricas.

Billé e Arbia (2019) também definem uma função utilidade geral para o caso multinomial, assumindo que os termos de erros seguem uma distribuição normal conforme as equações (28), (29) e (30).

$$U_{ij}^* = \rho \sum_{h=1}^n w_{1,ih} U_{hj}^* + \mathbf{x}_{ij}' \boldsymbol{\beta} + \sum_{h=1}^n w_{2,ih} \tilde{\mathbf{x}}_{hj}' \boldsymbol{\theta} + u_{ij} \quad (28)$$

$$u_{ij} = \lambda \sum_{k=1}^J m_{jk} u_{ik} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathbf{x}(0, \sigma^2_{\varepsilon_j}) \quad (29)$$

$$y_{ij} = 1 \text{ se } U^*_{ij} > U^*_{ik} \text{ para } k \neq j \quad (30)$$

$$y_{ij} = 0 \text{ caso contrário}$$

em que $i, h = 1, \dots, n$ são as observações, $j, k = 1, \dots, J$ são as alternativas do conjunto de escolha em que a j -ésima alternativa é escolhida se sua utilidade é a maior em relação a todas as outras alternativas;

U^*_{ij} é a utilidade (ou preferência) do indivíduo i para a alternativa j ;

$\mathbf{x}_{ij}$ são os regressores exógenos específicos do indivíduo com vetor de parâmetros $\boldsymbol{\beta}$;

y_{ij} é a variável multinomial observada correspondente;

$\rho \sum_{h=1}^n w_{1,ih} U^*_{hj}$ sumariza a estrutura de dependência entre a preferência de indivíduos;

$\lambda \sum_{k=1}^J m_{jk} u_{ik}$ sumariza a estrutura de dependência entre atributos não observados e alternativas selecionadas;

$\sum_{h=1}^n w_{2,ih} \tilde{\mathbf{x}}'_{hj} \boldsymbol{\theta}$ sumariza os efeitos de *spillovers* locais específicos do indivíduo; e $\sigma^2_{\varepsilon_j} = 1$ para identificação do modelo.

Os autores chamam atenção para esse modelo pois considera a existência de correlação espacial entre alternativas, além de ser interessante checar se atributos específicos do indivíduo não observados podem ser espacialmente correlacionados. Outro ponto a destacar é que os modelos Probit multinomiais não apresentam função de probabilidade fechada, sendo a estimação feita em sua maioria por simulação (TRAIN, 2009). Assim, a especificação dos modelos SAR e SEM dificultam ainda mais a estimação desses modelos.

A partir desse modelo generalizado, diversos outros também podem ser especificados, analogamente ao modelo binário. Outro ponto interessante é que os autores consideram os regressores $\mathbf{x}_{ij}$ como sendo específicos do indivíduo. No entanto, o termo $\sum_{h=1}^n w_{2,ih} \tilde{\mathbf{x}}'_{hj} \boldsymbol{\theta}$ também pode sumarizar os efeitos de *spillovers* locais advindos da dependência espacial de características específicas de uma alternativa, como tempo de viagem, custos monetários entre outros.

Na especificação de diferentes modelos a partir do generalizado, Billé e Arbia (2019) também derivam os modelos SAR ($\boldsymbol{\theta} = 0$ e $\lambda = 0$) e SEM ($\boldsymbol{\theta} = 0$ e $\rho = 0$), mas nenhuma menção é feita ao modelo SLX ($\rho = 0$ e $\lambda = 0$). Um dos motivos para os modelos do tipo SLX

não serem alvo de discussões decorre de sua especificação não implicar em problemas encontrados nos modelos autorregressivos, como a presença de heterocedasticidade e a inversão de matrizes de grande dimensão inúmeras vezes, elevando significativamente o esforço computacional (CALABRESE; ELKINK, 2014; SMIRNOV, 2010; VEGA; ELHORST, 2015). Assim, o estimador de máxima verossimilhança de modelos de escolha discreta não espaciais pode ser utilizado para estimação desses modelos (SLX) utilizando softwares já disponíveis.

Na literatura, alguns autores reivindicam que o uso dos modelos SLX são negligenciados devido ao uso pioneiro dos modelos SAR e SEM e a existência de testes de especificação como o teste de Multiplicador de Lagrange (ELHORST; VEGA, 2017; VEGA; ELHORST, 2015). Ruttenauer (2019) mostra por meio de simulação *Markov Chain Monte Carlo* (MCMC) que modelos como o SLX ou SDM podem ter um desempenho melhor que os autorregressivos no contexto linear, e por isso é importante a sua estimação. Ruttenauer (2019) também argumenta que a maioria dos trabalhos empíricos de análise de dependência espacial faz uso indiscriminado de modelos do tipo SAR e SEM sem motivações teóricas que os justifiquem, e que no caso de modelos lineares, a negligência da dependência nas variáveis independentes também é problemática, pois incorre em problemas de omissão de variável relevante, tornando os estimadores ineficientes e inconsistentes.

Além disso, outros autores como Gibbons e Overman (2012), Mcmillen (2012) e Partridge et al. (2012) argumentam que as aplicações empíricas têm apenas considerado especificações do tipo autorregressivo, e que as motivações teóricas que justificam tais proposições são fracas. O trabalho de Gibbons e Overman (2012) é o mais crítico nesse sentido, e argumenta que as aplicações empíricas devem considerar primeiro especificações do tipo SLX.

Caldeira, Isler e Giannotti (2021) discutem a aplicação de modelos de escolha discreta binários espaciais na literatura em geral e mostram que a maioria das aplicações empíricas se dão no contexto do Norte global utilizando modelos do tipo autorregressivo (SAR e SEM). Os autores analisam diferentes modelos espaciais (SAR, SEM, SLX, SDM e SDEM) no contexto de escolha de transporte público e automóvel na cidade de São Paulo, Brasil. Os resultados apontam que no contexto analisado os modelos com melhor desempenho foram os que consideraram a dependência espacial nas variáveis independentes (SLX e suas combinações) e o modelo SEM.

3.3 MATRIZ DE VIZINHANÇA E PESOS ESPACIAIS

A matriz de vizinhança é um dos temas mais importantes e controversos na literatura de econometria espacial. Como discutido por Anselin (2002), os efeitos de dependência espacial são totalmente dependentes da definição das relações entre as observações e a intensidade dessa relação. A matriz de vizinhança é uma matriz $n \times n$ em que n é o tamanho da amostra. Assim, define quais observações são vizinhas umas das outras e, a partir dessa vizinhança, são definidos os pesos que cada observação tem sobre seus vizinhos.

Na economia, o uso da matriz de vizinhança recebe muitas críticas (CORRADO; FINGLETON, 2012; GIBBONS; OVERMAN, 2012), principalmente porque na maioria dos estudos é considerada como conhecida. Nesse caso, assume-se de antemão que as relações e suas intensidades são exógenas, e não endógenas. Essa prática é comum, pois a maioria dos estudos busca analisar processos autorregressivos, e a estimação de parâmetros que definem as relações no espaço se torna complicada. Porém, outros autores como Vega e Elhorst (2015) destacam como uma das vantagens do modelo SLX a possibilidade de parametrizar a matriz de vizinhança e estimá-la com os demais parâmetros do modelo. Mais recentemente esse tema tem recebido mais atenção nos modelos de regressão linear (KELEJIAN; PIRAS, 2017) e também no caso de modelos de escolha discreta binários (ZHOU; WANG; HOLGUÍN-VERAS, 2016).

3.3.1 Critérios de vizinhança

No que se refere aos critérios de vizinhança, os mais frequentes são os de contiguidade e baseados em distância. Como a aplicação da econometria espacial é em sua maioria nas ciências regionais, as unidades de análise são representadas por regiões, e o uso do critério de contiguidade é o mais simples e mais utilizado. Assim, uma observação é considerada vizinha se compartilha uma fronteira com a observação considerada, ou se tem uma parcela de suas fronteiras em comum. Os critérios de contiguidade também são denominados por critérios do tipo *queen* (rainha) e *rock* (torre) em alusão aos movimentos do jogo de xadrez. De forma geral, atribui-se pesos binários tal que para as unidades consideradas vizinhas atribui-se peso um e para as unidades que não possuem relação atribui-se peso zero conforme a equação (31).

$$w_{ij} = \begin{cases} 1, & \text{se } i \text{ e } j \text{ são contíguos} \\ 0, & \text{caso contrário} \end{cases} \quad (31)$$

em que w_{ij} é o peso do indivíduo j sobre o indivíduo i .

Outros critérios frequentes são os baseados em distância, em que se utiliza a distância inversa entre as observações, de modo que quanto maior a distância entre as observações, menor o efeito de uma observação sobre a outra, atendendo ao que preconiza a primeira lei de Tobler. A equação (32) mostra a especificação dos pesos da matriz em função da distância inversa entre as observações.

$$w_{ij}(d_{ij}) = \begin{cases} 1/d_{ij}^\alpha, & \text{se } d_{ij} \leq \bar{d} \\ 0, & \text{se } d_{ij} > \bar{d} \end{cases} \quad (32)$$

em que $w_{ij}(d_{ij})$ é o $\langle i, j \rangle$ -ésimo elemento da matriz de vizinhança W ;

d_{ij} é a distância entre as observações i e j ;

$\bar{d}$ é a distância de corte na qual um indivíduo é considerado vizinho; e

α é um parâmetro que define a intensidade da relação entre as observações, assim que quanto maior seu valor, menor a intensidade da relação entre as observações.

Além desses critérios discutidos, outros podem ser assumidos a depender do fenômeno em análise e das premissas. Por exemplo, é possível definir distâncias entre indivíduos através de distâncias sociais, conforme realizado em estudos sobre influência social (MANESS; CIRILLO; DUGUNDJI, 2015; SIDHARTHAN *et al.*, 2011).

Além disso, para a definição do peso espacial considera-se a padronização dos pesos na linha. No caso de matrizes binárias, por exemplo, a padronização na linha da matriz de vizinhança produz uma matriz de pesos espaciais em que o efeito de cada vizinho sobre a unidade considerada é um peso médio ponderado entre todos os vizinhos daquela observação.

3.3.2 Propriedades da matriz de vizinhança

Além da decisão sobre as características das relações e intensidades entre as observações, a matriz de pesos espaciais precisa atender alguns pressupostos para que seu uso seja viável do ponto de vista da econometria. Almeida (2012) discute essas propriedades necessárias para a matriz de pesos espaciais conforme apresentado a seguir:

- $0 \leq w_{ij} \leq \infty$: pesos não negativos e finitos. A propriedade de não negatividade é importante pois, segundo Almeida (2012), não há sentido econômico entre uma relação negativa de duas regiões. A necessidade de um peso finito está associada ao uso de propriedades assintóticas dos estimadores, como a máxima verossimilhança.

- $\sum_j w_{ij} \neq 0$: impede que a matriz de vizinhança tenha observações sem vizinhos, ou seja, isoladas. A existência de observações isoladas pode prejudicar a análise principalmente se essa observação tem importância. Além disso, no caso de modelos autorregressivos, a existência de observações isoladas impede a inversão da matriz $(I - \rho W)^{-1}$ dado que se torna singular;
- $w_{ii} = 0$: convenção assumida de que uma observação não é vizinha de si mesma;
- $E(w_{ij}\varepsilon) = 0$: pesos exógenos. Como discutido anteriormente, a maioria dos estudos assume que a intensidade (peso) entre as relações é exógena, o que não pode representar a realidade, dado que o fenômeno estudado e a relação entre as observações possam ser simultaneamente dependentes.

Vale destacar que para o caso de dependência espacial nas variáveis independentes, a segunda propriedade não é necessária dado que a matriz de vizinhança não é invertida, e consequentemente problemas de singularidade não ocorrem. Ainda, para os casos em que as observações representam indivíduos, por exemplo, o pesquisador pode assumir que uma observação está isolada no espaço por indisponibilidade de dados em sua vizinhança, podendo retirá-la da análise sem maiores consequências.

3.4 ESTIMADORES

A estimação de modelos de escolha discreta espacial é desafiadora devido aos efeitos de heterocedasticidade e da necessidade de inversão da matriz de pesos espaciais em modelos autorregressivos, que impõe grandes esforços computacionais e procedimentos de estimação específicos (MCMILLEN, 1992; SMIRNOV, 2010). Na literatura encontra-se três grandes grupos de abordagens de estimação de modelos de escolha discreta espacial, quais sejam: Método de Máxima Verossimilhança (MV); Métodos Bayesianos; e Método Generalizado de Momentos (GMM em inglês) (SILVEIRA SANTOS; PROENÇA, 2019).

Entre os estimadores MV tem-se: o estimador *Expectation-Maximization* (EM), que foi o primeiro para modelos espaciais autorregressivos proposto por McMillen (1992); estimadores baseados em técnicas de simulação, como o *Recursive Importance Sampling* (RIS) (BERON; VIJVERBERG, 2004) e o *Geweke-Hajivassiliou-Keene* (GHK) (PACE; LESAGE, 2016), e métodos MV parciais (BHAT, 2011; BILLÉ; LEORATO, 2020; WANG; IGLESIAS;

WOOLDRIDGE, 2013). Além disso, os métodos de aproximação apresentados por Martinetti e Geniaux (2017) e Mozharovskiy e Vogler (2016) utilizam o estimador *Composite Maximum Likelihood* (CML) para lidar com a inversão da matriz de vizinhança e pesos espaciais.

LeSage (2000) propôs um estimador bayesiano que faz uso do procedimento de amostragem de *Gibbs*, enquanto Zhou, Wang e Holguín-Veras (2016) propuseram um estimador baseado em *Markov Chain Monte Carlo* (MCMC) para um modelo Probit espacial com uma matriz de vizinhança e pesos espaciais endógena. Finalmente, em relação aos estimadores GMM, Pinkse e Slade (1998) foram os primeiros a aplicarem-no para um modelo de escolha discreta do tipo SEM, e Klier e McMillen (2008) propuseram uma versão linearizada do estimador de Pinkse e Slade (1998) para estimação de modelos Logit. Mais recentemente, Carrión-Flores, Flores-Lagunes e Guci (2018) propuseram um estimador baseado em GMM aplicado a um modelo multinomial.

Calabrese e Elkind (2014) compararam estimadores existentes (EM, RIS, amostragem de Gibbs, GMM e GMM linearizado) por meio de simulação MCMC em modelos binários do tipo SAR, apontando suas vantagens e desvantagens.

3.5 MEDIDAS PARA AVALIAÇÃO DE POLÍTICAS DE TRANSPORTES

O uso de modelos de escolha discreta espaciais para derivação de medidas para avaliação de políticas não é análogo aos modelos não espaciais. LeSage e Pace (2009) discutem a derivação dos efeitos marginais para o modelo binário Probit do tipo SAR. A derivação dos efeitos marginais desse modelo não é direta como nos modelos não espaciais e nos modelos SEM e SLX, pois o processo autorregressivo implica em efeito de *feedback* das variáveis independentes na variável dependente, ou seja, a variável independente tem um efeito direto sobre sua vizinhança, ao passo que esse efeito retorna de forma indireta para observação inicial. Assim, os efeitos de uma variável independente podem ser caracterizados como de efeitos espaciais diretos e indiretos (LESAGE; PACE, 2009).

As aplicações empíricas analisadas mostram que a maioria dos trabalhos que usam modelos de escolha discreta espaciais utilizam efeitos marginais para discussão de análise dos efeitos das variáveis de interesse (CALDEIRA; ISLER; GIANNOTTI, 2021). Como discutido no capítulo 2, outra medida bastante utilizada para análise de políticas de transporte é o VTT. No entanto, a derivação dessa medida no contexto de modelos de escolha discreta espaciais não foi encontrada. Como o foco desta dissertação é em modelos de escolha discreta espaciais do tipo SLX, em que a derivação de medidas como elasticidades e VTT são diretas como em

modelos não espaciais, não são apresentadas discussões mais aprofundadas para os modelos autorregressivos.

3.6 APLICAÇÕES

No que tange aos trabalhos empíricos, grande parte das análises de modelos de escolha discreta espaciais têm enfoque em análises de uso/parcelamento do solo, escolha de local de residência e problemas de localização de firmas usando modelos de escolha discreta do tipo SAR e SEM (CARRIÓN-FLORES; FLORES-LAGUNES; GUCI, 2018; GUO; BHAT, 2004; MIYAMOTO *et al.*, 2004; WANG; KOCKELMAN; LEMP, 2012). No que tange a especificação de modelos SLX, os trabalhos de Alamá-Sabater, Artal-Tur e Navarro-Azorín (2011) e Autant-Bernard (2007) analisaram problemas de localização de instalações industriais.

Na área de transportes já existe quantidade considerável de trabalhos que consideram efeitos de dependência espacial usando modelos lineares em dados do tipo transversal, e em menor escala com dados em painel. Nesta dissertação as aplicações empíricas discutidas são relacionadas a modelos de escolha discreta espaciais com foco em escolha de modos de transportes.

Goetzke (2008) estimou um modelo Logit binário espacial do tipo SAR condicional para analisar a existência de efeitos de interação espacial, que o autor define como efeitos sociais de rede. O modelo caracteriza escolhas entre transporte público e automóvel para viagens a trabalho na cidade de Nova York, Estados Unidos. A matriz de vizinhança foi definida por dois critérios de vizinhança: *k*-vizinhos mais próximos e distância. Foram considerados vizinhos os 40 indivíduos mais próximos ou aqueles que estão até 1,2 km de distância.

Gotzke (2008) apresenta o resultado de três modelos que são estimados por MV: um modelo Logit sem o efeito de interação espacial; um modelo considerando o termo de interação espacial (Wy); e o último sendo o segundo modelo mais a inserção de uma variável *dummy* para controlar o acesso ao automóvel ao centro da cidade. Os resultados encontrados indicam a existência do efeito de dependência espacial (efeito social de rede) nos dois modelos espaciais, com parâmetro espacial com sinal positivo e estatisticamente significativo. Os dois modelos que consideram a presença dos efeitos de dependência têm um desempenho melhor que o modelo não espacial em termos de log-verossimilhança.

O trabalho de Gotzke (2008) foi um dos primeiros na literatura sobre o efeito de dependência espacial na análise de escolhas de modos de transportes, sendo resultado de sua tese de Doutorado (GOETZKE, 2006). O autor destaca a importância da consideração desses

efeitos no processo de planejamento de transportes, dado que o negligenciamento prejudica a previsão de demanda, assim como prejudica a proposição de políticas de transporte que levam em consideração a existência desse efeito na população.

No trabalho de Sidharthan *et al.* (2011), os autores estimaram um modelo Probit multinomial espacial do tipo SAR para escolha de modos de transportes de crianças à escola na cidade São Francisco, Estados Unidos. O modelo é estimado considerando efeitos de dependência espacial e social. Para isso, os autores consideraram dois critérios para definir a matriz de vizinhança: um geográfico e um social. Para o critério geográfico, vários testes foram realizados considerando matrizes de distância inversa, matriz de distância inversa exponencial e matriz de distância e pesos binários.

Para determinar a distância social os autores consideraram uma distância de proximidade demográfica utilizando variáveis de renda e idade. Os resultados mostraram que para todos os critérios geográficos o parâmetro espacial foi estatisticamente significativo, mas não foi para nenhuma especificação do critério social. A matriz de vizinhança com melhor resultado em termos de log-verossimilhança foi a matriz de distância inversa, utilizada para apresentar os resultados, além disso, todas as matrizes foram padronizadas na linha.

O modelo foi estimado pelo estimador *Maximum Approximate Composite Marginal Likelihood* (MACML) proposto por Bhat (2011) e os resultados indicaram que existe o efeito de dependência espacial na escolha de modos de transporte de crianças em viagens escolares, de forma que a propensão de uma criança escolher determinado modo influencia outras crianças a fazerem a mesma escolha. O modelo espacial de Sidharthan *et al.* (2011) também apresenta melhora no desempenho em termos do valor de log-verossimilhança em relação ao modelo não espacial.

Wang *et al.* (2015) desenvolveram um modelo Probit binário espacial do tipo SAR para escolha de bicicleta em viagens realizadas por estudantes e funcionários da Universidade do Estado de Ohio, Estados Unidos. Dois critérios de vizinhança foram utilizados para definir a matriz de vizinhança, ambas baseadas em critério de distância. Definiu-se um valor de corte igual a 1.000 metros. Definiu-se uma matriz com pesos binários e outra com peso espacial igual à distância inversa. A estimação foi realizada por um estimador bayesiano por simulação MCMC. Os resultados indicaram a existência de efeito de dependência espacial, com o parâmetro espacial com sinal positivo indicando a existência do efeito de dependência no uso da bicicleta, de forma que a propensão de um indivíduo utilizar esse modo de transporte influencia seus vizinhos a também usarem a bicicleta.

No contexto de viagens a trabalho e efeitos de dependência espacial, no estudo de Sheng e Sharp (2019), um modelo Probit binário espacial do tipo SAR foi estimado considerando as viagens realizadas na região metropolitana de Auckland, Nova Zelândia. As alternativas foram definidas como transporte público e transporte privado. A matriz de vizinhança foi baseada em critério de distância, sendo a distância corte para definir um vizinho igual a 1.000 metros. Os pesos foram definidos como a distância inversa e a matriz foi padronizada na linha.

Os autores encontraram que o parâmetro espacial é estatisticamente significativo e positivo, indicando que a escolha de um indivíduo pelo transporte público influencia a probabilidade de indivíduos na sua vizinhança a também escolher essa alternativa.

No Brasil foram identificados apenas dois trabalhos que incorporaram dependência espacial em modelos de demanda por transporte. O trabalho de Assirati (2018) incorporou a dependência espacial através do indicador SivarG aplicado às probabilidades de pertencimento a grupos com determinado comportamento determinado por um classificador *k-means*. Os resultados indicaram que o modelo que incorporou as variáveis espaciais teve melhor desempenho que o modelo não espacial. O trabalho de Lopes, Brondino e da Silva (2014) considerou a presença de dependência espacial em modelos de geração de viagens, onde foram estimados modelos do tipo SAR e SEM, comparados ao modelo de regressão não espacial. Os resultados indicaram que os modelos espaciais tiveram melhor desempenho.

Assim, é válido destacar que em todas as pesquisas que incorporam efeitos de dependência espacial em modelos de escolha discreta encontradas assume-se que a abordagem subjacente ao fenômeno em análise é o de maximização da utilidade, não sendo encontrado até o presente momento discussões sobre efeitos de dependência espacial em outras abordagens como a *RRM*.

4 ÁREA DE ESTUDO, DADOS E MÉTODO

Esse capítulo dedica-se à discussão da área de estudo, banco de dados e o processo de tratamento e caracterização dos atributos das alternativas. Por fim são apresentadas as especificações dos modelos.

4.1 ÁREA DE ESTUDO

A cidade de São Paulo situa-se no estado de São Paulo, Brasil, sendo a maior da América Latina em termos populacionais com cerca de 12,3 milhões de habitantes e extensão territorial de 1.521 km² (IBGE, 2021). Além disso, São Paulo é a cidade sede da Região Metropolitana de São Paulo (RMSP), formada por 39 municípios com aproximadamente 20,8 milhões de habitantes, sendo a maior região metropolitana e aglomeração urbana do Brasil (METRÔ-SP, 2019).

O sistema viário da cidade de São Paulo é complexo, com aproximadamente 20.000 km de extensão (SÃO PAULO, 2021) e cerca 6 mil cruzamentos semaforizados (CET, 2021). A frota de veículos é a maior do país, com cerca de 8,8 milhões de veículos, dos quais 67,5% são automóveis (DENATRAN, 2021).

O sistema de transporte coletivo por ônibus de São Paulo é um dos maiores do mundo, com frota de aproximadamente 14 mil ônibus em 2020 (SMT, 2021). Em termos de infraestrutura de transporte coletivo sobre pneus, a cidade conta com cerca de 131,2 km de corredores exclusivos e 500 km de faixas exclusivas (SPTRANS, 2021a), 32 terminais de integração (SPTRANS, 2021c) e aproximadamente 19,5 mil pontos de parada (SPTRANS, 2021d).

Além disso, a cidade de São Paulo conta com o Bilhete Único (BU) para pagamento da tarifa e acesso aos sistemas de transporte coletivo. O usuário que utiliza o BU tem direito a até 4 embarques por um período de 3 horas sem custo adicional, além de oferecer desconto na integração de viagens entre o sistema de ônibus e o sistema metroferroviário (SPTRANS, 2021a). A Figura 2 mostra a distribuição das faixas exclusivas e dos corredores na cidade.

Figura 2 – Mapa de corredores e faixas exclusivas de ônibus de São Paulo.

Infraestrutura — Corredor — Faixas exclusivas

Esri, HERE, Garmin, INCREMENT P, © OpenStreetMap contributors, and the GIS user community

Fonte: São Paulo (2021).

O sistema metroferroviário conta com linhas de trens suburbanos operadas pela Companhia Paulista de Trens Metropolitanos – CPTM e linhas de metrô operadas em sua maioria pela Companhia do Metropolitano de São Paulo – METRÔ, além das linhas 4-Amarela

operada por meio de Parceria Público-Privada (PPP) e 5-Lilás que são concessionadas pelo poder público para a iniciativa privada (METRÔ-SP, 2021).

As linhas de trens operadas pela CPTM contam com 271 km de extensão e 95 estações na RMSP, dos quais 137,4 km estão dentro do município de São Paulo, distribuídas em 7 linhas: Linha 7 – Rubi, Linha 8 – Diamante, Linha 9 – Esmeralda, Linha 10 – Turquesa, Linha 11 – Coral, Linha 12 – Safira e Linha 13 – Jade (CPTM, 2021). Além disso, as linhas de trens da CPTM atendem 23 municípios do Estado de São Paulo, transportando em média cerca de 3 milhões de passageiros em dias úteis (CPTM, 2021)

O Metrô, por sua vez, conta com cerca de 101,1 km de extensão no município de São Paulo, com 89 estações distribuídas em 6 Linhas: Linha 1 – Azul, Linha 2 – Verde, Linha 3 – Vermelha, Linha 4 – Amarela, Linha 5 – Lilás e Linha 15 – Prata, atendendo uma demanda diária de cerca de 5 milhões de passageiros (METRÔ-SP, 2021). A Figura 3 mostra a distribuição das linhas de trem e metrô na cidade de São Paulo.

Em termos de demanda, segundo os dados da Pesquisa Origem e Destino 2017 (OD 2017), na RMSP são realizadas cerca de 42 milhões de viagens diárias, representando um aumento de 10,3% no total de viagens em relação ao ano de 2007. Além disso, dessas viagens 36,4% são feitas por transporte coletivo, 30,9% por transporte individual motorizado, 31,8% a pé e 0,9% por bicicleta (METRÔ-SP, 2019). Vale destacar que das 42 milhões de viagens na RMSP, cerca de 57% são feitas exclusivamente na cidade de São Paulo (METRÔ-SP, 2017a), indicando a alta demanda por transportes na cidade.

Figura 3 – Mapa de linhas de trens suburbanos e metrô de São Paulo.

Fonte: São Paulo (2021).

4.2 BASE DE DADOS

Os dados utilizados nesta pesquisa são oriundos da OD 2017 realizada pelo METRO (METRÔ-SP, 2017b). A Pesquisa OD vem sendo realizada a cada dez anos desde 1967, sendo

a última no período 2017/2018, com um método de coleta de dados consolidado e que busca caracterizar os deslocamentos na RMSP. Tem como objetivo principal melhorar o planejamento dos sistemas de transporte da região, com foco principalmente nos sistemas de transporte sobre trilhos.

A pesquisa OD 2017 coleta dados do tipo PR acerca de deslocamentos realizados pelos entrevistados em um dia típico da semana (dia anterior a realização da pesquisa) e coleta informações que caracterizam o domicílio em termos socioeconômicos e os indivíduos em termos de suas características sociodemográficas. Além disso, os 39 municípios da RMSP são divididos em 517 zonas de tráfego, das quais 342 estão dentro do município de São Paulo conforme a Figura 4.

Figura 4 – Zonas de tráfego da Pesquisa OD 2017.

Fonte: Metrô-SP (2017).

A base de dados disponibilizada pela Pesquisa OD 2017 é formada a partir de uma amostra estratificada por renda da população da RMSP. Ao todo, cerca de 32 mil domicílios foram visitados somando aproximadamente 100 mil pessoas entrevistadas. Como a pesquisa não representa um censo da população, um fator de expansão é calculado para as viagens. Assim, uma observação com fator de expansão de viagem igual a 200 representa 200 viagens realizadas na RMSP. No total, 183.092 observações de viagens foram caracterizadas na

Pesquisa OD 2017, representando cerca de 42 milhões de viagens na RMSP em um dia útil da semana, a qual é utilizada nessa pesquisa.

Como já mencionado na seção 2.1.1 sobre tipos de dados para modelagem, dados de PR apresentam vantagens e desvantagens em relação aos dados de PD. Nesta seção são discutidos os procedimentos representados no fluxograma da Figura 5 para obter o banco de dados utilizado na estimação dos modelos propostos. Esta etapa é caracterizada pela eliminação de observações que inviabilizam a estimação dos modelos e pela caracterização dos atributos de nível de serviço.

Figura 5 – Processo de tratamento dos dados da OD 2017.

Fonte: O autor (2021).

4.2.1 Motivo de viagem

Como o objetivo desta pesquisa é analisar viagens por motivo trabalho, filtrou-se a variável “motivo no destino” para as categorias: “Trabalho Indústria”, “Trabalho Comércio” e “Trabalho Serviços”. Nessa etapa 76,17% das observações foram eliminadas, representando outros motivos de viagem (“Escola/Educação”, “Compras”, “Médico/Dentista/Saúde”, “Recreação/Visitas/Lazer”, “Residência”, “Procurar Emprego”, “Assunto Pessoais”, e “Refeição”), resultando em 39.631 observações de viagens por motivo trabalho, representando 23,82% do total de viagens na RMSP.

4.2.2 Conjunto de alternativas

A definição do conjunto de escolha de modos de transporte para viagens a trabalho é uma etapa importante para a modelagem pois define as alternativas disponíveis aos indivíduos. Nesta pesquisa optou-se por considerar apenas modos motorizados, sendo eles: Ônibus (*BUS*), equivalentes aos modos principais “Ônibus/micro-ônibus/perua do município de São Paulo”, “Ônibus/micro-ônibus/perua de outros municípios” e “Ônibus/micro-ônibus/perua metropolitano” representando cerca de 23,54% do total de viagens a trabalho; Trilhos (*RAIL*), equivalente aos modos “Metrô”, “Trem”, representando 15,52% do total de viagens por motivo trabalho; e Automóvel (*CAR*), equivalente ao modo “Dirigindo Automóvel”, representando 24,72% do total de viagens a trabalho na RMSP.

A alternativa Motocicleta foi desconsiderada por dois motivos: dificuldade de caracterização adequada dos atributos, como o tempo e custo de viagem; e pequena quantidade de viagens amostradas na pesquisa, representando 4,18% das viagens a trabalho na RMSP.

As alternativas de modos ativos (a pé e bicicleta) foram desconsideradas pois acredita-se que a definição de disponibilidade de tais modos para toda a amostra é complexa e depende de diversos fatores subjetivos, implicando na necessidade de definição de medidas como limite máximo de distância, por exemplo, para que sejam consideradas pelos indivíduos, e que não necessariamente representam a realidade.

A caracterização do modo escolhido pelo indivíduo foi feita com base na variável “Modo Principal” estabelecida na OD 2017, que indica o modo de transporte com maior capacidade utilizado na viagem. Assim, uma pessoa que tenha utilizado os modos “Ônibus”, “Bicicleta” e “Metrô” é classificada como tendo realizado uma viagem pelo modo “Metrô”.

Nesta etapa foram eliminadas 37,98% das observações da amostra resultante da primeira filtragem, constituídas pelos demais modos (“Dirigindo Motocicleta”, “Transporte Fretado”, “Transporte Escolar”, “Passageiro de automóvel”, “Táxi Convencional”, “Táxi não Convencional”, “Passageiro de moto”, “Bicicleta”, “A pé” e “Outros”), resultando em 24.578 observações, representando 15,19% do total de viagens na RMSP.

A definição da disponibilidade das alternativas é uma etapa importante pois estabelece o real conjunto de escolha do indivíduo, que é uma etapa sensível na abordagem *RRM*. A fim de evitar a presença de indivíduos com apenas duas alternativas (*BUS* e *RAIL*) disponíveis (veja discussão das seções 2.2 e 2.2.4), apenas os indivíduos com automóvel no domicílio foram selecionados. Assim, para a alternativa *CAR* considerou-se que o modo está disponível se a

quantidade de automóveis na residência é maior que zero ou se o indivíduo realizou a viagem pelo modo “Dirigindo Automóvel”. Essa premissa pode não ser necessariamente realista, como por exemplo os domicílios que possuem apenas um automóvel, mas que este não está disponível para todos os membros da família. Além disso, no Brasil a permissão para dirigir é condicionada a indivíduos que tenham mais de 18 anos e possuam habilitação para dirigir, sendo essa última informação indisponível na OD 2017. Para as alternativas *BUS* e *RAIL* foi considerado que ambas estão disponíveis para todos os indivíduos da amostra.

4.2.3 Viagens no município de São Paulo

A aplicação desse filtro tem como objetivo selecionar as viagens que foram iniciadas e terminaram na cidade de São Paulo. A aplicação desse procedimento foi necessária pois os tempos de viagens para as alternativas *BUS* e *RAIL* não são simples de serem caracterizados fora dos limites da cidade em termos de oferta e estrutura tarifária (ausência de dados de oferta e diferentes políticas tarifárias). Apesar da escala metropolitana ser importante na análise dos sistemas de transportes e do estudo da mobilidade urbana, acredita-se que tal limitação não prejudica a análise proposta. Nesta etapa foram selecionadas 16.259 observações, representando 8,69% do total de viagens na RMSP.

4.2.4 Atributos sociodemográficos

Para este procedimento, definiu-se um limite inferior de idade, considerando-se apenas indivíduos com 16 anos ou mais com o objetivo de eliminar observações com idade abaixo da idade mínima para trabalhar. Além disso, foram selecionados os indivíduos que realizaram viagem pelo modo “Dirigindo Automóvel” ou tinham pelo menos um automóvel no domicílio. Para estes e os demais atributos sociodemográficos utilizados na modelagem foram removidos os indivíduos com valores faltantes.

A eliminação de observações duplicadas foi realizada dadas as limitações associadas a aplicação das técnicas de econometria espacial. A Pesquisa OD 2017 coleta todas as viagens realizadas por um indivíduo em um dia útil. Assim, indivíduos que realizam mais de uma viagem com motivo trabalho por dia apresentam mais de uma observação no banco de dados.

Uma das premissas adotadas na econometria espacial é de que um indivíduo não é vizinho de primeira ordem de si mesmo (ANSELIN, 1988) e, para não violar tal premissa, apenas uma viagem do mesmo indivíduo pode ser representada geograficamente na amostra.

Assim, para os indivíduos com mais de uma viagem na amostra, foi selecionado o primeiro registro para esta observação, eliminando-se o restante das observações associadas. Apesar de qualquer outro critério pudesse ser utilizado, este foi utilizado dado que em muitas situações a primeira viagem tem maior probabilidade de coincidir com o domicílio do indivíduo. O banco de dados desta etapa resultou em 10.352 observações, representando 4,99% do total de viagens e 20,94% daquelas por motivo trabalho na RMSP.

4.3 CARACTERIZAÇÃO DE ATRIBUTOS DE NÍVEL DE SERVIÇO

Dado que as informações reportadas pelos indivíduos entrevistados na OD 2017 referem-se apenas aos modos utilizados nas viagens realizadas, não é possível identificar precisamente as alternativas que estavam disponíveis para escolha e os atributos de cada uma delas. Além disso, os atributos associados às viagens reportadas na base de dados são aqueles percebidos pelos indivíduos e não se referem a medidas objetivas desses atributos.

Neste contexto, Koppelman e Bhat (2006) recomendam que o procedimento mais adequado é modelar todos os atributos de nível de serviço relacionados às alternativas, inclusive aquelas escolhidas de fato pelos indivíduos, estabelecendo as disponibilidades das alternativas a partir de regras de acesso ao modo de transporte, como por exemplo, distância máxima de acesso aos terminais de transporte público e/ou posse de habilitação para dirigir automóveis.

Os níveis dos atributos calculados nessa etapa foram tempo de viagem, custo monetário e distâncias de acesso, egresso e total. A Figura 6 ilustra o procedimento de caracterização dos níveis dos atributos de nível de serviço.

Figura 6 – Fluxograma de caracterização de atributos de nível de serviço

Fonte: O Autor (2022).

4.3.1 Nível dos atributos

Atualmente, sensores e dispositivos móveis têm sido utilizados para o levantamento e caracterização do deslocamento de veículos em áreas urbanas, cujos dados transmitidos podem ser utilizados para fins de modelagem e planejamento de transportes. Para o cálculo dos níveis dos atributos específicos das alternativas foram utilizadas duas Interfaces de Programação de Aplicações (*Application Programming Interface* - API) de roteamento de viagens.

Uma API é um conjunto de rotinas e padrões de programação para acesso a um aplicativo de software ou plataforma baseada na Web (GRUS, 2016). De forma geral, uma API é desenvolvida quando uma empresa deseja que outros criadores de softwares desenvolvam produtos associados a seus serviços. Neste trabalho foram utilizadas a API de roteamento multimodal OTP (*Open Trip Planner*) e a de roteamento da empresa TomTom.

O OTP é uma API de roteamento multimodal desenvolvida em *Java* e tem como vantagem o uso livre (OTP, 2020). Para tanto, a API utiliza uma base de dados de rede do *OpenStreetMap* (OpenStreetMap, 2021) e faz o cálculo de rotas de viagem por diversos modos de transporte. Para os modos individuais o cálculo é feito pelas rotas de menor custo e para os modos coletivos o roteamento é feito com base em dados que seguem o padrão GTFS (*General Transit Feed Specification*), que foram disponibilizados pela SPTrans (SPTrans, 2017).

O uso do OTP para cálculo do tempo de viagem tem sido recorrente em estudos de acessibilidade (PEREIRA, 2018; PEREIRA *et al.*, 2019; PRITCHARD *et al.*, 2019; SLOVIC *et al.*, 2019) e também para estudo de comportamento relacionado a viagens utilizando modelos de escolha discreta (MARCUS; BLAINEY, 2016). Neste trabalho o OTP foi utilizado para o cálculo do tempo de viagem por ônibus, trilhos e as distâncias de acesso, egresso e total.

A segunda API utilizada é comercializada pela empresa TomTom, tal que o cliente/usuário precisa fazer um requerimento e a API disponibiliza um conjunto de informações acerca da viagem requisitada. O tempo de viagem por automóvel foi calculado com apoio dessa API utilizando sua versão gratuita. Para caracterização das rotas utilizou-se as origens e destinos, a data de realização e o horário de partida das viagens.

4.3.2 Distâncias de acesso, egresso e total

A partir dos dados georreferenciados determinou-se as estações de trem/metrô mais próximas da origem e destino das viagens. Para isso, realizou-se uma operação espacial (*nearest object*) a partir dos dados das estações de trem/metrô disponibilizados na rede do GTFS para o ano de 2017. É importante salientar que essa operação considera a distância euclidiana e não a distância em rede entre as coordenadas do indivíduo e as estações.

Após a caracterização das estações mais próximas da origem e do destino, as distâncias de acesso, egresso e total foram calculadas pela API do OTP. Além disso, nesta etapa foram eliminadas as observações que apresentaram a mesma estação de origem e destino, dado que nessa situação a alternativa *RAIL* não é considerada pelo indivíduo. O banco de dados dessa etapa resultou em 9.380 observações, representando 4,40% do total de viagens e 18,50% do total de viagens por motivo trabalho na RMSP.

Para as distâncias de acesso, egresso e total das viagens não roteadas pela API, tais atributos foram definidos pela multiplicação de um fator de correção, medido pela razão entre a distância total em rede e a distância total euclidiana das viagens roteadas pela API, conforme a equação (33), e sua respectiva distância euclidiana, conforme equações (34), (35) e (36) para as distâncias de acesso, egresso e total respectivamente. Todas as medidas têm unidade em metros.

$$C_{CORR} = \frac{d_{TOTrede}}{d_{TOTeuclidiana}} \quad (33)$$

$$d_{esto} = d_{esto_{euclidiana}} \cdot C_{CORR} \quad (34)$$

$$d_{estd} = d_{estd_{euclidiana}} \cdot C_{CORR} \quad (35)$$

$$d_{total} = d_{total_{euclidiana}} \cdot C_{CORR} \quad (36)$$

em que C_{CORR} é o fator de correção;

$d_{TOT_{rede}}$ é a distância total em rede das viagens roteadas pelo OTP;

$d_{TOT_{euclidiana}}$ é a distância euclidiana total das viagens roteadas pelo OTP;

d_{esto} é a distância da origem até a estação mais próxima da origem calculada para as viagens não roteadas pelo OTP;

$d_{esto_{euclidiana}}$ é a distância euclidiana da origem até a estação mais próxima da origem das viagens não roteadas pelo OTP;

d_{estd} é a distância da estação mais próxima do destino até o destino calculada para as viagens não roteadas pelo OTP;

$d_{estd_{euclidiana}}$ é a distância euclidiana da origem até a estação mais próxima da origem das viagens não roteadas pelo OTP;

d_{total} é a distância total calculada para as viagens não roteadas pelo OTP; e

$d_{total_{euclidiana}}$ é a distância euclidiana total das viagens não roteadas pelo OTP.

4.3.3 Tempo e custo de viagem *BUS*

O tempo de viagem para a alternativa *BUS* foi calculado pela API do OTP, caracterizado pelas coordenadas de origem e destino da viagem. Dentre as opções de modo de transporte disponíveis na API (*transit*, *walk*, *bicycle*, *car*, *bus*, *rail*, ou *transit* e *bicycle*) foi selecionada a opção *bus*, que considera o trecho de caminhada até o ponto de ônibus mais próximo e a rota exclusivamente pelo sistema de ônibus até o destino, também considerando o tempo de caminhada no destino.

Para as viagens cujos tempos não foram obtidas pelo OTP, este atributo foi calculado a partir da velocidade média, definida pela razão entre a distância total em rede e o tempo de viagem por ônibus das viagens roteadas pela API, conforme a equação (37), e sua respectiva distância total, conforme a equação (38).

$$VM_{BUS} = \frac{d_{TOT_{ROTEADO}}}{tt_{BUS_{ROTEADO}}} \quad (37)$$

$$tt_{BUS} = \frac{d_{total}}{VM_{BUS}} \quad (38)$$

em que VM_{BUS} é a velocidade média da alternativa *BUS*;

$d_{TOT_{ROTEADO}}$ é a distância total das observações roteadas pelo OTP;
 $tt_{BUS_{ROTEADO}}$ é o tempo por ônibus das observações roteadas pelo OTP;
 d_{total} é a distância total para as observações não roteadas pelo OTP; e
 tt_{BUS} é o tempo por ônibus calculado para as observações não roteadas pelo OTP.

Por fim, foram eliminadas as viagens com tempo de viagem superior a 300 minutos. Quanto ao custo da alternativa *BUS*, dado que o BU garante integração temporal com até quatro embarques no sistema de ônibus, definiu-se o valor unitário da tarifa no período de realização da OD 2017 de R\$3,80 (SPTRANS, 2021a), sendo esse valor constante para toda a amostra.

4.3.4 Tempo e custo de viagem *RAIL*

Para determinar o tempo de viagem para a alternativa *RAIL* foram calculados tempos para diferentes trechos. Esse procedimento foi utilizado pois não é possível inferir a forma de acesso às estações para os indivíduos que não escolheram a alternativa *RAIL*. Calculou-se o tempo a pé da origem até a estação de trem/metrô mais próxima e da estação de trem/metrô mais próxima do destino ao destino; o tempo de ônibus da origem até a estação de trem/metrô mais próxima da origem e da estação de trem/metrô mais próxima do destino e o destino; e o tempo por trem/metrô da estação mais próxima da origem até a estação mais próxima do destino.

Para identificar a forma de acesso às estações, definiu-se um limite de distância em rede baseado nas distâncias dos indivíduos que escolheram a alternativa *RAIL*. Assim, para esses indivíduos que acessaram a sua respectiva estação a pé, calculou-se o 85º percentil de distância, que resultou em um valor igual a 1.316 metros e que foi aproximado para 1.300 metros. Assim adotou-se o limite máximo no qual o indivíduo acessa a estação de trem/metrô a pé tanto na origem quanto no destino. Para as distâncias superiores a esses limites, considerou-se o tempo por ônibus. O tempo de viagem para a alternativa *RAIL*, então, foi calculado conforme a equação (39).

$$tt_{RAIL} = \begin{cases} tt_{ape_o_esto} + tt_{esto_estd} + tt_{ape_estd_d}, & dist_{esto} \leq 1300 \text{ e } dist_{estd} \leq 1300 \\ tt_{ape_o_esto} + tt_{esto_estd} + tt_{bus_estd_d}, & dis_{esto} \leq 1300 \text{ e } dist_{estd} > 1300 \\ tt_{bus_o_esto} + tt_{esto_estd} + tt_{ape_estd_d}, & dis_{esto} > 1300 \text{ e } dist_{estd} \leq 1300 \\ tt_{bus_o_esto} + tt_{esto_estd} + tt_{ape_estd_d}, & dis_{esto} > 1300 \text{ e } dist_{estd} > 1300 \end{cases} \quad (39)$$

em que tt_{RAIL} é o tempo de viagem para a alternativa *RAIL*;

$tt_{ape_o_esto}$ é o tempo de viagem a pé da origem até a estação de trem/metrô mais próxima da origem

$tt_{bus_o_esto}$ é o tempo viagem de ônibus da origem até a estação de trem/metrô mais próxima da origem;

tt_{esto_estd} é o tempo de viagem da estação de trem/metro mais próxima da origem até a estação de trem/metrô mais próxima do destino;

$tt_{ape_estd_d}$ é o tempo de viagem a pé da estação de trem/metrô mais próxima do destino até o destino;

$tt_{bus_estd_d}$ é o tempo viagem de ônibus da estação de trem/metrô mais próxima do destino até o destino;

$dist_{esto}$ é a distância da origem até a estação de trem/metrô mais próxima da origem; e

$dist_{estd}$ é a distância da estação de trem/metrô mais próxima do destino até o destino.

No OTP, para o cálculo dos componentes $tt_{ape_o_esto}$ e $tt_{ape_estd_d}$ foi selecionado o modo *walk*, para os componentes $tt_{bus_o_esto}$ e $tt_{bus_estd_d}$ foi selecionada a opção *bus* e para o componente tt_{esto_estd} utilizou-se a opção *transit*. Para cada trecho de viagem cujos tempos a pé, por ônibus e trem que não foram roteados pelo OTP, tais atributos foram calculados a partir da velocidade média, definida pela razão entre a distância total em rede e o tempo de viagem para cada trecho e cada modo roteado pela API, conforme procedimento anterior. Por fim, foram eliminadas as viagens com tempo superior a 300 minutos.

O custo de viagem da alternativa *RAIL* foi estimado em duas etapas. Para os indivíduos que escolheram a alternativa *RAIL* são conhecidos todos os modos de transporte utilizados na viagem, sendo possível inferir um custo aproximado. Para os indivíduos que escolheram as alternativas *BUS* ou *CAR* o custo foi inferido utilizando o critério de distância limite de acesso pelo modo a pé discutido anteriormente para definição de tarifas de integração. Assim, para o indivíduo a uma distância maior que 1.300 metros da estação tanto na origem quanto no destino, considerou-se o custo como o valor de integração entre o sistema de ônibus e o sistema metroferroviário no valor de R\$6,80 no período de realização da OD 2017 (METRÔ-SP, 2017a). Para indivíduos com distância até a estação na origem ou no destino menor que 1.300 metros definiu-se o custo como o valor de uma tarifa igual a R\$3,80 no período de realização da OD 2017, conforme indicado na equação (40).

$$CO_{RAIL} = \begin{cases} 3,80, & dist_{est} \leq 1300m \\ 6,80, & dist_{est} > 1300m \end{cases} \quad (40)$$

em que CO_{RAIL} é o custo da viagem pelo modo trilhos; e

$dist_{est}$ é a distância de acesso à estação mais próxima da origem ou distância de egresso da estação mais próxima do destino.

4.3.5 Tempo e custo de viagem CAR

O tempo de viagem para a alternativa CAR foi calculado pela API da TomTom e caracterizado pelas coordenadas de origem e destino, data de realização e horário de partida da viagem. Para as viagens cujos tempos não foram roteados pela API da TomTom, este atributo foi calculado a partir da velocidade média e a distância total da viagem, de forma similar aos atributos já discutidos anteriormente.

O custo da alternativa Automóvel foi estimado com base em Gomide e Morato (2011), corrigido para os valores do ano de realização da Pesquisa OD 2017 com base no IPCA (Índice de Preços ao Consumidor Amplo) (BANCO CENTRAL DO BRASIL, 2021). Esse custo é estimado com base nos dados levantados pela Associação Nacional dos Transportes Públicos (ANTP) para as 27 capitais brasileiras e os 16 municípios que, à época, tinham mais de 500 mil habitantes, sendo composto por um componente fixo que compreende manutenção, depreciação e tributos, e um componente variável que inclui combustível e estacionamento (GOMIDE; MORATO, 2011).

Apesar deste custo não representar o custo direto percebido pelo usuário dessa alternativa como é comum em estudos de demanda, acredita-se que esse método atende ao propósito desta pesquisa. Além disso, essa métrica também foi utilizada por Dias (2013) para aplicação de um modelo de escolha discreta para análise de políticas tarifárias na cidade de São Paulo usando dados da Pesquisa Origem e Destino de 2007.

A partir da correção, o custo de viagem da alternativa CAR foi estimado conforme a equação (41).

$$CO_{AUTO} = 2,88 + \left(\frac{2,31 \cdot dist_{tot}}{7000} \right) \quad (41)$$

em que CO_{auto} é o custo médio de viagem por automóvel; e

$dist_{tot}$ é a distância entre a origem e destino em metros.

4.4 ATRIBUTOS PARA MODELAGEM

Os atributos de nível de serviço considerados na modelagem para a alternativa *CAR* foram tempo de viagem (TT_{CAR}) e custo monetário (CO_{CAR}) e para as alternativas *BUS* e *RAIL* foram considerados o tempo de viagem (TT_{BUS} e TT_{RAIL} , respectivamente) e custo monetário na forma de tarifas (CO_{BUS} e CO_{RAIL}).

Os atributos sociodemográficos considerados foram: idade (*IDADE*), sexo (*SEXO*), com a categoria “masculino” como referência; número de moradores da família (*HOUSE*); número total de viagens realizadas pelo indivíduo (*TRIPS*); classe social em três categorias, Classe A (referência), Classe B e Classe C/D/E; grau de instrução em três categorias, até Fundamental II incompleto como referência, Fundamental II completo até Superior incompleto (GI_2) e Superior completo (GI_3); e se o indivíduo iniciou a viagem no horário de pico da manhã ou não (*PICO*), com a categoria não adotada como referência.

O processo de tratamento dos dados e caracterização dos atributos resultou em um banco de dados composto por 9.335 observações, representando cerca de 18,36% do total de viagens por motivo trabalho na RMSP. A Tabela 1 apresenta as estatísticas descritivas dos atributos considerados, enquanto a Figura 7 mostra a espacialização das observações da amostra utilizada e suas respectivas escolhas entre as alternativas.

Tabela 1 – Estatística descritiva dos atributos utilizados para estimação.

Atributo	Unidade	Média	Desvio Padrão	Mínimo	Máximo
Modo de Transporte					
<i>BUS</i>	%	21,26	-	-	-
<i>RAIL</i>	%	23,15	-	-	-
<i>CAR</i>	%	55,57	-	-	-
<i>TT_{BUS}</i>	minutos	64.75	32.84	0.75	276.35
<i>TT_{RAIL}</i>	minutos	86.50	38.71	14.49	334.27
<i>TT_{CAR}</i>	minutos	22.13	12.56	1.13	107.68
<i>CO_{BUS}</i>	R\$	3,80	3,80	3,80	0
<i>CO_{RAIL}</i>	R\$	6,24	1,17	3,80	6,80
<i>CO_{CAR}</i>	R\$	6,43	2,76	2,98	25,31
<i>IDADE</i>	anos	42,79	13,73	16	98
<i>CLASSE_A</i>	%	30	-	-	-
<i>CLASSE_B</i>	%	56	-	-	-
<i>CLASSE_{C/D/E}</i>	%	14	-	-	-
<i>SEXO</i>	%	45	-	-	-
<i>HOUSE</i>	Moradores	3,13	1,28	1	12
<i>TRIPS</i>	viagens	2,90	1,37	1	17
<i>GI₁</i>	%	9	-	-	-
<i>GI₂</i>	%	29	-	-	-
<i>GI₃</i>	%	62	-	-	-
<i>PICO</i>	%	56	-	-	-

Figura 7 – Escolha das alternativas da amostra de modelagem.

Esri, HERE, Garmin, INCREMENT P, © OpenStreetMap contributors, and the GIS user community

Fonte: O autor (2021).

4.5 MÉTODO

4.5.1 Especificação da matriz de vizinhança e pesos espaciais

Neste trabalho foram especificadas matrizes de vizinhança e pesos espaciais a partir de dois critérios de vizinhança. O primeiro é baseado em um limite de distância entre a unidade considerada e seus vizinhos, com pesos binários conforme a equação (42). O segundo é baseado no k -ésimo vizinho mais próximo (knn – *k-nearest neighbor*), também com pesos binários conforme a equação (43).

$$w_{ij}(d_{ij}) = \begin{cases} 1, & \text{se } d_{ij} \leq d_{max} \\ 0, & \text{caso contrário} \end{cases} \quad (42)$$

$$w_{ij}(N) = \begin{cases} 1, & \text{se } knn \leq N \\ 0, & \text{caso contrário} \end{cases} \quad (43)$$

em que w_{ij} é o $\langle i, j \rangle$ -ésimo elemento da matriz de vizinhança W ;

d_{ij} é a distância entre as observações; e

d_{max} é a distância máxima que um indivíduo é considerado vizinho;

knn é o número de vizinhos mais próximos do indivíduo i ; e

N é o número máximo de indivíduos considerados vizinhos.

Para analisar a robustez das matrizes adotadas variou-se o limite de distância entre 1.000 metros e 2.000 metros em intervalos de 100 metros para as matrizes baseadas nesse critério e para as matrizes baseadas em knn variou-se o número de vizinhos entre 10 e 100 em intervalos de 10 vizinhos.

Dessa forma, é possível avaliar o impacto de diferentes valores limites sobre a estabilidade do efeito de dependência espacial nos modelos. A definição dos valores máximos e mínimos seguiu como referência aqueles mais recorrentes encontrados na literatura e discutidos na seção 3.6. No caso de pesquisas de comportamento relacionado a viagens, o limite de 1.000 metros é o mais utilizado, enquanto a única aplicação do critério knn encontrado foi o trabalho de Goetzke (2008), que adotou o valor k igual 40.

As matrizes foram padronizadas na linha, ou seja, tiveram a soma na linha igualadas a um, sendo este o método mais utilizado na literatura e também o mais simples, garantindo as propriedades desejadas e discutidas na seção 3.3.2.

Cabe destacar que casos de observações isoladas podem ocorrer para as matrizes de vizinhança baseadas em critérios de distância. No entanto, os modelos do tipo *SLX* não

envolvem inversão de matrizes e, portanto, não incorrem em problemas de singularidade observados nos modelos autorregressivos. Nesta dissertação, 18 observações foram retiradas da amostra por não terem pelo menos um vizinho a 1.000 metros de distância, representando menos de 0,002% da amostra. Esse procedimento foi adotado para permitir que trabalhos futuros que queiram adotar especificações autorregressivas possam realizar uma comparação com os modelos propostos nesta dissertação. A discussão na seção 4.4 sobre o banco de dados final utilizado para modelagem já apresenta o resultado com a eliminação dessas observações.

É importante salientar também que poderiam ser utilizadas diferentes matrizes de vizinhança e pesos espaciais tanto para os atributos de um modo quanto para atributos entre os diferentes modos. Como destacado no capítulo 3, a aplicação de modelos de escolha discreta espacial ainda é pequena, e como não foram encontradas evidências de que diferentes relações de vizinhanças são significativas para diferentes modos de transporte ou diferentes atributos, as mesmas matrizes para todos os modos e atributos foram utilizadas nos modelos propostos.

4.6 ESPECIFICAÇÃO DOS MODELOS

Modelos Logit foram estimados para avaliar o comportamento de viagens a trabalho na cidade de São Paulo sob as abordagens *RUM* e *RRM*, com e sem a consideração de efeitos de dependência espacial nas variáveis independentes.

Foram estimados modelos sob as abordagens *RUM* e *RRM* com apenas variáveis de nível de serviço (*LOS*), com e sem efeitos de dependência espacial nas variáveis independentes (atributos de nível de serviço apenas). Na sequência, modelos *RUM* e *RRM* foram estimados com atributos de nível de serviço mais os atributos sociodemográficos (*LOS + SOCIO*), com e sem efeitos de dependência espacial (atributos de nível de serviço apenas).

A equação de probabilidade do modelo Logit sob ambas as abordagens foi discutida no capítulo 2 e não é apresentada aqui, sendo discutidas somente as especificações das funções utilidade e arrependimento. O teste de razão de verossimilhança foi utilizado para especificação do modelo *RUM* e os modelos *RRM* utilizaram a mesma especificação para fins de comparação.

4.6.1 Modelos *RUM* não espaciais

As especificações das funções utilidades para as alternativas *CAR*, *BUS* e *RAIL* para o modelo sob a abordagem *RUM* com apenas atributos de nível de serviço (RUM_{LOS}) são apresentadas nas equações (44), (45) e (46) respectivamente. Para simplificação de leitura, o

subscrito n referente ao indivíduo não é apresentado em nenhuma função utilidade ou arrependimento no restante do texto.

$$U_{CAR} = \beta_{TT_{CAR}} * TT_{CAR} + \beta_{CO_{CAR}} * CO_{CAR} \quad (44)$$

$$U_{BUS} = ASC_{BUS} + \beta_{TT_{BUS}} * TT_{BUS} + \beta_{CO_{GEN}} * CO_{BUS} \quad (45)$$

$$U_{RAIL} = ASC_{RAIL} + \beta_{TT_{RAIL}} * TT_{RAIL} + \beta_{CO_{GEN}} * CO_{RAIL} \quad (46)$$

As especificações das funções utilidade para as alternativas CAR , BUS e $RAIL$ para o modelo sob a abordagem RUM com atributos de nível de serviço e sociodemográficos ($RUM_{LOS+SOCIO}$) são apresentadas nas equações (47), (48) e (49) respectivamente.

$$U_{CAR} = \beta_{TT_{CAR}} * TT_{CAR} + \beta_{CO_{CAR}} * CO_{CAR} \quad (47)$$

$$\begin{aligned} U_{BUS} = & ASC_{BUS} + \beta_{TT_{BUS}} * TT_{BUS} + \beta_{CO_{GEN}} * CO_{BUS} + \beta_{IDADE} * IDADE \\ & + \beta_{GI_2} * GI_2 + \beta_{GI_3} * GI_3 + \beta_{SEX_{BUS}} * SEXO + \beta_{CLASSE_B} \\ & * CLASSE_B + \beta_{CLASSE_{C/D/E}} * CLASSE_{C/D/E} + \beta_{HOUSE} * HOUSE \\ & + \beta_{TRIP_{BUS}} * TRIP + \beta_{PICO} * PICO \end{aligned} \quad (48)$$

$$\begin{aligned} U_{RAIL} = & ASC_{RAIL} + \beta_{TT_{RAIL}} * TT_{RAIL} + \beta_{CO_{GEN}} * CO_{RAIL} + \beta_{IDADE} * IDADE \\ & + \beta_{GI_2} * GI_2 + \beta_{GI_3} * GI_3 + \beta_{SEX_{RAIL}} * SEXO + \beta_{CLASSE_B} \\ & * CLASSE_B + \beta_{CLASSE_{C/D/E}} * CLASSE_{C/D/E} + \beta_{HOUSE} * HOUSE \\ & + \beta_{TRIP_{RAIL}} * TRIP + \beta_{PICO} * PICO \end{aligned} \quad (49)$$

4.6.2 Modelos RRM e RMU não espaciais

A especificação da função arrependimento para as alternativas CAR , BUS e $RAIL$ para o modelo sob a abordagem RRM com atributos de nível de serviço (RRM_{LOS}) é apresentada nas equações (50), (51) e (52), respectivamente.

A especificação das constantes específicas das alternativas (ASC_{BUS} e ASC_{RAIL}) nos modelos RRM é baseada na discussão de Hess, Beck e Chorus (2014), que apontam que a consideração da constante analogamente à abordagem RUM (adição linear) ou na sua forma não linear $= -\ln(1 + \exp(ASC_{BUS} - ASC_{CAR}))$ não tem impacto no desempenho do modelo,

ficando a critério do analista a sua especificação. Nesta dissertação a especificação linear é utilizada.

$$\begin{aligned}
 R_{CAR} = & -\ln(1 + \exp(\beta_{TT_{BUS}} * TT_{BUS} - \beta_{TT_{CAR}} * TT_{CAR})) \\
 & - \ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{CAR}} * TT_{CAR})) \\
 & - \ln(1 + \exp(\beta_{CO_{GEN}} * CO_{BUS} - \beta_{CO_{CAR}} * CO_{CAR})) \\
 & - \ln(1 + \exp(\beta_{CO_{GEN}} * CO_{RAIL} - \beta_{CO_{CAR}} * CO_{CAR}))
 \end{aligned} \tag{50}$$

$$\begin{aligned}
 R_{BUS} = & ASC_{BUS} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{BUS})) \\
 & - \ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{BUS}} * TT_{BUS})) \\
 & - \ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{BUS})) \\
 & - \ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{RAIL} - CO_{BUS})))
 \end{aligned} \tag{51}$$

$$\begin{aligned}
 R_{RAIL} = & ASC_{RAIL} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{BUS})) \\
 & - \ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{BUS}} * TT_{BUS})) \\
 & - \ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{RAIL})) \\
 & - \ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{BUS} - CO_{RAIL})))
 \end{aligned} \tag{52}$$

A especificação do modelo *RMU* utilizada nesta dissertação é baseada na formulação de Chorus (2010) (modelo *CRRM*) e Chorus, Rose e Hensher (2013) que contempla heterogeneidade de estratégias de decisão no processamento de atributos, ou seja, um modelo híbrido *RUM-RRM* discutido na seção 2.2.2. As funções utilidade modificadas para as alternativas *CAR*, *BUS* e *RAIL* são apresentadas nas equações (53), (54) e (55) respectivamente.

$$\begin{aligned}
 R_{CAR} = & -\ln(1 + \exp(\beta_{TT_{BUS}} * TT_{BUS} - \beta_{TT_{CAR}} * TT_{CAR})) \\
 & - \ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{CAR}} * TT_{CAR})) \\
 & - \ln(1 + \exp(\beta_{CO_{GEN}} * CO_{BUS} - \beta_{CO_{CAR}} * CO_{CAR})) \\
 & - \ln(1 + \exp(\beta_{CO_{GEN}} * CO_{RAIL} - \beta_{CO_{CAR}} * CO_{CAR}))
 \end{aligned} \tag{53}$$

$$\begin{aligned}
R_{BUS} = & ASC_{BUS} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& - \ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& - \ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{BUS})) \\
& - \ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{RAIL} - CO_{BUS}))) + \beta_{IDADE} \\
& * IDADE + \beta_{GI_2} * GI_2 + \beta_{GI_3} * GI_3 + \beta_{SEXO_{BUS}} * SEXO \\
& + \beta_{CLASSE_B} * CLASSE_B + \beta_{CLASSE_{C/D/E}} * CLASSE_{C/D/E} \\
& + \beta_{HOUSE} * HOUSE + \beta_{TRIP_{BUS}} * TRIP + \beta_{PICO} * PICO
\end{aligned} \tag{54}$$

$$\begin{aligned}
R_{RAIL} = & ASC_{RAIL} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& - \ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& - \ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{RAIL})) \\
& - \ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{BUS} - CO_{RAIL}))) + \beta_{IDADE} \\
& * IDADE + \beta_{GI_2} * GI_2 + \beta_{GI_3} * GI_3 + \beta_{SEX_{RAIL}} * SEXO \\
& + \beta_{CLASSE_B} * CLASSE_B + \beta_{CLASSE_{C/D/E}} * CLASSE_{C/D/E} \\
& + \beta_{HOUSE} * HOUSE + \beta_{TRIP_{RAIL}} * TRIP + \beta_{PICO} * PICO
\end{aligned} \tag{55}$$

4.6.3 Modelos *RUM* espaciais

A especificação da função utilidade para as alternativas *CAR*, *BUS* e *RAIL* se diferem das funções utilidades dos modelos RUM_{LOS} e $RUM_{LOS+SOCIO}$ pela adição das defasagens espaciais dos atributos de nível de serviço apenas, que são calculados conforme a equação (56).

Nesta dissertação as defasagens de tempo e custo de viagem das alternativas *CAR*, *BUS* e *RAIL* são representadas por WTT_{CAR} , WCO_{CAR} , WTT_{BUS} , WCO_{BUS} , WTT_{RAIL} , e WCO_{RAIL} e são utilizadas para os demais modelos espaciais sob ambas as abordagens. Além disso, a especificação dos parâmetros espaciais seguiu a mesma para os parâmetros não espaciais, não sendo adotado nenhum procedimento de especificação segundo o teste de razão de verossimilhança.

$$z_{ikn} = \sum_{n=1}^N w_{pq} x_{ikn} \tag{56}$$

em que z_{ikn} é a defasagem espacial do atributo k da alternativa i para o indivíduo n ;

x_{ikn} é atributo k da alternativa i para o indivíduo n ;

w_{pq} é o elemento $\langle p, q \rangle$ da matriz de pesos espaciais de dimensão $(n \times n)$; e
 N é o número de indivíduos na amostra.

A Figura 8 ilustra o efeito de dependência espacial exógeno da variável tempo de viagem por ônibus, em que o tempo de viagem na vizinhança do indivíduo analisado apresenta valores similares, causando um efeito de aglomeração (dependência espacial positiva) que influencia a observação em questão na escolha pela alternativa ônibus.

Figura 8 – Efeito de dependência espacial do tempo de viagem por ônibus (RUM).

Fonte: O autor (2021).

As respectivas funções utilidade para o modelo *RUM* com atributos de nível de serviço e suas respectivas defasagens espaciais para as alternativas *CAR*, *BUS* e *RAIL* são apresentadas nas equações (57), (58) e (59) respectivamente. Como são estimados modelos espaciais com diferentes matrizes, os modelos serão denominados $RUM_{LOS_{D1000}}$, $RUM_{LOS_{D1100}}$, e assim por diante até $RUM_{LOS_{D2000}}$ para os modelos com matriz de vizinhança baseada no critério de distância. Para os modelos espaciais com matriz baseada em *knn* se utiliza a denominação $RUM_{LOS_{KNN10}}$, $RUM_{LOS_{KNN20}}$, sucessivamente até $RUM_{LOS_{KNN100}}$.

$$U_{CAR} = \beta_{TT_{CAR}} * TT_{CAR} + \beta_{CO_{CAR}} * CO_{CAR} + \beta_{ESP_{TT_{CAR}}} * WTT_{CAR} + \beta_{ESP_{CO_{CAR}}} * WCO_{CAR} \quad (57)$$

$$U_{BUS} = ASC_{BUS} + \beta_{TT_{BUS}} * TT_{BUS} + \beta_{CO_{GEN}} * CO_{BUS} + \beta_{ESP_{TT_{BUS}}} * WTT_{BUS} + \beta_{ESP_{CO_{GEN}}} * WCO_{BUS} \quad (58)$$

$$U_{RAIL} = ASC_{RAIL} + \beta_{TT_{RAIL}} * TT_{RAIL} + \beta_{CO_{GEN}} * CO_{RAIL} + \beta_{ESP_{TT_{RAIL}}} * TT_{RAIL} + \beta_{ESP_{CO_{GEN}}} * WCO_{RAIL} \quad (59)$$

As respectivas funções utilidade das alternativas *CAR*, *BUS* e *RAIL* para o modelo *RUM* com atributos de nível de serviço e suas respectivas defasagens espaciais mais os atributos sociodemográficos são apresentadas nas equações (60), (61) e (62) respectivamente. Os modelos são denominados $RUM_{LOS+SOCIO_{D1000}}$, $RUM_{LOS+SOCIO_{D1100}}$, e assim por diante até $RUM_{LOS+SOCIO_{D2000}}$ para os modelos com matriz de vizinhança baseada no critério de distância. Para os modelos espaciais com matriz baseada em *knn* se utiliza a denominação $RUM_{LOS+SOCIO_{KNN10}}$, $RUM_{LOS+SOCIO_{KNN20}}$, sucessivamente até $RUM_{LOS+SOCIO_{KNN100}}$.

$$U_{CAR} = \beta_{TT_{CAR}} * TT_{CAR} + \beta_{CO_{CAR}} * CO_{CAR} + \beta_{ESP_{TT_{CAR}}} * WTT_{CAR} + \beta_{ESP_{CO_{CAR}}} * WCO_{CAR} \quad (60)$$

$$\begin{aligned} U_{BUS} = & ASC_{BUS} + \beta_{TT_{BUS}} * TT_{BUS} + \beta_{CO_{GEN}} * CO_{BUS} + \beta_{ESP_{TT_{BUS}}} * WTT_{BUS} \\ & + \beta_{ESP_{CO_{GEN}}} * WCO_{BUS} + \beta_{IDADE} * IDADE + \beta_{GI_2} * GI_2 + \beta_{GI_3} \\ & * GI_3 + \beta_{SEX_{BUS}} * SEXO + \beta_{CLASSE_B} * CLASSE_B + \beta_{CLASSE_{C/D/E}} \\ & * CLASSE_{C/D/E} + \beta_{HOUSE} * HOUSE + \beta_{TRIP_{BUS}} * TRIP + \beta_{PICO} \\ & * PICO \end{aligned} \quad (61)$$

$$\begin{aligned}
U_{RAIL} = & ASC_{RAIL} + \beta_{TT_{RAIL}} * TT_{RAIL} + \beta_{CO_{GEN}} * CO_{RAIL} + \beta_{ESP_{TT_{RAIL}}} \\
& * WTT_{RAIL} + \beta_{ESP_{CO_{GEN}}} * WCO_{RAIL} + \beta_{IDADE} * IDADE + \beta_{GI_2} \\
& * GI_2 + \beta_{GI_3} * GI_3 + \beta_{SEX_{RAIL}} * SEXO + \beta_{CLASSE_B} * CLASSE_B \\
& + \beta_{CLASSE_{C/D/E}} * CLASSE_{C/D/E} + \beta_{HOUSE} * HOUSE + \beta_{TRIP_{RAIL}} \\
& * TRIP + \beta_{PICO} * PICO
\end{aligned} \tag{62}$$

4.6.4 Modelos RRM e RMU espaciais

O modelo proposto nesta dissertação sob a abordagem RRM incorpora efeitos de dependência espacial exógeno, dado que tais efeitos podem ser comparados entre a alternativa considerada e as demais alternativas disponíveis. A Figura 9 apresenta esquematicamente o efeito de dependência do tempo de viagem por ônibus, e dessa vez com a consideração do efeito de dependência espacial do tempo de viagem por metrô. Nesse caso, para além da comparação do tempo de ônibus com o tempo de metrô, os efeitos de dependência de ambas as alternativas também são considerados.

Figura 9 – Efeito de dependência espacial do tempo de viagem por ônibus e metrô (RRM).

Fonte: O autor (2021).

A equação (63) apresenta a função arrependimento que incorpora, então, os efeitos diretos das variáveis de nível de serviço e os efeitos da comparação dos efeitos de dependência espacial desses respectivos atributos.

$$R_i = \sum_{i=1}^J \sum_{k=1}^K \ln(1 + \exp[\beta_k \cdot (x_{jk} - x_{ik})]) + \sum_{i=1}^J \sum_{k=1}^K \ln(1 + \exp[\beta_{k_{espacial}} \cdot (z_{jk} - z_{ik})]) \quad (63)$$

em que R_i é o arrependimento relativo à alternativa i ;

x_{ik} e x_{jk} são os níveis dos $k = 1, \dots, K$ atributos relativos às alternativas i e j , respectivamente; e

β_k e $\beta_{k_{espacial}}$ são parâmetros estimados.

São propostos então dois modelos sob a abordagem *RRM* que incorporam efeitos de dependência espacial exógena, sendo um com apenas atributos de nível de serviço e suas respectivas defasagens espaciais, e outro com atributos de nível de serviço, suas respectivas defasagens espaciais mais os atributos sociodemográficos.

As respectivas funções arrependimento para o modelo *RRM* com atributos de nível de serviço e as defasagens espaciais são apresentadas nas equações (64), (65) e (66) para as alternativas *CAR*, *BUS* e *RAIL* respectivamente. Os modelos são denominados RRM_{D1000} , RRM_{D1100} , e assim por diante até RRM_{D2000} para os modelos com matriz de vizinhança baseada no critério de distância. Para os modelos espaciais com matriz baseada em *knn* se utiliza a denominação RRM_{KNN10} , RRM_{KNN20} , sucessivamente até RRM_{KNN100} .

$$\begin{aligned}
R_{CAR} = & -\ln(1 + \exp(\beta_{TT_{BUS}} * TT_{BUS} - \beta_{TT_{CAR}} * TT_{CAR})) \\
& -\ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{CAR}} * TT_{CAR})) \\
& -\ln(1 + \exp((\beta_{CO_{GEN}} * CO_{BUS} - \beta_{CO_{CAR}} * CO_{CAR}))) \\
& -\ln(1 + \exp(\beta_{CO_{GEN}} * CO_{RAIL} - \beta_{CO_{CAR}} * CO_{CAR})) \\
& -\ln(1 + \exp(\beta_{ESP_{TT_{BUS}}} * WTT_{BUS} - \beta_{ESP_{TT_{CAR}}} * WTT_{CAR})) \\
& -\ln(1 + \exp(\beta_{ESP_{TT_{RAIL}}} * WTT_{RAIL} - \beta_{ESP_{TT_{CAR}}} * WTT_{CAR})) \\
& -\ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * WCO_{BUS} - \beta_{ESP_{CO_{CAR}}} * WCO_{CAR})) \\
& -\ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * WCO_{RAIL} - \beta_{ESP_{CO_{CAR}}} * WCO_{CAR}))
\end{aligned} \tag{64}$$

$$\begin{aligned}
R_{BUS} = & ASC_{BUS} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& -\ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& -\ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{BUS})) \\
& -\ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{RAIL} - CO_{BUS}))) \\
& -\ln(1 + \exp(\beta_{ESP_{TT_{CAR}}} * WTT_{CAR} - \beta_{ESP_{TT_{BUS}}} * WTT_{BUS})) \\
& -\ln(1 + \exp(\beta_{ESP_{TT_{RAIL}}} * WTT_{RAIL} - \beta_{ESP_{TT_{BUS}}} * WTT_{BUS})) \\
& -\ln(1 + \exp(\beta_{ESP_{CO_{CAR}}} * WCO_{CAR} - \beta_{ESP_{CO_{GEN}}} * WCO_{BUS})) \\
& -\ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * (WCO_{RAIL} - WCO_{BUS})))
\end{aligned} \tag{65}$$

$$\begin{aligned}
R_{RAIL} = & ASC_{RAIL} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& -\ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{BUS}} * TT_{BUS})) \\
& -\ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{RAIL})) \\
& -\ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{BUS} - CO_{RAIL}))) \\
& -\ln(1 + \exp(\beta_{ESP_{TT_{CAR}}} * TT_{CAR} - \beta_{ESP_{TT_{BUS}}} * TT_{BUS})) \\
& -\ln(1 + \exp(\beta_{ESP_{TT_{RAIL}}} * WTT_{RAIL} - \beta_{ESP_{TT_{BUS}}} * WTT_{BUS})) \\
& -\ln(1 + \exp(\beta_{ESP_{CO_{CAR}}} * WCO_{CAR} - \beta_{ESP_{CO_{GEN}}} * WCO_{RAIL})) \\
& -\ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * (WCO_{BUS} - WCO_{RAIL})))
\end{aligned} \tag{66}$$

As respectivas funções utilidade-arrependimento para o modelo *RMU* com atributos de nível de serviço, as defasagens espaciais mais os atributos sociodemográficos são apresentadas nas equações (67), (68) e (69) para as alternativas *CAR*, *BUS* e *RAIL* respectivamente. Os modelos são denominados RMU_{D1000} , RMU_{D1100} , e assim por diante até RMU_{D2000} para aqueles com matriz de vizinhança baseada no critério de distância. Para os modelos com matriz baseada em *knn* se utiliza a denominação RMU_{KNN10} , RMU_{KNN20} , sucessivamente até RMU_{KNN100} .

$$\begin{aligned}
 R_{CAR} = & -\ln(1 + \exp(\beta_{TT_{BUS}} * TT_{BUS} - \beta_{TT_{CAR}} * TT_{CAR})) \\
 & -\ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{CAR}} * TT_{CAR})) \\
 & -\ln(1 + \exp(\beta_{CO_{GEN}} * CO_{BUS} - \beta_{CO_{CAR}} * CO_{CAR})) \\
 & -\ln(1 + \exp(\beta_{CO_{GEN}} * CO_{RAIL} - \beta_{CO_{CAR}} * CO_{CAR})) \\
 & -\ln(1 + \exp(\beta_{ESP_{TT_{BUS}}} * WTT_{BUS} - \beta_{ESP_{TT_{CAR}}} * WTT_{CAR})) \\
 & -\ln(1 + \exp(\beta_{ESP_{TT_{RAIL}}} * WTT_{RAIL} - \beta_{ESP_{TT_{CAR}}} * WTT_{CAR})) \\
 & -\ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * WCO_{BUS} - \beta_{ESP_{CO_{CAR}}} * WCO_{CAR})) \\
 & -\ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * WCO_{RAIL} - \beta_{ESP_{CO_{CAR}}} * WCO_{CAR}))
 \end{aligned} \tag{67}$$

$$\begin{aligned}
 R_{BUS} = & ASC_{BUS} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{BUS})) \\
 & -\ln(1 + \exp(\beta_{TT_{RAIL}} * TT_{RAIL} - \beta_{TT_{BUS}} * TT_{BUS})) \\
 & -\ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{BUS})) \\
 & -\ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{RAIL} - CO_{BUS}))) \\
 & -\ln(1 + \exp(\beta_{ESP_{TT_{CAR}}} * WTT_{CAR} - \beta_{ESP_{TT_{BUS}}} * WTT_{BUS})) \\
 & -\ln(1 + \exp(\beta_{ESP_{TT_{RAIL}}} * WTT_{RAIL} - \beta_{ESP_{TT_{BUS}}} * WTT_{BUS})) \\
 & -\ln(1 + \exp(\beta_{ESP_{CO_{CAR}}} * WCO_{CAR} - \beta_{ESP_{CO_{GEN}}} * WCO_{BUS})) \\
 & -\ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * (WCO_{RAIL} - WCO_{BUS}))) + \beta_{IDADE} \\
 & * IDADE + \beta_{GI_2} * GI_2 + \beta_{GI_3} * GI_3 + \beta_{SEXO_{BUS}} * SEXO \\
 & + \beta_{CLASSE_B} * CLASSE_B + \beta_{CLASSE_{C/D/E}} * CLASSE_{C/D/E} \\
 & + \beta_{HOUSE} * HOUSE + \beta_{TRIP_{BUS}} * TRIP + \beta_{PICO} * PICO
 \end{aligned} \tag{68}$$

$$\begin{aligned}
R_{RAIL} = & ASC_{RAIL} - \ln(1 + \exp(\beta_{TT_{CAR}} * TT_{CAR} - \beta_{TT_{BUS}} * TT_{RAIL})) \\
& - \ln(1 + \exp(\beta_{TT_{BUS}} * TT_{BUS} - \beta_{TT_{RAIL}} * TT_{RAIL})) \\
& - \ln(1 + \exp(\beta_{CO_{CAR}} * CO_{CAR} - \beta_{CO_{GEN}} * CO_{RAIL})) \\
& - \ln(1 + \exp(\beta_{CO_{GEN}} * (CO_{BUS} - CO_{RAIL}))) \\
& - \ln(1 + \exp(\beta_{ESP_{TT_{CAR}}} * WTT_{CAR} - \beta_{ESP_{TT_{RAIL}}} * WTT_{RAIL})) \\
& - \ln(1 + \exp(\beta_{ESP_{TT_{BUS}}} * WTT_{BUS} - \beta_{ESP_{TT_{RAIL}}} * WTT_{RAIL})) \quad (69) \\
& - \ln(1 + \exp(\beta_{ESP_{CO_{CAR}}} * WCO_{CAR} - \beta_{ESP_{CO_{GEN}}} * WCO_{RAIL})) \\
& - \ln(1 + \exp(\beta_{ESP_{CO_{GEN}}} * (WCO_{BUS} - WCO_{RAIL}))) + \beta_{IDADE} \\
& * IDADE + \beta_{GI_2} * GI_2 + \beta_{GI_3} * GI_3 + \beta_{SEX_{RAIL}} * SEXO \\
& + \beta_{CLASSE_B} * CLASSE_B + \beta_{CLASSE_{C/D/E}} * CLASSE_{C/D/E} \\
& + \beta_{HOUSE} * HOUSE + \beta_{TRIP_{RAIL}} * TRIP + \beta_{PICO} * PICO
\end{aligned}$$

4.6.5 VTT e elasticidades

O cálculo do VTT para a abordagem *RUM* é simples e resulta da razão entre os parâmetros de tempo e de custo de viagem. O VTT para a abordagem *RRM* é calculado com base na formulação de Chorus (2012d) discutida na seção 2.2.6. A determinação do erro associado à essa medida pode ser feita pelo método Delta discutido por Daly, Hess e de Jong (2012). Esse método é utilizado nesta dissertação para determinação do nível de confiança do VTT apenas para os modelos da abordagem *RUM*. Não se encontrou na literatura discussão sobre o uso do método Delta aplicado no contexto da abordagem *RRM*, indo além do escopo desta pesquisa analisar com maior profundidade essa questão.

As elasticidades diretas são calculadas para os atributos de nível de serviço, ou seja, o efeito desses atributos em suas respectivas alternativas. Para efeitos de comparação entre modelos são calculadas as elasticidades diretas no ponto, que são agregadas pelo método de enumeração amostral. Para a abordagem *RUM* o leitor pode encontrar mais detalhes sobre elasticidades e o método de enumeração amostral em Ben-Akiva e Lerman (1985). No caso da abordagem *RRM* as formulações para o cálculo das elasticidades para o modelo clássico *CRRM* pode ser encontrado em Hensher, Greene e Chorus (2013) conforme discutido na seção 2.2.6.

4.7 SOFTWARE

Os procedimentos de tratamento de dados e de estimação dos modelos foram realizados no software estatístico R (R CORE TEAM, 2019). Para o georreferenciamento dos dados e operações espaciais utilizou-se o pacote *sf* (PEBESMA, 2018); a coleta de dados via API do OTP foi realizada usando o pacote *otpr* (YOUNG, 2019) para roteamento das viagens das alternativas de transporte público e distâncias; os pacotes *httr* (WICKHAM, 2019) e *jsonlite* (OOMS, 2014) foram utilizados para realizar os requerimentos à API da empresa TomTom. A construção das matrizes de vizinhança foram realizadas pelo pacote *spdep* (BIVAND; PEBESMA; GÓMEZ-RUBIO, 2013; BIVAND; PIRAS, 2015).

Os modelos de escolha discreta, os testes de razão de verossimilhança e o teste de B&S foram estimados pelo pacote *apollo* (HESS; PALMA, 2019a, 2019b), uma ferramenta de estimação de modelos de escolha discreta bastante flexível, que permite a estimação de modelos simples como Logit multinomial até estruturas mais complexas, como modelos híbridos e por diferentes estimadores, como estimação bayesiana e algoritmo EM.

5 RESULTADOS

Este capítulo apresenta os resultados dos modelos estimados, inicialmente pela discussão das estimativas e medidas de aderência, seguindo para discussões sobre os parâmetros estimados, medidas de VTT e elasticidades. Por fim discute-se implicações dos modelos para a formulação de políticas públicas de transporte.

5.1 RESULTADOS GERAIS

Esta subseção apresenta os resultados dos modelos em termos de ajuste (Log-Verossimilhança) e da medida AIC a fim de analisar o desempenho entre as abordagens propostas (*RUM* e *RRM*) e a incorporação de dependência espacial exógena.

5.1.1 Modelos com atributos de nível de serviço

A Figura 10 e Figura 11 apresentam o resultado de log-verossimilhança dos modelos *RUM* e *RRM* estimados apenas com as variáveis de nível de serviço (*LOS*), com e sem a presença de efeitos de dependência espacial exógena.

Figura 10 – Log-Verossimilhança dos modelos *LOS* para matriz baseada em distância.

Figura 11 – Log-Verossimilhança dos modelos LOS para matriz baseada em knn .

O modelo RRM_{LOS} apresenta desempenho melhor que o modelo RUM_{LOS} . Dado que não incorporam efeitos de dependência espacial, o resultado de log-verossimilhança é constante para todos os limites de distância e número de vizinhos mais próximos estabelecidos. A diferença de log-verossimilhança entre esses modelos é de 27,54 pontos, indicando que o RRM_{LOS} tem desempenho melhor em termos de ajuste que o RUM_{LOS} . Essa diferença é estatisticamente significativa a 95% de confiança segundo o teste de B&S ($z = 27,54, p - value < 0,00$). Além disso, esse resultado condiz com as análises de Wong *et al.* (2020) que indicam a abordagem RRM com melhor desempenho em relação à abordagem RUM em dados de PR.

A análise dos modelos sob a abordagem RUM com e sem a presença de efeitos de dependência espacial indica que todos aqueles que incorporaram efeitos de dependência espacial têm desempenho melhor em termos de log-verossimilhança que o modelo RUM_{LOS} . Para os modelos espaciais baseados na matriz de distância, a menor diferença entre o modelo RUM_{LOS} se dá com o $RUM_{LOS_{D1100}}$ (modelo RUM com matriz baseada em distância e limite máximo de distância entre vizinhos igual 1.100 metros), com valor igual 145,90 pontos de log-verossimilhança, diferença estatisticamente significativa a 95% de confiança pelo teste de razão de verossimilhança ($\chi^2(5) = 326,08, p - value < 0,00$). O modelo espacial com matriz

baseada em distância de melhor desempenho foi o modelo $RUM_{LOS_{D_{2000}}}$, com diferença de 167,68 pontos de log-verossimilhança em relação ao RUM_{LOS} .

Quanto aos modelos espaciais baseados no critério knn , observa-se que estes também têm desempenho melhor que o RUM_{LOS} , sendo a menor diferença para o $RUM_{LOS_{KNN10}}$ com valor igual a 139,03 pontos de log-verossimilhança, e estatisticamente significativo a 95% de confiança pelo teste de razão de verossimilhança ($\chi^2(5) = 326,08, p - value < 0,00$). O modelo baseado na matriz knn com melhor desempenho é o $RUM_{LOS_{KNN80}}$ com 182,64 pontos de log-verossimilhança de diferença com o RUM_{LOS} .

A comparação entre os modelos espaciais pelas diferentes especificações de matrizes indica que o modelo com melhor desempenho é o baseado na matriz knn ($RUM_{LOS_{KNN80}}$) com diferença de 14,95 pontos de log-verossimilhança comparado com o melhor modelo para a matriz baseada em distância ($RUM_{LOS_{D_{2000}}}$).

Além disso, cabe destacar que apesar do melhor resultado, observa-se na Figura 10 que a partir do limite de 1.400 metros os resultados são estáveis e os modelos com limite de distância máxima acima desse valor não apresentam melhora significativa. Para os modelos espaciais com matriz baseada em knn (Figura 11) observa-se estabilização a partir de k igual a 80 vizinhos.

Esses resultados indicam que a incorporação de efeitos de dependência espacial nos modelos baseados na abordagem RUM melhora o desempenho em termos de ajuste, sendo essas diferenças estatisticamente significantes.

A comparação da abordagem RRM entre o modelo não espacial (RRM_{LOS}) com os modelos espaciais, por sua vez, também indica que a incorporação de dependência espacial nos atributos de nível de serviço melhora o desempenho dos modelos em termos de log-verossimilhança.

No caso da abordagem RRM , a menor diferença entre o modelo RRM_{LOS} e os modelos espaciais com matriz baseada em critério de distância foi para o $RRM_{LOS_{D_{1100}}}$ com valor igual a 147,81 pontos de log-verossimilhança, diferença estatisticamente significativa pelo teste de razão de verossimilhança ($\chi^2(5) = 295,62, p - value < 0,00$). A maior diferença, por sua vez, se deu com o modelo $RRM_{LOS_{D_{2000}}}$ com valor de 165,72 pontos de log-verossimilhança.

No caso dos modelos RRM espaciais com matriz baseada em knn , a menor diferença com o modelo RRM_{LOS} se deu com o $RRM_{LOS_{KNN10}}$ com valor igual a 142,49 pontos de log-verossimilhança, significativa estatisticamente a 95% de confiança ($\chi^2(5) = 284,98, p - value < 0,00$). A maior diferença se deu com o modelo $RRM_{LOS_{KNN80}}$ com valor igual 181,90

pontos de log-verossimilhança. A comparação entre os modelos espaciais sob a abordagem *RRM* indica que os modelos com matriz baseada em *knn* apresentaram os melhores resultados, com diferença de 16,18 pontos de log-verossimilhança entre o $RRM_{LOS_{KNN80}}$ e o $RRM_{LOS_{D2000}}$.

De maneira similar ao verificado para os modelos sob a abordagem *RUM*, observa-se que os resultados dos modelos espaciais sob a abordagem *RRM* se estabilizam a partir de 1.400 metros para os modelos com matriz baseada em distância e a partir de *k* igual a 80 vizinhos para os modelos com matriz baseada em *knn*.

A comparação entre as diferentes abordagens mostra que os modelos espaciais sob a abordagem *RRM* têm um desempenho melhor em termos de log-verossimilhança que os modelos baseados na abordagem *RUM* independentemente do critério de vizinhança adotado. A menor diferença entre os modelos espaciais sob a abordagem *RUM* e *RRM* com matriz baseada em distância se deu entre o $RRM_{LOS_{D2000}}$ e o $RUM_{LOS_{D2000}}$ com valor igual a 25,72 pontos de log-verossimilhança, sendo essa diferença estatisticamente significativa pelo teste B&S ($z = 25,72$, $p - value < 0,00$). A maior diferença para os modelos espaciais com matriz baseada em *knn* se deu entre $RRM_{LOS_{KNN80}}$ e $RUM_{LOS_{KNN80}}$ com valor de 41,76 pontos de log-verossimilhança.

Essa última diferença também representa a maior entre o melhor modelo espacial sob a abordagem *RRM* com o melhor modelo espacial sob a abordagem *RUM*. Pode-se concluir, portanto, que a incorporação de dependência espacial exógena através das defasagens dos atributos de nível de serviço melhora o desempenho dos modelos em termos de log-verossimilhança, indicando a presença de dependência espacial nesses atributos.

Além disso, pode-se afirmar que a incorporação de dependência espacial no contexto da abordagem *RRM* resultou em melhora significativa em termos de ajuste, sendo a diferença do melhor modelo ($RRM_{LOS_{KNN80}}$) para o pior modelo (RUM_{LOS}) com valor de 209,44 pontos de log-verossimilhança ao custo de 5 parâmetros adicionais.

Para escolha do melhor modelo, no entanto, a estatística AIC é recomendada pois considera a complexidade dos modelos propostos em termos de parcimônia (número de parâmetros). A Figura 12 e Figura 13 apresentam o resultado dessa estatística para os modelos *LOS*. O modelo com menor valor da estatística é o mais recomendado pois indica parcimonia em termos de ajuste e número de parâmetros.

Observa-se que, de maneira similar à análise feita anteriormente em termos de log-verossimilhança, a estatística AIC indica que os modelos com melhor desempenho são os baseados na abordagem *RRM* e aqueles que incorporam dependência espacial nas variáveis

independentes. Entre os modelos espaciais, observa-se que aqueles com matriz baseada em *knn* apresentam o menor valor da estatística, indicando que têm melhor desempenho. Observa-se também que a partir do limite de distância de 1.400 metros e de *k* igual a 80 vizinhos os resultados são estáveis e bastante próximos entre suas respectivas categorias.

Figura 12 – AIC dos modelos *LOS* para matriz baseada em distância.

Figura 13 – AIC dos modelos *LOS* para matriz baseada em k vizinhos mais próximos.

Assim, pode-se afirmar que o modelo com melhor desempenho em termos de log-verossimilhança e parcimônia é o espacial sob a abordagem *RRM*. A Tabela 2 apresenta os resultados gerais dos modelos RUM_{LOS} e RRM_{LOS} e os melhores modelos espaciais sob ambas as abordagens. Observa-se que os valores de aderência (ρ^2 e ρ_{ajust}^2) indicam o melhor desempenho dos modelos espaciais em relação aos não espaciais e daqueles sob a abordagem *RRM* em relação aos modelos sob a abordagem *RUM*.

Tabela 2 – Resultados gerais dos modelos *LOS*.

Modelo	$LL_{inicial}$	LL_{final}	ρ^2	ρ_{ajust}^2	AIC	BIC	#Param	Tempo
RUM_{LOS}	-10.255,55	-7.779,52	0,2414	0,2407	15.573,05	15.623,04	7	2s
RRM_{LOS}		-7.751,98	0,2441	0,2434	15.517,97	15.567,96	7	5s
$RUM_{LOS_{KNN80}}$		-7596.88	0.2592	0.2581	15217.77	15303.47	12	3s
$RRM_{LOS_{KNN80}}$		-7.570,08	0,2619	0,2607	15.164,16	15.249,86	12	12s

Assim como a medida AIC, a medida BIC (que tem maior penalização por parâmetro em grandes amostras) indica o melhor desempenho do modelo espacial sob a abordagem *RRM*. Outro ponto importante é que a diferença de tempo de estimação dos modelos entre as diferentes abordagens não foi expressiva, indicando que ambos os modelos são simples de estimar do ponto de vista prático.

5.1.2 Modelos com atributos de nível de serviço e sociodemográficos

A Figura 14 e Figura 15 apresentam os resultados de ajuste dos modelos com atributos de nível de serviço e sociodemográficos (*LOS + SOCIO*) sob a abordagem *RUM* e *RRM*, com e sem efeitos de dependência espacial, para as matrizes baseadas em critério de distância e *knn*, respectivamente. Diferente dos modelos *LOS*, aqueles sob a abordagem *RRM* são híbridos *RUM-RRM* (*RMU*), dada a estimação dos atributos sociodemográficos sob a abordagem *RUM* discutida na seção 2.2.2 e capítulo 4.

Figura 14 – Log-Verossimilhança dos modelos *LOS + SOCIO* para matriz de distância.

Figura 15 – Log-Verossimilhança dos modelos *LOS + SOCIO* para matriz baseada em *knn*.

Pelas figuras, o modelo *RMU* tem desempenho melhor que o modelo $RUM_{LOS+SOCIO}$, com diferença de 21,93 pontos de log-verossimilhança, estatisticamente significativa a 95% de confiança pelo teste de B&S ($z = 21,93, p - value < 0,00$). Esse resultado também é análogo a estudos que encontraram que modelos híbridos têm melhor desempenho que aqueles apenas sob a abordagem *RUM* ou *RRM* (HENSHER; GREENE; CHORUS, 2013; LEONG; HENSHER, 2015).

Os modelos espaciais baseados na abordagem *RUM* apresentaram melhor desempenho comparado ao modelo $RUM_{LOS+SOCIO}$. A menor diferença entre o modelo $RUM_{LOS+SOCIO}$ em relação aos modelos espaciais com matriz baseada em distância foi de 13,20 pontos de log-verossimilhança para o modelo $RUM_{LOS+SOCIO_{D1100}}$, sendo estatisticamente significativa a 95% de confiança pelo teste de razão de verossimilhança ($\chi^2(5) = 26,42, p - value < 0,00$). A maior diferença entre o modelo $RUM_{LOS+SOCIO}$ e os modelos espaciais com matriz baseada em distância foi de 17,77 pontos de log-verossimilhança para o modelo $RUM_{LOS+SOCIO_{D2000}}$.

No caso dos modelos espaciais *RUM* com matriz baseada em *knn*, os resultados também mostram que têm desempenho melhor que o modelo $RUM_{LOS+SOCIO}$. Nesse caso, a menor diferença foi de 12,65 pontos de log-verossimilhança do modelo $RUM_{LOS+SOCIO}$ para o modelo $RUM_{LOS+SOCIO_{KNN10}}$, sendo essa diferença estatisticamente significativa a 95% de confiança pelo teste de razão de verossimilhança ($\chi^2(5) = 25,30, p - value < 0,00$). A maior

diferença foi de 26,38 pontos de log-verossimilhança entre o modelo $RUM_{LOS+SOCIO}$ e o modelo $RUM_{LOS+SOCIO_{KNN90}}$.

Analisando os modelos sob a abordagem híbrida (modelo RMU), observa-se que a inclusão dos efeitos de dependência espacial exógena também melhora o desempenho, com menor diferença entre o não espacial e os modelos espaciais com matriz baseada em distância no valor de 13,86 pontos de log-verossimilhança para o modelo RMU_{D1100} , estatisticamente significativa a 95% de confiança ($\chi^2(5) = 27,72, p - value < 0,00$).

Entre os modelos híbridos espaciais com matriz baseada em knn e o modelo RMU , a menor diferença se deu com o modelo RMU_{KNN10} com valor igual a 13,68 pontos de log-verossimilhança, estatisticamente significativa pelo teste de razão de verossimilhança ($\chi^2(5) = 27,36, p - value < 0,00$). A maior diferença se deu com o modelo RMU_{KNN90} com valor igual a 25,71 pontos de log-verossimilhança.

Analisando os modelos espaciais sob a abordagem RUM ($LOS + SOCIO$), diferentemente dos modelos espaciais RUM (LOS), se observa que aqueles com matriz baseada em distância apresentaram resultado abaixo do modelo RMU para todos os limites de distância. No caso dos modelos espaciais RUM com matriz baseada em knn , só apresentam melhor desempenho que o RMU a partir de k igual a 60 vizinhos, com maior diferença para o modelo $RUM_{LOS+SOCIO_{KNN90}}$ com valor igual a 4,46 pontos de log verossimilhança, porém não estatisticamente significativa a 95% de confiança ($\chi^2(5) = 8,92, p - value = 0,11$).

Dessa forma, se observa que nenhum modelo espacial sob a abordagem RUM ($LOS + SOCIO$) apresentou resultado melhor que a abordagem RRM , até mesmo em comparação com o modelo RMU não espacial com 5 parâmetros a menos.

A análise comparativa entre modelos espaciais sob as diferentes abordagens indica que tem melhor performance em termos de ajuste os modelos espaciais sob a abordagem RRM . A maior diferença entre os modelos espaciais RMU com os modelos espaciais $RUM_{LOS+SOCIO}$ com matriz baseada em distância se deu entre o RMU_{D2000} e o RUM_{D2000} com valor igual a 20,55 pontos de log-verossimilhança. No caso dos modelos com matriz baseada em knn , a maior diferença se deu entre o RMU_{KNN100} e o RUM_{KNN100} com valor igual a 21,21 pontos de log-verossimilhança. Ambas as diferenças são estatisticamente significativas a 95% de confiança segundo o teste de B&S.

Assim, de forma similar aos modelos LOS , a abordagem RRM com atributos sociodemográficos (modelo híbrido), apresentou desempenho melhor que a abordagem RUM . No caso dos modelos com atributos de nível de serviço e sociodemográficos, a diferença

absoluta entre os modelos espaciais sob a abordagem *RUM* e *RRM* é similar em magnitude ao encontrado nos modelos *LOS* com matrizes baseadas em distância, e diferentes para os modelos baseados em *knn* (sendo essa diferença aproximadamente metade dos modelos *LOS*).

O melhor modelo com atributos de nível de serviço e sociodemográficos foi o híbrido, que incorpora efeitos de dependência espacial com matriz baseada em *knn* (RMU_{KNN90}), com diferença de 47,63 pontos de log-verossimilhança para o modelo com pior desempenho ($RUM_{LOS+SOCIO}$) e diferença significativa a 5% de significância segundo o teste de B&S ($z(5) = 47,63, p - value < 0,00$). Assim, incorporar heterogeneidade no processamento dos atributos (*RUM-RRM*) e incorporar efeitos de dependência espacial sob a abordagem *RRM* também melhoraram o desempenho dos modelos de maneira significativa do ponto de vista de ajuste.

A partir da Figura 14 e Figura 15 ainda é possível observar que, de maneira similar aos modelos com atributos de nível de serviço, os resultados de ajuste dos modelos espaciais para ambas as abordagens se estabilizam a partir de 1.400 metros no caso das matrizes baseadas em critério de distância e a partir de k igual a 80 vizinhos para os modelos com matrizes baseadas em *knn*.

A Figura 16 e Figura 17 apresentam os resultados da estatística AIC para os modelos *LOS + SOCIO*. Como se observa, e de maneira similar aos resultados de log-verossimilhança, os modelos com melhor desempenho são os híbridos *RMU*, em comparação aos modelos baseados na abordagem *RUM* com ou sem efeitos de dependência espacial. O melhor modelo entre todos os estimados é o RMU_{KNN90} . É válido destacar, no entanto, que a partir do limite de distância de 1.400 metros e k igual a 80 vizinhos, os resultados são estáveis para cada critério de vizinhança e abordagem.

Figura 16 – AIC dos modelos *LOS + SOCIO* para matriz baseada em distância.

Figura 17 – Log-Verossimilhança dos modelos *LOS + SOCIO* para matriz baseada em *knn*.

A Tabela 3 apresenta resultados gerais de aderência dos modelos sem efeitos de dependência espacial e os melhores modelos espaciais de cada abordagem. De maneira

simplificada, os resultados de aderência (ρ^2 e ρ_{ajust}^2) e medidas de ajuste (AIC e BIC) indicam o melhor resultado da abordagem *RRM* híbrida tanto entre modelos que não incorporam efeitos de dependência espacial quanto entre aqueles que incorporam. Outro fator importante de destaque é o tempo de estimação entre as abordagens, que para estes modelos também foi insignificante.

Tabela 3 – Resultados gerais dos modelos *LOS + SOCIO*.

Modelo	$LL_{initial}$	LL_{final}	ρ^2	ρ_{ajust}^2	AIC	BIC	#Param	Tempo
$RUM_{LOS+SOCIO}$	-10.255,55	-6.872,69	0,3299	0,3279	13.785,39	13.928,22	20	4s
RMU		-6.850,77	0,3320	0,3300	13.741,54	13.884,37	20	9s
$RUM_{LOS+SOCIO}$		-6.846,31	0,3324	0,3300	13.742,63	13.921,16	25	5s
RMU_{KNN90}		-6.825,06	0,3345	0,3321	13.700,11	13.878,65	25	18s

5.2 RESULTADOS DOS PARÂMETROS ESTIMADOS

A Tabela 4 e a Tabela 5 apresentam os resultados dos parâmetros estimados dos modelos *LOS* e *LOS + SOCIO*, respectivamente. Em síntese, apresentam-se os parâmetros dos modelos não espaciais e apenas os parâmetros dos melhores modelos espaciais discutidos anteriormente. Os resultados para os demais modelos são apresentados no Apêndice A.

Tabela 4 – Parâmetros estimados dos modelos *LOS*.

Atributo	RUM_{LOS}		RRM_{LOS}		$RUM_{LOS_{KNN80}}$		$RRM_{LOS_{KNN80}}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	1.2725	10.99*	0.4491	6.45*	0.3016	0.80	0.1648	0.49
ASC_{RAIL}	2.0194	14.87*	1.5978	15.77*	1.6248	3.38*	2.0330	4.90*
TT_{CAR}	-0.0453	-14.43*	-0.0335	-15.86*	-0.0294	-8.80*	-0.0231	-9.79*
TT_{BUS}	-0.0306	-26.04*	-0.0199	-25.74*	-0.0322	-25.10*	-0.0208	-25.71*
TT_{RAIL}	-0.0137	-12.80*	-0.0124	-16.16*	-0.0135	-10.88*	-0.0131	-15.99*
CO_{CAR}	-0.2561	-15.71*	-0.1804	-22.87*	-0.2512	-14.50*	-0.1709	-20.78*
CO_{GEN}	-0.7514	-30.79*	-0.4282	-33.61*	-0.7545	-28.49*	-0.4325	-31.20*
Parâmetros defasados espacialmente								
TT_{CAR}					-0.1084	-10.54*	-0.0522	-10.01*
TT_{BUS}					0.0062	2.24*	0.0088	3.66*
TT_{RAIL}					-0.0030	-1.10	0.0031	2.01*
CO_{CAR}					0.2020	4.40*	0.2046	3.98*
CO_{GEN}					0.0116	0.11	-0.0637	-1.20

* Significativo a 95% de confiança

Tabela 5 – Parâmetros estimados dos modelos $LOS + SOCIO$.

Atributo	$RUM_{LOS+SOCIO}$		RMU		$RUM_{LOS+SOCIO_{KNN90}}$		$RRM_{LOS_{KNN90}}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	3,1166	13,68*	2,0536	10,06*	3,1807	7,04*	2,8322	6,91*
ASC_{RAIL}	3,3302	12,77*	2,7140	11,40*	3,4937	6,22*	3,5482	7,01*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0413	-12,21*	-0,0305	-12,93*	-0,0340	-9,54*	-0,0264	-10,41*
TT_{BUS}	-0,0344	-27,69*	-0,0221	-27,72*	-0,0349	-25,81*	-0,0222	-26,36*
TT_{RAIL}	-0,0157	-14,07*	-0,0136	-17,70*	-0,0154	-12,03*	-0,0141	-16,63*
CO_{CAR}	-0,2254	-13,24*	-0,1607	-19,28*	-0,2462	-13,56*	-0,1683	-19,51*
CO_{GEN}	-0,8081	-31,27*	-0,4524	-33,74*	-0,8011	-28,82*	-0,4536	-31,46*
Parâmetros defasados espacialmente								
TT_{CAR}	-	-	-	-	-0,0770	-6,89*	-0,0393	-6,88*
TT_{BUS}	-	-	-	-	-0,0001	-0,02	0,0020	0,86
TT_{RAIL}	-	-	-	-	-0,0028	-0,96	0,0023	1,21
CO_{CAR}	-	-	-	-	0,2768	5,59*	0,2604	4,80*
CO_{GEN}	-	-	-	-	0,0365	0,34	-0,0225	-0,37
Parâmetros Sociodemográficos								
$PICO$	-0,2812	-5,45*	-0,2771	-5,41*	-0,2544	-4,89*	-0,2504	-4,86*
$IDADE$	-0,0298	-15,14*	-0,0296	-15,13*	-0,0298	-15,08*	-0,0295	-15,05*
$CLASSE_B$	0,8309	13,27*	0,8251	13,29*	0,7897	12,45*	0,7903	12,58*
$CLASSE_{C/D/E}$	1,0442	11,22*	1,0318	11,17*	0,9782	10,26*	0,9686	10,23*
$SEXO_{BUS}$	0,9482	15,46*	0,9397	15,33*	0,9503	15,45*	0,9412	15,31*
$SEXO_{RAIL}$	0,7054	10,84*	0,6852	10,54*	0,7045	10,80*	0,6828	10,49*
$HOUSE$	0,1869	9,16*	0,1854	9,16*	0,1774	8,62*	0,1764	8,63*
$TRIPS_{BUS}$	-0,2773	-10,69*	-0,2725	-10,55*	-0,2785	-10,72*	-0,2738	-10,59*
$TRIPS_{RAIL}$	-0,2097	-7,88*	-0,2070	-7,81*	-0,2118	-7,95*	-0,2089	-7,86*
GI_{2BUS}	-0,1210	-1,19	-0,1070	-1,05	-0,0892	-0,87	-0,0756	-0,73
GI_{2RAIL}	-1,0591	-9,97*	-1,0344	-9,76*	-0,9841	-8,97*	-0,9665	-8,84*
GI_{3BUS}	0,2756	2,20*	0,2848	2,26*	0,3043	2,41*	0,3209	2,53*
GI_{3RAIL}	-0,1846	-1,46	-0,1613	-1,27	-0,1368	-1,05	-0,0988	-0,75

* Significativo a 95% de confiança

5.2.1 Parâmetros dos atributos de nível de serviço

Observa-se na Tabela 4 que o parâmetro ASC_{BUS} é significativo a 95% de confiança apenas nos modelos não espaciais (RUM_{LOS} e RRM_{LOS}) e não significativo nos modelos

espaciais. O sinal positivo indica que, em média, a alternativa *BUS* é preferida em relação à alternativa *CAR*. O parâmetro ASC_{RAIL} , por sua vez, é positivo e significativo a 95% de confiança em todos os modelos *LOS*, indicando a preferência da alternativa *RAIL* em relação à alternativa *CAR*. No caso dos modelos *LOS + SOCIO*, ambos os parâmetros ASC_{BUS} e ASC_{RAIL} são positivos e significativos a 95% de confiança, indicando a preferência dos indivíduos por essas alternativas em relação à *CAR*, em média.

O parâmetro de tempo de viagem da alternativa *CAR* apresentou sinal esperado (negativo), além de ser significativo a 95% de confiança em todos os modelos estimados, *LOS* e *LOS + SOCIO*. Assim, pode-se afirmar que o tempo de viagem por automóvel afeta negativamente a propensão de escolha da alternativa *CAR*. É válido lembrar também que a comparação em termos absolutos entre os parâmetros sob as diferentes abordagens não tem sentido, visto que representam processos distintos como discutido na seção 2.2.5 acerca da interpretação de parâmetros sob a abordagem *RRM*.

A análise entre os modelos para cada abordagem, no entanto, pode ser realizada. Observa-se que aqueles com a presença de efeitos de dependência espacial apresentam parâmetros com menor valor comparado com suas contrapartes não espaciais, tanto na abordagem *RUM* quanto na abordagem *RRM* para o parâmetro de tempo de viagem da alternativa *CAR*.

Os parâmetros de tempo de viagem da alternativa *BUS* e *RAIL* (TT_{BUS} e TT_{RAIL}) também apresentam sinal esperado (negativo) e são estatisticamente significativos a 95% de confiança para todos os modelos *LOS* e *LOS + SOCIO* sob ambas as abordagens. Assim, o tempo de viagem por transporte público também tem efeito negativo na probabilidade de escolha dessas alternativas como modo de transporte.

De forma geral, para cada abordagem, o parâmetro do atributo TT_{BUS} tem efeito negativo maior que o parâmetro de tempo de viagem da alternativa *RAIL* (TT_{RAIL}) em todos os modelos. Comparando os três atributos de tempo de viagem, observa-se que o parâmetro do atributo TT_{CAR} é o com maior efeito negativo sobre a probabilidade de escolha, com a exceção dos modelos $RUM_{LOS_{KNN80}}$ e $RUM_{LOS+SOCIO_{KNN90}}$.

Os parâmetros de custo de viagem (CO_{CAR} e CO_{GEN}) apresentam sinal esperado (negativo) e são todos estatisticamente significativos a 95% de confiança para todos os modelos *LOS* e *LOS + SOCIO* em ambas as abordagens. Assim, é possível afirmar que o custo de viagem tem efeito negativo na propensão em escolher uma dessas alternativas. Para cada abordagem e entre todos os modelos, observa-se que o parâmetro de custo das alternativas *BUS* e *RAIL*

(CO_{GEN}) tem um efeito maior que o parâmetro de custo da alternativa CAR , indicando que tais atributos têm efeito negativo maior na escolha pelo transporte público em relação ao automóvel.

Além disso, diferente do atributo TT_{CAR} , os parâmetros de custo em cada abordagem não apresentaram variação significativa de valor na comparação dos modelos com e sem efeitos de dependência espacial, com exceção dos modelos LOS sob a abordagem RRM .

5.2.2 Parâmetros de atributos sociodemográficos

Os parâmetros relacionados às características sociodemográficas dos indivíduos são estimados apenas nos modelos $LOS + SOCIO$. Como em ambos os modelos RUM e RMU os parâmetros são estimados pela mesma abordagem, podem ser comparados entre si. Os únicos parâmetros que não apresentaram significância estatística a 5% foram aqueles relacionados às categorias de grau de instrução nas alternativas BUS (GI_{2BUS}) e $RAIL$ (GI_{3RAIL}) em todos os modelos.

O parâmetro do atributo $PICO$ apresentou sinal negativo, indicando que indivíduos que iniciam sua viagem no período de pico da manhã têm menor probabilidade de escolher as alternativas de transporte público em relação à alternativa CAR . Esse efeito pode decorrer principalmente do fato de que durante esse horário os sistemas de transporte público coletivo apresentam alta demanda de passageiros, com níveis de serviço e atributos de conforto menores, fazendo com que o modo individual seja preferido.

O parâmetro do atributo $IDADE$ por sua vez apresentou sinal negativo indicando que quanto maior a idade do indivíduo, maior a preferência pelo automóvel em relação às alternativas de transporte público. Esse resultado acompanha aqueles encontrados em estudos geracionais que discutem o aumento do uso de alternativas de transporte motorizado individual com o avanço da idade (DELBOSC; NAZNIN, 2019; KLEIN; SMART, 2017).

Os parâmetros relacionados às classes sociais dos indivíduos ($CLASSE_B$ e $CLASSE_{C/D/E}$) apresentaram sinal positivo, ou seja, pertencer a classes sociais mais baixas aumenta a probabilidade das alternativas de transporte público serem escolhidas.

Os parâmetros relacionados ao sexo dos indivíduos ($SEXO_{BUS}$ e $SEXO_{RAIL}$) têm sinal positivo, de maneira que ser do sexo feminino aumenta a propensão de escolha pelos modos de transporte público coletivo em relação ao automóvel. A questão de gênero e mobilidade já vem sendo discutida na literatura de transportes. No Brasil, por exemplo, o estudo de Macêdo *et al.* (2020) discute a diferenciação no padrão de mobilidade por gênero em diferentes regiões do

país, incluindo a cidade de São Paulo. Assim, os resultados encontrados por aqueles autores corroboram os identificados nesta pesquisa, com maior propensão de indivíduos do sexo feminino de utilizarem o transporte público em relação a modos individuais motorizados.

O parâmetro do atributo *HOUSE* indica que o aumento do número de moradores no domicílio aumenta a propensão dos indivíduos em optar pelo transporte público coletivo em relação à alternativa *CAR*. Isso decorre principalmente da necessidade de divisão dos gastos familiares com transporte. Assim, quanto maior o número de pessoas na família, mais difícil é realizar todas as viagens por automóvel, fazendo com que alguns membros passem a utilizar outras alternativas.

O aumento do número de viagens feitas diariamente diminui a probabilidade de escolha das alternativas *BUS* e *RAIL* em relação à alternativa *CAR* conforme indicado pelos parâmetros $TRIPS_{BUS}$ e $TRIPS_{RAIL}$. Esse resultado está relacionado principalmente à classe social, tal que indivíduos pertencentes a classes sociais mais altas tem maior probabilidade de escolher modos individuais motorizados e, segundo a pesquisa OD 2017, realizam mais viagens que indivíduos de classes sociais mais baixas.

Por fim, a análise dos atributos relacionados ao grau de instrução indica que pertencer à categoria GI_2 (Fundamental II completo até Superior incompleto) diminui a propensão de escolha da alternativa *RAIL* em relação à categoria de referência (até Fundamental II incompleto), enquanto pertencer a categoria GI_3 (Superior completo) aumenta a probabilidade de escolher a alternativa *BUS* em relação à categoria de referência.

5.2.3 Parâmetros espaciais

Os parâmetros relacionados aos efeitos de dependência variaram para cada abordagem nos modelos *LOS*, mas foram similares em termos de significância nos modelos *LOS + SOCIO*. É importante ressaltar que, apesar de os parâmetros espaciais não estarem restritos entre -1 e +1 como aqueles dos modelos autorregressivos, os valores dos parâmetros associados às defasagens das variáveis independentes não se apresentaram fora desse intervalo.

No modelo $RUM_{LOS_{KNN80}}$ somente as defasagens de TT_{CAR} , TT_{BUS} e CO_{CAR} são estatisticamente significativas, indicando a presença de dependência espacial nessas variáveis. Nesse modelo, TT_{BUS} e CO_{CAR} apresentam sinal positivo, indicando efeitos positivos de dependência, ou efeitos de aglomeração, indicando que indivíduos próximos espacialmente experimentam valores similares para esses atributos. A variável TT_{CAR} por sua vez apresentou

sinal negativo, indicando que essa variável está associada ao afastamento de valores similares em sua vizinhança.

No modelo $RRM_{LOS_{KNN80}}$ apenas a variável de defasagem de CO_{GEN} não é significativa a 95% de confiança, indicando forte presença de dependência espacial nos atributos de nível de serviço. As defasagens de TT_{BUS} , TT_{RAIL} e CO_{CAR} indicam efeito de dependência espacial positiva e a variável TT_{CAR} indica a presença de dependência espacial negativa. Nos modelos $LOS + SOCIO$, somente as defasagens associadas a TT_{CAR} e CO_{CAR} são significativas a 95% de confiança em ambas as abordagens. Nesses, a defasagem de TT_{CAR} é negativa em ambos os modelos, enquanto a defasagem de CO_{CAR} é positiva.

5.3 VALORES DE TEMPO DE VIAGEM E ELASTICIDADES

A Tabela 6 apresenta os resultados das medidas de VTT para os modelos apresentados anteriormente. Conforme discutido no Capítulo 2, a medida de VTT é a mais utilizada para avaliação de projetos de transportes, pois indica a disposição de indivíduos a pagar pela diminuição do tempo de viagem de suas respectivas alternativas. Para a abordagem RUM todas as medidas são estatisticamente significativas a 95% de confiança de acordo com o método Delta.

Tabela 6 – VTT dos modelos LOS e $LOS + SOCIO$.

Modelos	CAR		BUS		RAIL	
	VTT (R\$/h)	Desvio Padrão	VTT (R\$/h)	Desvio Padrão	VTT (R\$/h)	Desvio Padrão
RUM_{LOS}	10,60*	-	2,44*	-	1,09*	-
RRM_{LOS}	16,01	5,65	7,24	4,10	1,52	0,37
$RUM_{LOS_{KNN80}}$	7,01*	-	2,56*	-	1,07*	-
$RRM_{LOS_{KNN80}}$	14,67	5,64	7,95	4,59	1,71	0,42
$RUM_{LOS+SOCIO}$	10,99*	-	2,55*	-	1,16*	-
RMU	17,36	6,65	8,37	5,00	1,62	0,41
$RUM_{LOS+SOCIO_{KNN90}}$	8,30*	-	2,61*	-	1,15*	-
RMU_{KNN90}	16,14	6,27	8,53	5,12	1,69	0,43

* Significativo a 95% de confiança pelo método Delta

Para os modelos sob a abordagem RUM observa-se que a alternativa CAR tem o maior valor em relação às demais alternativas, indicando que indivíduos estão dispostos a pagar mais pela diminuição do tempo de viagem dessa alternativa em relação às de transporte público.

Acredita-se que esse resultado ocorre pois indivíduos que utilizam o automóvel pertencem de forma geral a classes sociais mais altas, tendo maior capacidade de pagar para melhorar o desempenho do automóvel, enquanto usuários do transporte público, por outro lado, têm menor capacidade de pagar por tais melhorias. Além disso, como se discutiu na seção anterior, o custo do transporte público tem efeito negativo maior na escolha dessas alternativas em comparação ao efeito do custo de viagem da alternativa *CAR*.

Outros estudos calcularam o VTT no contexto da cidade de São Paulo, como por exemplo Dias (2013) que estimou um modelo binário (Automóvel e Transporte Público) com parâmetro de tempo e custo genéricos, e encontrou o valor de R\$1,83 (R\$3,32 corrigido pelo IPCA para o ano de 2017) usando os dados da Pesquisa OD2007; e Carvalho (2013) que usou dados de PD para cenários de diferentes linhas de ônibus e encontrou o valor de R\$1,44 (R\$1,88 corrigido pelo IPCA para o ano de 2017) para essa alternativa.

Os resultados encontrados nesta dissertação têm valores próximos aos identificados por aqueles autores para o transporte público, sendo a média do VTT das alternativas *BUS* e *RAIL* iguais a R\$1,76 e R\$1,85 para os modelos RUM_{LOS} e $RUM_{LOS+SOCIO}$ respectivamente. Como o estudo de Dias (2013) não apresentou valor específico para a alternativa *CAR*, não é possível identificar se a diferença do VTT entre as alternativas automóvel e transporte público naquele período é da mesma ordem de grandeza da encontrada neste trabalho.

Um modelo com parâmetros genéricos de tempo e custo estimado com os dados desta dissertação indicou um valor de VTT de R\$4,36, também próximo ao estimado por Dias (2013) com valores corrigidos. No entanto, esse modelo apresentou valor de log-verossimilhança muito pior comparado ao modelo RUM_{LOS} proposto (mais de 200 pontos de log-verossimilhança de diferença).

A comparação entre os modelos RUM_{LOS} e $RUM_{LOS_{KNN80}}$ mostra que os valores de VTT para as alternativas *BUS* e *RAIL* não variam significativamente, com diferenças de R\$0,12 e \$0,02 respectivamente. No entanto, para a alternativa *CAR* essa diferença é significativa, sendo o valor absoluto de R\$3,59, com o VTT do modelo RUM_{LOS} 33% maior que o $RUM_{LOS_{KNN80}}$. Cabe destacar que no modelo $RUM_{LOS_{KNN80}}$ as defasagens de TT_{CAR} e CO_{CAR} são estatisticamente significativas indicando a presença de dependência espacial nessas variáveis.

Assim, nota-se que não controlar efeitos de dependência espacial em modelos de escolha discreta sob a abordagem *RUM* afeta o resultado do VTT, que nesta pesquisa superestimou o valor da alternativa *CAR* em 33%, o que não foi observado para as demais

alternativas em que não se verificou presença de dependência espacial em seus respectivos atributos, com exceção da defasagem de TT_{BUS} . Esse resultado confirma a discussão na literatura de que o negligenciamento da dependência espacial enviesaria o resultado dos parâmetros estimados.

A comparação do modelo RUM_{LOS} com o modelo $RUM_{LOS+SOCIO}$ mostra que não há variação significativa no valor do VTT para as três alternativas, ou seja, a ausência dos atributos sociodemográficos não altera significativamente o valor dessa medida. Este resultado é importante para as aplicações práticas de modelos de escolha discreta em projetos de transporte, pois facilita o processo de sistematização e coleta de dados.

No entanto, quando se compara as medidas de VTT entre o modelo $RUM_{LOS+SOCIO}$ e o modelo $RUM_{LOS+SOCIO_{KNN90}}$, o mesmo padrão dos modelos LOS é observado para a alternativa CAR . Nesse caso, o VTT da alternativa CAR do modelo $RUM_{LOS+SOCIO}$ é 24,4% maior que o VTT do modelo $RUM_{LOS+SOCIO_{KNN90}}$. Neste modelo espacial, os parâmetros relacionados às defasagens de tempo e custo de viagem da alternativa CAR são as únicas defasagens espaciais significativas, indicando a existência de dependência espacial nessas variáveis.

A análise das medidas de VTT para a abordagem RRM também evidencia que usuários de automóvel estão mais dispostos a pagar pela diminuição do tempo de viagem que usuários do transporte público coletivo. Além disso, os VTTs das alternativas CAR e BUS sob a abordagem RRM apresentaram valores maiores que suas contrapartes RUM , o que não se observa para a alternativa $RAIL$.

Em termos absolutos, a diferença entre os modelos RUM_{LOS} e RRM_{LOS} , por exemplo, foi de R\$5,41 para a alternativa CAR e R\$4,80 para a alternativa BUS , sendo a diferença para a alternativa $RAIL$ de apenas R\$0,80. Em termos relativos, os VTTs das alternativas CAR e BUS do modelo RRM_{LOS} são 51,03% e 196,72% maiores respectivamente que essas medidas no modelo RUM_{LOS} . Nos modelos $LOS + SOCIO$ essas diferenças relativas são da ordem de 57,96% e 228,23% para as alternativas CAR e BUS , respectivamente.

Comparando os modelos LOS sob a abordagem RRM com e sem a presença de dependência espacial, observa-se que o modelo RRM_{LOS} superestima o VTT da alternativa CAR em 8,36% em comparação com a respectiva medida no modelo $RRM_{LOS_{KNN80}}$. Por outro lado, assim como nos modelos LOS sob a abordagem RUM , as medidas de VTT para as alternativas BUS e $RAIL$ não apresentaram diferenças significativas entre os modelos RRM_{LOS} e $RRM_{LOS_{KNN80}}$.

Nos modelos híbridos RMU e RMU_{KNN90} , a diferença entre o VTT da alternativa CAR é pequena, sendo que o primeiro superestima o valor do VTT da alternativa CAR em 7% e a diferença entre as alternativas BUS e $RAIL$ não é expressiva.

É válido destacar que no modelo $RRM_{LOS_{KNN80}}$ apenas a defasagem espacial do custo de viagem do transporte público não apresentou significância estatística. No caso do modelo RMU_{KNN90} , apenas as variáveis de tempo e custo da alternativa CAR apresentaram resultados estatisticamente significativos. No caso do modelo $RRM_{LOS_{KNN80}}$, a presença de dependência espacial nas variáveis de serviço das alternativas BUS e $RAIL$ não afetou o resultado das suas respectivas medidas de VTT em comparação com o modelo RRM_{LOS} .

Assim, cabe destacar que a abordagem RRM apresentou maior robustez na estimação da medida de VTT na ausência de controle de dependência espacial nas variáveis de nível de serviço, superestimando seu valor em 8%, em contraposição a 33% no caso dos modelos sob a abordagem RUM . Além disso, é importante notar a diferença entre as medidas para ambas as abordagens. Por exemplo, a diferença entre o pior modelo (RUM_{LOS}) e o melhor modelo (RMU_{KNN90}) é de R\$5,54 para a alternativa CAR e R\$6,09 para a alternativa BUS . Em termos relativos representam diferenças de 52,26% e 249,59%, respectivamente.

A Tabela 7 apresenta os resultados das elasticidades diretas agregadas no ponto para as três alternativas e para os parâmetros de tempo e custo de viagem dos modelos discutidos anteriormente.

Tabela 7 – Elasticidade agregada direta no ponto dos modelos LOS e $LOS + SOCIO$.						
Modelos	TT			CO		
	CAR	BUS	$RAIL$	CAR	BUS	$RAIL$
RUM_{LOS}	-0,35	-1,45	-0,63	-0,60	-2,14	-2,37
RRM_{LOS}	-0,53	-1,77	-1,09	-0,88	-1,89	-3,45
$RUM_{LOS_{KNN80}}$	-0,33	-1,39	-0,63	-0,58	-2,06	-2,37
$RRM_{LOS_{KNN80}}$	-0,30	-1,87	-1,22	-0,79	-1,89	-3,52
$RUM_{LOS+SOCIO}$	-0,27	-1,42	-0,69	-0,45	-1,99	-2,52
RMU	-0,41	-1,95	-1,22	-0,67	-1,95	-3,75
$RUM_{LOS+SOCIO_{KNN90}}$	-0,29	-1,26	-0,60	-0,50	-1,84	-2,34
$RMU_{LOS+SOCIO_{KNN90}}$	-0,33	-1,97	-1,29	-0,72	-1,94	-3,74

Em geral, as elasticidades para as alternativas não apresentam variações significativas entre modelos sob a mesma abordagem e sob abordagens diferentes, com exceção da elasticidade de tempo de viagem da alternativa $RAIL$ entre a abordagem RUM e RRM . A

elasticidade do tempo de viagem da alternativa *CAR* não é elástica em nenhum dos modelos, ou seja, tem valor absoluto menor que 1. Esse resultado indica que o aumento em 1% do tempo de viagem dessa alternativa diminui a probabilidade de escolha da alternativa *CAR* em 0,35% para o modelo RUM_{LOS} e 0,53% para o modelo RRM_{LOS} , por exemplo.

A elasticidade do tempo de viagem para alternativa *BUS* por sua vez é elástica, indicando que o aumento em 1% no tempo de viagem dessa alternativa diminui em mais de 1% a probabilidade de escolha dessa alternativa. No caso do tempo de viagem da alternativa *RAIL*, os resultados da abordagem *RUM* indicam que essa variável não é elástica, enquanto sob a abordagem *RRM* observa-se o oposto.

Para o atributo custo de viagem, no caso da alternativa *CAR* observa-se que as elasticidades variam pouco entre as abordagens e entre os modelos não espaciais e espaciais. Além disso, esse atributo não é elástico, de modo que o incremento em 1% no custo de viagem por automóvel implica na redução de 0,60% da probabilidade de escolha dessa alternativa segundo o modelo RUM_{LOS} e redução de 0,88% segundo o modelo RRM_{LOS} , por exemplo.

O custo de viagem da alternativa *BUS*, similar ao tempo de viagem dessa alternativa, também é elástico, indicando que o aumento dessa variável acarreta redução significativa da probabilidade de escolha dessa alternativa. Por fim, o custo de viagem da alternativa *RAIL* foi o atributo com maior elasticidade entre todos os atributos e alternativas. Além disso, a diferença entre abordagens foi a maior em termos absolutos. Assim, o aumento em 1% do custo dessa alternativa acarreta a redução em 2,37% e 3,45% da probabilidade de escolha da alternativa *RAIL* segundo os modelos RUM_{LOS} e RRM_{LOS} , respectivamente.

5.4 IMPLICAÇÃO DAS ABORDAGENS PARA FORMULAÇÃO DE POLÍTICAS PÚBLICAS DE TRANSPORTE

No atual contexto de limitação de recursos públicos para investimentos em grandes projetos, entender os fatores e as dinâmicas associadas à escolha da população quanto aos seus deslocamentos tem impactos diretos para a administração pública e na alocação de recursos. Historicamente, e em sua grande maioria, a análise de projetos de transportes baseada em modelos de escolha discreta usou como princípio norteador a premissa de que indivíduos buscam maximizar a utilidade das alternativas a ele disponíveis.

Como já apontado por diversos autores (CHORUS, 2012d; HESS; BIERLAIRE; POLAK, 2005) a abordagem *RUM* é fundada em teorias microeconômicas que permitem a

derivação de diversos indicadores econômicos para avaliação de projetos e análise de bem-estar social. No entanto, mais recentemente, tais premissas têm sido questionadas, e tão antigas quanto a abordagem *RUM*, abordagens de modelagem de comportamento fundadas em diferentes teorias comportamentais tem se apresentado na literatura de transportes, como é o caso da abordagem *RRM* discutida ao longo deste trabalho (CHORUS, 2012a; HANCOCK *et al.*, 2020b; HANCOCK; HESS; CHOUDHURY, 2018; HESS *et al.*, 2012).

Além disso, a partir das análises apresentadas é possível afirmar que o negligenciamento de dependência espacial nas variáveis independentes (modelo SLX) sob a abordagem *RUM* tem impactos no cálculo da medida de VTT, indicador utilizado para avaliação de viabilidade de projetos de transporte. Assim, apesar de incomum em aplicações científicas e no âmbito prático, recomenda-se a consideração desses efeitos para análise de demanda por transportes.

Os resultados apresentados indicam que a abordagem *RRM* tem um melhor desempenho que a abordagem *RUM* em termos de ajuste dos modelos estimados, sendo o modelo com melhor desempenho aquele que incorpora efeitos de dependência espacial sob a abordagem *RRM*. Além disso, os resultados de VTT para os modelos sob a abordagem *RRM* se mostraram mais robustos que as medidas sob a abordagem *RUM* na ausência de controle de dependência espacial nas variáveis independentes de nível de serviço, indicando que o negligenciamento de tais efeitos nos modelos *RRM* tem efeito menor do que na abordagem *RUM* no banco de dados utilizado.

Por fim, é válido destacar que a abordagem *RRM* busca representar comportamentos e decisões humanas, sem necessariamente se basear em fundamentos e axiomas microeconômicos, tendo como principal propriedade o efeito semi-compensatório (CHORUS, 2012d). No campo prático de planejamento e estabelecimento de políticas públicas de transporte, comportamentos compensatórios podem ser considerados em políticas de atrair e afastar (BANCO INTERAMERICANO DE DESARROLLO, 2013), por exemplo, que partem do princípio de que o incentivo ao uso do transporte público está associado à melhora dos atributos desse modo de transporte, mas também pelo desincentivo do uso dos modos individuais, ou seja, piora no desempenho dos atributos associados a esses modos (alternativas competidoras).

6 CONSIDERAÇÕES FINAIS

A evolução das relações e interações nas cidades resulta na necessidade do planejamento de sistemas de transportes que atendam as demandas pelas atividades derivadas dessas interações. Além disso, as cidades do Sul Global experimentaram crescimento demográfico associado a um planejamento focado no estímulo ao uso do automóvel, trazendo desafios que municípios, como no caso do Brasil, ainda não conseguiram enfrentar. Consequentemente, as externalidades negativas advindas de um processo intenso de motorização e incentivo de uso dos modos individuais motorizados são observadas nas cidades brasileiras através de extensos congestionamentos, alto número de sinistros de trânsito e péssima qualidade do ar, além de baixa qualidade de vida dada a desigualdade de acesso a oportunidades e serviços na cidade.

No campo do planejamento de transportes, diversas técnicas têm sido utilizadas para análise da demanda a fim de planejar sistemas mais eficientes e eficazes do ponto de vista econômico e social. Técnicas baseadas em modelos de escolha discreta são utilizadas para derivação de medidas econômicas para avaliação de viabilidade de projetos de infraestrutura e para análise de bem-estar social resultante de políticas públicas propostas.

A abordagem *RUM*, baseada na premissa de que indivíduos fazem escolhas sob um processo de maximização da utilidade das alternativas disponíveis, é a mais frequente em análise de projetos de transporte, seja essa análise realizada através da técnica sequencial de 4 etapas ou através de abordagens baseadas em atividades.

No entanto, outras teorias e abordagens comportamentais se apresentam como ferramentas de análise alternativas à abordagem *RUM*, como é o caso da abordagem *RRM*, que se baseia na Teoria Econômica do Arrependimento e preconiza que indivíduos derivam sua escolha pela minimização do arrependimento resultante da comparação binária dos atributos da alternativa considerada com as alternativas competidoras.

Além disso, a complexidade dos fenômenos sociais resultantes das interações no espaço geográfico e social fez reconhecer que as premissas de independência entre regiões ou indivíduos pode não se sustentar nos modelos econométricos utilizados para análise de fenômenos físicos e sociais. Dessa forma, os resultados dos modelos econométricos baseados nessa relação de independência podem ser enviesados pela existência de dependência espacial.

Assim, o objetivo geral desta pesquisa foi analisar a incorporação de efeitos de dependência espacial exógena (modelo SLX) sob a abordagem *RRM*, usando como estudo de caso a cidade de São Paulo, Brasil, no contexto de escolha de modo de transporte em viagens

urbanas por motivo trabalho. Além disso, como objetivos secundários buscou-se analisar o desempenho da abordagem *RRM* com a abordagem *RUM* no contexto não espacial, o que também se mostrou um tema a ser explorado na literatura no contexto do Sul Global. Além disso, a incorporação de dependência espacial em modelos de escolha discreta baseados na abordagem *RUM* ainda é um tópico pouco explorado em diversas áreas de maneira geral, mas em particular em estudos de comportamento relacionado a viagens.

Utilizou-se a pesquisa OD 2017 do METRÔ-SP para fins de modelagem, que contém dados de PR de viagens feitas na RMSP em um dia útil da semana entre o período de 2017 e 2018. Dados sociodemográficos dos indivíduos e dos domicílios foram utilizados como variáveis preditoras. Um procedimento para caracterização de atributos de nível de serviço por meio de APIs de roteamento multimodal foi sistematizado, tendo sido o tempo e custo de viagem calculados para fins de modelagem.

Modelos Logit sob as abordagens *RUM* e *RRM* foram estimados considerando apenas variáveis de nível de serviço (*LOS*) e incluindo atributos sociodemográficos (*LOS + SOCIO*). Além disso, modelos que incorporam efeitos de dependência espacial foram estimados sob ambas as abordagens, sendo apenas consideradas as variáveis de nível de serviço como variáveis defasadas espacialmente.

Os resultados dos modelos estimados indicaram que a abordagem *RRM* apresentou desempenho melhor que a abordagem *RUM* em termos de ajuste (log-verossimilhança) para os modelos *LOS* e *LOS + SOCIO*, com ou sem a presença de efeitos de dependência espacial. Além disso, os modelos espaciais apresentaram melhor desempenho que suas contrapartes não espaciais, com diferenças estatisticamente significativas. O modelo com o melhor desempenho foi o híbrido *RUM-RRM*, cujos atributos *LOS* e suas defasagens espaciais foram estimados sob a abordagem *RRM* e os atributos sociodemográficos sob a abordagem *RUM*.

As análises entre as duas especificações das matrizes de vizinhança propostas, baseadas em distância e *knn*, indicaram que os modelos baseados em *knn* apresentaram melhor desempenho em termos de ajuste que os modelos com matriz baseada em distância. Além disso, a variação sistemática para os limites máximos para caracterização de vizinhos indicou que a partir de 1.400 metros (limite de distância) e de $k = 80$ (número de vizinhos) os resultados de ajuste são estáveis e não apresentaram melhoria significativa acima desses limites.

A análise dos parâmetros indicou que tempo e custo de viagem têm efeitos negativos sobre a probabilidade de escolher uma alternativa, sendo tais parâmetros significativos estatisticamente em todos os modelos estimados.

Os parâmetros associados às características sociodemográficas dos indivíduos indicaram que ser do sexo feminino, pertencer a classes sociais mais baixas e o número moradores no domicílio aumentam a probabilidade de escolha das alternativas de transporte público coletivo (*BUS* e *RAIL*) em relação ao automóvel, e pertencer ao grupo com nível de instrução de ensino superior completo aumenta a probabilidade de escolha pelo *BUS* em relação à sua categoria de referência. Por outro lado, idade mais avançada, iniciar viagem no pico da manhã e número de viagens diárias são características que diminuem a escolha pelas alternativas *BUS* e *RAIL* em relação ao automóvel, enquanto pertencer ao grupo com grau de instrução de ensino fundamental completo até ensino superior incompleto diminui a probabilidade de escolha em relação a sua categoria de referência.

A análise dos parâmetros espaciais não foi uniforme entre os modelos *LOS*. Sob a abordagem *RUM*, as defasagens de TT_{BUS} e CO_{CAR} indicaram presença de dependência espacial positiva, enquanto a defasagem de TT_{CAR} apresentou sinal negativo e as demais variáveis não apresentaram significância estatística a 95% de confiança. No modelo RRM_{LOS} , ao contrário, apenas a defasagem de CO_{GEN} não é estatisticamente significativa a 95% de confiança, sendo positivas as defasagens de TT_{BUS} , TT_{RAIL} e CO_{CAR} e com efeito negativo a defasagem de TT_{CAR} . Para os modelos $LOS + SOCIO$, os resultados foram similares em termos de sinais e significância entre as abordagens, sendo as defasagens de TT_{CAR} e CO_{CAR} significativas a 95% de confiança em ambas as abordagens. Além disso, em ambas o efeito da defasagem de TT_{CAR} é negativo, enquanto é positivo para a defasagem de CO_{CAR} .

A análise das medidas de VTT mostrou que o negligenciamento de dependência espacial pode afetar o valor estimado pelos modelos quando há dependência espacial nos atributos utilizados para derivação dessa métrica. Nesta pesquisa, os modelos que não incorporaram dependência espacial superestimaram o valor do VTT da alternativa *CAR* em até 33%, com efeito maior nos modelos sob a abordagem *RUM* que sob a abordagem *RRM*, indicando a importância de consideração desses efeitos principalmente naquela abordagem.

Tais resultados são relevantes pois a desconSIDERAÇÃO desses efeitos pode prejudicar a valoração de projetos de transportes, fazendo com que investimentos possam não ter seus retornos esperados, dada a superestimação da disposição de indivíduos a pagar por melhorias no sistema de transporte sob avaliação. Além disso, diferenças significativas foram encontradas entre as abordagens *RUM* e *RRM* para as alternativas *CAR* e *BUS* para os valores do VTT.

Os resultados das elasticidades das variáveis de nível de serviço indicaram que tempo de viagem por ônibus e custo do transporte público (*BUS* e *RAIL*) são atributos elásticos em

todos os modelos sob ambas as abordagens analisadas, e tempo de viagem por trilhos é elástico apenas na abordagem *RRM*, indicando que políticas com enfoque nessas características têm maior capacidade de alterar a demanda por esses sistemas de transporte.

Por fim, os resultados indicaram que a abordagem *RRM* é promissora na análise do comportamento de escolha de modos de transporte no contexto da cidade de São Paulo. Ainda, do ponto de vista de políticas públicas de transporte, a abordagem *RRM* apresenta propriedades importantes para representar o comportamento humano, como o efeito compensatório. Além disso, políticas públicas como o atrair-afastar exemplificam estratégias que consideram o comportamento de comparação de atributos entre alternativas pelos indivíduos, sendo a abordagem *RRM* uma estratégia promissora para modelagem desse tipo de política.

6.1 LIMITAÇÕES E RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Como discutido anteriormente, esta pesquisa se limitou a analisar a escolha de modos de transportes apenas na cidade de São Paulo devido ao método adotado para caracterização dos níveis dos atributos de nível de serviço.

Além disso, critérios como a presença de pelo menos um automóvel no domicílio para modelagem exclui da análise parcelas socialmente vulneráveis que ainda dependem exclusivamente de transporte público e modos ativos para se deslocar na cidade. Embora Pereira *et al.* (2021) apontem que o gasto familiar com modos individuais motorizados cresceu mais entre as parcelas mais pobres que entre as parcelas mais ricas durante o período de 2002-2017, acredita-se que essa é uma limitação desta pesquisa.

Por fim, acredita-se que uma análise mais holística da abordagem *RRM* em relação à abordagem *RUM* poderia englobar a capacidade preditiva dos modelos propostos, principalmente em análises com dados de fora da amostra. No entanto, preferiu-se não realizar tal procedimento pelo fato de a amostragem no contexto espacial não ser um procedimento consolidado no que tange a aplicação de modelos de escolha discreta.

O autor acredita que muitos esforços ainda são necessários para a ampliação da compreensão das limitações, vantagens e desvantagens das abordagens e efeitos considerados nesta dissertação com foco em melhorar o entendimento sobre comportamentos relacionados e a viagens que gerem subsídios para formulação de políticas públicas mais eficazes.

Quanto à abordagem *RRM*, estudos futuros podem explorar aspectos empíricos, analisar o desempenho dessa abordagem em diferentes contextos no Sul Global, além de considerar também diferentes modos de transporte, como os modos ativos e modos de *ride-*

hailing, por exemplo. Além disso, novas formas funcionais do arrependimento já foram propostas na literatura, como os modelos P -RRM e μ RRM, que também podem ser exploradas. Além disso, como discutido no capítulo 2, a derivação da formulação para cálculo de significância de medidas de VTT para a abordagem RRM ainda não foi realizada.

No que tange a consideração de efeitos de dependência espacial em modelos de escolha discreta, e em particular em estudos de comportamento relacionado a viagens, trabalhos futuros podem explorar diferentes especificações de dependência espacial (do tipo SAR, SEM etc.) no contexto multinomial, um tópico ainda pouco explorado na área de transportes, discutido recentemente por Bahamonde-Birke (2021). Além disso, no caso de modelos do tipo SLX, a parametrização da matriz espacial é um tópico ainda não explorado, podendo ser analisado também o efeito de diferentes matrizes de vizinhança para diferentes atributos, assim como para diferentes modos de transporte. Do ponto de vista metodológico, acredita-se que trabalhos futuros devam buscar aprimorar técnicas para análise preditiva de modelos de escolha discreta que incorporam efeitos de dependência espacial, principalmente em observações fora da amostra de estimação.

No que tange ao objetivo principal desta pesquisa de incorporar efeitos de dependência espacial em modelos de escolha discreta sob a abordagem RRM, acredita-se que trabalhos futuros podem explorar a especificação do arrependimento no contexto de dependência espacial da variável dependente, como já apontado por outros autores no contexto de efeitos de influência social (MANESS; CIRILLO; DUGUNDJI, 2015), além de explorar tais efeitos em diferentes contextos empíricos.

REFERÊNCIAS

ALAMÁ-SABATER, L.; ARTAL-TUR, A.; NAVARRO-AZORÍN, J. M. Industrial location, spatial discrete choice models and the need to account for neighbourhood effects. **Annals of Regional Science**, v. 47, n. 2, p. 393–418, 2011.

ANSELIN, L. **Spatial econometrics: methods and models**. 1. ed. [s.l.] Kluwer Academic Publishers, 1988.

ANSELIN, L. Under the hood: issues in the specification and interpretation of spatial regression models. **Agricultural Economics**, v. 27, n. 3, p. 247–267, 2002.

ANSELIN, L. Spatial externalities, spatial multipliers, and spatial econometrics. **International Regional Science Review**, v. 26, n. 2, p. 153–166, abr. 2003.

ANSELIN, L. Thirty years of spatial econometrics. **Papers in Regional Science**, v. 89, n. 1, p. 3–25, 1 mar. 2010.

ARENTZE, T. A.; TIMMERMANS, H. J. P. Parametric action decision trees: Incorporating continuous attribute variables into rule-based models of discrete choice. **Transportation Research Part B: Methodological**, v. 41, n. 7, p. 772–783, 1 ago. 2007.

ASSIRATI, L. **Análise da influência da vizinhança no comportamento individual relativo a viagens através de dados em painel**. [s.l.] Universidade de São Paulo, 2018.

AUTANT-BERNARD, C. Where Do Firms Choose to Locate Their R&D? A Spatial Conditional Logit Analysis on French Data. v. 14, n. 9, p. 1187–1208, 2007.

BAHAMONDE-BIRKE, F. A brief discussion on the treatment of spatial correlation in multinomial discrete models. **Journal of Transport and Land Use**, v. 14, n. 1, p. 521–535, 20 abr. 2021.

BANCO CENTRAL DO BRASIL. **BCB - Calculadora do cidadão**. Disponível em: <<https://www3.bcb.gov.br/CALCIDADAOPublico/exibirFormCorrecaoValores.do?method=>

exibirFormCorrecaoValores&aba=1>. Acesso em: 8 out. 2021.

BANCO INTERAMERICANO DE DESARROLLO. **Guía Práctica: políticas de estacionamientos y reducción de congestión en América Latina**. [s.l: s.n.]. Disponível em: <<http://www.despacio.org/portfolio/estacionamientos-en-america-latina/>>.

BELGIAWAN, P. F. et al. Context-dependent models (CRRM, MuRRM, PRRM, RAM) versus a context-free model (MNL) in transportation studies: a comprehensive comparisons for Swiss and German SP and RP data sets. **Transportmetrica A: Transport Science**, v. 15, n. 2, p. 1487–1521, 2019.

BELL, D. E. Regret in Decision Making Under Uncertainty. **Operations Research**, v. 30, n. 5, p. 961–981, 1982.

BEN-AKIVA, M. E. et al. Hybrid Choice Models: Progress and Challenges. **Marketing Letters**, v. 13, n. 3, p. 163–175, 2002.

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

BEN-AKIVA, M. E.; SWAIT, J. Akaike Likelihood Ratio Index. **Transportation Science**, v. 20, n. 2, p. 133–136, 1986.

BERON, K. J.; VIJVERBERG, W. P. M. Probit in a Spatial Context: A Monte Carlo Analysis. In: ANSELIN, L.; FLORAX, R. J. G. M.; REY, S. J. (Eds.). . **Advances in Spatial Econometrics: Methodology, Tools and Applications**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 169–195.

BHAT, C. R. The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. **Transportation Research Part B: Methodological**, v. 42, n. 3, p. 274–303, 1 mar. 2008.

BHAT, C. R. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. **Transportation Research**

Part B: Methodological, v. 45, n. 7, p. 923–939, 2011.

BILLÉ, A. G.; ARBIA, G. Spatial Limited Dependent Variable Models: a Review Focused on Specification, Estimation, and Health Economics Applications. **Journal of Economic Surveys**, v. 33, n. 5, p. 1531–1554, 29 dez. 2019.

BILLÉ, A. G.; LEORATO, S. Partial ML estimation for spatial autoregressive nonlinear probit models with autoregressive disturbances. **Econometric Reviews**, v. 39, n. 5, p. 437–475, 2020.

BIONDI, B. et al. Modelling consumer choice through the random regret minimization model: An application in the food domain. **Food Quality and Preference**, v. 73, p. 97–109, 1 abr. 2019.

BIVAND, R. S.; PEBESMA, E.; GÓMEZ-RUBIO, V. **Applied Spatial Data Analysis with R**. New York: Springer, 2013.

BIVAND, R. S.; PIRAS, G. Comparing implementations of estimation methods for spatial econometrics. **Journal of Statistical Software**, v. 63, n. 18, p. 1–36, 2015.

BOERI, M. et al. Site choices in recreational demand: a matter of utility maximization or regret minimization? **Journal of Environmental Economics and Policy**, v. 1, n. 1, p. 32–47, 2012.

BOERI, M.; LONGO, A. The importance of regret minimization in the choice for renewable energy programmes: Evidence from a discrete choice experiment. **Energy Economics**, v. 63, p. 253–260, 1 mar. 2017.

BOERI, M.; MASIERO, L. Regret minimisation and utility maximisation in a freight transport context. **Transportmetrica A: Transport Science**, v. 10, n. 6, p. 548–560, 2014.

BROWNSTONE, D.; BUNCH, D. S.; TRAIN, K. E. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. **Transportation Research Part B: Methodological**, v. 34, n. 5, p. 315–338, 1 jun. 2000.

CALABRESE, R.; ELKINK, J. A. Estimators of binary spatial autoregressive models: A Monte

Carlo study Online Appendix. **Journal of Regional Science**, v. 54, n. 4, p. 664–687, 2014.

CALDEIRA, G. P.; ISLER, C. A. Random Regret Minimization to commuting mode choice in São Paulo, Brazil. **Under Review**, 2021.

CALDEIRA, G. P.; ISLER, C. A.; GIANNOTTI, M. Space in the right place: Travel mode choice in spatial binary discrete choice models. **To be Published**, 2021.

CARRIÓN-FLORES, C. E.; FLORES-LAGUNES, A.; GUCI, L. An estimator for discrete-choice models with spatial lag dependence using large samples, with an application to land-use conversions. **Regional Science and Urban Economics**, v. 69, n. December 2017, p. 77–93, 1 mar. 2018.

CARVALHO, E. C. S. **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.** [s.l.] Universidade de São Paulo, 2013.

CET. **Sinal Verde.** Disponível em: <<http://cetsp1.cetsp.com.br/sinalverde/osSemaforosAgora.aspx#>>. Acesso em: 8 out. 2021.

CHORUS, C. G. A New Model of Random Regret Minimization. **EJTIR**, v. 10, n. 10, p. 181–196, 1 jun. 2010.

CHORUS, C. G. et al. Modeling politicians' preferences for road pricing policies: A regret-based and utilitarian perspective. **Transport Policy**, v. 18, n. 6, p. 856–861, 1 nov. 2011.

CHORUS, C. G. What about behaviour in travel demand modelling? An overview of recent progress. **Transportation Letters**, v. 4, n. 2, p. 93–104, 2012a.

CHORUS, C. G. Regret theory-based route choices and traffic equilibria. **Transportmetrica**, v. 8, n. 4, p. 291–305, 2012b.

CHORUS, C. G. **Random Regret-based Discrete Choice Modeling: A tutorial.** 1. ed. [s.l.] Springer, 2012c.

CHORUS, C. G. Random Regret Minimization: An Overview of Model Properties and Empirical Evidence. **Transport Reviews**, v. 32, n. 1, p. 75–92, jan. 2012d.

CHORUS, C. G. A Generalized Random Regret Minimization model. **Transportation Research Part B: Methodological**, v. 68, p. 224–238, 1 out. 2014.

CHORUS, C. G.; ARENTZE, T. A.; TIMMERMANS, H. J. P. A Random Regret-Minimization model of travel choice. **Transportation Research Part B: Methodological**, v. 42, n. 1, p. 1–18, 1 jan. 2008.

CHORUS, C. G.; ROSE, J. M.; HENSHER, D. A. Regret minimization or utility maximization: It depends on the attribute. **Environment and Planning B: Planning and Design**, v. 40, n. 1, p. 154–169, 1 fev. 2013.

CHORUS, C. G.; VAN CRANENBURGH, S.; DEKKER, T. Random regret minimization for consumer choice modeling: Assessment of empirical evidence. **Journal of Business Research**, v. 67, n. 11, p. 2428–2436, 1 nov. 2014.

CORRADO, L.; FINGLETON, B. Where is the economics in spatial econometrics? **Journal of Regional Science**, v. 52, n. 2, p. 210–239, maio 2012.

CPTM. **Visão Geral | CPTM - Companhia Paulista de Trens Metropolitanos**. Disponível em: <<https://www.cptm.sp.gov.br/a-companhia/Pages/a-companhia.aspx>>. Acesso em: 10 out. 2021.

DALY, A.; HESS, S.; DE JONG, G. Calculating errors for measures derived from choice modelling estimates. **Transportation Research Part B: Methodological**, v. 46, n. 2, p. 333–341, 2012.

DAVIS, M. **Planeta favela**. [s.l.] Boitempo Editorial, 2006.

DE BEKKER-GROB, E. W.; CHORUS, C. G. Random regret-based discrete-choice modelling: An application to healthcare. **PharmacoEconomics**, v. 31, n. 7, p. 623–634, 26 jul.

2013.

DEKKER, T. Indifference based value of time measures for Random Regret Minimisation models. **Journal of Choice Modelling**, v. 12, p. 10–20, 1 set. 2014.

DEKKER, T.; CHORUS, C. G. Consumer surplus for random regret minimisation models. **Journal of Environmental Economics and Policy**, v. 7, n. 3, p. 269–286, 2018.

DELBOSC, A.; NAZNIN, F. Future life course and mobility: A latent class analysis of young adults in Victoria, Australia. **Transport Policy**, 2019.

DENATRAN. **Frota de Veículos - 2021 — Frota por Município e Tipo**. Brasília: [s.n.]. Disponível em: <<https://www.gov.br/infraestrutura/pt-br/assuntos/transito/conteudo-denatran/frota-de-veiculos-2021>>. Acesso em: 8 out. 2021.

DENNIS, E. J. et al. Consumer Reactions to E. Coli and Antibiotic Residue Recalls: Utility Maximization vs. Regret Minimization. **Frontiers in Veterinary Science**, v. 7, p. 611, 3 set. 2020.

DIAS, F. F. **Contribuições ao Estudo de Implantação de Pedágio Urbano em São Paulo São Paulo**. [s.l.] Universidade de São Paulo, 2013.

ELHORST, J. P. Applied Spatial Econometrics: Raising the Bar. **Spatial Economic Analysis**, v. 5, n. 1, 2010.

ELHORST, J. P.; VEGA, S. H. El modelo SLX: ampliación de la forma general, y sensibilidad de los desbordamientos espaciales a la especificación de la W. **Papeles de Economía Española**, v. 152, p. 34–50, 2017.

FORUM APOLLO. **Delta method for RRM models - ApolloChoiceModelling forum**. Disponível em: <<http://www.apollochoicemodelling.com/forum/viewtopic.php?f=16&t=340>>. Acesso em: 9 out. 2021.

GEURS, K. T.; VAN WEE, B. Accessibility evaluation of land-use and transport strategies:

Review and research directions. **Journal of Transport Geography**, v. 12, n. 2, p. 127–140, 1 jun. 2004.

GIBBONS, S.; OVERMAN, H. G. Mostly pointless spatial econometrics? **Journal of Regional Science**, v. 52, n. 2, p. 172–191, maio 2012.

GOETZKE, F. Are Travel Demand Forecasting Models Biased Because of Uncorrected Spatial Autocorrelation? **50th Annual North American Meetings of the Regional Science Association International**, 2003.

GOETZKE, F. **Network effects and spatial autoregression in mode choice models: Three essays in urban transportation economics**. Morgantown: West Virginia University, 2006.

GOETZKE, F. Network effects in public transit use: Evidence from a spatially autoregressive mode choice model for New York. **Urban Studies**, v. 45, n. 2, p. 407–417, 2008.

GOETZKE, F.; WEINBERGER, R. Separating contextual from endogenous effects in automobile ownership models. **Environment and Planning A**, v. 44, p. 1032–1046, 2012.

GOMIDE, A. D. Á.; MORATO, R. Instrumentos de desestímulo ao uso do transporte individual motorizado: lições e recomendações IEMA. **Iema**, p. 1–64, 2011.

GREENE, W. H. **Econometric Analysis**. 7th. ed. [s.l.] Prentice Hall, 2012.

GREENE, W. H.; HENSHER, D. A. A latent class model for discrete choice analysis: Contrasts with mixed logit. **Transportation Research Part B: Methodological**, v. 37, n. 8, p. 681–698, 1 set. 2003.

GRUS, J. **Data Science do Zero**. [s.l.] Alta Books Editora, 2016.

GUO, J. Y.; BHAT, C. R. **Modifiable areal units problem or perception in modeling of residential location choice?** Transportation Research Record. **Anais...National Research Council**, 2004

HAGHANI, M.; BLIEMER, M. C. J.; HENSHER, D. A. The landscape of econometric discrete choice modelling research. **Journal of Choice Modelling**, v. 40, p. 100303, 24 jun. 2021.

HANCOCK, T. O. et al. Quantum choice models: A flexible new approach for understanding moral decision-making. **Journal of Choice Modelling**, v. 37, p. 100235, 2020a.

HANCOCK, T. O. et al. Quantum probability: A new method for modelling travel behaviour. **Transportation Research Part B: Methodological**, v. 139, p. 165–198, 1 set. 2020b.

HANCOCK, T. O.; HESS, S.; CHOUDHURY, C. F. Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. **Transportation Research Part B: Methodological**, v. 107, p. 18–40, 1 jan. 2018.

HENSHER, D. A.; GREENE, W. H. The Mixed Logit model: The state of practice. **Transportation**, v. 30, n. 2, p. 133–176, maio 2003.

HENSHER, D. A.; GREENE, W. H.; CHORUS, C. G. Random regret minimization or random utility maximization: an exploratory analysis in the context of automobile fuel choice. **Journal of Advanced Transportation**, v. 47, n. 7, p. 667–678, 1 nov. 2013.

HESS, S. et al. Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. **Transportation**, v. 39, n. 3, p. 565–591, 24 maio 2012.

HESS, S.; BECK, M. J.; CHORUS, C. G. Contrasts between utility maximisation and regret minimisation in the presence of opt out alternatives. **Transportation Research Part A: Policy and Practice**, v. 66, n. 1, p. 1–12, 1 ago. 2014.

HESS, S.; BIERLAIRE, M.; POLAK, J. W. Estimation of value of travel-time savings using mixed logit models. **Transportation Research Part A: Policy and Practice**, v. 39, n. 2- 3 SPEC. ISS., p. 221–236, 1 fev. 2005.

HESS, S.; DALY, A.; BATLEY, R. Revisiting consistency with random utility maximisation: theory and implications for practical work. **Theory and Decision**, v. 84, p. 181–204, 2 jan. 2018.

HESS, S.; PALMA, D. **Apollo : a flexible, powerful and customisable freeware package for choice model estimation and application: User manual, version 0.1.0**. Leeds: [s.n.].

HESS, S.; PALMA, D. Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. **Journal of Choice Modelling**, v. 32, p. 100170, 1 set. 2019b.

IBGE. **IBGE | Cidades | São Paulo | São Paulo | Panorama**. Brasília: [s.n.]. Disponível em: <<https://cidades.ibge.gov.br/brasil/sp/sao-paulo/panorama>>. Acesso em: 7 out. 2021.

INMAN, J. J.; DYER, J. S.; JIANMIN, J. A generalized utility model of disappointment and regret effects on post-choice valuation. **Marketing Science**, v. 16, n. 2, p. 97–111, 1997.

JANG, S.; RASOULI, S.; TIMMERMANS, H. Incorporating psycho-physical mapping into random regret choice models: model specifications and empirical performance assessments. **Transportation**, v. 44, n. 5, p. 999–1019, 11 set. 2017.

JING, P. et al. Travel mode and travel route choice behavior based on Random Regret Minimization: A systematic review. **Sustainability (Switzerland)**, v. 10, n. 4, p. 1185, 14 abr. 2018.

KAPLAN, S.; PRATO, C. G. The application of the random regret minimization model to drivers' choice of crash avoidance maneuvers. **Transportation Research Part F: Traffic Psychology and Behaviour**, v. 15, n. 6, p. 699–709, 1 nov. 2012.

KELEJIAN, H. H.; PIRAS, G. Endogenous Weighting Matrices: Specifications and Estimation. **Spatial Econometrics**, n. 2009, p. 283–292, 2017.

KEYA, N.; ANOWAR, S.; ELURU, N. Freight Mode Choice: A Regret Minimization and Utility Maximization Based Hybrid Model. **Transportation Research Record**, v. 2672, n. 9, p. 107–119, 1 dez. 2018.

KLEIN, N. J.; SMART, M. J. Millennials and car ownership: Less money, fewer cars.

Transport Policy, v. 53, p. 20–29, 1 jan. 2017.

KLIER, T.; MCMILLEN, D. P. Clustering of auto supplier plants in the United States: Generalized method of moments spatial logit for large samples. **Journal of Business and Economic Statistics**, v. 26, n. 4, p. 460–471, 2008.

KOPPELMAN, F. S.; BHAT, C. R. **A Self Instructing Course in Mode Choice Modeling : Multinomial and Nested Logit Models**. Austin: [s.n.].

LEONG, W.; HENSHER, D. A. Contrasts of Relative Advantage Maximisation with Random Utility Maximisation and Regret Minimisation. **Journal of Transport Economics and Policy (JTEP)**, v. 49, n. 1, p. 167–186, 2015.

LESAGE, J. P. Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models. **Geographical Analysis**, v. 32, n. 1, p. 19–35, 3 set. 2000.

LESAGE, J. P.; PACE, R. K. **Introduction to Spatial Econometrics**. 1. ed. [s.l.] Taylor & Francis, 2009.

LESSA, D. A.; LOBO, C.; CARDOSO, L. Accessibility and urban mobility by bus in Belo Horizonte/Minas Gerais – Brazil. **Journal of Transport Geography**, v. 77, p. 1–10, 1 maio 2019.

LI, D. et al. Hybrid Random Regret Minimization and Random Utility Maximization in the Context of Schedule-Based Urban Rail Transit Assignment. **Journal of Advanced Transportation**, v. 2018, p. 28, 2018.

LI, M.; HUANG, H. J. A regret theory-based route choice model. **Transportmetrica A: Transport Science**, v. 13, n. 3, p. 250–272, 2017.

LOOMES, G.; SUGDEN, R. Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. **The Economic Journal**, v. 92, n. 368, p. 805, 1982.

LOPES, S. B.; BRONDINO, N. C. M.; DA SILVA, A. N. R. GIS-based analytical tools for

transport planning: Spatial regression models for transportation demand forecast. **ISPRS International Journal of Geo-Information**, v. 3, n. 2, p. 565–583, 2014.

LUCE, R. D. **Individual Choice Behavior: A Theoretical Analysis**. New York: Wiley, 1959.

MACÊDO, B. et al. Caracterização das diferenças no padrão de mobilidade de mulheres e homens em grandes cidades brasileiras. **TRANSPORTES**, v. 28, n. 4, p. 89–102, 16 nov. 2020.

MANESS, M.; CIRILLO, C.; DUGUNDJI, E. R. Generalized behavioral framework for choice models of social influence: Behavioral and data concerns in travel behavior. **Journal of Transport Geography**, v. 46, p. 137–150, 1 jun. 2015.

MARCUS, Y.; BLAINEY, S. **Defining probability-based rail station catchments for demand modelling** Defining probability-based rail station catchments for demand modelling. 48th Annual UTSG Conference. **Anais...**Bristol: 8 jan. 2016Disponível em: <https://eprints.soton.ac.uk/384539/1/UTSG_paper_young_blaineyv2.pdf>. Acesso em: 8 maio. 2020

MARTINETTI, D.; GENIAUX, G. Approximate likelihood estimation of spatial probit models. **Regional Science and Urban Economics**, v. 64, n. February, p. 30–45, 1 maio 2017.

MCFADDEN, D. P. Conditional logit analysis of qualitative choice behavior. In: **Analysis of Qualitative Choice Behavior**. 1. ed. [s.l: s.n.].

MCMILLEN, D. P. Probit With Spatial Autocorrelation. **Journal of Regional Science**, v. 32, n. 3, p. 335–348, ago. 1992.

MCMILLEN, D. P. Perspectives on spatial econometrics: Linear smoothing with structured models. **Journal of Regional Science**, v. 52, n. 2, p. 192–209, 2012.

METRÔ-SP. **Pesquisa Origem e Destino 2017**. [s.l: s.n.]. Disponível em: <<http://www.metro.sp.gov.br/pesquisa-od/resultado-das-pesquisas.aspx>>. Acesso em: 17 jul. 2020a.

METRÔ-SP. **Tarifa de integração de ônibus, CPTM e Metrô terá novo valor a partir do dia 15 de abril.** Disponível em: <<http://www.metro.sp.gov.br/noticias/11-04-2017-tarifa-de-integracao-de-onibus-cptm-e-metro-tera-novo-valor-a-partir-do-dia-15-de-abr.fss>>. Acesso em: 17 jul. 2020b.

METRÔ-SP. **Pesquisa Origem Destino 2017 - Relatório Síntese.** São Paulo: [s.n.].

METRÔ-SP. **Metrô São Paulo | Quem somos - saiba sobre transporte metropolitano.** Disponível em: <<http://www.metro.sp.gov.br/metro/institucional/quem-somos/index.aspx>>. Acesso em: 10 out. 2021.

MIYAMOTO, K. et al. Discrete choice model with structuralized spatial effects for location analysis. **Transportation Research Record**, n. 1898, p. 183–190, 2004.

MORIKAWA, T. Correcting state dependence and serial correlation in the RP/SP combined estimation method. **Transportation**, v. 21, p. 153–165, 1994.

MOZHAROVSKIY, P.; VOGLER, J. Composite marginal likelihood estimation of spatial autoregressive probit models feasible in very large samples. **Economics Letters**, v. 148, p. 87–90, 1 nov. 2016.

OOMS, J. The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. 12 mar. 2014.

OpenStreetMap. Disponível em: <<https://www.openstreetmap.org/#map=4/-15.13/-53.19>>. Acesso em: 6 out. 2021.

PACE, R. K.; LESAGE, J. P. Fast simulated maximum likelihood estimation of the spatial probit model capable of handling large samples. **Advances in Econometrics**, Advances in Econometrics. v. 37, p. 3–34, 1 jan. 2016.

PARTHAN, K.; SRINIVASAN, K. K. Investigation of Alternate Behavioural Frameworks for Mode Choice Decisions of Workers in Chennai City. **Procedia - Social and Behavioral Sciences**, v. 104, p. 573–582, 2 dez. 2013.

PARTRIDGE, M. D. et al. Introduction: Whither spatial econometrics? **Journal of Regional Science**, v. 52, n. 2, p. 167–171, maio 2012.

PAUL, P. et al. A comparison of regret-based and utility-based discrete choice modelling—an empirical illustration with hospital bed choice. **Applied Economics**, v. 50, n. 40, p. 4295–4305, 27 ago. 2018.

PEBESMA, E. Simple features for R: Standardized support for spatial vector data. **R Journal**, v. 10, n. 1, p. 439–446, 1 jul. 2018.

PEREIRA, R. H. M. Transport legacy of mega-events and the redistribution of accessibility to urban destinations. **Cities**, v. 81, p. 45–60, 1 nov. 2018.

PEREIRA, R. H. M. et al. Distributional effects of transport policies on inequalities in access to opportunities in Rio de Janeiro. **Journal of Transport and Land Use**, v. 12, n. 1, p. 741–764, 2019.

PIKE, S.; LUBELL, M. Geography and social networks in transportation mode choice. **Journal of Transport Geography**, v. 57, p. 184–193, 1 dez. 2016.

PINJARI, A. R.; BHAT, C. R. Activity-based travel demand analysis. In: ANDRÉ DE PALMA et al. (Eds.). . **A Handbook of Transport Economics**. [s.l.] Edward Elgar Publishing Ltd., 2011. p. 213–248.

PINKSE, J.; SLADE, M. E. Contracting in space: An application of spatial statistics to discrete-choice models. **Journal of Econometrics**, v. 85, n. 1, p. 125–154, 1998.

PRITCHARD, J. P. et al. An international comparison of equity in accessibility to jobs: London, São Paulo and the Randstad Transport Findings. **Transport Findings**, 2019.

R CORE TEAM. **R: The R Project for Statistical Computing** ViennaR Foundation for Statistical Computing, , 2019. Disponível em: <<https://www.r-project.org/>>. Acesso em: 17 jul. 2020

RASOULI, S.; TIMMERMANS, H. J. P. Activity-based models of travel demand: promises, progress and prospects. **International Journal of Urban Sciences**, v. 18, n. 1, p. 31–60, 2014.

RASOULI, S.; TIMMERMANS, H. J. P. **Bounded Rational Choice Behaviour: Applications in Transport**. 1. ed. [s.l.] Emerald Group Publishing Limited, 2015.

RASOULI, S.; TIMMERMANS, H. J. P. Specification of regret-based models of choice behaviour: formal analyses and experimental design based evidence. **Transportation**, v. 44, n. 6, p. 1555–1576, 2016.

RÜTTENAUER, T. Spatial Regression Models: A Systematic Comparison of Different Model Specifications Using Monte Carlo Experiments. **Sociological Methods & Research**, p. 1–32, 2019.

SÃO PAULO. **Sistema de Consulta do Mapa Digital da Cidade de São Paulo**. Disponível em: <http://geosampa.prefeitura.sp.gov.br/PaginasPublicas/_SBC.aspx#>. Acesso em: 9 out. 2021.

SHENG, M.; SHARP, B. Commuter's transport mode preferences and social network effects in New Zealand. **Journal of Transport Economics and Policy**, v. 53, n. January, p. 19–46, 2019.

SIDHARTHAN, R. et al. Model for children's school travel mode choice. **Transportation Research Record**, n. 2213, p. 78–86, 2011.

SILVEIRA SANTOS, L.; PROENÇA, I. The inversion of the spatial lag operator in binary choice models: Fast computation and a closed formula approximation. **Regional Science and Urban Economics**, v. 76, n. October 2018, p. 74–102, 2019.

SIMONSON, I. The Influence of Anticipating Regret and Responsibility on Purchase Decisions. **Journal of Consumer Research**, v. 19, n. 1, p. 105, 1992.

SLOVIC, A. D. et al. The long road to achieving equity: Job accessibility restrictions and

overlapping inequalities in the city of São Paulo. **Journal of Transport Geography**, v. 78, p. 181–193, 1 jun. 2019.

SMIRNOV, O. A. Modeling spatial discrete choice. **Regional Science and Urban Economics**, v. 40, n. 5, p. 292–298, 2010.

SMT. **Frota Contratada | Secretaria Municipal de Mobilidade e Trânsito | Prefeitura da Cidade de São Paulo**. Disponível em: <https://www.prefeitura.sp.gov.br/cidade/secretarias/mobilidade/institucional/sptrans/acesso_a_informacao/index.php?p=245214>. Acesso em: 9 out. 2021.

SPTRANS. **Tarifas | SPTrans**. Disponível em: <<https://www.sptrans.com.br/tarifas/>>. Acesso em: 10 out. 2021a.

SPTRANS. **Corredores e Faixas Exclusivas | SPTrans**. Disponível em: <<https://www.sptrans.com.br/corredores-e-faixas-exclusivas/>>. Acesso em: 9 out. 2021b.

SPTRANS. **Terminais | SPTrans**. Disponível em: <<https://www.sptrans.com.br/terminais>>. Acesso em: 9 out. 2021c.

SPTRANS. **Pontos de Parada | SPTrans**. Disponível em: <<https://www.sptrans.com.br/terminais-corredores-e-pontos-de-parada/pontos-de-parada/>>. Acesso em: 9 out. 2021d.

SWAIT, J. A non-compensatory choice model incorporating attribute cutoffs. **Transportation Research Part B: Methodological**, v. 35, n. 10, p. 903–928, 1 nov. 2001.

TAYLOR, K. A. A Regret Theory Approach to Assessing Consumer Satisfaction. **Marketing Letters**, v. 8, n. 2, p. 229–238, 1997.

THIENE, M.; BOERI, M.; CHORUS, C. G. Random Regret Minimization: Exploration of a New Choice Model for Environmental and Resource Economics. **Environmental and Resource Economics**, v. 51, n. 3, p. 413–429, 8 set. 2012.

TOBLER, W. R. A Computer Movie Simulating Urban Growth in the Detroit Region. **Economic Geography**, v. 46, p. 234, 1970.

TRAIN, K. E. **Discrete Choice Methods with Simulation**. 2nd. ed. New York: CAMBRIDGE UNIVERSITY PRESS, 2009.

VAN CRANENBURGH, S.; GUEVARA, C. A.; CHORUS, C. G. New insights on random regret minimization models. **Transportation Research Part A: Policy and Practice**, v. 74, p. 91–109, 1 abr. 2015.

VAN CRANENBURGH, S.; PRATO, C. G. On the robustness of random regret minimization modelling outcomes towards omitted attributes. **Journal of Choice Modelling**, v. 18, p. 51–70, 1 mar. 2016.

VASCONCELLOS, E. A. Urban transport policies in Brazil: The creation of a discriminatory mobility system. **Journal of Transport Geography**, v. 67, p. 85–91, 1 fev. 2018.

VEGA, S. H.; ELHORST, J. P. The slx model. **Journal of Regional Science**, v. 55, n. 3, p. 339–363, 2015.

WANG, C. H.; AKAR, G.; GULDMANN, J. M. Do your neighbors affect your bicycling choice? A spatial probit model for bicycling to The Ohio State University. **Journal of Transport Geography**, v. 42, p. 122–130, 2015.

WANG, H.; IGLESIAS, E. M.; WOOLDRIDGE, J. M. Partial maximum likelihood estimation of spatial probit models. **Journal of Econometrics**, v. 172, n. 1, p. 77–89, 1 jan. 2013.

WANG, X. C. (CARA); KOCKELMAN, K. M.; LEMP, J. D. The dynamic spatial multinomial probit model: Analysis of land use change using parcel-level data. **Journal of Transport Geography**, v. 24, p. 77–88, 1 set. 2012.

WICKHAM, H. **Tools for Working with URLs and HTTP [R package httr version 1.4.1]**Comprehensive R Archive Network (CRAN), , 2019. Disponível em: <<https://cran.r-project.org/package=httr>>. Acesso em: 17 jul. 2020

WONG, S. D. et al. A Revealed Preference Methodology to Evaluate Regret Minimization with Challenging Choice Sets: A Wildfire Evacuation Case Study. **Travel Behaviour and Society**, v. 20, p. 331–347, 1 jul. 2020.

XU, Y.; ZHOU, J.; XU, W. Regret-based multi-objective route choice models and stochastic user equilibrium: a non-compensatory approach. **Transportmetrica A: Transport Science**, v. 16, n. 3, p. 473–500, 2020.

YOUNG, M. An R Wrapper for the “OpenTripPlanner” REST API [R package **otpr version 0.4.1**]Comprehensive R Archive Network (CRAN), , 2019. Disponível em: <<https://cran.r-project.org/package=otpr>>. Acesso em: 17 jul. 2020

ZEELLENBERG, M. The use of crying over spilled milk: A note on the rationality and functionality of regret. **Philosophical Psychology**, v. 12, n. 3, p. 325–340, 1999.

ZEELLENBERG, M.; PIETERS, R. A theory of regret regulation 1.1. **Journal of Consumer Psychology**, v. 17, n. 1, p. 29–35, 1 jan. 2007.

ZHANG, J. et al. Modeling traveler choice behavior using the concepts of relative utility and relative interest. **Transportation Research Part B: Methodological**, v. 38, n. 3, p. 215–234, 1 mar. 2004.

ZHOU, Y.; WANG, X.; HOLGUÍN-VERAS, J. Discrete choice with spatial correlation: A spatial autoregressive binary probit model with endogenous weight matrix (SARBP-EWM). **Transportation Research Part B: Methodological**, v. 94, p. 440–455, 2016.

APÊNDICE 1 – RESULTADOS DOS MODELOS ESPACIAIS

A Tabela A1 e Tabela A2 apresentam os resultados dos modelos espaciais baseados na abordagem *RUM* e *RRM* com apenas variáveis de nível de serviço (*LOS*) e com atributos de nível de serviço e sociodemográficos (*LOS + SOCIO*) para as matrizes baseadas no critério de distância e *knn* respectivamente.

Tabela A1 – Modelos espaciais *LOS*.

Atributo	<i>RUM_{LOS_{D1000}}</i>		<i>RRM_{LOS_{D1000}}</i>		<i>RUM_{LOS_{D1100}}</i>		<i>RRM_{LOS_{D1100}}</i>	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
<i>ASC_{BUS}</i>	0,033	0,10	-0,4830	-1,28	0,1124	0,35	-0,3441	-0,98
<i>ASC_{RAIL}</i>	1,27	3,12*	1,2625	2,63*	1,3742	3,24*	1,4439	3,22*
Parâmetros de nível de serviço								
<i>TT_{CAR}</i>	-0,0319	-9,58*	-0,0250	-10,69*	-0,0323	-9,75*	-0,0252	-10,80*
<i>TT_{BUS}</i>	-0,0328	-25,56*	-0,0213	-26,18*	-0,0327	-25,50*	-0,0213	-26,18*
<i>TT_{RAIL}</i>	-0,0138	-10,90*	-0,0136	-16,37*	-0,0137	-10,86*	-0,0136	-16,34*
<i>CO_{CAR}</i>	-0,2523	-14,47*	-0,1730	-21,17*	-0,2502	-14,38*	-0,1724	-21,12*
<i>CO_{GEN}</i>	-0,7617	-28,74*	-0,4372	-30,79*	-0,7589	-28,70*	-0,4345	-30,96*
Parâmetros defasados espacialmente								
<i>TT_{CAR}</i>	-0,0632	-8,02*	-0,0299	-7,54*	-0,0627	-7,72*	-0,0296	-7,26*
<i>TT_{BUS}</i>	0,0082	3,68*	0,0097	5,07*	0,0080	3,56*	0,0098	5,09*
<i>TT_{RAIL}</i>	-0,0018	-0,84	0,0031	2,20*	-0,0015	-0,69	0,0036	2,53*
<i>CO_{CAR}</i>	0,0649	1,75	0,0667	2,12*	0,0619	1,63	0,0705	2,17*
<i>CO_{GEN}</i>	0,0619	0,76	-0,0043	-0,07	0,0377	0,44	-0,0300	-0,53
<i>LL_{inicial}</i>	-10255,55		-10255,55		-10255,55		-10255,55	
<i>LL_{final}</i>	-7629,42		-7600,68		-7633,62		-7604,16	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	<i>RUM_{LOS_{D1200}}</i>		<i>RRM_{LOS_{D1200}}</i>		<i>RUM_{LOS_{D1300}}</i>		<i>RRM_{LOS_{D1300}}</i>	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
<i>ASC_{BUS}</i>	0,0878	0,26	-0,3427	-0,96	0,0769	0,22	-0,3467	-0,94
<i>ASC_{RAIL}</i>	1,3483	3,05*	1,4383	3,16*	1,3466	2,94*	1,4611	3,13*
Parâmetros de nível de serviço								
<i>TT_{CAR}</i>	-0,0316	-9,51**	-0,0247	-10,58*	-0,0313	-9,40*	-0,0245	-10,48*
<i>TT_{BUS}</i>	-0,0325	-25,39*	-0,0211	-26,03*	-0,0326	-25,40*	-0,0212	-26,07*
<i>TT_{RAIL}</i>	-0,0136	-10,76*	-0,0135	-16,22*	-0,0137	-10,84*	-0,0135	-16,26*
<i>CO_{CAR}</i>	-0,2501	-14,38*	-0,1719	-21,06*	-0,2506	-14,41*	-0,1719	-21,06*
<i>CO_{GEN}</i>	-0,7586	-28,74*	-0,4346	-31,11*	-0,7590	-28,80*	-0,4347	-31,21*

Parâmetros defasados espacialmente

TT_{CAR}	-0,0689	-8,06*	-0,0324	-7,60*	-0,0726	-8,22*	-0,0337	-7,67*
TT_{BUS}	0,0074	3,21*	0,0094	4,74*	0,0076	3,23*	0,0098	4,82*
TT_{RAIL}	-0,0022	-0,96	0,0032	2,27*	-0,0021	-0,91	0,0033	2,27*
CO_{CAR}	0,0753	1,90	0,0785	2,27*	0,0875	2,14*	0,0871	2,42*
CO_{GEN}	0,0448	0,50	-0,0271	-0,47	0,0465	0,50	-0,0274	-0,46
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7627,28		-7599,51		-7622,81		-7594,77	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOSD1400}$		$RRM_{LOSD1400}$		$RUM_{LOSD1500}$		$RRM_{LOSD1500}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,1259	0,35	-0,3138	-0,87	0,0211	0,05	-0,4100	-1,04
ASC_{RAIL}	1,4512	3,04*	1,5581	3,37*	1,3098	2,64*	1,4432	2,89*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0312	-9,39*	-0,0244	-10,44*	-0,0312	-9,39*	-0,0243	-10,43*
TT_{BUS}	-0,0327	-25,46*	-0,0212	-26,13*	-0,0326	-25,39*	-0,0212	-26,08*
TT_{RAIL}	-0,0136	-10,83*	-0,0134	-16,18*	-0,0135	-10,76*	-0,0134	-16,12*
CO_{CAR}	-0,2494	-14,36*	-0,1712	-20,96*	-0,2479	-14,29*	-0,1708	-20,90*
CO_{GEN}	-0,7567	-28,78*	-0,4335	-31,28*	-0,7594	-28,93*	-0,4347	-31,38*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0749	-8,21*	-0,0348	-7,64*	-0,0777	-8,27*	-0,0358	-7,69*
TT_{BUS}	0,0080	3,35*	0,0102	4,92*	0,0076	3,10*	0,0103	4,83*
TT_{RAIL}	-0,0020	-0,84	0,0033	2,27*	-0,0028	-1,15	0,0028	1,97*
CO_{CAR}	0,0926	2,21*	0,0918	2,45*	0,0980	2,29*	0,0930	2,44*
CO_{GEN}	0,0235	0,24	-0,0399	-0,68	0,0567	0,57	-0,0226	-0,36
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7616,55		-7588,81		-7616,48		-7589,45	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOSD1600}$		$RRM_{LOSD1600}$		$RUM_{LOSD1700}$		$RRM_{LOSD1700}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,0954	0,24	-0,3075	-0,79	-0,0431	-0,10	-0,4699	-1,12
ASC_{RAIL}	1,4086	2,73*	1,5772	3,17*	1,2369	2,31*	1,3894	2,59*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0316	-9,53*	-0,0246	-10,55*	-0,0316	-9,53*	-0,0245	-10,54*
TT_{BUS}	-0,0326	-25,42*	-0,0212	-26,12*	-0,0326	-25,35*	-0,0212	-26,10*

TT_{RAIL}	-0,0135	-10,75*	-0,0133	-16,11*	-0,0135	-10,76*	-0,0133	-16,13*
CO_{CAR}	-0,2472	-14,28*	-0,1706	-20,88*	-0,2453	-14,19*	-0,1699	-20,78*
CO_{GEN}	-0,7577	-29,02*	-0,4336	-31,54*	-0,7608	-29,16*	-0,4352	-31,65*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0792	-8,30*	-0,0363	-7,66*	-0,0787	-8,10*	-0,0360	-7,54*
TT_{BUS}	0,0077	3,11*	0,0103	4,83*	0,0076	3,02*	0,0106	4,82*
TT_{RAIL}	-0,0027	-1,07	0,0030	2,07*	-0,0031	-1,21	0,0028	1,87
CO_{CAR}	0,1073	2,47*	0,1023	2,55*	0,0977	2,21*	0,0920	2,38*
CO_{GEN}	0,0411	0,39	-0,0363	-0,58	0,0708	0,66	-0,0170	-0,25
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7618,75		-7591,68		-7616,81		-7589,81	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOSP1800}$		$RRM_{LOSP1800}$		$RUM_{LOSP1900}$		$RRM_{LOSP1900}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,0240	0,05	-0,3692	-0,90	0,0249	0,05	-0,3707	-0,88
ASC_{RAIL}	1,3435	2,41*	1,5410	2,91*	1,3616	2,35*	1,5646	2,88*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0318	-9,60*	-0,0246	-10,59*	-0,0318	-9,61*	-0,0246	-10,58*
TT_{BUS}	-0,0326	-25,36*	-0,0212	-26,12*	-0,0326	-25,37*	-0,0212	-26,13*
TT_{RAIL}	-0,0134	-10,74*	-0,0133	-16,09*	-0,0133	-10,68*	-0,0132	-16,00*
CO_{CAR}	-0,2441	-14,14*	-0,1693	-20,70*	-0,2430	-14,09*	-0,1688	-20,62*
CO_{GEN}	-0,7585	-29,18*	-0,4337	-31,69*	-0,7582	-29,28*	-0,4334	-31,82*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0781	-7,91*	-0,0356	-7,31*	-0,0814	-8,04*	-0,0370	-7,41*
TT_{BUS}	0,0077	3,06*	0,0106	4,84*	0,0077	3,01*	0,0108	4,84*
TT_{RAIL}	-0,0028	-1,10	0,0031	2,04*	-0,0032	-1,21	0,0029	1,89
CO_{CAR}	0,0944	2,12*	0,0923	2,32*	-0,0032	2,27*	0,0985	2,38*
CO_{GEN}	0,0469	0,42	-0,0388	-0,58	0,0464	0,40	-0,0388	-0,57
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7617,40		-7590,17		-7613,63		-7587,23	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOS_{D2000}}$		$RRM_{LOS_{D2000}}$		$RUM_{LOS_{KNN10}}$		$RRM_{LOS_{KNN10}}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,0822	0,18	-0,3069	-0,72	-0,2169	-0,89	-0,9645	-3,01
ASC_{RAIL}	1,4448	2,40*	1,6604	3,01*	0,8862	2,91*	0,5662	1,47

Parâmetros de nível de serviço

TT_{CAR}	-0,0317	-9,60*	-0,0245	-10,57*	-0,0330	-9,77*	-0,0256	-10,89*
TT_{BUS}	-0,0326	-25,40*	-0,0212	-26,12*	-0,0323	-25,38*	-0,0211	-25,89*
TT_{RAIL}	-0,0133	-10,66*	-0,0131	-15,91*	-0,0137	-10,97*	-0,0137	-16,40*
CO_{CAR}	-0,2433	-14,10*	-0,1689	-20,61*	-0,2485	-14,26*	-0,1735	-21,13*
CO_{GEN}	-0,7565	-29,35*	-0,4326	-31,93*	-0,7810	-29,13*	-0,4491	-31,60*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0852	-8,24*	-0,0386	-7,52*	-0,0426	-6,74*	-0,0220	-6,63*
TT_{BUS}	0,0076	2,96*	0,0108	4,78*	0,0068	3,40*	0,0078	4,46*
TT_{RAIL}	-0,0033	-1,23	0,0027	1,78	-0,0026	-1,35	0,0020	1,60
CO_{CAR}	0,1180	2,54*	0,1098	2,50*	-0,0141	-0,45	0,0121	0,57
CO_{GEN}	0,0351	0,29	-0,0457	-0,66	0,1392	2,30*	0,0748	1,56
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7611,83		-7586,26		-7640,49		-7609,49	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOS_{KNN20}}$		$RRM_{LOS_{KNN20}}$		$RUM_{LOS_{KNN30}}$		$RRM_{LOS_{KNN30}}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	-0,0345	-0,12	-0,5824	-1,79	-0,0294	-0,10	-0,4535	-1,31
ASC_{RAIL}	1,1908	3,46*	1,0937	2,76*	1,1975	3,21*	1,2733	3,04*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0309	-9,19*	-0,0244	-10,32*	-0,0300	-8,92*	-0,0237	-10,02*
TT_{BUS}	-0,0323	-25,22*	-0,0210	-25,76*	-0,0322	-25,11*	-0,0209	-25,68*
TT_{RAIL}	-0,0135	-10,72*	-0,0135	-16,24*	-0,0134	-10,66*	-0,0134	-16,17*
CO_{CAR}	-0,2517	-14,41*	-0,1728	-21,06*	-0,2529	-14,46*	-0,1731	-21,04*
CO_{GEN}	-0,7671	-28,51*	-0,4407	-30,72*	-0,7659	-28,47*	-0,4388	-30,70*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0612	-8,19*	-0,0301	-7,91*	-0,0741	-9,02*	-0,0358	-8,71*
TT_{BUS}	0,0068	3,01*	0,0083	4,28*	0,0063	2,65*	0,0087	4,14*
TT_{RAIL}	-0,0026	-1,24	0,0026	1,92	-0,0034	-1,51	0,0025	1,82
CO_{CAR}	0,0383	1,07	0,0475	1,70	0,0810	2,09*	0,0817	2,42*
CO_{GEN}	0,0707	1,01	0,0072	0,14	0,0804	1,05	-0,0032	-0,06
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7625,26		-7596,56		-7618,60		-7591,43	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOS_{KNN40}}$		$RRM_{LOS_{KNN40}}$		$RUM_{LOS_{KNN50}}$		$RRM_{LOS_{KNN50}}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,0431	0,13	-0,2851	-0,83	0,1104	0,33	-0,1531	-0,44

ASC_{RAIL}	1,3050	3,25*	1,4924	3,57*	1,3698	3,22*	1,6254	3,86*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0297	-8,86*	-0,0235	-9,94*	-0,0294	-8,76*	-0,0233	-9,83*
TT_{BUS}	-0,0322	-25,09*	-0,0209	-25,67*	-0,0321	-25,04*	-0,0208	-25,61*
TT_{RAIL}	-0,0134	-10,69*	-0,0133	-16,11*	-0,0134	-10,73*	-0,0208	-16,07*
CO_{CAR}	-0,2524	-14,46*	-0,1724	-20,94*	-0,2520	-14,46*	-0,1718	-20,87*
CO_{GEN}	-0,7625	-28,45*	-0,4364	-30,83*	-0,7608	-28,46*	-0,4353	-30,95*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0820	-9,40*	-0,0395	-9,02*	-0,0900	-9,73*	-0,0431	-9,27*
TT_{BUS}	0,0065	2,60*	0,0088	4,06*	0,0060	2,36*	0,0085	3,77*
TT_{RAIL}	-0,0032	-1,37	0,0028	1,96*	-0,0033	-1,33	0,0029	1,95*
CO_{CAR}	0,1073	2,65*	0,1065	2,80*	0,1340	3,17*	0,1291	3,07*
CO_{GEN}	0,0601	0,72	-0,0260	-0,50	0,0507	0,58	-0,0369	-0,71
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7613,61		-7586,32		-7608,96		-7582,04	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOS_{KNN60}}$		$RRM_{LOS_{KNN60}}$		$RUM_{LOS_{KNN70}}$		$RRM_{LOS_{KNN70}}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,1196	0,34	-0,0998	-0,28	0,2203	0,61	0,0409	0,12
ASC_{RAIL}	1,3931	3,16*	1,7070	3,98*	1,5161	3,29*	1,8704	4,44*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0294	-8,76*	-0,0232	-9,81*	-0,0294	-8,81*	-0,0232	-9,82*
TT_{BUS}	-0,0322	-25,09*	-0,0208	-25,66*	-0,0322	-25,07*	-0,0208	-25,66*
TT_{RAIL}	-0,0134	-10,78*	-0,0132	-16,04*	-0,0135	-10,82*	-0,0132	-16,01*
CO_{CAR}	-0,2521	-14,48*	-0,1717	-20,84*	-0,2512	-14,47*	-0,1712	-20,78*
CO_{GEN}	-0,7597	-28,47*	-0,4348	-31,02*	-0,7573	-28,49*	-0,4337	-31,13*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0955	-9,94*	-0,0456	-9,43*	-0,1006	-10,13*	-0,0484	-9,62*
TT_{BUS}	0,0062	2,37*	0,0088	3,79*	0,0061	2,26*	0,0086	3,66*
TT_{RAIL}	-0,0035	-1,36	0,0028	1,85	-0,0032	-1,22	0,0030	1,95*
CO_{CAR}	0,1558	3,58*	0,1499	3,30*	0,1739	3,89*	0,1719	3,55*
CO_{GEN}	0,0536	0,58	-0,0374	-0,70	0,0306	0,32	-0,0528	-0,99
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7605,50		-7579,11		-7603,65		-7576,95	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOS_{KNN80}}$		$RRM_{LOS_{KNN80}}$		$RUM_{LOS_{KNN90}}$		$RRM_{LOS_{KNN90}}$	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,3016	0,80	0,1648	0,49	0,3572	0,92	0,2189	0,65
ASC_{RAIL}	1,6248	3,38*	2,0330	4,90*	1,6899	3,40*	2,0998	5,00*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0293	-8,80*	-0,0231	-9,79*	-0,0297	-8,94*	-0,0233	-9,92*
TT_{BUS}	-0,0322	-25,10*	-0,0208	-25,71*	-0,0321	-25,08*	-0,0207	-25,70*
TT_{RAIL}	-0,0135	-10,88*	-0,0131	-15,99*	-0,0135	-10,91*	-0,0131	-15,96*
CO_{CAR}	-0,2512	-14,50*	-0,1709	-20,78*	-0,2493	-14,43*	-0,1701	-20,68*
CO_{GEN}	-0,7544	-28,49*	-0,4325	-31,20*	-0,7529	-28,52*	-0,4318	-31,24*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,1084	-10,54*	-0,0522	-10,01*	-0,1110	-10,56*	-0,0534	-9,99*
TT_{BUS}	0,0062	2,24*	0,0088	3,66*	0,0060	2,15*	0,0087	3,55*
TT_{RAIL}	-0,0030	-1,10	0,0031	2,01*	-0,0029	-1,03	0,0032	1,98*
CO_{CAR}	0,2020	4,40*	0,2046	3,98*	0,2112	4,51*	0,2144	4,05*
CO_{GEN}	0,0116	0,11	-0,0637	-1,20	-0,0001	-0,00	-0,0700	-1,28
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-7596,88		-7570,07		-7597,29		-7570,57	

* Significativo a 95% de confiança

Tabela A1 – Continuação.

Atributo	$RUM_{LOS_{KNN100}}$		$RRM_{LOS_{KNN100}}$	
	Est.	razão-t	Est.	razão-t
ASC_{BUS}	0,4188	1,05	0,2607	0,77
ASC_{RAIL}	1,7567	3,43*	2,1293	5,02*
Parâmetros de nível de serviço				
TT_{CAR}	-0,0300	-9,05*	-0,0234	-10,01*
TT_{BUS}	-0,0321	-25,04*	-0,0207	-25,64*
TT_{RAIL}	-0,0136	-10,99*	-0,0130	-15,98*
CO_{CAR}	-0,2483	-14,38*	-0,1694	-20,58*
CO_{GEN}	-0,7519	-28,56*	-0,4316	-31,30*
Parâmetros defasados espacialmente				
TT_{CAR}	-0,1133	-10,59*	-0,0546	-10,02*
TT_{BUS}	0,0057	2,01*	0,0082	3,30*
TT_{RAIL}	-0,0026	-0,92	0,0032	1,96*
CO_{CAR}	0,2164	4,56*	0,2176	4,05*
CO_{GEN}	-0,0163	-0,15	-0,0761	-1,36
$LL_{inicial}$	-10255,55		-10255,55	
LL_{final}	-7597,39		-7570,76	

* Significativo a 95% de confiança

Tabela A2 – Modelos espaciais *LOS*.

Atributo	$RUM_{LOS+SOCIO_{D1000}}$		RMU_{D1000}		$RUM_{LOS+SOCIO_{D1100}}$		RMU_{D1100}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,74	7,25*	2,0640	4,97*	2,8570	7,35*	2,1746	5,33*
ASC_{RAIL}	3,17	6,62*	2,8825	5,66*	3,2909	6,67*	2,9887	5,93*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0357	-9,98*	-0,0274	-10,81*	-0,0363	-10,13*	-0,0277	-10,96*
TT_{BUS}	-0,0356	-26,40*	-0,0227	-26,82*	-0,0354	-26,33*	-0,0227	-26,79*
TT_{RAIL}	-0,0155	-11,92*	-0,0144	-16,74*	-0,0155	-11,90*	-0,0144	-16,74*
CO_{CAR}	-0,2441	-13,37*	-0,1687	-19,51*	-0,2422	-13,30*	-0,1680	-19,46*
CO_{GEN}	-0,8102	-29,02*	-0,4577	-31,54*	-0,8071	-28,98*	-0,4562	-31,54*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0376	-4,61*	-0,0192	-4,36*	-0,0357	-4,24*	-0,0183	-4,03*
TT_{BUS}	0,0040	1,71	0,0046	2,48*	0,0037	1,55	0,0043	2,34*
TT_{RAIL}	-0,0016	-0,72	0,0023	1,50	-0,0012	-0,55	0,0026	1,69
CO_{CAR}	0,1433	3,73*	0,1256	3,25*	0,1373	3,48*	0,1239	3,17*
CO_{GEN}	0,0733	0,87	0,0178	0,32	0,0481	0,55	0,0014	0,02
Parâmetros Sociodemográficos								
$PICO$	-0,2632	-5,07*	-0,2584	-5,02*	-0,2636	-5,08*	-0,2593	-5,04*
$IDADE$	-0,0297	-15,05*	-0,0293	-15,01*	-0,0297	-15,08*	-0,0294	-15,03*
$CLASSE_B$	0,8045	12,74*	0,8020	12,81*	0,8051	12,74*	0,8026	12,82*
$CLASSE_{C/D/}$	1,0011	10,55*	0,9890	10,50*	1,0007	10,54*	0,9886	10,49*
$SEXO_{BUS}$	0,9508	15,45*	0,9408	15,31*	0,9501	15,45*	0,9402	15,30*
$SEXO_{RAIL}$	0,7081	10,88*	0,6830	10,50*	0,7067	10,86*	0,6818	10,49*
$HOUSE$	0,1803	8,78*	0,1783	8,75*	0,1807	8,80*	0,1788	8,77*
$TRIPS_{BUS}$	-0,2768	-10,67*	-0,2723	-10,55*	-0,2772	-10,68*	-0,2726	-10,55*
$TRIPS_{RAIL}$	-0,2109	-7,92*	-0,2085	-7,85*	-0,2115	-7,94*	-0,2091	-7,87*
GI_{2BUS}	-0,0841	-0,82	-0,0725	-0,70	-0,0919	-0,89	-0,0792	-0,77
GI_{2RAIL}	0,2819	2,24*	0,3056	2,41*	0,2792	2,22*	0,3028	2,39*
GI_{3BUS}	-0,9737	-8,93*	-0,9605	-8,83*	-0,9849	-9,03*	-0,9699	-8,91*
GI_{3RAIL}	-0,1763	-1,36	-0,1270	-0,97	-0,1793	-1,38	-0,1308	-1,00
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6856,70		-6834,70		-6859,49		-6836,91	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{D1200}}$		RMU_{D1200}		$RUM_{LOS+SOCIO_{D1300}}$		RMU_{D1300}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,8397	7,09*	2,1741	5,18*	2,7989	6,81*	2,1575	4,94*
ASC_{RAIL}	3,2546	6,39*	2,9585	5,73*	3,1936	6,09*	2,9333	5,48*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0356	-9,94*	-0,0273	-10,79*	-0,0353	-9,83*	-0,0272	-10,71*
TT_{BUS}	-0,0353	-26,24*	-0,0226	-26,69*	-0,0354	-26,24*	-0,0226	-26,69*
TT_{RAIL}	-0,0154	-11,82*	-0,0143	-16,66*	-0,0155	-11,88*	-0,0144	-16,72*
CO_{CAR}	-0,2425	-13,30*	-0,1678	-19,44*	-0,2435	-13,34*	-0,1681	-19,45*
CO_{GEN}	-0,8067	-29,02*	-0,4564	-31,65*	-0,8079	-29,11*	-0,4570	-31,73*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0412	-4,65*	-0,0210	-4,47*	-0,0445	-4,86*	-0,0224	-4,65*
TT_{BUS}	0,0031	1,27	0,0040	2,08*	0,0031	1,24	0,0042	2,11*
TT_{RAIL}	-0,0018	-0,79	0,0023	1,44	-0,0019	-0,81	0,0023	1,43
CO_{CAR}	0,1517	3,65*	0,1340	3,27*	0,1637	3,84*	0,1437	3,36*
CO_{GEN}	0,0588	0,64	0,0102	0,18	0,0727	0,77	0,0172	0,29
Parâmetros Sociodemográficos								
$PICO$	-0,2608	-5,02*	-0,2569	-4,99*	-0,2590	-4,99*	-0,2554	-4,96*
$IDADE$	-0,0298	-15,10*	-0,0294	-15,06*	-0,0298	-15,10*	-0,0294	-15,05*
$CLASSE_B$	0,8027	12,70*	0,8008	12,79*	0,8026	12,70*	0,8008	12,79*
$CLASSE_{C/D/}$	0,9934	10,45*	0,9832	10,42*	0,9910	10,43*	0,9807	10,39*
$SEXO_{BUS}$	0,9497	15,44*	0,9396	15,29*	0,9492	15,43*	0,9391	15,28*
$SEXO_{RAIL}$	0,7067	10,86*	0,6818	10,48*	0,7058	10,84*	0,6811	10,47*
$HOUSE$	0,1798	8,76*	0,1780	8,73*	0,1798	8,75*	0,1780	8,73*
$TRIPS_{BUS}$	-0,2769	-10,67*	-0,2725	-10,55*	-0,2769	-10,67*	-0,2724	-10,54*
$TRIPS_{RAIL}$	-0,2111	-7,93*	-0,2089	-7,86*	-0,2113	-7,93*	-0,2090	-7,86*
GI_{2BUS}	-0,0934	-0,91	-0,0817	-0,80	-0,0902	-0,88	-0,0787	-0,77
GI_{2RAIL}	0,2799	2,22*	0,3024	2,39*	0,2835	2,25*	0,3056	2,41*
GI_{3BUS}	-0,9836	-9,01*	-0,9706	-8,92*	-0,9789	-8,97*	-0,9660	-8,87*
GI_{3RAIL}	-0,1741	-1,34	-0,1266	-0,97	-0,1680	-1,29	-0,1215	-0,93
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6857,66		-6835,68		-6856,44		-6834,40	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{D1400}}$		RMU_{D1400}		$RUM_{LOS+SOCIO_{D1500}}$		RMU_{D1500}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,8303	6,70*	2,1805	5,04*	2,7158	6,28*	2,0839	4,48*

ASC_{RAIL}	3,2763	6,05*	3,0083	5,64*	3,1050	5,58*	2,8761	5,05*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0353	-9,84*	-0,0271	-10,68*	-0,0353	-9,86*	-0,0271	-10,71*
TT_{BUS}	-0,0354	-26,27*	-0,0227	-26,75*	-0,0353	-26,20*	-0,0226	-26,67*
TT_{RAIL}	-0,0155	-11,88*	-0,0143	-16,72*	-0,0154	-11,84*	-0,0143	-16,68*
CO_{CAR}	-0,2428	-13,33*	-0,1677	-19,42*	-0,2422	-13,31*	-0,1677	-19,40*
CO_{GEN}	-0,8059	-29,09*	-0,4560	-31,76*	-0,8085	-29,23*	-0,4569	-31,88*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0457	-4,84*	-0,0229	-4,60*	-0,0479	-4,93*	-0,0238	-4,66*
TT_{BUS}	0,0036	1,41	0,0046	2,26*	0,0030	1,17	0,0045	2,16*
TT_{RAIL}	-0,0016	-0,69	0,0024	1,51	-0,0025	-1,03	0,0020	1,21
CO_{CAR}	0,1665	3,81*	0,1482	3,35*	0,1725	3,86*	0,1496	3,29*
CO_{GEN}	0,0530	0,54	0,0070	0,11	0,0937	0,92	0,0281	0,44
Parâmetros Sociodemográficos								
$PICO$	-0,2566	-4,94*	-0,2530	-4,91*	-0,2564	-4,93*	-0,2531	-4,91*
$IDADE$	-0,0298	-15,09*	-0,0294	-15,04*	-0,0298	-15,10*	-0,0294	-15,05*
$CLASSE_B$	0,7996	12,65*	0,7984	12,75*	0,8013	12,68*	0,8001	12,78*
$CLASSE_{C/D/}$	0,9847	10,36*	0,9754	10,34*	0,9873	10,39*	0,9784	10,37*
$SEXO_{BUS}$	0,9495	15,43*	0,9395	15,29*	0,9495	15,43*	0,9394	15,29*
$SEXO_{RAIL}$	0,7058	10,84*	0,6811	10,48*	0,7057	10,84*	0,6809	10,47*
$HOUSE$	0,1794	8,74*	0,1777	8,72*	0,1794	8,74*	0,1778	8,72*
$TRIPS_{BUS}$	-0,2764	-10,65*	-0,2719	-10,52*	-0,2766	-10,66*	-0,2722	-10,54*
$TRIPS_{RAIL}$	-0,2110	-7,92*	-0,2086	-7,85*	-0,2108	-7,92*	-0,2085	-7,85*
GI_{2BUS}	-0,0891	-0,87	-0,0771	-0,75	-0,0894	-0,87	-0,0777	-0,76
GI_{2RAIL}	0,2829	2,25*	0,3050	2,41*	0,2828	2,24*	0,3045	2,40*
GI_{3BUS}	-0,9716	-8,89*	-0,9580	-8,79*	-0,9736	-8,90*	-0,9602	-8,80*
GI_{3RAIL}	-0,1668	-1,28	-0,1204	-0,92	-0,1676	-1,29	-0,1221	-0,94
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6855,66		-6833,68		-6855,54		-6834,09	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{D1600}}$		RMU_{D1600}		$RUM_{LOS+SOCIO_{D1700}}$		RMU_{D1700}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,7860	6,26*	2,1818	4,68*	2,6465	5,78*	2,0289	4,04*
ASC_{RAIL}	3,1962	5,56*	2,9993	5,21*	3,0104	5,07*	2,8100	4,58*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0355	-9,93*	-0,0272	-10,77*	-0,0356	-9,94*	-0,0273	-10,78*
TT_{BUS}	-0,0354	-26,22*	-0,0226	-26,71*	-0,0353	-26,16*	-0,0227	-26,66*

TT_{RAIL}	-0,0154	-11,85*	-0,0143	-16,70*	-0,0154	-11,88*	-0,0143	-16,74*
CO_{CAR}	-0,2425	-13,35*	-0,1679	-19,45*	-0,2413	-13,30*	-0,1675	-19,37*
CO_{GEN}	-0,8067	-29,32*	-0,4559	-31,97*	-0,8096	-29,46*	-0,4572	-32,10*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0494	-5,01*	-0,0244	-4,73*	-0,0487	-4,87*	-0,0239	-4,58*
TT_{BUS}	0,0032	1,22	0,0047	2,23*	0,0031	1,17	0,0049	2,24*
TT_{RAIL}	-0,0024	-0,94	0,0021	1,30	-0,0027	-1,07	0,0020	1,17
CO_{CAR}	0,1836	4,05*	0,1624	3,43*	0,1767	3,85*	0,1528	3,24*
CO_{GEN}	0,0818	0,77	0,0196	0,30	0,1161	1,06	0,0394	0,57
Parâmetros Sociodemográficos								
$PICO$	-0,2566	-4,94*	-0,2534	-4,92*	-0,2570	-4,95*	-0,2540	-4,93*
$IDADE$	-0,0298	-15,12*	-0,0295	-15,07*	-0,0298	-15,12*	-0,0295	-15,08*
$CLASSE_B$	0,8035	12,72*	0,8025	12,82*	0,8044	12,73*	0,8030	12,82*
$CLASSE_{C/D/}$	0,9900	10,41*	0,9813	10,39*	0,9904	10,41*	0,9814	10,39*
$SEXO_{BUS}$	0,9496	15,43*	0,9395	15,29*	0,9499	15,44*	0,9398	15,29*
$SEXO_{RAIL}$	0,7054	10,84*	0,6807	10,47*	0,7049	10,83*	0,6804	10,46*
$HOUSE$	0,1794	8,74*	0,1780	8,73*	0,1787	8,70*	0,1772	8,69*
$TRIPS_{BUS}$	-0,2769	-10,67*	-0,2724	-10,55*	-0,2768	-10,67*	-0,2725	-10,55*
$TRIPS_{RAIL}$	-0,2113	-7,93*	-0,2091	-7,87*	-0,2112	-7,93*	-0,2089	-7,86*
GI_{2BUS}	-0,0932	-0,91	-0,0812	-0,79	-0,0930	-0,90	-0,0807	-0,79
GI_{2RAIL}	0,2782	2,21*	0,2998	2,37*	0,2780	2,21*	0,3008	2,37*
GI_{3BUS}	-0,9786	-8,95*	-0,9646	-8,85*	-0,9764	-8,93*	-0,9625	-8,83*
GI_{3RAIL}	-0,1741	-1,34	-0,1296	-0,99	-0,1714	-1,32	-0,1258	-0,96
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6855,34		-6833,82		-6855,83		-6834,29	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{D1800}}$		RMU_{D1800}		$RUM_{LOS+SOCIO_{D1900}}$		RMU_{D1900}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,6908	5,71*	2,0942	4,16*	2,6899	5,55*	2,0891	4,08*
ASC_{RAIL}	3,0852	5,03*	2,9114	4,67*	3,1047	4,90*	2,9286	4,61*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0358	-10,03*	-0,0274	-10,85*	-0,0357	-10,00*	-0,0273	-10,82*
TT_{BUS}	-0,0353	-26,14*	-0,0226	-26,65*	-0,0354	-26,16*	-0,0226	-26,67*
TT_{RAIL}	-0,0153	-11,85*	-0,0143	-16,69*	-0,0152	-11,78*	-0,0142	-16,63*
CO_{CAR}	-0,2403	-13,25*	-0,1671	-19,34*	-0,2400	-13,25*	-0,1671	-19,32*
CO_{GEN}	-0,8075	-29,48*	-0,4558	-32,08*	-0,8067	-29,56*	-0,4554	-32,19*

Parâmetros defasados espacialmente

TT_{CAR}	-0,0472	-4,66*	-0,0230	-4,36*	-0,0502	-4,83*	-0,0244	-4,51*
TT_{BUS}	0,0032	1,18	0,0049	2,23*	0,0032	1,18	0,0051	2,27*
TT_{RAIL}	-0,0027	-1,03	0,0021	1,22	-0,0030	-1,11	0,0019	1,10
CO_{CAR}	0,1714	3,70*	0,1498	3,15*	0,1801	3,82*	0,1581	3,21*
CO_{GEN}	0,1005	0,88	0,0244	0,34	0,0990	0,84	0,0267	0,36
Parâmetros Sociodemográficos								
$PICO$	-0,2576	-4,96*	-0,2546	-4,94*	-0,2565	-4,94*	-0,2536	-4,92*
$IDADE$	-0,0298	-15,13*	-0,0295	-15,08*	-0,0298	-15,12*	-0,0295	-15,08*
$CLASSE_B$	0,8043	12,72*	0,8027	12,81*	0,8024	12,69*	0,8015	12,79*
$CLASSE_{C/D/}$	0,9906	10,41*	0,9815	10,39*	0,9872	10,38*	0,9792	10,37*
$SEXO_{BUS}$	0,9495	15,43*	0,9395	15,29*	0,9502	15,44*	0,9402	15,30*
$SEXO_{RAIL}$	0,7047	10,83*	0,6802	10,46*	0,7048	10,83*	0,6805	10,47*
$HOUSE$	0,1788	8,70*	0,1773	8,69*	0,1783	8,67*	0,1769	8,67*
$TRIPS_{BUS}$	-0,2765	-10,66*	-0,2721	-10,53*	-0,2765	-10,66*	-0,2720	-10,53*
$TRIPS_{RAIL}$	-0,2113	-7,93*	-0,2088	-7,86*	-0,2114	-7,94*	-0,2089	-7,86*
GI_{2BUS}	-0,0941	-0,91	-0,0815	-0,79	-0,0910	-0,88	-0,0789	-0,77
GI_{2RAIL}	0,2768	2,20*	0,2988	2,36*	0,2798	2,22*	0,3006	2,37*
GI_{3BUS}	-0,9777	-8,94*	-0,9632	-8,83*	-0,9720	-8,89*	-0,9581	-8,78*
GI_{3RAIL}	-0,1746	-1,34	-0,1299	-1,00	-0,1714	-1,32	-0,1283	-0,98
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6856,88		-6835,46		-6855,59		-6834,61	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{D2000}}$		RMU_{D2000}		$RUM_{LOS+SOCIO_{KNN10}}$		RMU_{KNN10}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,7393	5,49*	2,1486	4,18*	2,3928	7,43*	1,4088	3,79*
ASC_{RAIL}	3,1878	4,86*	3,0205	4,70*	2,7151	7,00*	2,1121	4,85*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0357	-10,02*	-0,0273	-10,83*	-0,0367	-10,15*	-0,0278	-10,93*
TT_{BUS}	-0,0354	-26,20*	-0,0226	-26,69*	-0,0352	-26,30*	-0,0226	-26,70*
TT_{RAIL}	-0,0152	-11,78*	-0,0142	-16,59*	-0,0155	-12,03*	-0,0144	-16,66*
CO_{CAR}	-0,2404	-13,26*	-0,1672	-19,33*	-0,2356	-12,95*	-0,1661	-19,17*
CO_{GEN}	-0,8043	-29,62*	-0,4544	-32,27*	-0,8288	-29,34*	-0,4678	-31,86*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0522	-4,91*	-0,0253	-4,56*	-0,0212	-3,15*	-0,0116	-3,02*
TT_{BUS}	0,0033	1,21	0,0052	2,28*	0,0031	1,47	0,0036	2,19*
TT_{RAIL}	-0,0030	-1,11	0,0018	1,05	-0,0021	-1,10	0,0011	0,87
CO_{CAR}	0,1901	3,95*	0,1675	3,28*	0,0555	1,69	0,0491	2,01*

CO_{GEN}	0,0866	0,71	0,0202	0,27	0,1417	2,27*	0,0842	1,84
Parâmetros Sociodemográficos								
$PICO$	-0,2562	-4,93*	-0,2534	-4,92*	-0,2630	-5,07*	-0,2579	-5,01*
$IDADE$	-0,0298	-15,12*	-0,0295	-15,08*	-0,0296	-15,02*	-0,0293	-14,98*
$CLASSE_B$	0,8024	12,69*	0,8019	12,80*	0,8056	12,77*	0,8001	12,80*
$CLASSE_{C/D/}$	0,9887	10,40*	0,9805	10,38*	0,9928	10,49*	0,9772	10,40*
$SEXO_{BUS}$	0,9502	15,44*	0,9405	15,30*	0,9501	15,45*	0,9407	15,31*
$SEXO_{RAIL}$	0,7049	10,83*	0,6807	10,47*	0,7052	10,84*	0,6800	10,46*
$HOUSE$	0,1784	8,67*	0,1771	8,68*	0,1799	8,78*	0,1772	8,71*
$TRIPS_{BUS}$	-0,2760	-10,64*	-0,2716	-10,51*	-0,2752	-10,61*	-0,2705	-10,48*
$TRIPS_{RAIL}$	-0,2115	-7,94*	-0,2088	-7,86*	-0,2099	-7,88*	-0,2068	-7,79*
GI_{2BUS}	-0,0914	-0,89	-0,0800	-0,78	-0,0810	-0,79	-0,0693	-0,67
GI_{2RAIL}	0,2796	2,22*	0,2990	2,36*	0,2865	2,27*	0,3100	2,44*
GI_{3BUS}	-0,9705	-8,87*	-0,9577	-8,78*	-0,9730	-8,94*	-0,9574	-8,82*
GI_{3RAIL}	-0,1719	-1,32	-0,1309	-1,00	-0,1662	-1,28	-0,1162	-0,89
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6854,92		-6834,37		-6860,04		-6837,09	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{KNN20}}$		RMU_{KNN20}		$RUM_{LOS+SOCIO_{KNN30}}$		RMU_{KNN30}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,7204	7,81*	1,9517	5,11*	2,7516	7,47*	2,0994	5,08*
ASC_{RAIL}	3,1143	7,37*	2,7253	6,01*	3,0925	6,89*	2,8322	5,78*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0354	-9,84*	-0,0271	-10,66*	-0,0346	-9,62*	-0,0267	-10,48*
TT_{BUS}	-0,0351	-26,08*	-0,0225	-26,51*	-0,0350	-25,94*	-0,022	-26,37*
TT_{RAIL}	-0,0154	-11,79*	-0,0143	-16,53*	-0,0153	-11,78*	-0,0142	-16,54*
CO_{CAR}	-0,2406	-13,20*	-0,1671	-19,31*	-0,2437	-13,34*	-0,1682	-19,38*
CO_{GEN}	-0,8131	-28,74*	-0,4596	-31,19*	-0,8136	-28,75*	-0,4597	-31,22*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0357	-4,48*	-0,0188	-4,34*	-0,0477	-5,45*	-0,0246	-5,29*
TT_{BUS}	0,0026	1,10	0,0036	1,91	0,0016	0,65	0,0032	1,58
TT_{RAIL}	-0,0020	-0,94	0,0018	1,22	-0,0029	-1,26	0,0016	1,04
CO_{CAR}	0,1155	3,06*	0,1000	2,95*	0,1619	3,93*	0,1383	3,39*
CO_{GEN}	0,0709	0,98	0,0222	0,46	0,0932	1,18	0,0275	0,52
Parâmetros Sociodemográficos								
$PICO$	-0,2597	-5,00*	-0,2551	-4,96*	-0,2581	-4,97*	-0,2541	-4,94*
$IDADE$	-0,0297	-15,04*	-0,0293	-15,00*	-0,0297	-15,03*	-0,0294	-15,01*

$CLASSE_B$	0,8001	12,66*	0,7964	12,71*	0,8012	12,67*	0,7986	12,75*
$CLASSE_{C/D/}$	0,9873	10,41*	0,9745	10,34*	0,9889	10,41*	0,9776	10,37*
$SEXO_{BUS}$	0,9484	15,42*	0,9391	15,29*	0,9482	15,42*	0,9389	15,28*
$SEXO_{RAIL}$	0,7036	10,81*	0,6796	10,45*	0,7031	10,80*	0,6796	10,45*
$HOUSE$	0,1793	8,73*	0,1770	8,69*	0,1793	8,73*	0,1774	8,70*
$TRIPS_{BUS}$	-0,2761	-10,64*	-0,2717	-10,52*	-0,2767	-10,65*	-0,2723	-10,53*
$TRIPS_{RAIL}$	-0,2102	-7,89*	-0,2078	-7,82*	-0,2106	-7,90*	-0,2081	-7,83*
GI_{2BUS}	-0,0837	-0,81	-0,0724	-0,70	-0,0923	-0,90	-0,0799	-0,78
GI_{2RAIL}	0,2905	2,31*	0,3106	2,45*	0,2883	2,29*	0,3081	2,43*
GI_{3BUS}	-0,9753	-8,94*	-0,9606	-8,83*	-0,9840	-9,00*	-0,9683	-8,88*
GI_{3RAIL}	-0,1595	-1,23	-0,1133	-0,87	-0,1595	-1,23	-0,1136	-0,87
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6858,19		-6836,05		-6854,17		-6832,88	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{KNN40}}$		RMU_{KNN40}		$RUM_{LOS+SOCIO_{KNN50}}$		RMU_{KNN50}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	2,8547	7,37*	2,2941	5,60*	2,9451	7,34*	2,4419	5,96*
ASC_{RAIL}	3,2182	6,77*	3,0557	6,24*	3,2829	6,62*	3,1714	6,42*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0346	-9,61*	-0,0267	-10,46*	-0,0344	-9,56*	-0,0266	-10,42*
TT_{BUS}	-0,0351	-25,92*	-0,0224	-26,39*	-0,0350	-25,86*	-0,0224	-26,35*
TT_{RAIL}	-0,0154	-11,83*	-0,0142	-16,59*	-0,0154	-11,90*	-0,0142	-16,63*
CO_{CAR}	-0,2451	-13,42*	-0,1687	-19,47*	-0,2455	-13,45*	-0,1686	-19,47*
CO_{GEN}	-0,8095	-28,72*	-0,4575	-31,28*	-0,8077	-28,73*	-0,4567	-31,34*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0521	-5,59*	-0,0270	-5,50*	-0,0574	-5,82*	-0,0296	-5,72*
TT_{BUS}	0,0018	0,67	0,0033	1,60	0,0013	0,48	0,0029	1,36
TT_{RAIL}	-0,0026	-1,07	0,0020	1,26	-0,0024	-0,96	0,002	1,35
CO_{CAR}	0,1837	4,28*	0,1652	3,71*	0,2045	4,56*	0,1851	3,90*
CO_{GEN}	0,0712	0,83	0,0079	0,15	0,0621	0,69	-0,001	-0,01
Parâmetros Sociodemográficos								
$PICO$	-0,2575	-4,96*	-0,2534	-4,92*	-0,2565	-4,94*	-0,2529	-4,91*
$IDADE$	-0,0297	-15,03*	-0,0294	-15,00*	-0,0297	-15,05*	-0,0294	-15,03*
$CLASSE_B$	0,8004	12,65*	0,7983	12,74*	0,7968	12,59*	0,7955	12,69*
$CLASSE_{C/D/}$	0,9885	10,40*	0,9773	10,36*	0,9853	10,36*	0,9743	10,32*
$SEXO_{BUS}$	0,9478	15,41*	0,9385	15,27*	0,9475	15,40*	0,9383	15,27*
$SEXO_{RAIL}$	0,7035	10,80*	0,6804	10,46*	0,7034	10,80*	0,6809	10,47*

<i>HOUSE</i>	0,1791	8,71*	0,1775	8,70*	0,1779	8,65*	0,1764	8,64*
<i>TRIPS_{BUS}</i>	-0,2774	-10,68*	-0,2728	-10,55*	-0,2777	-10,70*	-0,2731	-10,57*
<i>TRIPS_{RAIL}</i>	-0,2109	-7,91*	-0,2082	-7,83*	-0,2111	-7,93*	-0,2083	-7,84*
<i>GI_{2BUS}</i>	-0,0870	-0,84	-0,0748	-0,73	-0,0890	-0,86	-0,0770	-0,75
<i>GI_{2RAIL}</i>	0,2945	2,33*	0,3124	2,46*	0,2964	2,35*	0,3135	2,47*
<i>GI_{3BUS}</i>	-0,9793	-8,95*	-0,9638	-8,84*	-0,9827	-8,98*	-0,9673	-8,86*
<i>GI_{3RAIL}</i>	-0,1550	-1,19	-0,1118	-0,86	-0,1504	-1,16	-0,1088	-0,83
<i>LL_{inicial}</i>	-10255,55		-10255,55		-10255,55		-10255,55	
<i>LL_{final}</i>	-6853,76		-6832,07		-6852,91		-6831,11	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	<i>RUM_{LOS+SOCIO_{KNN60}}</i>		<i>RMU_{KNN60}</i>		<i>RUM_{LOS+SOCIO_{KNN70}}</i>		<i>RMU_{KNN70}</i>	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
<i>ASC_{BUS}</i>	2,9610	7,17*	2,5195	6,07*	3,0840	7,22*	2,6668	6,56*
<i>ASC_{RAIL}</i>	3,3024	6,46*	3,2557	6,50*	3,4271	6,48*	3,4026	6,85*
Parâmetros de nível de serviço								
<i>TT_{CAR}</i>	-0,0342	-9,51*	-0,0264	-10,37*	-0,0341	-9,53*	-0,0264	-10,39*
<i>TT_{BUS}</i>	-0,0350	-25,88*	-0,0224	-26,37*	-0,0350	-25,83*	-0,0223	-26,36*
<i>TT_{RAIL}</i>	-0,0154	-11,92*	-0,0142	-26,37*	-0,0154	-11,97*	-0,0141	-16,64*
<i>CO_{CAR}</i>	-0,2467	-13,52*	-0,1690	-19,53*	-0,2465	-13,54*	-0,1688	-19,53*
<i>CO_{GEN}</i>	-0,8069	-28,77*	-0,4564	-31,40*	-0,8038	-28,76*	-0,4551	-31,41*
Parâmetros defasados espacialmente								
<i>TT_{CAR}</i>	-0,0632	-6,19*	-0,0324	-6,10*	-0,0675	-6,39*	-0,0347	-6,36*
<i>TT_{BUS}</i>	0,0013	0,46	0,0030	1,35	0,001	0,30	0,0027	1,17
<i>TT_{RAIL}</i>	-0,0027	-1,06	0,0021	1,24	-0,0025	-0,92	0,0023	1,32
<i>CO_{CAR}</i>	0,2290	4,96*	0,2093	4,20*	0,2449	5,16*	0,2287	4,44*
<i>CO_{GEN}</i>	0,0686	0,73	0,0012	0,02	0,0437	0,44	-0,0134	-0,23
Parâmetros Sociodemográficos								
<i>PICO</i>	-0,2557	-4,92*	-0,2519	-4,89*	-0,2549	-4,90*	-0,2511	-4,87*
<i>IDADE</i>	-0,0297	-15,06*	-0,0294	-15,03*	-0,0297	-15,06*	-0,0294	-15,03*
<i>CLASSE_B</i>	0,7957	12,57*	-0,0763	12,67*	0,7932	12,52*	0,7928	12,63*
<i>CLASSE_{C/D/I}</i>	0,9860	10,36*	-0,9671	10,32*	0,9843	10,33*	0,9737	10,29*
<i>SEXO_{BUS}</i>	0,9477	15,40*	0,3139	15,27*	0,9490	15,43*	0,9396	15,29*
<i>SEXO_{RAIL}</i>	0,7035	10,80*	-0,1109	10,47*	0,7044	10,81*	0,6823	10,49*
<i>HOUSE</i>	0,1778	8,64*	0,9383	8,63*	0,1775	8,63*	0,1763	8,63*
<i>TRIPS_{BUS}</i>	-0,2779	-10,70*	0,6812	-10,57*	-0,2781	-10,70*	-0,2734	-10,57*
<i>TRIPS_{RAIL}</i>	-0,2111	-7,93*	0,7950	-7,84*	-0,2115	-7,94*	-0,2086	-7,85*
<i>GI_{2BUS}</i>	-0,0880	-0,86	0,9754	-0,74	-0,0861	-0,84	-0,0739	-0,72

GI_{2RAIL}	0,2971	2,35*	0,1764	2,47*	0,3017	2,39*	0,3176	2,50*
GI_{3BUS}	-0,9816	-8,96*	-0,2733	-8,86*	-0,9820	-8,97*	-0,9661	-8,85*
GI_{3RAIL}	-0,1522	-1,17	-0,2083	-0,85	-0,1451	-1,11	-0,1061	-0,81
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	
LL_{final}	-6850,43		-6829,00		-6849,48		-6827,84	

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIO_{KNN80}}$		RMU_{KNN80}		$RUM_{LOS+SOCIO_{KNN90}}$		RMU_{KNN90}	
	Est.	razão-t	Est.	razão-t	Est.	razão-t	Est.	razão-t
ASC_{BUS}	3,1305	7,10*	2,7575	6,77*	3,1807	7,04*	2,8322	6,91*
ASC_{RAIL}	3,4610	6,33*	3,4831	6,97*	3,4937	6,22*	3,5482	7,01*
Parâmetros de nível de serviço								
TT_{CAR}	-0,0338	-9,44*	-0,0262	-10,32*	-0,0340	-9,54*	-0,0264	-10,41*
TT_{BUS}	-0,0349	-25,83*	-0,0223	-26,36*	-0,0349	-25,81*	-0,0222	-26,36*
TT_{RAIL}	-0,0154	-12,00*	-0,0141	-16,63*	-0,0154	-12,03*	-0,0141	-16,63*
CO_{CAR}	-0,2472	-13,59*	-0,1687	-19,54*	-0,2462	-13,56*	-0,1683	-19,51*
CO_{GEN}	-0,8021	-28,78*	-0,4544	-31,45*	-0,8011	-28,82*	-0,4536	-31,46*
Parâmetros defasados espacialmente								
TT_{CAR}	-0,0745	-6,83*	-0,0381	-6,83*	-0,0770	-6,89*	-0,0393	-6,88*
TT_{BUS}	0,0003	0,12	0,0024	1,01	-0,0000	-0,02	0,0020	0,86
TT_{RAIL}	-0,0027	-0,97	0,0022	1,24	-0,0028	-0,96	0,0023	1,21
CO_{CAR}	0,2678	5,51*	0,2515	4,74*	0,2768	5,59*	0,0023	4,80*
CO_{GEN}	0,0408	0,39	-0,0159	-0,27	0,0365	0,34	-0,0225	-0,37
Parâmetros Sociodemográficos								
$PICO$	-0,2540	-4,89*	-0,2500	-4,85*	-0,2544	-4,89*	-0,2504	-4,86*
$IDADE$	-0,0298	-15,08*	-0,0294	-15,05*	-0,0298	-15,08*	-0,0294	-15,05*
$CLASSE_B$	0,7893	12,45*	0,7890	12,57*	0,7897	12,45*	0,7903	12,58*
$CLASSE_{C/D/}$	0,9785	10,27*	0,9679	10,22*	0,9782	10,26*	0,9686	10,23*
$SEXO_{BUS}$	0,9493	15,43*	0,9399	15,29*	0,9503	15,45*	0,9412	15,31*
$SEXO_{RAIL}$	0,7038	10,80*	0,6818	10,48*	0,7045	10,80*	0,6828	10,49*
$HOUSE$	0,1769	8,59*	0,1758	8,60*	0,1774	8,62*	0,1764	8,63*
$TRIPS_{BUS}$	-0,2783	-10,71*	-0,2735	-10,58*	-0,2785	-10,72*	-0,2738	-10,59*
$TRIPS_{RAIL}$	-0,2116	-7,94*	-0,2085	-7,85*	-0,2118	-7,95*	-0,2089	-7,86*
GI_{2BUS}	-0,0855	-0,83	-0,0728	-0,71	-0,0892	-0,87	-0,0756	-0,73
GI_{2RAIL}	0,3048	2,41*	0,3206	2,52*	0,3043	2,41*	0,3209	2,53*
GI_{3BUS}	-0,9792	-8,93*	-0,9627	-8,81*	-0,9841	-8,97*	-0,9665	-8,84*
GI_{3RAIL}	-0,1379	-1,06	-0,0999	-0,76	-0,1368	-1,05	-0,0988	-0,75
$LL_{inicial}$	-10255,55		-10255,55		-10255,55		-10255,55	

LL_{final}	-6846,51	-6825,11	-6846,31	-6825,05
--------------	----------	----------	----------	----------

* Significativo a 95% de confiança

Tabela A2 – Continuação.

Atributo	$RUM_{LOS+SOCIOKNN100}$		RMU_{KNN100}	
	Est.	razão-t	Est.	razão-t
ASC_{BUS}	3,2218	6,96*	2,8714	6,94*
ASC_{RAIL}	3,5084	6,08*	3,5500	6,93*
Parâmetros de nível de serviço				
TT_{CAR}	-0,0342	-9,59*	-0,0265	-10,48*
TT_{BUS}	-0,0348	-25,77*	-0,0222	-26,31*
TT_{RAIL}	-0,0155	-12,11*	-0,0140	-16,66*
CO_{CAR}	-0,2457	-13,55*	-0,1678	-19,44*
CO_{GEN}	-0,8007	-28,87*	-0,4535	-31,52*
Parâmetros defasados espacialmente				
TT_{CAR}	-0,0787	-6,93*	-0,0402	-6,92*
TT_{BUS}	-0,0006	-0,22	0,0014	0,60
TT_{RAIL}	-0,0027	-0,93	0,0023	1,19
CO_{CAR}	0,2801	5,58*	0,2610	4,78*
CO_{GEN}	0,0309	0,28	-0,0261	-0,41
Parâmetros Sociodemográficos				
$PICO$	-0,2543	-4,89*	-0,2503	-4,85*
$IDADE$	-0,0297	-15,07*	-0,0294	-15,03*
$CLASSE_B$	0,7900	12,45*	0,7902	12,58*
$CLASSE_{C/D/}$	0,9795	10,27*	0,9691	10,23*
$SEXO_{BUS}$	0,9507	15,45*	0,9415	15,32*
$SEXO_{RAIL}$	0,7047	10,81*	0,6830	10,49*
$HOUSE$	0,1778	8,64*	0,1769	8,66*
$TRIPS_{BUS}$	-0,2787	-10,72*	-0,2737	-10,58*
$TRIPS_{RAIL}$	-0,2119	-7,95*	-0,2088	-7,86*
GI_{2BUS}	-0,0908	-0,88	-0,0777	-0,76
GI_{2RAIL}	0,3059	2,42*	0,3221	2,54*
GI_{3BUS}	-0,9865	-9,00*	-0,9692	-8,86*
GI_{3RAIL}	-0,1314	-1,01	-0,0948	-0,72
$LL_{inicial}$	-10255,55		-10255,55	
LL_{final}	-6846,37		-6825,16	

* Significativo a 95% de confiança