

USO DA INFERÊNCIA CAUSAL NA ANÁLISE DOS DETERMINANTES DA DEMANDA POR TRANSPORTES

Matheus Fontenelle Siqueira

Carlos Felipe Grangeiro Loureiro

Programa de Pós-Graduação em Engenharia de Transportes
Universidade Federal do Ceará

RESUMO

Apesar do uso recorrente de modelos desagregados na compreensão dos determinantes da demanda por transportes, ainda surgem questionamentos quanto à interpretação causal dos coeficientes estimados, dada a distinção entre correlação e causalidade. Nessa pesquisa de mestrado, desenvolve-se uma proposta metodológica que incorpora o arcabouço teórico da Inferência Causal na identificação de efeitos causais em decisões de viagem. Partindo de uma representação conceitual do fenômeno, constrói-se um diagrama causal que traduz as hipóteses de causa-efeito. Essa linguagem própria da Inferência Causal já permite identificar lacunas nos esforços atuais de modelagem, decorrentes sobretudo da existência de processos causais com efeitos diretos e indiretos. As hipóteses de causalidade são testadas através da construção de Modelos de Equações Estruturais, que permitem relaxar diversas premissas existentes em outras abordagens de modelagem. O método desenvolvido trará avanços na compreensão das diferenças na mobilidade entre grupos socioeconômicos, através da aplicação baseada em dados de pesquisas OD domiciliares.

1. INTRODUÇÃO

Os esforços de modelagem de transportes se baseiam no entendimento de que a demanda por transportes não é um fim em si própria, sendo derivada da necessidade de acessar as oportunidades distribuídas espacialmente (Cascetta, 2009). Essas necessidades se materializam nos fluxos de viagem, frequentemente visualizados como linhas de desejo, por diversos motivos, em diferentes horários do dia e por meio de distintos modos de transporte. A demanda por transportes pode ser interpretada como a agregação de uma série de decisões de viagem (Ortúzar e Willumsen, 2011), como: realização (produção) de viagens, frequência, horário de partida, destino, modo(s) utilizado(s), rota e encadeamento.

As abordagens desagregadas de modelagem vêm sendo adotadas há décadas para explicar as decisões de viagens. Métodos tradicionais usualmente se baseiam em modelos de regressão, muito utilizados na modelagem da produção de viagens (Ortúzar e Willumsen, 2011). O desenvolvimento da teoria da utilidade aleatória e dos modelos de escolha discreta (Ben-Akiva e Lerman, 1985) possibilitou avanços consideráveis, permitindo modelar o comportamento dos indivíduos como um processo de escolha entre alternativas discretas. Mais recentemente, destaca-se a utilização de modelos de equações estruturais (Golob, 2003) e modelos híbridos (Ben-Akiva *et al.*, 2002), abordagens multivariadas que incorporam as variáveis latentes na representação de variáveis contínuas e discretas, respectivamente.

O poder explicativo desses modelos se fundamenta na interpretação dos resultados da calibração de equações de regressão ou funções de utilidade. Os coeficientes estimados possuem interpretação de efeito parcial “*ceteris paribus*” (Wooldridge, 2013), representando a associação entre cada variável *explicativa* e a variável *dependente*, quando controlados os efeitos das demais variáveis. Isso permite aprofundar a compreensão sobre um fenômeno através da identificação do tipo de associação (positiva ou negativa), das variáveis de maior importância relativa e das elasticidades. Entretanto, surgem questionamentos quanto ao significado causal dos coeficientes estimados, simbolizados na clássica afirmação “correlação não é causalidade”, especialmente ao se tratar de dados observacionais.

A investigação de efeitos causais em decisões de viagem não é trivial, e o ferramental acima citado mostra-se insuficiente para distinguir relações causais de relações associativas espúrias. Essas lacunas, compartilhadas por outras áreas do conhecimento, vem sendo preenchidas com a incorporação de um arcabouço teórico recente, denominado Inferência Causal. Pearl (2001) instituiu, com essa abordagem, as bases teóricas para o estudo das relações de causa-efeito através de objetos matemáticos com semântica bem definida e lógica bem justificada. Essa teoria mostra, dentre outros, os motivos pelos quais associação e causalidade são conceitos distintos, e porque associação pode existir mesmo sem haver causalidade. Esse robusto paradigma para investigar relações de causa-efeito ainda permanece inexplorado na compreensão das decisões de viagem, o que reacende questionamentos sobre os reais determinantes da demanda por transportes e sobre os seus efeitos relativos. Diante disso, deriva a principal questão de pesquisa desse trabalho: *como investigar relações de causa-efeito nas decisões de viagem sob a perspectiva da inferência causal?*

2. OBJETIVOS

A partir da discussão anterior, estabelece-se que o objetivo geral desse esforço de pesquisa de mestrado consiste em desenvolver uma proposta metodológica de análise das relações de causa-efeito nas decisões de viagem, incorporando o paradigma da inferência causal. O objetivo geral será alcançado por meio dos seguintes objetivos específicos:

- a) Constituir uma representação conceitual dos determinantes das decisões de viagem, identificando os fatores explicativos e suas possíveis inter-relações;
- b) Identificar lacunas na modelagem da demanda por transportes sob o ponto de vista da representação do fenômeno e da teoria da Inferência Causal;
- c) Desenvolver um método de modelagem que incorpore o paradigma da Inferência Causal no estudo da causalidade na demanda por transportes;
- d) Mostrar a eficácia do método na investigação de hipóteses de causalidade, aplicando-o na compreensão das diferenças na mobilidade entre grupos socioeconômicos.

3. REVISÃO DA LITERATURA E PROPOSTA METODOLÓGICA

A proposta metodológica aqui desenvolvida é construída em torno da teoria da Inferência Causal desenvolvida por Pearl (2001). Essa abordagem assume que a demanda por transportes possui, em sua essência, um processo causal desconhecido – e possivelmente complexo – que *explica* os conjuntos de dados observados empiricamente em pesquisas de campo, por exemplo. Nesse esforço confirmatório, as hipóteses de causalidade formuladas *a priori* sobre os determinantes da demanda por transportes são corroboradas a partir da utilização de pesquisas OD domiciliares, que podem ser interpretadas como a materialização, ou “retrato”, dos processos causais existentes. Para tanto, distinguem-se três principais etapas metodológicas, descritas a seguir, que partem da inferência causal e culminam na inferência estatística, com a possível confirmação das hipóteses formuladas.

A primeira etapa dessa pesquisa envolveu a construção de um mapa mental que representasse os determinantes das decisões de viagem, com base em uma revisão da literatura. Partiu-se da representação de van Wee (2002), que propõe uma explicação com base em três grupos de fatores: (i) localização das atividades; (ii) desejos e necessidades das pessoas; e (iii) resistências (impedâncias) de viagem. As *necessidades e desejos* agregam as decisões de atividade dos indivíduos, relativas à vida profissional, às atribuições do lar ou a necessidades individuais. Sua complexidade envolve elementos socioeconômicos (notadamente o padrão de vida), culturais, intra-relações familiares, mas também às limitações nos orçamentos de tempo (van Wee *et al.*, 2013). A *localização das atividades* e as *resistências de viagem* representam duas dimensões

importantes da acessibilidade. Enquanto a primeira traduz a proximidade às oportunidades, o que é condicionado pelas decisões locais no ambiente urbano, a segunda pode ser interpretada como custo generalizado de transporte, agregando a disponibilidade de veículos privados, fatores monetários, tempos de viagem, além da valoração de aspectos qualitativos, como conforto e a confiabilidade das alternativas (Annema, 2013).

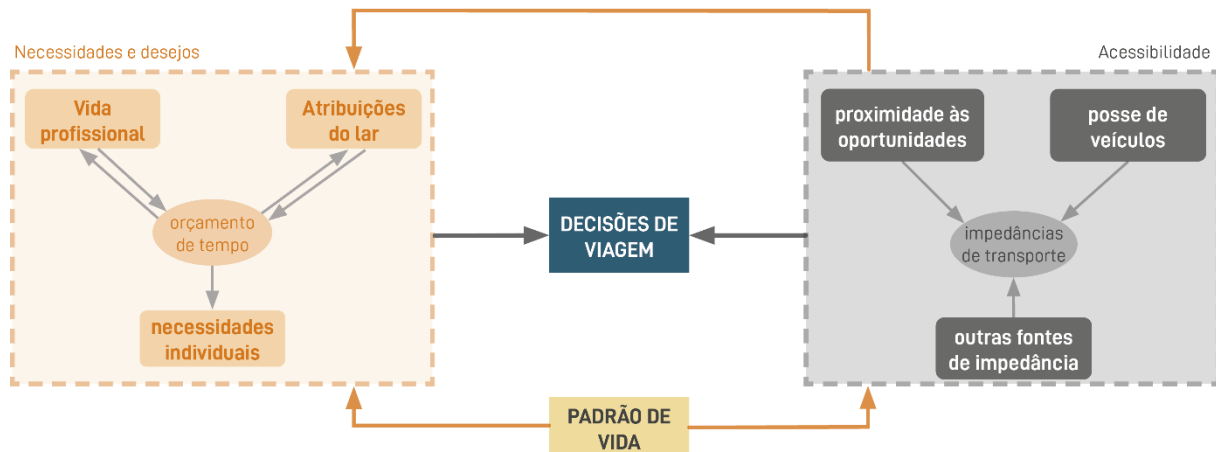


Figura 1: Mapa mental dos determinantes das decisões de viagem

O mapa mental forneceu os subsídios necessários para a segunda etapa metodológica de *construção do diagrama causal*, que é a linguagem própria para o estudo da causalidade (Pearl, 2001). O diagrama causal é uma representação gráfica de um processo causal hipotetizado, onde os vértices correspondem a variáveis e as setas representam o sentido da causalidade entre duas variáveis. Busca-se *explicar* as relações associativas entre os conjuntos de variáveis partindo de três situações básicas: se há correlação entre X e Y; ou X causa Y, ou Y causa X; ou X e Y possuem uma causa em comum. Nesse esforço, o diagrama foi construído de modo que não houvesse relações de retroalimentação, compondo um diagrama direcionado acíclico, que possui propriedades matemáticas que favorecem a identificação dos efeitos causais (Shiple, 2016). Isso implicou em assumir como exógenas, por exemplo, as decisões locais, o que não é limitante dado o horizonte temporal de longo prazo dos efeitos de intervenções no sistema de transportes sobre o uso do solo.

A partir da representação causal desenvolvida, identificam-se lacunas importantes no atual processo de modelagem da demanda. Os aspectos socioeconômicos, como estágio de vida e posse veicular, possuem impacto *indireto* nas decisões de viagem, atuando sobre as decisões de atividade e sobre a acessibilidade. As abordagens baseadas em escolha discreta não só assumem esses efeitos como diretos (Cascetta, 2009), como ignoram as relevantes dinâmicas familiares e as negociações involuntárias realizadas pelos indivíduos sobre o seu orçamento de tempo. Outra implicação importante no estudo das decisões de mobilidade é reconhecer que a acessibilidade tem um componente locacional (domiciliar), mas também modal (individual), o que implica em possíveis diferenças na acessibilidade entre indivíduos de uma mesma família. Por fim, identificam-se efeitos indiretos da acessibilidade, que decorrem da compreensão da sua função social, limitando ou induzindo o acesso a oportunidades (van Wee *et al.*, 2013).

Outra contribuição da construção do diagrama causal foi reconhecer que parte da associação entre as decisões de viagem e os seus determinantes se dá devido à existência de causas em comum, que geram correlações espúrias e podem trazer viés às estimativas. Uma das principais fontes desse *confundimento* é o padrão de vida, que influencia tanto as necessidades e desejos

(induzindo atividades) quanto a acessibilidade, através da posse e uso de veículos motorizados. Para identificar efeitos causais, torna-se imprescindível controlar esse tipo de variável, o que pode ser feito por controle estatístico ou segmentação amostral (Shipley, 2016).

A terceira, e última, etapa metodológica consiste na tradução do diagrama causal em um conjunto de testes estatísticos, ou seja, realiza-se a ponte entre a inferência causal (diagramas causais) e a inferência estatística (testes de hipótese). Para confirmar as *explicações* hipotetizadas através do diagrama causal, dois métodos principais emergem: os testes de *d-separation* (Shipley, 2016) e o uso de modelos estatísticos. A opção pela modelagem estatística e, em particular, pela Modelagem de Equações Estruturais decorreu, primeiramente, do seu caráter multivariado, possibilitando identificar efeitos diretos e indiretos, como os descritos anteriormente. A modelagem de variáveis latentes, que buscam representar conceitos abstratos ou de difícil mensuração, permite incorporar elementos como “orçamento de tempo” e “atribuições do lar”. Por fim, esse ferramental é versátil e robusto o suficiente para relaxar premissas da regressão clássica (Bollen, 1989), permitindo, com a estimação dos modelos, interpretações causais sobre os coeficientes estruturais, tendo em vista o controle prévio das fontes de endogeneidade e de correlações espúrias (não-causais).

4. RESULTADOS ESPERADOS

Espera-se, com esse esforço de pesquisa, desenvolver um método que permita, através da incorporação da Inferência Causal, preencher lacunas metodológicas na investigação de relações de causa-efeito nas decisões de viagem. O método construído terá a sua eficácia demonstrada através de uma aplicação, a qual será voltada para a compreensão das diferenças na mobilidade entre grupos socioeconômicos. Acredita-se que as contribuições metodológicas desse trabalho têm implicações importantes sobre o processo de planejamento, fornecendo um ferramental operacional para o diagnóstico de causalidade em diversas problemáticas. Adicionalmente, espera-se preencher lacunas fenomenológicas na compreensão dos determinantes da demanda por transportes, evidenciando possíveis heterogeneidades de processos causais entre grupos socioeconômicos, notadamente estratos de gênero e renda. Entende-se que esse conhecimento pode subsidiar a formulação de políticas de transporte mais assertivas no combate às inequidades socioeconômicas das grandes cidades brasileiras.

REFERÊNCIAS BIBLIOGRÁFICAS

- Annema, J. A. (2013) Transport resistance factors: time, money and effort. B. van Wee, J. A. Annema, & D. Banister (Eds), The transport system and transport policy: an introduction. Edward Elgar Publishing Ltd.
- Ben-Akiva, M., e Lerman, S. R. (1985) Discrete Choice Analysis: Theory and Application to Travel Demand. Journal of Business & Economic Statistics (1a. Ed., Vol. 6). The MIT Press, Cambridge, MA.
- Ben-Akiva, M., Mcfadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D. S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A., e Munizaga, M. A. (2002) Hybrid Choice Models: Progress and Challenges. Marketing Letters, 13(3), 163–175.
- Bollen, K. A. (1989) Structural Equations with Latent Variables. Contemporary Sociology (Vol. 20). John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Cascetta, E. (2009) Transportation Systems Analysis, Models and Applications (2a. Ed., Vol. 10). Springer.
- Golob, T. F. (2003) Structural Equation Modeling for Travel Behavior Research. Transportation Research Part B: Methodological, 37(1), 1–25.
- Ortúzar, J. de D., e Willumsen, L. G. (2011) Modelling Transport (4th Ed.). John Wiley and Sons, Ltd.
- Pearl, J. (2001) Causality: Models, Reasoning, and Inference (First Ed.). Cambridge University Press, New York.
- Shipley, B. (2016) Cause and Correlation in Biology. (2nd Ed.). Cambridge University Press, Cambridge.
- Van Wee, B. (2002) Land use and transport: Research and policy challenges. Journal of Transport Geography, 10(4), 259–271.
- Van Wee, B., Annema, J. A., e Banister, D. (2013) The transport system and transport policy: an introduction., 399.
- Wooldridge, J. M. (2013) Introductory econometrics: a modern approach. Cengage Learning., 899.