

Uma abordagem markoviana para a análise da dinâmica da pobreza multidimensional

Letícia Verona¹, Maria Luiza Machado Campos¹ e Marcos Borges¹

¹ Instituto de Computação – Universidade Federal do Rio de Janeiro
Avenida Athos da Silveira Ramos, 274 - Cidade Universitária,
Rio de Janeiro, RJ, Brasil - CEP: 21941-916

leticiaverona@ufrj.br, mluiza@ppgi.ufrj.br, mborges@ppgi.ufrj.br

Abstract. *Research analyzing social dynamics using multidimensional data often faces a trade-off between the complexity of variable dependencies and the need for straightforward insights for decision-makers. To bridge this gap, this research integrates the Multidimensional Poverty Index with Markov Chains to model household trajectories within the state space of poverty deprivations, enabling the simulation of policy applications and the forecasting of their cascading effects on deprivations over time. Applied to a longitudinal panel of over one million households from the Brazilian Continuous National Household Sample Survey, the framework identified attractor states, forecasted poverty trends, and simulated the cumulative impact of interventions, providing actionable policy guidance grounded in probabilistic analysis.*

Resumo. *Pesquisas que analisam dinâmicas sociais por meio de dados multidimensionais frequentemente enfrentam um contraponto entre a complexidade das interdependências entre variáveis e a necessidade de informações objetivas para tomadores de decisão. Para preencher esta lacuna, esta pesquisa integra o Índice de Pobreza Multidimensional com Cadeias de Markov para modelar trajetórias familiares no espaço de estados das privações de pobreza, permitindo simular a aplicação de políticas e prever seu efeito em cascata sobre as privações ao longo do tempo. Aplicado a um painel longitudinal de mais de um milhão de domicílios da Pesquisa Nacional por Amostra de Domicílios Contínua, o método identificou estados atratores, previu tendências de pobreza e simulou o impacto cumulativo de intervenções, oferecendo diretrizes de ação baseadas em análises probabilísticas.*

1. Introdução

Mais de 1,1 bilhão de pessoas em todo o mundo vivem em pobreza extrema, enfrentando privações complexas que vão além da escassez de renda [UNDP; OPHI, 2024]. A pobreza foi tradicionalmente quantificada por medidas monetárias, unidimensionais. Entretanto, a partir dos anos 1980, Amartya Sen decompôs os componentes do sistema complexo que define a pobreza na medição das *capacidades*. Para Sen (1981, 2008), a capacidade de uma pessoa corresponde à liberdade que ela tem de viver um tipo de vida ou outro e depende de uma variedade de fatores, incluindo características pessoais e arranjos sociais. A pobreza deve ser medida pela identificação de níveis minimamente aceitáveis de capacidades básicas (ex.: capacidade de escapar da morte evitável), abaixo dos quais as pessoas são consideradas em situação de privação. As *privações*, portanto, atuam como barreiras e afetam diretamente a liberdade das pessoas. Com base nos conceitos de sua

obra, foram desenvolvidos métodos para medir a pobreza multidimensional. O método Alkire-Foster [Alkire; Foster, 2011] tornou-se prevalente e o Índice de Pobreza Multidimensional (IPM) é adotado por mais de 80 países e por instituições como o Banco Mundial e as Nações Unidas. Os blocos fundamentais do IPM são indicadores binários de privação (privado/não privado), agregados em um índice. Duas famílias podem compartilhar a mesma intensidade de pobreza e vivenciar privações distintas, com soluções variadas. As privações têm efeito em cascata, e antecipar como o alívio de uma afeta a mobilidade em outras pode otimizar recursos. A ampla adoção do IPM gerou muitos dados sobre a pobreza multidimensional disponíveis em diversos países, despertando interesse na literatura científica quanto à sua interpretação e uso. Assim como a pobreza, esses dados são complexos, envolvendo variáveis interdependentes e difíceis de analisar. Para lidar com esses volumes e entender as relações entre privações ao longo do tempo, são necessárias ferramentas computacionais que capturem a mobilidade social e convertam em conhecimento útil para políticas. Desse desafio emerge a questão central desta pesquisa: dada a dinâmica interconectada e não determinística da pobreza multidimensional, como a eficácia de médio prazo de diferentes políticas pode ser prevista e comparada para melhor informar a alocação de recursos e ações futuras?

Estudos avançados a respeito da dinâmica da pobreza multidimensional tentam inferir a dinâmica do sistema a partir de mudanças ao longo do tempo na contagem agregada de privações, e a estrutura a partir das relações entre indicadores individuais [Amornbunchornvej et al., 2023; García Arancibia e Girela, 2024]. Processos estocásticos são frequentemente usados para modelar a incerteza, e a modelagem do espaço de estados determina a profundidade da análise. Alguns estudos utilizam Cadeias de Markov, capturando as probabilidades de transição estocásticas, porém comprimem as privações individuais em faixas amplas de IPM [Ramírez; Kadelbach; Mata, 2021; Wang et al., 2022]. Embora essas abordagens permitam cálculos preditivos, elas obscurecem a composição específica da pobreza. Suppa (2025) deu passos significativos ao analisar a dinâmica de entrada e saída de privações por meio de uma abordagem de regressão em painel longitudinal. O estudo preserva a interpretabilidade ao nível do indicador, mas acompanha cada um separadamente, monitorando entradas e saídas de forma independente, sem analisar a configuração conjunta. Neste preciso ponto que combina a interpretabilidade individual dos indicadores e a captura da estrutura conjunta das privações é que o presente estudo se posiciona. Para preencher essa lacuna, desenvolvemos um método analítico baseado em Cadeias de Markov, em que o espaço de estados é definido pelos perfis de privação vivenciados pelos domicílios. A abordagem permite: (i) analisar a dinâmica da pobreza e identificar combinações de privações em que as famílias tendem a se concentrar; (ii) prever as trajetórias futuras das famílias e sua saída ou persistência na pobreza; e (iii) simular intervenções de políticas públicas e avaliar seus efeitos. Ao definir o espaço de estados a partir de perfis de privação, em vez de contagens ou categorias, o modelo fornece informações para políticas de intervenção mais eficientes, combinando complexidade preditiva e explicabilidade para os decisores.

Para demonstrar a utilidade do método, ele foi aplicado à análise da pobreza laboral no Brasil, entre 2019 e 2025, em um painel longitudinal com dados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC), conduzida pelo IBGE (Instituto Brasileiro de Geografia e Estatística). Analisando mais de 1 milhão de domicílios, identificamos combinações de privações que atuam como atratores do sistema, identificamos tendências da pobreza no médio prazo e simulamos três políticas

diferentes para comparar seu potencial de impacto. O trabalho insere-se na e-Ciência aplicada às Ciências Sociais, ligando o processamento de grandes volumes de dados públicos à formulação de políticas baseadas em evidências. Embora baseado em dados do Brasil, o método é aplicável a qualquer IPM nacional ou regional, sendo flexível a diferentes definições de privação. A Seção 2 formaliza a metodologia e sua validação. Na Seção 3, descrevemos o experimento; na Seção 4, descrevemos e discutimos os resultados. A Seção 5 conclui.

2. Metodologia

O modelo parte de duas entradas: (i) a estrutura de modelagem Alkire-Foster (que define indicadores, pesos e limiares de pobreza); e (ii) dados longitudinais com múltiplas observações das mesmas unidades familiares ao longo do tempo.

Modelamos a variável aleatória X como a combinação dos indicadores de privação que uma família enfrenta, denominada perfil de privação. Para cada família ω a variável X fornece um vetor $(X_1, X_2, \dots, X_n)$ onde $X_i = 1$ se a família estiver privada no i -ésimo indicador, $X_i = 0$ caso contrário. O espaço amostral de X é definido na Equação 1.

$$(1) X = \Omega \rightarrow \{0,1\}^n$$

O espaço amostral Ω contém todas as 2^n possíveis combinações de privação e não privação para os n indicadores. Partindo da premissa que a dinâmica da variável X pode ser aproximada por um processo estocástico discreto com a propriedade de falta de memória, caracterizamos uma Cadeia de Markov $\{X_t\}_{t \in N}$. O espaço de estados S coincide com o espaço amostral Ω e cada estado $s_i \in S$ é definido pelo vetor binário de privação observado no instante t . Seja P a matriz de transição onde P_{ij} representa a probabilidade de uma família transitar do estado i para o estado j em um passo de tempo. P é uma matriz quadrada de dimensão $2^n \times 2^n$ e estocástica por linhas ($\sum_j P_{ij} = 1$). Cada elemento é estimado a partir dos dados com a Equação 2.

$$(2) P_{ij} = \frac{\text{Número de transições observadas de } i \text{ para } j}{\text{Número de transições a partir de } i}$$

De posse da matriz de transição P , é possível calcular o vetor de distribuição de estado estacionário π resolvendo a Equação 3. O vetor π indica a probabilidade de uma família estar em cada estado em um momento aleatório de observação no longo prazo, assumindo que a dinâmica atual se mantenha. Estados com alta probabilidade no vetor estacionário são de alto tráfego, onde as famílias tendem a se concentrar. Estados de pobreza com alta probabilidade estacionária indicam armadilhas estruturais de pobreza do sistema.

$$(3) \pi = \pi P$$

3. Experimento: Pobreza Laboral no Brasil (2019-2025)

Optamos por especializar o método Alkire-Foster em uma formulação de pobreza laboral, restrita às privações ligadas ao trabalho, à proteção social e à formação de capital humano. A escolha alinha o experimento à vocação da PNADC, cuja estrutura de painel rotativo trimestral oferece a granularidade necessária. Além disso, estes indicadores capturam as capacidades produtivas que sustentam a saída estrutural da pobreza a longo prazo.

3.1. Dados utilizados e contexto

A PNADC tem como objetivo central aferir as condições de inserção da população no mercado de trabalho. Opera com um esquema de painel rotativo, em que cada domicílio selecionado é entrevistado cinco vezes consecutivas, com periodicidade trimestral. Essa estrutura acompanha a trajetória de uma mesma família ao longo de 15 meses, criando janelas de observação suficientes para a modelagem proposta. O experimento abrange o período do 1º trimestre de 2019 ao 4º trimestre de 2025. A escolha deste recorte submete o modelo a testes de estresse sob condições macroeconômicas e políticas distintas. O período abrange o choque da pandemia de COVID-19 (2020-2021) e a subsequente recuperação, além de uma transição significativa na orientação político-econômica do país, incluindo as administrações de Bolsonaro (2019-2022) e de Lula (2023-presente), o que oferece um cenário rico para testar a sensibilidade da matriz de transição¹. O conjunto de dados final contém aproximadamente 4,5 milhões de transições de estado.

3.2. Índice de Pobreza Laboral seguindo o método Alkire-Foster

A seleção de indicadores de privação para medir a pobreza laboral no Brasil foi um exercício discricionário, fundamentado em métricas oficiais do IPM utilizadas na América Latina [Santos; Villatoro, 2018]. Analisamos todos os índices nacionais de pobreza multidimensional empregados em países latino-americanos (Chile, Colômbia, Costa Rica, El Salvador, Equador, Guatemala, Honduras, Panamá, Paraguai, República Dominicana e Uruguai) e selecionamos os indicadores relacionados ao trabalho, à proteção social e à formação de capital humano. Verificamos a viabilidade de apuração de cada indicador com base nos dados da PNADC, e a lista final resultante inclui cinco indicadores de privação: Alfabetização, Escolaridade de Adultos, Emprego, Acesso à Previdência Social e Informalidade. A partir da seleção, definimos o ponto de corte de privação para cada indicador. Esses limites determinam se um domicílio é classificado como não privado, abrangendo uma regra de agregação que considera um domicílio privado se pelo menos uma pessoa for privada, ou somente se todos os indivíduos forem privados. A formulação final dos indicadores está na Tabela 1.

Tabela 1: Indicadores de privação de pobreza laboral.

Indicadores de privação	Privado se...
i_1 - Alfabetização	Algum morador com 15 anos ou mais não sabe ler ou escrever um bilhete simples.
i_2 - Escolaridade de adultos	Nenhum adulto concluiu o ensino fundamental (9 anos de estudo, conforme os padrões brasileiros).
i_3 - Emprego	Um indivíduo é privado se estiver desocupado (em busca de trabalho, sem encontrar), desalentado (disponível para trabalhar, mas parou de procurar por falta de esperança) ou subocupado (trabalhando menos horas do que as disponíveis e necessárias). Um domicílio é privado se qualquer membro adulto o for.
i_4 - Acesso à previdência social	Nenhum membro contribui para o sistema de previdência social, está protegido por condições especiais (como militares ou trabalhadores da agricultura de subsistência), e nenhum membro idoso recebe aposentadoria.
i_5 - Informalidade	Nenhum adulto possui emprego formal ou empresa registrada.

Dado que nosso modelo possui 5 indicadores, nosso espaço de estados teórico consiste em $2^5 = 32$ perfis de privação possíveis. Cada estado é representado pelo vetor

¹ Ressalva metodológica: devido às restrições sanitárias, o IBGE realizou a transição das entrevistas presenciais para um sistema de coleta telefônica entre o segundo trimestre de 2020 e meados de 2021. Essa mudança pode ter introduzido um viés específico de não resposta, visto que domicílios nos estratos socioeconômicos mais baixos são mais difíceis de contatar e reter por telefone.

ordenado $X_t = (i_1, i_2, i_3, i_4, i_5)$ onde cada posição corresponde a um indicador específico. Por exemplo, estar no estado **01011** indica que no domicílio: não há pessoas sem alfabetização ($i_1 = 0$); nenhum adulto concluiu o ensino fundamental ($i_2 = 1$), existem pessoas ocupadas ($i_3 = 0$), não há filiação à previdência social ($i_4 = 1$) e todos são trabalhadores informais ($i_5 = 1$).

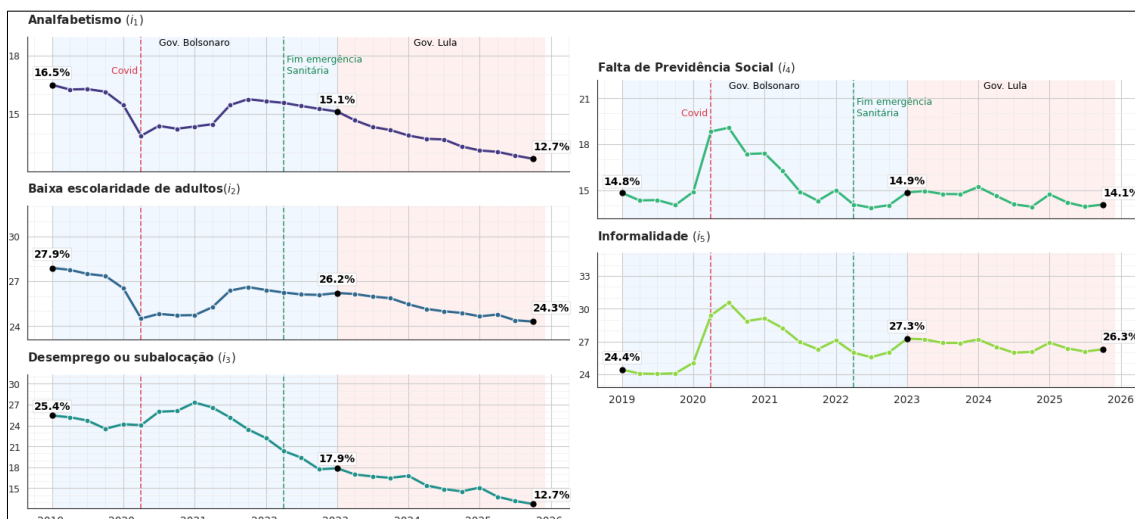
3.3. Cálculo da Matriz de Transição

Com base nos dados históricos, foi calculada a matriz de transição P . Cada linha da matriz estocástica indica a probabilidade de transição do estado (linha) para o estado (coluna). Em nosso contexto, trata-se de uma matriz quadrada 32×32 que representa a probabilidade de um domicílio migrar de um perfil de privação para qualquer outro em um passo de tempo correspondente a um trimestre. Nesta matriz, aplicamos o teste da propriedade de Markov. O teste de razão de verossimilhança, ao comparar o modelo de primeira ordem com uma alternativa de segunda ordem, rejeitou a hipótese de *markovianidade* estrita, confirmando a dependência de trajetória esperada nas dinâmicas sociais. Essa rejeição, por si só, é pouco informativa. Com o alto volume de transições, o teste consegue detectar desvios arbitrariamente pequenos. A questão substantiva é se a dependência de segunda ordem detectada é suficientemente grande para comprometer o uso do modelo. Defendemos a utilidade do modelo pela sua validade preditiva [Shmueli, 2010]. Criamos 20 janelas temporais no conjunto de dados e estimamos a matriz de transição para cada uma. Projetamos o vetor de distribuição para os próximos quatro trimestres e comparamos essas projeções com a distribuição observada nos dados. O erro médio absoluto foi de 0,38% (dp. 0,09%), indicando que as estimativas de probabilidade para os perfis de privação desviam-se apenas ligeiramente dos dados reais. Além disso, o modelo previu com precisão de 94% dos perfis de privação da população com quatro trimestres de antecedência. Isso demonstra que a matriz de transição de primeira ordem é uma aproximação eficaz e, portanto, útil para o planejamento de políticas de médio prazo.

4. Resultados

A análise da série histórica (Figura 1) mostra os efeitos da pandemia e da recuperação subsequente.

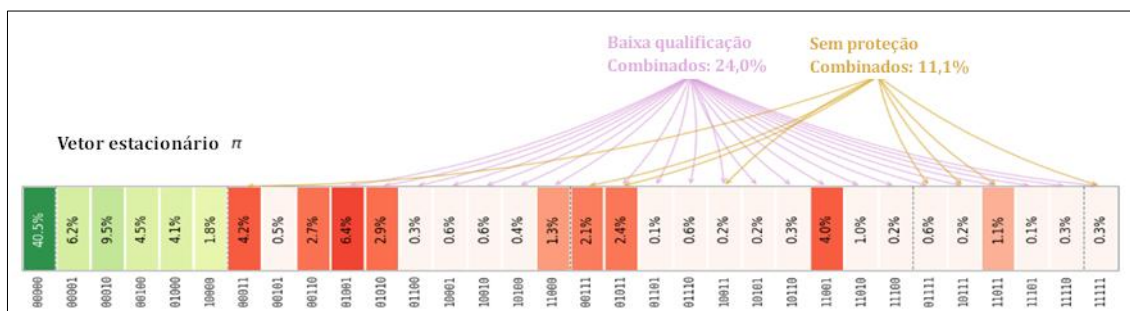
Figura 1. Evolução da incidência de cada indicador (% de famílias com privação, 2019-2025)



A pandemia agravou a vulnerabilidade dos trabalhadores, levando muitos deles à informalidade. O posterior aquecimento da economia reduziu a privação de emprego, mas manteve a informalidade acima do nível pré-crise, em um mercado de trabalho mais flexível e menos protegido. Em 2025, 24,7% dos domicílios não têm adultos com ensino fundamental completo, e 30,1% das famílias não têm previdência social.

Calculamos o vetor de estado estacionário com base na matriz de transição P para a Pobreza Laboral Multidimensional no Brasil (2025). Este vetor representa o equilíbrio de longo prazo do sistema. A distribuição é independente do estado inicial e revela o destino estrutural da população, ditado pelas políticas atuais e pelas condições de mercado. Os resultados, ilustrados na Figura 2, indicam uma probabilidade de 40,5% das famílias atingirem o estado sem nenhuma privação. Entre os estados com apenas uma privação, destacam-se aqueles que representam a precariedade do trabalho, com probabilidade de 9,5% de um domicílio possuir apenas a privação de acesso à previdência social e de 6,2% apenas a informalidade. Podemos agrupar as famílias mais vulneráveis, que possuem múltiplas privações, em dois grupos: (a) baixa qualificação (24,0%), com privações relacionadas ao analfabetismo e à baixa escolaridade, indicando força de trabalho presa em condições precárias por déficit de capital humano; (b) desproteção (11,1%), com famílias privadas de acesso à previdência social e sujeitas à informalidade, totalmente vulneráveis a imprevistos e às oscilações do mercado.

Figura 2. Probabilidades de cada estado no vetor estacionário (% de famílias com privação, 2019-2025)

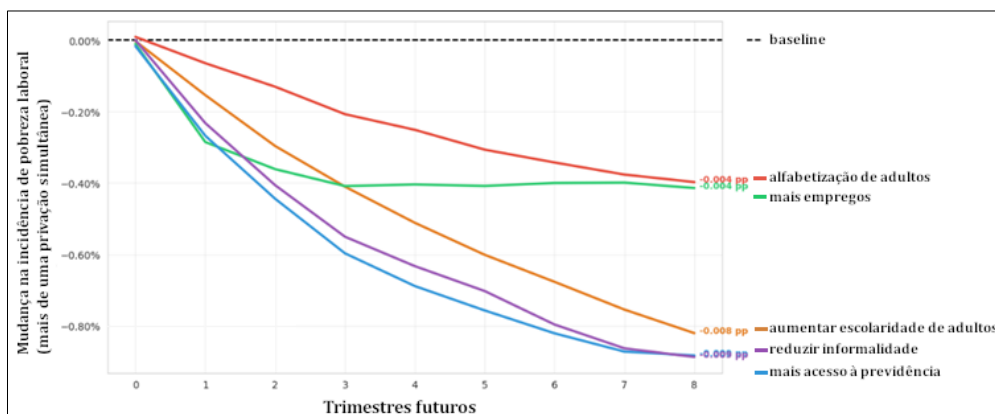


Ver a pobreza laboral como um sistema de transições entre perfis de privação, ao invés de um conjunto de tendências de indicadores paralelos, revela estruturas que as medidas de incidência unidimensionais não mostram: (a) a privação de emprego é a que menos retém famílias; em contraste, a informalidade é a mais persistente: conseguir trabalho é diferente de conseguir trabalho com proteção; (b) leituras isoladas dos indicadores sugeriam que o mercado de trabalho era o principal motor de melhora; no entanto, a análise dinâmica ano a ano mostra que há uma desaceleração destes ganhos e uma aceleração nas probabilidades de resolução nos indicadores educacionais nos últimos dois anos, um contraste que não era visível na série de incidência individual dos indicadores; (c) a análise dinâmica revela um gradiente: as probabilidades de melhorias nos indicadores diminuem à medida que as privações se acumulam. A estrutura conjunta das privações não é um pano de fundo passivo, e sim uma restrição concreta à eficácia das intervenções, atenuando seus resultados exatamente onde a desvantagem acumulada é maior.

Para traduzir as dinâmicas de transição em horizontes temporais concretos, realizamos simulações de Monte Carlo projetando a evolução dos indicadores para os próximos 8 trimestres. Adotamos uma abordagem de microsimulação: o algoritmo

modelou a trajetória individual de 100.000 domicílios virtuais, submetendo cada um ao mecanismo estocástico definido pela matriz de transição P , respeitando a distribuição de estados iniciais observada nos dados reais. Esta simulação foi repetida 1.000 vezes para quantificar a incerteza. As projeções revelam um ponto de saturação: apesar dos avanços recentes, o sistema tende rapidamente a um ponto de estabilidade resistente, em torno de 33,5% da população em pobreza laboral (múltiplas privações simultâneas). Após os ganhos cíclicos de emprego, o desafio atual de erradicar a pobreza laboral no Brasil passa pela adoção de novas políticas para aprimorar o capital humano e a proteção social, que o mercado e o passar do tempo sozinhos não resolverão. Diante do diagnóstico de estagnação, a última fase da modelagem consistiu em simular intervenções na matriz de transição. Para identificar o indicador com maior elasticidade na redução da pobreza, aplicamos perturbações positivas, gerando cinco matrizes de transição, uma para cada indicador, com um aumento de 10% na probabilidade de superação das privações e executamos simulações de Monte Carlo, comparando as curvas resultantes (Figura 3).

Figura 3. Comparação da incidência de pobreza laboral entre cenários de intervenção e baseline (sem intervenção).



Melhorias no emprego saturam rapidamente, indicando que mais empregos, por si sós, não superam as barreiras da privação múltipla. O combate ao analfabetismo também apresenta um limite de evolução, com ganhos iniciais que se estabilizam. Políticas de aumento da escolaridade de adultos por outro lado, mostram uma queda estável na pobreza laboral. Intervenções na Previdência Social e na formalização são ainda mais eficazes, rompendo a “armadilha da desproteção” do vetor estacionário. O estudo evidencia que erradicar a pobreza laboral no Brasil exige o desmantelamento das armadilhas que levam as famílias a se concentrar em perfis de privação comuns.

5. Conclusão

Ao aplicar este método ao mercado de trabalho brasileiro (2019-2025), demonstramos que o mapeamento da dinâmica do espaço de estados de privação, em toda a sua granularidade, revela mecanismos invisíveis às variáveis agregadas, fornecendo uma base científica para o desenho de políticas sociais de precisão. Três características reforçam o potencial de adoção do método: (i) parsimônia, ao condensar a dinâmica de um sistema social complexo em uma matriz de transição; (ii) interpretabilidade, uma vez que cada estado corresponde a um perfil concreto de privações, tornando o mecanismo auditável por tomadores de decisão; e (iii) simplicidade de implementação: todo o *pipeline* (estimação da matriz, cálculo do vetor estacionário e simulações) executa em minutos em um *notebook* Python local sobre o painel com mais de um milhão de domicílios.

O estudo apresenta limitações que apontam para caminhos futuros: (a) a premissa de Markov, embora empiricamente robusta, simplifica a dependência da trajetória da pobreza. Estudos futuros podem explorar modelos de ordem superior para capturar efeitos de longo prazo com maior precisão; (b) dados longitudinais que capturem múltiplas privações simultaneamente não são comuns, o que pode ser um obstáculo à aplicação do método em uma avaliação mais ampla da pobreza multidimensional. A construção de painéis sintéticos a partir de dados transversais para uso deste método é uma estratégia a ser investigada, abrindo caminho para aplicação do método em contextos com dados longitudinais escassos.

Referências

- ALKIRE, Sabina; FOSTER, James. Counting and multidimensional poverty measurement. **Journal of Public Economics**, v. 95, n. 7–8, p. 476–487, ago. 2011.
- AMORNBUNCHORNVEJ, Chainarong *et al.* Framework for inferring empirical causal graphs from binary data to support multidimensional poverty analysis. **Heliyon**, v. 9, n. 5, p. e15947, maio 2023.
- GARCÍA ARANCIBIA, Rodrigo; GIRELA, Ignacio. Graphical Representation of Multidimensional Poverty: Insights for Index Construction and Policy Making. **Social Indicators Research**, v. 172, n. 2, p. 595–634, mar. 2024.
- RAMÍREZ, José Carlos; KADELBACH, Vivian; MATA, Leovardo. The dynamics of social deprivation in Mexico. **Latin American Economic Review**, v. 30, n. Latin American Economic Review, Vol. 30 (2021), p. 1–20, 4 jun. 2021.
- SANTOS, Maria Emma; VILLATORO, Pablo. A Multidimensional Poverty Index for Latin America. **Review of Income and Wealth**, v. 64, n. 1, p. 52–82, mar. 2018.
- SEN, Amartya. Issues in the Measurement of Poverty. In: STRØM, Steinar (Org.). **Measurement in Public Choice**. London: Palgrave Macmillan UK, 1981. p. 144–166.
- SEN, Amartya. **Commodities and capabilities**. 13th impr ed. New Delhi: Oxford Univ. Press, 2008.
- SHMUELI, Galit. To Explain or to Predict? **Statistical Science**, v. 25, n. 3, 1 ago. 2010.
- SUPPA, Nicolai. Deprivations Rarely Come Alone. Multidimensional Poverty Dynamics in Europe. **Review of Income and Wealth**, v. 71, n. 4, p. e70031, nov. 2025.
- UNDP; OPHI. **Global Multidimensional Poverty Index 2024: Poverty amid conflict**. United Nations Development Programme (UNDP), and Oxford Poverty and Human Development Initiative (OPHI), University of Oxford, 2024. Disponível em: <<https://ophi.org.uk/Publications/GMPI15-2024>>. Acesso em: 2 fev. 2026
- WANG, Bingbing *et al.* Differences and dynamics of multidimensional poverty in rural China from multiple perspectives analysis. **Journal of Geographical Sciences**, v. 32, n. 7, p. 1383–1404, jul. 2022.