

Vulnerabilidade Socioeconômica de Indígenas e Quilombolas no Brasil: uma proposta de análise do Cadastro Único

Claudelino Martins Dias Junior¹
Rodrigo Speckhahn Soares da Silva²
Rodrigo Paranhos Faleiro³

Introdução

Pensar as categorias de grupos sociais que são objeto de análise de vulnerabilidade impõe um recorte pragmático e um exercício reflexivo. Um recorte pragmático por serem justamente aqueles cidadãos que, de acordo com as instituições públicas responsáveis, estão cadastrados em bases de dados como beneficiários das políticas públicas (PP) para minorias e vulneráveis. Porém, descrever um conjunto de critérios que permitam incluir essas categorias de cidadãos de um país continental, diverso cultural e racialmente, é antes de qualquer coisa uma pretensão ambiciosa e ousada.

Pensar em métodos classificatórios de seres humanos por recortes sociais costumava ser fácil quando a geopolítica mundial os mantinha separados. Mas, hoje, em vários países do mundo e em especial no Brasil, tais possibilidades classificatórias seriam aplicáveis em um pouco mais de uma centena de grupos que vivem em isolamento voluntário. Em geral, povos nativos que vivem em áreas ainda independentes do modelo de desenvolvimento ocidental e alguns grupos sociais que buscam manter alguma singularidade genética ou cultural – ainda que possamos considerar tal objetivo como sendo, no mínimo, pretensioso e, quiçá, pouco efetivo.

Até mesmo os esforços em caracterizar raças e seus atributos com o objetivo de classificá-las em uma hierarquia foram abandonados após a Segunda Guerra Mundial. Desde então, tais atributos de fenótipos passaram a ser utilizados apenas em levantamentos estatísticos que, hoje, no Brasil, passaram a ser meramente auto declaratórios. Isso se deve a própria dificuldade em classificar indivíduos em um ou outro grupo racial, posto que além dos

¹ Professor do Departamento de Gestão, Mídias e Tecnologia, do Programa de Pós-Graduação em Administração e do PósDesign (Programa de Pós Graduação em Design) da Universidade Federal de Santa Catarina (UFSC) – Brasil – claudelino@gmail.com ou claudelino.junior@ufsc.br

² Doutorando do Programa de Pós-Graduação em Administração da Universidade Federal de Santa Catarina (UFSC) – Brasil – speckhahn@gmail.com

³ Pesquisador do Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) – Brasil – rodrigo.paranhos@gmail.com

padrões originais terem sido relativizados, destaca-se que a identidade racial deixou de ser uma característica imposta pela aparência e passou a ser uma identidade que indivíduos detêm com esta ou aquela cultura. Não raro, observam-se famílias constituídas por membros de fenótipos diferentes.

Porém, se há uma dificuldade em estabelecer critérios biológicos para classificar pessoas em grupos sociais com base em sua origem, por que teríamos PP específicas para esses grupos? Historicamente, vários grupos valiam-se dos territórios onde viviam para realizarem sua reprodução social e econômica (CARMO, GUIZARDI, 2018). Registra-se esse entendimento desde o Reinado, quando a Coroa editou normativas que reconheciam os direitos dos índios aos seus territórios tradicionais. Outros buscaram na constituição de territórios uma forma de resistir à escravidão ou a perda de sua identidade enquanto grupo social, o que não é necessariamente uma identidade étnica e sim racial ou de excluídos e escravizados. Em outras palavras, parte das PP destinadas ao reconhecimento de territórios tradicionais para os índios ainda no Reinado subsidiou as políticas que persiste ainda hoje e foram elas que serviram de base para as atuais políticas de reconhecimento de quilombolas (PARANHOS FALEIRO, 2015).

Apesar da legislação nacional atribuir à Fundação Nacional dos Povos Indígenas (FUNAI) a competência quanto ao reconhecimento de povos indígenas e seus territórios, à Fundação Cultural Palmares de certificar as comunidades quilombolas e ao Instituto Nacional de Colonização e Reforma Agrária (INCRA) a regularização destes territórios de quilombo, a sociedade brasileira vem acatando o autorreconhecimento como instrumento que define os povos e as comunidades tradicionais que possuem organização social própria. E que, igualmente, ocupam territórios e recursos naturais para sua reprodução cultural, social, religiosa, ancestral e econômica, valendo-se do conhecimento tradicionalmente transmitido (Decreto nº 6.040, de 7 de fevereiro de 2007, que “institui a Política Nacional de Desenvolvimento Sustentável dos Povos e Comunidades Tradicionais”).

Se, por um lado, o reconhecimento dos territórios dessas populações subsidiou as PP que hoje são dedicadas às minorias sociais – índios e quilombolas em termos amplos – e étnicas – índios isolados, de recente contato ou que mantêm sua reprodução social independente de outros grupos sociais –, por outro, notabiliza-se que muitos desses indivíduos são enquadrados como vulneráveis – e, portanto, objeto do conjunto de PP

desenvolvidas com o avanço das políticas de direitos humanos – e também, são considerados excluídos – inclusive como beneficiários das políticas de combate à exclusão, miséria, fome e etc.

Pode-se, portanto, pensar que PP dedicadas às minorias sociais – no caso, indígenas e quilombolas – são políticas orientadas para fortalecer e manter a identidade coletiva do grupo. Enquanto, as políticas dedicadas ao combate da exclusão e vulnerabilidade, tendem a priorizar a manutenção da vida do indivíduo e de sua família – ainda que esta seja beneficiada através de um membro da família. Embora o sombreamento das duas políticas gere “melhores condições de vida” para algumas famílias, esse impacto não é apenas positivo considerando-se o estabelecimento de desigualdades internas no grupo pela inserção do comércio em substituição aos tradicionais mecanismos de troca da comunidade. Pensando nessas questões, em um projeto de pesquisa mais amplo no qual os autores tomam parte, em alguns casos analisados ficou explícito que a inserção de benefícios dessas políticas impôs à comunidade desafios para evitar sua desagregação. Ao mesmo tempo, em outros casos, registrou-se que, algumas comunidades ao optarem por gerir esses benefícios de forma coletiva, lograram êxitos na destinação dos recursos para fortalecimento do grupo.

O presente estudo propõe definir faixas de vulnerabilidade socioeconômica para famílias indígenas e quilombolas, retratando, sob determinada medida, a efetividade da assistência social dada com o Programa Bolsa Família (PBF). Para tanto, valendo-se das amostras desidentificadas da base de dados “Famílias” presente no CadÚnico relativos ao ano de 2018, disponibilizada pelo antigo Ministério da Cidadania, atual Ministério do Desenvolvimento e Assistência Social, Família e Combate à Fome, e de uma abordagem analítica quantitativa. A escolha do CadÚnico deve-se, além do quantitativo de famílias cadastradas em todo o país (ver Figura 1), do seu uso na implementação de importantes programas sociais, desde sua criação em 2003, destinados a minimização da condição de vulnerabilidade socioeconômica de grupos minoritários.

Propõe-se um modelo avaliativo alicerçado nos pressupostos da Teoria de Resposta ao Item (TRI) para a criação de uma escala que subsidia as análises com vistas a definição de extratos ou faixas de vulnerabilidade socioeconômica das populações indígenas e quilombolas do CadÚnico. Ressalta-se que essa aplicação de métodos quantitativos nas políticas de vulnerabilidade para grupos minoritários é um investimento fundamental sendo

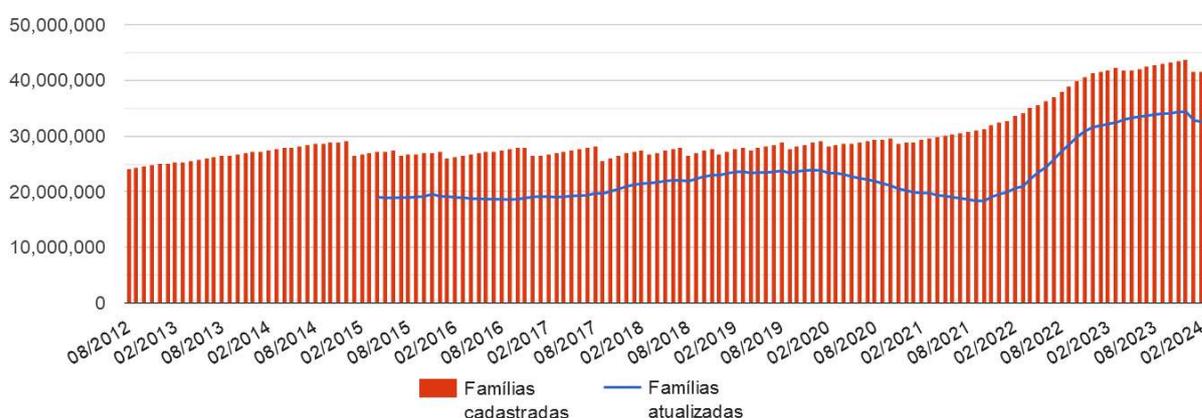
recorte relativos às populações indígenas e quilombolas essencial para o aprimoramento das PP e o fortalecimento da gestão dos territórios por esses grupos minoritários.

População de interesse

A população de interesse do presente estudo é constituída por famílias brasileiras registradas no CadÚnico. Esse cadastro é um registro que permite ao Governo Federal saber quem são e como vivem as famílias brasileiras de baixa renda. É um instrumento que busca demonstrar as características de famílias e indivíduos em estado de vulnerabilidade socioeconômica e sua ideia é unificar, numa mesma base de dados, os atuais e os potenciais beneficiários de programas sociais. Assim sendo, reúne dados que permitem caracterizar grupos sociais menos favorecidos (DIREITO et al., 2016; SEDES, 2020).

O CadÚnico é elaborado e atualizado anualmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Fazem parte do CadÚnico dois grandes conjuntos de dados – de Famílias e de Indivíduos – ambos atualizados anualmente, em função do número de indivíduos e famílias incluídas e excluídas. A base de famílias, no ano de 2012, possuía mais de 20 milhões de famílias cadastradas à época e ultrapassou a marca de 40 (quarenta) milhões de famílias ao final de 2022, conforme apresentado na Figura 1.

Figura 1 – Número de Famílias Cadastradas e Atualizadas no CadÚnico.



Fonte: Adaptada de Brasil (2024).

O CadÚnico constitui a porta de entrada para inúmeros outros programas sociais como exemplos, têm-se: o Programa Auxílio Brasil; o Programa Tarifa Social de Energia

Elétrica; Isenção de Taxas em Concursos Públicos; o ID Jovem; a Carteira do Idoso; o Programa Casa Verde e Amarela e o Programa Bolsa Verde. O CadÚnico é ainda utilizado pelos estados e municípios brasileiros como base para seus projetos e programas sociais subsidiários (DIREITO et al., 2016).

A base de dados do CadÚnico constitui uma das mais importantes referências na distribuição de benefícios sociais do governo federal e, em especial, para o PBF retomado em março de 2023. Particularmente, o grupo de famílias de baixa renda dentro do CadÚnico engloba aquelas que ganham até meio salário-mínimo por pessoa, ou ainda as que percebem até três salários-mínimos de renda mensal total. Dentre essas, as famílias beneficiárias do PBF distribuem-se por renda per capita entre R\$ 0 (zero) e R\$ 477,00 (quatrocentos e setenta e sete reais), valores referentes ao salário-mínimo no ano de 2018 (R\$ 954,00). Os dados disponibilizados pelo CadÚnico possibilitam, por meio de uma estratificação ou escolha de perfis, identificar e caracterizar conjuntos de famílias, indivíduos ou ainda domicílios com características semelhantes (MDS, 2020).

De acordo com Barros et al. (2010), o CadÚnico tem abrangência quase censitária de seu público-alvo, seja pelo quantitativa de registros quanto pela diversidade de dados que armazena, tornando-se uma ferramenta importante para análise e diagnóstico das condições de vida de seus integrantes, bem como para o planejamento e a conformação de programas sociais voltados a essa população.

Os programas sociais contemplam indivíduos ou grupos que se encaixam dentro do perfil socioeconômico alvo de determinada política governamental, atendendo usualmente os mais pobres ou ainda os socialmente mais vulneráveis, fornecendo meios de melhorar suas condições de vida, alimentação, saúde, educação, habitação, inclusão social e trabalho, tomando-se como referência a vulnerabilidade de famílias com renda média *per capita*, beneficiárias ou não de demais programas sociais.

Amostragem e medida da vulnerabilidade socioeconômica

Para o estabelecimento de uma medida, o presente estudo propõe a criação de uma escala e conseqüentemente a estimativa de uma “nota” ou “escore” associada à vulnerabilidade socioeconômica das famílias. Essa classificação é obtida de forma semelhante à classificação de estudantes do Ensino Médio no Brasil quando da realização

das provas do ENEM. Onde esses são ranqueados através do seu desempenho ao responder os itens avaliativos que visam avaliar diferentes conjuntos de conhecimentos, permitindo, da mesma forma, classificar e comparar grupos de famílias do CadÚnico. (SARTES; SOUZA-FORMIGONI, 2013, RODRIGUES; MATOS; FERREIRA, 2017). Nesse sentido a metodologia empregada no estudo é a Teoria de Resposta ao Item (TRI), que adiante, discorre-se sobre seus pressupostos e caracterização.

Teoria de Resposta ao Item

A TRI é um tipo de avaliação baseado em métodos estatísticos e modelos matemáticos, onde obtém-se uma estimativa do traço latente e um erro associado a partir das respostas dadas por indivíduos a um grupo de itens avaliativos. Ou seja, é uma modelagem estatística desenvolvida para mensurar um atributo que não pode ser medido diretamente, denominado traço latente. Assim, com fundamento na observação de outras variáveis relacionadas a essa medida (variáveis secundárias) pode-se obter uma estimativa de uma grandeza de interesse (ALEXANDRE et al., 2002, REISE; AINSWORTH e HAVILAND, 2005, VIEIRA et al., 2011 e SILVA, 2018).

A TRI e seus modelos são tipicamente utilizados em avaliações educacionais, sendo que no Brasil, suas principais aplicações são realizadas no Sistema de Avaliação da Educação Básica (SAEB) e no Exame Nacional do Ensino Médio (ENEM). No contexto da Ciência da Administração, as aplicações concentram-se na criação de escalas de avaliação para apoio à tomada de decisão, *benchmarking*, mensuração de atitude, análise de perfis de decisão, avaliação da satisfação de usuários, avaliação de infraestrutura e análise de fatores críticos (SOARES, 2005; SPENESSATO, 2017; TEZZA et al., 2016 e MEC, 2020).

A TRI prima por representar a probabilidade de ser obtida uma resposta esperada em função da habilidade de um indivíduo a um item avaliativo. As diferentes representações da relação entre a probabilidade desse indivíduo responder a um item avaliativo do modo esperado como função de um traço latente e das características dos itens ocorre porque, quanto maior a habilidade do respondente, maior será a probabilidade de responder do modo esperado.

Tezza e Bornia (2009), Embretson e Reise (2000) e Silva (2018), sugerem as seguintes vantagens da TRI frente à Teoria Clássica dos Testes⁴ (TCT) em processos de avaliação de desempenho:

- a) posicionar indivíduos ou grupos diferentes em uma escala comum;
- b) avaliar com maior precisão as propriedades dos itens do instrumento de avaliação;
- c) desenvolver indicadores mais eficientes na avaliação de diferenças individuais de processos, práticas, sistemas ou indivíduos;
- d) maior robustez aos resultados pretendidos;
- e) comparar o desempenho entre populações distintas ou não e que tenham sido submetidos a avaliações parcialmente diferentes (exige-se uma porcentagem mínima de 20% em termos de itens comuns);
- f) é centrada nos itens e não no teste/questionário como um todo;
- g) fornece informações mais precisas do desempenho dos respondentes;
- h) permite comparar os indivíduos que responderam questionários compostos por itens parcialmente distintos, a fim de mensurar o mesmo traço latente, sendo estes colocados em uma mesma escala;
- i) determinar o erro padrão associado à medida de cada respondente, tornando a estimação do erro mais precisa;
- j) permite utilizar itens com formatos diferentes (dicotômicos e politômicos) não causando desequilíbrio no escore total do teste.

Por outro lado, observam-se algumas dificuldades, sendo: as escalas criadas pela TRI podem ocasionar erros de percepção e de significado percebido; além da complexidade e da necessidade de grandes amostras (muito maiores que nas avaliações tradicionais) com respondentes nos diferentes níveis do traço latente. Além disso, a TRI também exige o uso de *softwares* específicos para realização de estimativas dos parâmetros (TEZZA, 2016; TEZZA; BORNIA, 2009; VEY, 2011).

A TRI apresenta uma série de modelos estatísticos de variáveis latentes que devem ser selecionados conforme: o tipo de função matemática a ser adotada; o tipo de item(s) em

⁴ A Teoria Clássica dos Testes (TCT) compreende um conjunto de conceitos e técnicas para construir instrumentos de medida e tem como referência abordagens de medidas psicométricas (SARTES; SOUZA-FORMIGONI, 2013).

análise; os parâmetros dos itens a serem investigados; e o tipo de processo de resposta. Sua característica principal reside na relação da probabilidade da resposta esperada de um indivíduo a itens em relação às características desses itens (ANDRADE; TAVARES; VALLE, 2000).

De outra parte, os modelos podem ser classificados como cumulativos ou não acumulativos. Nos cumulativos o crescimento da probabilidade é condicionado ao crescimento do traço latente e das características dos parâmetros do item escolhido. E nos modelos não cumulativos, a probabilidade de obter a resposta esperada pode ser dada, por exemplo, em função da distância entre os parâmetros e o item na escala, não dependendo diretamente dos parâmetros do item e do traço latente (ARAUJO; ANDRADE; BORTOLOTTI, 2009).

Ao analisar as bases de dados do CadÚnico, observa-se a existência de itens dicotômicos e politômicos. Assim, adota-se como referência os modelos acumulativos: Modelo Logístico de 2 (dois) parâmetros (ML2P) e o Modelo de Escala Gradual de Samejima (SAMEJIMA, 1969; VAN DER LINDEN; HAMBLETON, 1997).

O ML2P baseia-se no fato de que os respondentes mais favoráveis a certa afirmação (traço latente de interesse) terem maior probabilidade de responder “sim” entre as alternativas “sim” e “não”. Sendo definido na Equação 1.

$$P_{ij} = P(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_i)}} \quad (1)$$

Os termos da Equação 1, suas definições e valores adequados são apresentados na Tabela 1.

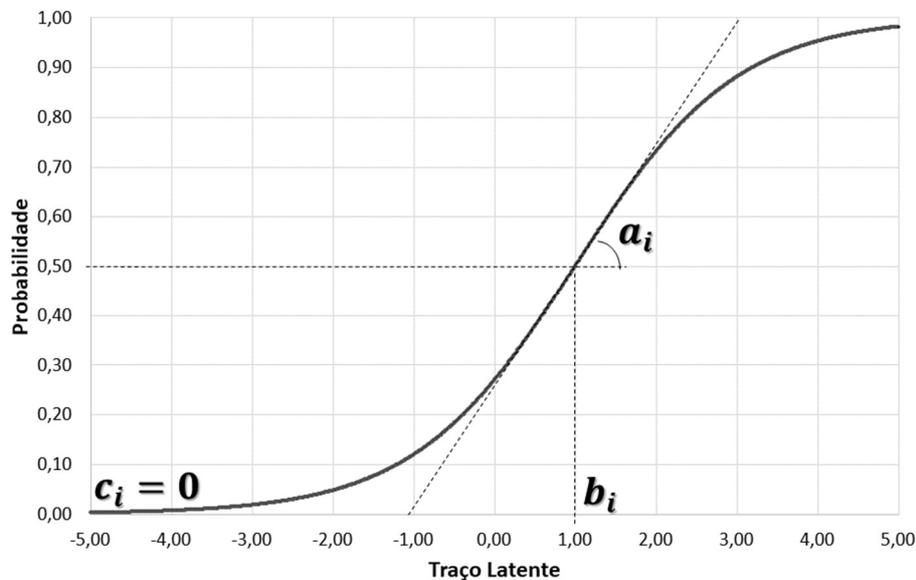
Tabela 1 – Termos, definições e valores adequados para o ML2P.

Termo da Equação	Definição	Valores Adequados	Referências
θ_j	Representa o traço latente (ou habilidade) do j -ésimo ($j = 1, 2, 3, \dots, n$) indivíduo	Teoricamente de $-\infty$ a $+\infty$	
U_{ij}	Variável dicotômica que assume o valor 1 (um) quando o avaliado j responde “sim” ao item i , ou 0 (zero) quando responde “não”.	0 ou 1	(ANDRADE; TAVARES; VALLE, 2000)
j	Representa o traço latente do j -ésimo avaliado.	1 a j	
$P(U_{ij} = 1 \theta_j)$	Função de Resposta do Item: probabilidade de um respondente j com traço latente θ responder “sim” ao item i .	$0 \leq P_i \leq 1$	
D	É um fator de escala constante.	$D = 1$	
a_i	Medida da capacidade do item i de distinguir dois respondentes distintos.	$0,7 \leq a_i \leq 4,0$	(BAKER, 2002)
b_i	Representa o traço latente (ou habilidade) do j -ésimo ($j = 1, 2, 3, \dots, n$) indivíduo	$-4,0 \leq b_i \leq 4,0$	

Fonte: Elaborada pelos autores (2024).

Usualmente a escala criada pela TRI para avaliar o traço latente θ é centrada na média (igual a 0) com desvio padrão igual a 1, seguindo uma distribuição supostamente normal. Na Figura 3, a curva característica do ML2P que representa a probabilidade de se responder afirmativamente a um item em função do traço latente de interesse (θ).

Figura 3 – Curva característica de um item no ML2P.



Fonte: Adaptada de Andrade, Tavares e Valle (2000).

Já o Modelo de Resposta Gradual (MRG), proposto por Samejima (1969), é útil na análise de itens politômicos com categorias ordenadas (por exemplo, em escalas Likert). Nesse modelo também são estimados os parâmetros a_i e b_{ik} caracterizadores dos itens.

O parâmetro b_{ik} é denominado “parâmetro de limiar”, ou seja, é o parâmetro de dificuldade da k -ésima categoria do item i . Esse parâmetro refere-se ao nível do traço latente (θ) no qual a probabilidade de um indivíduo selecionar uma categoria de resposta ou uma categoria mais alta ordenada (parcial, ou total) é 0,5. Representa-se, assim, o ponto na escala do traço latente com os limiares entre as categorias de resposta possíveis.

Já o parâmetro a_i , também é denominado parâmetro de discriminação entre os respondentes de diferentes níveis do traço latente. Assim, quanto maior o valor deste parâmetro, melhor é o item em termos da discriminação entre indivíduos com diferentes níveis de traço latente.

No MRG a discriminação de uma categoria específica de resposta depende dos parâmetros (a_i), que é comum a todas as categorias do item, e da distância entre as categorias de dificuldade (b_{ik}) adjacentes. É necessário que haja um ordenamento entre as categorias de um mesmo item, de acordo com a classificação de seus escores, ou seja, $b_{i1} \leq b_{i2} \leq \dots \leq b_{im_i}$.

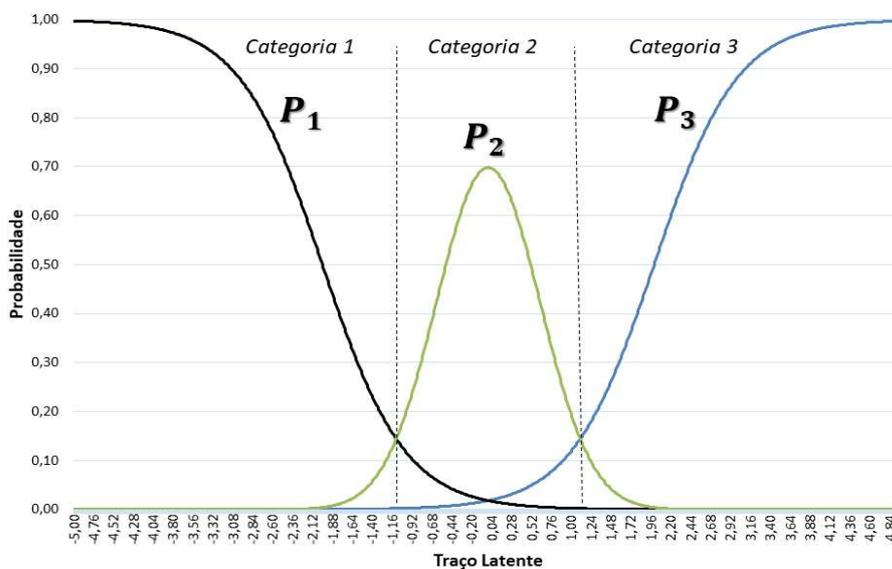
Assim, a probabilidade de um respondente j receber um escore k no item i é dado pela Equação 2.

$$P_{i,k}(\theta_j) = P_{i,k}^+(\theta_j) - P_{i,k+1}^+(\theta_j) = \frac{1}{1+e^{-D\alpha_i(\theta_j-b_{i,k})}} - \frac{1}{1+e^{-D\alpha_i(\theta_j-b_{i,k+1})}} \quad (2)$$

Onde: $P_{i,0}^+(\theta_j) = 1$ e $P_{i,m_i+1}^+(\theta_j) = 0$.

E para um item com $(m_i + 1)$ categorias, m_i valores de dificuldades (b_{ik}) precisam ser estimados, além do parâmetro de discriminação (α_i) do item. Em outras palavras, o número de parâmetros a serem estimados é o número de categorias de respostas de um item. A Figura 4 representa um item politômico com 3 (três) categorias de respostas.

Figura 4 – Representação gráfica de um item politômico com três categorias de respostas observáveis segundo o MRG.



Fonte: Elaborada pelos Autores (2024).

Andrade, Tavares e Valle (2000) destacam a centralidade dos itens nos instrumentos avaliativos sob a ótica da TRI, e não em todo o teste como no caso da TCT, permitindo a comparação de características entre populações distintas ou ainda a comparação de

indivíduos de uma mesma população que tenham sido avaliados com itens avaliativos parcialmente distintos (EMBRETSON; REISE, 2000).

Tavares-Zukowsky (2013) enfatiza que, como metodologia para a construção e análise de avaliações de grande porte, a TRI permite a construção de escalas que agregam matizes capazes de conter, na mesma métrica, milhares de itens e a estimativa de escores de milhões de respondentes em momentos distintos. Ainda que tenham respondido a testes parcialmente distintos, é possível associar e comparar os dados obtidos pois estão na mesma escala. Destaca ainda que a confiabilidade da escala depende diretamente da qualidade dos itens. Qualidade relacionada a itens com altos níveis de discriminação posicionados em diferentes níveis de dificuldade.

Com a TRI ainda é possível a obtenção de estimativas para cada um item avaliativo (função de informação do item) e de todo o teste (função de informação do teste) com maior precisão. Pois, as análises e as interpretações estão sempre associadas aos itens e não ao grupo de indivíduos que os responderam (KLEIN, 2003).

O Modelo em Desenvolvimento

a) Descrição, seleção e adaptação dos itens selecionados

A base FAMÍLIAS do CadÚnico é composta por 32 (trinta e dois) itens avaliativos. Esses itens são categorizados em 4 (quatro) grupos, a saber:

- O **primeiro grupo** de itens é constituído por **códigos de identificação**: Código IBGE do Município; nome do estabelecimento EAS/MS; código do estabelecimento EAS/MS; nome do CRAS/CREAS; código do CRAS/CREAS; e identificador único da família para pareamento com a base de pessoas;

- O **segundo grupo** compõe-se das variáveis não utilizadas na caracterização dos perfis socioeconômicos, mas possíveis de serem utilizadas em agrupamentos específicos: data do cadastramento da família; data da última alteração em qualquer campo da família (variável utilizada nos anos de 2014, 2015, 2016 e 2017); data da última atualização da família, dados esse considerado sensível à manutenção do cadastro no formato YYYY-MM-DD (Variável utilizada somente nos anos de 2014, 2015, 2016 e 2017); data da última alteração em qualquer campo da família (variável utilizada somente nos anos de 2012 e 2013); quantidade de cômodos do domicílio; quantidade de cômodos servindo como

dormitório do domicílio; peso calculado da família; grandes grupos de municípios, de acordo com a quantidade de famílias cadastradas; subdivisão pela Unidade Federativa (UF) e divisão administrativa; quantidade de pessoas utilizada no cálculo da renda *per capita* familiar;

- O **terceiro grupo** é constituído das variáveis que potencialmente possam caracterizar o perfil socioeconômico da base de dados FAMÍLIAS. Dentre elas a variável Renda e Valor da renda média familiar (*per capita*), que é utilizada, posteriormente, como parâmetro de comparação para a medida criada com a TRI. Essas variáveis são representadas por: características do local onde está situado o domicílio e grupos tradicionais e específicos que não qualificam gradualmente os perfis pretendidos;

- O **quarto grupo** é constituído de 11 (onze) variáveis efetivamente aproveitadas na caracterização dos perfis socioeconômicos por serem itens diretamente relacionados à caracterização do perfil socioeconômico, e são compatíveis com os pressupostos do modelo logístico de 2 parâmetros (ML2P) e/ou modelo de escala gradual da TRI. Os itens são: espécie do domicílio; material predominante no piso do domicílio; material predominante nas paredes externas do domicílio) – se o domicílio tem água encanada; forma de abastecimento de água; existência de banheiro; forma de escoamento sanitário); forma de coleta do lixo; tipo de iluminação; calçamento; família indígena; família quilombola; marcação se a família é beneficiária do PBF.

Dentre os itens selecionados têm-se itens dicotômicos e politômicos. O traço latente de interesse, denominado de **vulnerabilidade socioeconômica** das famílias, apesar de ser dependente de uma série de elementos distintos, ser um índice multidimensional⁵, foi tido como unidimensional nesse estudo. Ou seja, pressupõe-se a existência de um traço latente dominante que pode ser medido indiretamente.

Outra suposição essencial no processo de estimação dos parâmetros do modelo é a independência local (ou independência condicional). Assume-se que para determinada habilidade as respostas aos diferentes itens do questionário são independentes. Essa

⁵ O conceito de pobreza multidimensional tem origem nos estudos de Amartya Sen (SEN, 2021) quanto à limitação da dimensão unicamente financeira (fluxos de renda) para avaliar as privações as quais uma pessoa está sujeita. Nessa abordagem tem-se em conta outros elementos como saúde, educação, habitação, participação política, igualdade entre os sexos. Apoiar-se nos indicadores utilizados para construir o Índice de Pobreza Multidimensional bem como em algumas dimensões dos Objetivos do Desenvolvimento Sustentável (SANTOS; VILLATORO, 2018).

independência é consequência da correta determinação da dimensionalidade dos dados. Assim, os itens selecionados são apresentados no Quadro 1.

Quadro 1 – Características dos itens selecionados do CadÚnico.

Nº	Código	Descrição	Tipo
8	<i>cod_especie_domic_fam</i>	Espécie do domicílio	Politômico
11	<i>cod_material_piso_fam</i>	Material predominante no piso do domicílio	Politômico
12	<i>cod_material_domic_fam</i>	Material predominante nas paredes externas do domicílio	Politômico
13	<i>cod_agua_canalizada_fam</i>	Se o domicílio tem água encanada	Dicotômico
14	<i>cod_abaste_agua_domic_fam</i>	Forma de abastecimento de água	Politômico
15	<i>cod_banheiro_domic_fam</i>	Existência de banheiro	Dicotômico
16	<i>cod_escoa_sanitario_domic_fam</i>	Forma de escoamento sanitário	Politômico
17	<i>cod_destino_lixo_domic_fam</i>	Forma de coleta do lixo	Politômico
18	<i>cod_iluminacao_domic_fam</i>	Tipo de iluminação	Politômico
19	<i>cod_calçamento_domic_fam</i>	Calçamento	Politômico
32	<i>marc_pbf</i>	Marcação se a família é beneficiária do Programa Bolsa Família	Dicotômico

Fonte: Elaborado pelos Autores (2024).

Algumas das opções apresentadas aos respondentes na base FAMÍLIAS são reagrupadas, considerando o que Andreoli e Souza-Formigoni (2013) sugerem no uso concomitante de itens categóricos com 2 (dois) ou mais níveis. Nesse caso, tomando-se como exemplo a variável 11 (material predominante no piso do domicílio), tem-se 7 (sete) categorias distintas, apresentadas nessa ordem: terra; cimento; madeira aproveitada; madeira aparelhada; cerâmica, lajota ou pedra; carpete; outro material. Nesse exemplo seguiu-se o ordenamento do material mais simples até o mais sofisticado, conforme demonstra-se na Figura 5.

Figura 5 – Ordenamento dos materiais predominantes no piso dos domicílios.



Fonte: Elaborada pelos Autores (2024).

b) Calibração dos itens

A calibração dos parâmetros dos itens selecionados foi feita através de uma amostragem estratificada proporcional à população de cada unidade da federação (UF) disponível na base de dados para garantir a representatividade de cada UF na calibração dos parâmetros dos 11 (onze) itens selecionados (JELIHOVSCHI, 2014). Para tanto foi utilizado o pacote “*sampler*”, disponível no ambiente computacional R (BALDASSARO, 2019). Na Tabela 2 estão os números de famílias de cada unidade da federação selecionadas de forma aleatória e proporcional em cada UF.

Tabela 2 – N^o de famílias selecionadas aleatoriamente para calibração dos parâmetros dos itens em cada ano.

Código	UF	2012	2013	2014	2015	2016	2017	2018
11	RO	4.371	4.548	4.596	4.416	4.510	4.201	4.140
12	AC	2.090	2.198	2.312	2.295	2.413	2.325	2.295
13	AM	9.059	10.453	10.664	10.757	11.533	11.243	11.625
14	RR	1.000	1.411	1.557	1.748	1.842	1.697	1.829
15	PA	23.640	24.784	25.156	25.924	27.720	27.646	28.854
16	AP	1.573	1.523	1.770	1.769	1.955	2.030	2.319
17	TO	4.793	4.371	4.961	5.163	5.365	5.188	5.181
21	MA	27.221	27.621	27.117	26.826	28.626	27.440	27.100
22	PI	13.786	12.987	13.112	13.006	13.227	12.429	12.288
23	CE	35.567	34.591	33.217	33.470	34.872	33.673	33.627

Código	UF	2012	2013	2014	2015	2016	2017	2018
24	RN	12.357	11.828	11.938	11.584	11.757	11.146	11.180
25	PB	15.668	14.935	15.111	14.824	15.333	14.742	14.541
26	PE	37.260	36.668	35.483	34.313	35.550	34.781	35.980
27	AL	12.892	12.606	12.305	12.235	12.589	12.071	11.996
28	SE	8.097	8.154	8.100	8.206	8.666	8.353	8.556
29	BA	58.448	57.768	56.547	56.014	58.019	56.414	55.521
31	MG	50.110	48.629	48.612	47.041	47.078	45.654	44.339
32	ES	8.383	8.398	8.355	8.327	8.376	7.678	7.656
33	RJ	29.878	30.493	30.794	30.265	30.814	30.293	31.820
35	SP	60.678	63.098	62.248	65.641	53.790	70.393	72.161
41	PR	22.417	21.673	21.640	22.047	22.598	21.693	20.862
42	SC	7.937	7.627	8.365	8.025	7.976	7.092	6.567
43	RS	19.607	19.824	20.561	19.856	19.419	17.730	16.718
50	MS	6.184	6.824	7.237	7.694	7.900	7.271	7.054
51	MT	7.661	7.840	8.537	8.973	9.070	8.622	8.564
52	GO	14.217	14.214	14.934	15.179	15.282	14.858	14.259
53	DF	5.102	4.934	4.772	4.403	3.719	3.335	2.968
Total		499.996	500.000	500.001	500.001	499.999	499.998	500.000

Fonte: Elaborada pelos Autores (2024).

Nunes e Primi (2005) citam que, corroborando resultados de estudos simulados de Embretson e Reise (2000); Hambleton e Swaminathan (1985); e Muñoz Fernández (1990), os cálculos dos parâmetros de dificuldade (b_i) variam bastante para amostras menores do que 0,5% do total. Já os parâmetros de discriminação (a_i), em suma, aumentam com o aumento da amostra: a partir de 0,5 % mantêm-se nas suas posições relativas (parâmetro b). As amostras com no mínimo 1,0% do total fornecem estimativas muito próximas daquelas amostras contendo em torno de 50% da amostra completa

Após realizado o processo de estimativa dos parâmetros dos itens, os resultados obtidos, os parâmetros dos 11 itens selecionados da base de famílias do ano de 2018, são apresentados na Tabela 3.

Tabela 3 – Parâmetros dos itens obtidos para o ano de 2018.

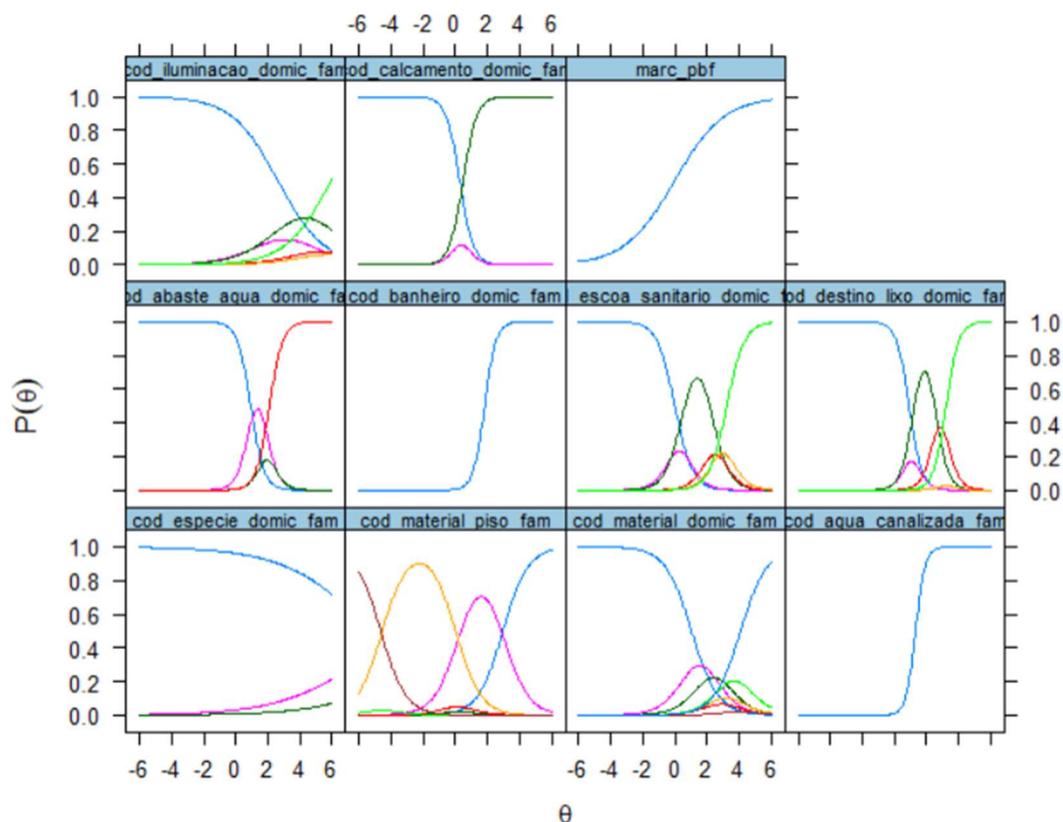
Código do Item	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
<i>cod_especie_domic_fam</i>	0,378	8,447	12,811	-	-	-	-	-
<i>cod_material_piso_fam</i>	1,282	2,973	0,251	0,189	0,036	-4,551	-4,642	-
<i>cod_material_domic_fam</i>	1,211	1,041	2,046	2,799	3,015	3,352	4,029	4,101
<i>cod_agua_canalizada_fam</i>	3,386	1,313	-	-	-	-	-	-

Código do Item	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
<i>cod_abaste_agua_domic_fam</i>	2,421	0,935	1,797	2,093	-	-	-	-
<i>cod_banheiro_domic_fam</i>	2,863	1,829	-	-	-	-	-	-
<i>cod_escoa_sanitario_domic_fam</i>	1,823	-0,003	0,517	2,271	2,748	3,232	-	-
<i>cod_destino_lixo_domic_fam</i>	2,542	0,918	1,190	2,561	3,178	3,220	-	-
<i>cod_iluminacao_domic_fam</i>	0,701	2,623	3,493	5,134	5,585	5,936	-	-
<i>cod_calcamento_domic_fam</i>	2,359	0,277	0,483	-	-	-	-	-
<i>marc_pbf</i>	0,649	-0,030	-	-	-	-	-	-

Fonte: Elaborada pelos Autores (2024).

Na Figura 6, as Curvas Características dos itens (CCI) permitem verificar graficamente o comportamento dos itens.

Figura 6 – Curvas características dos itens selecionados do CadÚnico para o ano de 2018.



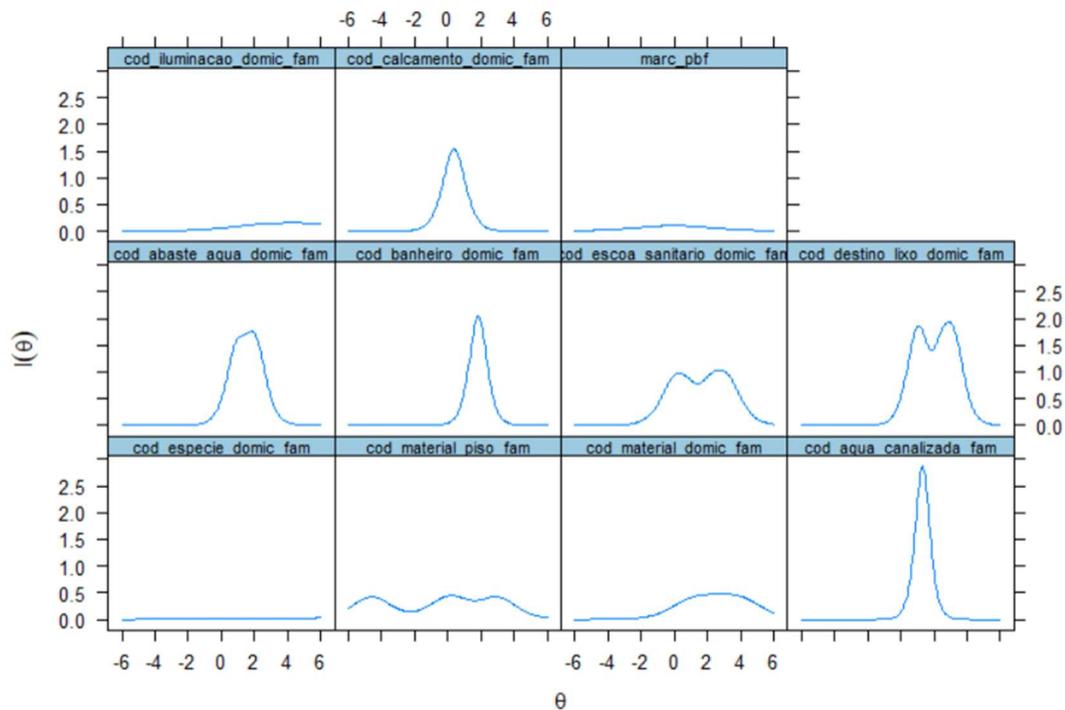
Fonte: Elaborada pelos Autores (2024).

Observando os parâmetros obtidos (ver Tabela 3) e a Figura 6 evidencia-se que algumas categorias de itens politômicos (*cod_material_domic_fam*,

cod_abastec_aqua_domic_fam, *cod_escoa_sanitario_domic_fam*,
cod_destino_lixo_domic_fam) apresentam sobreposição nas curvas características de algumas categorias em função do reduzido número de respondentes que apresentam, mesmo diante de um número significativo de famílias na amostra de calibração. Diferentemente do comportamento obtido no item (*cod_material_piso_fam*) em que se observa uma separação nítida de quatro categorias. Assim, pode-se questionar se tais itens avaliativos e suas categorias contemplam situações enfrentadas por famílias brasileiras de modo a permitir uma avaliação consistente da vulnerabilidade socioeconômica diante da realidade atual da população de famílias brasileiras.

Na Figura 7 são apresentadas as contribuições dos itens para a medida do traço latente de interesse e a região da escala em que se encontram através das curvas de informação de cada item. Nos itens com categorias mais bem definidas observa-se mais de um pico de informação contribuindo com a caracterização do traço latente. Para definir a escala, foi necessário recategorizar àqueles itens que se sobrepuseram de modo a se obter resultados mais adequados para os parâmetros dos itens. Além disso, itens que pouco contribuem com caracterização do traço latente, poderiam, eventualmente, ser eliminados sem grandes prejuízos do ponto de vista de análise com a TRI.

Figura 7 – Curvas de informação itens selecionados do CadÚnico para o ano de 2018.

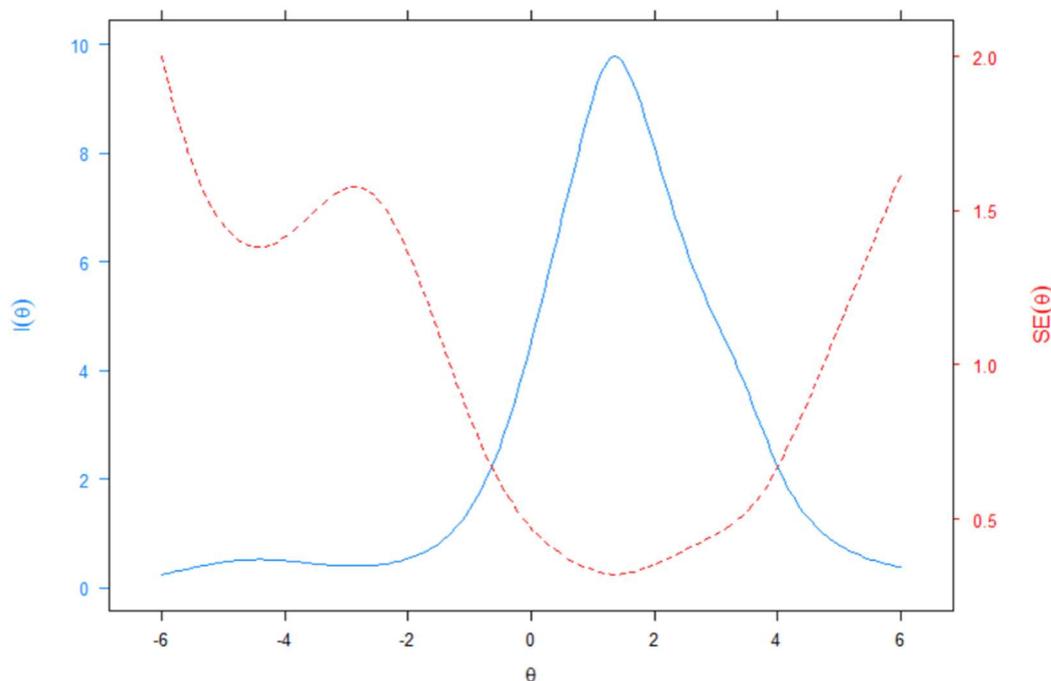


Fonte: Elaborada pelos Autores (2024).

Observa-se que apenas 3 (três) itens apresentaram pequenas contribuições com relação ao traço latente: *cod_especie_domic_fam*; *cod_iluminacao_domic_fam*; e *marc_pbf*. A região em que um item apresenta maior quantidade de informação sobre o traço latente (vulnerabilidade socioeconômica) permite diferenciar famílias com nível (θ) próximos nessa mesma região da escala. Nesse sentido, apresenta-se o exemplo do item *cod_banheiro_domic_fam* com parâmetro $b = 1,829$, ou seja, permite avaliar bem as famílias nessa região da escala criada pela TRI.

Na Figura 8 é apresentada a função de informação total (em azul), ou seja, a soma das contribuições de cada um dos 11 (onze) itens selecionados, bem como o erro associado à estimativa (em vermelho).

Figura 8 – Curva de informação dos itens (CIT) do CadÚnico.



Fonte: Elaborada pelos Autores (2024).

Conforme evidencia em Figura 8, a região da escala em que se obtém maior informação relativa aos níveis de vulnerabilidade socioeconômica das famílias e o erro associado, sendo as das famílias situadas aproximadamente no intervalo $[-1,0; 4,0]$. Para além desse intervalo, abaixo ou acima, o erro associado à medida cresce substancialmente.

c) Cálculo dos escores das famílias

A partir do posicionamento dos itens na escala (parâmetro b) é possível estimar a medida do traço latente para cada uma das famílias da amostra. E a partir dessa estimativa, é possível “ranquear” as mais de 4,8 milhões de famílias, ou seja, posicioná-las na mesma escala. Da mesma forma, tendo acesso a totalidade dos dados, realizar essa estimativa para todas as famílias cadastradas no CadÚnico.

A partir deste “*ranking*” pode-se selecionar grupos de famílias definindo-se faixas de vulnerabilidade a partir de critérios estabelecidos e da interpretação da escala criada. Assim, não são utilizados apenas critérios puramente econômico-financeiros (como a renda familiar) ou estatísticos, mas também qualitativos na seleção de grupos de famílias em situação de maior vulnerabilidade. Na Tabela 3 apresenta-se uma contagem do número de famílias por faixas de intervalo.

Tabela 3– Número de famílias por faixa de escore (θ).

Nº	Faixa de Escore	Nº de Famílias
1] $-\infty$; $-1,5$ [6.426
2	[$-1,5$; $-1,0$ [698.130
3	[$-1,0$; $-0,5$ [776.128
4	[$-0,5$; $0,0$ [1.027.923
5	[$0,0$; $0,5$ [944.273
6	[$0,5$; $1,0$ [683.079
7	[$1,0$; $1,5$ [402.611
8	[$1,5$; $2,0$ [217.618
9	[$2,0$; $2,5$ [48.538
10	[$2,5$; $3,0$ [3.036
11	[$3,0$; $+\infty$ [234

Fonte: Elaborada pelos Autores (2024).

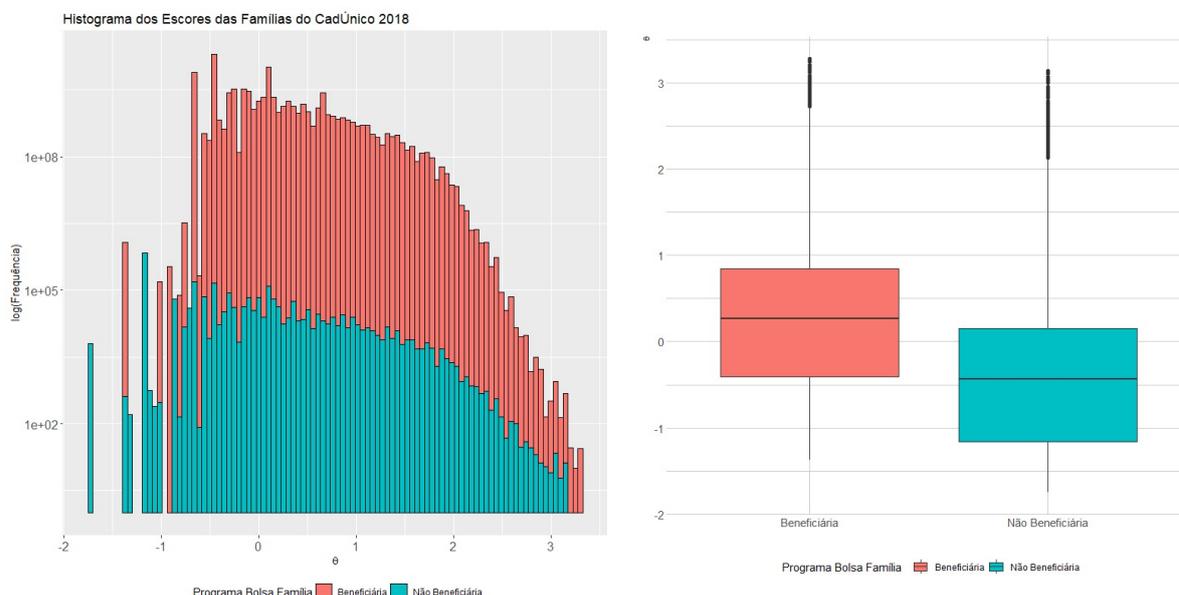
Observando a distribuição dos escores das famílias têm-se 2.508.607 (dois milhões, quinhentos e oito mil e seiscentos e sete) famílias situadas abaixo da média ($\theta = 0$) e 2.299.389 (dois milhões, duzentos e noventa e nove mil e trezentos e oitenta e nove) famílias situadas acima da média. Ao remover da análise o elemento renda média, e tendo em conta apenas as respostas dadas pelas famílias aos demais itens do CadÚnico, novas possibilidades analíticas surgem, com os resultados obtidos. Para tanto, foram elaboradas algumas visualizações para avaliar os escores dos seguintes públicos: beneficiários ou não do PBF; famílias indígenas e famílias não indígenas; famílias quilombolas e famílias não quilombolas. Na Tabela 4 são apresentados os quantitativos para as famílias de cada grupo.

Tabela 4 – Quantitativo de Famílias Beneficiárias e Não do PBF, Famílias Indígenas e Não Indígenas, Famílias Quilombolas e Não Quilombolas.

Família	Sim	Não	NA
Beneficiária do PBF	2.424.434	2.383.562	-
Indígena	25.168	4.782.826	2
Quilombola	29.310	4.753.516	25.170

Fonte: Elaborada pelos Autores (2024).

Na Figura 9, a partir do histograma e *boxplots* é possível visualizar a distribuição dos escores obtidos para os diferentes grupos de famílias descritos e como esses se posicionaram ao longo da escala definida pela TRI.

Figura 9– Histograma e *boxplots* dos escores das Famílias do CadÚnico de 2018.

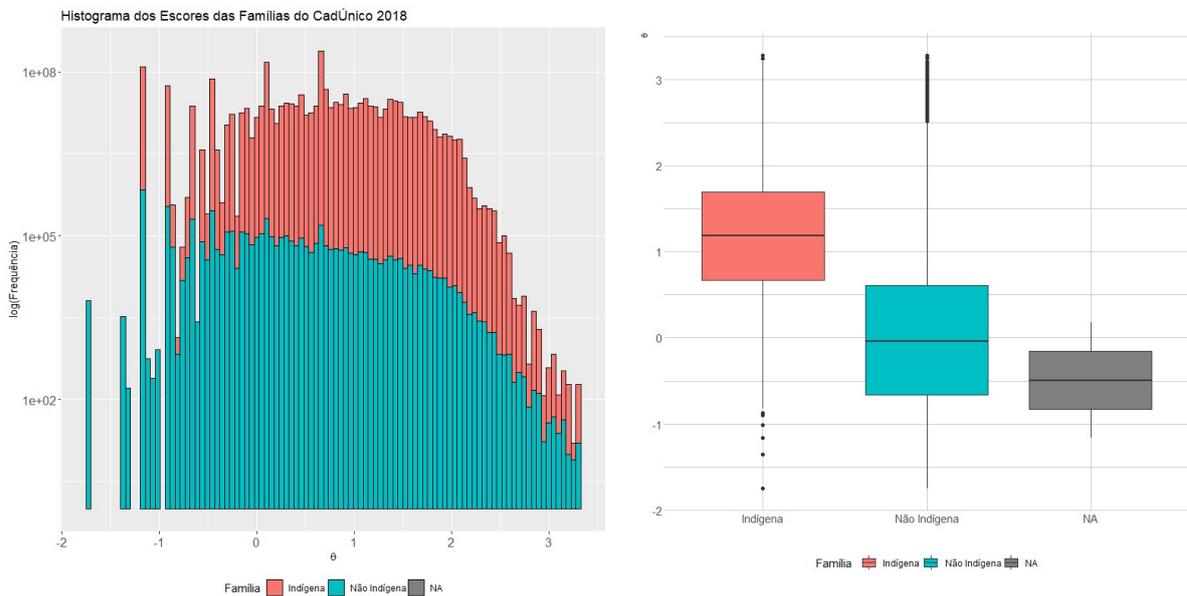
Fonte: Elaborada pelos Autores (2024).

O grupo de Famílias Beneficiárias e Não Beneficiárias do PBF encontram-se relativamente bem-posicionadas na escala, contudo na região inferior da escala à esquerda ($\theta < -0,5$) observa-se a existência de *gaps* entre as barras do histograma. Tal situação ocorre em função da pequena quantidade de itens avaliativos posicionados nessa região da escala (ver Tabela 3) pois os itens, majoritariamente, apresentaram parâmetros b_i com valores acima de zero ($\theta > 0$).

Outro aspecto que pode ser observado nos *boxplots* (à direita, na Figura 9), é a simetria entre os grupos dentro do intervalo interquartílico, bem como o posicionamento das Famílias Não Beneficiárias do PBF que apresentaram escores médios abaixo daquelas famílias beneficiárias do Programa.

Na Figura 10 têm-se os escores agrupados que destacam as Famílias Indígenas e Não Indígenas.

Figura 10 – Histograma e *boxplots* dos escores das Famílias Indígenas e Não Indígenas do CadÚnico de 2018.

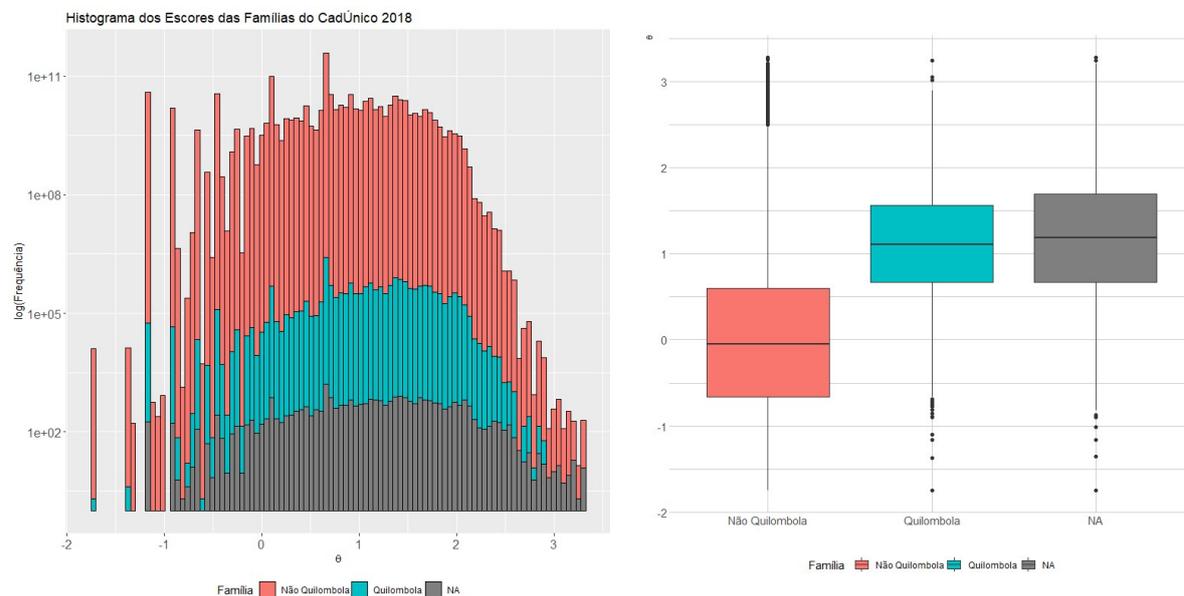


Fonte: Elaborada pelos Autores (2024).

Observa-se na Figura 10 que a distribuição é praticamente simétrica no intervalo interquartil e 75% das Famílias Indígenas obtiveram escores maiores do que as Não Indígenas, ou seja, posicionaram-se na escala acima das demais famílias do CadÚnico. Principalmente, na região da escala em que se obteve maior precisão e menor erro na medida da vulnerabilidade socioeconômica.

Na Figura 11 são avaliados o histograma e o *boxplots* das Famílias Quilombolas (vermelho) e Não Quilombolas (azul).

Figura 11 – Histograma e *boxplots* dos escores das Famílias Quilombolas e Não quilombolas do CadÚnico de 2018.



Fonte: Elaborada pelos Autores (2024).

Na Figura 10, observa-se a distribuição dos escores das Famílias Quilombolas e Não Quilombolas, sendo que as Quilombolas apresentam escores superiores às Famílias Não Indígenas e Não Quilombolas.

Não obstante, é oportuno observar que as famílias mais vulneráveis, beneficiárias ou não do PBF, indígenas ou quilombolas, precisam ser mais bem avaliadas, isto porque encontram-se poucos itens avaliativos na extremidade inferior da escala e, portanto, sugere-se a inclusão de novos itens avaliativos no CadÚnico. Com o aperfeiçoamento desse cadastro é possível obter um melhor detalhamento e caracterização das famílias mais vulneráveis desse conjunto de populações. Da mesma forma a inclusão de mais itens específicos que permitam caracterizar com maior detalhamento as famílias indígenas e quilombolas.

Agradecimentos

As fases preliminares desse estudo no formato de projeto de extensão receberam apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Conclusões e Considerações Finais

Os resultados apresentados mostram o potencial de classificar e distinguir famílias em situação de vulnerabilidade socioeconômica. Porque como os parâmetros a_i apresentaram uma amplitude que varia entre 0,378 e 3,386 (Tabela 4). Ou seja, apenas o item *cod_especie_domic_fam* ($a = 0,378$) apresentou baixo poder de discriminação. Ademais, avaliou-se a CIT (Figura 8), demonstrando que o instrumento se mostra adequado na avaliação de famílias com nível de vulnerabilidade no intervalo $-1,0 < \theta < 4,0$. Contudo, os itens selecionados do CadÚnico apresentam limitações, ao passo que não permitem avaliar de modo mais detalhado as famílias em situação de maior vulnerabilidade socioeconômica ($\theta \leq -0,5$), que deveriam ser o alvo prioritário das PP de distribuição de benefícios como o PBF.

Como resultado obteve-se uma escala de avaliação de níveis de vulnerabilidade socioeconômica que permite retratar diferentes grupos de famílias cadastradas no CadÚnico, em específico, para as famílias indígenas e quilombolas. Essa análise permite explicitar em termos quantitativos as condições desses grupos com vistas ao fornecimento de novos elementos ao poder público quanto as PP. Evidentemente, essa demonstração aponta caminhos a serem percorridos, entre os quais, o impacto dos benefícios na comunidade e sua capacidade de resposta. Também aponta a necessidade de uma aproximação gradual dos contextos dessas comunidades por recortes regionais, urbano-rural, subsistência, dieta ou atividades econômicas, oportunizando o diálogo com análises qualitativas e com bases georreferenciadas.

A importância concreta desta pesquisa está na desmistificação da percepção desses grupos minoritários pela sociedade e na oferta de bases para seu fortalecimento enquanto coletividade coesa e capaz de administrar os meios que detém. A desmistificação pode contribuir com a revisão da ambiguidade histórica (MELATTI, 1987) que, de um lado, insinua serem os índios – por exemplo – um estágio a ser superado em prol do desenvolvimento ou, do outro, os remete ao símbolo romântico e edênico que não pode ser alterado em prol da manutenção da nossa imagem de mundo perdido.

Complementarmente, ao ofertar as bases para o fortalecimento da capacidade dessas comunidades, essa pesquisa permite revisar criticamente os discursos políticos da

última década que insinuam estarem os “índios edênicos desviados de sua natureza pura” e, portanto, serem hoje “apenas usurpadores de recursos públicos e naturais que deveriam servir ao Brasil”. Discurso esse que, conscientemente ou não, ao pregar leituras discriminatórias se contrapõe ao conjunto de PP dedicadas às minorias, e, com isso, demonstra interesse na revisão das bases legais dessas políticas para incorporação dos territórios desses grupos em ativos no mercado de capitais sem, antes, combinar com os povos que aí vivem.

Nesse sentido, os dados produzidos constituem contribuições importantes, ao permitirem classificar e definir níveis de vulnerabilidade socioeconômica a partir de intervalos ou faixas de vulnerabilidade socioeconômica baseada em parâmetros além da renda para essas populações e, não obstante, corroborando com a ação estatal de suporte a essas populações por meio de seus representantes.

Paralelamente, sugere-se que uma avaliação qualitativa desses grupos familiares ao longo dos extremos da escala criada, por meio de uma pesquisa de campo a ser realizada por bolsistas e pesquisadores – serviço social, antropologia, sociologia, economistas etc. – e por profissionais que prestam assistência nas comunidades – saúde e assistência – no sentido de validar a medida proposta, definindo-se preliminarmente algumas cidades como referência ou ainda um ponto de concentração regional. Em tempo, sugere-se ainda a inclusão de novos itens no CadÚnico que permitam avaliar sob um espectro maior a escala de vulnerabilidade socioeconômica criada com a testagem de novos itens de avaliativos.

Embora a medida do traço latente (vulnerabilidade socioeconômica) definida com a TRI se mostre consistente, não deve ser tomada como absoluta se existir variação em torno das médias de referência. No caso de inclusão de novos dados na base FAMÍLIAS ou ainda novos itens avaliativos, as posições entre famílias com escores muito próximos podem sofrer pequenas alterações, considerando que a medida do escore da família (θ) precisa estar associada à medida do erro da estimativa (SE).

Referências Bibliográficas

ALEXANDRE, J. W. C. et al. “Uma proposta de análise de um construto para medição dos fatores críticos da gestão pela qualidade por intermédio da Teoria da Resposta ao Item”. *Gestão & Produção*, v. 9, n. 2, 2002.

ANDREOLI, L. M.; SOUZA-FORMIGONI, M. L. O. “Avanços na psicometria: da teoria clássica dos testes à teoria de resposta ao item” in ANDREOLI, L. M., SOUZA-FORMIGONI, M. L. O. *Psicologia: Reflexões e Crítica*, v. 26, n. 2, 2013.

ANDRADE, D. F.de; TAVARES, H. R.; VALLE, R. da C. *Teoria da Resposta ao Item: conceitos e Aplicações*. São Paulo: ABE – Associação Brasileira de Estatística, 2000. Disponível em: <<http://egov.ufsc.br/portal/sites/default/files/livrotri.pdf>>.

ARAUJO, E. A. C.de; ANDRADE, D. F.de; BORTOLOTTI, S. L. V. “Teoria da Resposta ao Item”. *Revista da Escola de Enfermagem da USP*, v. 43, n. spe, p. 1000–1008, Dez.2009. Disponível em: <www.ee.usp.br/reeusp/>.

BAKER, Frank B. *The basics of item response theory*. 2nd ed. College Park: ERIC Clearinghouse on Assessment and Evaluation, 2002.

BALDASSARO, Michael. **sampler: Sample Design, Drawing & Data Analysis Using Data Frames**. [s.l.: s.n.], 2019. Disponível em: <<https://CRAN.R-project.org/package=sampler>>.

BARROS, R. P.; CARVALHO, M.; MENDONÇA, R. S. P. “Sobre as utilidades do Cadastro Único”. In: CASTRO, J. A.; MODESTO, L. (Org.). *Bolsa Família 2003-2010: avanços e desafios*. Brasília: Ipea, 2010, v. 1, p. 179-212.

BRASIL, Ministério do Desenvolvimento Social e Combate à Fome. Secretaria Nacional de Assistência Social. Brasília, mar. 2024. Disponível em: <https://cecad.cidadania.gov.br/painel03.php#top>. Acesso em: 20 abr. 2024.

CARMO, M. E.; GUIZARDI, F. L. “O Conceito de vulnerabilidade e seus sentidos para as políticas públicas de saúde e assistência social”. *CSP - Cadernos de Saúde Pública*. 34 (3), 2018.

DIREITO, D. C.; KOGA, N. M.; LICIO, E. C.; CHAVES, J. C. P. N. “O Cadastro Único como instrumento de articulação de políticas sociais”. n. 145, p. 24, 2016. Disponível em: <https://ipcig.org/pub/port/WP145PT_Cadastro_Unico_como_instrumento_de_articulacao.pdf>. Acesso em 27 de fev. de 2020.

EMBRETSON, S. E. e REISE, S. P. “Item Response Theory for Psychologists”. New Jersey: *Lawrence Erlbaum Associates*, 2000.

HAMBLETON, R. K.; SWAMINATHAN, H. “Item Response Theory - Principles and

Applications”. [S.l.]: Springer, 1985.

JELIHOVSCHI, E. *Análise Exploratória de Dados usando o R*. Ilhéus: Editora da UESC, 2014.

KLEIN, R. “Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB)”. *Revista Ensaio, [S. l.]*, v. 11, n. 40, p. 283–296, 2003. DOI: <http://dx.doi.org/10.22347/2175-2753v1i2.38>.

MDS – Ministério da Cidadania. Disponível em: mds.gov.br/assuntos/cadastro-único/. Acesso em 26 de fev. 2020.

MEC - MINISTÉRIO DA EDUCAÇÃO. Disponível em: <http://portal.mec.gov.br/component/tags/tag/34530>. Acesso em 15 de março de 2020.

MELATTI, J. C. *Índios do Brasil*. São Paulo: Editora Hucitec, 1987.

MUÑIZ FERNÁNDEZ, J. *Teoría de respuesta a los ítems: un nuevo enfoque en la evolución psicológica y educativa*. [S.l.: s.n.], 1990.

NUNES, C. H. S. da S.; PRIMI, R. “Impacto do Tamanho da Amostra na Calibração de Itens e Estimativa de Escores por Teoria de Resposta ao Item”. *Avaliação Psicológica*, v. 4, n. 2, p. 141–153, 2005.

PARANHOS FALEIRO, R. *Alegorias das Políticas de Territorialidade: um estudo sobre a afirmação de direitos étnicos dos Ingarikó na sobreposição entre o Parque Nacional do Monte Roraima e a Terra Indígena Raposa Serra do Sol em comparação com as etnias Taironas de Sierra Nevada em Santa Marta, Colômbia*. Brasília: UnB/CEPPAC, 2015.

REISE, S.; AINSWORTH, A.; HAVILAND, M. “Item Response Theory”. *Current Directions in Psychological Science*, 14(2), 95-101, 2005.

RODRIGUES, E. C.; MATOS, D. A. S.; FERREIRA, A. dos S. “Nível socioeconômico e ensino superior: cálculo e aplicações”. *Revista da Avaliação da Educação Superior*. São Paulo, Campinas, v. 22, n. 2, p. 494–511, ago. 2017. <https://doi.org/10.1590/s1414-40772017000200013>.

SAMEJIMA, F. “Estimation of latent ability using a response pattern of graded scores”. *Psychometrika*, v. 34, n. S1, p. 1–97, 1969.

SANTOS, M. E.; VILLATORO, P. A Multidimensional Poverty Index for Latin America. **Review of Income and Wealth**, v. 64, n. 1, p. 52–82, mar. 2018.

SARTES, L. M. A.; DE SOUZA-FORMIGONI, M. L. O. “Avanços na psicometria: da teoria clássica dos testes à teoria de resposta ao item”. *Psicologia: Reflexão e Crítica*, v. 26, n. 2, p. 241–250, 2013. <https://doi.org/10.1590/S0102-79722013000200004>.

SEDES – Secretaria de Desenvolvimento Social. *O que é Cadastro Único*. Disponível em: <<http://www.sedes.df.gov.br/cadastro-unico/>>. Acesso em 01 de março de 2020.

SEN, A. **Desenvolvimento como liberdade**. [s.l.] Companhia das Letras, 2021.

SILVA, R. S. S. *Avaliação de desempenho na administração pública municipal brasileira pela TRI*. Dissertação. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Administração, Florianópolis, 2018.

SOARES, T. M. “Utilização da teoria de resposta ao item na produção de indicadores sócio-econômicos”. *Pesquisa Operacional*, v.25, n. 1, p.83-112, abr. 2005. <https://doi.org/10.1590/S0101-74382005000100006>.

SPENESSATO, D. *Manutenção do Banco de Itens para Testes Adaptativos Computadorizados Aplicados em Avaliações de Alto Impacto*. Tese. 2017. Universidade Federal de Santa Catarina, [S. l.], 2017.

TEZZA, R.; BORNIA, A. C. Teoria de Resposta ao Item: Vantagens e Oportunidades para a Engenharia de Produção. In XXIX ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO 2009, Salvador. Anais. Salvador: ABEPRO, 2009. p. 13.

TEZZA, R. et al. “Empirical Comparison of the Multidimensional Models of Item Response Theory in E-Commerce”. *Pesquisa Operacional*, v. 36, n. 3, p. 503–532, 2016.

VAN DER LINDEN, W. J.; HAMBLETON, R. K. “Item Response Theory: brief History, common Models and Extensions”. In VAN DER LINDEN, Wim J.; HAMBLETON, Ronald K. *Handbook of Modern Item Response Theory*. New York, NY: Springer, 1997. b. p. 1–28. DOI: 10.1007/978-1-4757-2691-6_1. Disponível em: http://link.springer.com/10.1007/978-1-4757-2691-6_1.

VEY, I. H. *Avaliação de Desempenho Logístico no Serviço ao Cliente baseada na Teoria da Resposta ao Item*. Tese. 2011. Universidade Federal de Santa Catarina, [S. l.], 2011.

VIEIRA, M. J.; RIBEIRO, R. B.; ALMEIDA, L.; PRIMI, R. “Comparação de modelos da Teoria de Resposta ao Item (TRI) na validação de uma prova de dependência-independência de campo”. *Avaliação Psicológica*, vol.10, no.1. Porto Alegre, abr. 2011.

ZUKOWSKY-TAVARES, C. “Teoria de Resposta ao Item: Uma Análise Crítica dos Pressupostos Epistemológicos”. *Estudos em Avaliação Educacional*, [S. l.], v. 24, n. 54, p. 56–76, 2013.

Vulnerabilidade Socioeconômica de Indígenas e Quilombolas no Brasil: uma proposta de análise do Cadastro Único (CadÚnico)

Resumo: O artigo propõe uma amostragem da base de dados “Famílias” presente no Cadastro Único (CadÚnico). O CadÚnico tem o registro de famílias beneficiárias e não beneficiárias do PBF (Programa Bolsa Família) residentes em todo o território nacional. A metodologia de análise constitui um modelo proposto a partir da TRI, com scripts estabelecidos de importação, de tratamento dos dados, que possibilitaram determinar faixas de vulnerabilidade socioeconômica (agrupamentos) de famílias indígenas e quilombolas, considerando que esses grupos assumam uma condição de vulnerabilidade dentro de um árduo processo histórico de superação de desigualdades e, em alguns casos, de tentativa de preservação de suas identidades. Os resultados sugerem que tais famílias se encontram bem posicionadas em relação à efetividade da política pública de distribuição de renda promovida pelo PBF no período analisado.

Palavras-chave: Vulnerabilidade; CadÚnico; TRI (Teoria de Resposta ao Item); Famílias Indígenas; Famílias Quilombolas.

Vulnerabilidad socioeconómica de indígenas y quilombolas en Brasil: una propuesta de análisis de CadÚnico

Resumen: El artículo propone una muestra de la base de datos “Famílias” presente en el Cadastro Único (CadÚnico). CadÚnico registra familias beneficiarias y no beneficiarias del PBF (Programa Bolsa Família) residentes en todo el territorio nacional. La metodología de análisis constituye un modelo propuesto desde el TRI, con guiones establecidos de importación y tratamiento de datos, que permitió determinar rangos de vulnerabilidad socioeconómica (agrupaciones) de familias indígenas y quilombolas, considerando que estos grupos asumen una condición de vulnerabilidad dentro de un arduo proceso histórico de superación de desigualdades y, en algunos casos, de intento de preservación de sus identidades. Los resultados sugieren que estas familias se encuentran en una buena posición en relación a la efectividad de la política pública de distribución del ingreso impulsada por el PBF en el período analizado.

Palabras clave: Vulnerabilidad; CadÚnico; TRI (Teoría de la Respuesta al Ítem); Familias Indígenas; Familias Quilombolas.

Socioeconomic Vulnerability of Indigenous and Quilombola in Brazil: a CadÚnico analysis proposal

Abstract: The article proposes a sampling of the “Families” database present in the Cadastro Único (CadÚnico). CadÚnico registers beneficiary and non-beneficiary families of the PBF (Bolsa Família Program) residing throughout the national territory. The analysis methodology constitutes a model proposed from the TRI, with established import and data treatment scripts, which made it possible to determine ranges of socioeconomic vulnerability (groupings) of indigenous and quilombola families, considering that these groups assume a condition of vulnerability within of an arduous historical process of overcoming inequalities and, in some cases, of trying to preserve their identities. The results suggest that these families are in a good position in relation to the effectiveness of the public income distribution policy promoted by the PBF in the analyzed period.

Keywords: Vulnerability; CadÚnico; IRT (Item Response Theory); Indigenous Families; Quilombolas Families.