ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mostafa, Joana; dos Santos, Thuany

Working Paper

Limitações de um teste de meios via predição de renda: Evidências de uma aplicação no Programa Bolsa Família

Texto para Discussão, No. 2238

Provided in Cooperation with:
Institute of Applied Economic Research (ipea), Brasília

Suggested Citation: Mostafa, Joana; dos Santos, Thuany (2016) : Limitações de um teste de meios via predição de renda: Evidências de uma aplicação no Programa Bolsa Família, Texto para Discussão, No. 2238, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília

This Version is available at:
https://hdl.handle.net/10419/177454

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

2238
TEXTO PARA DISCUSSÃO

LIMITAÇÕES DE UM TESTE DE MEIOS VIA PREDIÇÃO DE RENDA: EVIDÊNCIAS DE UMA APLICAÇÃO NO PROGRAMA BOLSA FAMÍLIA

Joana Mostafa
Thuany dos Santos

ipea

2238
TEXTO PARA DISCUSSÃO
Rio de Janeiro, outubro de 2016

LIMITAÇÕES DE UM TESTE DE MEIOS VIA PREDIÇÃO DE RENDA: EVIDÊNCIAS DE UMA APLICAÇÃO NO PROGRAMA BOLSA FAMÍLIA[1]

Joana Mostafa[2]
Thuany dos Santos[3]

1. Este trabalho não poderia ter sido realizado sem o apoio e a dedicação da equipe do Departamento do Cadastro Único (CadÚnico) do Ministério do Desenvolvimento Social, das diversas coordenações municipais do CadÚnico e do Ipea, que permitiram a formulação, a aplicação e a análise dos dados da predição de renda aplicada no âmbito do processo de averiguação cadastral de 2013. Igualmente gostaríamos de agradecer o incentivo do então secretário de Renda de Cidadania, Luis Henrique da Silva de Paiva, para que esta aplicação fosse levada a cabo e registrada pelas autoras, bem como a sua revisão rigorosa e cuidadosa. Erros, omissões e opiniões são de inteira responsabilidade das autoras.
2. Técnica de planejamento e pesquisa da Diretoria de Estudos e Políticas Sociais (Disoc) do Ipea.
3. Estatística do Comando da Aeronáutica (Comaer).

Governo Federal

Ministério do Planejamento, Desenvolvimento e Gestão
Ministro interino Dyogo Henrique de Oliveira

ipea **Instituto de Pesquisa Econômica Aplicada**

Fundação pública vinculada ao Ministério do Planejamento, Desenvolvimento e Gestão, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidente
Ernesto Lozardo

Diretor de Desenvolvimento Institucional, Substituto
Carlos Roberto Paiva da Silva

Diretor de Estudos e Políticas do Estado, das Instituições e da Democracia
João Alberto De Negri

Diretor de Estudos e Políticas Macroeconômicas
Claudio Hamilton Matos dos Santos

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais
Alexandre Xavier Ywata de Carvalho

Diretora de Estudos e Políticas Setoriais de Inovação, Regulação e Infraestrutura
Fernanda De Negri

Diretora de Estudos e Políticas Sociais
Lenita Maria Turchi

Diretora de Estudos e Relações Econômicas e Políticas Internacionais
Alice Pessoa de Abreu

Chefe de Gabinete
Márcio Simão

Assessora-chefe de Imprensa e Comunicação
Regina Alvarez

Ouvidoria: http://www.ipea.gov.br/ouvidoria
URL: http://www.ipea.gov.br

Texto para Discussão

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pelo Ipea, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões.

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2016

Texto para discussão / Instituto de Pesquisa Econômica Aplicada.- Brasília : Rio de Janeiro : Ipea , 1990-

ISSN 1415-4765

1.Brasil. 2.Aspectos Econômicos. 3.Aspectos Sociais. I. Instituto de Pesquisa Econômica Aplicada.

CDD 330.908

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério do Planejamento, Desenvolvimento e Gestão.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

JEL: I3.

SUMÁRIO

SINOPSE

ABSTRACT

1 INTRODUÇÃO: HIPÓTESES PARA A BOA COBERTURA DO CADASTRO ÚNICO.....7

2 UM PREDITOR DE RENDA: POSSÍVEIS USOS E METODOLOGIA 13

3 RAZÕES DA DISTÂNCIA ENTRE A PREDIÇÃO E A REALIDADE DAS FAMÍLIAS REGISTRADAS NO CADÚNICO 26

4 CONCLUSÃO 36

REFERÊNCIAS 39

APÊNDICE 40

SINOPSE

O impacto do Programa Bolsa Família (PBF) na redução da pobreza extrema se deveu, basicamente, à sua precisão e ao seu alcance. Ou seja, além da audácia de alcançar praticamente todas as famílias pobres do país, o programa o fez com grande precisão, de fato beneficiando quem dele necessita. Seu impacto só não é ainda maior devido ao modesto valor dos benefícios em relação à renda brasileira, transferindo apenas 0,5% do produto interno bruto (PIB) aos 25% mais pobres. Não obstante as evidências de seu sucesso, apontadas por analistas nacionais e internacionais, o espectro da má focalização ronda novamente sua gestão, trazido à baila pela crise política do governo da presidenta Dilma Rousseff. Nada mais distante da realidade. O PBF, um raro exemplo na política pública brasileira, não só realiza profundas auditorias cadastrais há mais de dez anos, como foi pioneiro na aplicação de testes de meios mediante a predição estatística em nível nacional. Assim, a aplicação de um *proxy means test* (PMT) no âmbito da auditoria de 2013 trouxe inúmeros aprendizados sobre as dificuldades de se romper com as situações de pobreza, bem como com as sérias limitações deste tipo de modelo estatístico para se determinar a elegibilidade ou realizar a revisão de benefícios essenciais à sobrevivência destas famílias. Esta aplicação inédita e seus aprendizados são registrados neste *Texto para Discussão*.

Palavras-chave: Programa Bolsa Família; focalização; teste de meios; *proxy means test*.

ABSTRACT

The impact of Programa Bolsa Família in reducing extreme poverty has its roots on the program's precision as well as coverage. Its results are only hindered by the modest values transferred, totaling only 0,5% of the Brazilian GDP in the hands of the poorest quartile. Notwithstanding the national and international evidence of the Program's success, the specter of bad targeting reappeared as the political crisis of Dilma Rousseff's government thrived. Nothing further from reality. Programa Bolsa Família is a rare example in Brazilian public policy of constant auditing and benefit revision. Additionally, its responsible managers have applied a proxy means test methodology in the context of the 2013 audit, shedding more light on the determinants of poverty in Brazil and revealing serious limitations of these instruments, be it to determine

eligibility or revise benefits so essential to these families' survival. These findings are registered in this discussion paper.

Keywords: Bolsa Família; targeting; means test; proxy means test.

1 INTRODUÇÃO: HIPÓTESES PARA A BOA COBERTURA DO CADASTRO ÚNICO

O Cadastro Único (CadÚnico) para programas sociais do governo federal tem por objetivo registrar os dados pessoais e o perfil socioeconômico de todos os brasileiros com renda familiar abaixo de meio salário mínimo *per capita*[1] além das famílias selecionadas por serviços e programas específicos. Em janeiro de 2015, havia um total de 26,7 milhões de famílias cadastradas, das quais 19,7 milhões tinham atualizado suas informações pelo menos uma vez nos últimos dois anos, o que indica maior aderência da informação à situação vivida pelas famílias. Assim, o cadastro fornece informações razoavelmente atualizadas de aproximadamente um terço da população brasileira.

Até o momento, a maior prova de que o CadÚnico reflete de maneira fidedigna as condições de vida da população mais pobre do país é o impacto do Programa Bolsa Família (PBF) na redução da desigualdade brasileira (Ipea, 2013, p. 27).[2] De toda a queda do índice de Gini entre 2002 e 2012 – que passou de 58,7% em 2002 para 52,6% em 2012 –, 17% deveram-se ao PBF, sendo que o gasto com o programa representa apenas 0,5% do produto interno bruto (PIB) brasileiro. Esta queda, apesar de parecer pequena em termos de Gini, significa que, se continuássemos neste ritmo, chegaríamos, em cerca de vinte anos, ao nível de países notoriamente igualitários, como o Canadá.

O impacto do PBF se deveu, basicamente, à sua precisão e ao seu alcance. Ou seja, além da audácia de alcançar praticamente todas as famílias pobres do país, o programa o fez com grande precisão, de fato beneficiando quem dele necessita. Seu impacto só não é ainda maior devido ao modesto valor dos benefícios em face da renda brasileira, transferindo apenas 0,5% do PIB aos mais pobres. As famílias beneficiárias recebem, em média, R$ 167 ao mês. As evidências coletadas por meio da Pesquisa Nacional por Amostra de Domicílios (Pnad) do Instituto Brasileiro de Geografia e Estatística (IBGE) mostram ainda que a concentração das transferências do programa

1. As outras possibilidades não censitárias de registro no CadÚnico são: famílias com renda total de até três salários mínimos; e público de programas e serviços sociais que utilizem o CadÚnico como instrumento de seleção ou acompanhamento, independentemente da renda.
2. Sobre evidências históricas e internacionais ver Soares (2008).

para a população mais pobre aumentou (Ipea, 2012, p. 28).[3] Ao longo dos últimos onze anos, isso ocorreu mesmo nas ocasiões em que o programa se aproximava do topo de sua meta de atendimento.

A princípio, o alcance da meta de atendimento do PBF poderia tornar cada vez mais difícil a tarefa de encontrar e incluir famílias realmente elegíveis que ainda não tenham sido contempladas. De fato, trata-se de um desafio, mas este tem sido superado pelas gestões municipais, em especial após o lançamento da ideia força da Busca Ativa. Ademais, o aumento dos benefícios instituído pela estratégia do programa Brasil sem Miséria – que assegurou que cada família recebesse o suficiente para superar a extrema pobreza – supostamente aumentaria os incentivos para a subdeclaração ou a omissão de renda. Como veremos adiante, nada disso ocorreu.

Existem vários mecanismos e regras da gestão do CadÚnico que colaboraram para que a cobertura das famílias de baixa renda seja alcançada, até mesmo sem intenção explícita de seu formulador. Não sabemos exatamente a parte que cabe a cada mecanismo, mas consideramos importante enumerar algumas conjecturas. Espera-se que as hipóteses aqui levantadas apresentem aos leitores um rico panorama do tema, para além de definições clássicas de erro de inclusão e exclusão de famílias.

Então, quais as principais hipóteses que explicam esses bons resultados?

Primeiramente, a expansão do CadÚnico se deu, basicamente, pela *autosseleção* das famílias. Do ponto de vista da cobertura cadastral, a estratégia de comunicação massiva e eficaz em torno do mais importante programa do popular e recém-eleito presidente Lula possibilitou a expansão do CadÚnico como mecanismo de identificação e seleção do PBF. As famílias mobilizaram-se para o cadastramento sabendo, em linhas gerais, do objetivo do programa, então sob o guarda-chuva do Fome Zero: aliviar a pobreza. A identificação de uma parte expressiva da população com as origens humildes e nordestinas do novo presidente impulsionou a autosseleção. Ou seja, de um lado, a efetiva pobreza deste segmento, que torna atrativo o cadastramento no PBF mesmo

3. O coeficiente de concentração do PBF nada mais é que um Gini, ou seja, a desigualdade somente da renda do PBF, a qual em 2011 era de -52,5. O programa é, portanto, absurdamente concentrado nos pobres, contrapondo-se, por exemplo, à renda do trabalho, que tem sinal inverso e era concentrador em 53,9 no mesmo ano, explicando grande parte da desigualdade (índice de Gini) da renda brasileira em 2011, de 52,7.

para receber um benefício modesto, e, de outro, o chamamento crível por parte do governo federal confluíram para uma demanda por cadastramento gigantesca entre 2003 e 2006.[4]

Do ponto de vista da oferta do programa, o chamamento do governo federal, balizado por metas de atendimento ousadas e por transferências automáticas e diretas do sistema bancário às famílias, mesmo utilizando-se de estruturas municipais, possibilitou uma expressiva autonomização em relação aos poderes locais. Do ponto de vista da demanda das famílias, a autosseleção apenas daquelas em situação de pobreza se dá, primordialmente, porque os benefícios comumente associados ao CadÚnico não são expressivos. Como resultado, a população de maiores rendimentos não se vê motivada a enfrentar um processo de cadastramento árduo e cometer a ilegalidade de subdeclarar sua condição socioeconômica.

Assim, o CadÚnico passa de cerca de 5 milhões de famílias em 2003, na sua maioria beneficiárias do antigo programa Bolsa Escola, para 14 milhões de famílias em 2006. Os cadastros realizados por meio de visita domiciliar, entre as inúmeras estratégias de cadastramento, representam apenas uma pequena parte das famílias registradas no CadÚnico, seja nos cadastros novos, seja nas atualizações cadastrais. Desde que este número pôde ser acompanhado, a partir da implantação do novo formulário, em 2011, a porcentagem das famílias registradas variou de 4,4% em dezembro de 2011 para 7,3% em julho de 2014, passando de um volume de 485 para 1.787 famílias visitadas. Nesse sentido a autosseleção foi, e continua sendo, o mecanismo predominante de cobertura das famílias pobres.

Segundo, a divulgação e o monitoramento de *metas municipais* de cobertura e a atualização cadastral serviram como parâmetros para a expansão. As metas foram

4. É importante fazer referência ao contexto que possibilitou o reconhecimento da "questão" da pobreza e da fome. Nos anos 1990 e início dos anos 2000, foram inúmeras as manifestações artísticas, culturais e sociais que delinearam a questão social no Brasil e construíram um razoável consenso acerca das sérias privações vividas por parte expressiva de nossa sociedade. Configurou-se um caldo cultural de inconformismo diante da situação de fome destas populações e da violência que sofriam, o que gerou um apelo por mínimos direitos sociais. O exemplo mais contundente é o movimento social Ação da Cidadania, que nasce em 1993 liderado pelo expoente nacional Betinho, mas cita-se toda uma safra de filmes nacionais sobre a questão: *Central do Brasil* (1998), de Walter Salles; *Domésticas* (2001), de Fernando Meirelles; *Abril Despedaçado* (2001), de Walter Salles; *Amarelo Manga* (2002), de Cláudio Assis; *Cidade de Deus* (2002), de Fernando Meirelles; e Carandiru (2003), de Hector Babenco, que é baseado no massacre de Carandiru, ocorrido em 1992.

utilizadas, principalmente, como estimativas municipais do erro de exclusão, no sentido de incentivar os municípios que fizessem maiores esforços de cobertura com repasses orçamentários. Este mecanismo legitimou e apoiou, financeiramente, a estruturação da rede de atendimento para atender à demanda de cadastramento das famílias, além de, como vimos, aumentar a realização de visitas domiciliares.

É importante ressaltar que o CadÚnico instituiu, desde o início, a meta de inclusão cadastral de todas as famílias com renda de até meio salário mínimo por pessoa, o que significava cerca de o dobro da linha de pobreza visada pelo PBF no início dos anos 2000. Esse é um elemento importante porque todas as alterações da linha de atendimento do PBF puderam contar com uma demanda de famílias previamente identificadas e dimensionadas pelo CadÚnico, não havendo corridas por atendimento ou descontinuidades na cobertura do programa. Além disso, a própria meta de atendimento do Bolsa Família segue uma metodologia de cálculo muito mais inclusiva que a simples foto do número de pobres, dada uma certa linha de pobreza (*headcount*). O programa considera fatores como a volatilidade de renda para incrementar a foto e tentar alcançar as famílias que são ou serão pobres ao longo de um filme de dois anos. Assim, as metas amplas de inclusão no CadÚnico proporcionam aos programas, usuários de suas informações, a possibilidade de implementarem desenhos sofisticados e inclusivos de proteção social.

À medida que o CadÚnico chegou à sua meta de cobertura de famílias, bem como o PBF chegou à sua de atendimento, houve uma pressão na ponta municipal pela qualificação, atualização e até exclusão das informações cadastrais. Diariamente, famílias entram na pobreza e na extrema pobreza e demandam cadastramento e benefícios. Se o município não conseguir identificar as famílias que saíram da pobreza na mesma velocidade, fica em situação delicada diante dos novos demandantes. Assim, no atual estágio de cobertura do Bolsa Família (100%), o sistema de metas municipais atua como um mecanismo de focalização, no sentido de amarrar o erro de exclusão ao de inclusão. Este é o estágio deste mecanismo atualmente.

Terceiro, para a correção dos erros de inclusão, outros mecanismos podem ter surtido efeitos líquidos positivos. De um lado, cruzamentos sistemáticos com outros *registros administrativos* com informação de rendimentos do trabalho formal, de óbito e dos benefícios do Instituto Nacional de do Seguro Social (INSS) geram efeito imediato

sobre as pessoas com indícios de subdeclaração ou omissão de renda. Isto inibe a atualização cadastral das famílias diretamente afetadas e gera efeito demonstração para as que não são de baixa renda e não estavam cadastradas. Estes cruzamentos ocorrem desde 2005 e são de amplo conhecimento da população cadastrada. De outro lado, a responsabilização das famílias e da gestão local do CadÚnico pela declaração e pelo registro de informações verídicas e atuais torna crível a aplicabilidade das sanções cabíveis.[5]

Tendo o governo local tomado as devidas providências para cumprir com os preceitos normativos e operativos exarados pelo nível central, e tendo instituído práticas de responsabilização das famílias pela informação declarada, como o armazenamento de formulários assinados e a própria atualização cadastral, foi possível desresponsabilizar os poderes locais e instituir a *autodeclaração* das informações cadastrais como preceito básico do CadÚnico.

Um dos grandes entraves à construção prévia e à manutenção de mecanismos como o CadÚnico é a dificuldade do Estado de instituir a autodeclaração para as parcelas desfavorecidas da população, apesar deste instituto já existir para as mais favorecidas, como no caso da declaração do imposto de renda da pessoa física. Para incluir, de forma transparente e republicana, a parcela informal da população no cadastro, seria preciso acreditar na palavra dos próprios pobres, sem lançar mão do empregador para registrar os eventos trabalhistas. Assim, em alguma medida, o CadÚnico permitiu quebrar a histórica cultura política de tutela entre o cidadão pobre e o Estado, majorada pela herança paternalista do campo da assistência social. A autodeclaração acoplada a um grande esforço de padronização dos conceitos cadastrais possibilitaram, em grau diverso em cada município, a redução do poder de "indicação" dos gestores locais.

Por fim, contribuiu para a focalização do CadÚnico a circunscrição da atividade de cadastramento, na grande maioria dos municípios, ao campo da *assistência social*. Historicamente, a assistência social, até 1988, proveu serviços e transferências aos mais

5. As famílias assinam o formulário em que se comprometem com a veracidade das informações declaradas e com a sua atualização. A rastreabilidade da declaração das famílias é uma grande preocupação das gestões locais, que, para se protegerem de imputação equivocada, têm a prática de armazenar todos os formulários de cadastramento em papel, com a assinatura do responsável familiar. Este modelo de responsabilização das famílias e das gestões locais tem sido bem-sucedido, mas representa um fardo bastante oneroso para a rede de assistência social, elevando a tensão das relações com seu público.

pobres nas brechas das políticas públicas até então contributivas, como a saúde e a previdência, ou ainda nas lacunas das políticas educacionais, caso da educação infantil, rural, de jovens e adultos e da qualificação profissional.[6]

Está claro, no que nos interessa aqui, que a relação do CadÚnico com o campo da assistência social possibilitou uma abordagem das famílias, não inteiramente, porém mais apropriada a sua realidade, linguagem e cultura, reduzindo os preconceitos institucionais, as fricções de comunicação, as exigências documentais. Ou seja, de modo geral, reduziram-se as barreiras à entrada das famílias mais pobres ao CadÚnico.[7]

O CadÚnico e o PBF desenvolveram-se concomitantemente à efetivação e à delimitação do campo de atuação da assistência social como política pública no país. Assim, as estruturas de atendimento do CadÚnico e do Bolsa Família foram montadas *pari passu* à construção dos equipamentos públicos da assistência social em nível municipal e estadual – Centro de Referência de Assistência Social (Cras), Centro de Referência Especializado de Assistência Social (Creas), Centro de Referência Especializado para População em Situação de Rua (Centros POP) etc. –, à regulamentação e à estruturação de suas equipes de referência, e ao mesmo tempo que se delimita, tipifica, o próprio "fazer" dos atendimentos e serviços socioassistenciais. Estas estruturas, equipamentos, normativos, sintaxes, conceitos, enfim, sistemas, foram sendo construídos concomitantemente, entremeando-se, influenciando-se mutuamente, sem perder, logicamente, as especificidades próprias de cada campo.

Este texto se propõe a contribuir para apenas uma pequena parte dessa discussão: após a alteração do PBF, em 2011, que passa a cobrir o hiato de extrema pobreza

6. O campo da assistência ganha contornos próprios após a Constituição de 1988, com a progressiva regulamentação do Sistema Único de Assistência Social (Suas) e a importante decisão de se criar o Ministério do Desenvolvimento Social e Combate à Fome (MDS), sob o qual o campo da assistência social conseguiu superar a fragilidade institucional anterior.

7. Isso fica mais claro quando se evidencia a dificuldade de adequação das políticas previdenciárias e do campo da habitação, por exemplo, à população mais pobre, na recente articulação do CadÚnico com o plano previdenciário Facultativo de Baixa Renda e com o Programa Minha Casa Minha Vida (PMCMV). O distanciamento semântico das regras, a exigência de conhecimentos prévios, além da própria geografia das estruturas de atendimento, inibe a expansão destas políticas para a população pobre. O emaranhado de dissonâncias entre a população e o estado, bem como entre as agências executoras destas políticas em nível municipal e em nível federal, tornam sua operação custosa, custo que recai sobre o próprio cidadão. O viés informacional destas políticas, o insulamento burocrático das instituições que as operam, toda a carga de nomenclaturas, os sistemas informacionais, as exigências formuladas inicialmente para outros públicos e adaptadas para um recorte de menor renda levam a enormes dificuldades de articulação com o CadÚnico e o seu público.

entre a renda declarada e a linha delimitada pelo programa, podemos dizer que as famílias estão agindo estrategicamente, levando a um maior erro de inclusão do Bolsa Família? Havendo erro na renda declarada, qual a maior causa deste erro? Desatualização cadastral? Volatilidade da renda? Omissão por parte das famílias?

Esse exercício será dividido em duas partes e uma conclusão. Primeiramente far-se-á um breve relato metodológico do modelo estatístico aqui utilizado, desenvolvido pelo Ipea e aprimorado pelo MDS para melhor adaptação aos dados do CadÚnico. Em seguida, pretende-se explorar as razões dos possíveis erros de medida da renda familiar *per capita* a partir de uma pesquisa levada à campo pelo próprio MDS e pelas gestões municipais.

2 UM PREDITOR DE RENDA: POSSÍVEIS USOS E METODOLOGIA

Preditor de renda é o nome genérico de uma família variada de modelos estatísticos que associam um número de características de uma família ou pessoa a uma probabilidade de se ter certo nível de renda. Este tipo de instrumento se aproxima dos procedimentos de seleção de famílias utilizados por programas de transferências de renda em outros países da América Latina, como é o caso do programa Oportunidades, do México, do Más Famílias en Acción, da Colômbia – por meio do sistema de beneficiários –, bem como da pontuação da Ficha de Proteccíon Social, utilizada pelo Chile Solidario.

Trata-se, entre os diversos mecanismos de seleção de beneficiários, de um instrumento de medição da elegibilidade das famílias em nível individual que vai além de recortes geográficos ou demográficos. No campo das transferências de renda, a adoção de estimadores estatísticos de pobreza pode concorrer, a depender do desenho, com outros mecanismos de seleção em nível individual, ou complementá-los, como acontece com a seleção baseada na indicação de setores organizados de uma dada comunidade (*community based targeting*) ou com a seleção baseada na procura da própria família: a autosseleção. São os famosos *proxy means tests* (PMT).[8]

8. Ver a esse respeito: Coady, Grosh e Hoddinott (2004); Kidd e Wylde (2011) e Kidd (2013).

O Bolsa Família é classificado internacionalmente como parte dos programas que executam um teste de meios ou recursos por aproximação (*means-test*) para a seleção de uma família, ou seja, que coleta informações, verificáveis ou não, que possibilitem a avaliação da vulnerabilidade ou a baixa renda. Esta coleta é justamente o registro no CadÚnico. A informação utilizada para a elegibilidade ao programa não é uma *proxy* de pobreza, mas a insuficiência de renda em si, estabelecida a partir da renda autodeclarada.

Como vimos anteriormente, ao rever a história de expansão do CadÚnico, na prática, houve preponderância da autosseleção das famílias acoplada à verificação, *a posteriori*, dos rendimentos formais, de óbito e dos benefícios contributivos. Neste sentido, o modelo brasileiro se diferencia bastante do que é correntemente chamado de *proxy means test*, no qual a coleta de informações permite inferir, com algum grau de certeza, a situação de vulnerabilidade das famílias. O PMT pode variar substancialmente de país para país, indo desde um modelo de predição de renda bastante complexo até um simples indicador multidimensional de pobreza com pesos iguais entre suas dimensões.

No modelo brasileiro, a elegibilidade é determinada tão somente por meio da renda declarada pela família, que, devido à informalidade de seu engajamento no mercado de trabalho, é na maioria das vezes anunciada sem comprovação documental. A verificação da renda após o cadastramento é feita em nível nacional com base nos registros de rendimentos formais do trabalho (Relação Anual de Informações Sociais, do Ministério do Trabalho e Emprego – Rais/MTE) ou nos benefícios da previdência social. A verificação é compreendida, entretanto, apenas como um indício de inconsistência, devido à alta rotatividade do mercado formal de trabalho brasileiro, ao atraso da captura e da sistematização, e aos possíveis erros dos registros de trabalho, óbito e benefícios no Brasil. A partir deste indício, divulgam-se para a rede de gestores municipais do CadÚnico as listas das famílias, para que seja realizada a atualização cadastral, com envio simultâneo e direto aos cidadãos cadastrados de cartas e mensagens no extrato de benefícios. Por fim, se a atualização do cadastro da família não ocorrer no prazo divulgado, ou se houver alteração cadastral que provoque o desenquadramento, as famílias podem ter seus benefícios cancelados.

Mesmo com um modelo de autodeclaração com verificação *a posteriori*, seria possível utilizar estes PMTs para, por exemplo, incrementar as verificações da renda declarada? Este método permite de fato apontar, com uma alta probabilidade, famílias

que estariam subdeclarando renda? Para ensaiar respostas a estas perguntas, foi construído e aplicado[9] um preditor de renda desenvolvido pelo Ipea que utiliza os dados do Censo Demográfico do Brasil de 2010 e do CadÚnico. A especificação e a aplicação são feitas em quatro etapas.

1) Selecionam-se as variáveis que são compatíveis ao CadÚnico e ao Censo 2010. Apenas as variáveis com preenchimento consistente e que tenham conceitos bastante próximos, comparando-se uma base com a outra, são escolhidas para compor o modelo.
2) É feita uma análise de agrupamento (*cluster*) no Censo 2010 com o intuito de diferenciar as famílias "pobres" das "não pobres" entre aquelas cujo rendimento total é igual a zero. As famílias não pobres, ou seja, aquelas cujas características observáveis mais estruturais apontam para um perfil de renda razoavelmente maior que zero, são excluídas da construção do modelo. Isto é feito para que sua boa situação não influencie, erroneamente, os parâmetros que associam as características observáveis à situação de pobreza.
3) Para as observações de todas as famílias do Censo 2010, exceto para aquelas que tinham renda igual a zero e foram classificadas como não pobres no agrupamento indicado no item anterior, é calibrado e especificado um modelo de predição de renda.
4) Por fim, com os parâmetros do Censo 2010 aplicados às informações coletadas no CadÚnico, estima-se a probabilidade de a renda estar acima da declarada no cadastro.

Na *primeira etapa*, a partir das variáveis compatíveis e do bom preenchimento no Censo 2010 e no CadÚnico, foram construídos os seguintes atributos:

- número de cômodos;
- número de cômodos que servem como dormitório;
- material predominante nas paredes externas do domicílio;
- forma do abastecimento de água;

9. A posição da base de dados do CadÚnico utilizada neste artigo foi de 19 de julho de 2014. Muitos estimadores foram testados pela equipe da diretoria social do Ipea ao longo dos anos. O estimador aqui utilizado foi essencialmente desenvolvido pelo técnico Fábio Monteiro Vaz, a quem agradecemos pelas diversas críticas e o apoio ao trabalho do MDS e de suas equipes técnicas.

- forma de escoamento sanitário;
- tipo de iluminação;
- destino do lixo;
- número de homens por faixa etária;
- número de mulheres por faixa etária;
- se o responsável familiar tem cônjuge;
- proporção de pessoas na família por local de nascimento (migração);
- proporção de pessoas por grau de instrução e faixa etária;
- proporção de pessoas alfabetizadas por faixa etária;
- proporção de pessoas de até 17 anos por rede de ensino e faixa etária; e
- número de pessoas ocupadas por sexo e faixa etária.

Ainda neste passo, são definidos filtros importantes na base do CadÚnico e na do Censo 2010 para descartar informações inconsistentes ou não pertinentes. Ei-los:

1) São considerados apenas domicílios particulares permanentes, tanto no CadÚnico quanto no Censo 2010.
2) São consideradas apenas famílias do CadÚnico com estado cadastral regular, ou seja, "cadastrado", ou cuja data da última atualização seja mais recente que 19 de julho de 2010 (48 meses anteriores à data de extração do CadÚnico). As famílias com cadastros desatualizados há mais tempo não aportam a qualidade informacional necessária e figuram, para fins de estatísticas do CadÚnico, como inativas.
3) São consideradas apenas famílias atualizadas ou incluídas na versão 7 do aplicativo do CadÚnico, implementada desde agosto de 2010. Esta é a versão que possibilita a captura das informações utilizadas pelo modelo do preditor e num padrão compatível com o Censo 2010.
4) No Censo 2010 excluíram-se os pensionistas (moradores de pensão ou que alugam quarto da residência), empregados domésticos, parente do empregado doméstico e individual de domicílio coletivo, por não constarem no conceito de família do CadÚnico.
5) Para apurar-se a ocupação dos membros da família, foi considerada a informação declarada para os membros com 12 anos ou mais. Todas as pessoas abaixo disso foram consideradas não ocupadas. Este corte etário justifica-se pelo alto número de não preenchimento da variável ocupação para as crianças entre 10 e 12 anos no CadÚnico.

6) Para apurar-se o grau de instrução dos membros da família, foi considerada a informação declarada para os membros com 5 anos ou mais. Todas as pessoas abaixo desta idade foram consideradas sem instrução. Este corte etário se justifica pelo fato de a coleta do Censo 2010 investigar o grau de instrução apenas destas pessoas.

Na *segunda etapa*, o método de agrupamento (também chamado de conglomerado ou *cluster*) utilizado para o procedimento aqui descrito foi o método de *k*-means, que consiste em particionar as observações em *k* grupos com a maior distinção possível entre eles. O procedimento tenta alocar as observações de uma maneira a minimizar a variabilidade das variáveis descritoras dentro do grupo (intragrupo), a fim de maximizar a variabilidade destas entre os diferentes grupos.

A identificação dos grupos é feita pelas variáveis: tamanho da família, educação do chefe da família, existência de telefone fixo, telefone celular, máquina de lavar e microcomputador no domicílio. Baseado em Osório, Soares e Souza (2011), utilizando-se estas variáveis, consegue-se caracterizar dois grupos bastante distintos entre as famílias com renda zero: um considerado pobre; e outro apenas temporariamente sem renda, muito provavelmente não pobre. A tabela 1 mostra a média ou a proporção destas variáveis para os dois grupos.

TABELA 1
Média e proporção das variáveis associadas aos grupos de famílias de renda zero identificadas pelo método de agrupamento (*cluster*)

Variáveis utilizadas para agrupamento	Famílias tipicamente pobres	Famílias temporariamente pobres
Número médio de anos de estudo do chefe	2,53	10,74
Número médio de membros na família	2,86	2,62
Existência de máquina de lavar roupa (%)	24,6	51,9
Existência de telefone celular (%)	57,5	86,2
Existência de telefone fixo (%)	19,8	42,3
Existência de microcomputador (%)	11,7	43,8

Fonte: Censo 2010.

Assim, observa-se que, no Censo 2010, entre aquelas famílias com renda *per capita* zero, existem dois grupos bastante distintos. No primeiro, composto por 1,5 milhão de famílias identificadas como pobres pelo agrupamento, a média de anos de estudo do chefe é de 2,5 anos, enquanto no segundo grupo, de 1,0 milhão de famílias, a média de anos de estudo do chefe é de 10,7 anos. O modelo de agrupamento pelo

método de *k*-means consegue explicar 81,2% da variância entre os grupos (medida pela estatística *R*2), tendo alto poder explicativo. Assim, cria-se um filtro consistente para excluir o segundo grupo para fins de estimação do modelo do preditor.

Na *terceira etapa*, ajusta-se o modelo do preditor de renda no Censo 2010. O modelo escolhido pelo Ipea foi o de duas partes (Two-Part model), devido à grande quantidade de domicílios com renda zero. A ocorrência de muitas observações com valor zero na variável de interesse leva à adoção de técnicas que pretendem modelar dois obstáculos (*double-hurdle*): o primeiro diz respeito à decisão ou à situação de trabalhar ou não, de obter ou não renda, de consumir ou não um produto etc. Só então, depois da entrada em atividade, ou de um mercado consumidor, é que se pode modelar o valor potencialmente recebido ou gasto (segundo obstáculo). Assim, existem duas situações a serem modeladas: a primeira é de "participação" (trabalhar ou não, consumir um produto ou não etc.); e a segunda diz respeito ao nível de renda ou consumo.

O evento de renda zero têm inúmeras explicações. Mesmo que se procure trabalho ou que se esteja trabalhando, a renda ainda pode ser zero por fatores alheios à decisão individual: fatores climáticos, no caso da agricultura; pagamentos feitos em serviços ou produtos (escambo); produção para consumo próprio; ou consumo da renda em dívida com o empregador. Em função destas múltiplas explicações, o primeiro passo é modelar a probabilidade de a renda ser diferente de zero e conjugá-la à estimação do nível de renda, dado que a renda é maior que zero.

O modelo de duas partes pressupõe que exista dominância da primeira situação sobre a segunda, ou seja, a primeira situação (decisão de participação, ou, no nosso caso, a renda ser ou não maior que zero) limita as possibilidades da renda predita na segunda situação, a qual será sempre superior a zero. Se escolhe-se fumar, o nível predito de consumo de fumo será sempre maior que zero. Se há probabilidade de a pessoa trabalhar, o rendimento do trabalho predito será sempre maior que zero, e assim por diante. Divide-se assim, o modelo em duas partes, a probabilidade de haver renda maior que zero, e o valor estimado desta renda, dado que esta é maior que zero.

Assim, iniciamos o modelo de duas partes com a estimação da variável indicadora (*R*), que indica quando a renda familiar *per capita* predita é positiva ou igual à zero. Estima-se *R* a partir de:

$$R = 1 \text{ se } y > 0 \Rightarrow y^* = x_1^T \tau + e_1 \text{ sendo } e_1 \sim N(0,1);$$

$$R = 0 \text{ se } y = 0 \Rightarrow y^* = 0.$$

Onde *R* é o indicador predito por um modelo probit; *y* é a renda observada; e y^* é a renda predita. O conjunto *T* de variáveis independentes x_1 explica parte da renda predita, e o vetor de parâmetros τ associa o nível das variáveis independentes x_1 ao nível de renda predita. A parte não explicada da renda predita está contemplada no erro e_1, com distribuição aproximadamente normal. A variável *R* será igual a zero se a probabilidade de renda zero for significativa.

No segundo estágio, usa-se um modelo de efeito fixo para estimar os parâmetros β da equação a seguir, que relaciona as variáveis independentes ao nível de renda. Nela, a renda deve ser maior que zero. Assim:

Quando R=1,

$$E[y|R = 1, x_2] = E[y|y > 0, x_2] = x_2^T \beta + E[e_2|y > 0, x_2].$$

O modelo de efeito fixo permite que se estime um parâmetro fixo para cada Unidade da Federação (UF), de modo que a estimação carregue um efeito "de nível" dos determinantes estruturais dos estados onde a família reside. Isto é útil na medida em que as dinâmicas locais determinam sobremaneira o nível de renda das famílias, podendo, assim, isolar as diferenças estaduais do efeito das outras variáveis explicativas.

Assumindo-se que o erro é aleatório, e sua média é igual a zero $E[e_2|y > 0, x_2] = 0$, tem-se que:

$$E[y|y > 0, x_2] = x_2^T \beta$$

Portanto, para calcular o valor esperado da renda ($E(y)$), podemos decompô-la em:

$$E[y] = P(y > 0)\, E[y|y > 0] + P(y = 0)E(y|y = 0)$$

Sabendo que $E(y|y = 0) = 0$,

$$E[y] = P(y > 0)\, E[y|y > 0]$$

$$E[y] = P(y > 0)\, x_2^T \beta$$

A probabilidade de a renda ser maior que zero ($P(y > 0)$), dada pelo modelo probit, é a função distribuição acumulada de $y > 0$, que é $\Phi(x_1^T \tau)$. Então, a renda predita será:

$$E[y] = \Phi(x_1^T \tau)\, x_2^T \beta$$

No caso do modelo especificado pelo Ipea, o conjunto x_1 e x_2 de determinantes das duas partes do modelo são os mesmos. Adicionalmente, é estimada uma equação para cada UF, novamente dando conta dos determinantes locais e regionais da renda familiar (como um efeito fixo). Além disso, a renda estimada não é a média e sim a mediana, por ser mais simples de ser calculada e menos suscetível aos valores extremos. A tabela 2 traz os parâmetros estimados no Censo 2010 para o estado de São Paulo.

TABELA 2
Estado de São Paulo: estimativa dos parâmetros que associam os atributos das famílias à renda familiar *per capita* mediana

Grupo de atributos	Nome da variável	Descrição	Estimativa do parâmetro que associa o atributo à renda	*P*-valor da estimativa do parâmetro: se menor que 0,05, é significativa
Número de cômodos	ncomodos		0,11	<.0001
Número de cômodos que servem como dormitório	prdormitorios		-0,26	<.0001
Material predominante nas paredes externas do domicílio	matparede1	Alvenaria com revestimento	0,02	0.1360
	matparede2	Alvenaria sem revestimento	-0,11	<.0001
	matparede3	Madeira aparelhada	-0,08	<.0001
	matparede5	Taipa não revestida	-0,05	0.1258
	matparede6	Madeira aproveitada	-0,12	<.0001
Forma do abastecimento de água	canodagua1	Água encanada em pelo menos 1 cômodo	0,02	<.0001
	abast1	Rede geral de distribuição	0,05	<.0001
	abast2	Poço ou Nascente	0,04	0.0002
Forma de escoamento sanitário	esgotamento1	Rede coletora de esgoto ou pluvial	-0,06	0.0037
	esgotamento2	Fossa séptica	-0,14	<.0001
	esgotamento3	Fossa rudimentar	-0,19	<.0001
	esgotamento6	Outra forma	-0,23	<.0001
Destino do lixo	destlixo1	Coletado diretamente	-0,01	0.3711
	destlixo2	Coletado indiretamente	0,00	0.9860
	destlixo3	Queimado ou enterrado	-0,10	<.0001
	destlixo5	Jogado em rio, lago ou mar	-0,08	0.0012
Tipo de iluminação	iluminacao1	Elétrica com medidor próprio	0,18	<.0001
	iluminacao2	Elétrica com medidor comunitário	0,10	<.0001
	iluminacao3	Elétrica sem medidor	0,06	0.0021
Número de homens por faixa etária	h_0_4	Número de homens de 0 a 4 anos	-0,32	<.0001
	h_5_9	Número de homens de 5 a 9 anos	-0,22	<.0001
	h_10_14	Número de homens de 10 a 14 anos	-0,21	<.0001
	h_15_17	Número de homens de 15 a 17 anos	-0,26	<.0001
	h_18_24	Número de homens de 18 a 24 anos	-0,11	<.0001
	h_25_34	Número de homens de 25 a 34 anos	-0,13	<.0001
	h_35_44	Número de homens de 35 a 44 anos	-0,22	<.0001
	h_45_54	Número de homens de 45 a 54 anos	-0,18	<.0001
	h_55_64	Número de homens de 55 a 64 anos	-0,09	<.0001
	h_65_mais	Número de homens com 65 anos ou mais	-0,12	<.0001

(Continua)

(Continuação)

Grupo de atributos	Nome da variável	Descrição	Estimativa do parâmetro que associa o atributo à renda	*P*-valor da estimativa do parâmetro: se menor que 0,05, é significativa
Número de mulheres por faixa etária	m_0_4	Número de mulheres de 0 a 4 anos	-0,32	<.0001
	m_5_9	Número de mulheres de 5 a 9 anos	-0,22	<.0001
	m_10_14	Número de mulheres de 10 a 14 anos	-0,22	<.0001
	m_15_17	Número de mulheres de 15 a 17 anos	-0,27	<.0001
	m_18_24	Número de mulheres de 18 a 24 anos	-0,18	<.0001
	m_25_34	Número de mulheres de 25 a 34 anos	-0,23	<.0001
	m_35_44	Número de mulheres de 35 a 44 anos	-0,27	<.0001
	m_45_54	Número de mulheres de 45 a 54 anos	-0,28	<.0001
	m_55_64	Número de mulheres de 55 a 64 anos	-0,25	<.0001
	m_65_mais	Número de mulheres com 65 anos ou mais	-0,24	<.0001
Se o responsável familiar tem cônjuge	cônjuge		-0,08	<.0001
Proporção de pessoas na família por local de nascimento (migração)	nasc1	Nascido no município onde reside	-0,15	<.0001
	nasc2	Nascido em outro município da mesma UF	-0,03	0.0015
	nasc3	Nascido em outra UF	-0,17	<.0001
Número de pessoas alfabetizadas por faixa etária	alfab_18_24	Pessoas alfabetizadas de 18 a 24 anos	0,05	0.0020
	alfab_25_34	Pessoas alfabetizadas de 25 a 34 anos	0,08	<.0001
	alfab_35_44	Pessoas alfabetizadas de 35 a 44 anos	0,10	<.0001
	alfab_45_54	Pessoas alfabetizadas de 45 a 54 anos	0,12	<.0001
	alfab_55_64	Pessoas alfabetizadas de 55 a 64 anos	0,10	<.0001
	alfab_65_mais	Pessoas alfabetizadas de 65 anos ou mais	0,04	<.0001
Número de pessoas, por grau de instrução mais alto que alcançou e faixa etária	instr0_18_24	Alfabetização completa, incompleta ou jamais frequentada	-0,23	<.0001
	instr0_25_34		-0,26	<.0001
	instr0_35_44		-0,24	<.0001
	instr0_45_54		-0,22	<.0001
	instr0_55_64		-0,20	<.0001
	instr0_65_mais		-0,11	<.0001
	instr1_18_24	Fundamental incompleto	-0,30	<.0001
	instr1_25_34		-0,36	<.0001
	instr1_35_44		-0,31	<.0001
	instr1_45_54		-0,28	<.0001
	instr1_55_64		-0,21	<.0001
	instr1_65_mais		-0,11	<.0001
	instr2_18_24	Fundamental completo	-0,29	<.0001
	instr2_25_34		-0,32	<.0001
	instr2_35_44		-0,25	<.0001
	instr2_45_54		-0,21	<.0001
	instr2_55_64		-0,11	<.0001
	instr2_65_mais		0,02	0.0032

(Continua)

(Continuação)

Grupo de atributos	Nome da variável	Descrição	Estimativa do parâmetro que associa o atributo à renda	*P*-valor da estimativa do parâmetro: se menor que 0,05, é significativa
Número de pessoas, por grau de instrução mais alto que alcançou e faixa etária	instr3_18_24	Médio completo ou superior incompleto	-0,23	<.0001
	instr3_25_34		-0,13	<.0001
	instr3_35_44		-0,02	0.0037
	instr3_45_54		0,06	<.0001
	instr3_55_64		0,21	<.0001
	instr3_65_mais		0,33	<.0001
Número de pessoas, por rede de ensino e faixa etária	redeensino1_5_9	Crianças de 5 a 9 anos na rede pública	-0,21	<.0001
	redeensino1_10_14	Crianças de 10 a 14 anos na rede pública	-0,22	<.0001
	redeensino1_15_17	Jovens de 15 a 17 anos na rede pública	-0,12	<.0001
	redeensino2_5_9	Crianças de 5 a 9 anos na rede particular	0,09	<.0001
	redeensino2_10_14	Crianças de 10 a 14 anos na rede particular	0,08	<.0001
	redeensino2_15_17	Jovens de 15 a 17 anos na rede particular	0,16	<.0001
Número de pessoas ocupadas, por sexo e faixa etária	ocuph_15_17	Número de homens ocupados de 15 a 17 anos	0,24	<.0001
	ocuph_18_24	Número de homens ocupados de 18 a 24 anos	0,28	<.0001
	ocuph_25_34	Número de homens ocupados de 25 a 34 anos	0,40	<.0001
	ocuph_35_44	Número de homens ocupados de 35 a 44 anos	0,50	<.0001
	ocuph_45_54	Número de homens ocupados de 45 a 54 anos	0,43	<.0001
	ocuph_55_64	Número de homens ocupados de 55 a 64 anos	0,31	<.0001
	ocuph_65_mais	Número de homens ocupados de 65 anos ou mais	0,38	<.0001
	ocupm_15_17	Número de mulheres ocupadas de 15 a 17 anos	0,19	<.0001
	ocupm_18_24	Número de mulheres ocupadas de 18 a 24 anos	0,28	<.0001
	ocupm_25_34	Número de mulheres ocupadas de 25 a 34 anos	0,38	<.0001
	ocupm_35_44	Número de mulheres ocupadas de 35 a 44 anos	0,32	<.0001
	ocupm_45_54	Número de mulheres ocupadas de 45 a 54 anos	0,24	<.0001
	ocupm_55_64	Número de mulheres ocupadas de 55 a 64 anos	0,19	<.0001
	ocupm_65_mais	Número de mulheres ocupadas de 65 anos ou mais	0,22	<.0001

Fonte: Censo 2010.

Observa-se na tabela que, para o estado de São Paulo, apenas quatro dos parâmetros estimados, entre os 95 atributos, não foram significativos em sua associação com a renda familiar *per capita*, a um nível de confiança de 5%: o domicílio ter paredes de alvenaria com revestimento, taipa não revestida, coleta de lixo direta ou indireta. Todos os outros têm parâmetros estimados significativos. Percebe-se que o sinal da associação é intuitivo para algumas variáveis, sendo menor para outras. Observa-se, por exemplo, que quanto maior o número de cômodos, maior a renda associada, mas quanto maior a proporção de cômodos servindo de dormitório, menor a renda associada. Da mesma forma, quanto maior o número de pessoas, de qualquer idade ou sexo, menor a renda

familiar *per capita*, e quanto maior o número de pessoas ocupadas, de qualquer idade ou sexo, maior a renda associada. Ainda assim, o valor da associação, mesmo que negativo, se reduz à medida que níveis de escolaridade mais altos se verificam junto com uma maior incidência de idosos nos domicílios, indicando, possivelmente, recebimento de aposentadorias e pensões.

As estatísticas usuais de teste de robustez do modelo foram avaliadas e satisfizeram aos padrões usuais desse tipo de estimação. Pode-se resumir a robustez do modelo pela medida *R*2, que mostra a parcela da variância total associada às variáveis incluídas no modelo. Como foi estimado um conjunto de parâmetros para cada UF, temos 27 equações, todas com *R*2 entre 52,3%, para Santa Catarina, e 60,1%, para o Distrito Federal. Ou seja, as variáveis utilizadas explicam cerca de metade do nível de renda domiciliar *per capita*.

TABELA 3
Estatística *R* do modelo de duas partes que associa a renda familiar *per capita* aos atributos das famílias, dos domicílios e das pessoas
(Em %)

UF	*R*2 do preditor de renda	UF	*R*2 do preditor de renda
Santa Catarina	52,3	Pará	57,3
Rondônia	53,1	Roraima	57,3
Mato Grosso	53,9	Minas Gerais	57,8
Goiás	54,0	Ceará	57,9
Rio Grande do Sul	54,6	Amazonas	58,1
São Paulo	55,7	Pernambuco	58,3
Piauí	55,9	Amapá	58,5
Maranhão	56,2	Rio Grande do Norte	58,5
Rio de Janeiro	56,3	Paraíba	58,7
Mato Grosso do Sul	56,6	Alagoas	59,7
Paraná	56,9	Sergipe	60,2
Tocantins	57,2	Acre	61,2
Bahia	57,2	Distrito Federal	60,1
Espírito Santo	57,3		

Fonte: Censo 2010.

Sabe-se que outros tantos fatores determinam a renda *per capita* das famílias, fatores que não estão contemplados no modelo, seja por não haver captura destas informações (como a qualidade do ensino, a existência de ativos ou a produção de gêneros agrícolas e pecuários para consumo próprio), seja por serem fatores imponderáveis (a exemplo

de desastres naturais, mortes familiares ou volatilidade do mercado de trabalho). Neste sentido, no contexto de uma economia de mercado, o modelo, apesar de robusto, tem limitações bastante importantes.

As considerações anteriores são respaldadas pela literatura internacional. Kidd e Wylde (2011) fazem um apanhado desta literatura e replicam diferentes PMTs nos dados de quatro países (Bangladesh, Indonésia, Sri Lanka e Ruanda), revelando diversas falhas desses modelos. Assim, além de darem conta de explicar cerca de apenas 50%, 60% da variância da renda, os PMTs têm ainda pior *performance* quando se trata de avaliar os dois extremos da distribuição de renda (muito pobres ou muito ricos), como é o nosso caso aqui. Por isso, quando se compara a renda predita pelo modelo à renda efetivamente captada por uma pesquisa domiciliar ou pelo censo, o modelo categoriza como não elegíveis cerca de 44% a 55% dos efetivamente elegíveis, a depender do país considerado e para um nível de cobertura compatível com o PBF (*idem, ibidem*). Ou seja, os modelos excluem erroneamente muitas famílias. Como consequência, quando aplicarmos o modelo aqui exposto ao PBF, possivelmente ele apontará para famílias pobres classificando-as como não pobres. Mas em que medida isso ocorrerá? É isto que investigaremos.

Além disso, em nível individual, um modelo pode apontar famílias diferentes a depender das escolhas metodológicas, em grande medida arbitrárias, feitas ao longo da construção do modelo e da preparação dos dados. De acordo com Kidd e Wylde (2011), o tratamento de valores nulos ou *missings* e a escolha das variáveis podem mudar em até 30% as famílias escolhidas pelo modelo para serem classificadas como elegíveis ou inelegíveis. Esse nível de arbitrariedade é preocupante quando se trata de elegibilidade a benefícios essenciais à subsistência material das famílias.

Por fim, estimados os parâmetros que associam as variáveis independentes à renda do Censo 2010, na *quarta etapa*, aplicam-se estes parâmetros às variáveis do CadÚnico para encontrar a renda predita para as famílias do CadÚnico e suas características.

3 RAZÕES DA DISTÂNCIA ENTRE A PREDIÇÃO E A REALIDADE DAS FAMÍLIAS REGISTRADAS NO CADÚNICO

O benefício de superação da extrema pobreza (BSP), lançado pelo programa Brasil Carinhoso a partir de junho de 2012 como parte da estratégia do plano Brasil sem Miséria, varia com a distância da renda declarada e da linha de extrema pobreza, à época de R$ 70,00, depois atualizada para R$ 77,00, em julho de 2014. Assim, quanto menor a renda declarada, maior o benefício, podendo-se chegar a quase R$ 77,00 por pessoa, no caso de não haver qualquer outro rendimento no domicílio.

Assim, a partir do lançamento do BSP, existiria um incentivo à subdeclaração da renda, embora não haja um diagnóstico de que isto esteja de fato ocorrendo. Pelo contrário, o padrão de atualização das rendas no CadÚnico, que é monitorado pelo MDS, revela que desde 2011, entre as famílias que atualizaram sua renda familiar, 57% declararam renda mais alta do que havia registrado anteriormente. Este padrão não se alterou com o novo benefício.

Ademais, ao agruparmos as famílias segundo o mês de seu cadastramento, a entrada no CadÚnico daquelas com renda média abaixo da linha de extrema pobreza tem seguido um padrão de crescimento mensal moderado, em média de 4,9% após o lançamento do BSP. Este incremento é absolutamente compatível com o esforço de busca ativa e autosseleção das famílias após a divulgação ampla da meta de erradicação da miséria.

Com o objetivo de investigar a efetividade do preditor de renda em apontar efetivas subdeclarações de renda, em 2013 o MDS realizou uma ação conjunta com as gestões municipais para pesquisar uma amostra de famílias para as quais o preditor a princípio teria ainda maior probabilidade de acerto. Após o lançamento do BSP, foi escolhida uma amostra das famílias que se tornaram elegíveis ao benefício ou a valor maior, por terem reduzido a sua renda média cadastrada entre maio de 2012 e janeiro de 2013 (última base disponível à época).

O número de famílias que variou sua renda e acabou recebendo o benefício, ou mais dele, foi de 501 mil entre maio de 2012 e janeiro de 2013, para as quais aplicou-se o preditor de renda. Destas, foi escolhida uma amostra daquelas que tinham probabilidade maior que 90% de ter renda acima de R$ 70,00 (extrema pobreza à

época), e, no caso das famílias de renda zero, todas as famílias com probabilidade de 90% ou mais de ter renda diferente de zero, pelo modelo probit. Ou seja, escolheu-se investigar as famílias que, após o lançamento do BSP, supostamente teriam maior probabilidade de terem sucumbido à tentação de subdeclarar a renda, pois a sua situação "estrutural" estimada é melhor que a média dos pobres.

Esta intencionalidade na escolha das famílias a serem investigadas deriva justamente do contexto da aplicação do preditor. A intenção do MDS não era fazer pesquisa, mas encontrar um instrumento muito efetivo para apontar erros de inclusão e, a partir da mobilização dos gestores locais, já nesta primeira aplicação, tomar providências, como o cancelamento de benefícios. Nesse sentido, os achados deste exercício, se bem estão "contaminados" pelo seu propósito de servir como instrumento de verificação das rendas declaradas, também revelam a possibilidade ou não de acerto líquido do mecanismo, utilizando-se a capacidade institucional local e nacional existente.

Neste exercício, decidiu-se que todas as famílias de renda zero, com 90% de chance ou mais de possuírem renda acima de zero, teriam questionário aplicado pelo gestor municipal do CadÚnico. Para as famílias de renda média maior que zero, só aquelas que caíram na amostra – totalizando 5% do total de famílias com mais de 90% de probabilidade de terem subdeclaração – participariam do exercício.

O plano amostral utilizado nesta investigação aplicou uma amostragem probabilística e estratificada por UF e região metropolitana (RM). Para o cálculo do tamanho da amostra usou-se a variabilidade máxima de 50%, tendo em vista que a variância populacional é desconhecida e que o nível de confiança estabelecido foi de 95%.

Assim, foram encaminhadas para os municípios listagens de famílias para atualização cadastral com necessidade de preenchimento de um questionário suplementar (apêndice). Estas listagens continham 3.224 famílias com renda familiar *per capita* maior que zero (5% de 64.430), e 7.936 famílias com renda igual a zero, cuja atualização deveria ser acompanhada das respostas ao questionário suplementar. A distribuição destas famílias por UF é apresentada a seguir.

Ao todo, foram respondidos 3.839 questionários suplementares, dos quais 1.085 correspondiam à amostra do público com renda positiva, e 2.754 correspondiam a famílias com renda zero. No geral, 34% dos questionários foram respondidos; 33,7% da amostra de famílias com renda positiva tiveram seus questionários respondidos; e 34,7% das famílias do público-alvo possuíam renda zero. Antes de analisar as respostas, foi importante estabelecer se estas poderiam ser tomadas como representação do universo de questionários. Ou seja, é preciso ter alguma mensuração do viés de resposta.

Como as realidades em cada UF no Brasil diferem muito entre si, tanto o preditor de renda como o plano amostral em seus cálculos consideraram estas diferenças. Portanto, primeiramente avaliou-se a hipótese de as proporções dos respondentes serem equivalentes às proporções da população por UF. Se estas fossem equivalentes, a possibilidade de haver viés de resposta ficaria reduzida, devido às diferentes realidades existentes no país, pois todas as UFs estariam representadas na mesma proporção da população pesquisada.

Observa-se que, para vinte das 27 UFs, a proporção de questionários respondidos foi estatisticamente igual à proporção de questionários na amostra. Para outras três unidades, houve sub-representação (São Paulo, Alagoas e Amazonas), e para quatro unidades houve sobrerrepresentação nas respostas (Ceará, Maranhão, Piauí e Sergipe) em relação à distribuição da amostra. Conclui-se que houve pouco viés territorial nas respostas.

Resolveu-se também testar as proporções das variáveis estruturais da população e dos respondentes para anular a possibilidade de haver viés devido a particularidades dos respondentes. Se os responsáveis das famílias respondentes fossem, em média, mais estudados que os responsáveis familiares do público-alvo, talvez houvesse um viés nas respostas que indicasse, equivocadamente, uma menor vulnerabilidade das famílias apontadas pelo preditor. As variáveis estruturais analisadas foram: anos de estudo do responsável familiar; domicílio rural ou urbano; sexo do responsável e renda *per capita* da família. Seguem os resultados para a escolaridade.

Observa-se que a diferença entre o número de anos de estudo dos respondentes e dos não respondentes é mínima, tanto entre as famílias com renda zero quanto entre as com renda positiva. A medida do viés, também em anos de estudo, aponta para uma

diferença ponderada, pela não resposta de 0,02 ano de estudo entre os respondentes e os não respondentes no caso das famílias com renda zero, ou seja, cerca de uma semana de estudo de diferença. Assim, em todas as variáveis estruturais analisadas, não houve viés significativo na resposta do questionário suplementar, podendo as repostas serem generalizadas para o universo: famílias que teriam respondido ao incentivo de reduzir a renda por ocasião do lançamento do BSP o fizeram, a despeito de serem estruturalmente não pobres (ou seja, há chance de mais de 90% de elas terem subdeclarado renda).

TABELA 4
Viés da média de anos de estudo do responsável familiar

Grupo	Média de anos de estudo		Desvio-padrão de anos de estudo		Viés da média de anos de estudo
	Respondentes	Não respondentes	Respondentes	Não respondentes	
Renda zero	5,45	5,42	3,79	3,77	0,02
Renda maior que zero	6,47	6,70	3,72	3,64	0,15

Fonte: Amostra e questionário suplementar do preditor.
Elaboração das autoras.

No que tange ao questionário suplementar, foram elaboradas nove perguntas para que o gestor municipal, ao visitar a família apontada, pudesse identificar dois aspectos.

1) Falhas nos dados de entrada do preditor: tentou-se captar o grau de precariedade do domicílio e do ambiente de moradia da família, o que refina a precariedade mais genérica captada pelas variáveis de domicílio do CadÚnico. Assim, por mais que a família possa viver em domicílio com alvenaria e piso de cimento, com estrutura urbana de coleta de lixo, abastecimento de água etc., o domicílio em si pode estar em alto grau de deterioração, com fiação elétrica exposta, buracos e goteiras, ter água apenas em um cômodo etc.
2) Erro no próprio apontamento do modelo por situação de volatilidade de renda ou outras vulnerabilidades não observáveis, imprevisíveis ou não capturadas: tentou-se investigar principalmente as causas dos baixos rendimentos declarados ao CadÚnico, como inatividade, tipo de inserção no mercado de trabalho, omissão de renda ou membro etc. Assim, a intenção foi apontar fatores não levados em conta pelo preditor que possam explicar a situação de extrema pobreza, ou que possam ratificar o achado do preditor. Ou seja, de que a família estaria subdeclarando a renda.

Do total de famílias que responderam ao questionário, 56% tinham moradia em bom estado, 34% em estado precário e apenas 9% em estado muito precário. Assim, considerando-se o perfil extremamente pobre da população cadastrada, as famílias apontadas pelo preditor, além de terem melhores condições de moradia nos quesitos

do CadÚnico, também têm, em média, moradias em bom estado. Conclui-se que a dependência do modelo em relação aos quesitos genéricos do Cadúnico não representa um grande problema.

Claramente, o padrão de inserção dessa população no mercado de trabalho é um fator determinante para sua situação de vulnerabilidade. No total, 62% dos responsáveis familiares, em sua grande maioria mulheres, não estavam trabalhando no momento da pesquisa (questão 2, apêndice). Os motivos mais frequentes para não estarem trabalhando (questões 3 e 4, apêndice) foram: *i)* por não encontrarem trabalho, apesar da procura (29%); *ii)* por doença ou outra condição que lhes impossibilite o trabalho (26%) *iii)* ou por terem que prover cuidados a crianças e outros dependentes (23%).

De forma geral, essa situação é igual à média encontrada no CadÚnico à época de aplicação do questionário (agosto de 2013). Para as responsáveis familiares acima de 18 anos, em média 62% não trabalharam na semana anterior ao registro no CadÚnico. A seguir as respostas do questionário suplementar, para as famílias com renda acima de zero e aquelas com renda igual a zero. Como se vê, não houve diferenças significativas nas respostas destes dois públicos.[10]

Para além do desemprego, a resposta à terceira variável (motivo de o responsável familiar não estar trabalhando) revela, a uma só tempo, o lugar socialmente concedido à mulher, circunscrito ao cuidado dos dependentes em detrimento da participação no mercado de trabalho, bem como a situação precária de saúde da mulher que vive em extrema pobreza. Tal situação tem fundamentos na carência material e no desalento, que muitas vezes geram distúrbios psíquicos, além de físicos, os quais reforçam a delimitação de sua vida social ao espaço doméstico.

10. Como a investigação foi inédita, aplicada a poucas famílias e movida pelo esforço de realizar uma averiguação cadastral com possíveis repercussões em benefícios, optou-se por um questionário simples, sucinto, mas com duas questões abertas (questões 4 e 7, apêndice), para captar respostas não previstas, justamente pela novidade do tema. Assim, o MDS fez uma extensa codificação das respostas às perguntas abertas para, quando pertinente, reagrupá-las ao domínio preestabelecido ou às novas categorias das respostas às questões 3 e 6, respectivamente.

TABELA 5
Respostas ao questionário suplementar[1]

Agrupamento de respostas ao questionário suplementar		Famílias com renda igual a zero		Famílias com renda maior que zero		Total de respostas	
		Número de famílias	%	Número de famílias	%	Número de famílias	%
Variável 1 – condições da moradia (estrutura e acabamentos).	Em bom estado	1.518	55	642	59	2.160	56
	Precárias	975	35	339	31	1.314	34
	Muito precárias	258	9	101	9	359	9
	Não informou	3	0	3	0	6	0
Variável 2 – responsável familiar está trabalhando no momento?	Não	1.784	65	603	56	2.387	62
	Sim	951	35	479	44	1.430	37
	Não informou	19	1	3	0	22	1
Variável 3 – motivo de o responsável familiar não estar trabalhando.	Não consegue emprego	514	19	184	17	698	18
	Doença ou condição que impossibilitou trabalho	468	17	145	13	613	16
	Tem que prover cuidados	380	14	161	15	541	14
	Aposentado, pensionista ou benefício de prestação continuada	190	7	31	3	221	6
	Inatividade – nunca trabalhou, não procura	70	3	22	2	92	2
	Atividade agrícola	35	1	9	1	44	1
	Cônjuge trabalha ou tem outro rendimento	12	0	9	1	21	1
	Precisa de qualificação ou experiência	3	0	2	0	5	0
	Outro motivo	50	2	22	2	72	2
	Não se aplica ou não informou	1.032	37	500	46	1.532	40
Variável 5 – atividade ou setor em que o responsável familiar habitualmente trabalha.	Serviço doméstico	741	27	400	37	1.141	30
	Agricultura, pecuária, pesca, extrativismo	599	22	111	10	710	18
	Outros serviços	323	12	133	12	456	12
	Comércio	214	8	150	14	364	9
	Serviço público	107	4	43	4	150	4
	Indústria	74	3	23	2	97	3
	Cuidador	51	2	19	2	70	2
	Construção civil	43	2	13	1	56	1
	Serviços de beleza	42	2	27	2	69	2
	Transporte e armazenamento	10	0	6	1	16	0
	Nunca trabalhou	337	12	109	10	446	12
	Não informou	213	8	51	5	264	7
Variável 6 – motivo para a renda familiar estar tão baixa na última atualização.	Não encontrou emprego ou perdeu emprego	1.285	47	598	55	1.883	49
	Omissão de renda ou membro	702	25	278	26	980	26
	Doença ou condição que impossibilitou trabalho	384	14	123	11	507	13
	Exclusão de membro com renda	118	4	42	4	160	4
	Atividade rural	77	3	21	2	98	3
	Morte ou catástrofe	72	3	13	1	85	2
	Não sabe ou erro no cadastro	60	2	22	2	82	2
	Tem que prover cuidados	16	1	9	1	25	1
	Outro motivo	208	8	74	7	282	7
	Não se aplica ou não informou	23	1	18	2	41	1

(Continua)

(Continuação)

Agrupamento de respostas ao questionário suplementar		Famílias com renda igual a zero		Famílias com renda maior que zero		Total de respostas	
		Número de famílias	%	Número de famílias	%	Número de famílias	%
Variável 8 – renda familiar é instável?	Sim	1.592	58	682	63	2.274	59
	Não	933	34	344	32	1.277	33
	Não informou	229	8	59	5	288	8
Variável 9 – família é vulnerável?	Sim	1.466	53	535	49	2.001	52
	Não	1.050	38	490	45	1.540	40
	Não informou	238	9	60	6	298	8
Total de respostas		**2.754**	**100**	**1.085**	**100**	**3.839**	**100**
Total de famílias na amostra		**7.936**		**3.224**		**11.160**	
Famílias respondentes (%)		**35**		**34**		**34**	

Fonte: Questionário suplementar do preditor.
Elaboração das autoras.
Nota: [1] Ao aplicar-se o peso de cada família na amostra, a distribuição das respostas muda apenas na segunda casa decimal, por isso optou-se por trazer aqui apenas as respostas brutas.

No que tange à atividade ou ao setor em que a responsável familiar atua, a maior parte trabalha ou trabalhou em emprego doméstico. Para as famílias de renda positiva, 37% costumam trabalhar no emprego doméstico, e para as com renda igual a zero, apenas 27% o fazem. Se agregarmos a atividade de cuidador, chega-se a 39% e 29%, respectivamente. A diferença é que, para as famílias com renda igual a zero, o peso da atividade de agricultura, pecuária, pesca e extrativismo chega a 22%, enquanto para as famílias com renda positiva, representa apenas 10%. Faz sentido, na medida em que tal atividade tende a impor maior volatilidade de renda por sua sazonalidade, levando à situação de renda zero. Além disso, a atividade permite a subsistência sem renda monetária.

Comparando-se com a média das responsáveis familiares acima de 18 anos no CadÚnico, percebe-se uma diferença bastante importante, pois apenas 6% das que trabalhavam eram identificadas como empregadas domésticas. Já nas famílias que responderam ao questionário suplementar, em média 29% das que trabalhavam no momento da pesquisa exerciam atividade doméstica. Novamente, a precariedade do vínculo de emprego doméstico, função notadamente marcada pela informalidade, pelo desrespeito profissional, pela instabilidade e pelos baixos rendimentos, tem fundamentos nas relações de gênero de nossa sociedade.

Quanto ao trabalho na agricultura, na criação de animais, na pesca, na coleta ou no extrativismo, a pergunta 8.03[11] do CadÚnico revela que 28% dos responsáveis pela unidade familiar acima de 18 anos têm essa atividade, contra 18% destes responsáveis pesquisados no questionário suplementar, que à época trabalhavam, ou seja, aquém da média cadastral. Nesse sentido, a concentração em tal atividade não parece ser uma explicação especialmente representativa do grupo apontado pelo preditor para a volatilidade de seus rendimentos, ou para o baixo nível destes.

A questão 6, respondida pelo entrevistador a partir da interação com a família, é sobre o motivo da baixa renda no momento da última atualização cadastral. As respostas à esta pergunta trazem novos elementos para a análise das famílias apontadas pelo preditor de renda. Apesar da dificuldade de inserção e rotatividade no mercado de trabalho, confirmadas mais uma vez pela predominância das respostas de "não encontrou emprego ou perdeu o emprego" (43% do total de respostas) e "doença ou condição que impossibilitou trabalhar" (12% do total de respostas), 24% das famílias têm algum indício de subdeclaração de renda familiar pela omissão de membros da família, de rendimentos ou benefícios.

Por fim, as questões 8 e 9 tentam captar de forma sintética a percepção do entrevistador sobre a volatilidade de renda da família e a sua vulnerabilidade. Este último é um conceito amplamente trabalhado pela assistência social e que se refere a uma avaliação mais ampla que a situação de renda, inclusive para designar situações de potencial violação de direitos e desestruturação de vínculos familiares e sociais.

Em média, 64% das famílias entrevistadas que responderam à questão 8 (apêndice) tinham renda familiar instável, e 56% delas eram vulneráveis. Entretanto, somente 43% eram vulneráveis e tinham renda instável, e 22% não eram nem uma coisa nem outra. Assim, se fosse possível estabelecer um critério de erro e acerto do preditor, levando-se em consideração uma avaliação mais ampla que a ausência ou a volatilidade de renda, pode-se dizer que em 43% dos casos o procedimento apontou, equivocadamente, uma família que é vulnerável e tem renda instável (erro), e em 22% apontou para uma família que não tem renda instável e nem é vulnerável (acerto). Isso

11. A pergunta 8.03 do caderno principal do CadÚnico é: "Esse trabalho principal que (nome) exerceu foi na agricultura, criação de animais, pesca ou coleta (extração vegetal)?" As respostas possíveis são "sim" ou "não".

já dentro do subgrupo com 90% ou mais de chance de ter renda acima do recorte de extrema pobreza.

Quando se cruzam essas respostas é possível perceber a associação entre os conceitos de vulnerabilidade e instabilidade de renda. A tabela 6 mostra que, para 65% das famílias que responderam às questões 8 e 9, houve coincidência entre instabilidade de renda e vulnerabilidade (soma de 22% de famílias não instáveis e não vulneráveis com 43% de famílias instáveis e vulneráveis). Para 14% delas, apesar da vulnerabilidade em outros quesitos, no que diz respeito à renda, não havia instabilidade. No entanto, um desafio para a interpretação do conceito de vulnerabilidade refere-se aos 21% de famílias que, apesar da instabilidade de renda, não são vulneráveis. Neste sentido, vulnerabilidade não é apenas uma expansão do conceito estrito de potencial carência material, mas tem outras apropriações possíveis.

TABELA 6
Famílias que responderam às questões 8 e 9 do questionário suplementar (apêndice), segundo instabilidade de renda e vulnerabilidade
6A – Em números absolutos

		Variável 9 – vulnerabilidade		**Total**
		Não	Sim	
Variável 8 – instabilidade da renda	Não	793	482	**1.275**
	Sim	747	1.517	**2.264**
Total		**1.540**	**1.999**	**3.539**

6B – Em porcentagem

		Variável 9 – vulnerabilidade		**Total**
		Não	Sim	
Variável 8 – instabilidade da renda	Não	22	14	**36**
	Sim	21	43	**64**
Total		**44**	**56**	**100**

Fonte: Questionário suplementar do preditor.
Elaboração das autoras.

Como os critérios de "vulnerabilidade" e "instabilidade de renda" são pouco objetivos, torna-se interessante modular as respostas com critérios mais claros, como o estado da moradia ou a percepção, pelo entrevistador, de omissão de renda. Neste sentido, apesar de 22% dos respondentes não terem renda instável, nem serem vulneráveis, são apenas 19% que, além disso, têm moradia em bom estado. Em apenas

10% dos respondentes identificou-se o caso mais grave de omissão de renda conjugada à não vulnerabilidade e à estabilidade de renda (acerto do preditor).

TABELA 7
Respostas cruzadas ao questionário suplementar (apêndice), segundo instabilidade de renda e vulnerabilidade

7A – Respostas às questões 1, 8 e 9

Variável 1 – moradia	Variável 9 – vulnerabilidade	Variável 8 – instabilidade da renda	Número de respostas	%
Em bom estado	Não	Não	660	19
		Sim	616	17
	Sim	Não	159	4
		Sim	551	16
Precárias	Não	Não	121	3
		Sim	122	3
	Sim	Não	251	7
		Sim	726	21
Muito precárias	Não	Não	11	0
		Sim	7	0
	Sim	Não	72	2
		Sim	238	7
Total			**3.534**	**100**

7B – Respostas às questões 6, 7, 8 e 9

Variáveis 6 e 7	Variável 9 – vulnerabilidade	Variável 8 – instabilidade da renda	Número de respostas	%
Houve omissão de renda ou membro	Não	Não	353	10
		Sim	237	7
	Sim	Não	97	3
		Sim	212	6
	Subtotal		899	25
Não houve omissão de renda ou membro	Não	Não	440	12
		Sim	510	14
	Sim	Não	385	11
		Sim	1.305	37
	Subtotal		2.640	75
Total			**3.539**	**100**

Fonte: Formulário suplementar do preditor de renda.
Elaboração das autoras.

De maneira geral, as respostas ao questionário suplementar comprovam a hipótese de que o preditor tem um erro grande devido aos fenômenos não incluídos no modelo, os quais se referem aos choques sofridos pelas famílias: desemprego, doença etc. Quando trabalhando, estas estão inseridas em atividades altamente informais,

instáveis e duras fisicamente, como o emprego doméstico e a agricultura, reforçando a baixa proteção ante o desemprego, a doença e os baixos rendimentos. Estes choques levam as famílias a caírem abaixo da linha da pobreza, apesar de suas condições de moradia não serem as piores possíveis.

Assim, apesar de o preditor identificar algumas famílias com alta probabilidade de não serem extremamente pobres, em apenas 22% dos casos individuais apontados há elementos que poderiam confirmar alguma subdeclaração, pois não se trata de famílias nem vulneráveis nem com instabilidade de renda. Visto de outro ângulo, cerca de 78% dos apontamentos de subdeclaração do preditor são resultados de outros fenômenos associados a choques no âmbito do mercado de trabalho e da vida familiar, da vinculação precária com o mercado de trabalho e da posição das mulheres, em especial daquelas em situação de pobreza na vida social. Assim, são variáveis não observáveis ou não capturadas pelo questionário do CadÚnico e pelo modelo.

4 CONCLUSÃO

A opção por aplacar a ausência de renda está absolutamente condizente com o papel complementar do Bolsa Família à matriz de garantia de renda brasileira por meio de um pilar não contributivo, o qual tem por objetivo dar segurança monetária mínima a todas as famílias brasileiras que dele precisarem. Não se pretende, com um instrumento monetário, aplacar os diversos determinantes da pobreza, ainda que as transferências tenham suas externalidades e, se acompanhadas pelo reforço dos direitos à saúde, à educação, à moradia, entre outros, podem ter efeitos mais decisivos nesta direção. Para aplacar os determinantes da pobreza, são necessárias políticas específicas que tratem das variadas ausências e das carências das famílias em situação de pobreza, ou seja, tratar a baixa escolaridade com oferta de escolas, professores mais engajados e bem-remunerados, melhor ambiente escolar etc.

Se o objetivo da transferência é promover a complementação de renda quando esta, por algum motivo, for insuficiente, a aplicação de um preditor de renda, cujo erro é bastante significativo, aumentará o erro de exclusão, ao barrar o atendimento de famílias com situação estrutural supostamente não tão ruim, mas sofrendo de severa insuficiência de renda. Entretanto, o procedimento pode vir a beneficiar famílias em

situação estrutural muito precária, mas com renda acima da linha de pobreza, levando a erros de inclusão.

Nesse sentido, como já apontava Sen (1995), talvez a característica mais óbvia de um mecanismo de seleção de beneficiário, mais importante, não obstante frequentemente esquecida pelos especialistas, é que a medida da privação exprima de forma mais exata possível a carência que se quer suprir. Se a proteção em tela é uma transferência de renda, então, sem dúvida, a melhor medida para expressar a carência desta proteção é a ausência de renda. Mas tornar esta medida o mais exata possível requer muito mais que uma simples pergunta de renda ou um modelo estatístico, demanda treinamento constante de equipes, exige normativos e padronizações, atualização cadastral, constantes verificações, boas técnicas de formulário, excelentes sistemas, financiamento de estruturas e equipes, fortalecimento burocrático das gestões locais, entre tantos determinantes, inclusive políticos, muito mais poderosos que a beleza estatística dos modelos.

Assim, se se pretendesse, por meio do Bolsa Família, suprir as múltiplas carências que tornam uma família "vulnerável" medida por um preditor ou um indicador multidimensional, então faria sentido selecionar e monitorar o resultado da política por esta mesma medida. Mas como sabemos, o objetivo é, antes de tudo, aplacar uma carência monetária básica, que impede o acesso adequado aos alimentos, transporte, vestimenta etc. Se esta é a carência que se quer suprir, a melhor medida é a renda.

Qual a armadilha de medir renda para aplacar a ausência de renda com renda? É a própria volatilidade da renda, que obriga a um esforço enorme de atualização cadastral por parte do estado e das famílias. Dando, *a posteriori*, racionalidade ao movimento de gestão do CadÚnico e ao PBF, esta fez esforços, nos anos recentes, para fugir desta armadilha. A regra de atualização a cada dois anos, a estimativa de cobertura levando-se em conta a volatilidade da renda, a captura da renda média dos últimos doze meses no formulário do CadÚnico, a regra de permanência de quem ultrapassa a linha de pobreza, mas permanece abaixo de meio salário mínimo *per capita*, bem como a regra do retorno garantido para aqueles que se desligam voluntariamente do programa são todos mecanismos que dialogam com a necessidade de se medir e ao mesmo tempo mitigar os efeitos gerenciais da volatilidade de renda. Ou seja, de sair da armadilha de ter que gerenciar a volatilidade mensal da renda de milhões de brasileiros.

A boa notícia é que os desenvolvimentos recentes do CadÚnico e do Bolsa Família imprimiram um viés protetivo à transferência, pois garantem a continuidade da transferência àquelas famílias que recentemente conseguiram superar a linha de pobreza, ou possibilitam a concessão para as que, na média do último ano, ficaram abaixo da linha de pobreza, mesmo estando acima no momento da inscrição no cadastro. Foram, sem dúvida, decisões salutares à gestão dessa política, bem como à vida das famílias pobres no Brasil.

Este trabalho pretendeu registrar a aplicação modesta de um preditor de renda (PMT) como instrumento de verificação *a posteriori* da renda declarada pelas famílias ao CadÚnico, no contexto de lançamento de uma nova regra de benefício do Bolsa Família. Assim, procurou-se investigar a possibilidade de uso deste instrumento para detectar omissão de renda por parte das famílias para além dos batimentos de registros administrativos ocorridos desde 2005. No entanto, enquanto os batimentos permitem cancelar até 60% das famílias beneficiárias com indício de inconsistência cadastral, o instrumento aqui investigado o faria com, no máximo, 25% das famílias consideradas pelo preditor com chance de mais de 90% de não serem extremamente pobres e que efetivamente omitiram renda. Dados os custos administrativos de capacitação das equipes, a realização de novas visitas domiciliares, entre outros fatores, no estado atual, este não parece ser um instrumento que tenha uma boa relação custo-benefício em sua aplicação.

No sentido de contribuir para a resposta de algumas das perguntas que fizemos no início, o exercício de validação do preditor de renda por meio da aplicação de um questionário suplementar revelou, de um lado, que este não é um instrumento com acurácia suficiente para ser utilizado na verificação das rendas declaradas ao CadÚnico. Por outro lado, a investigação mostrou quais os maiores determinantes dos possíveis erros na captura da renda das famílias. Desatualização cadastral? Sim. Desemprego? Sim. Vinculação precária e intermitente das mulheres no mercado de trabalho? Sim. Omissão por parte das famílias? Não.

REFERÊNCIAS

COADY, D.; GROSH, M.; HODDINOTT, J. **Targeting of transfers in developing countries**: review of lessons and experience. Washington: World Bank, 2004.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **A década inclusiva (2001-2011)**: desigualdade, pobreza e políticas de renda. Brasília: Ipea, set. 2012. (Comunicados do Ipea, n. 155).

_____. Duas décadas de desigualdade e pobreza medida pela Pnad/IBGE. Brasília: Ipea, out. 2013. (Comunicados do Ipea, n. 159).

KIDD, S. Rethinking 'Targeting' in international development. **Pathways' Perspectives on social policy and international development**, Oxfordshire, n. 11, out. 2013.

KIDD, S.; WYLDE, E. **Targeting the poorest**: an assessment of the proxy means test methodology. Canberra: Australian Aid, set. 2011.

OSÓRIO, R.; SOARES, S.; SOUZA, P. **Erradicar a pobreza extrema**: um objetivo ao alcance do Brasil. Brasília: Ipea, maio 2011. (Texto para Discussão, n. 1619).

SEN, A. The political economy of targeting. *In*: VAN DE WALLE, D.; NEAD, K. (Org.). **Public spending and the poor**: theory and evidence. Washington: World Bank, 1995.

SOARES, S. **O ritmo de queda na desigualdade no Brasil é adequado?** Evidências do contexto histórico e internacional. Brasília: Ipea, maio 2008. (Texto para Discussão, n. 1339).

APÊNDICE

QUADRO A.1
Questionário aplicado no âmbito da averiguação (2013)

Questão 1	Como você avalia as condições de moradia desta família (considere as condições do piso, das paredes, do teto e das instalações elétricas e hidráulicas)? a) Muito precárias. b) Precárias. c) Em bom estado.
Questão 2	O responsável pela unidade familiar (RF) está trabalhando no momento? a) Sim. b) Não.
Questão 3	Se o RF não está trabalhando, qual o motivo? a) Doença, privação de liberdade ou outra condição que impossibilite o trabalho. b) Está procurando, mas não consegue trabalho. c) Não tem com quem deixar os filhos ou outra pessoa da família que requeira cuidados. d) É aposentado, pensionista ou recebe BPC. e) Outro motivo. f) Não se aplica – assinale somente se marcou "sim" na questão 2.
Questão 4	Caso tenha assinalado "outro motivo" na questão anterior, descreva por qual razão o representante familiar não está trabalhando no momento.
Questão 5	Com que tipo de atividade ou setor de trabalho o RF trabalha (mesmo que não esteja trabalhando no momento)? a) Comércio (vendedores, ambulantes, etc.). b) Serviços domésticos (empregado doméstico, passadeira, jardineiro, limpeza etc.). c) Cuidadores de crianças, jovens, adultos ou idosos. d) Atividades de tratamento de beleza (manicure, cabelereiro, esteticista etc.). e) Construção civil (pedreiro, ajudante, mestre de obra etc.). f) Agricultura, criação de animais, pesca ou extrativismo. g) Transporte, armazenagem ou correio. h) Indústria. i) Serviço público ou militar. j) Outros serviços. k) O representante familiar nunca trabalhou.
Questão 6	Por que a família estava com renda tão baixa na atualização cadastral realizada em 2012? a) Algum componente da família não conseguiu trabalho ou não encontrou emprego naquele período. b) Algum componente da família que contribuía com a renda familiar perdeu o emprego. c) Algum componente da família que contribuía com renda foi excluído do cadastro da família. d) Algum componente da família ficou impossibilitado de trabalhar por condição de saúde, privação de liberdade ou outra condição que impossibilite o trabalho. e) Morte do RF ou de algum componente da família. f) A renda do RF ou de algum componente da família não foi informada. g) Não foram informadas pessoas que deveriam ter sido cadastradas como componentes da família. h) Foram informadas pessoas que não faziam parte da real composição familiar. i) Situações como catástrofes, calamidades, desastres naturais. j) Outra situação.
Questão 7	Caso tenha assinalado "outra situação" na questão anterior, descreva por qual outra razão a renda familiar estava tão baixa na atualização cadastral realizada em 2012.
Questão 8	A renda desta família é instável (considere o período entre a data da última atualização até hoje)? a) Sim. b) Não.
Questão 9	Esta família é vulnerável? a) Sim. b) Não.

Fonte: Ministério do Desenvolvimento Social.
Elaboração das autoras.

Ipea – Instituto de Pesquisa Econômica Aplicada

EDITORIAL

Coordenação
Cláudio Passos de Oliveira

Supervisão
Andrea Bossle de Abreu

Revisão
Carlos Eduardo Gonçalves de Melo
Elaine Oliveira Couto
Laura Vianna Vasconcellos
Luciana Nogueira Duarte
Mariana Silva de Lima
Vivian Barros Volotão Santos
Bianca Ramos Fonseca de Sousa (estagiária)
Thais da Conceição Santos Alves (estagiária)

Editoração
Aeromilson Mesquita
Aline Cristine Torres da Silva Martins
Carlos Henrique Santos Vianna
Glaucia Soares Nascimento (estagiária)
Vânia Guimarães Maciel (estagiária)

Capa
Luís Cláudio Cardoso da Silva

Projeto Gráfico
Renato Rodrigues Bueno

The manuscripts in languages other than Portuguese published herein have not been proofread.

Livraria Ipea

SBS – Quadra 1 - Bloco J - Ed. BNDES, Térreo.
70076-900 – Brasília – DF
Fone: (61) 2026-5336

Correio eletrônico: livraria@ipea.gov.br

Composto em adobe garamond pro 12/16 (texto)
Frutiger 67 bold condensed (títulos, gráficos e tabelas)
Impresso em offset 90g/m^2 (miolo)
Cartão supremo 250g/m^2 (capa)
Rio de Janeiro-RJ

Missão do Ipea
Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.

ISSN 1415-4765

ipea Instituto de Pesquisa Econômica Aplicada

MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO

BRASIL GOVERNO FEDERAL