

## **Programa de transferência de renda, perfil socioeconômico e probabilidades de classificação: avaliando o critério de seleção dos beneficiários do Bolsa família nas regiões Sul e Sudeste<sup>1</sup>**

Alessandra de Oliveira Teixeira<sup>♦</sup>  
 Admir Antônio Betarelli Junior<sup>♦♦</sup>  
 Marcel de Toledo Vieira<sup>♦♦</sup>  
 Ana Paula Kern<sup>°</sup>  
 Weslem Rodrigues Faria<sup>♦♦</sup>

### **RESUMO**

Este artigo contribui para os recentes debates econômicos e sociais sobre os critérios de seleção do Programa Bolsa Família (PBF). A análise do critério desse programa de transferência de renda é feita mediante a aplicação de um conjunto de técnicas estatísticas sobre o perfil socioeconômico dos indivíduos entrevistados no banco de dados da AIBF II, localizados nas regiões Sul e Sudeste em 2009. Com a finalidade de avaliar um padrão característico dos indivíduos entrevistados, primeiramente adotou-se as técnicas de Análise de Correspondência Múltipla (AC) e Análise de Componentes Principais (ACP), que também contribuíram para a construção de indicadores socioeconômicos. Em seguida, aplicou-se a Análise Discriminante (AD) para averiguar as probabilidades de classificação correta e incorreta dos indivíduos beneficiários e não beneficiários – registrados no Cadastro Único, de acordo com as respostas às questões socioeconômicas dos entrevistados. Os resultados indicam que o critério de inclusão dos indivíduos ao Bolsa Família pela renda *per capita* apresenta uma probabilidade de erro relativamente menor, mesmo ao considerar outras características de aspecto multidimensional da pobreza. Cerca de 25,88% dos beneficiários entrevistados não deveria receber o benefício do PBF. Por outro lado, dos 8.647 entrevistados que não são beneficiários, 40,45% deveria receber o benefício segundo sua classificação na análise realizada. Tais resultados conclusivos indicam que mais pessoas poderiam receber o benefício se fossem consideradas as demais características dos indivíduos.

**Palavras-chave:** Bolsa Família. Critério de seleção. Regiões Sudeste e Sul. Análise Multivariada.

### **ABSTRACT**

The paper contributes to the recent economic and social debates on the selection criteria of the Bolsa Família Program (PBF). The analysis of the criterion of this income transfer program is made through the application of a set of statistical techniques on the socioeconomic profile of the individuals interviewed in the AIBF II database, located in the Brazilian South and Southeast regions in 2009. In order to evaluate a characteristic pattern of the individuals interviewed, we first adopted the techniques of Multiple Correspondence Analysis (CA) and Principal Components Analysis (PCA), which also contributed to the construction of socioeconomic indicators. Then, Discriminant Analysis (AD) was applied to ascertain the probabilities of correct and incorrect classification of the beneficiary and non-beneficiary individuals - registered in the Cadastro Único, according to the answers to the socioeconomic questions of the interviewees. The results indicate that the criterion of inclusion of the individuals in the Bolsa Família by per capita income presents a relatively lower probability of error, even when considering other characteristics of multidimensional aspect of poverty. About 25.88% of the beneficiaries interviewed should not receive the benefit of the PBF. On the other hand, of the 8,647 interviewees who are not beneficiaries, 40.45% should receive the benefit according to their classification in the analysis performed. Such conclusive results indicate that more people could receive the benefit if the other characteristics of the individuals were considered.

**Key words:** Bolsa Família. Selection criteria. Southeast and South regions. Multivariate analysis.

**ÁREA ANPEC:** ÁREA 12 - ECONOMIA SOCIAL E DEMOGRAFIA ECONÔMICA

**Classificação JEL:** I38, C38

<sup>1</sup> Os autores agradecem à Fapemig, CAPES e CNPq pelo apoio financeiro.

<sup>♦</sup> Graduada em Ciências econômicas pela Universidade Federal de Juiz de Fora (UFJF).

<sup>♦♦</sup> Professores do programa de pós-graduação (PPGE), departamento de ciências econômicas, UFJF; <sup>°</sup> Doutoranda do PPGE/UFJF.

## 1 Introdução

No Brasil, os programas federais de transferência de renda datam do ano de 2000. Os programas implantados no país, que deram origem ao Bolsa Família, foram o Bolsa Escola, Cartão Alimentação, Auxílio-Gás e Bolsa Alimentação, estes criados no período entre 2001 e 2003. Com a criação do Programa Bolsa Família (PBF), providenciou-se a unificação de tais programas (IPEA, 2012). O PBF foi então criado pelo Governo Federal, pela Lei 10.836, de 09 de janeiro de 2004 e regulamentado pelo decreto nº. 5209 de 17/09/2004. O programa tem como objetivos reduzir a pobreza e desigualdade de renda, provendo um benefício mínimo para famílias pobres; e reduzir a transmissão intergeracional de pobreza, por meio de medidas situadas principalmente no campo das políticas de educação, saúde e assistência social, condicionando o recebimento dos benefícios a investimentos em capital humano pelos beneficiários. O foco do programa são as famílias pobres e extremamente pobres inscritas no Cadastro Único, segundo uma regra de elegibilidade relacionada à renda familiar per capita. Embora a administração seja feita pelo Ministério do Desenvolvimento Social e Agrário (MDSA), várias outras instituições estão envolvidas, como a Caixa Econômica Federal e as prefeituras dos municípios, entre outras, o que denota características de descentralização e intersetorialidade do programa (IPEA, 2012).

O MDSA destaca que as condicionalidades têm como objetivo garantir direitos sociais básicos à população para que as futuras gerações quebrem o ciclo da pobreza, graças a melhores oportunidades de inclusão social (BRASIL, 2016a). O poder público deve fazer o acompanhamento gerencial para identificar os motivos do não cumprimento das condicionalidades. Uma vez identificados esses motivos, são implementadas ações de acompanhamento das famílias em descumprimento, consideradas em situação de maior vulnerabilidade social. De acordo com o MDSA, utilizando dados de 2016, são beneficiárias do PBF 13,55 milhões de famílias, consideradas extremamente pobres e pobres, registradas no Cadastro Único para Programas Sociais do Governo Federal (CadÚnico), com renda per capita de até R\$ 170,00 mensais (BRASIL, 2016b). O Cadastro Único é um instrumento de identificação e caracterização socioeconômica das famílias brasileiras de baixa renda, que é obrigatoriamente utilizado para seleção de beneficiários e integração de programas sociais do Governo Federal voltados ao atendimento desse público. A coordenação do Cadastro Único é realizada pelo MDSA e suas informações podem ser utilizadas pelos governos municipais e estaduais, possibilitando o desenvolvimento de políticas sociais locais (BRASIL, 2016b). Para a seleção das famílias que receberão o benefício é utilizado um sistema informatizado, a partir dos dados que elas informam no Cadastro Único. A concessão do benefício depende de quantas famílias já foram atendidas no município, em relação à estimativa de famílias pobres feita para essa localidade. Além disso, o governo federal precisa respeitar o limite orçamentário do programa (BRASIL, 2016b). Portanto, o benefício não alcança todas as famílias pobres e extremamente pobres e a inscrição das famílias com renda elegível ao PBF não garante a entrada imediata no programa.

A pobreza é um fenômeno multidimensional. Sendo assim, desde 2010, o Programa das Nações Unidas para o Desenvolvimento (PNUD) passou a utilizar o Índice de Pobreza Multidimensional (IPM) como um de seus indicadores no Relatório de Desenvolvimento Humano. O IPM visa fornecer um retrato mais amplo sobre as pessoas que vivem com dificuldades. Este indicador aponta privações em educação, saúde e padrão de vida; e pode ajudar a canalizar recursos de forma mais eficaz. Reunidos, esses itens proporcionam um retrato mais completo de pobreza do que simples indicadores de renda (PNUD, 2010). De acordo com o Relatório de Desenvolvimento Humano de 2015, entre 2006 e 2013, o total de brasileiros que vive em condições de pobreza multidimensional caiu de 4% para 2,9%. A maior privação é na área de saúde, especificamente em relação à mortalidade infantil. Da mesma forma, o percentual de famílias que vivem próximas à pobreza multidimensional, ou seja, que estão mais vulneráveis a essa situação, reduziu quatro pontos percentuais descendo de 11,2% para 7,2%, no mesmo período. A PNAD foi a fonte nacional usada para o cálculo (PNUD, 2015). Além da valorização do salário mínimo e do controle da inflação, os programas de transferência de renda (PTR) também contribuíram para a redução da pobreza e da desigualdade no Brasil (HOFFMANN, 2014).

Neste mote de pesquisa, este artigo busca responder a um problema aplicado para a economia brasileira: *quando se toma como parâmetro as condições socioeconômicas para avaliar a pobreza, pode-se dizer que o PBF tem atingido seu público alvo?* Assim, pretende-se investigar se possivelmente há

famílias em condições de pobreza que necessitam do benefício e não estão recebendo o mesmo, ou o inverso dessa situação. A hipótese levantada por este trabalho é que o uso da renda como único balizador da elegibilidade para o PBF é bem restrito, pois a pobreza está relacionada a outros fatores que não a renda, como expectativa de vida, saneamento e acesso a bens e serviços públicos. Ademais, existe a possibilidade da ocorrência de erros devido à desatualização cadastral ou manobras por parte dos beneficiários para manter a renda *per capita* informada dentro da exigida para ser incluído ou permanecer no programa. De acordo com Senna *et al.* (2007), o uso da renda monetária como critério único de seleção das famílias deve ser melhor avaliado, uma vez que somente a renda não é suficiente para qualificar a pobreza. Além disso, segundo os mesmos autores, a definição arbitrária de um valor *per capita* muito baixo tende a impossibilitar a inclusão de famílias que, apesar de situadas em uma faixa de renda um pouco acima do valor definido, encontram-se também em situação de pobreza.

O estudo utiliza os dados da segunda rodada da pesquisa de Avaliação do Impacto do PBF (AIBF II), considerando especialmente os indivíduos entrevistados das regiões Sul e Sudeste (macrorregião) para o ano de 2009. A base de dados utilizada nesta pesquisa é resultante do estudo de Vieira e Freguglia (2015), no qual os autores realizaram um tratamento dos dados referentes à AIBF I e AIBF II. Além disso, articula-se um conjunto de técnicas de análise multivariada, entre elas a Análise de Correspondência (AC) e a Análise de Componentes Principais (ACP). Esses métodos permitem estudar a relação entre as muitas variáveis utilizadas e suas associações, e tornam possível a visão simultânea de todas as características em um mapa de baixa dimensionalidade (LOESCH; HOELTGEBAUM, 2012). Em seguida, aplica-se a Análise Discriminante (AD) para classificar as famílias. Esse trabalho utiliza dados relacionados à saúde, educação e condições de vida dos beneficiários; quanto aos domicílios das famílias será analisado o perfil por tipo de utilidade pública (acesso à água encanada, iluminação e coleta de lixo). Os resultados obtidos poderão colaborar com o debate em curso sobre o PBF e contribuir para os formuladores de políticas públicas e gestores dos programas de transferência de renda, para as discussões e ajustes de estratégias que reforcem a efetividade dos critérios estabelecidos.

## 2 Revisão da literatura

Até meados da década de 90, as políticas de proteção social na América Latina estruturavam-se ao redor do mercado formal de trabalho e do sistema de seguridade social. Os programas de transferência de renda (PTR) foram impulsionados neste período devido às dificuldades econômicas e seus impactos na esfera social, como destacam Nascimento e Reis (2009, p. 187). Ainda de acordo com os mesmos autores, na primeira década do século XXI, os programas de transferência de renda proliferaram-se por mais de uma dezena de países, e tinham como objetivo central erradicar a pobreza extrema e a fome. Além do Brasil, Chile e México se destacam por seus Programas de Transferência Condicionada de Renda. Para os governos, estes programas são uma saída para lidar com uma severa restrição fiscal concomitante à necessidade de avanços na área social – seria uma forma de se fazer muito gastando pouco (CARDOSO; 2016; SOARES *et al.*, 2007). Os autores descrevem os programas do Chile e do México e verificam seus impactos sobre a desigualdade. O “Chile Solidário” foi criado em maio de 2002 como um programa de proteção social com foco nas 225 mil famílias extremamente pobres, que se estimava existirem naquele país de 16 milhões de habitantes. O programa conta com três componentes, quais sejam: i) apoio à família por parte de um assistente social; ii) vários subsídios monetários; e iii) acesso prioritário a outros programas de proteção social. Já o programa mexicano “Oportunidades” é o mais conhecido internacionalmente. Sob o nome “Progresas”, o programa começou em 1997, como sucessor do controverso “Solidaridad”. No início cobria 300 mil famílias, expandindo-se rapidamente para 2,5 milhões em 2000. Nesse período, o foco recaiu sobre famílias pobres em localidades rurais com menos de 2,5 mil habitantes, que tinham acesso a escolas e postos de saúde, de modo a se poder exigir contrapartidas das famílias beneficiadas. Em 2001, o programa foi estendido às localidades urbanas com menos de 15 mil habitantes, passando a cobrir todo o país em 2002. O alcance do programa chegou a cinco milhões de famílias em 2004.

No Brasil, o Bolsa Família é um programa que contribui para o combate à pobreza e à desigualdade. Ele foi criado em 2004 e possui três eixos centrais: i) complemento da renda; ii) acesso a direitos; e iii) articulação com outras ações e integração de várias políticas sociais. Desde 2011, o PBF faz parte do

Plano Brasil Sem Miséria, que reuniu diversas iniciativas para permitir que as famílias deixassem a extrema pobreza, com efetivo acesso a direitos básicos e a oportunidades de trabalho e de empreendedorismo. A gestão do Bolsa Família é descentralizada, ou seja, tanto a União, quanto os estados, o Distrito Federal e os municípios têm atribuições em sua execução (BRASIL, 2016a). O valor que cada família recebe por mês é a soma de vários tipos de benefícios previstos no Programa Bolsa Família. O valor médio do benefício por família é de R\$ 183,78. Os tipos e as quantidades de benefícios que cada família recebe dependem da composição (número de pessoas, idades, presença de gestantes etc.) e da renda da família beneficiária. O Benefício Básico, no valor de R\$ 85,00, é pago apenas a famílias extremamente pobres (renda mensal por pessoa de até R\$ 85,00). Os Benefícios Variáveis, no valor de R\$ 39,00 cada um, tem um limite de até cinco benefícios por família e estão vinculados à criança ou ao adolescente de 0 a 15 anos, gestante ou nutriz. Benefício Variável Vinculado ao Adolescente, entre 16 e 17 anos, no valor de R\$ 46,00, sendo até dois por família. Havendo ainda o Benefício para Superação da Extrema Pobreza, em valor calculado individualmente para cada família. Este é pago às famílias que continuem com renda mensal por pessoa inferior a R\$ 85,00, mesmo após receberem os outros tipos de benefícios do Programa (BRASIL, 2016b).

O Brasil atingiu em 2011, pela Pesquisa Nacional por Amostra de Domicílios (PNAD), seu menor nível de desigualdade de renda desde os registros nacionais iniciados em 1960. Ainda assim, a desigualdade no país permanece entre as 15 maiores do mundo. O país levaria pelo menos 20 anos no ritmo de crescimento em que estava em 2012, para atingir níveis de países como os Estados Unidos, que ainda não são uma sociedade totalmente igualitária. Todavia, isso significa que existem consideráveis reservas de crescimento pró-pobre, que só começaram a ser exploradas na década passada (IPEA, 2012). O Relatório de Desenvolvimento Humano do PNUD, divulgado em dezembro de 2015, mostra o Brasil na 75ª colocação no ranking do Índice de Desenvolvimento Humano (IDH), entre 188 países, com índice de 0,755, acima da média da América Latina (0,748) e considerado alto. É o 13º melhor índice entre os latino-americanos. Contudo, se a desigualdade social for levada em conta, o país perde cerca de um quarto de seu IDH e acaba tendo desempenho pior que seus vizinhos na América Latina.

Segundo o PNUD (2010), o IDH é apenas uma média e não ilustra claramente a desigualdade na distribuição do desenvolvimento humano. O fato motivou a elaboração de outro índice, o IDH Ajustado à Desigualdade (IDH-D). Tal índice leva em conta a desigualdade humana em três dimensões (i. e., renda, educação e expectativa de vida). Inexiste um ranking mundial levando em conta o IDH-D porque, de acordo com o PNUD, parte dos países não tem dados suficientes para elaborar esse índice. De acordo com o IPEA, a inédita redução da desigualdade observada na década passada pode ser decomposta pelas diversas fontes de renda. Ao captar e retrabalhar tais fontes de renda obtidas da PNAD, obteve-se resultados interessantes. O trabalho foi responsável por cerca de 58% da redução da desigualdade, enquanto que a previdência correspondeu a 19%. A contribuição do Bolsa Família chegou a representar 13% na redução. Já o benefício de prestação continuada (BPC), 4% e outras rendas como aluguéis e juros com os restantes 6% (BRASIL, 2012). Desse modo, o PBF representa uma parcela a ser considerada na contribuição para redução da desigualdade.

Durante a elaboração do PBF diversas questões polêmicas foram levantadas, como por exemplo, a opção de operar com benefícios monetários para as famílias sem controle quanto à destinação que elas dariam aos recursos (CAMPELLO, 2013). Não obstante, decidiu-se que só a própria família poderia escolher a melhor forma de utilizar o dinheiro de acordo com suas características e necessidades. Na literatura relacionada aos PTR encontram-se trabalhos de pesquisa e verificação de diferentes aspectos associados a tais programas (e. g. Nascimento e Reis, 2009; Soares, 2010; Senna et al., 2007). Dentre eles, o mais estudado é o PBF, sendo encontrados inclusive estudos realizados por pesquisadores do exterior (e. g. Bastagli, 2008; Lignani, 2011; Brauw, 2012). Estudos que investigaram a geração de desincentivos para a oferta de trabalho e incentivos à fertilidade, por parte do PBF, apontam que, quando existem, estes efeitos são muito pequenos para serem considerados relevantes (NETO, 2014). Conforme Lavinhas (2000), avaliações dos programas brasileiros de garantia de renda mínima (PGRM) demonstram vantagem da transferência monetária direta na redução dos custos administrativos do programa e na redução da incidência de fraudes e desvios de recursos. Outra vantagem observada é a liberdade dos

beneficiários em escolher o que vão comprar, diferente da modalidade clássica de distribuição do benefício in natura.

Já o critério de seleção das famílias por meio da renda é amplamente criticado na literatura especializada. Como afirma Senna et al. (2007, p. 88), um dos pontos críticos enfrentados por programas focalizados como o PBF é a determinação dos critérios e mecanismos a serem utilizados para seleção dos beneficiários. Esta tarefa se torna complexa, uma vez que é necessário levar em consideração as vulnerabilidades sociais enfrentadas pela população que se pretende atender. Ainda no estudo realizado por Senna *et al.* (2007), faz-se uma crítica à atuação do PBF com foco na família, ou seja, tendo com prioridade os grupos já frequentemente atendidos por políticas sociais, quais sejam: as crianças, adolescentes, gestantes e nutrízes. De acordo com os autores, apesar do consenso de que estes grupos são mais vulneráveis, existe a preocupação para que sejam adotadas medidas voltadas à geração de emprego e renda que, em conjunto com estratégias socioeducativas junto às famílias, contribuiriam para romper com o ciclo de reprodução da pobreza no país. Para Magalhães *et al.* (2007), os mecanismos que promovem a inclusão dos grupos que necessitam de uma dada política serão mais efetivos quanto maior a capacidade do mecanismo de captar as distintas vulnerabilidades destes grupos. Aqui se observa novamente a importância e complexidade na determinação dos critérios e mecanismos a serem utilizados para seleção e manutenção dos beneficiários.

O PBF tem como característica foco nas famílias pobres e não exclusivamente no indivíduo. Este tipo de enfoque é resultado do processo de reformas das políticas sociais desenvolvidas no país desde os anos 1980. Segundo Carvalho (1997), a abordagem sobre o tema família adquire novas especificidades em razão do reconhecimento de uma inequívoca situação de desemprego estrutural, que afeta a já frágil capacidade das famílias enfrentarem os desafios de reprodução social de seus membros. De acordo com Draibe (1998), a decisão de se tomar a família como foco se baseia no argumento de que as políticas de proteção à família teriam maior potencial de impactar as condições de vida da população pobre. Assim, a família fica comprometida em cumprir as condicionalidades do programa, encaminhando os filhos à escola, cumprindo o calendário de vacinação e fazendo acompanhamento de saúde das gestantes e nutrízes. Ao analisar os programas de transferência de renda realizados nos anos 90, verificou-se que o foco na família foi a forma encontrada para atingir o público alvo que seriam as crianças e adolescentes. Sendo possível dessa forma, incluí-los em outras políticas, sobretudo na educação (DRAIBE *et al.*, 1998). Com o PBF, o objetivo da educação é alcançado em parte, pelo cumprimento das condicionalidades. Para Marques *et al.* (2007), além dos recursos monetários recebidos e da melhoria no acesso aos direitos sociais básicos, alcançados pelo cumprimento das condicionalidades do programa, uma questão levantada é que estes recursos ajudariam a movimentar a economia local. Ainda segundo os mesmos autores, o efeito multiplicador dos recursos aplicados ao PBF vai depender da propensão marginal a consumir de seus beneficiários. Como tal população possui alta propensão ao consumo, à economia local é movimentada e parte do recurso aplicado volta aos cofres públicos como forma de incremento na arrecadação de tributos.

Cabe mencionar que dentro do contexto deste estudo, há vários outros trabalhos aplicados que versam sobre Programas de Transferência de Renda (PTR). Por exemplo, Nascimento e Reis (2009) se propuseram a realizar uma reflexão acerca dos PTR desenvolvidos na América Latina e das diferentes condicionalidades exigidas aos beneficiários. A conclusão dos autores foi que esses programas de transferência de renda são ineficazes do ponto de vista do enfrentamento à pobreza, sobretudo porque não combatem as raízes da desigualdade na região. Já Soares (2010) buscou compreender os programas do Brasil, México, Chile e Uruguai. Segundo o autor, o formato que os PTR adotaram nos diferentes países tem mais relação com a história das políticas sociais em tais países. O autor avaliou o peso que os programas colocam nos dois objetivos primordiais dos PTR. Em sua conclusão, observou que existem mais diferenças do que similaridades entre os programas avaliados; tendo em comum, o fato de serem focalizados, fazerem transferências em dinheiro e exigirem em menor ou maior grau alguma contrapartida dos beneficiários. Senna *et al.* (2007) analisaram o desenho do PBF, buscando identificar possíveis inflexões nele contidas, a partir de eixos como foco e condições de acesso ao programa, condicionalidades, descentralização, intersetorialidade e controle social. O que se concluiu das análises é que o critério renda adotado pelo PBF é bem restrito, ao ponto de expor o programa ao risco de não

atingir os segmentos sociais que enfrentam situações diversas de vulnerabilidade. Por sua vez, Soares et al. (2010) discutiram o impacto do PBF tanto sobre a pobreza quanto sobre a desigualdade. Os impactos do PBF sobre a desigualdade foram estimados mediante a decomposição fatorial do coeficiente de Gini de acordo com cada fonte de renda. Já Souza e Osorio (2013) realizaram um estudo cujo objetivo foi analisar o perfil da pobreza no Brasil entre 2003 e 2011. Para tal foram investigadas uma série de características, entre elas geográficas, demográficas, educacionais e ocupacionais. Definiram quatro estratos a partir da renda domiciliar *per capita*, quais sejam: os extremamente pobres, pobres, vulneráveis e não pobres. Optou-se por apresentar as informações do modo mais simples e detalhado possível, recorrendo a tabulações cruzadas. A conclusão dos autores foi que se podia discutir se a melhora na vida dos mais pobres foi tão intensa quanto poderia ter sido, mas não restavam dúvidas de que a situação dos mais pobres melhorou. Camargo *et al.* (2013) traçaram o perfil socioeconômico dos beneficiários do PBF a partir da análise das informações do CadÚnico. A seleção de variáveis seguiu uma avaliação prévia da qualidade de preenchimento destas informações e, para análises históricas, da consistência da série, privilegiando como ponto de comparação os indicadores de 2007. Com a análise do perfil das famílias do PBF, as autoras concluíram que ainda existiam importantes vulnerabilidades em relação à população em geral, mas apontava-se para uma melhoria nas condições de vida em termos de acesso a serviços de infraestrutura domiciliar e escolaridade. Por sua vez, Neto (2014) analisou o efeito do PBF sobre o estado nutricional das crianças e adolescentes beneficiadas, no qual foram utilizados dados da Pesquisa de Orçamentos Familiares (POF, 2008/2009), conjuntamente com o método *Propensity Score Matching* (PSM); os resultados indicaram que o programa melhorou os indicadores nutricionais daqueles que pertenciam ao grupo de tratamento. Para o autor, além da transferência de renda, acreditava-se que as condicionalidades nas áreas de saúde e educação podiam explicar parte desse resultado. Todavia, para Lavinias (2004), a obrigatoriedade do cumprimento das condicionalidades representava uma ameaça aos princípios de cidadania, uma vez que é exigido o cumprimento de obrigações como condição para o exercício de um direito social. Já Arretche (2002) concluiu que, naquele período, os municípios brasileiros em geral ainda apresentavam dificuldades na oferta de serviços de educação e saúde, o que comprometeria o processo de implementação das condicionalidades do PBF.

Esta pesquisa possui algumas semelhanças com alguns dos artigos supracitados. No caso de Camargo et al. (2013), por exemplo, as autoras traçam o perfil socioeconômico dos beneficiários, porém utilizando o CadÚnico. No presente trabalho há também um diferencial com relação à metodologia utilizada, qual seja: o uso em conjunto da Análise de Correspondência (AC), da Análise de Componentes Principais (ACP) e da Análise Discriminante (AD). Aqui será utilizada a pesquisa AIBF II, assim como Jannuzzi e Pinto (2013) e não será realizada análise histórica. Jannuzzi e Pinto (2013) apresentaram evidências empíricas acerca dos efeitos do PBF. Já neste estudo, tratar-se-á das associações de acordo com as características dos indivíduos, por meio das técnicas de análise multivariada, com os dados de 2009. Além disso, o objetivo principal se diferencia dos demais. O propósito é avaliar se o PBF tem atingido seu público alvo, mas não adotando a renda como parâmetro, e sim as condições socioeconômicas observadas.

### 3 Metodologia

Esta seção tem como propósito descrever formalmente as técnicas de análise multivariada. Após a descrição da metodologia, discorre-se sobre a base de dados utilizada. As técnicas de Análise de Correspondência (AC) e Componentes Principais (ACP) permitem reduzir as dimensões originais dos dados em dimensões ortogonais. Além disso, é possível extrair escores que possam representar bem as associações das variáveis. A partir dos índices gerados com as duas primeiras técnicas, utiliza-se a Análise Discriminante (AD). O objetivo é separar a população entrevistada em dois grupos: aqueles que devem receber o benefício de acordo com suas características socioeconômicas e aqueles que não devem receber o benefício de acordo com as mesmas características. A inspiração para este tipo de análise vem da adoção pelo PNUD do Índice de Pobreza Multidimensional (IPM). Segundo as Nações Unidas, este índice avalia de forma mais ampla as condições de privação, nos aspectos mais básicos, que parte da população enfrenta. Tendo como foco as variáveis semelhantes às do IDH, quais sejam: educação, saúde e renda. O IPM mede dimensões de pobreza não relacionada com o rendimento. O IPM classifica a

população com base no número relativo de privações que as pessoas pobres sofrem simultaneamente, podendo haver pessoas próximas de uma situação de pobreza multidimensional, em situação de pobreza multidimensional, e em pobreza extrema. Os indicadores utilizados pelo IPM são relacionados à saúde (nutrição e mortalidade infantil), educação (anos de escolaridade e crianças matriculadas) e padrão de vida (gás de cozinha, sanitários, água, eletricidade, pavimento e bens domésticos). Cada indicador possui um peso proporcional a um terço do IPM e as famílias são consideradas multidimensionalmente pobres, se sofrem privação em pelo menos 30% dos indicadores (PNUD, 2016b). Com os dados coletados pretende-se analisar se os grupos que declaram renda elegível para que sejam beneficiárias do PBF possuem características socioeconômicas condizentes com as do público alvo do programa. Dessa forma, questiona-se se a renda, de forma isolada, é suficiente para determinar quem deve receber o benefício. Cabe mencionar que a escolha das regiões Sul e Sudeste é uma forma de delimitar o estudo, visto que a base de dados é representativa para o país e para três macrorregiões, sendo a macrorregião Sul+Sudeste, bastante heterogênea e representativa.

### 3.1 Análise de Correspondência

A escolha da técnica de análise de correspondência se deve ao grande número de informações que serão utilizadas e que possuem caráter qualitativo. Tais informações possuem estrutura composta por variáveis de respostas categóricas. A base de dados trabalhada possui grande parte de suas variáveis medidas em escala nominal, possibilitando a geração de uma tabela de contingências. Formalmente, na versão simples, essa técnica estatística decompõe em coordenadas uma matriz de participação ( $P$ ), constituída das proporções  $p_{ij} = \frac{n_{ij}}{n}$ , e oriunda de uma tabela de contingência. Assim, considere a matriz  $\tilde{P} = P - r c'$ , tal que o perfil das linhas é representado por  $r' = (\frac{n_{.1}}{n} \frac{n_{.2}}{n} \dots \frac{n_{.j}}{n})$ , sendo  $n_{.j} = \sum_{i=1}^I n_{ij}$ ; e o perfil das colunas das colunas é  $c' = (\frac{n_{.1}}{n} \frac{n_{.2}}{n} \dots \frac{n_{.j}}{n})$ , sendo  $n_{.j} = \sum_{i=1}^I n_{ij}$ . O posto da matriz  $\tilde{P}_{ixj}$  é dado por  $k = \min(p - 1, q - 1)$ . Aplicando a decomposição de valor singular sobre a matriz  $\tilde{P}_{ixj}$ , tem-se:

$$\tilde{P}_{ixj} = A \Lambda B' \quad (1)$$

em que  $A = D_r^{1/2} U_{ixk}$  é uma matriz de dimensão  $ixk$ ,  $B = D_c^{1/2} V_{jxk}$  é uma matriz de dimensão  $jxk$ ,  $U$  e  $V$  são matrizes ortogonais e  $\Lambda$  é a matriz de dimensão  $kxk$  contendo os autovalores da matriz  $\tilde{P}$  ordenados em ordem decrescente. A matriz  $U$  contém os autovetores da matriz  $\tilde{P}\tilde{P}'$ , e a matriz  $V$  contém os autovetores da matriz  $\tilde{P}'\tilde{P}$ . Como consequência dessa decomposição, as linhas da matriz  $\tilde{P}$  podem ser escritas como combinações lineares das linhas da matriz  $B'$ , e as colunas da matriz  $\tilde{P}$  podem ser escritas como combinações lineares das colunas da matriz  $A$ . Assim, são definidas as coordenadas principais das linhas ( $Y_{ixk}$ ) e das colunas ( $Z_{jxk}$ ) da matriz  $\tilde{P}$ :

$$Y_{ixk} = D_r^{-1} A_{ixk} \Lambda_{kxk} \quad (2)$$

$$Z_{jxk} = D_c^{-1} A_{jxk} \Lambda_{kxk} \quad (3)$$

Consequentemente, a matriz  $\tilde{P}$  pode ser expressa como função dos autovalores e das coordenadas principais, ou seja,  $\tilde{P} = P - r c' = \sum_{i=1}^k \hat{\lambda}_i \tilde{a}_i \tilde{b}_i'$ , tal que  $\tilde{a}_i$  denota a  $i$ -ésima coluna da matriz  $A$  e  $\tilde{b}_i$  denota a  $i$ -ésima coluna da matriz  $B$ , tal que  $k = \text{posto}(\tilde{P}) = \min(i - 1, j - 1)$ . As duas primeiras coordenadas principais das linhas e das colunas são as mais representativas em termos da associação total que existe entre as variáveis  $Z$  e  $Y$ , pois estão relacionadas aos maiores autovalores da matriz  $\tilde{P}$ . A variação total existente neste sistema, chamada também de inércia total, isto é,  $\sum_{i=1}^k \lambda_i^2$ , no qual  $\lambda_i$  são os autovalores não-nulos da diagonal da matriz  $\Lambda$ ,  $i = 1, 2, \dots, k$ . A proporção de representatividade da  $i$ -ésima coordenada principal em relação à inércia total é definida por  $\frac{\lambda_i^2}{\sum_{i=1}^k \lambda_i^2}$ . Essa técnica utilizada oferece uma representação geométrica das correspondências existentes na tabela de contingências. Dessa forma, pode-se fazer um gráfico de dispersão das coordenadas principais das linhas e das colunas com o intuito de analisar a associação entre as linhas e colunas. A análise de correspondência é mais efetiva se a matriz de dados é bastante grande, de modo que a inspeção visual ou análise estatística simples não consegue revelar sua estrutura (CZERMAINSKI; DIAS, 2004).

### 3.2 Análise de Componentes Principais

Essa técnica expressa a estrutura de variância e covariância por meio de (poucas) combinações lineares das variáveis originais em estudo. Essas combinações lineares chamadas de componentes principais, não são correlacionadas entre si (ortogonais), e são ordenadas pela parcela da informação total que cada componente retém. Assim, os componentes principais permitem reduzir e classificar os dados originais mantendo a maior parte das informações (MINGOTI, 2007). Considerando que a matriz de correlação  $C$  possui os pares de autovalores e autovetores,  $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$ , tal que

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ , e o vetor aleatório é  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$ , então os componentes principais são:

$$\begin{aligned} Y_1 &= \mathbf{a}'_1 \mathbf{Z} = a_{11}Z_1 + a_{12}Z_2 + \dots + a_{1p}Z_p \\ &\vdots \\ Y_p &= \mathbf{a}'_p \mathbf{Z} = a_{p1}Z_1 + a_{p2}Z_2 + \dots + a_{pp}Z_p \end{aligned} \quad (4)$$

sendo  $Var(\mathbf{a}'_i \mathbf{Z}) = \mathbf{a}'_i \mathbf{C} \mathbf{a}_i$ ,  $Cov(Y_i, Y_k) = \mathbf{a}'_i \mathbf{C} \mathbf{a}_k$ ,  $\forall i, k = 1, 2, \dots, p$ . A padronização de cada varável do vetor aleatório  $\mathbf{X}'$ ,  $Z_k = (X_k - \mu_k) / \sqrt{\sigma_{kk}}$ , retira a influência de sua escala na formação dos componentes.

As combinações lineares,  $Y_1, Y_2, \dots, Y_p$ , não são correlacionadas (ortogonais) e suas variâncias são as maiores possíveis (JOHNSON; WICHERN, 2002). Ou melhor, o  $i$ -ésimo componente ( $Y_i = \mathbf{a}'_i \mathbf{X}$ ) maximiza  $Var(\mathbf{a}'_i \mathbf{X})$  sujeito a  $\mathbf{a}'_i \mathbf{a}_i = 1$  e  $\mathbf{a}'_i \mathbf{a}_k = 0$ . Pelo resultado do teorema espectral, o somatório das variâncias do sistema padronizado de dados é o mesmo que o somatório da variância dos componentes

principais:  $\sum_{k=1}^p Var(Z_k) = \lambda_1 + \lambda_2 + \dots + \lambda_p = \sum_{i=1}^p Var(Y_i)$ . A proporção do  $i$ -ésimo componente mais

representativo é definida como:  $\%Var(Y_i) = \lambda_i / \sum_{i=1}^p \lambda_i$ . Assim, é possível selecionar um número mínimo

de componentes necessários para expressar uma grande parte da variância total das variáveis originais<sup>2</sup> (JOHNSON; WICHERN, 2002; MINGOTI, 2007).

### 3.3 Análise Discriminante

A Análise Discriminante (AD) é uma técnica que pode ser utilizada para classificação de elementos de uma amostra ou população. Para sua aplicação, é necessário que os grupos para os quais cada elemento amostral pode ser classificado sejam predefinidos, ou seja, conhecidos *a priori* considerando-se suas características gerais. Este conhecimento permite a elaboração de uma função matemática chamada de regra de classificação ou discriminação, que é utilizada para classificar novos elementos amostrais nos grupos já existentes (JOHNSON; WICHERN, 2002; MINGOTI, 2007). A discriminação é alcançada pela definição dos pesos das combinações lineares para cada variável, com o objetivo de maximizar a variância entre grupos em relação à variância dentro dos grupos. Neste artigo utilizar-se-á a função discriminante canônica de Fisher (MANLY, 2008). Apesar do método de Fisher conter em sua formulação básica a suposição de distribuição normal p-variada, na realidade o que se utiliza para elaboração das funções canônicas é a noção de análise de variância para construção das combinações lineares mais informativas. Assim, é um método que não depende intrinsecamente da normalidade, podendo ser considerado como não paramétrico e aplicável em situações nas quais a normalidade multivariada não é válida (MINGOTI, 2007; JOHNSON; WICHERN, 2002). Formalmente, a partir de um vetor aleatório ( $\mathbf{X}$ ), as funções discriminantes canônicas são representadas como:

$$Z = a_1 X_1 + a_2 X_2 + \dots + a_p X_p \quad (5)$$

Grupos podem ser bem separados usando  $Z$  se o valor médio desta variável muda consideravelmente de grupo para grupo, com os valores dentro do grupo sendo razoavelmente constantes. Uma maneira de

<sup>2</sup> Como regra básica deve-se selecionar um número mínimo que atinja pelo menos 70% da variância total. Além disso, é necessário que as variáveis sejam correlacionadas. Por essa razão, emprega-se o teste de esfericidade de Bartlett (JOHNSON; WICHERN, 2002; MINGOTI, 2007).



determinar os coeficientes  $a_1, a_2, \dots, a_p$  no índice envolve escolhê-los de modo a maximizar a razão  $F$  para uma análise de variância de um fator. Uma função adequada para separar os grupos pode ser definida como a combinação linear para a qual a razão  $F = M_B/M_W$  é tão grande quanto possível, como sugerido por Fisher (1936). Sendo  $M_B$  e  $M_W$ , em uma análise de variância nos índices  $Z$ , iguais ao quadrado médio entre grupos e dentro dos grupos, respectivamente. Quando essa abordagem é usada, acontece que pode ser possível determinar várias combinações lineares para separar grupos. Em geral, o número disponível,  $s$ , é o menor entre  $p$  e  $m-1$ . Encontrar os coeficientes das funções discriminantes canônicas vem a ser um problema de autovalor. A matriz de somas de quadrados e produtos cruzados dentro da amostra é definida como  $W$ , e a matriz amostral total de somas de quadrados e produtos cruzados define-se como  $T$ . Destas, a matriz entre grupos pode ser determinada.

$$B = T - W \quad (6)$$

A seguir, os autovalores e autovetores da matriz  $W^{-1}B$  têm que ser encontrados. Se os autovalores são  $\lambda_1 > \lambda_2 > \dots > \lambda_s$ , então  $\lambda_i$  é a razão da soma dos quadrados entre grupos e da soma dos quadrados dentro dos grupos para a  $i$ -ésima combinação linear,  $Z_i$ , enquanto que os elementos do correspondente autovetor,  $\mathbf{a}'_i = (a_{i1}, a_{i2}, \dots, a_{ip})$ , são os coeficientes das variáveis  $X$  para este índice. As funções discriminantes canônicas  $Z_1, Z_2, \dots, Z_s$  são combinações lineares das variáveis originais escolhidas de tal maneira que  $Z_1$  reflete diferenças de grupo tanto quanto possível,  $Z_2$  captura tanto quanto possível as diferenças de grupo não apresentadas por  $Z_1$ ;  $Z_3$  captura tanto quanto possível as diferenças de grupo não apresentadas por  $Z_1$  e  $Z_2$ , etc. A expectativa é que as primeiras poucas funções sejam suficientes para contar por quase todas as importantes diferenças de grupo. Em particular, se somente a primeira ou duas funções são necessárias para este propósito, então é possível uma representação gráfica simples do relacionamento entre os vários grupos representando os valores destas funções para os indivíduos da amostra (MANLY, 2008).

### 3.4 Base de Dados

Esta pesquisa tem como fonte a base de dados resultante do estudo de Vieira e Freguglia (2015). Os autores realizaram um tratamento dos dados obtidos da Pesquisa de Avaliação de Impactos do Programa Bolsa Família – primeira e segunda rodadas (AIBF I e AIBF II). Neste trabalho serão utilizados os dados referentes à segunda rodada. A coleta dos dados pela AIBF II foi realizada entre os meses de setembro e novembro de 2009, por meio de entrevistas a famílias beneficiárias e não beneficiárias do PBF. Segundo a ficha técnica da pesquisa, os itens avaliados pela AIBF II dizem respeito ao consumo das famílias, à antropometria de crianças, à vida escolar das crianças, entre outros, considerando condições de acesso à educação, saúde e mercado de trabalho (MDSA, 2016). A AIBF II possui caráter amostral e foi realizada em 269 municípios, tendo representatividade nacional e para três macrorregiões: Nordeste, Norte Urbano+Centro Oeste e Sul+Sudeste. Buscando um recorte regional, optou-se neste estudo por trabalhar os dados da macrorregião Sul+Sudeste, pelas suas características de grande número de habitantes, heterogeneidade entre a população e representatividade para o país (MDSA, 2016). O desenho amostral da pesquisa foi complexo, envolvendo estratificação, o que implicará na necessidade de consideração de mais características nas análises. Para o presente estudo, o peso amostral dos dados de Vieira e Freguglia (2015) foi utilizado tanto na aplicação da Análise de Correspondência Múltipla quanto na Análise de Componentes Principais. Os dados utilizados foram selecionados a partir do questionário dessas informações estatísticas da AIBF II, de forma a priorizar os indicadores utilizados pelo PNUD para calcular o IPM. Para tanto, este trabalho selecionou um conjunto de variáveis, separadas em três dimensões características do IPM, quais sejam: saúde, educação e padrão de vida (QUADRO 1).

A partir desse conjunto de variáveis, a estratégia empírica deste trabalho consistiu em dois passos sucessivos. O primeiro passo representou a elaboração de vários indicadores (ou índices gerais) das três dimensões IPM, aplicando duas tradicionais técnicas de estatística multivariada. Para a dimensão “Padrão de Vida”, três indicadores foram construídos pela Análise de Correspondência Conjunta (método joint), a saber: estrutura domiciliar, bens duráveis (que representam modo de transporte) e outro para bens duráveis (eletrodomésticos). Esses indicadores foram formados a partir de variáveis qualitativas nominais. Já para a dimensão “Saúde e Educação”, os indicadores envolvem variáveis quantitativas e nominais. Por essa razão, primeiramente construiu-se o indicador “Alimentos Básicos” pela Análise de Componentes

Principais. Por outro lado, um indicador “Saúde e Educação” foi também elaborado mediante ao uso da Análise de Correspondência Conjunta.

**QUADRO 1 – Dimensões, indicadores e variáveis**

Dimensões IPM	Indicadores	Variáveis	Descrição
Padrão de vida	Estrutura domiciliar	s01x16	Existe água canalizada dentro do domicílio?
		s01x22	Principal tipo de iluminação?
		s01x23	Principal tipo de combustível usado para cozinhar?
		s01x24	Principal destino do lixo domiciliar?
	Bens duráveis (eletrodomésticos)	s10a3001	Algum membro deste domicílio possui ou aluga fogão?
		s10a3002	Algum membro deste domicílio possui ou aluga freezer?
		s10a3003	Algum membro deste domicílio possui ou aluga geladeira?
		s10a3004	Algum membro deste domicílio possui ou aluga micro-ondas?
		s10a3009	Algum membro deste domicílio possui ou aluga ferro elétrico?
		s10a3010	Algum membro deste domicílio possui ou aluga máquina de lavar roupas?
		s10a3012	Algum membro deste domicílio possui ou aluga TV a cores?
		s10a3018	Algum membro deste domicílio possui ou aluga aparelho de som - cd?
		s10a3019	Algum membro deste domicílio possui ou aluga aparelho de dvd?
		s10a3025	Algum membro deste domicílio possui ou aluga chuveiro elétrico?
	Bens duráveis (modos de transporte)	s10a3026	Algum membro deste domicílio possui ou aluga micro-computador?
		s10a3027	Algum membro deste domicílio possui ou aluga antena parabólica?
		s10a3028	Você ou algum membro deste domicílio possui ou aluga automóvel?
s10a3029		Você ou algum membro deste domicílio possui ou aluga bicicleta?	
Saúde	Alimentos básicos	s09x5_1	Qde de arroz comprada, obtida ou produzida para consumo no domicílio?
		s09x5_2	Qde de feijão comprada, obtida ou produzida para consumo no domicílio?
		s09x5_501	Qde de óleo de soja comprada, obtida ou produzida para consumo no domicílio?
		s09x5_701	Qde de café comprada, obtida ou produzida para consumo no domicílio?
Educação	Saúde e Educação	s04a6	Tem cobertura de plano de saúde ou seguro saúde?
		s03a5	Frequenta escola ou creche atualmente?
	Bolsa Família	s15c1	Alguém do seu domicílio é beneficiário do Programa Bolsa Família?
	Peso amostral	weight	2009: peso amostral

Fo

## 4 Resultados e discussões

### 4.1 Análise dos indicadores construídos (índices gerais)

O método de AC utilizado parte do método Joint (ou conjunta) dado que tanto a matriz indicadora quanto à Burt inflam artificialmente as distâncias qui-quadrado entre os perfis coluna e a inércia total. Além disso, o número reduzido de dimensões deve atingir em torno de 90% da estrutura original dos dados para que represente bem a estrutura de dados original. Conforme se observa na Tabela 1, os valores das dimensões para os indicadores estão todos acima de 90%. Analisando o indicador “Estrutura domiciliar” o valor 0,089 apresentado na Tabela 2 é a inércia associada com a primeira dimensão. Essa inércia é de, aproximadamente, 67% da inércia total, ou seja, esta dimensão atinge cerca de 67% dos dados originais. A inércia associada com a segunda dimensão é 0,040, e esta alcança em torno de 31% da inércia total. Portanto, em conjunto, as duas dimensões são responsáveis por 98% da inércia total. Em suma, os indicadores construídos são representativos à estrutura original de dados.

As coordenadas sobre cada dimensão e a contribuição do padrão de resposta, atribuídas em relação à inércia total estão reportadas no Quadro 2, cuja ilustração está dividida por painéis que representam cada indicador: (A) Estrutura domiciliar; (B) Modos de transporte; . Praticamente todas as respostas nominais exibiram uma boa qualidade de aproximação, o que significa afirmar que o número de dimensões

selecionado representa bem as próprias respostas categóricas. Assim, por conveniência, os índices de qualidade de aproximação foram ocultados. Dentre as categorias listadas no indicador (A) Estrutura domiciliar, as respostas categóricas de “Qual o principal tipo de combustível usado para cozinhar?” (s01x23) foram as mais contributivas à inércia total. Ou seja, o indicador “Estrutura Domiciliar” é formado principalmente pelos tipos de combustível. Contudo, a resposta categórica sobre o destino do lixo “Queimado ou enterrado” também expressa bem a inércia total deste indicador. Em conjunto, esses padrões de resposta se destacam com grandes contribuições à inércia sobre a dimensão 1 (Dim 1). Das respostas que mais contribuem para a inércia total, em relação ao combustível usado para cozinhar, somente a resposta categórica “Gás” apresenta coordenadas positivas dentre elas. Desse modo, a dimensão 1 distingue basicamente os respondentes que usam esse tipo de combustível para cozinhar em seus domicílios, daqueles que não usam. Parece, portanto, que esse indicador procura isolar no estudo famílias típicas e que provavelmente usam o gás para cozinhar em seus domicílios.

**TABELA 1 – Resultados agregados dos indicadores por AC**

Dimensões IPM	Indicador	Dimensões	Inércia Principal	Part. %	Part. (%) acumulada
Padrão de vida	Estrutura domiciliar	Dim 1	0,089	67,43	67,43
		Dim 2	0,041	30,84	98,27
	Bens duráveis (meios de transporte)	Dim 1	0,029	71,13	71,13
		Dim 2	0,012	28,87	100,00
	Bens duráveis (eletrodomésticos)	Dim 1	0,025	86,01	86,01
		Dim 2	0,002	7,60	93,61
Saúde e Educação		Dim 1	0,129	73,97	73,97
		Dim 2	0,046	26,03	100,00

Fonte: Resultados da pesquisa.

Nota: Dimensão 1 (dim 1); Dimensão 2 (Dim 2).

Já para a dimensão 2 (Dim 2), a maior contribuição é dada pelos respondentes que usam essencialmente a “eletricidade” para cozinhar, cujo padrão de resposta exhibe coordenadas positivas. Observa-se também que esses respondentes não se desfazem do lixo domiciliar queimando-o. A variável que se refere ao destino dado ao lixo domiciliar foi a segunda mais representativa para explicar a inércia da dimensão 2. Em suma, enquanto a primeira dimensão procura captar, em alguma medida, a família típica brasileira, a segunda dimensão procura isolar famílias com um maior padrão em seus domicílios, representado pelo uso de eletricidade para cozinhar.

Por seu turno, no indicador “(B) Modos de transporte” as maiores contribuições para a inércia total foram dadas pelas respostas da variável “... possui ou aluga trator/caminhão?” (s10a3031). Isso demonstra que a resposta a essa questão tem grande influência sobre o indicador “Modos de Transporte”. Todavia, a resposta categórica à questão “... possui ou aluga moto?” (s10a3030) também contribui bastante para a inércia total deste indicador. Todos os que responderam “sim” para as questões das variáveis relacionadas a modos de transporte apresentam coordenadas positivas para a dimensão 1, destacando-se com maior valor a coordenada da variável s10a3031. Assim a dimensão 1 distingue os que possuem e os que não possuem modos de transporte, isolando os que aparentemente possuem uma condição socioeconômica mais favorável, ou seja aqueles que possuem trator/caminhão. Analisando a dimensão 2, a maior contribuição foi dada pela variável “... possui ou aluga bicicleta?” (s10a3029), cujo padrão de respostas possui coordenadas negativa para a resposta “sim” e positiva para a resposta “não”. Observa-se também que aqueles que responderam “não” nas quatro questões ficaram próximos graficamente e muito provavelmente são indivíduos em condições socioeconômicas menos favoráveis que os demais.

Já no indicador “(D) Eletrodomésticos”, contribuiu mais para a inércia total, a variável “... possui ou aluga microcomputador?” (s10a3026). Assim, o indicador “(D) Eletrodomésticos” é bastante influenciado por esta questão. Entretanto, a categoria “... possui ou aluga máquina de lavar roupas?” (s10a3010) também contribui consideravelmente para a inércia total desse indicador. Juntas, essas duas categorias se destacam com grandes contribuições à inércia sobre a dimensão 1. As coordenadas na dimensão 1 para a

resposta “sim” a estas questões, são fortemente positivas, assim aqueles que possuem microcomputador e/ou máquina de lavar roupas ficaram graficamente distantes dos que responderam “não” para qualquer variável deste indicador. Por outro lado, observando a dimensão 2, a maior contribuição é daqueles que responderam não possuir geladeira, cujo padrão de resposta exhibe coordenadas fortemente negativas. Esses indivíduos situaram-se distantes dos que responderam “sim” para qualquer variável listada na tabela acima. Dessa forma, entende-se que os que possuem microcomputador tem melhores condições socioeconômicas que os demais, e aqueles que não possuem geladeira são os que se encontram em condições associadas à pobreza e com maior necessidade de recebimento do Bolsa Família.

**QUADRO 2 – Coordenadas e composição dos indicadores formados**

	Respostas	Part. Inércia	Coordenadas		Respostas	Part. Inércia	Coordenadas	
			Dim 1	Dim 2			Dim 1	Dim 2
<b>(A) Estrutura Domiciliar</b>	<i>s01x16 - Existe água canalizada dentro do domicílio?</i>				<i>s10a3001 - Possui ou aluga fogão?</i>			
	Sim	0,00	<u>0,00</u>	<u>0,01</u>	Sim	0,00	<u>0,01</u>	<u>0,01</u>
	Nao	0,01	-0,06	-0,31	Nao	0,02	-0,36	-0,43
	<i>s01x22 - Qual é a principal iluminação utilizada?</i>				<i>s10a3003 - Possui ou aluga geladeira?</i>			
	Electrica	0,00	<u>0,00</u>	0,00	Sim	0,00	<u>0,02</u>	<u>0,02</u>
	Gerador	0,00	-0,44	<u>0,72</u>	Nao	0,05	-0,51	-0,36
	Lampiao	0,01	-1,00	-0,80	<i>s10a3009 - Possui ou aluga ferro elétrico?</i>			
	Vela ou lamparina	0,01	-0,90	-0,64	Sim	0,02	<u>0,09</u>	<u>0,03</u>
	Outro	0,00	<u>0,32</u>	-0,07	Nao	0,10	-0,42	-0,14
	<i>s01x23 - Qual o principal combustível usado para cozinhar?</i>				<i>s10a3010 - Possui ou aluga máquina de lavar?</i>			
	Electricidade	<u>0,29</u>	-0,60	<u>0,65</u>	Sim	0,09	<u>0,26</u>	-0,04
	Gas	<u>0,14</u>	<u>0,30</u>	-0,14	Nao	0,06	-0,19	<u>0,03</u>
	Querosene	0,00	<u>0,29</u>	<u>0,08</u>	<i>s10a3012 - Possui ou aluga TV a cores?</i>			
	Carvao ou lenha	<u>0,27</u>	-1,33	-0,45	Sim	0,00	<u>0,03</u>	<u>0,02</u>
	Outro	0,00	<u>0,32</u>	-0,01	Nao	0,05	-0,37	-0,22
	Nao usa	0,00	<u>0,29</u>	<u>0,08</u>	<i>s10a3018 - Possui ou aluga aparelho de som?</i>			
	<i>s01x24 - Qual o principal destino do lixo domiciliar?</i>				Sim	0,03	<u>0,13</u>	-0,02
	Coletado por serviço de limpeza	0,03	<u>0,11</u>	<u>0,07</u>	Nao	0,03	-0,14	<u>0,02</u>
Coletado indiretamente	0,01	-0,35	-0,15	<i>s10a3019 - Possui ou aluga aparelho de dvd?</i>				
Queimado ou enterrado	<u>0,21</u>	-0,94	-0,66	Sim	0,05	<u>0,16</u>	0,00	
Jogado em terreno baldio ou logradouro	0,01	-0,35	<u>0,04</u>	Nao	0,08	-0,25	0,00	
Jogado em rio, lago ou no mar	0,00	<u>0,31</u>	-0,10	<i>s10a3025 - Possui ou aluga chuveiro elétrico?</i>				
Outro	0,00	-0,12	<u>0,46</u>	Sim	0,01	<u>0,05</u>	<u>0,01</u>	
<i>s10a3028 - Possui ou aluga automóvel?</i>				Nao	0,05	-0,33	-0,10	
Sim	<u>0,10</u>	<u>0,23</u>	-0,04	<i>s10a3026 - Possui ou aluga micro-computador?</i>				
Nao	0,05	-0,10	<u>0,02</u>	Sim	<u>0,11</u>	<u>0,37</u>	-0,05	
<i>s10a3030 - Possui ou aluga moto?</i>				Nao	0,04	-0,14	<u>0,02</u>	
Sim	<u>0,17</u>	<u>0,38</u>	-0,35	<i>s10a3004 - Possui ou aluga micro-ondas?</i>				
Nao	0,02	-0,05	<u>0,04</u>	Sim	<u>0,11</u>	<u>0,34</u>	-0,10	
<i>s10a3031 - Possui ou aluga trator/caminhão?</i>				Nao	0,04	-0,14	<u>0,04</u>	
Sim	<u>0,49</u>	<u>1,73</u>	<u>0,82</u>	<i>s10a3002 - Possui ou aluga freezer?</i>				
Nao	0,01	-0,04	-0,02	Sim	0,04	<u>0,23</u>	-0,07	
<i>s10a3029 - Possui ou aluga bicicleta?</i>				Nao	0,01	-0,06	0,02	
Sim	0,10	<u>0,12</u>	-0,16	<i>s10a3027 - Possui ou aluga antena parabólica?</i>				
Nao	0,06	-0,08	<u>0,11</u>	Sim	0,01	0,05	0,02	
<i>s04a6 - Tem plano de saúde ou seguro saúde?</i>				Nao	0,01	-0,02	-0,01	
Sim	<u>0,41</u>	<u>0,78</u>	<u>0,46</u>					
Nao	0,09	-0,17	-0,10					
<i>s03a5 - Frequenta escola ou creche atualmente?</i>								
Sim, particular	<u>0,25</u>	<u>1,38</u>	-0,82					
Sim, municipal	0,09	-0,44	<u>0,26</u>					
Sim, estadual	0,10	-0,47	<u>0,28</u>					
Sim, federal	0,01	<u>1,05</u>	-0,63					
Não, já frequentou	0,04	<u>0,13</u>	-0,08					
Não, nunca	0,02	-0,21	<u>0,13</u>					

Fonte: Resultados da pesquisa.

Por fim, no indicador “(C) Saúde e Educação”, as duas variáveis listadas no Quadro 2 contribuem, cada uma, com 50% do valor da inércia total. Sendo que aqueles que responderam “sim” para a questão

“Tem cobertura de plano de saúde ou seguro saúde?” (s04a6) se destacam por apresentar uma contribuição de 41,2% de participação para a inércia total. Ou seja, o indicador “Saúde e Educação” é influenciado principalmente pelos respondentes que tem cobertura de plano de saúde ou seguro saúde. Contudo, a resposta categórica “Sim, particular” sobre a variável “Frequenta escola ou creche atualmente?” (s03a5) também expressa bem a inércia total deste indicador. Na dimensão 1, os respondentes que afirmaram frequentar a rede particular e a rede federal revelam coordenadas fortemente positivas, acompanhados daqueles que disseram possuir cobertura de plano de saúde ou seguro saúde. Destarte, a dimensão 1 distingue basicamente esses indivíduos, dos que frequentam a rede estadual ou municipal e os que não possuem cobertura de plano de saúde ou seguro saúde, os quais apresentam coordenadas negativas para a dimensão 1. Parece, portanto, que esse indicador procura isolar no estudo indivíduos que frequentam a rede particular e a rede federal, ficando esses dois grupos relativamente próximos entre si. Os que possuem plano de saúde ou seguro saúde também se distanciam dos demais. Para a dimensão 2, as contribuições curiosamente se repetem, reafirmando a situação indicada pela dimensão 1. Assim sendo, verificou-se certa aproximação entre os indivíduos que “frequentam creche ou escola na rede municipal” e os que “frequentam a rede estadual”, sendo o grupo mais próximo desses outros dois, o daqueles indivíduos que “nunca frequentaram escola ou creche”. Esses três grupos, juntamente com os indivíduos que “não frequentam creche ou escola, mas já frequentaram” estão associados com os que responderam não possuir cobertura de plano de saúde ou seguro saúde. Assim, seria razoável concluir que aqueles que frequentam a rede federal, a rede particular e os que possuem plano/seguro saúde encontram-se em situação socioeconômica mais favorável e, por isso, em termos de saúde e educação, poderiam ser considerados inelegíveis ao recebimento do benefício do PBF.

Uma vez que as informações de alimentos básicos são quantitativas – a quantidade de alimentos básicos consumidos – a construção de um indicador geral ocorreu mediante ao uso da Análise de Componentes Principais (ACP), que similarmente à AC, permite reduzir as dimensões originais dos dados em dimensões ortogonais. Em ACP busca-se estudar e reduzir a estrutura dos dados e não existe a necessidade de assumir uma distribuição normal. A Tabela 2 fornece os resultados da ACP. Obteve-se 30% para o componente 1, para o componente 2 cerca de 26%, para o componente 3 encontrou-se 24% e para o quarto componente 21%. Uma vez que os três primeiros componentes explicam 79% da variância total, optou-se por gerar um índice composto pela média ponderada (cujos pesos são representados pela porcentagem que cada componente explica da variância total) dos índices de cada um dos três componentes (PEROBELLI, 1999).

As variáveis referentes a “Alimentos básicos” representam especificamente a “Quantidade de... comprada, obtida ou produzida para consumo no domicílio?”. Portanto, o primeiro componente capta as direções semelhantes e opostas entre as variáveis aleatórias. Assim, de um lado, observa-se que o consumo de arroz é maior em famílias com maior consumo também de feijão. O consumo desses itens, sobretudo dentro do domicílio, é comum entre famílias que priorizam os alimentos básicos, que usualmente cozinham em casa e não costumam fazer suas refeições em restaurantes. Assim, os sinais negativos dos coeficientes das variáveis referentes ao consumo de óleo de soja e café podem indicar certa restrição ao consumo desses itens, o que afetaria negativamente o consumo do essencial (arroz e feijão).

**TABELA 2 – Resultados para Alimentos Básicos**

Variáveis	Componente 1	Componente 2	Componente 3	Componente 4
Arroz	0,67	0,18	0,15	0,70
Feijão	0,66	0,24		-0,71
Óleo de Soja	-0,13	0,74	-0,65	
Café	-0,30	0,60	0,74	
Proporção individual	0,30	0,26	0,24	0,21
Proporção acumulada	0,30	0,56	0,79	1,00

Fonte: Resultados da pesquisa

\* Os valores ocultos situam-se abaixo de 0,10.

Por seu turno, no segundo componente o consumo de óleo de soja e café é significativo e já o consumo de arroz e feijão se reduz. Desse modo os resultados parecem sinalizar que esse padrão de

consumo representa famílias que tomam café da manhã em casa, e com certa frequência realizam suas refeições em restaurantes ou no trabalho, e que ainda consomem alimentos pré-preparados (o que justifica o consumo do óleo de soja). Já o terceiro componente apresenta coeficientes que indicam um alto consumo somente de café, podendo indicar que as famílias com esse padrão de consumo não realizam suas refeições em seus domicílios.

#### 4.2 Análise discriminante dos grupos de indivíduos

Essa seção apresenta um conjunto de resultados gerados pela Análise Discriminante sobre dois grupos previamente definidos na amostra desta pesquisa: um grupo formado por indivíduos que recebem o benefício do PBF, conforme o critério adotado para torná-lo elegível ao programa; e aqueles que não recebem, conforme sua renda *per capita* ou por se localizar em um município com cotas que o restringe do acesso ao Bolsa Família. Deve-se esclarecer que a inscrição das famílias com renda elegível ao PBF não garante a entrada imediata no programa. A seleção das famílias é feita por um sistema informatizado, a partir dos dados que elas informam no Cadastro Único. A concessão do benefício depende de quantas famílias já foram atendidas no município, em relação à estimativa de famílias pobres feita para essa localidade. Além disso, o governo federal precisa respeitar o limite orçamentário do programa (BRASIL, 2016b). Portanto, essa questão orçamentária representa um limitador para esta pesquisa em questão, pois a base de dados utilizada não identifica o município dos indivíduos, cuja informação ausente poderia ser utilizada para filtrar os possíveis efeitos oriundos dessa cota municipal. Em suma, pode haver famílias elegíveis que não estão recebendo o benefício por falta de recursos. De qualquer maneira, ao apresentar uma função discriminante pelas características identificadas dos indivíduos listados na base de dados, este artigo oferece uma contribuição ao apontar se o critério previamente estabelecido do Bolsa Família revela uma taxa de erro aparente relativamente grande. Tal assertiva, se confirmada, pode auxiliar o planejador público no desenvolvimento de uma estratégia de revisão do critério de seleção pelo PBF, justamente por incluir outras características socioeconômicas, e não somente a renda *per capita*.

Para avaliar a qualidade dos resultados da análise discriminante é preciso inicialmente observar se as médias dos dois grupos prévios possuem distâncias estatisticamente significativas. Para isso, tradicionalmente existem o teste *t* de *student* bem como o “Lambda de Wilks”. O propósito deste procedimento é averiguar se os escores médios dos dois grupos são próximos, o que representaria possivelmente uma alta sobreposição entre ambos, comprometendo a qualidade dos resultados gerados pela referida técnica multivariada. Os testes foram satisfatórios e as médias dos grupos são estatisticamente diferentes<sup>3</sup>. Em seguida, estimou-se a função discriminante canônica de Fisher, cujo método prescinde de uma distribuição normal dos dados (JOHNSON; WICHERN, 2007)<sup>4</sup>. Uma vez que o trabalho acompanhou, em alguma medida, o conceito multidimensional da pobreza para avaliar os indivíduos da amostra, selecionando previamente as variáveis, essa pesquisa não aplicou os métodos de *stepwise*<sup>5</sup> em análise discriminante. Não seria de interesse da pesquisa de descartar possivelmente grande parte das variáveis selecionadas, mesmo com critério puramente estatístico, pois provocaria a perda de informações importantes dos indivíduos em ambos os grupos e poderia comprometer o poder de discriminação do método canônico de Fisher.

Após essas considerações e cuidados para o uso da AD, a Figura 1 fornece as matrizes de confusão pelo método de Lachenbruch, justamente por corrigir o viés das taxas de erro aparente pelo método tradicional (empírico). Entretanto, o método tradicional tende a subvalorizar as probabilidades de classificação incorreta, cujo problema é tratado pelo método Lachenbruch. De acordo com a ilustração, existem duas classificações, denominadas de *A* e *B*. Apesar desses resultados serem bastante próximos, a diferença entre eles está nas variáveis utilizadas para o cálculo. Em “*A*” foram utilizadas todas as variáveis, o que acarretou em uma redução no número de indivíduos devido às informações faltantes em

<sup>3</sup> Em ambos os testes, houve diferença estatística entre os escores médios dos dois grupos ao nível de significância de 1%, com *t* de Hotelling registrando um valor de 0,1501 ( $F = 313,64$ ) e Lambda de Wilks atingindo um valor de 0,8695.

<sup>4</sup> Realizou-se uma avaliação da distribuição dos indicadores construídos com auxílio de ferramentas gráficas e testes estatísticos (i.e., histograma, gráfico de Kernel, gráficos Q-Q, teste de Shapiro-Wilk (1956) e teste de simetria e curtose). Mesmo transformando as variáveis originais, a distribuição assimétrica permaneceu para a maioria das variáveis..

<sup>5</sup> Sobre os métodos de seleção de variáveis na análise discriminante, veja Mingoti (2007) e Johnson e Wichern (2007)

certas observações (*missings*), especialmente nas variáveis referentes à saúde e educação. Quando retiradas tais variáveis em “B”, os resultados são próximos e, conforme recomendações de Johnson e Wichern (2007), a estimação da matriz de covariância conjunta torna-se melhor com um maior número de observações. Conforme a classificação B, observa-se que, de um total de 14.636 entrevistados, 5.889 deles recebem o benefício, representando 40,2% da amostra desta pesquisa<sup>6</sup>.

**FIGURA 1 – Matrizes de classificações**

Classificação A				Classificação B			
Recebe PBF	sim	não	Total	Recebe PBF	sim	não	Total
sim	2995 74,54%	1023 25,46%	4018 100%	sim	4365 74,12%	1524 25,88%	5889 100%
não	2322 39,99%	3485 60,01%	5807 100%	não	3538 40,45%	5209 59,55%	8747 100%
Total	5317 54,12%	4508 45,88%	9825 100%	Total	7903 54,00%	6733 46,00%	14636 100%

Fonte: resultados da pesquisa

Por meio da função de classificação dos dois grupos, como podem ser visualizadas na Tabela 3, o escore de cada beneficiário do Bolsa Família foi avaliado. Assim, do total dos beneficiários, cerca de 25,88% estariam classificados incorretamente neste grupo, conforme suas características socioeconômicas. Essa taxa de erro aparente é relativamente menor quando comparada com a do outro grupo, aqueles indivíduos que não recebem o Bolsa Família. Ou melhor, dos 8.747 indivíduos que não são beneficiários, 40,45% deveria receber o benefício se considerados outros indicadores, além da renda *per capita*. Em números absolutos, 3.538 entrevistados se encontram em situação característica de pobreza e não recebem o PBF, enquanto 1.524 dos indivíduos entrevistados recebem o benefício mesmo em condições socioeconômicas que os classifica no grupo que não deveria receber. Assim, de acordo com os resultados encontrados, 54% dos entrevistados são classificados como em situação que os tornaria beneficiários do programa pelos critérios propostos.

**TABELA 3 – Funções de classificação e de discriminação para cada índice**

Variáveis	Função de classificação		Função de discriminação	
	Sim	Não	Não padronizada	Padronizada
ed1	-0,44	-0,06	0,48	0,17
ed2	-0,39	-0,65	-0,32	-0,06
mt2	-4,93	-2,59	2,95	0,25
mt1	-3,12	-2,05	1,35	0,19
e11	-3,82	-2,21	2,04	0,30
e12	2,36	2,31	-0,05	0,00
perkpta	0,00	0,01	0,00	0,77
Constante	-0,53	-0,92	-0,57	-

Fonte: Resultados da pesquisa.

Nota: Lambda de Wilks: 0.8695  $p < 0,001$ .

<sup>6</sup> A expansão amostral foi aplicada na análise de correspondência e análise de componentes principais, cujo objetivo foi construir os indicadores característicos.

Portanto, ao avaliar brevemente as taxas de classificação corretas e incorretas dos dois grupos da amostra, as mesmas parecem sugerir que o critério de seleção dos indivíduos ao PBF pela renda *per capita* apresenta uma probabilidade de erro relativamente menor, mesmo considerando as demais características dos indivíduos. Tal assertiva torna-se mais saliente se essa probabilidade de classificação incorreta for confrontada com aquela do grupo de pessoas não beneficiadas pelo programa social. Na verdade, os resultados sugerem que mais pessoas poderiam receber o benefício se o nível de pobreza for qualificado por outros atributos socioeconômicos em uma perspectiva multidimensional. Em outras palavras, quando se amplia o vetor de variáveis características, o PBF parece ser mais falho ao excluir indivíduos do que incluí-los como beneficiários, desconsiderando a cota municipal.

Reportadas na Tabela 3, as variáveis representam os índices criados a partir da Análise de Correspondência para estrutura domiciliar (ed1 e ed2), modos de transporte (mt1 e mt2) e eletrodomésticos (el1 e el2). A variável “perkpta” representa o rendimento *per capita* dos indivíduos dentro de cada família entrevistada. Os índices gerados para o indicador “Saúde e Educação” e para “Alimentos Básicos” não foram apresentados porque não foram considerados no resultado da Análise Discriminante. Para identificar a importância relativa de cada uma das variáveis características na classificação de um novo indivíduo em um dos grupos, comumente usa-se a função discriminante ou uma das funções de classificação. Por existirem apenas dois grupos, então haverá uma função discriminante, que classifica o novo indivíduo, dependendo se o escore médio se situa à direita ou à esquerda do ponto médio (intercepto). Ou melhor, a regra matemática ou de classificação é necessária para decidir qual é o grupo mais provável que o novo indivíduo pertencerá. Esse processo fundamenta-se na probabilidade de pertencer a um dos grupos de modo a minimizar o custo de classificação incorreta.

Entretanto, para interpretar quais variáveis são as mais importantes na função discriminante, Johnson e Wichern (2007) recomendam a padronização dos coeficientes (cargas discriminantes). Assim, todos os coeficientes são transformados para o intervalo [-1,1], facilitando a comparação dos coeficientes e a interpretação da função (MINGOTI, 2007). No caso da pesquisa em questão, verifica-se que a variável mais importante para a discriminação dos indivíduos é a renda *per capita* (0,77). O que parece confirmar a importância de levar em consideração este aspecto para definir a elegibilidade do indivíduo ao recebimento do benefício do PBF. No entanto, os coeficientes gerados para os índices dos indicadores “eletrodomésticos” (el1 – 0,30) e “modos de transporte” (mt1 – 0,19 e mt2 – 0,25) são consideráveis de maneira secundária. Se o objetivo for ampliar a avaliação das condições socioeconômicas dos indivíduos beneficiários ou aqueles potencialmente beneficiários do PBF, essas variáveis poderiam ser consideradas como um critério complementar.

Chama a atenção a variável ed2, que exibe um coeficiente padronizado com sinal negativo. Como mencionado na seção anterior, essa variável capta principalmente as pessoas que usam essencialmente a “eletricidade” para cozinhar, cuja forma de energia é relativamente cara para os indivíduos pobres e as famílias típicas no Brasil. Certamente, famílias com um maior padrão de vida em seus domicílios usam a eletricidade para cozinhar e, portanto, não as classificam como pobres e nem para serem beneficiadas ao PBF. Cabe mencionar que as funções de classificação, correspondentes a cada grupo da amostra, são uma forma prática e direta para classificar um novo indivíduo em um dos grupos. Ao utilizá-las, caso o escore do indivíduo for maior no grupo que recebem o Bolsa Família, então ele deveria também ser beneficiário, e vice-versa. Assim, esse indivíduo estaria mais próximo à média do grupo e, portanto, ele apresentaria maior (menor) probabilidade de classificação correta (incorreta).

## 5 Considerações finais

Este trabalho pretendeu oferecer contribuições acerca do critério e seleção do Programa Bolsa Família (PBF) a partir das características socioeconômicas dos indivíduos da região Sul e Sudeste em 2009, não somente a renda *per capita*. Essa preocupação tem sido recorrente nos debates desse programa de transferência de renda. Para tanto, adotou-se como premissa subjacente na seleção das variáveis o conceito de pobreza multidimensional da PNUD. A análise oferecida por essa pesquisa residiu sobre os entrevistados constantes da AIBF II e, portanto, trata-se de um estudo inédito nessa variante e por aplicar técnicas estatísticas exploratórias de classificação e associação. Esse trabalho utilizou dados relacionados à saúde, educação e condições de vida dos entrevistados; quanto aos domicílios das famílias foi analisado



o perfil por tipo de utilidade pública (acesso à água encanada, iluminação e coleta de lixo). Mais especificamente, após a elaboração dos indicadores socioeconômicos dos entrevistados pelas análises de correspondência múltipla (AC) e de componentes principais (ACP), os mesmos foram utilizados para estimar a função discriminante canônica de Fisher, fornecendo as probabilidades de classificação correta e incorreta dos indivíduos que recebem e não o benefício do programa.

Além disso, o objetivo desta pesquisa está em consonância com as discussões econômicas e sociais se confirma com a divulgação do resultado de uma espécie de auditoria no programa que teve início em junho de 2016, a pedido do MDSA. Os resultados foram obtidos após o cruzamento de dados com 6 bases do governo federal, dentre elas: Cadastro Geral de Empregados e Desempregados (Caged); Relação Anual de Informações Sociais (Rais); Sistema de Controle de Óbitos (Sisobi); Instituto Nacional do Seguro Social (INSS); Sistema Integrado de Administração de Recursos Humanos (Siape); e Cadastro Nacional de Pessoas Jurídicas (CNPJ). Com essa ação, o Governo encontrou irregularidades em 1,1 milhão de benefícios do Bolsa Família (PORTAL BRASIL, 2016). Portanto, a presente pesquisa vem colaborar com o debate em curso sobre o PBF e contribuir para os formuladores de políticas públicas e gestores dos programas de transferência de renda, para as discussões e ajustes de estratégias que reforcem a efetividade dos critérios estabelecidos. Além de sugerir uma ferramenta a mais, que não a simples apuração da renda *per capita*, para auxiliar na seleção e manutenção das famílias no programa.

A função discriminante estimada pelas características socioeconômicas permite indicar se o critério previamente estabelecido do Bolsa Família revela uma taxa de erro aparente relativamente grande, cuja taxa pode sinalizar uma provável revisão do mesmo, considerando, talvez, outras características, além da renda *per capita*. Observou-se que o critério de inclusão dos indivíduos ao Bolsa Família pela renda *per capita* apresenta uma probabilidade de erro relativamente menor, mesmo ao considerar outras características de aspecto multidimensional da pobreza. Em termos percentuais, de acordo com a análise realizada, 25,88% dos beneficiários entrevistados não deveria receber o benefício do PBF. Por outro lado, dos 8.647 indivíduos que não são beneficiários, 40,45% deveria receber o benefício segundo sua classificação na análise realizada. Tais resultados sugerem que mais pessoas poderiam receber o benefício se fossem consideradas as demais características dos indivíduos. Em suma, quando se amplia o vetor de variáveis características, o PBF parece ser mais falho ao excluir indivíduos do que incluí-los como beneficiários, desconsiderando a cota municipal.

Cabe ressaltar que a inscrição das famílias com renda elegível ao PBF não garante a entrada imediata no programa. A seleção das famílias é feita por um sistema informatizado, a partir dos dados que elas informam no Cadastro Único. A concessão do benefício depende de quantas famílias já foram atendidas no município, em relação à estimativa de famílias pobres feita para essa localidade. Além disso, o governo federal precisa respeitar o limite orçamentário do programa (BRASIL, 2016b). Como forma de avanço de pesquisa, um futuro estudo poderia realizar uma análise em que fosse possível anular a limitação orçamentária citada acima. Para tal, seria necessária uma base de dados que permitisse identificar os municípios a que cada indivíduo pertence. A Pesquisa Nacional por Amostra de Domicílios (PNAD) é uma boa fonte, porém possui uma infinidade de dados que deveriam passar por um refinado tratamento. Outro tipo de pesquisa interessante seria o de avaliar se a situação socioeconômica de quem recebe o benefício melhorou nos últimos anos e se tem se conseguido quebrar o ciclo da pobreza, ou se essa pobreza tem passado de geração para geração.

## Referências

- ARRETCHE, M. Federalismo e Relações Intergovernamentais no Brasil: A Reforma de Programas Sociais. **DADOS – Revista de Ciências Sociais**, Rio de Janeiro, vol. 45, n. 3, pp. 431 – 458, 2002.
- BRASIL. Ministério do Desenvolvimento Social e Agrário. **Programa Bolsa Família**. Disponível em: < <http://www.mds.gov.br/bolsafamilia> >. Acesso em 27, novembro, 2016b.
- BRASIL. Ministério do Desenvolvimento Social e Combate à Fome. **Programa Bolsa Família**. Disponível em: < <http://www.mds.gov.br/bolsafamilia> >. Acesso em 18, abril, 2016a.

- CAMARGO, C. F.; *et al.* Perfil socioeconômico dos beneficiários do Programa Bolsa Família: o que o Cadastro Único revela? In: CAMPELLO, T.; NERI, M.C. (org.). **Programa Bolsa Família: uma década de inclusão e cidadania**. Brasília: IPEA, 2013, pp. 157 - 177.
- CAMPELLO, T.; NERI, M. C.(Org.). **Programa Bolsa Família: uma década de inclusão e cidadania**. Brasília: Ipea, 2013.
- CARDOSO, D. F. **Capital e trabalho no Brasil no século XXI: O impacto de políticas de transferência e de tributação sobre desigualdade, consumo e estrutura produtiva**. 2016. 279 f. Tese (Doutorado em Economia) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.
- CZERMAINSKI, A. B. C.; DIAS, C. T. S. Análise de Correspondência. In: **Seminário apresentado na disciplina Análise Multivariada**, julho, 2004, Piracicaba, SP. Universidade de São Paulo, Escola Superior de Agricultura “Luiz de Queiroz”, 2004.
- DRAIBE, S. M. A construção institucional da política brasileira de combate à pobreza: perfis, processos e agenda. **Caderno de Pesquisa n. 34**. Campinas: NEPP/UNICAMP, 1998.
- HOFFMANN, R. “**Não pode haver dúvida de que programas de transferência de renda contribuem para reduzir a pobreza**”. 14, jul, 2014. Disponível em: <<http://www.ihu.unisinos.br>>. Acesso em 29, abril, 2016. Entrevista concedida ao Instituto Humanitas Unisinos.
- INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). Comunicados do Ipea. **A Década Inclusiva (2001-2011): Desigualdade, Pobreza e Políticas de Renda**. Brasília, n. 155, 2012.
- JANNUZZI, P. M.; PINTO, A. R. Bolsa Família e seus impactos nas condições de vida da população brasileira: uma síntese dos principais achados da Pesquisa de Avaliação de Impacto do Bolsa Família II. In: CAMPELLO, T.; NERI, M.C. (org.). **Programa Bolsa Família: uma década de inclusão e cidadania**. Brasília: IPEA, 2013, pp. 179 - 192.
- JANNUZZI, P. M.; PINTO, A. R. Bolsa Família e seus impactos nas condições de vida da população brasileira: uma síntese dos principais achados da Pesquisa de Avaliação de Impacto do Bolsa Família II. In: CAMPELLO, T.; NERI, M.C. (org.). **Programa Bolsa Família: uma década de inclusão e cidadania**. Brasília: IPEA, 2013, pp. 179 - 192.
- JOHNSON, R.; WICHERN, D. **Applied Multivariate Statistical Analysis**. 5th. ed. Englewood Cliffs: Prentice-Hall, 2007.
- LAVINAS, L. Universalizando direitos. **Revista Observatório de cidadania** – relatório 2004: medos e privações – obstáculos à segurança humana. Rio de Janeiro: IBASE, março de 2004. pp. 67 – 74.
- LOESCH, C.; HOELTGEBAUM, M. **Métodos Estatísticos Multivariados**. São Paulo: Saraiva, 2012.
- MAGALHÃES, R. *et al.* A implementação do programa Bolsa Família: as experiências de São Francisco de Itabapoana e Duque de Caxias. **Ciência & Saúde Coletiva**, Rio de Janeiro, vol. 12, n. 6, pp. 1513 – 1524, 2007.
- MANLY, B. F. J. **Métodos Estatísticos Multivariados: uma introdução**. Tradução Sara Ianda Carmona. 3ª edição. Porto Alegre: Bookman, 2008.
- MARQUES, R. M. *et al.* A importância do Bolsa Família nos Municípios Brasileiros. In: VAITSNAN, J.; PAES-SOUSA, R. (Org.). **Avaliação de Políticas e Programas do MDS – Resultados**. Volume 2 – Bolsa Família e Assistência Social. Brasília, DF: MDS; SAGI, 2007.
- MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2007.
- MINISTÉRIO DO DESENVOLVIMENTO SOCIAL E COMBATE À FOME. **Ficha Técnica da Avaliação de Impacto do Programa Bolsa Família: segunda rodada**. Disponível em: <<http://www.mds.gov.br/gestaodainformacao/avaliacao/>>. Acesso em 23, abr, 2016.

- NASCIMENTO, A. F.; REIS, C. N. Os Programas de Transferência Condicionada de Renda na América Latina: especificidades de uma realidade que se mantém. **Revista de Políticas Públicas**, São Luís, v. 13, n. 2, p. 183 – 193, jul./dez., 2009.
- NASCIMENTO, A. F.; REIS, C. N. Os Programas de Transferência Condicionada de Renda na América Latina: especificidades de uma realidade que se mantém. **Revista de Políticas Públicas**, São Luís, v. 13, n. 2, p. 183 – 193, jul./dez., 2009.
- NETO, V. R. P. **Transferências Condicionais de Renda e Nutrição: Uma avaliação do Programa Bolsa Família nas áreas rurais e urbanas do Brasil**. 2014. 33 f. Dissertação (Mestrado em Economia Social) – Fundação Getúlio Vargas, Rio de Janeiro, 2014.
- PEROBELLI, F.; OLIVEIRA, A.; NOVY, L.; FERREIRA, M. Planejamento Regional e Potenciais de desenvolvimento dos municípios de Minas Gerais na região em torno de Juiz de Fora: Uma aplicação de análise fatorial. **Revista Nova Economia**, Belo Horizonte, v. 9; nº 1, pp. 121 – 150, junho, 1999.
- PORTAL BRASIL. **Governo encontra irregularidades em 1,1 milhão de benefícios do Bolsa Família**. 07 nov. 2016. Disponível em: < <http://www.brasil.gov.br/cidadania-e-justica/2016/11/governo-encontra-irregularidades-em-1-1-milhao-de-beneficios-do-bolsa-familia> >. Acesso em: 20, novembro, 2016.
- PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. **Brasil mantém tendência de crescimento do IDH e reduz pobreza multidimensional, diz Relatório de Desenvolvimento Humano 2015**. 16 dez. 2015. Disponível em: <<http://www.pnud.org.br>>. Acesso em: 15, maio, 2016.
- PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. **Indicador avalia nova dimensão da pobreza**. 04 nov. 2010. Disponível em:<<http://www.pnud.org.br>>. Acesso em: 15, maio, 2016.
- SENN, M. C. M; *et al.* Programa Bolsa Família: nova institucionalidade no campo da política social brasileira? **Revista Katál**, Florianópolis, v. 10, n. 1, pp. 86 – 94, jan./jun.2007.
- SENN, M. C. M; *et al.* Programa Bolsa Família: nova institucionalidade no campo da política social brasileira? **Revista Katál**, Florianópolis, v. 10, n. 1, pp. 86 – 94, jan./jun.2007.
- SOARES, F. V. Para onde caminham os Programas de Transferência Condicionada? As experiências comparadas do Brasil, México, Chile e Uruguai. In: CASTRO, J.A.; MODESTO, L. (org.). **Bolsa Família 2003- 2010: avanços e desafios**. Brasília: IPEA, 2010, pp. 173 – 200.
- SOARES, S.; *et al.* Os impactos do benefício do Programa Bolsa Família sobre a desigualdade e a pobreza. In: CASTRO, J.A.; MODESTO, L. (org.). **Bolsa Família 2003- 2010: avanços e desafios**. Brasília: IPEA, 2010, pp. 27 - 52
- SOUZA, P.H.G.F.; OSORIO, R.G. O perfil da pobreza no Brasil e suas mudanças entre 2003 e 2011. In: CAMPELLO, T.; NERI, M.C. (org.). **Programa Bolsa Família: uma década de inclusão e cidadania**. Brasília: IPEA, 2013, pp. 139 – 155.
- SOUZA, P.H.G.F.; OSORIO, R.G. O perfil da pobreza no Brasil e suas mudanças entre 2003 e 2011. In: CAMPELLO, T.; NERI, M.C. (org.). **Programa Bolsa Família: uma década de inclusão e cidadania**. Brasília: IPEA, 2013, pp. 139 – 155.
- VIEIRA, M. D. T.; FREGUGLIA, R. S. **Análise e modelagem longitudinal dos dados da pesquisa de avaliação de impacto do Programa Bolsa Família (primeira e segunda rodadas)** Marcel Vieira e Ricardo Freguglia. Projeto de Pesquisa CNPQ. Juiz de Fora. 2015.