

Analfabetismo Isolado e Acesso a Programas Sociais no Brasil

Fernanda Dachi Carrets¹
Gibran da Silva Teixeira²
Felipe Garcia Ribeiro³

Resumo

Este trabalho investiga se a situação de analfabetismo isolado em domicílios brasileiros gera uma barreira ao conhecimento sobre o Cadastro Único do Governo Federal para indivíduos analfabetos. Desta forma, a proposta do estudo é checar se o fenômeno de analfabetismo isolado é capaz de reforçar as indesejadas exclusões indevidas de políticas públicas no Brasil. Com informações sobre o Cadastro Único e os Programas Sociais do Governo Federal contidas na PNAD 2014, a estratégia empírica do método de *Propensity Score Matching* combinada do algoritmo de Imbens (2015) encontrou evidências de que indivíduos analfabetos isolados têm em média uma menor probabilidade de conhecer o Cadastro Único do que indivíduos analfabetos não isolados. Os coeficientes encontrados para o Brasil, com e sem distinção de gêneros, foram, em média, de -9,06 a -13,03 pontos percentuais. Destacamos que o maior efeito foi encontrado para o gênero feminino. Estes resultados são robustos ao teste de análise de sensibilidade dos Limites de Rosenbaum.

Palavras-chave: Analfabetismo Isolado, Cadastro Único, *Propensity Score Matching*.

Área temática: Econometria

Classificação JEL: C31; I2

Abstract

This paper investigates whether the isolated illiteracy situation in Brazilian households generates a barrier to knowledge about the Cadastro Único for illiterate individuals. Thus, the purpose of the study is to verify if the phenomenon of isolated illiteracy is capable of reinforcing the undesired undue exclusion of public policies in Brazil. With information on the Cadastro Único and the Social Programs of the Federal Government contained in the PNAD 2014, the empirical strategy of the combined Propensity Score Matching method of the Imbens (2015) algorithm has found evidence that isolated illiterate individuals are on average less likely to know the Cadastro Único than non-isolated illiterate individuals. The coefficients found for Brazil, with and without distinction of genus, were, on average, from -9.06 to -13.03 percentage points. We emphasize that the greatest effect was found for the female gender. These results are robust to the sensitivity test of the Rosenbaum Limits.

Key words: Isolated Illiteracy, Cadastro Único, Propensity Score Matching.

Thematic area: Econometrics

JEL Classification: C31; I2

¹ Doutoranda em Economia Aplicada pelo Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Sul. E-mail: fe.dachi@gmail.com

² Professor Adjunto do Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande. E-mail: tgibran@hotmail.com

³ Professor Adjunto do Programa de Pós-Graduação em Organizações e Mercados da Universidade Federal de Pelotas. E-mail: felipe.garcia.rs@gmail.com

1. Introdução

O presente estudo investiga qual a relação do analfabetismo isolado com o acesso à informação sobre a existência do Cadastro Único do Governo Federal. Em última instância, o que se está propondo checar neste estudo é se o fenômeno de analfabetismo isolado é capaz de reforçar as indesejadas exclusões indevidas de políticas públicas. Este fato se daria em função da menor informação, ou mesmo do acesso a ela, visto que residir em um domicílio com apenas indivíduos analfabetos pode impor uma série de restrições ligadas ao acesso de informações e serviços.

O conceito de analfabetismo isolado foi apresentado, pela primeira vez, junto com a teoria de externalidades da alfabetização exposta por Basu e Foster (1998). Segundo eles, indivíduos analfabetos se apropriam do acesso às atividades que demandam o uso da escrita e leitura quando têm proximidade de indivíduos alfabetizados. Em função do conceito proposto pelos autores surgiram os termos Analfabeto Não Isolado (ANI) e Analfabeto Isolado (AI)⁴. O primeiro refere-se às pessoas analfabetas que residem em um domicílio com pelo menos uma pessoa alfabetizada, e o segundo representa os analfabetos que residem em um domicílio sem pessoas alfabetizadas.

De acordo com o 11º Relatório de Monitoramento Global de Educação para Todos de 2013 da UNESCO que avaliou a situação de 150 países ainda havia 774 milhões de adultos analfabetos no mundo. Entre os anos de 2005 a 2011 o Brasil estava entre os 10 países com a maior população de adultos analfabetos no mundo. Esta informação é corroborada pelos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2014, visto que a taxa de analfabetismo das pessoas de 15 anos ou mais de idade foi 8,3%, o que representa aproximadamente 13,2 milhões de pessoas.

Segundo o Ministério do Desenvolvimento Social e Agrário (MDSA) um dos grandes desafios para as políticas públicas no Brasil é alcançar a população considerada invisível, ou seja, aquelas pessoas elegíveis que não têm acesso aos serviços públicos e de proteção social. A população definida como invisível pelo MDSA, conhecida na literatura pelo termo “*hard-to-reach*”, descreve aqueles subgrupos da população que não são atingidos pelas redes de proteção social devido a desvantagens relacionadas a fatores geográficos, sociais e econômicos (FREIMUTH; METTGER, 1990). Sendo assim, as populações de difícil acesso são caracterizadas por serem socialmente invisíveis dado um conjunto de desvantagens, como o analfabetismo, e originando a criação de barreiras para o acesso aos seus direitos (SHAGHAGH; BHOPAL; SHEIKH, 2011).

O Cadastro Único (CadÚnico) é o instrumento utilizado para a identificação e caracterização socioeconômica das famílias brasileiras de baixa renda pelo Governo Federal, e empregado para seleção de beneficiários e integração de programas sociais no Brasil. No qual, é considerado de baixa renda aqueles indivíduos que possuem renda familiar *per capita* mensal de até ½ salário mínimo ou que possua renda familiar mensal de até três salários mínimos⁵. O CadÚnico vem se consolidando ao longo dos anos como uma ferramenta capaz de ampliar o número de cadastros de possíveis beneficiários, com abrangência nacional, e ajudar na manutenção de políticas públicas direcionadas ao público mais carente (DIREITO *et al*, 2016; PAES-SOUSA, 2013).

Baseados no estudo de Basu e Foster (1998) os trabalhos subsequentes seguiram por duas correntes de pesquisa. Na primeira o intuito foi aprimorar a medida tradicional de alfabetização com a inclusão dos conceitos de externalidades, analfabetismo isolado e analfabetismo não isolado. Já na segunda corrente, na qual o presente estudo está, o objetivo é avaliar empiricamente se, de fato, existem externalidades sociais da alfabetização. Uma vez que é reconhecido na literatura os retornos privados da alfabetização (FERRER; GREEN; RIDDELL, 2006; De BALDINI ROCHA; PONCZECK, 2011), ainda assim, os benefícios da alfabetização podem ultrapassar os ganhos privados e gerar externalidades para as pessoas analfabetas.

Como ferramenta de análise empírica utilizamos o *Propensity Score Matching* (PSM), e examinaremos o efeito do analfabetismo isolado sobre o conhecimento do Cadastro Único para indivíduos com renda *per capita* domiciliar de até ½ salário mínimo. Os dados utilizados são da Pesquisa Nacional

⁴Os termos analfabetos isolados (AI) e analfabetos não isolados (ANI) são uma livre tradução dos termos *Proximate Literate* e *Isolated illiterate* (BASU e FOSTER, 1998) feita por Ribeiro e Souza (2013).

⁵ Definição dada pelo Art. 4º do Decreto nº 6.135, de 26 de junho de 2007. Sendo a renda familiar mensal: a soma dos rendimentos brutos auferidos por todos os membros da família moradores de um mesmo domicílio.

por Amostra de Domicílios (PNAD) de 2014, os quais possuem informações sobre o Cadastro Único e os programas sociais do Governo Federal.

Os resultados que obtivemos indicam que o status de analfabetismo isolado atua no sentido de reduzir a probabilidade desses indivíduos conhecerem o Cadastro Único em relação a indivíduos analfabetos não isolados. Uma vez que esses indivíduos se encaixam no critério de renda para elegibilidade do Cadastro Único, e não tem conhecimento do mesmo, possivelmente, podem acabar fazendo parte da população considerada invisível que ainda vive fora de qualquer rede de proteção social dada a situação de analfabetismo isolado.

O presente artigo está estruturado em mais cinco seções além desta introdução. Na segunda seção, apresentamos a revisão de literatura que sintetiza os estudos que investigam a existência de externalidades da alfabetização, na terceira seção apresentamos uma breve discussão sobre o Cadastro Único além da questão da focalização de políticas públicas no Brasil sobre o ponto da exclusão indevida de indivíduos. Na quarta seção, retratamos a estratégia metodológica utilizada e a seleção amostral. Já na quinta seção examinamos os resultados obtidos e, por fim, na sexta seção, apresentamos as considerações finais do estudo com o apanhado dos principais resultados e sugestões para trabalhos futuros.

2. Externalidades da alfabetização

Numa literatura mais recente, pesquisadores passaram a admitir que os benefícios da alfabetização podem exceder os ganhos privados e gerar externalidades positivas para pessoas analfabetas que residem em domicílios com pelo menos uma pessoa alfabetizada, uma vez que a presença de uma pessoa capaz de desempenhar atividades de leitura e escrita proporciona que a alfabetização seja vista como um bem público para os membros analfabetos no domicílio (BASU; FOSTER, 1998).

Estudo como de Basu, Narayan e Ravallion (2001) avaliou, com base na *equação minceriana*, se o salário de um trabalhador analfabeto é afetado pela alfabetização de outros membros do domicílio em Bangladesh. Foram utilizados dados do *Household Income and Expenditure Survey* (HIES) captados entre os anos de 1995 e 1996. Os autores constataram que os rendimentos dos homens analfabetos são em média 15% maiores se eles fazem parte de um domicílio com pelo menos um dos membros alfabetizados. Para as mulheres, os autores encontraram que uma trabalhadora analfabeta não isolada do convívio de indivíduos alfabetizados no domicílio ganha cerca do dobro do salário de uma analfabeta que vive em domicílio onde todos os outros membros também são analfabetos. E ainda, as mulheres analfabetas solteiras apresentaram maiores retornos salariais, dado o fato de residir no mesmo domicílio que pessoas alfabetizadas, do que mulheres analfabetas casadas. Basu, Narayan e Ravallion (2001) também sugerem que os diferentes efeitos da alfabetização sobre o rendimento de homens e mulheres implicam que as mulheres podem ser mais eficientes como receptoras de externalidades de alfabetização no domicílio do que os homens.

Segundo Iversen e Palmer-Jones (2008) existe a possibilidade de que os efeitos encontrados para a externalidade da alfabetização no mercado de trabalho sejam contaminados por problemas de seleção, e características não observáveis tanto dos indivíduos analfabetos quanto dos indivíduos alfabetizados. Para isso, eles utilizaram a equação de salários minceriana e os dados do *Bangladesh Household Income and Expenditure Survey* (HIES) de 1995/6 e 2000/1 para explorar a robustez e discutir alguns resultados encontrados por Basu, Narayan e Ravallion (2001) com relação a seleção no casamento quanto ao nível educacional dos cônjuges. Segundo os autores, a seleção no casamento levanta a possibilidade de que o maior retorno salarial feminino que Basu, Narayan e Ravallion (2001) encontraram possa refletir os atributos não observados de mulheres analfabetas casadas com homens alfabetizados ao invés de compartilhamento de alfabetização domiciliar.

Para o Senegal, Sarr (2004) mediu o efeito do compartilhamento da alfabetização, no ambiente domiciliar, para o retorno salarial no mercado de trabalho. Os resultados foram obtidos através da equação minceriana com base nos dados da *Enquête Sénégalaise Auprès des Ménages* (ESAM) para o ano de 1994. Os autores comprovaram que há efeitos positivos da externalidade da alfabetização direcionados aos ganhos salariais de analfabetos não isolados. Segundo eles, os trabalhadores analfabetos, que vivem em uma família com pelo menos uma pessoa alfabetizada, recebem salários mais altos do que trabalhadores analfabetos que

pertencem a uma família em que todos os membros são analfabetos. E além disso, os resultados sugerem que as mulheres são melhores destinatárias e geradoras da externalidade da alfabetização do que os homens.

No campo da saúde infantil, Gibson (2001) utilizou como indicador de saúde a estatura de crianças de 0 a 5 anos de idade de Papua Nova Guiné, e mediu o impacto da externalidade da alfabetização para relação altura-idade das crianças decorrente da presença de membros alfabetizados no domicílio. Segundo os autores, a altura retrata o estado nutricional das crianças, além de ser um reflexo dos hábitos e cuidados dos adultos para com as crianças. Desta forma, a alfabetização dos indivíduos adultos no domicílio aumenta a capacidade de informação sobre cuidados relacionados a saúde das crianças. Empregando o método de mínimos quadrados ordinários a partir de dados do *Papua New Guinea Household Survey* (PNGHS) para o ano de 1996, os autores, encontraram resultados que revelam um aumento significativo na altura média das crianças que tem convívio com adultos alfabetizados em comparação com crianças que não convivem com pessoas alfabetizadas.

O trabalho de Borroah (2009) empregou um modelo *probit* bivariado para analisar os determinantes do risco de crianças de zonas rurais na Índia sofrerem de má nutrição infantil utilizando dados do *National Council of Applied Economic Research* (NCAER) no ano de 1994. Os resultados, de maneira geral, reportaram a importância da presença de mulheres alfabetizadas no domicílio sobre a redução do risco de má nutrição infantil. Nos domicílios em que a mãe era alfabetizada foram encontrados melhores índices nutricionais entre as crianças do que quando as mães não eram alfabetizadas. No entanto, quando o membro alfabetizado do domicílio era o pai o efeito da externalidade da alfabetização não se mostrou significativo.

Ainda na área de saúde, Husain e Dutta (2012) investigaram os efeitos da externalidade alfabetização entre casais no uso de métodos contraceptivos na Índia. Os autores realizaram as estimações via modelo *logit* com base nos dados levantados pelo *Demographic Health Survey* (DHS) para os anos de 2005 e 2006. O estudo investigou a hipótese de que a adoção de métodos contraceptivos entre mulheres analfabetas com parceiros alfabetizados é maior do que com parceiros analfabetos. Os resultados reportaram a existência do efeito de externalidade da alfabetização relacionada a questão do planejamento familiar de uma pessoa alfabetizada para seu parceiro analfabeto, porém, segundo os autores, o efeito é restrito a alguns grupos específicos (dados características socioculturais) e varia de acordo com o método contraceptivo adotado entre o casal.

Para o Brasil, o estudo feito por Ribeiro *et al.* (2016) utilizou regressões de Poisson e investigou o retorno social da alfabetização, as externalidades, em relação ao tabagismo para analfabetos que residem com indivíduos alfabetizados. Com base nos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2008, os autores encontram efeitos positivos da alfabetização das mulheres, ou seja, a alfabetização delas serviu como fator de proteção contra o tabagismo para indivíduos analfabetos, de ambos os sexos, residentes no mesmo domicílio.

3. Cadastro Único e Focalização de Políticas Públicas

Criado em 2001, através do Decreto nº 3.877, de 24 de julho de 2001⁶, o Cadastro Único, também conhecido como CadÚnico, só começou a ser utilizado após a unificação dos programas de transferência de renda do Governo Federal em 2003, sendo o Programa Bolsa Família (PBF) o único a utilizá-lo até então. Ao longo do tempo diversos Programas Federais foram agregados ao Cadastro Único, e a criação do Plano Brasil Sem Miséria (PBSM)⁷ em 2011 também contribuiu para o aumento no número de programas. (DIREITO *et al.*, 2016).

O Cadastro Único vem ampliando suas atividades na participação de processos de seleção de beneficiários e manutenção de políticas públicas direcionadas a população carente. Em setembro de 2016, cerca de 27,7 milhões de famílias estavam inscritas no Cadastro Único, e dessas, aproximadamente 13,5 milhões são beneficiárias do Programa Bolsa Família⁸. A inscrição no Cadastro Único é um importante

⁶Revogado pelo Decreto nº 6.135, de 26 de junho de 2007.

⁷ Conforme Decreto nº 7.492, de 2 de junho de 2011 “o Plano Brasil Sem Miséria, tem a finalidade de superar a situação de extrema pobreza da população em todo o território nacional, por meio da integração e articulação de políticas, programas e ações”.

⁸ Segundo dados do Ministério do Desenvolvimento Social.

passo na construção de uma base de dados para o registro de potenciais beneficiários, e que poderá efetivar o acesso dos mesmos a diversos programas de assistência social no Brasil (PAES-SOUSA, 2013).

Segundo Direito *et al.* (2016) uma das estratégias do Plano Brasil Sem Miséria é a Busca Ativa que tem a pretensão de minimizar os erros de exclusão no cadastro e inserir as famílias de baixa renda que ainda não foram cadastradas. Desde a concretização do PBSM, do período de dezembro de 2011 a janeiro de 2016, foram incluídas mais de 7 milhões de novas famílias no Cadastro Único. No entanto, segundo o Ministério do Desenvolvimento Social e Agrário (MDSA), ainda existe uma parcela de potenciais beneficiários que não tem acesso aos serviços públicos e de proteção social. Essa parcela é definida como “população invisível” pelo MDSA, e é conhecida na literatura como “*hard-to-reach*”. O termo de “difícil acesso” é utilizado para descrever subgrupos da população que são difíceis de atingir e incluir em programas sociais devido à sua localização ou mesmo devido a situação social e econômica. Sendo assim, as populações de difícil acesso são caracterizadas por serem socialmente invisíveis dado um conjunto de desvantagens, como o analfabetismo, e originando a criação de grandes barreiras para o acesso aos seus direitos (FREIMUTH; METTGER, 1990; SHAGHAGH; BHOPAL; SHEIKH, 2011).

Para a redução das indesejadas exclusões indevidas, segundo Kerstenetzky (2009), uma das estratégias dos programas de renda garantida é o instrumento de focalização nos mais pobres e as condicionantes com intuito de aumentar habilidades futuras dos beneficiários. Para a autora há alguns fatores determinantes na ocorrência de indivíduos elegíveis que não recebem benefícios do Programa Bolsa Família (PBF), como a falta de acesso a informação, dado que muitas dessas pessoas têm baixo ou nenhum grau de instrução, muitas vezes, faz com que elas nem saibam que tem direito ao benefício. O emprego de ferramentas de focalização também é justificado por uma questão de melhorar a eficiência da alocação dos recursos, e, portanto, a focalização é vista como um instrumento capaz de ampliar os impactos sobre a população mais pobre fazendo uma melhor distribuição dos recursos disponíveis, e assim tornando o programa mais efetivo (COADY; GROSH; HODDINOTT, 2004).

Tavares *et al.* (2009), a partir de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2004, avaliaram a situação da focalização do Programa Bolsa Família e qual seu impacto sobre a desigualdade de renda e dos níveis de pobreza no Brasil. Os autores encontraram uma grande disparidade no grau de focalização entre os estados, o que gerou heterogeneidade nos impactos gerados pelo programa. Segundo eles, o sucesso de focalização do programa depende do orçamento disponível e da eficiência nos processos de seleção dos beneficiários. Outro fator que influencia é a magnitude da pobreza em cada estado, assim, estados com populações mais pobres tem maior facilidade em identificar o público-alvo do programa apresentando melhores índices de focalização.

Os autores também propõem avaliar se mudanças no processo de seleção dos beneficiários influenciam os indicadores de focalização do programa. Para isso, foram realizadas duas simulações de seleção: a primeira utilizou o mecanismo de seleção de beneficiários com o método de *proxy means-test*, mantendo fixo o orçamento do programa para o ano de 2004; e a segunda simulação definiu uma ampliação do programa de 2004 para 2006, mas mantendo o critério atual de seleção via renda declarada. Os resultados das simulações, segundo Tavares *et al.* (2009), sugerem que a combinação do método de seleção via *proxy means-test*, haja vista, a expansão do programa aplicada na segunda simulação poderia obter melhores resultados de focalização.

Soares *et al.* (2010) estudaram a influência do Programa Bolsa Família na redução da desigualdade e da pobreza no Brasil a partir de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 1999 até 2009. Para o estudo sobre a desigualdade foi utilizada a metodologia de decomposição fatorial do coeficiente de Gini a fim de captar qual a contribuição do programa para a queda na desigualdade. E, segundo os autores, a renda proveniente do Programa Bolsa Família apresentou um efeito um tanto imprevisível, de acordo com os resultados, o benefício do Bolsa Família é responsável por 16% da queda global da desigualdade durante os anos analisados, uma vez que a renda do programa representa menos que 0,8% da renda total das famílias. Os autores afirmam que o principal motivo da renda transferida pelo Bolsa Família contribuir de modo tão determinante para a redução da desigualdade é a focalização nos mais pobres.

Sendo a focalização uma ferramenta importante na diminuição da desigualdade a presença de uma política bem focalizada é essencial para amparar a população mais carente, explicam os autores. Os

resultados encontrados indicaram uma tendência de declínio dos erros de inclusão e exclusão⁹ do Programa Bolsa Família durante os anos de 2005, 2007 e 2009, assim como as curvas de incidência, para todos os anos, mostraram uma boa focalização do programa concentrando mais de 80% dos benefícios do Bolsa Família entre os 40% mais pobres. Apesar disso, Soares *et al.* (2010) ressaltam que os erros de exclusão são ainda bastante preocupantes. Ao imaginar uma família que se encontra em situação de extrema vulnerabilidade social, e que por algum motivo não está incluída entre os beneficiários, um erro de exclusão deste porte resulta no desemprego da família. Dessa forma, para os autores, no aperfeiçoamento das ferramentas de focalização, a prioridade deve ser a redução dos erros de exclusão mesmo que, em contrapartida, isso gere o aumento dos erros de inclusão. Discutidas as questões de desigualdade e focalização, os autores partem para o efeito do Programa Bolsa Família sobre a pobreza. Os resultados reportam que a diferença entre as taxas de pobreza baseadas na renda domiciliar *per capita*, com e sem o benefício, comprovam o efeito atenuador da pobreza resultante do programa.

Barros *et al.* (2010) realizaram um estudo de focalização do Programa Bolsa Família em perspectiva comparada com diversos programas sociais em outros países e no Brasil. Os autores utilizaram curvas de incidência que indicam a porcentagem de beneficiários que pertencem a cada centésimo da distribuição de renda, desta forma, sendo um programa bem focalizado aquele que tem maior parcela de beneficiários nos centésimos mais pobres da população. Em comparação internacional, países como México, Chile e Panamá adotaram Programa de Transferência Condicionada de Renda (PTCR) semelhantes ao Brasil, porém, com diferentes sistemas de seleção dos beneficiários. Segundo Barros *et al.* (2010) é fundamental fazer uma avaliação dos programas internacionais de PTCR para verificar se o sistema de focalização do Bolsa Família está funcionando de forma adequada. Os autores encontraram resultados semelhantes para o desempenho da focalização do PBF e dos programas *Oportunidades* do México e do *Subsídio Único Familiar (SUF)* no Chile¹⁰. O *Red de Oportunidad* do Panamá apresentou um desempenho melhor que o PBF, o programa panamenho transferiu cerca de 80% dos seus recursos para os 20% mais pobres, já o programa brasileiro alcançou 80% das transferências somente entre os 40% mais pobres. Lembrando que o PTC panamenho utiliza um mecanismo de seleção baseado no método de prova de meios (*proxy means test*) que mostra a probabilidade das famílias cadastradas serem pobres levando em conta variáveis socioeconômicas.

Na comparação do grau de focalização do Programa Bolsa Família (PBF) e de outras políticas públicas do país¹¹, os autores encontraram resultados que mostram que o PBF estava em melhor situação de focalização que os outros programas no ano de 2003¹². Para as curvas de incidência obtidas, no PBF cerca de 75% dos beneficiários pertencem aos 40% mais pobres, e para os outros programas, menos de 40% dos beneficiários estavam entre os 40% mais pobres. E por fim, Barros *et al.* (2010) ainda compararam dois cenários possíveis para o Bolsa Família, um com o emprego de ferramentas de focalização e outro com a universalização das transferências. Buscando identificar como seria a renda dos mais pobres caso não fossem postos em prática os instrumentos de focalização; e acabaram confirmando que a focalização garantiu a maior eficiência nas transferências de renda para os centésimos mais pobres da população.

Sobre a análise do desempenho do Bolsa Família e de outros programas de transferência de renda na América Latina, Soares, Ribas e Osório (2010) comprovaram que há erros de direcionamento nos programas. Com base na Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2004 para o Brasil e na *Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)* de 2004 para o México, o Programa Bolsa Família apresentou um erro de exclusão menor em relação ao programa *Oportunidades* do México,

⁹O erro de inclusão é calculado como a razão entre o número de beneficiários não pobres e o total da população beneficiária. O erro de exclusão corresponde à razão entre os não beneficiários pobres (elegíveis) e o total de pobres (elegíveis).

¹⁰Os dados utilizados pelos autores foram: Pesquisa Nacional por Amostra de Domicílios, Brasil (2008), Encuesta de Hogares, Panamá (2008), Encuesta de Caracterización Socioeconómica Nacional (Casen), Chile (2003) e Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) México (2004).

¹¹Os outros programas e políticas públicas utilizados pelos autores foram: abono salarial, Fundo de Garantia de Tempo de Serviço (FGTS), seguro-desemprego, outras indenizações trabalhistas, auxílio alimentação e transporte, auxílio natalidade, maternidade e pensões e aposentadorias públicas.

¹²Como os autores usaram dados da Pesquisa de Orçamentos Familiares (POF) de 2003, avaliaram a focalização das políticas que deram origem ao PBF: Renda Mínima, Programa de Erradicação do Trabalho Infantil, Bolsa Escola, Agente Jovem, Auxílio-Gás e Auxílio Energia Elétrica.

enquanto o programa brasileiro deixava de atingir 59% dos pobres, o mexicano excluía 70% dos pobres. Por outro lado, o erro de inclusão do programa mexicano foi menor do que o brasileiro. Dos beneficiários do *Oportunidades*, 36% faziam parte dos não elegíveis, já para o Bolsa Família esse percentual era de 49% dos beneficiários. Para os autores, os resultados distintos entre os dois programas são causados pelos mecanismos de seleção dos mesmos. Enquanto o *Oportunidades* tem menor inclusão de não elegíveis ao custo de ter uma menor cobertura, o PBF tem maior cobertura, porém tem que sustentar um maior grau de vazamento de não elegíveis que recebem o benefício.

4. Estratégia empírica e seleção amostral

O método *Propensity Score Matching*, indicado por Rosenbaum e Rubim (1983), visa agrupar as características observáveis das unidades em um único escore que possibilite o pareamento entre tratados e controles lidando com o problema da multidimensionalidade. A técnica é amplamente utilizada na literatura de avaliação quando se refere a estimação de efeitos causais de tratamento (ANDRADE *et al.*, 2012; TAVARES, 2016; De SOUZA; WALTENBEG, 2016). Para a construção de modelos de avaliação que visam identificar o efeito do tratamento é preciso ter informações sobre as características observáveis dos indivíduos analisados, além do status de tratamento e seus resultados potenciais para a variável de interesse. No entanto, para indivíduos tratados não temos posse do resultado potencial caso eles não recebessem o tratamento, e é preciso encontrar uma maneira de supor esse resultado.

Surge assim o problema do contrafactual. Dado que só é possível observar o resultado potencial Y_{1t} do indivíduo tratado, e a informação sobre o resultado potencial do indivíduo tratado caso ele não sofresse intervenção não é observável, Y_{0t} . Torna-se necessária a construção do contrafactual para o grupo de tratamento, denotado como grupo de controle, formado por indivíduos que não sofreram a intervenção, mas apresentam características semelhantes as do grupo de tratamento para obtenção desse resultado potencial.

Para realizar a construção deste contrafactual é necessário a estimação de um escore de propensão, com base em características observáveis. O estimador do escore de propensão é baseado no modelo de regressão logística e mensura a probabilidade de cada indivíduo receber o tratamento ou não. A partir da escolha da função $h(x)$ e a estimação do parâmetro desconhecido γ por máxima verossimilhança. O estimador de *propensity score* segue a equação:

$$\hat{e}(x|W, X) = \frac{\exp(h(x)' \widehat{\gamma}_{ml}(W, X))}{1 + \exp(h(x)' \widehat{\gamma}_{ml}(W, X))} \quad (1)$$

A escolha das variáveis que vão formar a função $h(x)$ é realizada por meio do algoritmo de seleção proposto por Imbens (2015), o qual seleciona uma série de variáveis lineares e quadráticas, e ainda a interação entre as variáveis lineares, via teste de razão de máxima verossimilhança. Com X representando a matriz de características observáveis e W_i o indicador de tratamento. Sendo assim, se a unidade i receber o tratamento $W_i = 1$, caso contrário $W_i = 0$. O algoritmo oferece uma maior especificação na composição do escore de propensão por permitir a inclusão tanto de variáveis lineares como a interação entre elas, e realiza a escolha das mesmas via teste de razão de máxima verossimilhança e não por arbítrio do pesquisador. Isto posto, foi realizada a escolha das variáveis para construir o escore que estima a probabilidade de participação no tratamento via algoritmo de Imbens e calculado o *propensity score*, $\hat{e}(X_i)$.

Segundo Rosenbaum e Rubin (1983), com o *propensity score* já estimado o impacto do tratamento sobre a variável de interesse é obtido através do cálculo do efeito médio do tratamento sobre o tratado, representado pela expressão:

$$ATT = E\{E[Y_{1i}^{AI} | W = 1, \hat{e}(X)] - E[Y_{0i}^{AI} | W = 0, \hat{e}(X)] | W = 1\} \quad (2)$$

Onde, no nosso caso, o primeiro termo da igualdade é estimado com base nos indivíduos analfabetos isolados e o segundo termo através do resultado médio do grupo de controle de analfabetos não isolados pareados.

Para que as estimações sejam viáveis, a validade do método depende de duas hipóteses: (1) independência condicional ou de seleção nas observáveis, a qual assume que condicional ao vetor de características observáveis o resultado é independente do tratamento. Sendo a seleção baseada apenas em características observáveis dos indivíduos, admite-se que todas as variáveis que afetam o tratamento e o resultado são observadas na análise, dada por $Y_i \perp W_i | \hat{e}(X_i)$; (2) hipótese de suporte comum que implica que para cada valor de x exista observações de tratados e controle passíveis de comparação, na qual $0 < \hat{e}(W_i = 1 | X_i) < 1$.

Segundo Rosenbaum (2002) a hipótese de independência condicional restringe a análise do PSM, já que infere que todas as características que afetam o tratamento e o resultado são observadas no modelo. No entanto, a presença de fatores não observados pode interferir nos resultados. Ainda assim, a estimativa do *propensity score* não é suficiente para estimar o ATT, de modo que é uma variável contínua e a probabilidade de encontrarmos dois indivíduos com o mesmo valor de *propensity score* é praticamente zero (BECKER; ICHINO, 2002). Para lidar com esse problema vários métodos foram propostos na literatura como o *Nearest Neighbor Matching*, *Radius Matching*, *Kernel Matching* e *Stratification Matching*.

Os dados utilizados para a estimação são da Pesquisa Nacional por Amostra de Domicílios (PNAD) do ano de 2014, disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Além do suplemento da PNAD 2014 de Acesso ao Cadastro Único para Programas Sociais do Governo Federal e a Programas de Inclusão Produtiva que possibilitou a identificação dos indivíduos que conhecem o Cadastro Único de Programas Sociais do Governo Federal através da pergunta: “Algum morador já ouviu falar sobre o Cadastro Único de Programas Sociais do Governo Federal?”.

Para encontrarmos os indivíduos analfabetos isolados, primeiramente, identificamos cada domicílio da amostra com seus respectivos moradores, e realizamos a classificação desses indivíduos como alfabetizados ou não alfabetizados criando a variável da status de alfabetização dos indivíduos no domicílio a partir de determinada faixa de idade. A definição do status de alfabetização dos indivíduos nos domicílios foi construída com base na pergunta “sabe ler e escrever” que compõem a PNAD. Quando a resposta era “sim” o indivíduo foi caracterizado como alfabetizado, e quando era “não” os indivíduos eram considerados analfabetos.

A variável de alfabetização foi construída a partir do status de alfabetização dos indivíduos maiores de 15 anos. Por exemplo, na variável de alfabetização considerou-se apenas a alfabetização de indivíduos maiores de 15 anos, definindo o status de alfabetização de todos os moradores do domicílio maiores de 15 anos, atribuiu-se o valor um para os indivíduos alfabetizados maiores de 15 anos e zero para os demais indivíduos dentro do domicílio. Desta forma, através da variável de alfabetização, foi possível identificar aqueles domicílios em que para todos os moradores a variável alfabetização é igual a zero e, assim, especificar cada indivíduo desse domicílio como analfabeto isolado – AI15 - (tratados). Para aqueles domicílios em que pelo menos um morador foi identificado como alfabetizado, com base na idade especificada na definição da amostra, os indivíduos analfabetos foram classificados como não isolados (controles).

A variável dependente utilizada na análise empírica é binária, assume valor um para os indivíduos que conhecem o Cadastro Único e zero caso contrário. Além da variável sobre o conhecimento por parte de algum morador do domicílio da existência do Cadastro Único do Governo Federal, as covariadas utilizadas foram divididas em três grupos¹³:

- Atributos do indivíduo: variável de idade, *dummy* para raça, e *dummy* para estado civil.
- Atributos do domicílio: número de moradores menores de 9 anos, número de moradores maiores de 9 anos, número de quartos do domicílio, *dummy* para existência de banheiro no domicílio.
- Acesso à infraestrutura básica e localização do domicílio: *dummy* para identificar o acesso a rede de esgoto no domicílio e variável *dummy* para indicar a zona de localização do domicílio

¹³ As covariadas utilizadas foram escolhidas com base em estudos voltados para avaliação de políticas públicas, em especial do Programa Bolsa Família. Sendo assim, acreditamos que essas variáveis possam captar, de certa forma, as características de potenciais elegíveis aos programas sociais do Governo Federal (DE ARAÚJO; GOMES; LIMA, 2016; DE SOUZA; WALTENBEG, 2016).

Especificadas as variáveis empregadas, procedemos com as estimações a nível de Brasil, separando a amostra com e sem distinção de gênero. Quanto a especificação do nível de renda, optou-se por apresentar os resultados das estimações que consideram a amostra formada por indivíduos com renda *per capita* domiciliar de até ½ salário mínimo¹⁴ já que são esses indivíduos que atendem ao critério de renda para elegibilidade do Cadastro Único. Foram também realizadas estimações para o nível de renda de até 1 salário mínimo domiciliar *per capita* como forma de robustez para a análise, uma vez que ainda há uma parcela de indivíduos que são incluídos nas políticas públicas mesmo apresentando renda superior ao critério de elegibilidade (BARROS *et al.*, 2010; SOARES; RIBAS; OSÓRIO, 2010; SOARES *et al.*, 2010; SOUZA *et al.*, 2013).

Para que fosse possível avaliar o impacto do analfabetismo isolado para aqueles indivíduos que efetivamente são o público-alvo do programa, o critério de renda do Cadastro Único foi respeitado nas estimações. Acredita-se que ao construir uma amostra com indivíduos que de fato atendem ao critério de renda para participação do Cadastro Único, e que desta forma seriam potenciais beneficiários das políticas públicas que compõem o programa, seria estruturado um grupo mais equilibrado em relação as características observáveis.

4.1 Análise de sensibilidade (Limites de Rosenbaum)

Os estudos que avaliam o efeito de alguma intervenção com estimadores de correspondência partem da hipótese de independência condicional. Se os grupos de tratados e controles diferem devido a fatores não observados, e desta forma, as variáveis não observadas podem afetar simultaneamente o tratamento e variável de resultado, a especificação do modelo de pareamento pode gerar um viés na estimação do efeito médio do tratamento sobre os tratados. Uma estratégia para retratar este problema é através da abordagem proposta por Rosenbaum (2002). A análise de sensibilidade, os limites de Rosenbaum, não elimina a possibilidade de que fatores não observados existam, mas estabelece o quão prejudicial eles são para os resultados, a fim de contaminar as inferências de uma análise de correspondência.

A verificação da sensibilidade proposta por Rosenbaum (2002) foi aplicada em Aakvik (2001), DiPrete e Gangl (2004) e Becker e Caliendo (2007) para analisar se a inferência sobre os efeitos do tratamento pode ser alterada por fatores não observados. Desta forma, propõe uma verificação da robustez do modelo através da hipótese de viés de seleção. O método identifica dois indivíduos a partir de suas características observáveis, denominados i e j , dentro de uma distribuição logística. Para isso, é assumido que a probabilidade participação no tratamento é:

$$\tau_i = \hat{e}(W = 1|X_i = x) = F(\beta X_i + \gamma \mu_i) \quad (3)$$

Onde x_i é o vetor que contém as características observadas do indivíduo i , μ_i é o termo de erro estocástico, e γ mede o efeito de μ_i sobre a probabilidade do indivíduo ser analfabeto isolado. Evidentemente, se não houver nenhum viés de variável omitida, μ_i será zero e a probabilidade de ser tratado será determinada apenas pelo vetor de características observadas X_i . Caso contrário, constatado o viés de seleção devido a interferência de variáveis não observáveis na análise, dois indivíduos com características observáveis semelhantes terão diferentes probabilidades de receber o tratamento. Assumindo que $F(\cdot)$ é uma função com distribuição logística, a probabilidade de que o indivíduo i tem de ser tratado pode ser escrito:

$$\left(\frac{\tau_i}{1 - \tau_i} \right) = \exp(\beta X_i + \gamma \mu_i) \quad (4)$$

Logo, se fizermos a comparação de dois indivíduos pareados, i e j , que apresentam as mesmas características observáveis ($X_i=X_j$), a razão de probabilidade de receber o tratamento pode ser escrita como:

¹⁴O salário mínimo em 2014 era R\$ 724,00, logo a renda analisada nesse estudo compreende a faixa de renda domiciliar per capita de R\$ 362,00, ou seja, ½ salário mínimo.

$$\frac{\frac{\tau_i}{1-\tau_i}}{\frac{\tau_j}{1-\tau_j}} = \frac{\tau_i(1-\tau_j)}{\tau_j(1-\tau_i)} = \frac{\exp(\beta X_j + \gamma \mu_j)}{\exp(\beta X_i + \gamma \mu_i)} = \exp\gamma[(\mu_i - \mu_j)] \quad (5)$$

Se confirmada a presença de viés de seleção, os dois indivíduos com características observáveis semelhantes têm diferentes chances de serem submetidos ao tratamento, pois o efeito de μ é captado por γ que será diferente de zero. Não havendo diferenças relacionadas as características não observadas, nenhum viés é encontrado ($\gamma=0$) e a razão de chances dos indivíduos pareados receberem o tratamento é 1. Em resumo, a análise de sensibilidade avalia o quanto que o efeito médio do tratamento é modificado por alterações nos valores de γ e de $\mu_i - \mu_j$. Os limites da razão de probabilidades de participação no tratamento propostos por Rosenbaum (2002) são expressos pela desigualdade:

$$\frac{1}{\Gamma} \leq \frac{\tau_i(1-\tau_j)}{\tau_j(1-\tau_i)} \leq \Gamma \quad (6)$$

Para os indivíduos pareados que possuem a mesma probabilidade de participar do tratamento o valor de Γ será igual a um. Caso contrário, se por exemplo $\Gamma = 2$, os indivíduos que supostamente parecem ser semelhantes em termos de características observáveis podem diferir em suas chances de receber o tratamento por um fator de 2. Os limites de Rosenbaum indicam o quanto alterações nos valores de Γ modificam a inferência sobre do efeito do tratamento.

Então, com a finalidade de verificar a gravidade do viés de fatores omitidos, foi sugerido por Rosenbaum (2002) e Caliendo e Kopeinig (2008) a mensuração da sensibilidade das estimativas do efeito causal do tratamento em relação às características não-observadas por meio dos Limites de Rosenbaum. Os limites indicam o quanto alterações dos fatores não observáveis podem afetar a robustez dos resultados. A estatística de *Mantel-Haenszel* é a mais indicada para avaliar a hipótese do viés do efeito do tratamento, uma vez que a variável de resultado é binária. A estatística testa a hipótese nula de ausência de efeito do tratamento nas estimações (BECKER; CALIENDO, 2007; AAKVIK, 2001).

5. Resultados

Conforme os dados apresentados para a amostra, no Apêndice B, com e sem distinção de gênero, pode-se concluir por meio da análise realizada, que os indivíduos analfabetos isolados apresentam, em média, um menor conhecimento do Cadastro Único em relação aos analfabetos não isolados. Cerca de 50% dos indivíduos analfabetos isolados e 35% dos indivíduos analfabetos não isolados não sabem da existência do Cadastro Único. As informações reportam também as características dos indivíduos da amostra geral, e indicam que os indivíduos analfabetos isolados, em sua maioria, têm em torno de 39 anos, são não brancos, e não são casados. Decompondo os atributos do domicílio, observa-se que quase 20% das residências de indivíduos analfabetos isolados não possuem banheiro. Constata-se também a precariedade em relação ao serviço de saneamento básico. De acordo com os dados, em média, 19% dos domicílios com indivíduos do grupo de tratamento e 32% dos domicílios com indivíduos do grupo de controle apresentam rede de esgoto. Verifica-se ainda que os indivíduos tratados residem em domicílios, em média, com quantidade inferior de quartos e menor número de pessoas adultas e crianças do que os indivíduos do grupo de controle.

Posto isso, os resultados para o modelo *logit* que calculou a probabilidade de cada indivíduo ser analfabeto isolado com renda domiciliar per capita de até $\frac{1}{2}$ salário mínimo para o ano de 2014 foram utilizados como base para o pareamento entre os grupos. Cada indivíduo tratado (analfabeto isolado) foi pareado com um indivíduo do grupo de controle (analfabeto não isolado) que apresentou igual probabilidade de receber o tratamento. Tornando possível a comparação de indivíduos que diferem apenas

do fato de ser analfabeto isolado ou não, mas que compartilham de características observáveis muito próximas.

Após a estimação dos resultados do modelo *logit* é relevante verificar o balanceamento entre os grupos de tratamento e controle depois de realizado o pareamento, pois a qualidade desse balanceamento é de extrema importância para a obtenção de estimativas corretas do ATT. Nesse sentido, no Apêndice C, exibe os gráficos com a sobreposição do *propensity score* dos grupos de tratados e controles após pareamento para todos os métodos de pareamento. Em geral, os grupos amostrais analisados apresentaram um comportamento similar em termos de características observáveis após o pareamento para ambos os grupos. E entre os métodos de pareamentos utilizados, destacamos que o pareamento pelos três vizinhos mais próximos foi o que apresentou um melhor equilíbrio entre os grupos em todas as amostras, dado que registrou uma melhor sobreposição entre tratados e controles. Feitas as estimativas do *propensity score* pelo modelo *logit* e verificado o balanceamento entre os grupos de tratamento e controle, procedeu-se à estimação do efeito médio do tratamento sobre os tratados pelos três métodos de pareamento considerados, cujos resultados serão apresentados.

Na tabela 1 apresentamos as estimações do ATT para o Brasil. Os coeficientes mostram que para observações que contabilizam a condição de alfabetização no domicílio a partir dos 15 anos de idade sem distinção de gênero há evidências de que, em média, os indivíduos analfabetos isolados têm probabilidade menor de acesso a informação da existência do Cadastro Único em relação aos seus pares do grupo de controle na amostra de acordo com os métodos de pareamento empregados na análise. Esses resultados foram estatisticamente significativos e mostram que a diferença entre tratados e controles varia de -11,93 pontos percentuais com pareamento por raios, a -12,00 pontos percentuais através do pareamento pelos três vizinhos mais próximos.

Ainda na tabela 1, os coeficientes encontrados para as observações do sexo feminino, em domicílios que contabilizaram a condição de alfabetização apenas para maiores de 15 anos, e conforme todos os métodos de pareamento, mostram que o efeito do tratamento diminui a probabilidade de acesso a informação sobre a existência do Cadastro Único. Os resultados foram estatisticamente significativos e indicam que para as observações do sexo feminino a diferença entre o acesso a informação do programa de analfabetos isolados e não isolados varia, em média, de -9,60 pontos percentuais no método de pareamento pelos três vizinhos mais próximos, a -13,03 pontos percentuais, no método de pareamento de *Kernel*.

E por fim, ainda na tabela 1, encontramos resultados que indicam que o acesso a informação da existência do Cadastro Único é menor para indivíduos analfabetos isolados do sexo masculino da amostra, em domicílios que a alfabetização foi ponderada a partir dos 15 anos, em relação aos seus pares para todos os métodos de pareamento utilizados. Os coeficientes foram significativos e o efeito médio do tratamento sobre os tratados varia, em média, de -10,83 pontos percentuais para o método dos três vizinhos mais próximos, a -11,55 pontos percentuais para o método de pareamento por *Kernel*.

Tabela 1 - Estimativa do efeito de tratamento médio sobre os tratados: probabilidade de conhecer o Cadastro Único para indivíduos com nível de renda domiciliar per capita até ½ salário mínimo, Brasil 2014

Método de pareamento	AI	ANI	ATT	Erro-padrão (Bootstrap)
		AI15		
Vizinho mais próximo	0,5312	0,6513	-0,1200**	0,0277122
<i>Kernel Matching</i>	0,5312	0,6510	-0,1197**	0,0230082
<i>Raius Matching</i>	0,5312	0,6505	-0,1193**	0,0219944
		AI15 (Sexo Feminino)		
Vizinho mais próximo	0,5195	0,6156	-0,0960*	0,0447101
<i>Kernel Matching</i>	0,5195	0,6498	-0,1303**	0,0351121
<i>Raius Matching</i>	0,5195	0,6403	-0,1207**	0,0358849

AI15 (Sexo Masculino)				
Vizinho mais próximo	0,5360	0,6443	-0,1083**	0,0419485
Kernel Matching	0,5360	0,6515	-0,1155**	0,0282415
Raius Matching	0,5360	0,6471	-0,1111**	0,0306797

Fonte: Elaboração própria com dados da PNAD 2014. Nota: P-valor obtido pelo erro-padrão estimado por bootstrap com 50 repetições; **Significativo a 1% e * Significativo a 5%.

Verificado o efeito redutor do analfabetismo isolado sobre acesso a informação da existência do Cadastro Único dos indivíduos analfabetos isolados para a amostra sem distinção de gênero. Também foi analisado por gênero dos indivíduos, para os homens, destacamos o coeficiente de -11,55 pontos percentuais na probabilidade de conhecer o Cadastro Único para analfabetos isolados em relação a analfabetos não isolados na amostra para o sexo masculino. E para mulheres, o coeficiente de -13,03 pontos percentuais na probabilidade de conhecer o Cadastro Único para analfabetas isoladas em relação a analfabetas não isoladas, nas observações para o gênero do sexo feminino. Além disso, a maioria dos coeficientes absolutos mostram que há um efeito médio semelhante sobre indivíduos do sexo feminino e masculino no sentido de redução do acesso a informação devido a situação de analfabetismo isolado.

A comparação destes resultados com os outros trabalhos presentes na literatura sobre o tema, apresentados na seção de externalidades da alfabetização, nesse trabalho, fica prejudicada em razão da estratégia de análise ser diferente, e além do tema de políticas públicas também não ter sido abordado nessa linha de pesquisa. Neste trabalho, estimamos o efeito da ausência dessa externalidade no acesso a informação sobre políticas públicas, e no restante dos trabalhos é verificado a presença dessa externalidade.

5.1 Resultados da análise de sensibilidade

Os resultados encontrados na seção anterior levam em consideração a hipótese de independência condicional, mas ainda assim a presença de variáveis não observáveis na especificação do modelo pode resultar em um viés na estimação do efeito médio do tratamento sobre os tratados. Como forma de quantificar esse viés foi feita a análise de sensibilidade dos resultados, através do estudo dos Limites de Rosenbaum. A análise permitiu avaliar a sensibilidade dos resultados estimados em relação a possíveis falhas na suposição de independência condicional decorrentes de variáveis não observadas que afetam o status de tratamento e a variável de resultado, simultaneamente, e que podem gerar um viés de variável omitida nas estimações. Em virtude de ambas as variáveis, de tratamento e resultado, serem binárias o teste utilizado foi de *mhbounds*.

A estatística de teste *Mantel-Haenszel* verifica a hipótese nula de ausência do efeito do tratamento. O teste determina dois limites que podem ser interpretados da seguinte maneira: a estatística Q_{mh+} representa o limite para o caso de sobrestimação do efeito do tratamento e Q_{mh-} aponta o limite para o caso de subestimação do efeito do tratamento. O fator Γ mede a diferença de probabilidade de receber o tratamento entre os indivíduos com características observáveis semelhantes, para verificar qualquer alteração nas estimações devido a variáveis omitidas. Assim, indica o grau em que os indivíduos do grupo de tratamento e controle, que aparentemente são semelhantes em termos de características observáveis e dentro do suporte comum, diferem ou não em suas probabilidades de ser tratado. O teste leva em consideração um intervalo com os níveis críticos do fator Γ e os níveis de significância da hipótese de sobrestimação e subestimação. A consistência dos resultados é feita ao elevar a variação do Γ até o momento em que o nível de significância (p_{mh+} e p_{mh-}) é afetado. Desta forma, foi calculado o intervalo de nível crítico com as respectivas probabilidades de não-rejeição da hipótese nula de sobrestimação e subestimação dos resultados dentro de um intervalo de 1 a 2 para o fator Γ .

Conforme podemos observar, no Apêndice D, para todos os valores de Γ foi possível rejeitar a hipótese nula de sobrestimação dos resultados do ATT, ao nível de significância estatística de 1%, devido a influência de variáveis omitidas para as estimações com e sem distinção de gênero. Para valores de Γ referentes a hipótese de subestimação, de modo geral, os resultados mostram que há apenas alguns momentos de alteração além do intervalo de significância de 10%. Isso ocorre quando para a estatística

Q_{mh} - de *Mantel-Haenszel*, tendo como hipótese nula a subestimação do efeito do tratamento, o fator Γ assume valores entre 1,4 e 1,8 para alguns métodos de pareamento, mas acaba voltando a ter significância. Esta probabilidade está associada a uma possível subestimação do modelo no que tange ao efeito do tratamento, mas acreditamos que isso não prejudica a análise uma vez que foram alterações leves e, ainda assim, na maioria dos valores foi possível rejeitar a hipótese nula. Em geral, para os demais valores de Γ , os resultados da análise de sensibilidade indicam que as inferências a respeito do efeito do tratamento são válidas ao não ser observada influência de fatores não observáveis.

6. Considerações finais

O presente estudo teve como proposta avaliar se há relação entre o analfabetismo isolado e o acesso a informação sobre o Cadastro Único para indivíduos com renda domiciliar per capita de até $\frac{1}{2}$ salário mínimo, usando os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2014 e o método de *Propensity Score Matching* (PSM) aliado com o algoritmo de Imbens (2015). O estimador do PSM foi aplicado para a amostra com e sem distinção de gênero e para diferentes métodos de pareamento.

O tema de externalidades da alfabetização ganhou espaço após pesquisa de Basu e Foster (1998). Para os autores, uma pessoa capaz de desempenhar atividades de leitura e escrita proporciona que a alfabetização seja vista como um bem público para os membros analfabetos na residência. No entanto, analfabetos isolados não teriam acesso à informação, dada a ausência dessa externalidade dentro do domicílio, e estariam em uma situação de desvantagem em relação a analfabetos não isolados. Dada a importância da implantação de políticas públicas como estratégia dos governos no combate à pobreza e à desigualdade social (KERSTENETZKY, 2009). A pesquisa avaliou o desafio de direcionar o acesso de programas sociais aos indivíduos que realmente demandam desses benefícios, com base na focalização de políticas públicas justificadas pelo grau de exclusão de indivíduos dada a menor informação, ou mesmo a falta de acesso a ela, visto que residir em um domicílio com apenas indivíduos analfabetos pode impor uma série de restrições ligadas ao acesso de informações e serviços.

Os resultados confirmaram que a ausência de retorno social domiciliar da alfabetização, dado o fato de não residir com nenhum indivíduo alfabetizado, foi determinante na redução do acesso a informação sobre a existência do Cadastro Único. Visto que os mesmos seriam potenciais beneficiários, por se encaixarem no critério de renda para elegibilidade no Cadastro Único, os resultados levam a evidência de que a menor informação, ou mesmo a falta de acesso a ela, que a situação de analfabetismos isolado pode impor, afeta esses indivíduos criando um obstáculo para o conhecimento sobre seus direitos e, possivelmente, deixando-os de fora da rede de proteção social do governo. Quanto a análise de sensibilidade dos resultados encontrados, em geral, para a maioria dos valores de Γ , os resultados indicam que as inferências a respeito do efeito do tratamento são válidas ao não ser observada influência de fatores não observáveis nas estimações.

Evidenciada a importância do retorno social da alfabetização para o acesso a informação, uma vez que a falta de um indivíduo alfabetizado restringe o acesso às atividades que requerem o uso de leitura e escrita no domicílio. Fica a sugestão para pesquisas futuras a análise da influência relacionada a posição dos membros nos domicílios para externalidade social da alfabetização, assim como a verificação do efeito da externalidade em níveis mais abrangentes de contato dos indivíduos, e não só em âmbito familiar. Além também, de aumentar as discussões sobre a implantação de políticas públicas que consigam centrar em questões ligadas a analfabetismo, e em especial, do analfabetismo isolado.

REFERÊNCIAS BIBLIOGRÁFICAS

AAKVIK, Arild. Bounding a matching estimator: the case of a Norwegian training program. **Oxford bulletin of economics and statistics**, v. 63, n. 1, p. 115-143, 2001.

ANDRADE, Mônica Viegas et al. Income transfer policies and the impacts on the immunization of children: the Bolsa Família Program. **Cadernos de saúde pública**, v. 28, n. 7, p. 1347-1358, 2012

AUSTIN, Peter C.; JEMBERE, Nathaniel; CHIU, Maria. Propensity score matching and complex surveys. **Statistical Methods in Medical Research**, July 26, 2016.

DE BARROS, Ricardo Paes et al. A focalização do Programa Bolsa Família em perspectiva comparada. In: Jorge Abrahão de Castro, Lúcia Modesto. Brasília (Org.) **Bolsa família 2003-2010: avanços e desafios**. Brasília: IPEA, 2 v. (366p.), 2010.

BASU, Kaushik; FOSTER, James E. On measuring literacy. **The Economic Journal**, v. 108, n. 451, p. 1733-1749, 1998.

BASU, Kaushik; NARAYANC, Ambar; RAVALLION, Martin. Is literacy shared within households? Theory and evidence for Bangladesh. **Labour Economics**, v. 8, p. 649-665, 2001.

BECKER, Sascha O.; ICHINO, Andrea. Estimation of average treatment effects based on propensity scores. **The stata journal**, v. 2, n. 4, p. 358-377, 2002.

BECKER, Sascha O.; CALIENDO, Marco. Sensitivity analysis for average treatment effects. **The Stata Journal**. V. 7, N. 1, pp. 71-83, 2007.

BOROOAH, Vani. *The Role of Maternal Literacy in Reducing the Risk of Child Malnutrition in India*. In: Gender and Discrimination: Health, Nutritional Status and Role of Women in India. (Eds: Pal, Manoranjan, Bharati, Premananda, Ghosh, Bholanath and Vasulu, TS), **Oxford University Press**, pp. 141-162, 2009.

BRASIL. Decreto nº 3.877, de 24 de julho de 2001. Institui o Cadastramento Único para Programas Sociais do Governo Federal. Brasília – DF. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/2001/D3877.htm

_____. Decreto nº 6.135, de 26 de junho de 2007. Dispõe sobre o Cadastro Único para Programas Sociais do Governo Federal e dá outras providências. Brasília – DF. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Decreto/D6135.htm#art14

CALIENDO, Marco; KOPEINIG, Sabine. Some practical guidance for the implementation of propensity score matching. **Journal of economic surveys**, v. 22, n. 1, p. 31-72, 2008.

CAMERON, A. Colin; TRIVEDI, Pravin K. **Microeconometrics: methods and applications**. Cambridge university press, 2005.

COADY, David; GROSH, Margaret E.; HODDINOTT, John. **Targeting of transfers in developing countries: Review of lessons and experience**. World Bank Publications, 2004.

DE ARAÚJO, Aracy Alves; GOMES, Marília Fernandes Maciel; LIMA, João Eustáquio. Influência do Programa Bolsa Família na redução do trabalho infantil: evidências para o Nordeste brasileiro. **Revista Econômica do Nordeste**, v. 45, n. 3, p. 33-45, 2016.

- DE BALDINI ROCHA, Maúna Soares; PONCZEK, Vladimir. The effects of adult literacy on earnings and employment. **Economics of Education Review**, v. 30, n. 4, p. 755-764, 2011.
- DE SOUZA, Luana Passos; WALTEBEG, Fábio Domingues. Bolsa Família e assimetrias de gênero: reforço ou mitigação?. **Revista Brasileira de Estudos de População**, v. 33, n. 3, p. 517-539, 2016.
- DIPRETE, Thomas A.; GANGL, Markus. Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. **Sociological methodology**, v. 34, n. 1, p. 271-310, 2004.
- DIREITO, Denise do Carmo, et al. **O Cadastro Único como instrumento de articulação de políticas sociais**. No. 327. International Policy Centre for Inclusive Growth, 2016.
- DUGOFF, Eva H.; SCHULER, Megan; STUART, Elizabeth A. Generalizing observational study results: applying propensity score methods to complex surveys. **Health services research**, v. 49, n. 1, p. 284-303, 2014.
- FERRER, Ana; GREEN, David A.; RIDDELL, W. Craig. The effect of literacy on immigrant earnings. **Journal of Human Resources**, v. 41, n. 2, p. 380-410, 2006.
- FREIMUTH, Vicki S.; METTGER, Wendy. Is there a hard-to-reach audience? **Public health reports**, v. 105, n. 3, p. 232, 1990.
- GIBSON, J. (2001). Literacy and Intra household Externalities. **World Development**, vol. 29, p.p. 155-166, 2001.
- IMBENS, Guido W. Matching methods in practice: Three examples. **Journal of Human Resources**, v. 50, n. 2, p. 373-419, 2015.
- IVERSEN, Vegard; PALMER-JONES, Richard. Literacy sharing, assortative mating, or what? Labour market advantages and proximate illiteracy revisited. **The Journal of Development Studies**, v. 44, n. 6, p. 797-838, 2008.
- KERSTENETZKY, Celia Lessa. Redistribuição e desenvolvimento? A economia política do programa bolsa família. **Dados-Revista de Ciências Sociais**, v. 52, n. 1, p. 53-83, 2009. 54
- HUSAIN, Zakir; DUTTA, Mousumi. **Proximate illiteracy and modern contraceptive use in India: Analysis of DHS data**. MPRA Paper No. 39951, July 2012.
- PAES-SOUSA, Rômulo et al. **Plano Brasil Sem Miséria: Incremento e mudança na política de proteção e promoção social no Brasil**. IPC-IG Working Paper, n. 113, 2013.
- RIBEIRO, Felipe Garcia; SOUZA, André Portela Fernandes de. Uma investigação e aplicação das novas medidas de alfabetização no Brasil. **Nova Economia**, v. 23, n. 3, p. 623-656, 2013.
- RIBEIRO, Felipe Garcia et al. Retorno social domiciliar da alfabetização: análise da associação com o tabagismo de analfabetos corresidentes no Brasil. **Revista Panamericana de Salud Pública**, v. 39, n. 6, p. 316-321, 2016.
- ROSENBAUM, Paul R.; RUBIN, Donald B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, p. 41-55, 1983.

- ROSENBAUM, Paul R. Observational studies. In: **Observational Studies**. Springer New York, 2002.
- SARR, Leopold R. **The impact of family literacy on the earnings of illiterates: Evidence from Senegal**. Cornell Food and Nutrition Policy Program Working Paper No. 159, 2004.
- SHAGHAGHI, Abdolreza; BHOPAL, Raj S.; SHEIKH, Aziz. Approaches to Recruiting 'Hard-To-Reach' Populations into Re-search: A Review of the Literature. **Health Promotion**, v. 1, n. 2, p. 01-09, 2011.
- SOARES, Fábio Veras; RIBAS, Rafael Perez; OSÓRIO, Rafael Guerreiro. Evaluating the impact of Brazil's Bolsa Família: Cash transfer programs in comparative perspective. **Latin American Research Review**, v. 45, n. 2, p. 173-190, 2010.
- SOARES, Sergei; SOUZA, Pedro Herculano G. Ferreira de; OSÓRIO, Rafael Guerreiro e SILVEIRA, Fernando Gaiger. Os impactos do benefício do programa bolsa família sobre a desigualdade e a pobreza. In: Jorge Abrahão de Castro, Lúcia Modesto (Org.) **Bolsa família 2003-2010: avanços e desafios**. Brasília: Ipea, 2 v. (366p.), 2010.
- SOUZA, André Portela Fernandes de et al. **Uma investigação sobre a focalização do Programa Bolsa Família e seus determinantes imediatos**. *Working Paper*, Nov 2013.
- TAVARES, Priscilla Albuquerque et al. Uma avaliação do Programa Bolsa Família: focalização e impacto na distribuição de renda e pobreza. **Pesquisa e Planejamento Econômico**, v. 39, n. 1, 2009.
- TAVARES, Priscilla Albuquerque. Efeito do Programa Bolsa Família sobre a oferta de trabalho das mães. **Economia e sociedade**, v. 19, n. 3, p. 613-635, 2016. 55
- UNESCO. Relatório de Monitoramento Global de Educação para Todos. Publicado em 2014 pela Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Título original: **Teaching and Learning: Achieving Quality for All; EFA Global Monitoring Report**, 2013)

APÊNDICE

APÊNDICE A – Esquema de construção da variável de analfabetos isolado e analfabetos não isolados, Brasil 2014.



Fonte: Elaboração própria a partir dos dados da PNAD 2014.

APÊNCIDE B - Média e desvio-padrão das variáveis para indivíduos analfabetos com e sem distinção de gênero, Brasil 2014.

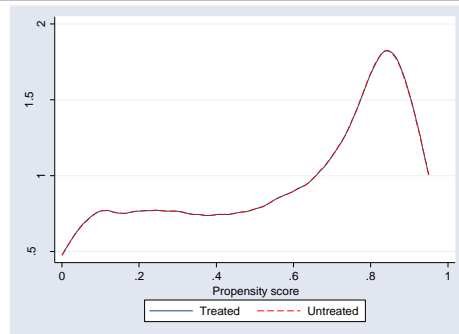
Variável	AI15		AI15 (Sexo Feminino)		AI15 (Sexo Masculino)	
	Analfabetos isolados (Tratado)	Analfabetos não isolados (Controle)	Analfabetos isolados (Tratado)	Analfabetos não isolados (Controle)	Analfabetos isolados (Tratado)	Analfabetos não isolados (Controle)
Variável de resultado						
Cadastro Único (conhecer=1)	0,5246 (0,4995)	0,6499 (0,4770)	0,5225 (0,4997)	0,6474 (0,4777)	0,5266 (0,4995)	0,6521 (0,4763)
Variáveis para construção do propensity score						
Características do indivíduo						
Idade do indivíduo	38,2288 (21,8612)	15,6776 (21,4146)	39,5889 (21,6167)	15,3912 (21,8108)	36,8981 (22,0264)	15,9403 (21,0423)
Dummy para raça (branco=1)	0,1969 (0,3977)	0,2799 (0,4489)	0,2230 (0,4164)	0,2902 (0,4539)	0,17132 (0,3769)	0,2704 (,4442)
Dummy de casado (casado=1)	0,4430 (0,4968)	0,1972 (0,3979)	0,4552 (0,4982)	0,1578 (0,3646)	0,4310 (0,4954)	0,2333 (0,4229)
Atributos do domicílio						
Número de quartos	1,7046 (0,7588)	2,0904 (0,8148)	1,7130 (0,7506)	2,0952 (,8117)	1,6963 (0,7670)	2,0860 (0,8177)
Dummy para possuir banheiro (banheiro=1)	0,7980 (0,4015)	0,9040 (0,2945)	0,8304 (0,3754)	0,9103 (0,2856)	0,7664 (0,4232)	0,8982 (0,3023)
Número de pessoas maiores de 9 anos	3,4507 (1,8790)	4,9596 (1,9677)	3,4516 (1,7980)	4,9596 (1,9629)	3,4499 (1,9558)	4,9597 (1,9722)
Número de pessoas menores de 9 anos	1,0127 (1,3079)	1,7121 (1,2095)	1,0211 (1,2758)	1,7312 (1,2025)	1,0045 (1,3391)	1,6946 (1,2157)
Acesso à infraestrutura básica e localização do domicílio						
Dummy para zona (urbana=1)	0,5273 (0,4993)	0,6940 (0,4608)	0,5778 (0,49412)	0,71604 (0,4509)	0,4779 (0,4997)	0,6739 (0,4687)
Dummy para possuir esgoto (esgoto=1)	0,1950 (0,3963)	0,3285 (0,4696)	0,2313 (0,4218)	0,34489 (0,4753)	0,1596 (0,3664)	0,3134 (0,4639)
Observações	2177	20794	1077	9945	1100	10849

Fonte: Elaboração própria com dados da PNAD 2014. Nota: P-valor obtido pelo erro-padrão estimado por bootstrap com 50 repetições; **Significativo a 1% e * Significativo a 5%.

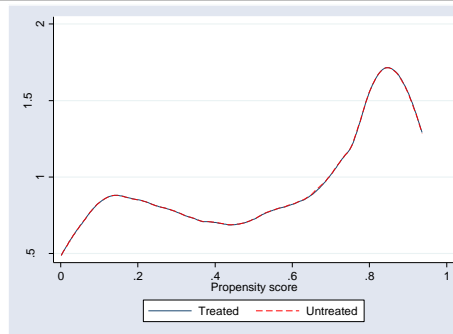
APÊNDICE C - Verificação da hipótese de sobreposição para a amostra, com e sem distinção de gênero, e por métodos de pareamento, Brasil 2014.

Pareamento pelos três vizinhos mais próximos

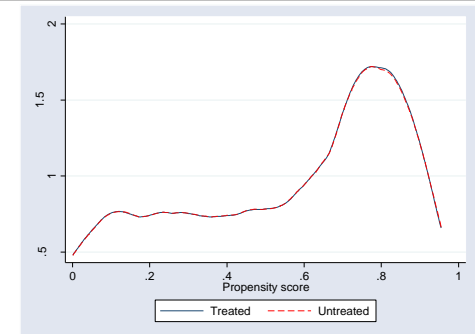
AI15



AI15 (Sexo Feminino)

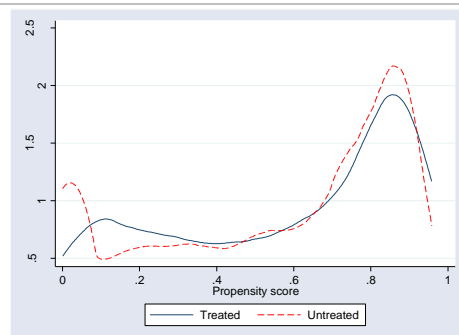


AI15 (Sexo Masculino)

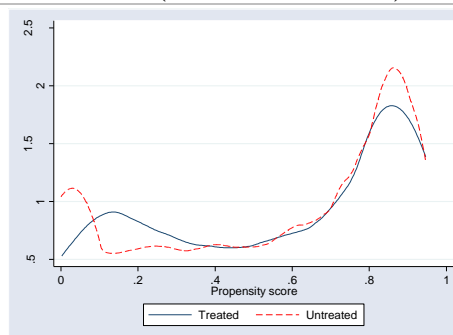


Pareamento por Kernel

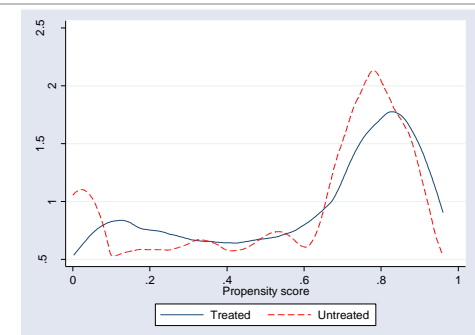
AI15



AI15 (Sexo Feminino)

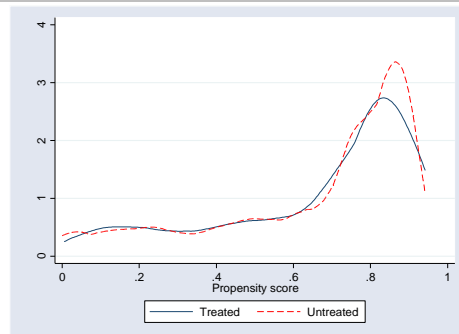


AI15

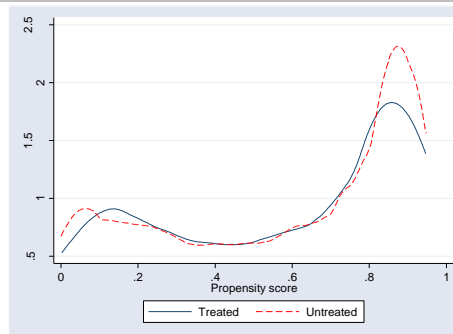


Pareamento por Raio

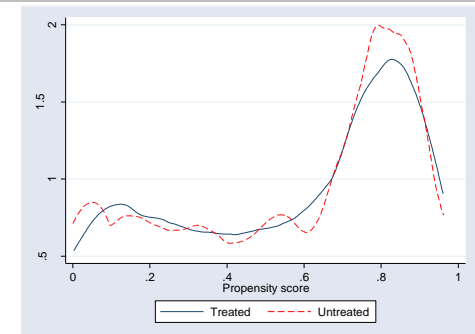
AI15



AI15 (Sexo Feminino)



AI15



Fonte: Elaboração própria a partir de dados da PNAD 2014.

APÊNDICE D - Análise de sensibilidade dos Limites de Rosenbaum para indivíduos analfabetos com e sem distinção de gênero, Brasil 2014

Vizinho mais próximo					<i>Kernel Matching</i>				<i>Radius matching</i>			
AI15												
Γ	Q_mh+	Q_mh-	p_mh+	p_mh-	Q_mh+	Q_mh-	p_mh+	p_mh-	Q_mh+	Q_mh-	p_mh+	p_mh-
1	6,8564	6,8564	3,5e-12	3,5e-12	10,4994	10,4994	0	0	10,4994	10,4994	0	0
1.2	9,66389	4,06616	0	0,000024	14,497	6,57033	0	2,5e-11	14,497	6,57033	0	2,5e-11
1.4	12,0536	1,71377	0	0,043286	17,9507	3,28174	0	0,000516	17,9507	3,28174	0	0,000516
1.6	14,1391	0,255907	0	0,399011	21,0123	0,445421	0	0,328008	21,0123	0,445421	0	0,328008
1.8	15,9928	2,05116	0	0,020126	23,7773	2,00779	0	0,022333	23,7773	2,00779	0	0,022333
2	17,6637	3,6587	0	0,000127	26,3097	4,24862	0	0,000011	26,3097	4,24862	0	0,000011
AI15 (Sexo Feminino)												
1	4,64655	4,64655	1,7e-06	1,7e-06	7,96919	7,96919	7,8e-16	7,8e-16	7,96919	7,96919	7,8e-16	7,8e-16
1.2	6,63437	2,6714	1,6e-11	0,003777	10,803	5,18714	0	1,1e-07	10,803	5,18714	0	1,1e-07
1.4	8,32525	1,00523	0	0,157394	13,2533	2,86075	0	0,002113	13,2533	2,86075	0	0,002113
1.6	9,80042	0,343956	0	0,36544	15,4268	0,856119	0	0,195966	15,4268	0,856119	0	0,195966
1.8	11,1111	1,61579	0	0,05307	17,3909	0,84286	0	0,199653	17,3909	0,84286	0	0,199653
2	12,2922	2,75474	0	0,002937	19,1906	2,42407	0	0,007674	19,1906	2,42407	0	0,007674
AI15 (Sexo Masculino)												
1	4,77744	4,77744	8,9e-07	8,9e-07	7,17147	7,17147	3,7e-13	3,7e-13	7,17147	7,17147	3,7e-13	3,7e-13
1.2	6,74704	2,82121	7,5e-12	0,002392	9,96435	4,42651	0	4,8e-06	9,96435	4,42651	0	4,8e-06
1.4	8,42307	1,17141	0	0,120717	12,3771	2,12847	0	0,016649	12,3771	2,12847	0	0,016649
1.6	9,88591	0,162439	0	0,43548	14,516	0,146033	0	0,441948	14,516	0,146033	0	0,441948
1.8	11,1862	1,42135	0	0,077607	16,4477	1,5347	0	0,062429	16,4477	1,5347	0	0,062429
2	12,3584	2,54863	0	0,005407	18,2171	3,10175	0	0,000962	18,2171	3,10175	0	0,000962

Fonte: Elaboração própria a partir dos dados da PNAD 2014. Nota: Q_mh+= estatística de *Mantel-Haenszel* tendo como hipótese nula a sobrestimação do efeito do tratamento; Q_mh- = estatística de *Mantel-Haenszel* tendo como hipótese nula a subestimação do efeito do tratamento; p+ = probabilidade de não rejeição da hipótese de sobrestimação; p- = probabilidade de não rejeição da hipótese de subestimação.