

ESTUDO TÉCNICO

N.º 24/ 2012

Metodologia de estimação de população em extrema pobreza: um estudo dos “Sem Declaração” e dos “Sem Rendimento” na PNAD

MDS

SAGI

MINISTÉRIO DO DESENVOLVIMENTO SOCIAL E COMBATE À FOME

SECRETARIA DE AVALIAÇÃO E GESTÃO DA INFORMAÇÃO

ESPLANADA DOS MINISTÉRIOS BLOCO A SALA 323

CEP 70054-960 BRASÍLIA DF TEL. 61 3433 1501

Estudo Técnico

No. 24/2012

Metodologias de estimação de população em extrema pobreza: um estudo dos “Sem Declaração” e dos “Sem Rendimento” na PNAD

Técnico responsável

Alexander Cambraia N. Vaz

Revisão

Paulo de Martino Jannuzzi

Estudos Técnicos SAGI é uma publicação da Secretaria de Avaliação e Gestão da Informação (SAGI) criada para sistematizar notas técnicas, estudos exploratórios, produtos e manuais técnicos, relatórios de consultoria e reflexões analíticas produzidas na secretaria, que tratam de temas de interesse específico do Ministério de Desenvolvimento Social e Combate à Fome (MDS) para subsidiar, direta ou indiretamente, o ciclo de diagnóstico, formulação, monitoramento e avaliação das suas políticas, programas e ações.

O principal público a que se destinam os Estudos são os técnicos e gestores das políticas e programas do MDS na esfera federal, estadual e municipal. Nesta perspectiva, são textos técnico-científicos aplicados com escopo e dimensão adequados à sua apropriação ao Ciclo de Políticas, caracterizando-se pela objetividade, foco específico e tempestividade de sua produção.

Futuramente, podem vir a se transformar em artigos para publicação na Cadernos de Estudos, Revista Brasileira de Monitoramento e Avaliação (RBMA) ou outra revista técnica-científica, para alcançar públicos mais abrangentes.

Palavras-chave: *extrema pobreza, análise multivariada, modelagem estatística*

Unidade Responsável

Secretaria de Avaliação e Gestão da Informação

Esplanada dos Ministérios | Bloco A | Sala 307

CEP: 70.054-906 Brasília | DF

Fone: 61 3433-1501 | Fax: 3433-1529

www.mds.gov.br/sagi

Secretário de Avaliação e Gestão da Informação

Paulo de Martino Jannuzzi

Secretária Adjunta

Paula Montagner

Apresentação

Este Estudo Técnico tem por objetivo apresentar duas metodologias- uma de natureza probabilística, outra determinística- para identificar a população em extrema pobreza dentre aquelas famílias que, no âmbito do Censo Demográfico Nacional (Censo) e a Pesquisa Nacional por Amostra de Domicílios (PNAD), optaram por não declarar suas respectivas rendas e, por isso, ou foram classificados como “Sem Declaração”, como no caso da PNAD, ou foram classificados como “Sem rendimento”, no caso do Censo. A classificação dos indivíduos nestas categorias tem sido recorrente ao longo do tempo e, mais importante, vem demonstrando, no caso da segunda, tendência de aumento, podendo levar a subestimação de população em extrema pobreza.

1. Contextualização

Um dos eixos de políticas públicas do governo da presidenta Dilma Rousseff consiste não apenas na manutenção das políticas sociais e de combate à pobreza do governo anterior, mas num aprofundamento da estratégia. E a tarefa tomou contornos claros já na época de campanhas eleitorais, transmutada na estratégia de superação e erradicação da extrema pobreza no Brasil. A meta correlata, divulgada em meados do ano de 2010, início de 2011, consistia em tirar da situação de extrema pobreza pelo menos 16 milhões de indivíduos - que se estimou estarem nesta situação.

Se a superação da pobreza é importante, não é difícil perceber que as dificuldades correlatas são significativas, especialmente num país que tem priorizado tal dimensão efetivamente somente nos últimos 10 anos. A primeira grande barreira, pode-se dizer, consiste justamente na estimação inicial do número de extremamente pobres. As pesquisas e bases de dados sóciodemográficos de referência do Brasil, quais sejam, os Censos e as PNADs, ambos do IBGE, são baseadas numa atividade de coleta que é empreendida num mês específico do ano. Essa periodicidade de captação pode gerar um problema concernente especificamente à renda dos indivíduos pesquisados, como bem observa Osorio et al (2011):

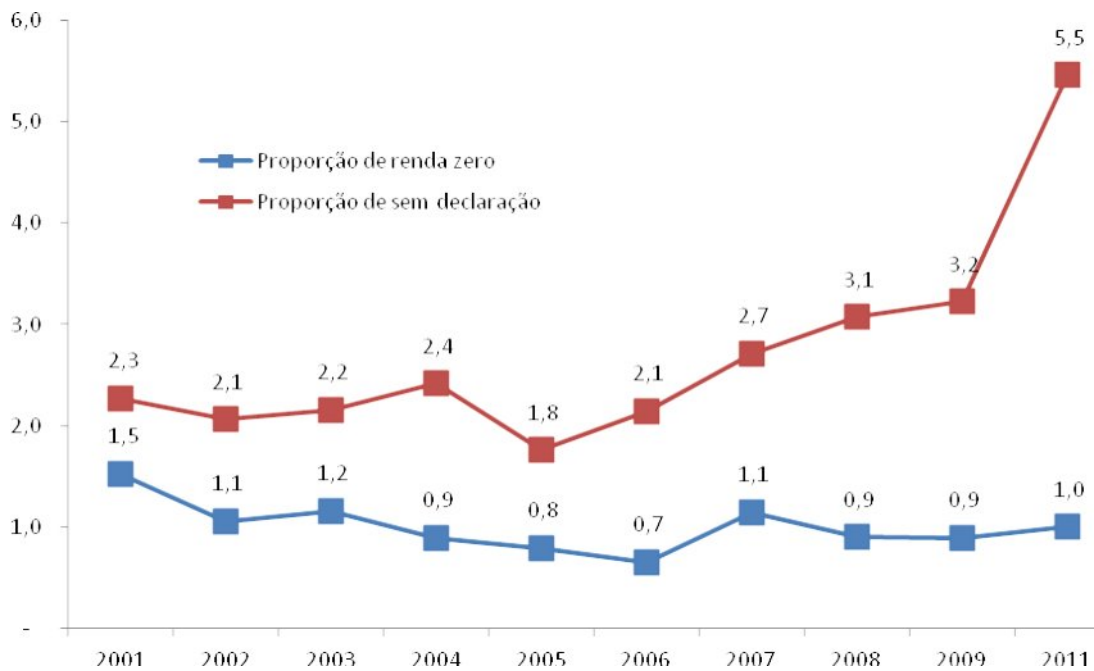
“Uma família pobre estar com renda zero no mês da observação por causa de decisões individuais dos provedores de renda, como investir tempo em qualificação ou mudança de emprego; ou também por causa de choques adversos e temporários, como falta de trabalho

para trabalhadores autônomos, incapacidade temporária ou desemprego. A causa não é relevante, desde que seus efeitos sejam temporários e estas famílias tenham renda permanente elevada, contando com outros mecanismos, como poupança ou mesmo a solidariedade de parentes, para se proteger da falta temporária de renda" (OSORIO, R. G.; SOARES, S. S. D.; SOUZA, P. H. G. F., 2011, Pg. 34).

A captação de dados em períodos específicos do ano, como se observa, pode levar à situação de encontrarmos indivíduos categorizados como "sem rendimento" nas bases de dados, os quais, tomando por base um critério de classificação de pobreza e/ou extrema pobreza calcado apenas na variável renda, seriam assim categorizados - ainda que, na realidade, não o fossem, mas estivessem apenas momentaneamente sem renda. Além disso, se considerarmos as pesquisas PNADs, as quais contam, ainda, com a categoria de classificação "Sem declaração", não é difícil perceber que uma categorização baseada apenas na renda tenda a subestimar o número real de pobres e extremamente pobres, já que indivíduos passíveis de classificação nesta situação eventualmente poderiam não declarar suas respectivas rendas.

Essa situação adquire contornos particularmente dramáticos na perspectiva de planejamento na área de políticas sociais e, especialmente, na estruturação ou fortalecimento de programas para superação da extrema pobreza, se considerarmos que, ao longo dos anos, a categoria "Sem declaração de rendimento", tem aumentado significativamente no âmbito das pesquisas realizadas. O Gráfico 1 confere um panorama claro dessa dinâmica: enquanto a parcela de pessoas com renda zero permanece em torno de 1% ao longo dos anos 2000, a parcela de pessoas em situação de "Sem rendimento" dobra entre 2001 e 2011, passando de 2,3% para 5,5%.

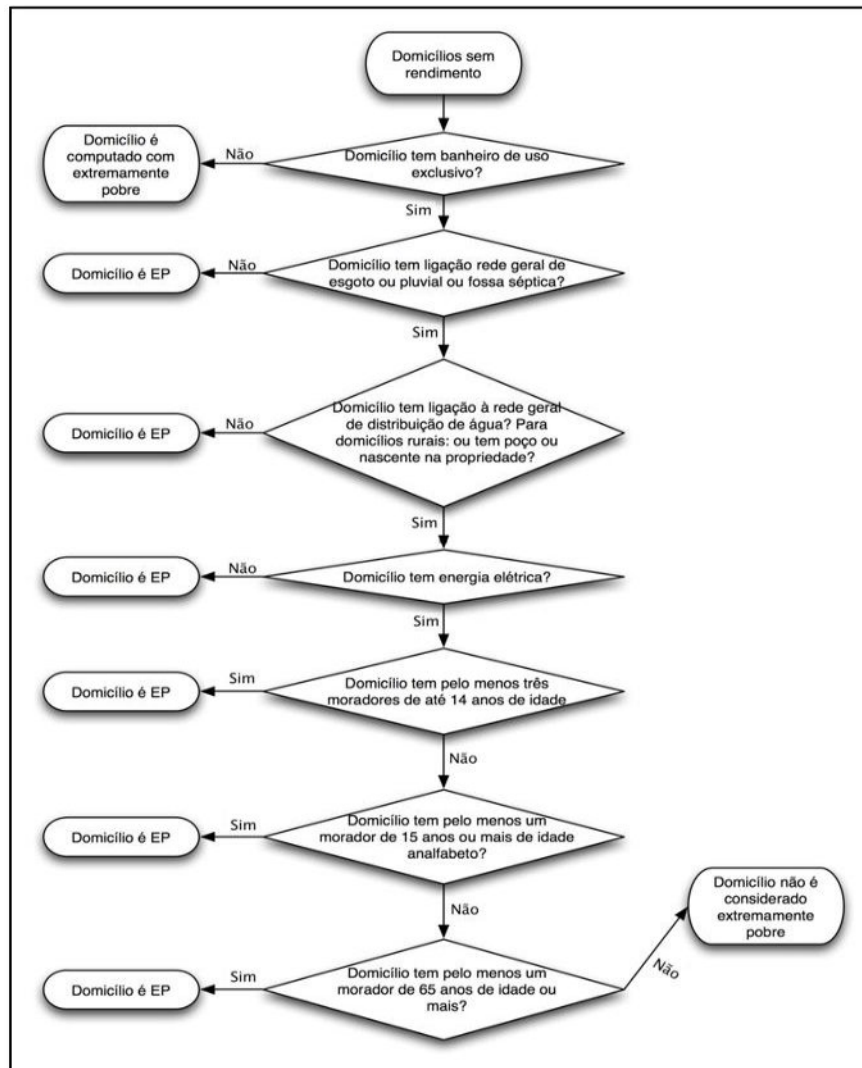
Gráfico 1 - Proporção da população residente com rendimento zero e sem declaração nas PNADs, Brasil 2001-2011



Fonte: PNADs IBGE – 2001-2011

A título ilustrativo, pode-se dizer que uma das formas de lidar com essa problemática consiste na utilização de um filtro preditor da categoria de extrema pobreza, levando em consideração outros tipos de variáveis para além da renda Sousa e Santos (2012). Nesse tipo de metodologia, prevê-se que, dentre uma série de variáveis, caso um indivíduo ou domicílio se adeque a pelo menos uma das condições estipuladas, será considerado extremamente pobre. A Figura 1 apresenta um exemplo de carta de variáveis e o fluxo de categorização de extrema pobreza passível de utilização nesse filtro.

Figura 1 – Variáveis e fluxo de cálculo utilizados no filtro para estimação da extrema pobreza no Censo Demográfico 2010



De fato, é possível afirmar que as variáveis trazidas à baila através da metodologia adotada no filtro são predictoras significativas da condição de pobreza, pois exercem significativa diferenciação entre os domicílios classificados na categoria “Sem rendimento” quando aplicadas seguindo o fluxo estipulado. Uma breve análise do filtro, especialmente no que tange à pertinência das variáveis elegidas como condicionais da situação de extrema pobreza, evidencia a capacidade da maioria das variáveis trabalhadas de discriminar de maneira significativa os indivíduos no que tange às suas condições socioeconômicas.

A seu turno, pode-se dizer que os resultados advindos da aplicação da técnica de restrição indicam claramente pelo menos duas questões. Primeiro, é inegável a necessidade de prospecção mais pormenorizada da categoria de extrema pobreza, notadamente no que

tange à consideração de outras variáveis para além da renda como discriminantes deste conjunto populacional. As técnicas utilizadas até este momento têm mostrado resultados significativos, mas é importante reconhecer a necessidade de se utilizar técnicas que, embora também discriminantes e determinísticas, contem com menor grau de arbitrariedade metodológica, especialmente na forma de correlação entre as variáveis tratadas. Neste sentido, sugere-se, a técnica de Análise fatorial pode cumprir importante papel, conforme se observará na próxima seção. A segunda consideração concerne exatamente à validação da cesta de variáveis trazidas à baila no filtro e que, portanto, sugere-se, podem ser mantidas nas análises e estudos posteriores.

2. Técnica probabilística de estimativa de renda: a utilização da Análise Fatorial

A principal hipótese deste estudo é de que a estruturação da categoria de extrema pobreza deve ser capaz de abarcar, nas pesquisas disponíveis, os indivíduos que ou não tem rendimentos, ou não declaram rendimentos. Para isso, parte do pressuposto tanto de que outras variáveis devem ser consideradas para além da renda, quanto de que a relação entre estas variáveis deve contar com o menor nível de arbitrariedade possível. Para tratar a hipótese trazida à baila respeitando os pressupostos aqui levantados, o estudo sugere, como primeira abordagem, a utilização da técnica de Análise Fatorial com fins de estabelecer perfis correlatos ao público extremamente pobre, ancorados em características não estritamente econômicas (como a renda), mas baseados, também, em variáveis de caráter socioeconômico e de condições de vida/domiciliares. Estes perfis são aplicados tanto aos indivíduos que se declaram sem rendimentos, quanto àqueles que não declaram rendimentos. A técnica é lastreada para os anos 2000 a 2012 e toma por base diferentes bases de dados, quais sejam, os Censos 2000 e 2010, as PNADS 2001 a 2011 e, por fim, a pesquisa PADS 2012.

Alguns procedimentos serão seguidos. Primeiro, a cesta de variáveis a serem trabalhadas. Neste caso, estamos nos apoiando tanto nos trabalhos já desenvolvidos pelo IBGE, quanto em publicações diversas da área, como Sousa e Santos (2012), dentre outros. O Quadro 1 descreve as variáveis utilizadas:

Quadro 1 – Variáveis utilizadas no filtro construído pelo IBGE para da extrema pobreza

Código da variável	Variável derivada	Categoria derivada	Abreviação da categoria derivada	Categoria original
Variáveis componentes do filtro de restrições do IBGE				
V8005	Existência de três moradores ou mais de até 14 anos de idade	Bom	Mor14_3_sim	Variável categórica da base de pessoas agregada para a base de domicílios
		Ruim	Mor14_3_não	
V6036	Existência de morador de 65 anos ou mais	Bom	Ido_sim	Variável categórica da base de pessoas agregada para a base de domicílios
		Ruim	Ido_não	
Variáveis adicionais para análise do filtro de restrições				
V0201	Condição de ocupação do domicílio - Domicílio próprio?	Bom	Imo_sim	Próprio de algum morador - já pago
				Próprio de algum morador - ainda pagando
		Ruim	Imo_não	Alugado
				Cedido por empregador
				Cedido de outra forma
Outra condição				
V0202	Material predominante nas paredes externas - Alvenaria ou madeira aparelhada?	Bom	Par_sim	Alvenaria com revestimento
				Alvenaria sem revestimento
				Madeira apropriada para construção (aparelhada)
		Ruim	Par_não	Taipa revestida
				Taipa não revestida
				Madeira aproveitada
				Palha
Outro material				
V6204	Densidade de morador por dormitório – Até 2,00 por cômodo	Bom	Den_sim	Até 2,00
		Ruim	Den_não	Acima de 2,00
V0210	Destino do lixo - Coletado por serviço de limpeza?	Bom	Lixo_sim	Coletado diretamente por serviço de limpeza
				Colocado em caçamba de serviço de limpeza
		Ruim	Lixo_não	Queimado (na propriedade)
				Enterrado (na propriedade)
				Jogado em terreno baldio ou logradouro
Jogado em rio, lago ou mar				
Tem outro destino				
V0213	Existência de rádio	Bom	Radio_sim	Sim
		Ruim	Radio_nao	Não
V0214	Existência de televisão	Bom	TV_sim	Sim
		Ruim	TV_nao	Não
Código da variável	Variável derivada	Categoria derivada	Abreviação da categoria derivada	Categoria original

Variáveis adicionais para análise do filtro de restrições				
V0215	Existência de máquina de lavar roupa	Bom	Mqlav_sim	Sim
		Ruim	Mqlav_nao	Não
V0216	Existência de geladeira	Bom	Gel_sim	Sim
		Ruim	Gel_nao	Não
			Aut_nao	Não
V6400	Nível de instrução	Ruim	NI_SI_FI	Sem instrução e fundamental incompleto
			NI_FC_MI	Fundamental completo e médio incompleto
		Bom	NI_MC_SI	Médio completo e superior incompleto
			NI_SC	Superior completo
			NI_ND	Não determinado

Fonte: IBGE (2011) apud Sousa e Santos (2012)

As variáveis são estruturadas em formato dummy, em termos de existência ou não de determinado atributo, tanto com relação aos domicílios, quanto com relação aos indivíduos. Assim, por exemplo, quando consideramos a presença ou não de geladeira no domicílio, temos duas variáveis: uma em que o sucesso (1) é ter geladeira e o fracasso (0) é não ter geladeira, chamada de Gel_Bom; e temos também outra variável, na qual a lógica é invertida, isto é, o sucesso (1) é não ter geladeira e o fracasso (0) é ter geladeira, chamada de Gel_Ruim.

Essa lógica de classificação das variáveis analisadas é importante em função da estrutura de correlação a ser estruturada entre cada qual. A análise fatorial permite o exame da estrutura correlações entre um grande número de variáveis, definindo um conjunto de dimensões latentes comuns, chamados de fatores (Triola, 2009) – estruturando-se pela estimativa de verossimilhança do i -ésimo indivíduo na l -ésima categoria da j -ésima variável (Sawyer et al, 2002). Os resultados apresentam o grau que cada variável é explicada por cada fator¹.

Os objetivos principais da análise fatorial são diminuir a dimensionalidade das variáveis e, ao mesmo tempo, gerar um novo instrumento de explicação, que será mais parcimonioso e poderá ser utilizado em substituição às variáveis individuais originais para se explicar um

¹ Os fatores baseiam-se, primeiro, na criação de perfis de referência, ou conjuntos de parâmetros puros, λ_{kjl} , que é a probabilidade (λ) da resposta (l) à variável (j) no perfil (k), que nada mais é do que os tipos de resposta que caracterizam dado perfil latente. Em segundo lugar, na estimativa, para cada caso da amostra, g_{ik} , de pertencimento (g) do indivíduo (i) ao perfil (k), que não tem caráter probabilístico, mas se refere especificamente às características apresentadas por cada indivíduo.

fenômeno. Além disso, permite também a organização hierarquizada da informação disponível. A vantagem principal da utilização da técnica é que os fatores finais são formados a partir das características compartilhadas dos públicos-alvo. Assim, é possível analisar as relações entre as variáveis, verificando seus respectivos graus de associação e peso relativo em cada fator criado, tendo por base as inter-relações estimadas entre os diferentes públicos da amostra (Triola, 2010).

Outra característica da análise fatorial é que ela busca explicar a variação comum, ou seja, aquela que é compartilhada com outras variáveis e não a variação total, como a análise de componentes principais². A vantagem é que os fatores são formados a partir das características compartilhadas entre os casos da amostra. A proporção da variância de cada variável explicada pelas componentes principais retidas designa-se comunalidade. As comunalidades iniciais são iguais a 1. Após a extração, variam entre 0 e 1, sendo zero quando os fatores comuns não explicam nenhuma variância da variável e 1 quando explicam toda a sua variância.

Neste trabalho, a aplicação da técnica concerne ao conjunto de variáveis e indicadores componentes da categoria de extrema pobreza. Assim, tomou-se por base central de análise o público com renda declarada que varia de R\$ 0,01 a R\$ 70,00, sendo o objetivo final a geração de subgrupos ou fatores correlatos a essa faixa de renda capazes de diferenciar os indivíduos ou domicílios em diferentes subconjuntos ditos extremamente pobres. Estes subconjuntos são, então, tomados como base para prospecção das mesmas características dentro dos grupos "Sem declaração" e "Sem rendimento".

Com base na geração dos fatores e/ou subconjuntos de extrema pobreza, a Análise Fatorial calcula, através de uma análise de confiabilidade e redundância, o grau de pertinência de cada caso da amostra considerada aos fatores construídos. A avaliação destes resultados foi realizada através da análise do Alfa de Cronbach, que é uma medida de confiabilidade que varia entre 0 e 1, sendo os valores de 0,60 a 0,70 considerados o limite inferior de aceitabilidade (HAIR et al, 2005). Neste trabalho, utilizou-se o valor de 0,70 como linha de corte para indicar pertencimento dos casos aos fatores construídos.

² A Análise de Componentes Principais é um modelo fatorial no qual os fatores são baseados na variância total, nesta análise as unidades (1s) são usadas na diagonal da matriz de correlação; o que implica que toda a variância é comum ou compartilhada (HAIR et al, 2005).

2.2 Resultados preliminares da Análise Fatorial

A aplicação da técnica de Análise Fatorial levou em consideração as linhas de pobreza e extrema pobreza definidas em 2011, que se baseiam no valor de R\$ 70,00 como “linha de corte” administrativa. Para todos os anos, estes valores foram deflacionados ou inflacionados tendo por base o INPC (Soares et al, 2011). Foram gerados três perfis de referência, como visto no Quadro 2:

Quadro 2 – Perfis de extrema pobreza gerados pela Análise Fatorial com base na faixa de renda R\$ 0,01 a R\$ 70,00 (Jun 2011), Brasil 2010

	Perfil 1	Perfil 2	Perfil 3
Características dos perfis	AGUA Ruim	AGUA Ruim	AGUA Bom
	ESGOTO Ruim	ESGOTO Ruim	ESGOTO Bom
	BANHEIRO Ruim	BANHEIRO Ruim	BANHEIRO Bom
	ENERGIA Ruim	ENERGIA Bom	ENERGIA Ruim
	PARDSEXT Ruim	PARDSEXT Ruim	PARDSEXT Bom
	DNSDOR Ruim	DNSDOR Ruim	DNSDOR Bom
	TV Ruim	TV Bom	TV Ruim
	GELADEIRA Ruim	GELADEIRA Bom	GELADEIRA Ruim
	MAQ_ROUPA Ruim	MAQ_ROUPA Ruim	MAQ_ROUPA Ruim
	NIVEL_INST Ruim	NIVEL_INST Ruim	NIVEL_INST Ruim
	EDUC_FAM Ruim	EDUC_FAM Ruim	EDUC_FAM Ruim

Fonte: Censo Demográfico IBGE 2010

No primeiro perfil, temos uma situação de indivíduos e/ou domicílios ruins em todos os aspectos considerados, desde características domiciliares específicas, como canalização de água e acesso à rede de esgoto, até no caso das características familiares, como nível de instrução do chefe e dos próprios familiares.

Tabela 2 - Grau de congruência das faixas "Sem declaração" e "Sem rendimento" aos fatores/perfis gerados na Análise Fatorial, por ano de pesquisa, Brasil 2000-2012

		Por Perfil			Congruência	Incongruência
		1	2	3	total	total
2000	Sem rendimento	15,48%	14,47%	9,13%	39,07%	60,93%
		846108	790842	499047	2135997	3330911
2001	Sem rendimento	20,53%	38,92%	5,85%	65,30%	34,70%
		420905	798066	119952	1338923	711536
	Sem declaração	62,79%	15,21%	3,86%	81,86%	18,14%
		2069980	501514	127116	2698610	597811
2002	Sem rendimento	9,43%	24,44%	37,98%	71,85%	28,15%
		143714	372209	578518	1094441	428765
	Sem declaração	6,89%	12,60%	69,36%	88,84%	11,16%
		217722	398180	2192486	2808388	352809
2003	Sem rendimento	6,47%	40,27%	6,03%	52,76%	47,24%
		108977	678697	101542	889216	796024
	Sem declaração	1,94%	70,21%	2,06%	74,20%	25,80%
		64290	2325550	68102	2457942	854466
2004	Sem rendimento	11,21%	7,92%	41,26%	60,38%	39,62%
		144845	102394	533228	780467	512020
	Sem declaração	1,28%	3,29%	69,76%	74,32%	25,68%
		49674	127953	2714934	2892561	999497
2005	Sem rendimento	15,81%	18,10%	9,79%	43,71%	56,29%
		194374	222565	120407	537346	691983
	Sem declaração	2,79%	32,92%	42,41%	78,12%	21,88%
		81365	960215	1236114	2276973	637694
2006	Sem rendimento	13,75%	44,08%	5,82%	63,64%	36,36%
		135118	433254	57213	625585	357349
	Sem declaração	22,03%	17,33%	4,87%	44,23%	55,77%
		818361	643863	181065	1643289	2072011
2007	Sem rendimento	8,01%	3,64%	58,05%	69,70%	30,30%
		172224	78230	1248437	1498891	651621
	Sem declaração	1,94%	3,24%	70,07%	75,25%	24,75%
		98594	164726	3565780	3829100	1259517
2008	Sem rendimento	6,50%	40,62%	41,70%	88,82%	11,18%
		110723	692393	710759	1513875	190501
	Sem declaração	1,55%	66,17%	23,51%	91,23%	8,77%
		90623	3856267	1370431	5317321	510903
2009	Sem rendimento	5,90%	53,18%	3,82%	62,90%	37,10%
		100470	906125	65078	1071673	632110
	Sem declaração	1,19%	76,73%	2,92%	80,83%	19,17%
		73431	4734076	180097	4987604	1182575
2010	Sem rendimento	33,47%	2,44%	10,00%	45,92%	54,08%
		2332439	170325	697077	3199841	3768701
2011	Sem rendimento	1,16%	44,77%	6,02%	51,94%	48,06%
		22524	872951	117347	1012822	937099
	Sem declaração	0,73%	31,99%	3,58%	36,30%	63,70%
		77549	3398364	380311	3856224	6768245
2012	Sem declaração	5,11%	47,90%	5,99%	59,00%	41,00%
		99036	928254	116052	1143342	794511

Fonte: PNADs IBGE 2001-2011; Censos Demográficos 2000 e 2010; PADS 2012

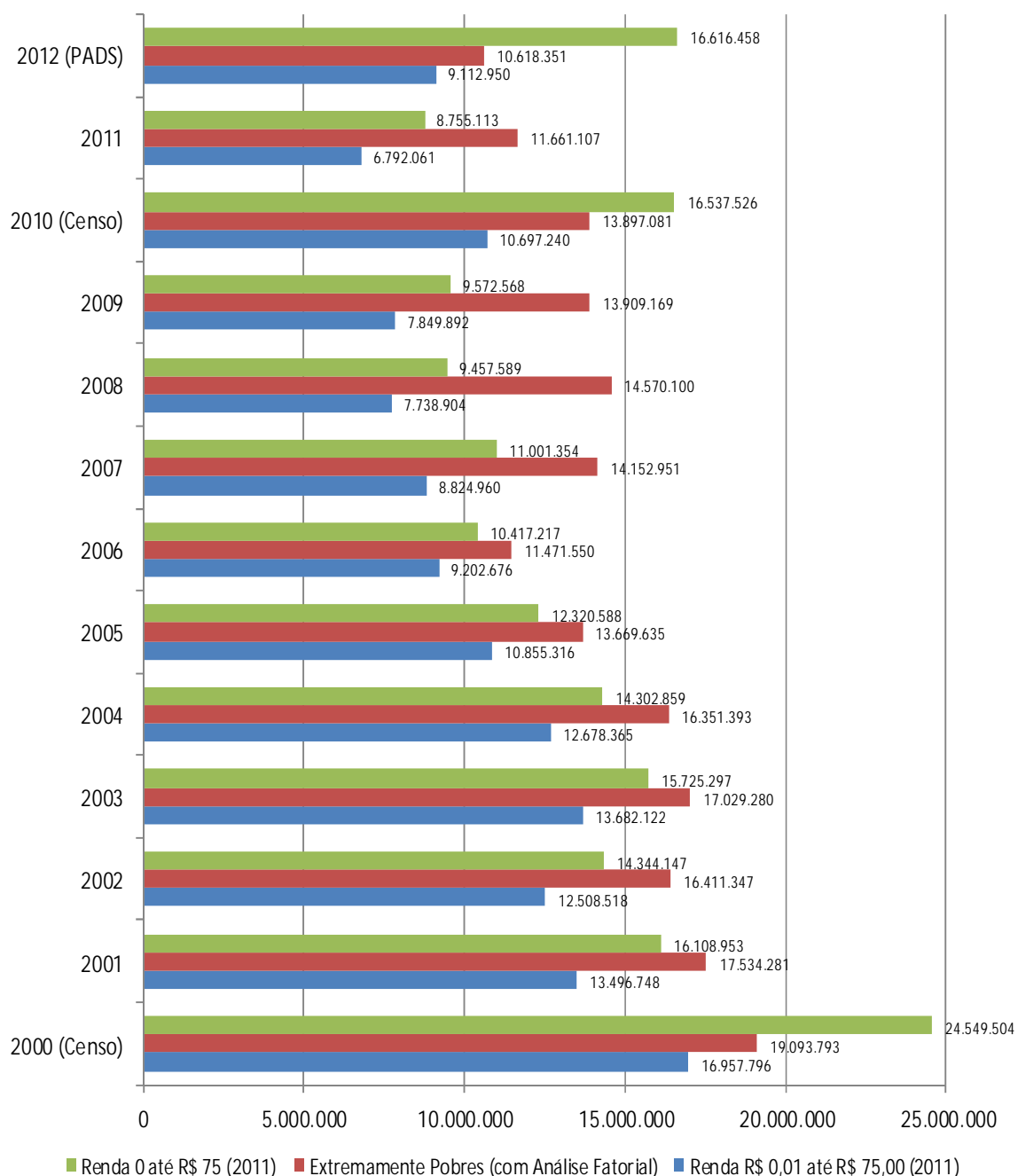
No segundo perfil, encontramos uma situação de melhoria no acesso a determinados atributos em relação ao primeiro perfil, como acesso a energia elétrica e alguns bens, notadamente geladeira e TV. No último perfil, temos também uma situação de acesso a determinados atributos, mas diferentes daqueles do perfil 2, como esgotamento sanitário, melhor condição das paredes externas do domicílio e melhor densidade do dormitório. Ao aplicarmos estes perfis nos demais anos de pesquisa e bases de dados e adotando o já

referido nível de confiabilidade de Cronbach de 0,7, obtivemos os graus de pertencimento dos casos contidos nas categorias “Sem rendimento” e “Sem Declaração” descritos na Tab 2. Se considerarmos o Censo 2000, por exemplo, observamos que, dentre aqueles considerados sem rendimento, a Análise Fatorial indicou que quase 40% podem ser categorizados como extremamente pobres. Considerando o Censo 2010, esse percentual sobe a quase 46%. Se pegarmos outro ano aleatoriamente, como 2006, baseado na PNAD, veremos que, dentre aqueles categorizados como Sem rendimento, mais de 60% podem ser classificados como extremamente pobres segundo a Análise Fatorial. Percentual que é de mais 44% no caso da categoria de Sem declaração.

Ao aplicarmos a Análise Fatorial, portanto, obtemos um incremento do número de extremamente pobres em relação ao número que seria obtido caso considerássemos apenas a variável renda. O Gráfico 2 realiza uma comparação, para todos os anos, entre três critérios de classificação de extrema pobreza: aquela que considera a renda declarada (que varia de R\$ 0,01 a R\$ 70,00 – valores de Junho de 2011); aquela baseada na técnica de Análise Fatorial, tendo por base os valores de congruência atestados na Tabela 2 para as categorias Sem rendimento e Sem declaração, somados aos valores encontrados na faixa de renda declarada de R\$ 0,01 a R\$ 70,00 (Junho 2011); e, por fim, o critério “puro” que considera a renda zero (Sem rendimento) até R\$ o valor de 70,00 (Junho de 2011).

O que podemos observar, em primeiro lugar, é que há significativa diferença entre os três critérios de classificação em todos os anos considerados. Se considerarmos o ano de 2000, por exemplo, a análise da categoria de extrema pobreza a partir apenas do critério de renda (R\$ 0,01 a R\$ 70,00 – Junho 2011) estipula a existência de aproximadamente 17 milhões de indivíduos nesta situação no país. Porém, quando aplicamos a análise fatorial, o número de indivíduos passíveis de serem classificados nesta definição é acrescido em mais de 2 milhões, passando para quase 19 milhões e 100 mil indivíduos. Nessa mesma linha, se considerarmos também a pesquisa Censo, mas, desta vez, o ano de 2010, observamos resultados semelhantes. Observamos, para o critério de renda R\$0,01 a R\$ 70,00, um número aproximado de extremamente pobres de 10 milhões e 700 mil indivíduos. No entanto, quando analisamos essa categoria à luz da análise fatorial, essa quantia sobe para quase 13 milhões e 900 mil indivíduos.

Gráfico 2 – Comparação entre número de extremamente pobres via três critérios de classificação: renda declarada, análise fatorial e renda total, Brasil 2000-2012



Fonte: PNADs IBGE 2001-2011; Censos Demográficos 2000 e 2010; PADS 2012

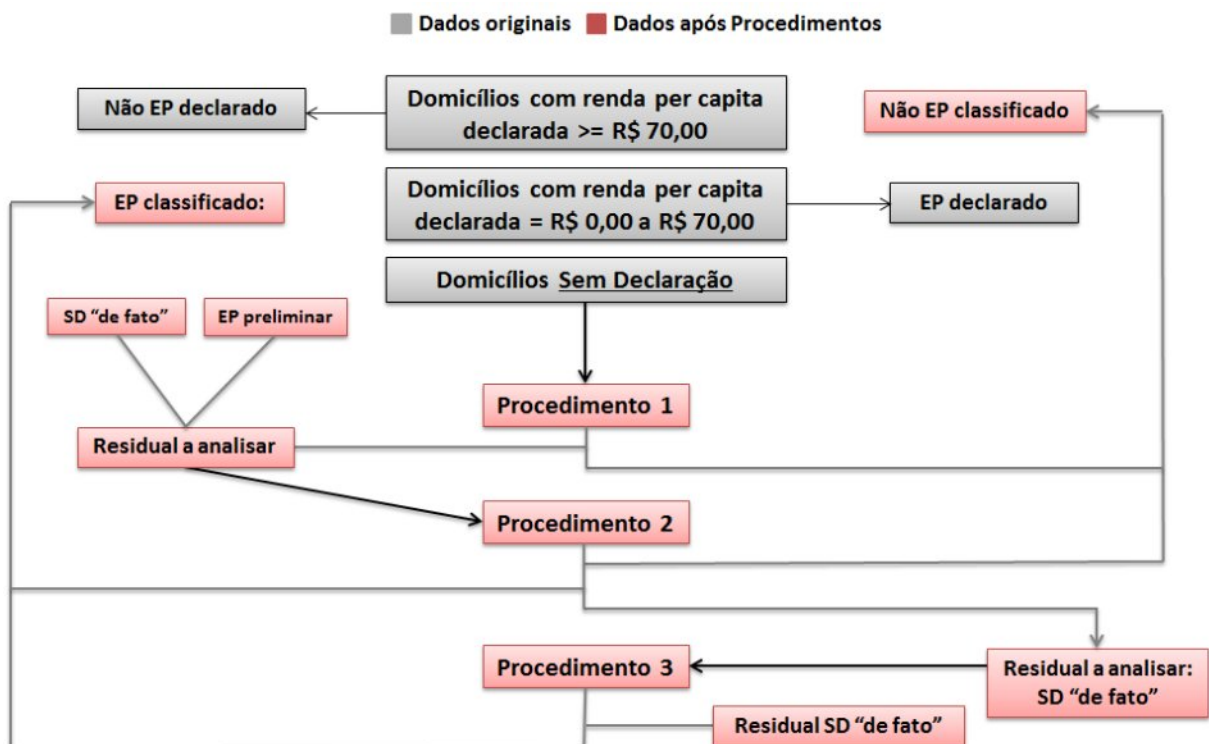
Por fim, quando consideramos uma base de dados diferente, notamos que há consistência nas diferenças. Se pegarmos, por exemplo, o ano de 2009, referente à PNAD, notamos que, enquanto o critério de renda R\$ 0,01 a R\$ 70,00 estima a existência de aproximadamente 7 milhões e 900 mil indivíduos extremamente pobres, o critério da análise fatorial estima a existência de mais de 9 milhões e 500 mil indivíduos nesta situação.

2. Técnica determinística de estimativa de renda

A segunda técnica de estimativa de renda focada neste estudo consiste na adoção de determinados pressupostos considerados válidos e plausíveis como aporte para imputação de renda aos indivíduos dos grupos “Sem Declaração” e “Sem rendimento”. São 3 (três) procedimentos a serem seguidos e que se encontram ilustrados na Figura 1.

O primeiro procedimento metodológico consiste na busca de outras fontes de renda declaradas pelo indivíduo. A estrutura metodológica das pesquisas aqui trabalhadas, especialmente Censo e PNAD, contém dada peculiaridade. Caso determinado indivíduo declare todas as rendas monetárias, mas deixa de fazê-lo em relação a pelo menos uma, ele já passa a ser automaticamente classificado dentre os grupos “Sem declaração” ou “Sem rendimento”. Em outras palavras, basta o indivíduo deixar de declarar apenas uma renda, mesmo tendo declarado todas as demais, para que seja classificado nos grupos supracitados.

Figura 1 – Fluxograma de aplicação dos procedimentos metodológicos de imputação de renda para os indivíduos “Sem declaração” e “Sem rendimento”



Fonte: elaboração própria

Assim, o primeiro procedimento de imputação determinística consiste na busca, para cada indivíduo localizado nestes grupos, de outras rendas eventualmente declaradas. Uma vez encontradas, essas rendas são, então, somadas e é calculada a renda per capita domiciliar. Se essa renda per capita ultrapassar os R\$ 70,00 (Junho 2011), então os indivíduos já não são considerados extremamente pobres, mesmo tendo sido, inicialmente, considerados “Sem declaração” ou “Sem rendimento”.

O segundo procedimento metodológico é realizado a partir da coorte ainda não tratada após o procedimento número 1. Neste caso, busca-se, dentre os indivíduos dos grupos “Sem declaração” e “Sem rendimento”, aqueles dotados das seguintes características:

- declararam receber renda monetária de Institutos de Previdência Oficiais;
- declararam possuir carteira de trabalho assinada;
- declararam ser militares ou estatutários com jornada de trabalho de 40 horas ou mais.

No caso dos indivíduos dotados destas características, julgou-se plausível atribuir-lhes uma renda mensal de pelo menos 1 salário mínimo, ainda que originalmente classificados como “Sem declaração” ou “Sem rendimento”. A partir da imputação da nova renda, foi calculada a renda domiciliar per capita e, assim como no procedimento metodológico anterior, caso essa renda per capita se localizasse acima da linha de corte de extrema pobreza (de R\$ 70,00 Junho 2011), os indivíduos já poderiam ser considerados como não extremamente pobres.

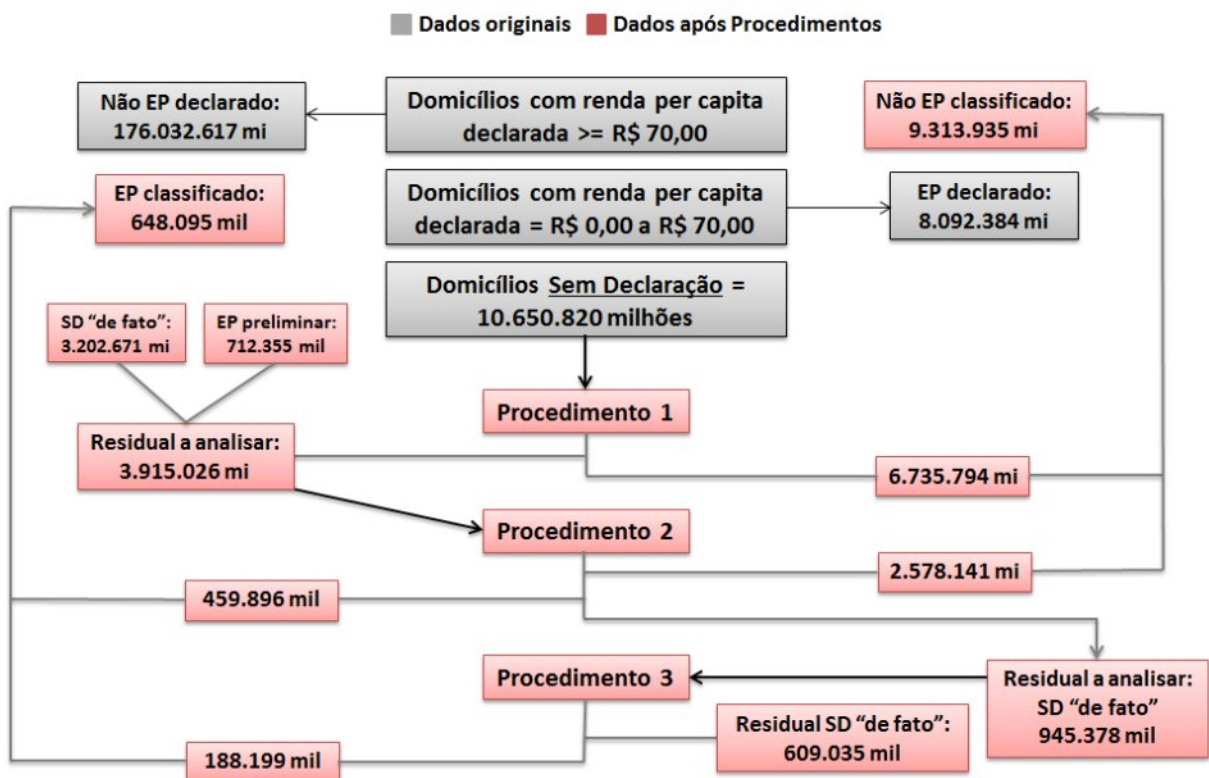
O terceiro procedimento metodológico consiste na discriminação dos grupos de extrema pobreza, dentre os grupamentos “Sem declaração” e “Sem rendimento”, a partir da adoção de quatro variáveis específicas dadas como altamente discriminantes destes grupos via testes nas metodologias anteriores. Estes variáveis compõem um índice que varia até 4 e os indivíduos que chegam a esse valor são considerados extremamente pobres.

A primeira variável discriminante adotada neste sentido é a que concerne à escolaridade do responsável pelo domicílio, sendo que estamos interessados em escolaridade fundamental incompleta ou sem instrução. As demais variáveis são trabalhador sem carteira assinada; domicílio sem acesso a rede geral de água; e domicílio sem acesso à rede de esgotamento sanitário.

3.2 Resultados preliminares da Imputação determinística

Após a adoção dos procedimentos de imputação de renda descritos, obtivemos resultados interessantes do ponto de vista analítico. A Figura 2 provê o esquema de fluxo de aplicação e os resultados para o ano de 2011.

Figura 2 – Fluxograma de aplicação dos procedimentos metodológicos de imputação de renda para os indivíduos “Sem declaração” e “Sem rendimento” – Brasil, 2011



Fonte: PNAD, 2011

Inicialmente, segundo a PNAD 2011, tínhamos aproximadamente 8,1 milhões de indivíduos extremamente pobres para este ano, a partir das informações declaradas de renda. O número de indivíduos “Sem declaração” é de mais de 10 milhões e é sobre este grupo que o método foi aplicado.

A partir do Procedimento 1, verificamos que 6,7 milhões de indivíduos possuíam alguma renda declarada e que o cômputo da renda per capita domiciliar a partir dessa renda os levou a serem classificados como não extremamente pobres. Ou seja, sua renda per capita ficou acima da linha de corte de R\$ 70,00 (Junho 2011). Dos 3,9 milhões restantes a serem

analisados, constatamos que para 3,2 milhões de fato não existia qualquer outra renda declarada e que para 712 mil a renda declarada os classificava como extremamente pobres.

O procedimento 2 foi aplicado sobre o grupo total de 3,9 milhões de indivíduos, já que, mesmo no caso dos 712 mil indivíduos para os quais se encontrou alguma renda no Procedimento 1 e essa renda foi baixa o suficiente para serem classificados como extremamente pobres, não se poderia descartar a possibilidade de encontrar os critérios e atributos do Procedimento 2. Assim, após aplicação deste último, dos 3,9 milhões de indivíduos, 2,578 milhões obtiveram renda que lhes permitiu sobrepujar a linha de extrema pobreza e aproximadamente 460 mil foram, a seu turno, classificados como extremamente pobres.

O procedimento 3 envolveu indivíduos que não foram “filtrados” nos procedimentos 1 e 2, ou seja, indivíduos caracterizados por não terem qualquer renda monetária declarada e cujas ocupações não se encaixam em nenhum dos critérios do procedimento 2. O procedimento 3, assim, buscou tratar um grupo restante de 945 mil indivíduos, dentre os quais 188,1 mil foram classificados como extremamente pobres e o restante, 609 mil, de fato não foram passíveis de classificação.

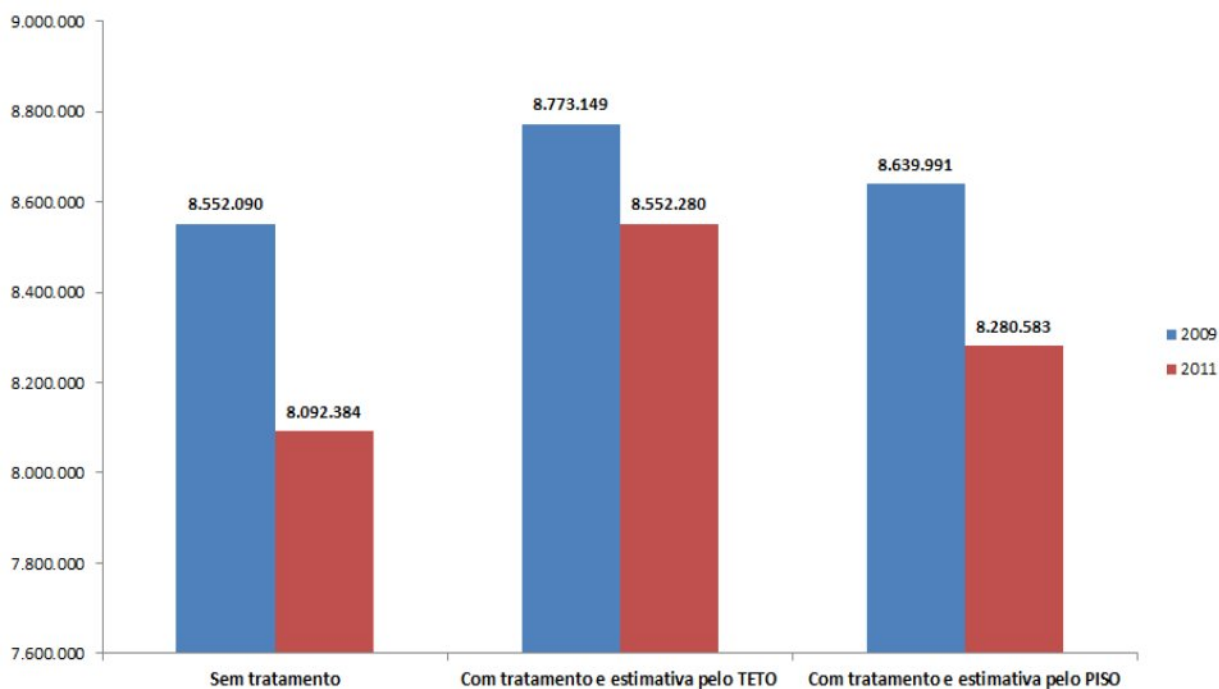
Assim, a partir da adoção dos procedimentos, somando os indivíduos extremamente pobres aferidos nas etapas 2 e 3, temos um total efetivo, em 2011, de 8,7 milhões de indivíduos em extrema pobreza, que é a soma da estimativa original de pobreza com o total de 648 mil indivíduos classificados originalmente como “Sem declaração”, mas dados como extremamente pobres após aplicação dos procedimentos metodológicos³.

A quantidade de extremamente pobres via adoção do modelo pode ser vista de maneira comparativa no Gráfico 3. Podemos perceber claramente as diferenças nos resultados. Em 2009, por exemplo, a aplicação método até a etapa 3 faz com o que o número de extremamente pobres varie de 8,55 milhões (que é a estimativa original baseada na renda declarada), para 8,77 milhões, que é a estimativa do Teto, isto é, considerando a soma do número de extremamente pobres dos procedimentos 2 e 3 em especial. Se considerarmos apenas o número de extremamente pobres do procedimento 3, teremos uma estimativa pelo Piso, que é de 8,63 milhões.

³ Exercício semelhante foi feito para o ano de 2009 e se encontra no Anexo A.

Para 2011, os valores variam de maneira semelhante. A estimativa original de extremamente pobres é de quase 8,1 milhões de indivíduos. Sua variação, após aplicação dos procedimentos, é de aproximadamente 500 mil, chegando a 8,55 milhões quando adotamos a estimativa pelo Teto. Já a estimativa pelo Piso gira em torno de 8,28 milhões de indivíduos extremamente pobres no país.

Gráfico 3 – Comparação entre Estimativas de População em Extrema Pobreza sem e com tratamento de Sem Declaração – PNAD 2009/2011 (R\$ 70 de Junho de 2011)

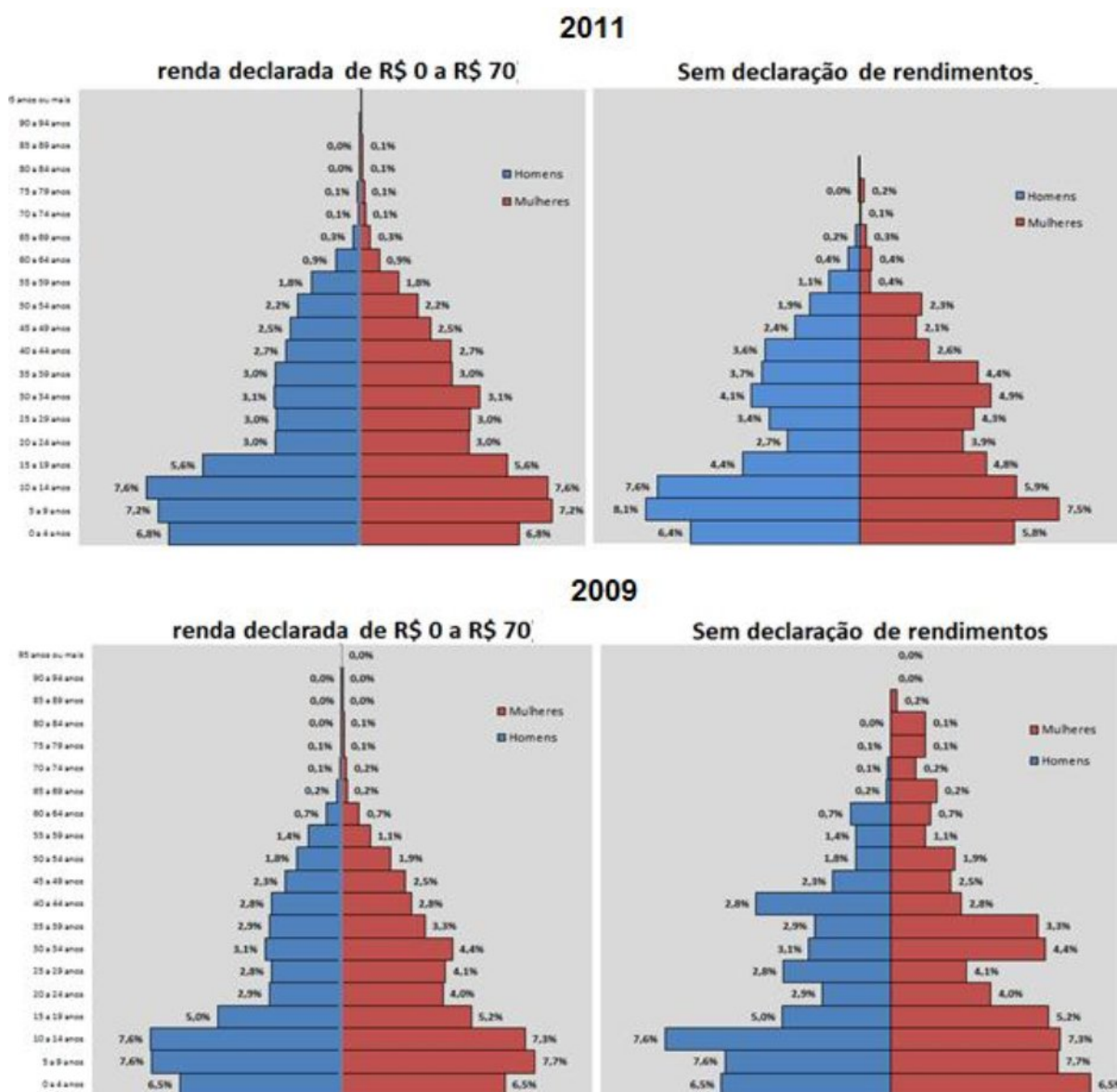


Fonte: PNAD 2011, 2009

Para verificar a adequação do modelo final e a plausibilidade de utilização do método determinístico, fizemos alguns testes comparativos entre as características de perfil demográfico do grupo classificado como extremamente pobre via renda declarada e daquele grupo que recebeu essa classificação após aplicação dos métodos.

A Figura 4 fornece panorama nesse sentido ao comparar o perfil etário dos dois grupos, considerando os anos de 2009 e 2011. Não é difícil perceber a semelhança entre as pirâmides etárias. No caso de 2011, é notável a concentração de indivíduos na base da pirâmide e a suavização ao longo de sua trajetória ascendente, sendo tal variação significativamente semelhante para os dois grupos comparados. No caso de 2009, observamos que, a despeito de algumas diferenças mais notáveis, tanto a tendência, quanto a caracterização geral de ambas as pirâmides guardam também significativa semelhança.

Figura 4 – Comparação entre perfis de População em Extrema Pobreza sem e com tratamento de Sem Declaração –Brasil, 2009/2011 (R\$ 70 de Junho de 2011)



Fonte: PNAD, 2009 – 2011

3. Considerações Finais

Este estudo teve por objetivo construir um modelo de estimação da categoria de extrema pobreza capaz de discriminar indivíduos e/ou domicílios desde uma perspectiva que não estivesse restrita à variável renda. A hipótese principal de trabalho é de que as categorias Sem rendimento e Sem declaração, contidas nas bases de dados disponíveis, consistem em repositórios significativos de indivíduos passíveis de classificação como extremamente

pobres e, por isso, estimativas baseadas apenas na variável renda tendem a subestimar o número efetivo desta categoria. Dois pressupostos foram tomados em consideração: primeiro, as variáveis trabalhadas e preditoras de extrema pobreza são aquelas definidas pelo IBGE e já testadas por Sousa e Santos (2012); segundo, a validade da estimativa está ligada, ao mesmo tempo, a um critério determinístico de aferição efetiva do número de extremamente pobres, mas, ao mesmo tempo, com o menor grau possível de arbitrariedade tanto na correlação, quanto na variância conjunta das variáveis elencadas.

Como forma de trabalhar a hipótese e, ao mesmo tempo, respeitar os pressupostos trazidos à baila, sugeriu-se a utilização da técnica multivariada de Análise Fatorial. Sua aplicação, feita com base na faixa de renda de R\$ 0,01 a R\$ 75,00, gerou 3 perfis de referência e a congruência de cada caso da amostra a estes perfis foi aferida através da análise de confiabilidade do Alpha de Cronbach, tendo por base a linha de corte de confiabilidade, para este estudo, o valor de 0,7. Três diferentes bases de dados foram tratadas: os Censos 2000 e 2010, as PNADs 2001-20011 e, por fim, a pesquisa PADS 2012.

A primeira observação do estudo é de que, de fato, há diferenças significativas no número de extremamente pobre, para todos os anos, dependendo do critério de restrição que utilizamos. A técnica fatorial, mais parcimoniosa e, portanto, menos arbitrária, mostrou-se consistente, especialmente ao obter níveis de classificação de extremamente pobres nas categorias Sem rendimento e/ou Sem declaração (a depender da base de dados) de mais de 40%. A segunda observação passível de ser feita é que, numa análise preliminar, nota-se que a extrema pobreza tem diminuído ao longo dos últimos anos e que sua concentração tem mudado da área rural para a área urbana, padrão que se mostrou consistente para um período de 10 anos em duas bases consideradas, tanto o Censo, quanto as PNADs.

Anexo A – Exercício de imputação de renda para o ano de 2009

