

Condicionantes fundamentais da pobreza no Brasil: um estudo com análise fatorial

Pier Francesco De Maria*

RESUMO

O presente trabalho busca estudar a pobreza a partir do método de análise fatorial. No trabalho, pretende-se obter um modelo que, a partir de uma série de variáveis considerados como fundamentais na análise de pobreza, obtenha uma lista mais sucinta de características, que teria a peculiaridade de ser suficientemente significativa para explicar a pobreza por si só. Ademais, a partir da revisão do estado da arte, o trabalho enseja criar um modelo de análise fatorial embasado no estudo desta metodologia, discutindo suas possibilidades e escolhendo as melhores alternativas para o problema proposto, a determinação dos principais condicionantes da pobreza no Brasil. Os resultados apontam, de forma generalizada, para uma alteração nos padrões de pobreza nacionais entre 2002 e 2011, com um aumento do peso do grupo “trabalho” e do grupo “bens”, em detrimento dos outros grupos de fatores. Analisando a subamostra extraída, percebe-se que o padrão de pobreza se altera de forma semelhante, com mudanças mínimas no tempo e nos fatores.

Palavras-Chaves: pobreza multidimensional; análise fatorial; mensuração da pobreza

ABSTRACT

This paper studies poverty using the factor analysis method. The aim of this research is to obtain a model, from a series of variables considered fundamental in poverty analysis, with a more concise list of factors, which have the peculiarity of being significant enough to explain the poverty alone. Moreover, after reviewing the literature on poverty, the aim of this work is creating a model of factor analysis based on the study of this methodology, discussing its possibilities and choosing the best alternatives to the proposed problem, the definition of the main determinants of poverty in Brazil. The results show, in a general way, a change in patterns of national poverty between 2002 and 2011, with an increase of the weight of the groups “work” and “goods” to the detriment of other groups of factors. Analyzing the subsample extracted, it is clear that the pattern of poverty changes similarly, with minimal changes in time and factors.

Key Words: multidimensional poverty; factor analysis; poverty measurement

JEL Classification:

C38 – Classification Methods; Cluster Analysis; Factor Models

I32 – Measurement and Analysis of Poverty

* Acadêmico do 4º ano do curso de Ciências Econômicas (IE/UNICAMP), bolsista de iniciação científica pelo PIBIC/CNPq. Contato: dpierf@gmail.com.

1. INTRODUÇÃO

Estudar a pobreza é uma tarefa árdua, e requer um elevado conhecimento dos elementos que serão estudados, analisados e selecionados. Esta pode ser analisada a partir de um sem-número de metodologias, cada qual com suas peculiaridades, que são escolhidas pelo pesquisador, a partir do seu interesse de pesquisa e dos objetivos fixados. A temática “análise de pobreza” é ampla e reconhecidamente multidisciplinar, incorporando disciplinas como a demografia, a economia, a história e a estatística.

Neste ensaio, o objetivo é criar um método de análise de pobreza que consiga resumir um conjunto extenso de variáveis em um menor número de elementos, de modo a permitir que se analise o mesmo fenômeno (a pobreza) a partir de um conjunto restrito – porém significativo – de elementos. Em outras palavras, o trabalho aqui desenvolvido busca criar um método de análise fatorial, capaz de nos apresentar os principais componentes da pobreza, que resuma seus principais componentes em poucos fatores, que sejam fieis aos originais e apresentem resultados confiáveis e válidos, tanto para a amostra quanto para possíveis subamostras.

Este trabalho se encontra dividido em mais cinco seções. A próxima seção trabalhará, de forma bastante detalhada, as três grandes vertentes de mensuração da pobreza, a subjetiva, a objetiva e a relativa, mostrando várias formas de análise e comparando-as entre si. Na terceira seção, será feito um estudo amplo sobre os princípios e as formulações teóricas dos sete passos que perfazem o modelo de análise fatorial. A quarta seção apresentará os resultados da análise fatorial, dividindo-se em dois grandes momentos: resultados para a amostra total; e resultados para uma subamostra selecionada. Esta seção abordará os seis primeiros estágios da análise fatorial, sendo que os estágios de 3 a 5 serão replicados tanto na amostra quanto na subamostra. A seção subsequente tratará a análise de resultados secundários da análise fatorial, que representa o sétimo e último estágio do modelo. Por fim, a última seção tecerá alguns comentários finais, a título de resumo e avaliação do trabalho.

2. MÉTODOS DE MENSURAÇÃO E ANÁLISE DA POBREZA

De forma geral, a pobreza pode assumir três variantes principais, de acordo com a metodologia escolhida e os objetivos da pesquisa. Estas são a base para a formulação de uma estrutura multidimensional, além de um instrumento para classificar as linhas de pobreza selecionadas. A necessidade de se enquadrar os indicadores em poucas classes vem do uso, pelos formuladores de políticas públicas e pesquisadores da área, de uma vasta gama de medidas de pobreza, como pôde ser observado pelo arcabouço teórico desenvolvido na seção anterior.

Resumidamente, Hagenaars e De Vos (1988) percebem que todos os possíveis indicadores de pobreza se enquadram em uma destas três classes: pobreza objetiva ou absoluta; pobreza

subjetiva ou individual; e pobreza relativa ou característica. A pobreza, quando absoluta, representa a definição normativa do nível suficiente, sendo que se determina um pobre quando seu nível de necessidades cumpridas está abaixo do normativo. Se a variante utilizada é a relativa, a pobreza é mensurada a partir da colocação do nível de necessidade da pessoa dentro da sociedade em que vive; neste caso, considera-se pobre aquele que se situa abaixo do nível crítico, desta vez estatisticamente definido. Por fim, caso se estude a pobreza pelo viés subjetivo, o pobre é definido por si mesmo – isto é, ele define se é pobre ou não, a partir da sua consideração acerca do seu nível de necessidade ser suficiente ou não.

Antes de seguir no estudo detalhado de cada variante, cabe destacar que o uso de uma variante em detrimento de outra implica, necessariamente, ou em resultados distintos, ou em resultados semelhantes com significados diferentes. Hagenaars e De Vos (1988) argumentam, ademais, que a escolha de qual método aplicar depende de variáveis que não são de domínio do pesquisador: disponibilidade de dados; decisões de cunho político; e argumentos históricos. Soares (2009) – embora creia que o conceito de pobreza não faça sentido – percebe que a determinação de qual variante utilizar deriva, incondicionalmente, da resposta à pergunta “O quê é pobreza?”. Tendo em vista que esta resposta deverá ser dada ao longo dos próximos dois capítulos – quando será definida a metodologia de trabalho – estudaremos, de forma detalhada, as três variantes da pobreza, apresentando seus pontos de destaque e suas carências perante a literatura atual.

Todo problema referente à pobreza se inicia com a falta de concordância sobre sua forma de análise. Drewnowski (1977), entretanto, apresenta uma solução não trivial para o problema: primeiramente, deve-se entender o quê se procura medir, para depois entender como isto deve ser medido. Segundo o autor, esta é a base dos embates na discussão acerca de pobreza. O principal problema, no que concerne à análise de pobreza, é entender a seguinte definição (DREWNOWSKI, 1977, p. 183, tradução e grifo próprios): “A pobreza se manifesta quando alguém tem suas necessidades satisfeitas em um nível inferior ao que se considera *suficiente*”. Esta condição de suficiência é o que gera a discussão acerca de como deve ser avaliada a pobreza, e é a partir do entendimento acerca desta suficiência que se dá a diferenciação da análise da pobreza nas três variantes descritas.

2.1. A OPÇÃO SUBJETIVA

Como dito anteriormente, a pobreza, na sua acepção subjetiva, é a condição pela qual o próprio pobre se determina como tal. Entretanto, esta colocação é vista como humilhante e ofensiva, além de ser potencial gerador de falsas respostas, em virtude do medo de ser excluído de possíveis programas de transferências de renda do governo (VAZ e SOARES, 2008). Na opinião de Ravallion (2012), a solução se dá por um conjunto de questões que contorna a pergunta direta. Deste modo, a

pobreza subjetiva pode ser a definição de pobre pelo próprio pobre, só que de forma indireta e não agressiva.

Segundo Marks (2007), o uso da medida subjetiva não deve ser deixado de lado, já que as pessoas sabem, melhor que ninguém, da sua situação – embora Rocha e Urani (2005) afirmem que só $\frac{1}{4}$ dos 20% mais ricos se reconheçam assim e $\frac{1}{3}$ dos mesmos se consideravam da classe C –, além destas opiniões refletirem condições passadas e o contexto social na qual as pessoas se situam. Embora seja uma medida bastante precisa da pobreza, Marks (2007) reconhece que este tipo de análise é deixado de lado, em relação à pobreza relativa e à objetiva. Corroborando com a tese acima, Pradhan e Ravallion (1997) afirmam que, por ser uma área nova na ciência econômica, poucos estudos foram realizados até hoje. Todavia, nos campos sociológico e psicológico, esta medida de análise já tem bom desenvolvimento, e já foi estudada de forma bastante intensa e extensa (QUINTERO e MEJÍA, 2006).

A percepção da necessidade de estudar a pobreza sob o enfoque subjetivo aparece claramente em Narayan *et al.* (2000), ao afirmar que, com base no estado da arte, percebe-se o desenvolvimento multifacetado daquilo que podemos chamar de “fenômeno pobreza”, no qual questões sociais, históricas, demográficas e econômicas – mas não somente – fazem parte do universo de análise. O crescimento da variante subjetiva se deu nos últimos anos, especialmente ao perceber que a variante objetiva conta com limitações nas suas medições, sendo, em sua maioria, restritas à análise via linhas de pobreza.

Assim sendo, pergunta-se porque trabalhar com esta variante da pobreza e não com outras. A resposta está nas ideias do Banco Mundial, a partir da percepção de que não tenha ninguém, melhor que o pobre, para definir o que os pobres precisam. Para o Banco Mundial (2000), ouvir os pobres é fundamental para a própria formulação de políticas públicas eficazes na solução da pobreza. Entretanto, ressaltam Quintero e Mejía (2006), este tipo de análise deve, necessariamente, ser acompanhado de um estudo objetivo, por aquela ser complemento deste. Como reforça Townsend (2004), a análise subjetiva complementa a objetiva e vice-versa, sendo que, se bem-feitas, resultam em uma análise mais acurada da qualidade da pobreza de um lugar.

A fim de mostrar o nível de avanço nesta variante de pesquisa em pobreza, apresentam-se alguns dados do estudo do Banco Mundial (NARAYAN, 2000), maior pesquisa em pobreza subjetiva realizada até hoje. No estudo, obtiveram-se informações que uma pesquisa objetiva não seria capaz de revelar de forma tão clara. A partir dos dados da pesquisa, viu-se que os pobres: percebem o Estado como ineficaz no combate à pobreza, além de ser corrupto; enxergam a alimentação e o emprego como suas questões principais; entendem as relações informais do *locus vivendi* como fundamentais para a sobrevivência – se semelhando a políticas públicas contencionistas –; veem todas as boas oportunidades de emprego como exclusivas dos ricos –

fortalecendo o círculo vicioso da pobreza –; e vivem em lugares onde a falta de segurança social predomina – com violência e corrupção.

Resolvido o porquê de trabalhar com esta variante, cabe a pergunta acerca dos métodos de investigação. Como já ressaltado por Vaz e Soares (2008), não é possível obter uma resposta direta do pobre, por ser potencialmente humilhante. A partir desta percepção, os pesquisadores da área desenvolveram perguntas indiretas, às quais se atribui o poder da mensuração subjetiva. A metodologia de análise subjetiva, segundo Ravallion (2012), se dá de duas formas, atualmente: por meio de questões de categorização qualitativa; e por meio de medidas paramétricas.

A análise subjetiva via categorização qualitativa se dá a partir de duas perguntas-chave, com as quais é possível medir a qualidade de satisfação das pessoas, em relação à sua condição de vida, e a sua colocação em uma escala gradativa rico-pobre. A primeira pergunta, a ELQ e suas variantes, questionam o entrevistado acerca do seu nível de renda: *“Imagine uma escala com N degraus, onde o degrau inferior é o mais pobre e o superior é o mais rico. Em qual dos N degraus você se colocaria?”*. A segunda pergunta, a SWL e variantes, investigam o grau de satisfação do entrevistado com a vida que leva: *“Sendo 1 ‘muito insatisfeito’ e 5 ‘muito satisfeito’, quão satisfeito você está com a vida que você leva atualmente?”*.

Ao utilizar-se o conjunto de medidas paramétricas, obtém-se uma visão subjetiva a partir de valores monetários. Neste conjunto, têm-se perguntas que investigam o quanto as pessoas consideram ser um nível de renda necessário para satisfazer algumas necessidades consideradas básicas. A pergunta mais comum é a MIQ, pela qual se obtém a “linha de pobreza subjetiva”: *“Qual é a renda mensal que seria a mínima para ter um padrão de vida digno?”*. Esta questão é uma derivação da proposta de Van Praag (1968), mais abrangente, na qual se pergunta: *“Qual renda mensal você consideraria...?”*, variando o final da pergunta com as alternativas: muito ruim; ruim; não bom; não ruim; bom; e muito bom. A esta pergunta, podemos dar o nome de ILE.

A partir dos dados obtidos, é possível traçar a linha de pobreza característica, determinando quem é ou não pobre. Até aqui, foram abordadas metodologias para a obtenção de dados para a análise subjetiva, não levando em conta as metodologias para a análise dos dados. No cenário atual, existem três métodos principais para a avaliação subjetiva da pobreza: o método de Antuérpia (DELEEK, 1985); o método da linha subjetiva (KAPTEYN, KOOREMAN, WILLEMSE, 1988); e o método de Leyden (FLIK e VAN PRAAG, 1991).

O método de Antuérpia tem por objetivo a identificação dos indivíduos que se situam mais próximos da linha de pobreza e, a partir disto, usar os dados, acerca da MIQ destas pessoas, para determinar a linha de pobreza. Para saber quem está mais próximo desta linha de pobreza desconhecida, Deleek (1985) recorre a uma pergunta qualitativa, que pode ser assim formulada:

“Com a renda mensal de que dispõe, com quê dificuldade chega ao fim do mês?”, variando as respostas de “muita dificuldade” até “muita facilidade”.

O método da linha subjetiva trabalha exclusivamente com a MIQ, montando uma comparação direta entre dois casos: a renda mínima definida na MIQ é igual à renda da pessoa; a renda mínima definida na MIQ é função linear da renda da pessoa. O ponto onde os dois casos se equivalem corresponde à linha subjetiva de pobreza. Vaz e Soares (2008) chamam atenção para a praticidade deste método, por permitir a inclusão de “controles para tipos familiares diferentes e diferenças regionais” (p. 6), bastando incluir elementos na regressão definida no segundo caso, acima exposto.

Por fim, o método de Leyden é uma evolução do método LS, no qual se busca aumentar o leque de informações da MIQ, a partir da ILE. A grande diferença entre este método e o da linha subjetiva reside na necessidade da definição de um nível de bem-estar mínimo, abaixo do qual se considera pobre uma pessoa. Este nível corresponde, necessariamente, a um dos níveis da ILE. Embora seja preciso normatizar um parâmetro a mais – o nível crítico de bem-estar – esta metodologia apresenta um maior leque de dados, trazendo uma maior precisão no resultado obtido, além de um nível de robustez maior.

2.2. A OPÇÃO OBJETIVA¹

Em oposição ao exposto até agora, a variante objetiva traz consigo a bagagem da metodologia subsistencialista e das *basic needs*. A solução objetiva se dá na maior parte dos casos, especificamente por parte de órgãos internacionais, como a ONU e o Banco Mundial. O estudo desta vertente parte também das críticas feitas à subjetiva; para Coudouel, Hentsche e Wodon (2002), há problemas fundamentais no uso da análise subjetiva, pois: a amostra tende a ser pequena e não representativa; necessita de investimentos e tempo maiores; as perguntas são passíveis de interpretação; os entrevistados podem sobreavaliar suas necessidades; e as respostas podem não refletir a realidade, senão as expectativas dos entrevistados. Ademais, os autores ressaltam que a pesquisa subjetiva pode não mostrar os padrões de marginalização e/ou discriminação, caso isto seja percebido como algo normal.

Em outro ponto da curva, Soares (2009) vê a análise subjetiva como uma possibilidade que precisa de detalhes demais. Ravallion (2012) é ainda mais criterioso, e cita diversas fontes de problemas para esta vertente. Especialmente, o autor chama atenção para três problemas fundamentais: a inegável heterogeneidade das respostas subjetivas entre pessoas do mesmo grupo;

¹ Trataremos aqui, como vertente objetiva, as seguintes opções, tratadas de forma distinta por Soares (2009): medidas absolutas; medidas objetivas; medidas administrativas.

estrutura das pesquisas; e erros de mensuração. Percebe-se a série de problemas que a vertente subjetiva traz, embora todos – de uma forma ou de outra – resolvíveis².

Neste compasso, apresenta-se a vertente objetiva. Rowntree (1901) foi pioneiro nesta vertente, sendo um dos primeiros a estudar metodologias quantitativas para a determinação do nível de pobreza. Para o autor, caracteriza-se o pobre quando este não tem entradas suficientes para realizar todas as necessidades essenciais de sustentação. Embora, como afirma Velázquez (2009), outras duas noções de pobreza são importantes – a saber, se sentir pobre; e ganhar tão pouco que necessita de programas de transferências de renda –, a determinação de níveis objetivos de pobreza se faz importante na construção, de certa forma rápida e eficiente, de uma linha de pobreza que consiga determinar quem precisa de ajuda.

Uma ideia da necessidade de uma linha objetiva de pobreza está em Sen (1981), mostrando-nos que há um conjunto de capacidades humanas consideradas absolutas, pelas quais o conceito de pobreza passa³. Segundo ele, contrariando Townsend (1979) – o qual afirma que as necessidades humanas são relativizáveis no espaço e no tempo, isto é, necessariamente se alteram –, a visão relativa de pobreza é apenas complementar à variante objetiva – absoluta, nas palavras de Sen. Podemos estender esta visão à análise subjetiva, considerando que esta é parte da percepção pública da pobreza e é passível de vieses.

Altimir (1979) deixa claro que a pobreza, em sua vertente absoluta, busca estabelecer o nível de pobreza com base em alguns *benchmarks* normativos. Entretanto, relata o grande problema desta abordagem: não necessariamente retrata o nível de vida que determinado lugar tem – problemas ao qual damos o nome de “problema do distanciamento”. Sendo assim, poder-se-ia pensar que esta abordagem é totalmente descartável. Ao que se possa pensar isto, o autor chama atenção para alguns pontos fortes desta vertente: a normatização se dá com base em valores universais de bem-estar; a própria normatização, especialmente em países subdesenvolvidos, mostra o caminho pelo qual o país deve seguir para ter uma condição reduzida de pobreza; por fim, mesmo com o problema do distanciamento, pode-se normatizar com base em aspectos socioculturais, evitando a generalização dos padrões de desenvolvimento.

O empasse da variante objetiva é sua normatização, como visto acima. Para este problema, poucas soluções estão disponíveis, por ter sempre, em algum momento da pesquisa, a necessidade de uma definição acerca de qual nível deve ser estabelecido como crítico. Embora esta normatização necessária e perigosa, Orshansky (1965) e Sawyer (1975) atribuem a vantagem da variante objetiva ao foco que esta dá à pobreza em si, distinguindo-a do conceito de desigualdade de

² Sugere-se a leitura de Ravallion (2012, p. 15-28) para a solução do problema da heterogeneidade.

³ Percebemos, como tais capacidades absolutas: a alimentação; o acesso ao ensino; o acesso a condições à saúde; o vestuário decente; a moradia e suas condições de saneamento; e a qualidade do trabalho. Seguindo Altimir (1979), podemos definir este núcleo como o conjunto de direitos humanos aos quais todos têm direito.

renda: este relembra que a definição europeia do bem-estar se dá a partir de conceitos objetivos; aquele traz à tona as definições absolutas de pobreza dos EUA. Em poucas palavras, a grande vantagem da vertente objetiva é sua praticidade na aplicação.

A aplicação da metodologia objetiva é possível a partir de um vasto conjunto de informações, obtidas em pesquisas nacionais – como os censos demográficos. De posse dos dados, o pesquisador tem a possibilidade de escolher entre um sem-número de medidores de pobreza, cada qual com sua peculiaridade e utilidade. Hagenaars e De Vos (1988) citam quatro tipos de medidores, que Maria e Maia (2013) chamam de RCP. Estes indicadores – dos quais um oriundo da metodologia centenária de Rowntree (1901), o indicador de necessidades básicas a partir da cesta alimentícia – buscaram medir o nível de pobreza a partir de características de consumo⁴.

Excetuando o medidor trabalhado por Rowntree (1901), os outros três são: o IAR, que trabalha diretamente com a relação entre gastos alimentícios e renda; o ICFR, cujo objetivo é medir todos os custos normais da família em relação à renda; e o IGR, que se ocupa de estudar os gastos alimentícios e outros (como higiene e ensino). O IAR é um indicador sintético primeiramente apresentado por Orshansky (1965) e Watts (1967). Este trabalha com a ideia da Lei de Engel⁵, a partir da qual se define um nível crítico de consumo com alimentos. Se a família consumir mais que este nível crítico, considera-se a família como pobre, caso contrário não o será. O ICFR é um medidor dos gastos com custos fixos da família, e, da mesma forma que o IAR, gastando-se mais que o nível crítico determinado, a família será considerada pobre por este quesito. Por fim, o IGR trabalha com os gastos de uma família, considerando-se pobre a família que tem gastos que não sejam sustentados pelos rendimentos mensais totais.

Outro conjunto de indicadores objetivos de pobreza está diretamente relacionado à renda, de forma única e exclusiva. O exemplo clássico é a medição da linha de pobreza do Banco Mundial. Revista em 2008, hoje é definida como US\$ 1,25/dia PPC – que, para efeito didático, será considerado como R\$ 2,50/dia –, e é a medida oficial de pobreza do órgão. Nesta toada, seguem os indicadores de necessidades básicas, pelos quais se determina um valor crítico – considerado “adequado”, segundo Ravallion (2008) – de consumo em bens de primeira necessidade. A definição deste nível pode seguir critérios estatísticos, mas costuma ser feito de forma normativa e, dependendo das perguntas das pesquisas censitárias, pode ser amparada em respostas subjetivas, como as indicadas na seção anterior.

⁴ Um ponto fundamental é entender quando estes medidores são utilizados para determinação do limiar de pobreza e quando são utilizados para constituir um indicador do nível de pobreza. No primeiro caso, o medidor será trabalhado a fim de obter o valor crítico dele – pelo qual, acima ou abaixo dele, conforme a especificação, a pessoa será considerada pobre –; no segundo, o medidor, a partir do nível crítico estabelecido, nos dirá quantos são pobres naquele aspecto.

⁵ Hamilton (2001) define esta Lei como a relação inversa entre gastos com alimentação e renda auferida pela família. Já Ogaki (1992) generaliza esta Lei, superando a falácia da composição que existiria na afirmação anterior, apresentando-a como a relação inversa entre consumo alimentício e crescimento econômico.

Um último conjunto de indicadores reflete o método mais antigo de mensuração objetiva da pobreza. O método calórico é ainda o mais utilizado dentro da esfera objetiva, e se situa como um método razoavelmente complexo, por incluir dimensões biológicas – no caso, os alimentos –, econômicas e socioculturais. Seguindo a nomenclatura dada por Soares (2009), os chamamos de métodos calóricos direto e indireto, sendo o primeiro estritamente relacionado com o consumo de calorias, enquanto o segundo relaciona o consumo alimentício e outros bens, considerados necessários. Estas metodologias – especialmente o método calórico direto – permitem definir um nível de indigência⁶, pelo qual abaixo dele é impensável se levar uma vida digna.

2.3. A OPÇÃO RELATIVA

A última vertente a analisar, a relativa, é um avanço em termos tanto de medição quanto de interpretação. A opção de se trabalhar com a vertente relativa surge nas entranhas da vertente objetiva, pois, como aponta Soares (2009), há três graves problemas com as linhas de pobreza desta: não se tem concordância na definição, devido ao viés ideológico de cada pesquisador; não são medidas diretamente comparáveis a nível nacional, devido a fatores como preços relativos oscilantes; e estas linhas de pobreza objetiva falharam na OCDE do pós-guerra, o que levou à reformulação dessas linhas de pobreza.

Ao discutirem a metodologia relativa, Ravallion e Chen (2009) apontam algumas outras diferenças básicas entre esta e a objetiva, especialmente no que concerne à mudança de padrões de pobreza. Reproduzindo e ajustando o exemplo oferecido, suponham-se dois países – A desenvolvido, B subdesenvolvido –, nos quais duas pessoas detêm a mesma renda mensal, US\$ 1.000,00. Para as definições objetivas de pobreza, as duas pessoas, independentemente do país em que morem, terão o mesmo resultado – pobre ou não-pobre –, pois a renda auferida é a mesma⁷. Supondo que a linha de pobreza relativa seja equivalente à metade da renda média de cada país, A ou B serão considerados pobres se, e somente se, o valor US\$ 1.000,00 for inferior a este parâmetro.

A partir deste exemplo simples, podemos, seguindo os autores, apresentar três argumentos em sustento do uso das linhas relativas de pobreza. Primeiramente, percebe-se que a linha de pobreza é concebida como uma medida monetária da utilidade e, assim sendo, de bem-estar; a linha de pobreza atrelada ao nível médio local de renda permite avaliar o nível de utilidade alcançado pelo detentor da renda, avaliando sua condição de bem-estar em relação à media da sua sociedade. O segundo argumento utilizado remete à visão das *capabilities* de Sen (1983; 1985), pelo qual a linha de pobreza deve ser flexível ao que Ravallion e Chen (2009) chamam de “custo da inclusão

⁶ Observando Barros, Henriques e Mendonça (2001), tem-se que a linha de indigência reflete o consumo mínimo de alimentos, em termos de calorias, sem o qual não se tem condições de seguir adiante; a linha de pobreza se define como um múltiplo desta, englobando o consumo de itens primários.

⁷ Esta afirmação só é válida à medida que existe paridade de poder de compra.

social”. Retomando o conceito de *capabilities*, sabemos que estas são as condições que permitem, ao indivíduo, de ter o padrão de vida desejado.

Expandindo o conceito à segunda ordem – aquela dos bens obtidos –, obtém-se aquilo que as *capabilities* permitem ter: as *commodities*. A pobreza é relativa, pois não necessariamente, a partir de certo nível de privação de *capabilities*, se tem pobreza; a pobreza é sinônimo da falta de *capabilities*, mas é a falta de *capabilities* que gera a pobreza, e o nível crítico de cada *capability* depende do nível de desenvolvimento do país. Assim sendo, nas palavras de Sen (1983, p. 168, grifo próprio): “uma abordagem *absoluta* na esfera das *capabilities* se translada em uma abordagem *relativa* na esfera das *commodities*”. Temos, assim, a pobreza relativizada ao contexto no qual se dá a falta de *capabilities*.

Suponham-se dois indivíduos dos países A e B do exemplo anterior. Supondo o mesmo nível de acesso à saúde (30%) como única *capability* de cada indivíduo, esta gera uma pobreza muito maior no indivíduo do país A, desenvolvido, do que no de B, subdesenvolvido. Isto acontece, pois o acesso médio à saúde em B é historicamente menor que em A, o que retorna um *gap* de *capability* muito menor em B do que em A. Isto mostra como a mesma *capability*, em nível absoluto, translada-se em níveis diferentes na esfera das *commodities*.

Por fim, o terceiro argumento favorável é o possível uso, destas medidas relativas, em comparações supranacionais. Embora Ravallion e Chen (2009) só vejam esta possibilidade quando se enxerga a linha de pobreza como “*welfarist*” – oriunda do bem-estar –, reconhecem que, caso seja adotada tal hipótese, a metodologia relativa consegue a anulação das diferenças nas metodologias de cada país, permitindo a comparação de resultados entre países. Acatando a hipótese, temos a possibilidade de estudar, sob um mesmo método, vários países, com origens históricas e socioculturais distintas.

Para que se entenda o surgimento das linhas relativas, vale a pena remontar ao período entre os anos 40 e 50 do século XX. Neste período, desenvolveram-se as bases do pensamento relativista; especialmente, Duesenberry (1949) elaborou a HRR, dividida em duas hipóteses principais. Pela primeira, todo indivíduo tem seu consumo e sua poupança atrelados à relação de sua renda com as de seus símiles, mais do que a padrões de vida estabelecidos. A segunda, mais específica, relaciona o consumo atual com três fatores: a renda absoluta obtida hoje; a relação da renda atual com a renda dos símiles na sociedade; e o nível de consumo tido em períodos anteriores. Como chama a atenção Palley (2008), esta hipótese – que envolver elementos de cunho social – seria capaz de interligar Keynesianos e Institucionalistas⁸; entretanto, o Modelo de Ciclo de Vida do Modigliani e Brumberg

⁸ Para uma abordagem mais profunda da Economia Institucionalista acerca das dimensões sociais do consumo, ver Patten (1889) e Mitchell (1910).

(1954) e a Hipótese da Renda Permanente de Friedman (1957) sepultaram a visão relativa e casaram o consumo às teorias da corrente ortodoxa.

Outra teoria, desenvolvida nos anos 50 e que contribuiu no desenvolvimento da abordagem relativa, foi a TPR. Podemos atribuir sua origem a Festinger (1954), quando elaborou a Teoria da Comparação Social, e a Davis (1959), o qual formalizou a TPR nas suas hipóteses fundamentais⁹. O pressuposto inicial desta teoria é a possibilidade de particionamento da população em vários grupos dicotômicos – homem e mulher, jovem e velho, casado e solteiro, entre outros. A partir desta suposição válida, apresenta-se um conjunto de seis teoremas, divididos em três categorias: genéricos; intra-grupos; e extra-grupos.

Primeiramente, a teoria supõe a existência de uma categoria, em cada grupo dicotômico, preferível à outra – por exemplo, jovem pode ser preferido a velho – e, embora as opiniões variem de pessoa pra pessoa, tem pelo menos um grupo dicotômico no qual a preferência é inequívoca: é preferível não sofrer de privações ao sofrê-las. De forma bem simples, este teorema apresenta uma das bases fundamentais da teoria relativa, pois todas as comparações feitas terão por base a preferência de cada um em relação a determinados elementos de bem-estar. O segundo teorema permite a análise probabilística da população estudada. Segundo este teorema, as comparações são feitas seguindo uma distribuição gaussiana¹⁰, sendo fundamental para a elaboração das probabilidades de um evento acontecer. Sendo assim, a possibilidade de uma pessoa fazer N comparações será dada pelo produtório, de 1 a N, da probabilidade de cada evento acontecer. No limite, a probabilidade de todas as comparações serem feitas por uma pessoa tende a zero, o que mostra que limita-se o leque de investigações feitas por cada indivíduo.

Os dois teoremas seguintes, da categoria intra-grupos, são não só fundamentais para a análise relativa, mas também para o estudo subjetivo. O primeiro deles remete à comparação entre dois indivíduos – X e Y –, de modo que estes tenham níveis de privação diferentes; isto acontecendo, X e Y terão sensações subjetivas – de maior privação ou maior “gratificação”, nas palavras de Davis (1959) – em sentidos diametralmente opostos. O segundo teorema intra-grupos auxilia o primeiro: supondo o estudo de X, este terá um sentimento de tratamento diferencial em relação ao seu grupo. Se X sentir uma privação relativa em relação a Y, ele se sentirá mais pobre que o resto da sociedade, e vice-versa, se sentindo sempre como especial em relação aos outros do seu grupo. Ademais, todos os indivíduos do grupo se sentirão especiais, de uma forma ou de outra, acreditando que seu *status* é diferencial.

⁹ A TPR foi originada no estudo de Stouffer *et al.* (1949), embora não apareça com o formalismo do autor estudado.

¹⁰ Nas palavras de Davis (1959, p. 282, grifo próprio): “Na população, as comparações são *aleatórias*”. Entretanto, acreditamos que este teorema seja falho, pois, como apontam Coudouel, Hentsche e Wodon (2002), problemas podem não ser vistos como tais, caso sejam percebidos como normais. Isto mostra o direcionamento das comparações das pessoas, levando muito em conta o contexto socioeconômico no qual se situam.

Ao cabo da análise da TPR, os dois teoremas extra-grupos também têm cunho relativo e subjetivo, sendo análogos aos seus congêneres intra-grupos. Pelo primeiro teorema, se duas pessoas de grupos diferentes – M e N – se comparam, tendo elas níveis de privação diferentes, M terá um sentimento, em relação ao grupo de N, diametralmente oposto àquele que N tem. Isto posto, supondo faixas de renda, tem-se que uma pessoa de um grupo poderá experimentar subordinação ou superioridade relativa, em relação ao grupo ao qual se comparou. Pelo segundo teorema, que complementa a análise, tem-se que, havendo tal sentimento relativo, a pessoa de um grupo terá uma sensação de que o seu grupo tem uma privação diferenciada – tanto para mais, quanto para menos – em relação ao outro grupo, gerando o chamado “distanciamento social” (DAVIS, 1959, p. 284, tradução própria).

Percebeu-se, pela exposição acima, as qualidades que a vertente relativa traz em si. Cabe agora analisar as principais metodologias empregadas pelos defensores desta estrutura, sendo que, atualmente, existem três abordagens distintas: o estudo de bens relacionais; a visão das *capabilities*; e a fração de renda média ou mediana. Cada uma destas metodologias trabalha tendo por base as estruturas de Duesenberry (1949) e Davis (1959), e o objetivo é mostrar as relações existentes entre os indivíduos da sociedade, não dando apenas um nível crítico de pobreza ou avaliando, exclusivamente, a visão das pessoas.

O primeiro método citado, o de bens relacionais, prevê um conjunto de bens os quais dão bem-estar. Entretanto, este bem-estar – como chama atenção Soares (2009) – é relacionado ao quanto se consome dele em relação à sociedade, e não ao seu consumo absoluto. Veblen (1934) reitera a pobreza por este aspecto, lembrando-nos que o nível de consumo, destes bens relacionais, tem a ver com a visão que os componentes do grupo social do indivíduo têm a respeito dele. Assim sendo, consumir um bem relacional em quantidade menor à média da população representa um nível de pobreza relativo no que diz respeito a este bem. Tendo-se este indicador para todos os bens, tem-se o valor, em moeda corrente, que deve ser gasto, com bens relacionais, para não ser pobre nesta vertente. Seguindo Townsend (1954, 1962), temos ainda uma segunda observação a fazer: esta medida é dinâmica, mudando, ao longo do tempo, com o desenvolvimento da sociedade.

A segunda metodologia utilizada, a das *capabilities* e *commodities* de Amartya Sen, é um instrumento que, de acordo com Soares (2009), é longo-prazista. A fim de entender esta colocação, vamos analisar o método. Para Sen (1985), as capacidades humanas são constantes ao longo do tempo, isto é, não mudam variando o lugar e/ou a época, pois serão sempre necessárias para levar um padrão de vida digno e satisfatório. Entretanto, temos, como indivíduos dotados de vontade própria, a possibilidade de reduzir o nível de uso destas capacidades, por motivos específicos – podemos não fazer valer nosso exercício político ou podemos também deixar de comer o quanto comíamos antes. Deste modo, não é possível medir a pobreza, de forma direta, pelas *capabilities*,

sendo necessária a translação de um nível abstrato para um concreto, no qual seja possível a mensuração. Amartya Sen resolve o problema no plano das *commodities*, isto é, o conjunto de itens que são necessários para a sustentação do padrão de vida concebido no plano das *capabilities*¹¹. Em suma, a pobreza é relativa no plano concreto, pois o valor dos bens, necessários à realização de um padrão de vida digno, varia com o tempo, e não as necessidades humanas. Torna-se longo-prazista à medida que é necessária a avaliação de uma série de preços, para entender se é necessário o reajuste do valor mínimo ou a sua manutenção.

A última metodologia relativa, a fração da renda média – ou mediana – é empregada para uma definição rápida do nível crítico que define a pobreza. A partir da média ponderada dos rendimentos de cada família, tem-se o valor da renda média da população; de posse deste valor, pode-se determinar a fração de renda por dois métodos: normativo – o que costuma retornar valores de 50%, como na OCDE e na *Caritas Internationalis* –; ou estatístico. Variantes desta metodologia empregam a mediana da renda – como em Hagenaars (1986) – ou ainda a distribuição percentil da renda – resolvendo a linha de pobreza como um percentil inferior. Existe ainda outro método, menos utilizado e mais recente, que recorre à renda *per capita*; neste, obtida a renda *per capita* do lugar estudado, os domicílios com renda *per capita* inferior à geral são considerados pobres¹².

2.4. COMO ESCOLHER?

Ao longo das páginas anteriores, foi possível ver um conjunto de opções para mensuração do nível considerado como “linha de pobreza”. Não é uma tarefa fácil determinar qual vertente utilizar, pois todas contam com pontos fortes e estruturas válidas para a determinação da linha de pobreza. O que pode ser feito é uma análise acurada: dos dados de que se dispõe; dos empecilhos que cada metodologia pode trazer; e dos objetivos que se pretendem alcançar.

Percebe-se que o conceito de pobreza passa por um primeiro cunho importante, o de definir qual perspectiva será usada nas análises. A partir disto, surge um segundo momento na análise e definição do conceito, o viés ideológico do pesquisador e do idealizador de políticas públicas. O entendimento, por parte dos governos, de quem seja pobre e de como se lida com ele resulta em políticas sociais que afetam o nível de pobreza, para mais ou para menos. Atkinson (1987) afirma que políticas governamentais conservadoras – diante do cenário econômico da época – foram responsáveis pelo aumento dos níveis de pobreza. Todavia, há um grupo de pessoas – que são, em sua maioria, os que estão ou defendem o governo – que não acreditam nesta correlação, devido à

¹¹ Chamamos a atenção para a tradução do termo *commodity*: preferimos utilizar o termo nativo ao evitar um possível erro de tradução. Não se entenda, como fez erroneamente Soares (2009), a *commodity* como a comodidade alcançada por um bem, mas sim o bem que permite um nível de capacidade. Para tal, *commodity* poderia ser traduzido como “bem”, e não como “comodidade”.

¹² Para mais informações sobre este método, ver Beckerman (2002).

natureza pela qual se dá a mensuração da pobreza e à necessidade de se reconsiderar alguns conceitos básicos na área, a fim de aprimorar – segundo eles – os resultados obtidos.

A definição de pobreza depende de pelo menos três variáveis: perspectiva de análise; contexto e precedentes históricos; e viés político-ideológico¹³. A determinação de um conceito satisfatório para a pobreza é uma tarefa árdua e que, indubitavelmente, trará resultados válidos, porém diferentes, a partir de cada vertente. Entretanto, duas considerações “genéricas” podem ser feitas, a fim de auxiliar na tarefa de entender a pobreza.

Primeiramente, a partir da visão de Ugá (2004), o conceito de pobreza e seu estudo tornaram-se centrais, devido ao crescimento exponencial das pesquisas em torno da condição de privações de boa parte da população mundial, em especial no que tange a formulação e a avaliação de políticas públicas, além de ser objeto de estudo por parte de vários órgãos internacionais. Em segundo lugar, mas não menos importante, podemos definir pobreza de forma genérica, a fim de, mesmo usando o traçado teórico acima exposto, ter-se uma visão clara do que seja este fenômeno. Kageyama e Hoffmann (2006) introduzem o conceito de pobreza como sendo a privação em geral de uma pessoa, podendo ser de cunho material ou também sociocultural, podendo ter caráter subjetivo, objetivo – também chamado de absoluto – ou relativo.

3. ANÁLISE FATORIAL: PRINCÍPIOS E FORMULAÇÕES¹⁴

Para que possamos avançar na seleção de um método de mensuração, vamos primeiro selecionar os principais componentes (determinantes) da pobreza. Esta tarefa será feita por meio do procedimento de análise fatorial. Para que possamos proceder às análises, vamos estudar este método. A análise fatorial (HAIR *et al.*, 2006) é um método multivariado que tem por objetivo apresentar uma seleção sucinta de dados, seja por redução, seja por resumo de dados e identificação de estruturas subjacentes. A ideia da análise fatorial pode ser visualizada na figura a seguir.

¹³ Tem uma quarta variável a ser estudada, mas que remete, de forma mais específica, aos indicadores de pobreza – os quais têm o papel de definir a quantidade de pobres, e não a linha de pobreza em si. Hagenaars (1987) analisa a fundo esta questão, chamada de “exigências axiomáticas”.

¹⁴ Esta seção segue, fielmente, a estrutura e descrição apresentadas por Hair *et al.* (2006). Daqui em diante, salvo casos em contrário, esta é a referência para a seção, não sendo feitas novas referências a este manual.

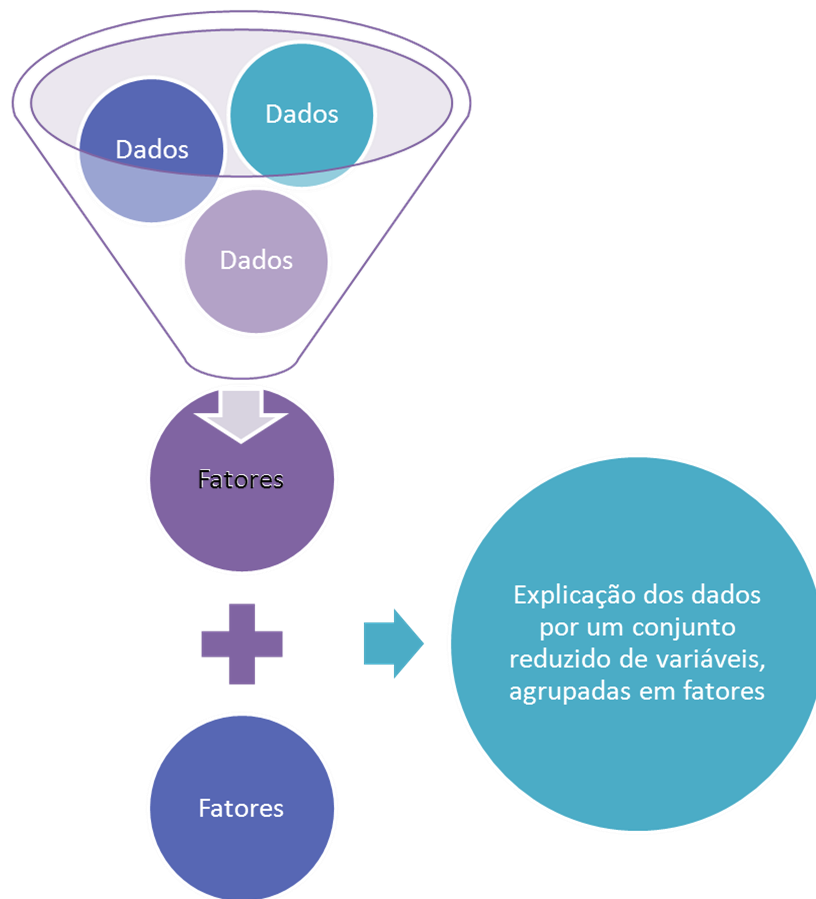


Figura 1 – O objetivo da análise fatorial

Fonte: elaboração própria.

O princípio da análise fatorial é, portanto, nos fornecer um conjunto reduzido de fatores, cada qual explicando um número de variáveis, de modo a entendermos um determinado fenômeno sem recorrer a um sem-número de variáveis. Os usos desta metodologia são variados, passando pelo uso para identificação de padrões semelhantes, pela determinação de fatores explicativos (em menor número frente às variáveis iniciais), e pela possibilidade de usar os resultados como preliminares em outras técnicas econométricas multivariadas. Neste trabalho, vamos associar a determinação de fatores explicativos ao uso dos resultados finais como ingredientes de outros procedimentos.

Como a abordagem matemática da análise fatorial é razoavelmente complexa, vamos nos limitar a apresentar os passos pelos quais deveremos passar, mostrados na figura a seguir, para depois aplicar os procedimentos apropriados em pacote estatístico próprio (podendo escolher entre SAS[®] e SPSS[®]). Resumidamente, temos sete estágios neste método multivariado: problematização; planejamento; elaboração; determinação; interpretação; validação; e reuso.

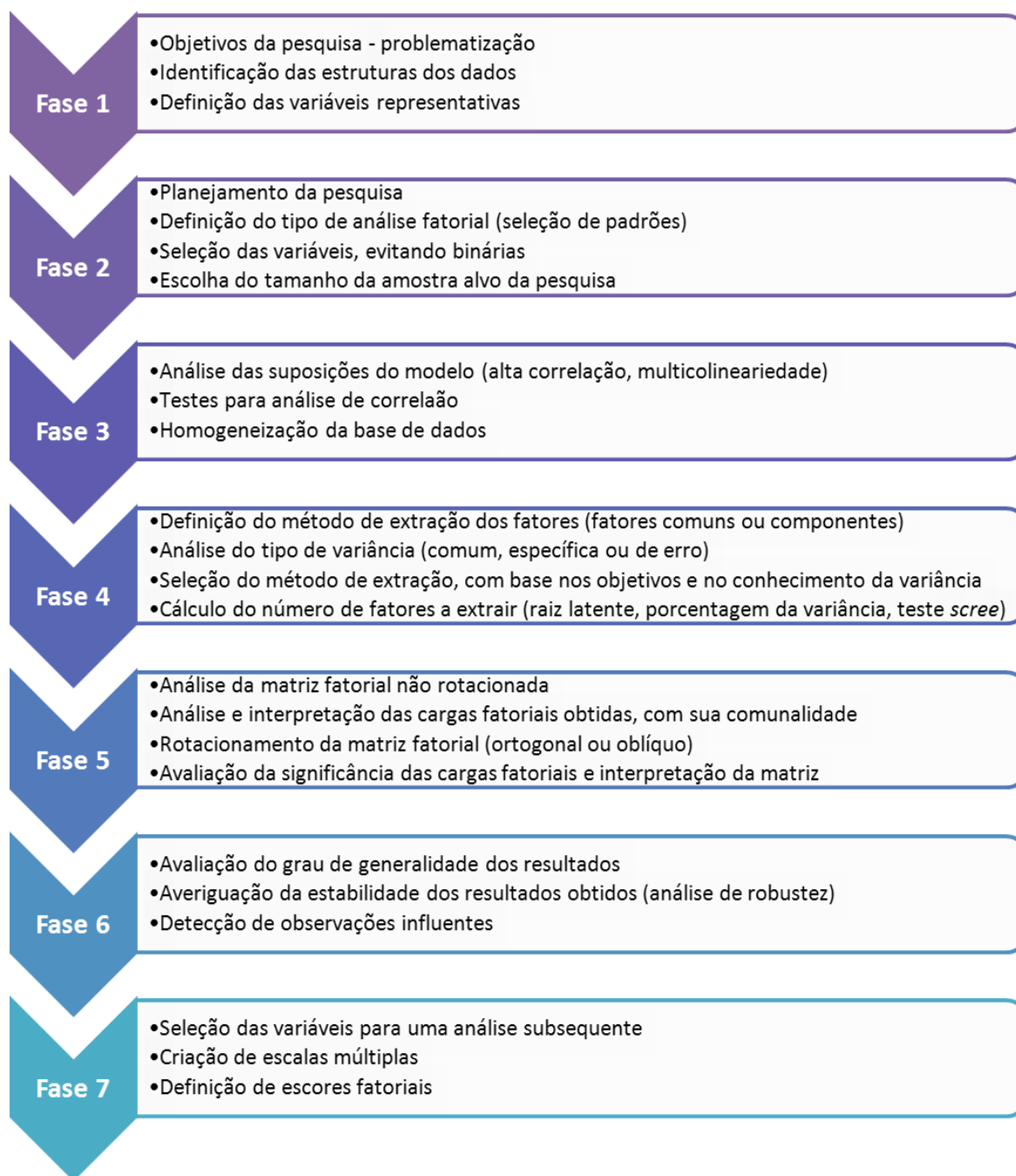


Figura 2 – O *step-by-step* do modelo de análise fatorial

Fonte: elaboração própria a partir de Hair *et al.* (2006).

O primeiro estágio é o da problematização. Nesta fase, devemos identificar a estrutura dos dados e os objetivos da pesquisa, de modo a selecionar as primeiras informações para o procedimento. Neste trabalho, o objetivo é obter um número reduzido de fatores, de modo a ter a explicação de um fenômeno (a pobreza) a partir destes. Deste modo, nossa análise é exploratória, com objetivo de resumir os dados e identificar as estruturas existentes nos dados.

O segundo estágio é o do planejamento. Por ele, devemos definir como trabalharemos a análise fatorial, isto é, o que está sendo agrupado, para depois fazer o planejamento da nossa pesquisa. O objetivo é agrupar variáveis em fatores relevantes, deste modo, estamos interessados no

uso de uma análise fatorial tipo-R, que objetiva o agrupamento dos dados de modo a identificar variáveis em grupo, e não analisando padrões repetidos ou semelhantes na nossa amostra.

O terceiro estágio da análise é o de elaboração das hipóteses estatísticas. Temos algumas necessidades, específicas deste método multivariado, que devem ser incluídas no modelo. Primeiramente, é importante que exista um mínimo de multicolinearidade, pois a análise fatorial objetiva encontrar padrões nos dados. Ademais, é fundamental ter um alto nível de correlação entre as variáveis, de modo que estas possam ser agrupadas a contento. Para fazer a análise das correlações, vamos recorrer a dois testes estatísticos, sendo o primeiro o Teste Bartlett de Esfericidade (TBE), e o segundo a Medida de Adequação da Amostra de Kaiser-Meyer-Olkin (KMO). O TBE, segundo Santana (2007), tem objetivo de testar se as variáveis são correlacionadas. Para tal, temos a seguinte especificação da estatística de teste:

$$\begin{cases} H_0: R_n = I_n \\ H_A: |R_n| \neq 1 \end{cases} \rightarrow \chi^2 = - \left[(n-1) - \frac{2p+5}{6} \right] \times \ln|R|, \quad gl = \frac{p(p-1)}{2} \quad (10)$$

A hipótese nula indica que o determinante da matriz R , de ordem n , é igual ao determinante de uma matriz identidade, de mesma ordem. A hipótese alternativa indica a correlação entre os elementos da matriz, de modo que o determinante seja diferente de 1. A estatística é uma χ^2 com n sendo o número de observações, p o número de variáveis e $|R|$ o determinante da matriz de correlação. Os graus de liberdade gl são determinados segundo a especificação acima. Para este teste, rejeitamos a hipótese nula caso a probabilidade de erro seja inferior ao nível crítico (que estabelecemos em 5%).

A medida KMO, seguindo a especificação de Santana (2007), tem objetivo de averiguar se o modelo de análise fatorial é o apropriado. Para tal, temos a seguinte especificação do teste:

$$KMO = \left[\sum_i \sum_j r_{ij}^2 \right] \left[\left(\sum_i \sum_j r_{ij}^2 \right) + \left(\sum_i \sum_j a_{ij}^2 \right) \right]^{-1}, \quad 0 \leq KMO \leq 1 \quad (11)$$

Nesta estatística, r_{ij} é o coeficiente de correlação entre x_i e x_j , enquanto a_{ij} é a correlação parcial entre as mesmas variáveis. A ideia é que se tenha um valor $a_{ij} \cong 0$, o que indica que há um fator comum entre tais variáveis e que a análise fatorial é um método apropriado. Para analisar o resultado obtido, temos a seguinte convenção, apresentada na tabela abaixo:

Tabela 1 – Avaliação do resultado obtido na estatística KMO

Valor KMO	Interpretação
> 0,9	Magnífico (é o desejável, difícil de obter)
> 0,8	Notável (é uma ótima amostra)
> 0,7	Razoável (é uma boa amostra)
> 0,6	Mediano (geralmente é o que se obtém)

$> 0,5$	Pobre (há outros métodos multivariados que podem se encaixar melhor)
$< 0,5$	Descartável (não se deve usar a análise fatorial)

Fonte: elaboração própria.

O quarto estágio diz respeito à determinação dos fatores e à avaliação do ajuste. Nesta fase do processo, já deve estar clara a proposta da análise fatorial, e deve se conhecer o objeto de estudo. Temos algumas importantes decisões a tomar, as quais são descritas a seguir: definição do método a ser usado para obter fatores; análise do tipo de variância; seleção dos fatores a extrair. Primeiramente, com base nos objetivos da análise fatorial, nos nossos conhecimentos sobre a variância dos dados e nos dados que temos, temos que determinar o método de obtenção dos fatores. Este método pode ser escolhido entre a análise de fatores comuns (AFC) e a análise de componentes principais (ACP).

O método AFC é utilizado quando se buscam dimensões latentes nas variáveis iniciais, e quando se tem pouco conhecimento acerca da variância das variáveis. O método ACP é mais usado quando a preocupação é obter o menor número de fatores que são necessários para explicar a maior variância possível, tendo também conhecimento que as variâncias específica e do erro são insignificantes perante a variância total.

Para entender a estrutura da variância, vamos analisar a figura a seguir, no qual mostramos que a variância total assume três formas: comum, que é a compartilhada, por uma variável, com as outras; específica, que é a que uma variável só tem; e a do erro, devida a erros de agrupamento de dados e de medição ou, ainda, oriunda de um processo *white noise* (erro aleatório produzido pro ruído branco). Como o valor da variância comum é desconhecido, trabalhamos com uma *proxy*, chamada de comunalidade.

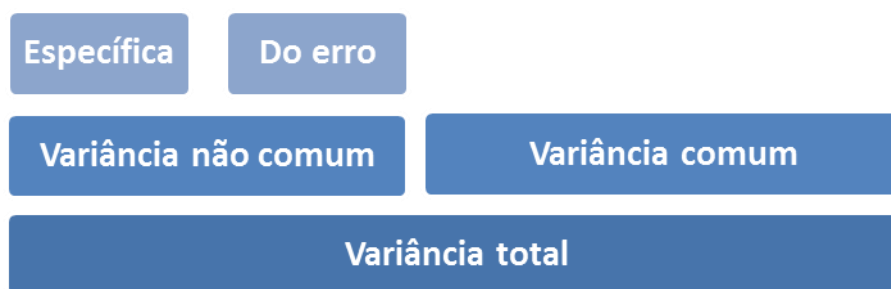


Figura 3 – Tipos de variância dos dados

Fonte: elaboração própria a partir de Hair *et al.* (2006).

Antes de passar para o quinto estágio, devemos ainda definir como extrair os fatores. Cada fator, após aplicação do método, tem um *eigenvalue* correspondente, que é o valor da variância total. Para selecionar o número de fatores adequado, tem vários critérios (sugere-se o uso de mais de

um procedimento, de modo a ter como fazer uma estimativa mais precisa), sendo que os principais são: o critério da raiz latente; o critério *a priori*; o *scree test*; e o critério da variância explicada. O critério da raiz latente é utilizado para ter um corte sumário – e sem intervenção do pesquisador – nos componentes a serem usados. O princípio é simples: todo e qualquer fator deve ser capaz de explicar a variância de pelo menos uma variável (MINGOTI, 2005). Deste modo, somente *eigenvalues* que entrem no critério $\lambda_f > 1$ (onde $1 \leq f \leq p$ é o número do fator) serão utilizados.

De forma oposta a este critério, o *a priori* não conta com uma regra, mas sim com uma normatização do pesquisador. Neste critério, há uma definição do número de fatores a extrair antes mesmo de se aplicar o processo de análise fatorial. Deste modo, sem mesmo saber os resultados, tem-se um critério neutro para a seleção de fatores. Na mesma linha, o *scree test* é um critério que pode ser considerado apriorístico, pois se define um método heurístico para seleção dos fatores. Descrito de forma algébrica, o critério nos diz que, se temos uma derivada próxima de zero, este é o ponto de corte:

$$\frac{\partial f}{\partial \lambda} \rightarrow \begin{cases} < 0 \Rightarrow \lambda \text{ significativo} \\ \cong 0^- \Rightarrow \lambda \text{ descartado} \end{cases} \quad (12)$$

O último critério é o da variância total, também sumário, como o da raiz latente. O objetivo é selecionar o número de fatores que, juntos, somem um mínimo de variância, de modo a reduzir as variáveis selecionadas sem perder o poder de explicatividade do modelo. O quanto de variância é explicado por cada variável é determinado pelo valor do *eigenvalue*, como segue (KLEFENS, 2009), de modo que a soma de pelo menos 60% da variância total σ_T^2 dará o número de fatores:

$$(\sigma_T^2)_i = \lambda_i \left[\sum_{i=1}^p \lambda_i \right]^{-1} \Rightarrow \sum_{i=1}^{f \leq p} (\sigma_T^2)_i \geq \frac{3\sigma_T^2}{5} \quad (13)$$

O quinto estágio tem a ver com a interpretação dos fatores extraídos. Primeiramente, temos que analisar a matriz fatorial não-rotacionada. Nela, cada variável terá um valor associado a cada fator, a carga fatorial, que é a correlação entre a variável inicial e seu respectivo fator. Se elevarmos ao quadrado, temos a variância, em uma variável, explicada por um fator específico, chamando isto de autovalor (λ). Se, por outro lado, elevarmos ao quadrado cada carga fatorial de uma variável, temos a comunalidade (ς), que nos dá quanto, da variância, uma variável inicial compartilha com as outras. Deste modo, temos:

$$\lambda_j = \sum_{i=1}^k [f(x_i)]^2 \quad (14a)$$

$$\varsigma_i = \sum_{j=1}^l [f_j(x_i)]^2 \quad (14b)$$

Do que podemos tirar: que o *eigenvalue* do j -ésimo fator (λ_j) é a soma dos quadrados das cargas fatoriais, naquele fator, de cada uma das i variáveis; e que o valor da comunalidade é dado pela soma dos quadrados da carga fatorial (da variável i) em cada um dos j fatores. Somando-se os *eigenvalues*, podemos ainda ter uma outra medida de análise, que é o percentual do traço τ , sendo que definimos o traço como o total de variáveis:

$$\sum_{i=1}^k \varsigma_i = \sum_{j=1}^l \lambda_j = x \Rightarrow \tau = \frac{x}{p} \leq 1 \quad (14c)$$

Derradeiramente, temos uma medida de confiabilidade (que mede a consistência, de um conjunto de variáveis, acerca do que se quer medir), o Alfa de Cronbach (1947; 1951). A ideia do coeficiente é apresentar uma estatística que nos dê a confiabilidade e a aceitabilidade do que se mediu. Esta estatística é definida da seguinte forma:

$$\alpha = \frac{f}{f-1} \times \left[\frac{1}{\sigma_T^2} \sum_{j=1}^f (\sigma_T^2)_i \right] \quad (15)$$

Da qual se deduz que este coeficiente é uma relação (ponderada pelo número de fatores f) entre a soma da variância de cada fator e a variância total. Adotamos, para este teste, nível mínimo de aceitação $\alpha = 0,65$.

Após a análise inicial da matriz, é hora de rotacioná-la. O objetivo da rotação é melhorar a visualização dos resultados alcançados, evidenciando relações que na matriz não-rotacionada eram latentes. Atualmente, tem-se dois modos de rotacionar uma matriz: métodos ortogonais; e métodos oblíquos. O grande interesse da rotação é fazer com que a carga fatorial de uma variável se concentre em um único fator, tentando zerar os outros, para conseguir visualizar quais variáveis pertencem a quais fatores. A diferença entre métodos oblíquos e ortogonais é que estes mantêm os eixos perpendiculares – garantindo a independência linear – enquanto aqueles variam o ângulo entre eixos. É muito mais fácil aplicar rotações ortogonais do que oblíquas, por pelo menos dois motivos: a rotação oblíqua necessita de um cuidado ainda maior na interpretação, pois se insere um fator de correlação entre cargas; e há poucas escolhas entre os métodos oblíquos.

Como o objetivo aqui é ter um número reduzido de fatores, que representem as variáveis iniciais, o método ortogonal se encontra mais adequado para os fins estabelecidos. Dentre os métodos ortogonais, o mais utilizado é o VARIMAX, que busca simplificar as colunas da matriz fatorial (fatores), de modo que a soma das variâncias seja maximizada. Depois de termos rotacionado a matriz, é importante avaliar a significância das cargas fatoriais. Adotaremos, dentre os vários métodos, o critério da garantia de significância prática, pelo qual temos três níveis de significância, mostrados na tabela a seguir.

Tabela 2 – Valor da significância prática para cargas rotacionadas por VARIMAX

Carga fatorial	Variância explicada	Interpretação
$ f_j(x_i) < 0,3$	$\zeta_{i,j} < 10\%$	Significância prática irrelevante (desprezível)
$ f_j(x_i) > 0,3$	$\zeta_{i,j} > 10\%$	Nível mínimo de significância
$ f_j(x_i) > 0,4$	$\zeta_{i,j} > 15\%$	Nível médio de significância
$ f_j(x_i) > 0,5$	$\zeta_{i,j} > 25\%$	Significância prática
$ f_j(x_i) > 0,7$	$\zeta_{i,j} > 50\%$	Nível elevado de significância
$ f_j(x_i) > 0,8$	$\zeta_{i,j} > 65\%$	Nível muito elevado de significância

Fonte: elaboração própria a partir de Hair *et al.* (2006).

O sexto estágio da análise é o da validação, que é o momento mais importante do processo, aquele no qual vamos analisar se os resultados são válidos para nossos intentos. O objetivo da validação é averiguar a estabilidade dos resultados, pois estes devem refletir a população, seja ela como for. Um método de validação é o particionamento aleatório, pelo qual se repete todo o processo, até agora desenvolvido, para as subamostras, a fim de verificar se os resultados são compatíveis com o da amostra inteira. Pode-se ainda analisar as observações influentes, a fim de avaliar a robustez do modelo; entretanto, os métodos de detecção de influência são mais difíceis de aplicar na análise fatorial. Deste modo, recorreremos aos métodos econométricos clássicos, eliminando os elementos com dados de variáveis influentes.

Finalmente, o sétimo estágio da análise fatorial é o reuso dos seus resultados. Temos, neste estágio, a possibilidade de selecionar formas de aproveitar os resultados obtidos, dentre as quais destacam-se: a seleção da variável substituta; a criação de escalas múltiplas; e a criação de *scores* fatoriais. A primeira técnica é bem simples e vai direto ao ponto: de cada fator, selecionamos a variável com maior carga fatorial rotacionada, de modo a termos uma variável substituta que represente determinado fator. A segunda técnica é a criação de escalas que sejam a média das cargas fatoriais de cada variável. Entretanto, quando tem-se variáveis com carga negativa, aplica-se um processo de inversão de *score*, pelo qual se faz uma inversão em relação à carga original, para evitar problemas de eliminação de variáveis com cargas de sinais invertidos. Vejamos:

$$\mathcal{E}_{n,j} = \frac{1}{k} \sum_{i=1}^k [f_j(x_i)]', \quad \begin{cases} f_j(x_i) \geq 0 \Rightarrow [f_j(x_i)]' = f_j(x_i) \\ f_j(x_i) < 0 \Rightarrow [f_j(x_i)]' = \eta - |f_j(x_i)| \end{cases} \quad (16)$$

Isto é, a escala múltipla n , composta por k cargas fatoriais do fator j é a média da soma das cargas fatoriais f_j de cada variável x_i . Se, entretanto, tal carga for menor que zero, temos que fazer um ajuste, a inversão de *score*, donde tiramos que a carga ajustada será o teto (η) da escala adotada menos a carga inicial. Por último, a terceira técnica que podemos adotar é a criação de *scores*

fatoriais, que, devido ao grau de complexidade algébrico, é calculado pelos pacotes estatísticos de forma direta, não apresentando aqui uma formulação específica.

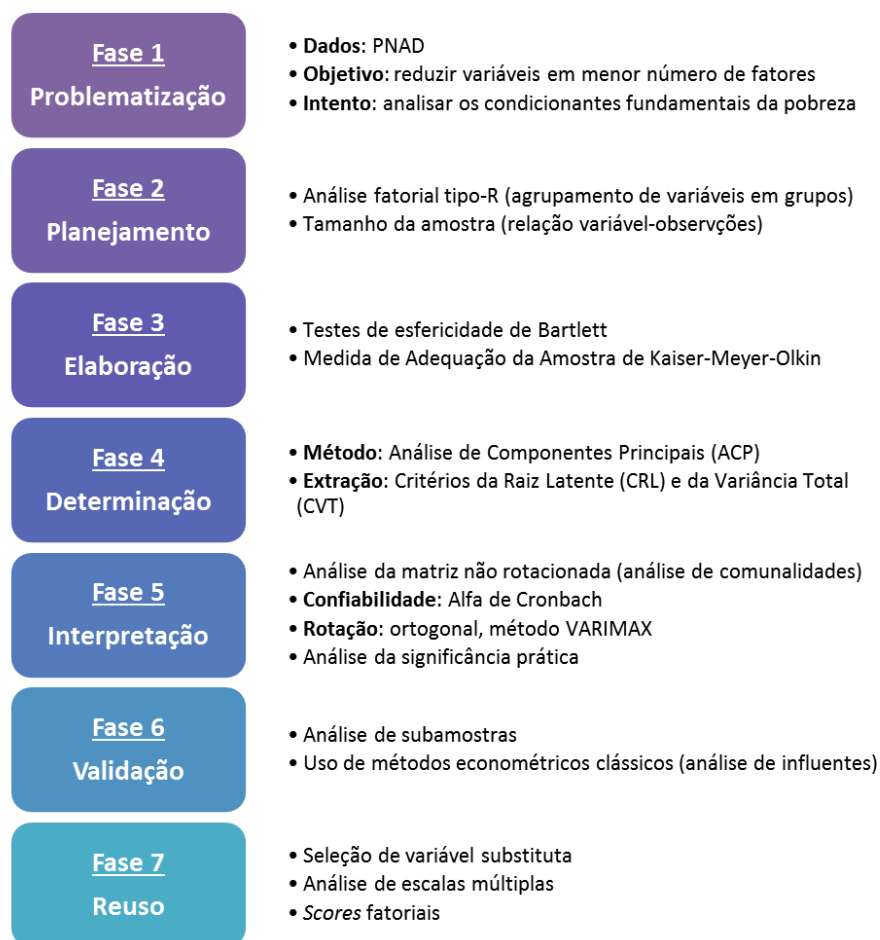


Figura 4 – Elementos, de análise fatorial, aplicados à pesquisa

Fonte: elaboração própria.

Apresentamos, acima, o processo de análise fatorial em sua forma genérica. Antes de seguir selecionando as variáveis, definimos, no quadro acima, o que será feito, neste ensaio, em cada estágio de análise, escolhendo entre as várias possibilidades apresentadas. O próximo passo é a definição das variáveis a serem usadas, dentro do conjunto existente na PNAD, sendo que os anos estudados serão de 2002 a 2011.

Como coloca Maria (2012), praticamente toda variável pode expressar alguma dimensão de pobreza. Isto pode levar a um problema de generalização excessiva, no qual ser baixo implica ser pobre, o que sabemos que é esdrúxulo. Deste modo, devemos ser criteriosos na seleção de variáveis, pois existem variáveis que não agregam um nível significativo de informações. Isto nos leva a outras perguntas, que se associam ao objetivo da análise fatorial: “Quantas variáveis escolher?”; “Quais variáveis são apropriadas para o estudo?”; “Com base em quê (ou em quem) selecionar as variáveis?”. Ademais, temos um último problema, que remete ao tipo de variável. É

desaconselhável, em análise fatorial, ter variáveis dicotômicas, sendo importante restringi-las ao menor número possível.

Feito este panorama geral, faremos a seleção das variáveis, com base nos dados coletados pela PNAD e na literatura pertinente¹⁵. É fundamental agregar um número considerável de variáveis, de modo que a análise fatorial faça sentido, mas lembrando que é importante ter-se uma relação mínima casos/variáveis. No Apêndice A, temos a lista de variáveis utilizadas, com o nome utilizado neste trabalho e o código correspondente utilizado na PNAD. Foram selecionadas 31 variáveis, as quais abrangem um leque extenso de temas, os quais buscamos resumir em um conjunto reduzido de fatores, por meio da análise fatorial descrita nesta seção.

O último momento que precede os cálculos é o estudo do procedimento fatorial no pacote estatístico selecionado. Usaremos o SAS[®], no qual o procedimento é denominado PROC FACTOR, para a análise fatorial em si. O manual para o usuário do SAS (2013, grifo próprio) define a *procedure* para análise fatorial como segue:

*“The FACTOR procedure performs a variety of **common factor** and **component analyses** and **rotations**. Input can be **multivariate data**, a correlation matrix, a covariance matrix, a factor pattern, or a matrix of scoring coefficients. The procedure can factor either the correlation or covariance matrix, and you can save most results in an output data set. (...) The methods for factor extraction are **principal component analysis**, **principal factor analysis**, **iterated principal factor analysis**, **unweighted least squares factor analysis**, **maximum likelihood (canonical) factor analysis**, **alpha factor analysis**, **image component analysis**, and **Harris component analysis**. A variety of methods for prior communality estimation is also available. (...) Specific methods for **orthogonal rotation** are **varimax**, **quartimax**, **biquartimax**, **equamax**, **parsimax**, and **factor parsimax**. **Oblique versions** of these methods are also available. In addition, **quartimin**, **biquartimin**, and **covarimin** methods for (direct) oblique rotation are available. (...) Output includes **means**, **standard deviations**, **correlations**, **Kaiser’s measure of sampling adequacy**, **eigenvalues**, a **scree plot**, **eigenvectors**, **prior and final communality estimates**, **the unrotated factor pattern**, **residual and partial correlations**, **the rotated primary factor pattern**, (...) and **scoring coefficients**.”*

Como se percebe, a *procedure* no SAS[®] é bastante sofisticada, permitindo análises profundas e com um leque amplo de opções.

Neste momento, já estamos prontos para excetuar os procedimentos estatísticos necessários para obter os resultados da análise fatorial. Na sessão a seguir, apresentaremos os resultados a partir de diversas perspectivas, almejando cobrir a maior parte dos resultados emitidos pelo pacote estatístico que utilizamos.

¹⁵ Para uma leitura acerca das principais variáveis utilizadas em análise de pobreza, ver Barros (2009), Barros, Henriques e Mendonça (2000; 2001), Rocha (1992; 1996; 2006) e Ribas (2005).

4. RESULTADOS ALCANÇADOS¹⁶

O procedimento de análise fatorial nos permitiu obter resultados para os anos selecionados, identificando padrões para a pobreza no Brasil. Nesta seção, vamos estudar os resultados para a amostra completa e para uma subamostra semi-aleatória e, finalmente, vamos discutir a possibilidade de reuso destes resultados em outros procedimentos econométricos.

4.1. RESULTADOS PARA A AMOSTRA COMPLETA

Para começar o estudo, vamos analisar as medidas de adequação da amostra. Conforme fica visível na tabela abaixo, praticamente 9 em cada 10 variáveis tem uma medida de adequação da amostra muito boa, tendo apenas uma ou outra variável abaixo de 0,6. Quando analisamos a medida de adequação para toda a amostra, temos um resultado muito satisfatório (sempre acima de 0,80), embora tenda a decrescer ao longo dos anos estudados¹⁷.

Tabela 3 – Análise da medida de adequação da amostra KMO para os anos analisados (amostra completa)

(%)	2002	2003	2004	2005	2006	2007	2008	2009	2011
Acima de 0,8	67,7	67,7	64,5	67,7	67,7	67,7	67,7	67,7	67,7
De 0,6 a 0,8	25,8	25,8	32,3	29,0	25,8	25,8	25,8	22,6	25,8
Abaixo de 0,6	6,5	6,5	3,2	3,2	6,5	6,5	6,5	9,7	6,5
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
KMO geral	0,845	0,842	0,846	0,848	0,841	0,838	0,832	0,830	0,832

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Como o SAS não nos fornece a estatística TBE, vamos nos restringir à medida de adequação de amostra. Como o resultado foi satisfatório, temos uma indicação de que o processo de análise fatorial é adequado para este estudo.

Após verificar a qualidade dos dados, aplicamos os procedimentos de análise fatorial. Para a análise dos dados, selecionamos o método ACP, usando o critério da variância total para extração dos fatores, com limiar de 60%. Isto resultou na extração, em cada um dos anos, de 9 fatores, com variância explicada entre 60,7% e 62,1%, além na possibilidade de calcular a medida de confiabilidade α (Alfa de Cronbach).

¹⁶ Como a análise fatorial se baseia em amostras, tomamos a decisão de não levar em conta os fatores de expansão, de modo a representarmos os resultados para a amostra em si (que é o objetivo do procedimento estatístico), e não o resultado expandido para a população. Um possível teste de validação é ponderar as observações pelo seu fator de expansão e analisar os resultados obtidos, tarefa que deixamos, nesta ocasião, a cargo do leitor.

¹⁷ No Apêndice B, trazemos as medidas KMO para todas as variáveis em todos os anos.

Tabela 4 – Análise de confiabilidade dos dados extraídos pré-rotação (amostra completa)

Variância	2002	2003	2004	2005	2006	2007	2008	2009	2011
Extraída	19,111	19,129	18,890	19,249	19,177	19,043	18,966	18,874	18,818
Total	31,000	31,000	31,000	31,000	31,000	31,000	31,000	31,000	31,000
Proporção	61,7%	61,7%	60,9%	62,1%	61,9%	61,4%	61,2%	60,9%	60,7%
Nº fatores	9	9	9	9	9	9	9	9	9
α	0,694	0,694	0,686	0,699	0,696	0,691	0,688	0,685	0,683

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Conforme a tabela abaixo, temos uma medida de confiabilidade razoável; em outros termos, o uso da análise fatorial vai gerar resultados que aderem, de forma satisfatória, à realidade estudada.

Tabela 5 – Análise estatística da comunalidade das variáveis (amostra completa)

Comunalidade		2002	2003	2004	2005	2006	2007	2008	2009	2011
Quartis	1º Quartil	0,51	0,49	0,52	0,50	0,52	0,53	0,53	0,53	0,53
	Mediana	0,60	0,60	0,59	0,60	0,59	0,61	0,62	0,63	0,61
	3º Quartil	0,72	0,73	0,67	0,74	0,74	0,73	0,73	0,73	0,72
Interquartis	Mínima	0,39	0,36	0,40	0,39	0,36	0,14	0,03	0,04	0,14
	Máxima	0,94	0,99	0,97	0,91	0,91	0,88	0,88	0,87	0,87
	Distância	0,55	0,63	0,58	0,52	0,55	0,74	0,85	0,83	0,73
	Média	0,66	0,67	0,68	0,65	0,64	0,51	0,45	0,46	0,51
Momentos	μ_1	0,616	0,617	0,609	0,621	0,619	0,614	0,612	0,609	0,607
	μ_2	0,021	0,024	0,017	0,021	0,021	0,025	0,029	0,028	0,024
	μ_3	0,435	0,412	0,827	0,311	0,285	-0,653	-1,213	-1,115	-0,652
	μ_4	-0,590	-0,316	0,587	-0,848	-0,715	1,228	3,365	2,960	1,333

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Já sabemos, neste ponto, que a amostra é boa e que os resultados a obter se adequam à realidade; é hora, portanto, de analisar a estrutura das comunalidades. As variáveis apresentam bons níveis de comunalidade, pois as variáveis na faixa interquartil têm comunalidades muito boas¹⁸. Analisando os terceiro e quarto momentos da distribuição, temos algumas informações mais

¹⁸ Para consultar o valor de comunalidade para cada variável em cada ano, ver o Apêndice C.

acuradas acerca das comunalidades. Quanto à *skewness* (μ_3), podemos ver que, até 2006, a maior parte das comunalidades se concentrava abaixo da média entre o primeiro e o último percentis, tendência que se reverteu em 2007 (com *skewness* negativa).

Isto nos diz que, para a série até 2006, a maior parte das comunalidades se concentrou abaixo de valores excessivamente altos, garantindo que não há problema de unicidade. Esta colocação é reforçada pela *kurtosis*, (μ_4), a qual é negativa para o mesmo período, apontando uma série platicúrtica. Em compensação, após 2006, a série tem *kurtosis* positiva, o que indica uma série leptocúrtica, pela qual a maioria das comunalidades se concentra ao redor da média (que, porém, como aponta a *skewness*, é muito maior que a média entre os limites percentis da série).

Uma medida de associação de resultados é o agrupamento de valores em classes de interesse. Para que a análise fatorial tenha sucesso, é importante que a maioria das comunalidades esteja entre 0,25 e 0,65 (conforme colocamos na tabela 2).

Tabela 6 – Análise de comunalidades por agrupamentos (amostra completa)

Comunalidade (%)	2002	2003	2004	2005	2006	2007	2008	2009	2011
Abaixo de 0,10	0,0	0,0	0,0	0,0	0,0	0,0	3,2	3,2	0,0
De 0,25 a 0,65	64,5	61,3	74,2	64,5	64,5	61,3	58,1	61,3	64,5
Acima de 0,80	16,1	16,1	9,7	12,9	12,9	16,1	12,9	12,9	9,7

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Como fica claro pelos dados, entre $\frac{1}{2}$ e $\frac{3}{4}$ das variáveis têm comunalidade no intervalo desejado, e confirma a qualidade das amostras utilizadas, nos motivando a seguir adiante com o rotacionamento da matriz fatorial inicial.

Valendo-nos do modo de rotação VARIMAX, procedemos ao rotacionamento da matriz fatorial inicial, de modo a identificar as variáveis que cada fator continha. Como o critério da variância total (CVT) nos informou que 9 fatores seriam retidos, as variáveis foram acomodadas, com base nas suas cargas fatoriais, em cada um dos fatores. O método utilizado tem por objetivo maximizar a carga em um dos fatores, tentando zerar a carga nos outros, de modo a delinear, mais claramente, que variável está em que fator.

Grosso modo, percebem-se três padrões diferenciados na análise, seguindo o que foi descrito na análise de comunalidades. Um primeiro padrão – mais homogêneo – segue de 2002 a 2005, um segundo – transitório, isto é, mutante ao longo dos anos – existe entre 2006 e 2009, e um terceiro, que é encontrado para o ano de 2011. O que muda é quais variáveis compõem as cargas fatoriais, embora quais variáveis estejam nas cargas fatoriais não mude substancialmente. Vamos analisar o quadro a seguir: como, para cada padrão, não sempre as variáveis presentes nos fatores são iguais,

os nomes atribuídos podem diferir entre os vários padrões. Vale ressaltar que os fatores estão ordenados de forma decrescente em relação à variância explicada.

Quadro 1 – Fatores extraídos por rotação VARIMAX, segundo o padrão observado (amostra completa)

Padrão I (2002 a 2005)	Padrão II-A (2006)	Padrão II-B (2007 e 2009)	Padrão II-C (2008)	Padrão III (2011)
Bens elementares	Bens elementares	Qualidade do trabalho	Qualidade do trabalho	Estrutura do trabalho
Qualidade do trabalho	Qualidade do trabalho	Bens indispensáveis	Bens da família	Bens indispensáveis
Bens da família	Bens da família	Bens da família	Bens indispensáveis	Bens da família
Forma de trabalho	Forma de trabalho	Forma de trabalho	Forma de trabalho	Bens de consumo
Estrutura familiar	Estrutura familiar	Bens de consumo	Bens de consumo	Estrutura familiar
Localização geográfica	Estrutura da casa	Estrutura familiar	Estrutura familiar	Localização geográfica
Estrutura da casa	Localização geográfica	Localização geográfica	Localização geográfica	Condições de vida
Condições de vida	Condições de vida	Estrutura da casa	Estrutura da casa	Estrutura da casa
Trabalho infantil	Trabalho infantil	Condições de vida	Condições de vida	Condições de vida infantil

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Pelo apresentado acima, percebe-se que as grandes categorias de análise podem ser ainda mais resumidas, em 5 grandes grupos: bens; trabalho; família; local; e vida. Partindo do quadro apresentado a seguir, vamos discernir a evolução dos grupos ao longo do tempo¹⁹.

No Grupo I (bens), percebe-se que há uma diferenciação importante nos “bens elementares”, no Padrão I, separando-se em bens indispensáveis (proveniência da água, água encanada, coleta de lixo e coleta de esgoto) e em bens de consumo (televisão, tipo de iluminação e geladeira), a partir do Padrão II-A. Os “bens de família”, fator assim chamado por incorporar itens gerais da família, no que diz respeito ao acesso à informação, à casa e à renda (computador e internet, número de cômodos, rendimento mensal per capita, máquina de lavar roupas, telefone e anos de estudo completos), se mantêm estáveis ao longo dos anos.

¹⁹ No Apêndice D, temos a divisão dos grupos por padrão, com os movimentos que se sucederam em cada período analisado.

No Grupo II (trabalho), a tendência é invertida. Ao longo dos anos, está se observando um movimento de aglutinação das condições de trabalho. Elementos de “qualidade do trabalho” (carteira assinada, contribuição previdenciária, recebimento de auxílios e tipo de contratação) e de “forma de trabalho” (tipo de trabalho, estar empregado, horas trabalhadas e alfabetização) passam a responder por um único fator, a “estrutura do trabalho”, da qual, todavia, sai a componente de alfabetização, a qual pode ser incluída como fator do Grupo V (vida).

O Grupo III (família) sempre foi composto pela “estrutura familiar”, fator que inclui o tipo de família, o número de pessoas e o número de banheiros disponíveis para uma determinada família²⁰. Na mesma direção, o Grupo IV (local) se mantém com os mesmos fatores, “localização geográfica” (unidade federativa e grupo étnico) e “estrutura da casa” (incluindo a qualidade do telhado e das paredes).

Por fim, no Grupo V (vida), temos uma reconfiguração dos componentes. Se, até 2006, havia dois fatores principais (“trabalho infantil” e “condições de vida”), entre 2007 e 2009 o trabalho infantil não tem significância prática mas, em 2011, volta a compor o grupo, substituindo a variável de migração, para compor o fator “condições de vida infantil”, que inclui os natimortos e o próprio trabalho infantil, enquanto o fator “condições de vida”.

Tendo em vista os resultados obtidos, decidimos ordenar os resultados para definir que grupos têm maior importância na definição da pobreza. Para tal, definimos um critério de pontuação baseado no seguinte raciocínio: quanto o fator f_j participa da variância explicada (σ_f^2) pelos f fatores extraídos?

$$\sigma_f^2 = \sum_{j=1}^{f \leq p} (\sigma_T^2)_j \geq \frac{3\sigma_T^2}{5} \Rightarrow \mathcal{P}(f_j) = 100 \times \frac{(\sigma_T^2)_j}{\sigma_f^2} \quad (17a)$$

Para cada grande grupo G_w criado, a participação relativa é dada pela soma dos fatores que participam, e o valor completo (do total dos h anos) é dado pela média das participações. Deste modo, somos capazes de comparar o peso que cada grupo teve para determinar a pobreza neste período analisado.

$$\mathcal{P}(G_w) = \frac{1}{h} \times \sum_{x=1}^h \left\{ \left[\frac{100}{(\sigma_f^2)_h} \right] \times \sum_{j_w=1}^n [(\sigma_T^2)_{j_w}]_h \right\} \quad (17b)$$

Os resultados obtidos estão na tabela abaixo, pela qual podemos ordenar os grandes grupos por importância na explicação da pobreza nos últimos anos. Como podemos perceber, os três grupos principais são os relacionados aos bens (Grupo I) e ao trabalho (Grupo II), seguidos pela

²⁰ De certa forma, faz sentido esta relação. O tipo da família tem relação estreita com seu tamanho, de modo que o número de banheiros tende a ser uma boa *proxy* do tamanho e da estrutura do arranjo familiar.

localização geográfica (Grupo IV). Com menor destaque, temos as condições de vida (Grupo V) e a estrutura familiar (Grupo III).

Analisando ano a ano, percebe-se que o Grupo I perdeu importância, enquanto o Grupo II se alavancou, em detrimento do Grupo V, que perdeu sensivelmente importância até a virada da década. Mais recentemente, com os deslocamentos entre os grupos, há um aumento na participação relativa do Grupo V e III, mantendo-se o Grupo I e o IV. O aumento do Grupo V e a queda do II serão reflexo da realidade se houver um movimento contínuo ao longo dos próximos anos.

Tabela 7 – Participação relativa (%) de cada grupo para composição da variância extraída (amostra completa)

Grupo	2002	2003	2004	2005	2006	2007	2008	2009	2011	Média
I	42,3	42,5	43,2	42,7	42,6	39,2	39,1	38,9	39,4	41,1
II	28,7	28,7	27,7	28,8	28,8	38,6	38,5	38,4	30,8	32,1
III	6,6	6,6	6,7	6,6	6,6	5,8	5,9	5,9	7,1	6,4
IV	11,8	11,8	11,7	11,5	11,5	11,1	11,3	11,4	11,5	11,5
V	10,5	10,5	10,6	10,4	10,4	5,3	5,3	5,4	11,2	8,8
Variância extraída (%)	61,7	61,7	60,9	62,1	61,9	61,4	61,2	60,9	60,7	61,7

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

4.2. RESULTADOS PARA A SUBAMOSTRA

Após esta análise detida das variáveis e dos fatores, vamos proceder à validação dos resultados. Para testarmos a validade do procedimento, decidimos aplicar um procedimento de extração de subamostra parcialmente aleatório, isto é, definimos, arbitrariamente, um critério para extração, mas não sabemos quem será extraído. Para seleção da subamostra, decidimos levar em conta os 5 maiores estados em termos de amostra (a maior amostra para cada macrorregião), como consta na tabela a seguir, após a exclusão dos não entrevistados, dos que não declararam renda e dos que não têm cômodos casa (considerando isto uma *proxy* dos que não têm moradia apropriada para viver).

Os estados selecionados, por este critério, foram, em todos os anos: Pará para o Norte; Bahia para o Nordeste; São Paulo para o Sudeste; Rio Grande do Sul para o Sul; e Goiás para o Centro-Oeste. Estes estados somam, no conjunto, entre 35 e 40% da amostra estudada nas PNAD de 2002 a 2011, sendo um bom conjunto subamostral para testarmos se as análises feitas valem para um outro conjunto semi-aleatoriamente selecionado. O objetivo, agora, é testar o mesmo modelo de análise

fatorial, reduzido a essa subamostra, ao longo dos anos, para verificar se os resultados para a amostra são compatíveis com uma realidade restrita como a de subamostras.

Tabela 8 – Tamanho da amostra total e da amostra do maior Estado de cada região (em milhares de pessoas)

	2002	2003	2004	2005	2006	2007	2008	2009	2011
Amostra	374,9	364,1	388,2	398,8	400,3	388,2	378,3	386,1	339,8
Subamostra	143,1	142,4	146,1	151,5	150,6	145,1	141,0	141,6	125,4
Razão (%)	38,2	39,1	37,6	38,0	37,6	37,4	37,3	36,7	36,9

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Os resultados obtidos para a subamostra podem ser divididos em alguns movimentos de análise: qualidade da subamostra e da comunalidade; variância extraída e confiabilidade; padrões fatoriais da subamostra *versus* da amostra. Começando com a análise da qualidade da subamostra e da comunalidade, temos que, assim como para a amostra completa, sempre mais de 60% das variáveis, por ano, tem teste KMO superior a 0,80, indicando que a nossa subamostra é tão boa quanto a amostra completa.

Tabela 9 – Análise da medida de adequação da amostra KMO para os anos analisados (subamostra)

(%)	2002	2003	2004	2005	2006	2007	2008	2009	2011
Acima de 0,8	64,5	64,5	58,1	67,7	67,7	67,7	67,7	67,7	67,7
De 0,6 a 0,8	29,0	29,0	29,0	22,6	22,6	22,6	22,6	22,6	22,6
Abaixo de 0,6	6,5	6,5	12,9	9,7	9,7	9,7	9,7	9,7	9,7
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
KMO amostra	0,845	0,842	0,846	0,848	0,841	0,838	0,832	0,830	0,832
KMO subamostra	0,837	0,838	0,830	0,844	0,841	0,841	0,833	0,830	0,835

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Embora este resultado favorável, temos que tomar cuidado com algumas variáveis, especificamente o número de pessoas na família e o tipo de família, pois o teste KMO deu valores inferiores a 0,60 e a 0,50, respectivamente, de modo que a análise fatorial, com tais variáveis, pode gerar resultados diferentes dos obtidos na amostra completa²¹. Comparativamente, percebemos que a medida KMO para a amostra é maior que para a subamostra até 2005, invertendo-se a tendência a

²¹ No apêndice E, temos os resultados para o teste KMO para a subamostra 1.

partir de 2006. Isto nos aponta que nossa subamostra é ligeiramente “mais realista” que a amostra em si, pois, embora a subamostra represente apenas o maior estado de cada região, temos que é um conjunto de observações mais adequado para o estudo. Em suma, podemos deduzir que esta subamostra influencia significativamente na qualidade da amostra.

Analisando a comunalidade das variáveis da nossa subamostra, temos que a maioria das variáveis, em todos os anos, tem comunalidade entre 0,50 e 0,75, indicando-nos que temos boas variáveis para esta subamostra.

Tabela 10 – Análise estatística da comunalidade das variáveis (subamostra)

Comunalidade		2002	2003	2004	2005	2006	2007	2008	2009	2011
Quartis	1º Quartil	0,50	0,52	0,54	0,53	0,53	0,51	0,51	0,51	0,54
	Mediana	0,59	0,60	0,57	0,60	0,60	0,59	0,63	0,60	0,60
	3º Quartil	0,73	0,74	0,72	0,76	0,77	0,75	0,73	0,76	0,73
Interquartis	Mínima	0,41	0,42	0,43	0,28	0,22	0,23	0,22	0,05	0,11
	Máxima	0,95	0,88	0,90	0,88	0,88	0,88	0,87	0,91	0,90
	Distância	0,54	0,47	0,47	0,61	0,66	0,65	0,65	0,86	0,79
	Média	0,68	0,65	0,66	0,58	0,55	0,56	0,55	0,48	0,51
Momentos	μ_1	0,618	0,621	0,615	0,624	0,622	0,615	0,615	0,610	0,610
	μ_2	0,022	0,018	0,015	0,020	0,022	0,023	0,023	0,031	0,027
	μ_3	0,453	0,307	0,426	-0,118	-0,301	-0,166	-0,341	-0,761	-0,680
	μ_4	-0,794	-1,057	-0,740	-0,359	0,106	-0,078	-0,043	1,860	1,540

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Estudando os momentos da distribuição de comunalidades, temos que a tendência da amostra completa se verifica parcialmente quando estudamos a subamostra, pois a *skewness* tem sinal invertido de 2004 para 2005, enquanto a amostra teve uma inversão em 2007. Em outras palavras, a maior parte das comunalidades, até 2004, se encontra abaixo da média entre os percentis extremos. Analisando a *kurtosis*, temos que a distribuição de comunalidades, quando não se encontra abaixo da média entre percentis extremos, tende a se concentrar em torno da média e da mediana, garantindo-nos, novamente, que não haverá risco de unicidade.

Discutindo agora o agrupamento de valores em classes de interesse, percebemos que, novamente, a maioria das variáveis tem comunalidade no intervalo desejado (entre 0,25 e 0,65), o

que só confirma a boa estrutura das variáveis selecionadas para o resultado desejado²². Deste modo, podemos novamente proceder ao rotacionamento da matriz fatorial, usando o mesmo procedimento (VARIMAX), com extração de 9 fatores pelo CVT, em todos os anos do estudo.

Tabela 11 – Análise de comunalidades por agrupamentos (subamostra)

Comunalidade (%)	2002	2003	2004	2005	2006	2007	2008	2009	2011
Abaixo de 0,10	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3,2	0,0
De 0,25 a 0,65	64,5	61,3	61,3	58,1	54,8	58,1	61,3	61,3	61,3
Acima de 0,80	12,9	16,1	3,2	12,9	16,1	16,1	9,7	16,1	12,9

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Os resultados obtidos por rotacionamento VARIMAX com extração de 9 fatores, para a subamostra, são semelhantes, mas não coincidem formalmente com os obtidos na amostra²³. Se, por um lado, os resultados para o ano de 2002 coincidem entre a amostra e a subamostra, mudanças na ordem dos fatores acontece em 2003, com redistribuições das variáveis, entre fatores, de 2004 a 2006. De 2006 a 2007, as diferenças são mínimas, enquanto que, entre 2008 e 2009, se dá uma nova reconfiguração dos fatores; por fim, em 2011, se tem uma nova reorganização dos fatores, para um novo padrão. Deste modo, podemos definir três grandes períodos, como foi feito para a amostra total, só que com estruturas diferenciadas: o primeiro vai de 2002 a 2007; o segundo, entre 2008 e 2009; e o terceiro, que é o atual (2011).

Feita esta divisão por padrões, vamos analisar os cinco grupos que compõem os fatores extraídos. Começando pelo Grupo I (bens), percebe-se que, até 2007, estes são divididos, como na amostra completa, em dois grandes subgrupos (“bens elementares” e “bens da família”). Este primeiro subgrupo, a partir de 2008, se divide em “bens indispensáveis” e “bens de consumo”, assim como aconteceu na amostra completa. No Grupo II (trabalho), temos os já citados elementos de “qualidade do trabalho” e de “forma de trabalho”, que se unem, em 2004 e 2011, em “estrutura do trabalho”, mostrando pouca diferença em relação à amostra completa.

Quanto ao Grupo III (família), este não sofreu maiores mudanças ao longo dos anos, mantendo o fator “estrutura familiar” como único componente da série. Por fim, os Grupos IV (local) e V (vida) são condições secundárias, mas constantes, na definição da pobreza. O primeiro grupo inclui, sistematicamente, o fator “localização geográfica”, o qual nos diz onde se vive e de quê gênero é o entrevistado, e o fator “estrutura da casa”. Por sua vez, o segundo grupo dá

²² O apêndice F reporta as comunalidades (por ano) das variáveis da subamostra 1.

²³ No apêndice G, pode-se observar a evolução da subamostra 1, ano-a-ano.

informações sobre as “condições de vida”, e “trabalho infantil”, sendo este último analisado ou a sós, ou junto à mortalidade infantil, compondo, neste caso, o fator “condições de vida infantil”.

Analisando os grupos por presença, vemos que o Grupo I é o que mais responde pela variância extraída pelo método ACP. Entretanto, mais recentemente (exceto em 2011), tal grupo divide a maior importância junto ao Grupo II. Os outros três grupos, em conjunto, respondem para cerca de ¼ da variância extraída (com máximo de 37,3% em 2004 e mínimo de 22,6% em 2008).

Tabela 12 – Participação relativa (%) de cada grupo para composição da variância extraída (subamostra)

Grupo	2002	2003	2004	2005	2006	2007	2008	2009	2011	Média
I	42,1	42,5	43,0	42,4	42,6	42,1	38,7	38,6	39,4	41,3
II	29,0	28,8	19,7	29,1	28,9	29,1	38,6	38,6	30,9	30,3
III	6,5	6,2	7,9	6,4	6,4	6,5	6,2	6,2	6,9	6,6
IV	11,9	11,7	18,4	11,6	11,6	11,7	11,2	5,8	11,7	11,7
V	10,5	10,8	10,9	10,4	10,5	10,6	5,2	10,9	11,0	10,1
Variância extraída (%)	61,8	62,1	61,5	62,4	62,2	61,5	61,5	61,0	61,0	61,8

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

Para encerrar o tópico, vamos avaliar se os resultados obtidos são confiáveis, por meio da análise da variância extraída (via Alfa de Cronbach). Pelo que a tabela abaixo nos reporta, podemos ver que a variância extraída é uma medida confiável para a representação das variáveis escolhidas, pois o Alfa de Cronbach se situa sempre acima do limiar estabelecido (0,65). Em outras palavras, podemos considerar nosso estudo confiável, por representar validamente, com um número reduzido de componentes, a situação existente em cada ano.

Tabela 13 – Análise de confiabilidade dos dados extraídos (subamostra)

Variância	2002	2003	2004	2005	2006	2007	2008	2009	2011
Extraída	19,152	19,257	19,079	19,335	19,275	19,075	19,066	18,923	18,918
Proporção	61,8%	62,1%	61,5%	62,4%	62,2%	61,5%	61,5%	61,0%	61,0%
Total	31,000	31,000	31,000	31,000	31,000	31,000	31,000	31,000	31,000
Nº fatores	9	9	9	9	9	9	9	9	9
α	0,695	0,699	0,692	0,702	0,699	0,692	0,692	0,687	0,687

Fonte: elaboração própria a partir das PNAD 2002 a 2011.

5. USOS SECUNDÁRIOS DOS RESULTADOS

Embora, de fato, os resultados não sejam completamente semelhantes, temos claro o resultado positivo da análise fatorial, pois a subamostra extraída tem semelhanças estruturais com os resultados para a amostra completa. As diferenças existentes podem ser explicadas de várias formas, quais sejam: os Estados selecionados estão entre os mais ricos de cada região; estes Estados não incluem toda a variedade existente na amostra; os Estados têm padrões de pobreza diferentes; e estes Estados participam de forma heterogênea da composição nacional da pobreza. Em outras palavras, a pobreza é estruturada nos seguintes três grandes fatores da pobreza, a saber: determinantes; causadores; condicionantes.

O determinante da pobreza é um fator endógeno, o qual traz a pobreza e é intrínseco à pessoa, e que nos diz a respeito da condição de uma pessoa. Por exemplo, o quanto uma pessoa ganha por mês nos diz (de forma rápida) se é pobre ou não, sendo um determinante da pobreza. Por sua vez, o causador da pobreza é um fator exógeno, que também traz pobreza, mas que não é pessoal, não é intrínseco à manifestação idiossincrática do indivíduo. É, nos termos de Sen (1985; 2001), uma privação de liberdade, uma *capability* faltante. Exemplos patentes são a falta de recursos médicos apropriados na região onde a pessoa mora, a má qualidade das escolas, a falta de segurança no bairro e a representatividade política. Por fim, o condicionante da pobreza é um fator também exógeno, sendo que, diferentemente dos outros dois, este é um fator que não necessariamente gera pobreza, mas que, certamente, é endogeneizando pelo indivíduo, por ser algo inescapável. Exemplo disto é a inflação, a qual corrói o poder real de compra do salário, podendo manifestar, indiretamente, a pobreza em um indivíduo.

Dito isto, antes de analisar o quadro abaixo, não de ser colocados pelo menos outros dois fatores da pobreza, quais sejam: os diferenciadores; e os geradores. Os fatores diferenciadores da pobreza são aqueles que se manifestam de forma generalizada, isto é, são os que não distinguem o indivíduo quando “atacam”. Todavia, são fatores que agredem diferentemente os indivíduos, dependendo de outros fatores (como os determinantes ou os causadores). Por exemplo, se há falta de hospitais em uma região (causador de pobreza), o indivíduo ser atacado por uma enfermidade tem efeitos diferenciados caso ele more ou não na região sem hospital. Quanto aos fatores geradores, são elementos que redundam em pobreza de forma “instantânea”, independente de todos os outros fatores. Um claro exemplo disso é o desemprego, pois estar desempregado gera pobreza (aqui cabe a hipótese de que não se vive de rendas, quando se procura trabalho).

Vejam, derradeiramente, o quadro abaixo, no qual estão relacionados os cinco grandes fatores da pobreza no mundo. Podemos instituir uma hierarquia entre os fatores da pobreza, assim discriminada: fatores estruturais, os quais são persistentes e de mais longo prazo; e fatores conjunturais, os quais se concentram no curto prazo e são (mais ou menos) facilmente sanáveis,

dependendo das condições conjunturais existentes. No primeiro grupo, se encaixam os fatores determinantes e os causadores, enquanto que, no segundo grupo, tem-se os condicionantes, os geradores e os diferenciadores.

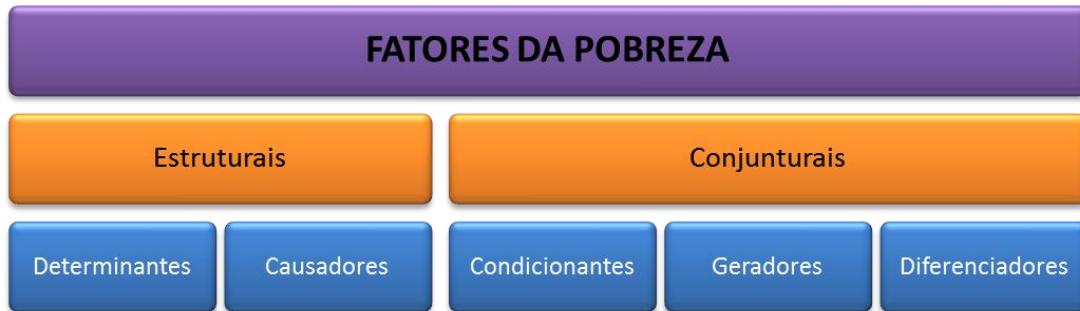


Figura 5 – Os cinco fatores da pobreza

Fonte: elaboração própria.

Todo este discurso visa preparar o terreno para a criação, a partir dos resultados da análise fatorial, de dois indicadores sintéticos de pobreza. O primeiro indicador analisa o peso que cada grupo tem na determinação da pobreza e, a partir disto, nos retorna o quanto uma unidade de análise é pobre, permitindo a comparação entre as diferentes unidades da amostra.

$$\rho_n = \frac{1}{\kappa_n} \left\{ \sum_{n=1}^m \kappa_n \left[\sum_{w=1}^5 (\rho_w \times \tilde{x}_w)_n \right] \right\} \quad (18a)$$

Sendo ρ_w o valor (normalizado entre 0 e 1) do nível de pobreza pelo grupo w de fatores, \tilde{x}_w sua ponderação normalizada (como apresentada na tabela 7), n a unidade de análise (podendo ser a pessoa ou o domicílio ou um estado inteiro) e κ_n a ponderação amostral desta unidade.

$$\rho_w = \sum_{j_w=1}^x \tilde{j}_w \left(\frac{1}{k} \times \sum_{i_j=1}^k \frac{i_j}{i_{j_{max}}} \right)_{j_w} \quad (18b)$$

Para calcular o nível normalizado ρ_w , devemos calcular a média normalizada (isto é, entre 0 e 1) das k variáveis i_j que compõem o fator j_w , multiplicando-a pela ponderação normalizada \tilde{j}_w (isto é, pelo peso que o fator j tem no grupo w)²⁴.

O segundo indicador analisa os diferentes grupos de fatores, e nos diz o quanto, em cada um dos grupos \mathcal{G}_w , há de pobreza em uma unidade de análise n . Para tal, definimos o indicador como uma escala fatorial, remetendo ao que foi definido na equação (16):

$$\varepsilon(\mathcal{G}_w) = \frac{1}{x} \left\{ \sum_{j_w=1}^x \left[\frac{10}{k} \left(\sum_{i_j=1}^k [f(i_j)]' \right) \right]_{j_w} \right\}, \quad \begin{cases} f(i_j) \geq 0 \Rightarrow [f(i_j)]' = f(i_j) \\ f(i_j) < 0 \Rightarrow [f(i_j)]' = 1 - |f(i_j)| \end{cases} \quad (19)$$

²⁴ No Apêndice H, temos as ponderações normalizadas dos fatores para cada grupo e ano.

Isto é, a escala fatorial de pobreza para o grupo G_w é dada pela média dos resultados obtidos para cada fator j_w , valor este calculado como a média das cargas fatoriais ajustadas $[f(i_j)]'$, multiplicadas por 10 (para obter uma escala entre 0 e 10). Para obtermos uma índice fatorial de pobreza (\mathfrak{F}_n), para uma determinada unidade de análise n , podemos proceder como segue:

$$\mathfrak{F}_n = \sum_{w=1}^5 \mathcal{E}(G_w) \left\{ \frac{1}{x} \left[\sum_{j_w=1}^x \left(\frac{1}{10k} \times \sum_{i_j=1}^k (10 - i_j) \right) \right]_{j_w} \right\} \quad (20a)$$

Em outros termos, afirmamos que o índice fatorial de pobreza é dado pela soma dos produtos entre a escala fatorial G_w e a média, ponderada pelo número de variáveis e pelo número de fatores por grupo, dos valores obtidos para cada variável i , do fator j , no grupo w . O valor do índice fatorial de pobreza varia entre 0 e a soma de todos os $\mathcal{E}(G_w)$; para termos um valor ajustado, vamos ter o seguinte arranjo:

$$(\mathfrak{F}_n)^* = \mathfrak{F}_n \left[\sum_{w=1}^5 \mathcal{E}(G_w) \right]^{-1} \quad (20b)$$

Para obter um indicador macrorregional, devemos recorrer às ponderações amostrais κ_n , de modo que o índice fatorial de pobreza para uma região, cidade ou um país é dado (em ano específico) por:

$$(\mathfrak{F}_t)^* = \frac{1}{m} \times \sum_{n=1}^m \kappa_n (\mathfrak{F}_n)^* \Big|_t \quad (20c)$$

sendo m o número de unidades de análise em um determinado espaço geográfico, e t o ano de referência para a análise.

O índice fatorial \mathfrak{F} e o índice ponderado de pobreza ρ se relacionam com os fatores da pobreza apresentados nesta seção, pois cada um dos fatores da pobreza, se alterado, vai ter efeitos em algumas componentes de cálculo de cada índice; vejamos como. Analisando as equações (18a) e (18b), vemos que, se qualquer um dos fatores da pobreza se alterar, poderão mudar: o valor da ponderação \tilde{x}_w de cada grupo (pois mudar um grupo afeta todos); o valor de cada variável i_j ; e o valor da ponderação \tilde{f}_w de cada fator extraído dentro do grupo w . Analisando as equações (19) e (20), temos que, mudando um dos fatores da pobreza, poderão mudar: a carga fatorial $f(i_j)$ (e, conseqüentemente, a escala fatorial $\mathcal{E}(G_w)$); e o valor de cada variável i_j . No limite, podemos ver que, com a mudança nos fatores da pobreza, os próprios fatores podem mudar, alterando completamente a configuração dos resultados.

Ficam registrados estes índices de pobreza, os quais podem ser calculados como resultados secundários da análise fatorial, sendo úteis para indicar o nível de pobreza e sua evolução ao longo dos anos. Temos, nestes indicadores, medidas objetivas de pobreza, pois estas são fruto de dados

obtidos de forma imediata (isto é, sem passar pela avaliação subjetiva do entrevistado). Embora existam críticas à vertente objetiva, é clara a qualidade e o poder de análise dos índices ξ e ρ , pois são frutos de uma análise fatorial que gerou resultados confiáveis²⁵, frutos de um modelo bem especificado e que levou em conta vários períodos de tempo (apresentando boa consistência temporal) e vários tipos de elementos (tendo assim boa consistência estrutural).

6. COMENTÁRIOS CONCLUSIVOS

Apresentou-se, neste trabalho, uma metodologia de análise dos elementos fundamentais da pobreza no Brasil, por meio da análise fatorial, via análise de componentes principais, com extração de fatores pelo critério da variância total e rotação pelo método ortogonal VARIMAX. O trabalho considerou um espectro temporal de 10 anos (de 2002 a 2011), utilizando 31 variáveis de interesse e a amostra completa da PNAD que, considerando-se apenas os entrevistados *de facto*, leva em conta entre 340 e 400 mil pessoas.

Os resultados obtidos foram plenamente satisfatórios, levando em conta que o estudo é feito usando observações que são influenciáveis. Como estamos lidando com dados sobre pessoas, uma pesquisa de ciências sociais aplicadas, é normal que os coeficientes estatísticos sejam um pouco menores do que o desejado (vejam-se os valores obtidos para o coeficiente Alfa de Cronbach), pois estamos trabalhando com variáveis que têm poder de análise e decisão. Diferentemente de ciências exatas e da terra, não podemos garantir a total fidedignidade dos dados coletados, pois os entrevistados são sensíveis a fatores tanto endógenos (a situação em que vive) quanto exógenos (as condições oferecidas pela sociedade e pelo Estado).

O momento crítico da análise é a análise de subamostras. Para que a validez dos resultados seja confirmada, é importante que se analise uma ou mais subamostras. Neste trabalho, optou-se por extrair uma amostra pelo método semi-aleatório, pelo qual escolhemos que Estados extrair, sem saber quais observações seriam extraídas. Esta escolha foi adotada por um motivo, em especial: era relevante fazer uma validação com subamostras significativas, tanto em termos econômicos, quanto em termos estatísticos. Tanto é que nossa subamostra representava entre 30% e 40% da amostra total, em todos os anos.

Pelos resultados obtidos na subamostra, que levou em conta o Estado, de cada macrorregião, com maior peso na amostra completa, podemos notar que há padrões de pobreza diferentes em cada parte do país, e a escolha da subamostra afeta o resultado desta e sua comparação com os resultados da amostra completa. Entretanto, de forma generalizada, os padrões encontrados na amostra são refletidos, com ligeiras variações temporais e de fatores, na subamostra, confirmando a validade da análise fatorial explicitada neste trabalho.

²⁵ Vejam-se os resultados para o Alfa de Cronbach, o teste KMO, a comunalidade das variáveis e a variância extraída.

Finalmente, trabalhou-se a criação de indicadores secundários, isto é, oriundos dos resultados obtidos na análise fatorial. Estes indicadores foram apenas explicitados, mostrando-se como estes são suscetíveis a mudanças nos cinco fatores da pobreza (determinantes, causadores, geradores, diferenciadores e condicionantes), e como podem ser utilizados como *proxy* do nível de pobreza existente no Brasil, nestes últimos dez anos. Pesquisas futuras se encarregarão de utilizar estes indicadores, como elementos-base para a análise da influência dos próprios fatores da pobreza sobre os resultados.

Em conclusão, pôde-se averiguar que o padrão de pobreza brasileiro sofreu mudanças consideráveis ao longo de uma década. Se, até recentemente, o trabalho infantil era uma variável que pesava, de forma independente, na determinação da pobreza, hoje é uma variável secundária. As condições e a estrutura do trabalho assumem, cada vez mais, o papel central na determinação do nível de pobreza, caminhando *pari passu* com a posse de bens básicos e necessidades indispensáveis. Em um plano mais secundário, a localização geográfica do domicílio e a estrutura étnica ainda são discriminantes da pobreza, enquanto a estrutura da casa e as condições de vida vêm perdendo importância na determinação, até mesmo devido aos avanços nos programas federais de erradicação da pobreza.

Sem dúvida nenhuma, discriminações regionais e de renda podem ser feitas, para que se analise o padrão de pobreza específico de cada discriminante. É patente que a localização geográfica é fundamental para determinar as necessidades de cada lugar, assim como a faixa de renda é um importante divisor de águas. Entretanto, dentro dos objetivos propostos para este trabalho, já fica clara esta segmentação sociodemográfica, que incorpora limites geográficos claros e latentes diferenciações étnicas e de gênero.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALTIMIR, O. **La dimensión de la pobreza en América Latina**. Cuaderno de la CEPAL n° 27. Santiago de Chile: CEPAL, 1979. 99p.
- ATKINSON, A. B. On the Measurement of Poverty. **Econometrica**. Malden: John Wiley & Sons Inc., v. 55, n. 4, 1987. pp. 749-764.
- BANCO MUNDIAL. **World Development Report 2000/2001: Attacking Poverty**. Oxford: Oxford University Press, 2000. 335p.
- BARROS, R. P. **Sobre a Evolução Recente da Pobreza e da Desigualdade**. Instituto de Pesquisa Econômica Aplicada. Brasília, 2009. Disponível em: <<http://bit.ly/14smM7s>>. Acesso em: 16 ago. 2012.
- _____; HENRIQUES, R.; MENDONÇA, R. Desigualdade e pobreza no Brasil: retrato de uma estabilidade inaceitável. **Revista Brasileira de Ciências Sociais**, São Paulo: ANPOCS, v. 15, n. 42, 2000. p. 124-142
- _____; HENRIQUES, R.; MENDONÇA, R. **A estabilidade inaceitável: desigualdade e pobreza no Brasil**. Texto para Discussão n° 800. Brasília: IPEA, 2001. 29p.
- BECKERMAN, W. **A poverty of reason: sustainable development and economic growth**. Oakland: The Independent Institute, 2002. 94p.
- COUDOUEL, A.; HENTSCHEL, J. S.; WODON, Q. T. Poverty Measurement and Analysis. In: KLUGMAN, J. (ed.) **A Sourcebook for Poverty Reduction Strategies** (Volume 1 – Core Techniques and Cross-Cutting Issues). Washington DC: The World Bank, 2002. pp. 27-74.
- CRONBACH, L. J. Test “reliability”: its meaning and determination. **Psychometrika**. Greensboro: The Psychometric Society, v. 12, n. 1, 1947. pp. 1-16.
- _____. Coefficient Alpha and the Internal Structure of Tests. **Psychometrika**. Greensboro: The Psychometric Society, v. 16, n. 3, 1951. pp. 297-334.
- DAVIS, J. A. A Formal Interpretation of the Theory of Relative Deprivation. **Sociometry**. Washington DC: American Sociological Association, v. 22, n. 4, 1959. pp. 280-296.
- DELEEK, H. La distribution inégale des dépenses sociales en Belgique. **Politiques et management public**. Cachan: Lavoisier, v. 3, n. 3, 1985. pp. 97-130.
- DREWNOWSKI, J. Poverty: Its Meaning and Measurement. **Development and Change**. The Hague: The International Institute of Social Studies, v. 8, n. 2, 1977. pp. 183-208.
- DUESENBERY, J. S. **Income, saving, and the theory of consumer behavior**. Cambridge (Massachusetts): Harvard University Press, 1949. 128p.
- FESTINGER, L. A Theory of Social Comparison Processes. **Human Relations**. London: The Tavistock Institute, v. 7, n. 2, 1954. pp. 174-140.

- FLIK, R. J.; VAN PRAAG, B. M. S. Subjective poverty line definitions. **De Economist**. Dordrecht: Kluwer Academic Publishers, v. 139, n. 3, 1991. pp. 311-330.
- FRIEDMAN, M. **A Theory of the Consumption Function**. Princeton: Princeton University Press, 1957. 259p
- HAGENAARS, A. J. M. **The Perception of Poverty**. Amsterdam: North-Holland, 1986. 301p.
- _____; DE VOS, K. The definition and measurement of poverty. **The Journal of Human Resources**. Madison: The University of Wisconsin Journals Division, v. 23, n. 2, 1988. pp. 211-221.
- HAIR, J. F.; TATHAM, R. L.; ANDERSON, R. E.; BLACK, W. **Análise multivariada de dados**. Porto Alegre: Bookman, 2006. 593p.
- HAMILTON, B. W. Using Engel's Law to Estimate CPI Bias. **The American Economic Review**. Pittsburgh: American Economic Association, v. 91, n. 3, 2001. pp. 619-630.
- KAGEYAMA, Â.; HOFFMANN, R. Pobreza no Brasil: uma perspectiva multidimensional. **Economia e Sociedade**. Campinas: IE/UNICAMP, v. 15, n. 1, 2006. pp. 79-112.
- KAPTEYN, A.; KOOREMAN, P.; WILLEMSE, R. Some Methodological Issues in the Implementation of Subjective Poverty Definitions. **The Journal of Human Resources**. Madison: The University of Wisconsin Journals Division, v. 23, n. 2, 1988. pp. 222-242.
- KLEFENS, P. C. O. **O Biplot na análise fatorial multivariada**. 2009. 102 f. Dissertação (Mestrado em Ciências) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2009.
- MARIA, P. F. **Determinantes da pobreza no Brasil: estudo com regressões logit**. Universidade Estadual de Campinas. Campinas, 2012. Disponível em: <<http://bit.ly/TSMmcT>>. Acesso em: 15 mar. 2013.
- _____; MAIA, A. G. Indicadores objetivos de pobreza: estudo da situação brasileira a partir da POF 2008-09. *In: Anais do 51º Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural*, Belém, 2013. Trabalho ainda não publicado.
- MARKS, G. N. **Income poverty, subjective poverty and financial stress**. Social Policy Research Paper nº 29. Melbourne: Melbourne Institute of Applied Economic and Social Research, 2007. 94p.
- MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005. 295p.
- MITCHELL, W. C. **The Rationality of Economic Activity**. Chicago: The University of Chicago Press, 1910. 37p.

- MODIGLIANI, F.; BRUMBERG, R. Utility Analysis and the Consumption Function: An Interpretation of Cross-section Data. *In: KURIHARA, K. K. (ed.) Post Keynesian economics*. New Brunswick: Rutgers University Press, 1954. pp. 388-436.
- NARAYAN, D.; PATEL, R.; SCHAFFT, K.; RADEMACHER, A.; KOCH-SCHULTE, S. **Voices of the Poor: can anyone hear us?** Oxford: Oxford University Press, 2000. 277p. Disponível em: <<http://bit.ly/ZPCoAU>>. Acesso em: 17 dez. 2012.
- OGAKI, M. Engel's Law and Cointegration. **Journal of Political Economy**. Chicago: The University of Chicago Press, v. 100, n. 5, 1992. pp. 1027-1046.
- ORSHANSKY, M. Counting the Poor: Another Look at the Poverty Profile. **Social Security Bulletin**. Woodlawn: U.S. Social Security Administration, v. 28, n. 1, 1965. pp. 3-29.
- PALLEY, T. **The Relative Income Theory of Consumption: a Synthetic Keynes-Duesenberry-Friedman Model**. PERI Working Paper Series n° 170. Amherst: Political Economy Research Institute, 2008. 27p. Disponível em: <<http://bit.ly/VMWnL6>>. Acesso em: 28 jan. 2013.
- PATTEN, S. N. **The Consumption of Wealth**. Philadelphia: University of Pennsylvania Press, 1889. 70p.
- PRADHAN, M.; RAVALLION, M. **Measuring Poverty Using Qualitative Perceptions of Welfare**. Policy Research Working Paper n° 2011. Washington DC: The World Bank, 1997. 42p.
- QUINTERO, L. F. A.; MEJÍA, A. M. O. Percepción subjetiva de los pobres: una alternativa a la medición de pobreza. **Reflexión Política**. Bucaramanga: Universidad Autónoma de Bucaramanga, v. 8, n. 15, 2006. pp. 26-40.
- RAVALLION, M. Poverty Lines. *In: DURLAUF, S. N.; BLUME, L. E. (orgs.) The New Palgrave Dictionary of Economics* (8 volumes). 2ª Edição. Basingstoke: Palgrave Macmillan, 2008. 7344p.
- _____. **Poor or Just Feeling Poor? On Using Subjective Data in Measuring Poverty**. Policy Research Working Paper n° 5968. Washington DC: The World Bank, 2012. 39p.
- _____.; CHEN, S. **Weakly Relative Poverty**. Policy Research Working Paper n° 4844. Washington DC: The World Bank, 2009. 31p.
- RIBAS, R. P. **Determinantes do risco de pobreza urbana no Brasil durante a década de 90**. Texto para Discussão n° 266. Belo Horizonte: CEDEPLAR, 2005. 24p.
- ROCHA, R.; URANI, A. Distribuição da Renda no Brasil: um Ensaio sobre a Desigualdade Desconhecida. *In: Anais do 33º Encontro Nacional de Economia*, Natal, 2008. Disponível em: <<http://bit.ly/sDiyx>>. Acesso em 26 jan. 2013.
- ROCHA, S. Pobreza no Brasil: parâmetros básicos e resultados empíricos. **Pesquisa e Planejamento Econômico**, Brasília: IPEA, v. 22, n. 3, 1992. pp. 541-560.

- _____. **Poverty studies in Brazil: a review.** Texto para Discussão nº 398. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada, 1996.
- _____. Pobreza e indignação no Brasil: algumas evidências empíricas com base na PNAD 2004. **Nova Economia**, Belo Horizonte: CEDEPLAR, v. 16, n. 2, 2006. pp. 265-299.
- ROWNTREE, B. S. **Poverty: a study of town life.** Londres: MacMillan, 1901. 437p.
- SANTANA, A C. Índice de desempenho competitivo das empresas de polpa de frutas do Estado do Pará. **Revista de Economia e Sociologia Rural**, Brasília: Sociedade Brasileira de Economia, Administração e Sociologia Rural, v. 45, n. 3, 2007. pp. 749-775.
- SAS Institute Inc. The FACTOR Procedure: Overview. **SAS/STAT® 9.2 User's Guide for the SAS System 9.2 for Windows ®, Second Edition.** 2013. Disponível em: <<http://bit.ly/10EWY5T>>. Acesso em: 10 abr. 2012.
- SAWYER, M. C. Poverty in some developed countries. *In: 14º General Conference of the International Association for Research in Income and Wealth*, Finlândia, 1975.
- SEN, A. K. **Poverty and Famines: an Essay on Entitlement and Deprivation.** Oxford: Oxford University Press, 1981. 257p.
- _____. Poor, relatively speaking. **Oxford Economic Papers.** Oxford: Oxford University Press, v. 35, n. 2, 1983. pp. 153-169.
- _____. **Commodities and Capabilities.** Amsterdam: North-Holland, 1985. 130p.
- _____. **Development as Freedom.** Oxford: Oxford University Press, 2001. 384p.
- SOARES, S. S. D. **Metodologias para estabelecer a linha de pobreza:** objetivas, subjetivas, relativas, multidimensionais. Texto para Discussão nº 1381. Brasília: IPEA, 2009. 53p.
- STOUFFER, S. A.; SUCHMAN, E. A.; DEVINNEY, L. C.; STAR, S. A.; WILLIAMS JR., R. M. **The American Soldier: Adjustment during Army Life (Studies in Social Psychology in World War II, Volume 1).** Princeton: Princeton University Press, 1949. 599p.
- TOWNSEND, P. Measuring Poverty. **The British Journal of Sociology.** London: London School of Economics, v. 5, n. 2, 1954. pp. 130-137.
- _____. The meaning of poverty. **The British Journal of Sociology.** London: London School of Economics, v. 13, n. 3, 1962. pp. 210-227.
- _____. **Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living.** Berkeley: University of California Press, 1979. 1216p.
- _____. Podemos medir la pobreza en términos de privaciones múltiples. **Punto de Equilibrio.** Lima: Universidad del Pacífico, v. 13, n. 84, 2004. pp. 8-10.
- UGÁ, V. D. A categoria “pobreza” nas formulações de política social do Banco Mundial. **Revista de Sociologia e Política.** Curitiba: Universidade Federal do Paraná, n. 23, 2004. pp. 55-62.

- VAN PRAAG, B. **Individual Welfare Functions and Consumer Behaviour: A theory of rational irrationality**. Amsterdam: North-Holland, 1968. 239p.
- VAZ, F. M.; SOARES, S. S. D. Linhas de pobreza subjetivas para o Brasil. *In: Anais do 36º Encontro Nacional de Economia*, Salvador, 2008. Disponível em: <<http://bit.ly/V9zDV5>>. Acesso em 23 jan. 2013.
- VEBLÉN, T. **Essays in Our Changing Order**. New York: The Viking Press, 1934. 470p.
- VELÁZQUEZ, J. J. N. Estado actual y nuevas aproximaciones a la medición de la pobreza. **Estudios de Economía Aplicada**. Cachan: Lavoisier, v. 27, n. 2, 2009. pp. 325-344.
- WATTS, H. W. The Iso-Prop Index: na Approach to the Determination of Differential Poverty Income Thresholds. **Journal of Human Resources**. Madison: The University of Wisconsin Journals Division, v. 2, n. 1, 1967. pp. 3-18.

APÊNDICE A – Variáveis selecionadas

Variável selecionada	Nome	Código PNAD
Água encanada	agenc	0211
Alfabetização	alfab	0601
Anos de estudo completos	anest	4738 ou 4838
Carteira assinada	cartass	9042
Coleta de esgoto	colesg	0217
Coleta de lixo	collix	0218
Computador/Internet	compint	0231 e 0232
Contribuição previdenciária	contpr	9059
Cor/Raça	corrac	0404
Geladeira	gelad	0228
Horas trabalhadas	hortr	4707
Máquina de lavar	maqlav	0230
Migração (por Estados)	migr	0502
Natimortos	natim	1111 e 1112
Número de banheiros	numb	2016
Número de cômodos	numc	0205
Proveniência da água	provag	0212
Qualidade das paredes	qualp	0203
Qualidade do telhado	qualt	0204
Recebimento de auxílios	recaux	9043 a 9047
Rendimento mensal per capita	rendpc	4742
Telefone (fixo/celular)	telef	0220 e 2020
Televisão	telev	0226 e 0227
Tipo de contratação	tipc	9041
Tipo de família	tipf	4723
Tipo de iluminação	tipi	0219
Tipo de trabalho	tipt	4817
Trabalha	trab	9001 a 9004
Trabalho infantil	trabinf	0701
Unidade da Federação	uf	UF
Número de pessoas que moram no domicílio	v0105	0105

APÊNDICE B – Medida de adequação da amostra de Kaiser-Mayer-Ohlin (amostra completa)

	2002	2003	2004	2005	2006	2007	2008	2009	2011
Geral	0,845	0,842	0,846	0,848	0,841	0,838	0,832	0,830	0,832
uf	0,868	0,863	0,898	0,897	0,904	0,898	0,891	0,885	0,894
v0105	0,629	0,610	0,665	0,622	0,579	0,583	0,541	0,531	0,541
qualp	0,879	0,885	0,892	0,903	0,910	0,907	0,900	0,902	0,908
qualt	0,582	0,568	0,639	0,643	0,631	0,632	0,603	0,595	0,624
numc	0,776	0,759	0,786	0,764	0,736	0,731	0,701	0,678	0,668
agenc	0,767	0,765	0,777	0,784	0,774	0,764	0,758	0,753	0,756
provag	0,764	0,761	0,770	0,776	0,767	0,758	0,752	0,747	0,755
colesg	0,942	0,940	0,946	0,946	0,944	0,937	0,934	0,934	0,930
collix	0,914	0,912	0,916	0,916	0,910	0,904	0,901	0,896	0,902
tipi	0,891	0,889	0,892	0,888	0,874	0,849	0,831	0,823	0,807
telef	0,926	0,927	0,930	0,930	0,928	0,922	0,918	0,920	0,919
telev	0,906	0,900	0,899	0,900	0,881	0,859	0,842	0,838	0,830
gelad	0,919	0,914	0,919	0,918	0,919	0,916	0,909	0,901	0,878
maqlav	0,922	0,923	0,916	0,922	0,922	0,921	0,920	0,917	0,918
compint	0,891	0,894	0,887	0,888	0,890	0,885	0,882	0,883	0,888
corrac	0,908	0,915	0,903	0,914	0,917	0,911	0,898	0,896	0,891
migr	0,868	0,869	0,880	0,875	0,889	0,858	0,872	0,869	0,860
alfab	0,931	0,928	0,932	0,924	0,924	0,920	0,914	0,913	0,908
trabinf	0,933	0,922	0,930	0,942	0,936	0,927	0,926	0,908	0,890
natim	0,696	0,693	0,671	0,691	0,707	0,734	0,734	0,741	0,769
numb	0,775	0,761	0,779	0,756	0,726	0,717	0,685	0,676	0,660
hortr	0,852	0,854	0,785	0,859	0,860	0,867	0,868	0,868	0,880
tipf	0,507	0,488	0,545	0,500	0,467	0,493	0,480	0,473	0,488
rendpc	0,890	0,877	0,885	0,880	0,876	0,869	0,865	0,871	0,861
tipt	0,837	0,837	0,891	0,835	0,836	0,836	0,837	0,834	0,841
anest	0,920	0,923	0,920	0,925	0,927	0,925	0,924	0,923	0,920
trab	0,768	0,770	0,758	0,773	0,775	0,776	0,777	0,779	0,786
tipc	0,876	0,875	0,872	0,883	0,882	0,882	0,885	0,885	0,881
cartass	0,747	0,749	0,753	0,757	0,758	0,764	0,765	0,770	0,776
recaux	0,955	0,956	0,954	0,954	0,952	0,954	0,954	0,951	0,952
contpr	0,812	0,815	0,824	0,824	0,825	0,831	0,834	0,841	0,851

APÊNDICE C – Comunalidade das variáveis selecionadas por ano (amostra completa)

	2002	2003	2004	2005	2006	2007	2008	2009	2011
Geral	19,111	19,129	18,890	19,249	19,177	19,043	18,966	18,874	18,818
uf	0,616	0,627	0,604	0,598	0,590	0,586	0,582	0,579	0,541
v0105	0,650	0,665	0,649	0,683	0,685	0,714	0,710	0,714	0,699
qualp	0,498	0,471	0,486	0,420	0,388	0,463	0,490	0,463	0,473
qualt	0,656	0,654	0,633	0,619	0,583	0,588	0,633	0,628	0,638
numc	0,598	0,600	0,601	0,603	0,595	0,605	0,616	0,635	0,613
agenc	0,708	0,709	0,719	0,714	0,719	0,740	0,729	0,724	0,694
provag	0,728	0,726	0,750	0,748	0,770	0,814	0,809	0,811	0,795
colesg	0,561	0,550	0,548	0,551	0,543	0,555	0,541	0,528	0,540
collix	0,533	0,537	0,551	0,562	0,545	0,559	0,562	0,565	0,563
tipi	0,417	0,408	0,443	0,456	0,484	0,643	0,648	0,631	0,614
telef	0,597	0,591	0,588	0,602	0,600	0,598	0,589	0,570	0,545
telev	0,504	0,493	0,522	0,508	0,540	0,669	0,673	0,648	0,609
gelad	0,513	0,495	0,511	0,492	0,494	0,534	0,533	0,527	0,551
maqlav	0,511	0,518	0,543	0,532	0,537	0,523	0,519	0,524	0,513
compint	0,565	0,570	0,580	0,592	0,601	0,606	0,605	0,601	0,583
corrac	0,476	0,461	0,501	0,496	0,486	0,496	0,517	0,516	0,528
migr	0,386	0,356	0,434	0,392	0,363	0,388	0,377	0,385	0,373
alfab	0,395	0,385	0,396	0,418	0,420	0,400	0,413	0,418	0,427
trabinf	0,936	0,986	0,974	0,915	0,911	0,143	0,029	0,041	0,142
natim	0,739	0,725	0,589	0,759	0,668	0,722	0,741	0,745	0,738
numb	0,692	0,732	0,681	0,736	0,771	0,776	0,785	0,789	0,792
hortr	0,829	0,827	0,792	0,819	0,819	0,825	0,819	0,816	0,818
tipf	0,454	0,442	0,465	0,436	0,456	0,463	0,443	0,430	0,439
rendpc	0,527	0,593	0,553	0,568	0,569	0,561	0,551	0,541	0,553
tipt	0,807	0,806	0,603	0,799	0,796	0,809	0,797	0,790	0,786
anest	0,488	0,488	0,498	0,501	0,506	0,507	0,504	0,504	0,511
trab	0,886	0,884	0,833	0,877	0,876	0,883	0,875	0,872	0,871
tipc	0,636	0,639	0,641	0,645	0,644	0,652	0,661	0,665	0,683
cartass	0,819	0,815	0,812	0,817	0,818	0,819	0,817	0,813	0,802
recaux	0,611	0,601	0,611	0,615	0,627	0,624	0,623	0,627	0,612
contpr	0,777	0,774	0,776	0,775	0,773	0,776	0,777	0,774	0,773

APÊNDICE D – Fatores distribuídos por padrões (amostra completa)

Grupo	Padrão I	Padrão II-A	Padrão II-B	Padrão II-C	Padrão III
I Bens	Bens elementares	Bens elementares	Bens indispensáveis	Bens indispensáveis	Bens indispensáveis
			Bens de consumo	Bens de consumo	Bens de consumo
	Bens da família	Bens de família	Bens de família	Bens de família	Bens de família
II Trabalho	Qualidade do trabalho	Qualidade do trabalho	Qualidade do trabalho	Qualidade do trabalho	Estrutura do trabalho
	Forma de trabalho	Forma de trabalho	Forma de trabalho	Forma de trabalho	
	Estrutura familiar	Estrutura familiar	Estrutura familiar	Estrutura familiar	Estrutura familiar
III Família	Localização geográfica	Localização geográfica	Localização geográfica	Localização geográfica	Localização geográfica
	Estrutura da casa	Estrutura da casa	Estrutura da casa	Estrutura da casa	Estrutura da casa
	Condições de vida	Condições de vida	Condições de vida	Condições de vida	Condições de vida
IV Local	Trabalho infantil	Trabalho infantil			Condições de vida infantil

APÊNDICE E – Medida de adequação da amostra de Kaiser-Mayer-Ohlin (subamostra)

	2002	2003	2004	2005	2006	2007	2008	2009	2011
Geral	0,837	0,838	0,830	0,844	0,841	0,841	0,833	0,830	0,835
uf	0,773	0,782	0,719	0,840	0,848	0,840	0,825	0,837	0,848
v0105	0,567	0,575	0,599	0,575	0,577	0,575	0,540	0,532	0,535
qualp	0,851	0,874	0,902	0,903	0,910	0,903	0,904	0,901	0,922
qualt	0,725	0,752	0,588	0,553	0,576	0,559	0,556	0,546	0,577
numc	0,695	0,706	0,698	0,689	0,703	0,690	0,668	0,648	0,631
agenc	0,769	0,773	0,788	0,795	0,780	0,776	0,772	0,754	0,764
provag	0,767	0,769	0,780	0,786	0,773	0,769	0,765	0,750	0,766
colesg	0,921	0,927	0,929	0,931	0,928	0,927	0,909	0,920	0,914
collix	0,917	0,916	0,923	0,922	0,918	0,915	0,914	0,903	0,920
tipi	0,896	0,896	0,904	0,891	0,877	0,859	0,840	0,834	0,821
telef	0,923	0,923	0,930	0,928	0,927	0,923	0,917	0,917	0,915
telev	0,906	0,904	0,907	0,905	0,886	0,870	0,854	0,843	0,842
gelad	0,922	0,919	0,928	0,923	0,923	0,921	0,914	0,902	0,883
maqlav	0,919	0,921	0,911	0,920	0,923	0,920	0,914	0,911	0,916
compint	0,893	0,897	0,889	0,883	0,886	0,880	0,879	0,876	0,882
corrac	0,879	0,889	0,843	0,899	0,895	0,898	0,883	0,885	0,878
migr	0,861	0,861	0,853	0,866	0,874	0,846	0,854	0,855	0,860
alfab	0,928	0,927	0,930	0,928	0,927	0,924	0,921	0,917	0,913
trabinf	0,907	0,841	0,919	0,943	0,926	0,913	0,906	0,886	0,889
natim	0,686	0,728	0,660	0,705	0,736	0,737	0,768	0,738	0,780
numb	0,728	0,729	0,719	0,718	0,711	0,701	0,669	0,661	0,637
hortr	0,854	0,857	0,765	0,861	0,862	0,868	0,869	0,870	0,881
tipf	0,452	0,460	0,447	0,442	0,446	0,494	0,481	0,479	0,497
rendpc	0,884	0,873	0,880	0,877	0,870	0,873	0,865	0,868	0,862
tipt	0,844	0,842	0,562	0,842	0,843	0,844	0,843	0,837	0,847
anest	0,921	0,925	0,921	0,928	0,930	0,929	0,929	0,928	0,925
trab	0,777	0,778	0,752	0,781	0,784	0,785	0,785	0,784	0,793
tipc	0,876	0,873	0,872	0,886	0,887	0,888	0,890	0,890	0,888
cartass	0,755	0,756	0,758	0,765	0,766	0,774	0,774	0,779	0,786
recaux	0,955	0,955	0,955	0,954	0,952	0,953	0,953	0,949	0,951
contpr	0,817	0,820	0,826	0,830	0,829	0,839	0,839	0,846	0,856

APÊNDICE F – Comunalidade das variáveis selecionadas por ano (subamostra)

	2002	2003	2004	2005	2006	2007	2008	2009	2011
Geral	19,152	19,257	19,079	19,335	19,275	19,075	19,066	18,923	18,918
uf	0,641	0,660	0,770	0,659	0,636	0,622	0,627	0,588	0,584
v0105	0,678	0,684	0,686	0,711	0,704	0,690	0,692	0,691	0,684
qualp	0,518	0,489	0,426	0,459	0,460	0,416	0,427	0,420	0,388
qualt	0,436	0,620	0,559	0,651	0,623	0,506	0,495	0,488	0,601
numc	0,622	0,657	0,639	0,648	0,652	0,639	0,650	0,646	0,631
agenc	0,717	0,740	0,712	0,754	0,757	0,731	0,735	0,742	0,686
provag	0,728	0,746	0,733	0,779	0,793	0,787	0,791	0,813	0,772
colesg	0,569	0,563	0,569	0,574	0,553	0,567	0,582	0,550	0,565
collix	0,492	0,504	0,509	0,510	0,476	0,482	0,464	0,476	0,501
tipi	0,463	0,468	0,467	0,542	0,553	0,566	0,628	0,616	0,616
telef	0,591	0,589	0,580	0,601	0,596	0,594	0,589	0,571	0,559
telev	0,509	0,517	0,537	0,570	0,590	0,592	0,641	0,637	0,606
gelad	0,527	0,526	0,539	0,547	0,530	0,506	0,508	0,518	0,560
maqlav	0,527	0,532	0,570	0,548	0,544	0,543	0,550	0,556	0,537
compint	0,564	0,571	0,573	0,591	0,603	0,606	0,604	0,598	0,592
corrac	0,528	0,524	0,539	0,524	0,521	0,504	0,521	0,528	0,543
migr	0,406	0,420	0,541	0,498	0,480	0,478	0,436	0,413	0,358
alfab	0,416	0,418	0,490	0,437	0,438	0,451	0,458	0,457	0,493
trabinf	0,948	0,811	0,900	0,278	0,222	0,229	0,225	0,053	0,112
natim	0,788	0,534	0,467	0,776	0,815	0,810	0,734	0,913	0,902
numb	0,738	0,768	0,759	0,762	0,783	0,775	0,800	0,793	0,797
hortr	0,817	0,825	0,748	0,822	0,819	0,821	0,814	0,812	0,807
tipf	0,437	0,434	0,437	0,433	0,433	0,474	0,415	0,395	0,428
rendpc	0,471	0,602	0,537	0,545	0,599	0,556	0,583	0,569	0,535
tipt	0,787	0,804	0,692	0,811	0,802	0,808	0,793	0,788	0,772
anest	0,491	0,493	0,505	0,499	0,502	0,509	0,507	0,500	0,521
trab	0,875	0,883	0,775	0,883	0,879	0,881	0,871	0,870	0,861
tipc	0,644	0,648	0,656	0,660	0,651	0,662	0,671	0,674	0,687
cartass	0,813	0,818	0,787	0,831	0,827	0,827	0,821	0,819	0,803
recaux	0,632	0,624	0,603	0,643	0,650	0,654	0,646	0,646	0,632
contpr	0,777	0,782	0,775	0,792	0,787	0,789	0,790	0,786	0,786

APÊNDICE G – Fatores extraídos por rotação VARIMAX, por padrão observado (subamostra)

Padrão I				
2002	2003	2004	2005-2006	2007
Bens elementares	Bens elementares	Bens elementares	Bens elementares	Bens elementares
Qualidade do trabalho	Qualidade do trabalho	Estrutura do trabalho	Qualidade do trabalho	Qualidade do trabalho
Bens da família	Bens da família	Bens da família	Bens da família	Bens da família
Forma de trabalho	Forma de trabalho	Estrutura familiar	Forma de trabalho	Forma de trabalho
Estrutura familiar	Localização geográfica	Estrutura sócio-geográfica	Estrutura familiar	Estrutura familiar
Localização geográfica	Estrutura familiar	Estrutura étnica	Localização geográfica	Localização geográfica
Estrutura da casa	Condições de vida	Condições de vida	Condições da casa	Estrutura da casa
Condições de vida	Estrutura da casa	Estrutura da casa	Condições de vida	Migrações
Trabalho infantil	Trabalho infantil	Trabalho infantil	Condições de vida infantil	Condições de vida infantil

Padrão II		Padrão III
2008	2009	2011
Qualidade do trabalho	Qualidade do trabalho	Estrutura do trabalho
Bens indispensáveis	Bens da família	Bens indispensáveis
Bens da família	Bens indispensáveis	Bens da família
Forma de trabalho	Forma de trabalho	Bens de consumo
Bens de consumo	Bens de consumo	Estrutura familiar
Estrutura familiar	Estrutura familiar	Localização geográfica
Localização geográfica	Localização geográfica	Condições de vida
Estrutura da casa	Condições de vida	Estrutura da casa
Condições de vida infantil	Condições de vida infantil	Condições de vida infantil

APÊNDICE H – Ponderações normalizadas dos fatores extraídos por grupo (amostra completa)

	2002	2003	2004	2005	2006	2007	2008	2009	2011
Fator 1	0,74	0,74	0,74	0,74	0,74	0,81	0,80	0,80	1,00
Fator 2	0,75	0,75	0,74	0,75	0,74	0,55	0,55	0,55	0,54
Fator 3	0,26	0,26	0,26	0,26	0,26	0,28	0,28	0,28	0,26
Fator 4	0,25	0,25	0,26	0,25	0,26	0,19	0,20	0,20	0,20
Fator 5	1,00	1,00	1,00	1,00	1,00	0,17	0,18	0,18	1,00
Fator 6	0,51	0,51	0,51	0,50	0,50	1,00	1,00	1,00	0,52
Fator 7	0,49	0,49	0,49	0,50	0,50	0,52	0,52	0,51	0,52
Fator 8	0,50	0,50	0,51	0,50	0,50	0,48	0,48	0,49	0,48
Fator 9	0,50	0,50	0,49	0,50	0,50	1,00	1,00	1,00	0,48

Legenda	Grupo I (bens)	Grupo II (trabalho)	Grupo III (família)	Grupo IV (local)	Grupo V (vida)
----------------	--------------------------	-------------------------------	-------------------------------	----------------------------	--------------------------