

Ministério do Planejamento, Orçamento e Gestão
Instituto Brasileiro de Geografia e Estatística – IBGE
Diretoria de Pesquisas
Coordenação de Métodos e Qualidade.

Textos para discussão

Diretoria de Pesquisas

número 55

Principais Aspectos de Amostragem das Pesquisas Domiciliares do IBGE - Revisão 2015

Sonia Albieri
Zélia Magalhães Bianchini

Rio de Janeiro
2015

ISSN 1518-675X **Textos para discussão. Diretoria de Pesquisas**

Divulga estudos e outros trabalhos técnicos desenvolvidos pelo IBGE ou em conjunto com outras instituições, bem como resultantes de consultorias técnicas e traduções consideradas relevantes para disseminação pelo Instituto. A série está subdividida por unidade organizacional e os textos são de responsabilidade de cada área específica.

ISBN 978-85-240-4349-9

© IBGE. 2015

Impressão

Gráfica Digital/Centro de Documentação e Disseminação de Informações - CDDI/IBGE, em 2015.

Capa

Gerência de Criação/CDDI

Albieri, Sonia

Principais aspectos de amostragem das pesquisas domiciliares do IBGE :
revisão 2015 / Sonia Albieri, Zélia Magalhães Bianchini . - Rio de Janeiro :
IBGE, Coordenação de Métodos e Qualidade, 2015.

46 p. - (Textos para discussão. Diretoria de Pesquisas, ISSN 1518-675X ;
n. 55)

Inclui bibliografia.

ISBN 978-85-240-4349-9

1. Levantamentos domiciliares – Métodos estatísticos. 2. Amostragem
(Estatística). I. Bianchini, Zélia Magalhães. II. IBGE. Coordenação de Métodos
e Qualidade. III. Título. IV. Série.

Gerência de Biblioteca e Acervos Especiais
R.J/2015-10

CDU 519.2:314.6
EST

Impresso no Brasil / Printed in Brazil

Sumário

Apresentação	5
Introdução	7
As Pesquisas Isoladas	8
<i>Pesquisa Nacional por Amostra de Domicílios (PNAD)</i>	8
<i>Pesquisa Mensal de Emprego (PME)</i>	11
De 1980 a 2001	11
Revisão 2002	12
<i>Pesquisa de Orçamentos Familiares (POF)</i>	13
Abrangência 11 áreas.....	13
Abrangência Nacional	15
<i>Pesquisa sobre Padrões de Vida (PPV)</i>	20
<i>Pesquisa de Economia Informal Urbana (ECINF)</i>	22
<i>Pesquisa Especial sobre Tabagismo (PETAB)</i>	24
<i>Pesquisa das Características Étnico-raciais da População (PCERP)</i>	26
O Sistema Integrado de Pesquisas Domiciliares - SIPD	28
<i>Amostra Mestra para o SIPD - Base 2000</i>	28
<i>Amostra Mestra para o SIPD - Base 2010</i>	29
<i>Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua)</i>	31
<i>Pesquisa Nacional de Saúde 2013 (PNS)</i>	33
Pesquisa por Amostragem na Coleta de Dados do Censo Demográfico	34
Resumo e conclusões	37
Referências	38
Anexo	41

Apresentação

Este documento é uma atualização dos artigos "Principais Aspectos de Amostragem das Pesquisas Domiciliares do IBGE - revisão 2002", publicado em Textos para Discussão da Diretoria de Pesquisas, nº 8, de 2003, e do artigo "Uma Revisão dos Principais Aspectos dos Planos Amostrais das Pesquisas Domiciliares Realizadas pelo IBGE", publicado nessa mesma série, Textos para Discussão nº91, de setembro de 1998.

Desde 1960, o IBGE vem usando amostragem probabilística na realização de suas principais pesquisas domiciliares. O artigo de 2002 descreve alguns aspectos metodológicos das principais pesquisas domiciliares por amostra realizadas pelo IBGE, a saber: a Pesquisa Nacional por Amostra de Domicílios (PNAD, iniciada em 1967); a Pesquisa Mensal de Emprego (PME, desde 1980); a Pesquisa de Orçamentos Familiares (POF, com as edições, de 1987-88, 1995-96 e 2002-2003); a Pesquisa sobre Padrões de Vida (PPV, realizada em 1996-97); a Pesquisa de Economia Informal Urbana (ECINF, realizada em 1997); e a amostra para a coleta do questionário detalhado do Censo Demográfico (de 1960 até 2000),

Além das pesquisas abordadas no artigo de 2002, esta atualização de 2015 contempla as seguintes pesquisas: Pesquisa Mensal de Emprego Revisada (PME Revisada); Economia Informal Urbana 2003 (ECINF); Pesquisa Especial sobre Tabagismo (PETAB), na forma de Suplemento da PNAD 2008; Pesquisa de Orçamentos Familiares 2008-2009 (POF), que passou a ser nacional e planejamento da Pesquisa de Orçamentos Familiares 2015-2016, no âmbito do Sistema Integrado de Pesquisas Domiciliares (SIPD); Pesquisa das Características Étnico-raciais da População (PCERP), realizada uma única vez, em 2008; três edições da investigação sobre o tema Saúde, como Suplemento da PNAD em 2003 e em 2008 e como pesquisa isolada, a Pesquisa Nacional de Saúde 2013 (PNS), esta última no âmbito do SIPD; a Amostra Mestra definida no final da década passada para o Sistema Integrado de Pesquisas Domiciliares - SIPD e sua atualização para a implantação da PNAD Contínua, no final de 2011. Além disso, são descritos os aspectos metodológicos referentes à investigação do questionário da amostra do Censo Demográfico 2010.

O artigo aborda as semelhanças e as diferenças principais entre as pesquisas, no que se refere a objetivos, população-alvo, abrangência geográfica, plano amostral, incluindo aspectos tais como: cadastros, estratificação, conglomeração, número de estágios, tamanho da amostra, seleção da amostra, taxas de não resposta e tratamentos adotados, estimação, avaliação de erros amostrais.

Sonia Albieri
Coordenação de Métodos e Qualidade

Introdução

O IBGE realiza várias pesquisas domiciliares com periodicidades diferentes (mensais, anuais, ou não definidas), com graus variados de abrangência geográfica (nacional, regional, regiões metropolitanas, apenas áreas urbanas, alguns municípios de capitais de estado) e complexidade, de acordo não só com os recursos disponíveis, mas também com a área temática e objetivos de cada pesquisa. Este artigo revisa alguns aspectos metodológicos das principais pesquisas domiciliares por amostra realizadas pelo IBGE, a saber:

- Pesquisa Nacional por Amostra de Domicílios (PNAD, iniciada em 1967);
- Pesquisa Mensal de Emprego (PME, desde 1980);
- Pesquisa de Orçamentos Familiares (POF, com as edições, de 1987-88, 1995-96, 2002-2003, 2008-2009 e 2015-2016, em fase de planejamento);
- Pesquisa sobre Padrões de Vida (PPV, realizada em 1996-97);
- Pesquisa de Economia Informal Urbana (ECINF, realizada em 1997 e em 2003);
- Pesquisa Especial sobre Tabagismo (PETAB), na forma de Suplemento da PNAD 2008;
- Pesquisa das Características Étnico-raciais da População (PCERP, realizada uma única vez, em 2008);
- Três edições da investigação sobre o tema Saúde, como Suplemento da PNAD em 2003 e em 2008 e como pesquisa isolada, a Pesquisa Nacional de Saúde 2013 (PNS);
- A Amostra Mestra definida no final da década passada para o Sistema Integrado de Pesquisas Domiciliares - SIPD e sua atualização para a implantação da PNAD Contínua, no final de 2011; e,
- A amostra para a investigação do questionário detalhado dos Censos Demográficos de 1960 a 2010.

Todas essas pesquisas adotam amostras probabilísticas de domicílios. Os desenhos amostrais das pesquisas listadas acima possuem várias semelhanças, as quais incluem amostragem de conglomerados em dois estágios (setores censitários + domicílios) ou mesmo três estágios (municípios + setores censitários + domicílios) com estratificação das unidades primárias de amostragem (UPAs). As UPAs (municípios ou setores censitários) são selecionadas com probabilidade proporcional a uma medida de tamanho – ppt (dentro de cada estrato). Os dados dos Censos Demográficos são frequentemente usados para construir medidas de tamanho. Os setores censitários possuem em média 300 domicílios na

área urbana e 200 domicílios na área rural. Para cada setor selecionado para a amostra de cada pesquisa é preparada uma listagem de todos os seus domicílios com o objetivo de preparar um cadastro atualizado para a seleção dos domicílios no último estágio de seleção. A precisão das estimativas é medida através dos coeficientes de variação (CVs) calculados para um conjunto de estimativas. O método do *ultimate cluster* de Hansen et al. (1953) é o que vem sendo usado na maioria das pesquisas para estimar a variância das estimativas de interesse.

As seções seguintes apresentam aspectos gerais de cada pesquisa aqui considerada, bem como uma breve descrição das principais diferenças entre as pesquisas, com relação a objetivos, população alvo, abrangência geográfica e plano amostral, incluindo aspectos tais como: cadastro, estratificação, conglomeração, número de estágios de seleção, tamanho da amostra, procedimento de seleção utilizado, taxas de não resposta e tratamentos adotados, método de estimação e avaliação da precisão das estimativas.

Além dessas pesquisas, o sistema de pesquisas domiciliares inclui o Censo Demográfico, realizado a cada 10 anos, que usa amostragem para obter dados sobre características selecionadas de pessoas, famílias e domicílios. O plano amostral e a metodologia de estimação adotados na pesquisa amostral do Censo são particularmente diferentes daqueles utilizados nas pesquisas realizadas no período intercensitário. Em uma seção isolada, é apresentada uma breve descrição da amostra usada para a coleta do questionário de amostra dos Censos Demográficos.

Vale ressaltar que este documento não tem por objetivo descrever com profundidade os planos amostrais das pesquisas. Os detalhes metodológicos em geral e específicos de amostragem de cada pesquisa podem ser encontrados nas referências relacionadas no final do documento.

As Pesquisas Isoladas

Pesquisa Nacional por Amostra de Domicílios (PNAD)

A PNAD foi implantada no País, gradativamente, a partir de 1967. O plano amostral da pesquisa foi inspirado no Plano Atlântida do *U.S. Bureau of the Census*. Desde então, essa pesquisa tornou-se a pesquisa domiciliar anual mais importante no Brasil. É uma pesquisa com múltiplos propósitos que investiga características econômicas e sociais, principalmente aquelas relativas à situação da força de trabalho. Até 2003, a não era realizada na área rural da região Norte. A partir de 2004, a pesquisa passou a contemplar as áreas rurais das Unidades da Federação Rondônia, Acre, Amazonas, Roraima, Pará e Amapá, tornando-se de fato com abrangência nacional. Além disso, ocasionalmente são aplicados questionários suplementares na mesma amostra e ao mesmo tempo em que se coleta a pesquisa tradicional, com o objetivo de investigar tópicos específicos, tais como saúde, educação, trabalho, seguridade social, fertilidade, acesso aos serviços de saúde, contracepção, participação política.

A pesquisa possui um plano amostral autoponderado¹ com três estágios de seleção, estratificado, popularizado na década de 60 (ver Kish, 1965), para uso em pesquisas domiciliares. As unidades primárias de seleção são os municípios, que são estratificados por tamanho em termos de população e selecionados sistematicamente com probabilidade proporcional ao tamanho. No segundo estágio, os setores censitários são selecionados também de forma sistemática e com ppt, sendo que nesse caso o tamanho é medido pelo número de domicílios. Uma amostra sistemática simples de domicílios é então selecionada no terceiro estágio.

Em 2004, a inclusão da área rural da região norte foi efetivada segundo metodologia que partiu dos municípios já selecionados no primeiro estágio do processo de seleção. Os setores rurais foram selecionados da mesma forma que os setores urbanos, mantendo a mesma fração de amostragem utilizada para os setores urbanos. Como, para alguns municípios, a aplicação direta dessa fração de amostragem resultaria em uma enorme quantidade de unidades domiciliares a serem entrevistadas por setor, sem o benefício equivalente no nível de precisão das estimativas, foi redefinida a fração de amostragem para a área rural de cada Unidade da Federação em questão e foram adotados fatores de ajuste de subamostragem variados para esses municípios.

O plano amostral adotado para a PNAD considera uma estratificação das unidades primárias (municípios), definida separadamente em cada unidade da federação, com seleção de duas unidades por estrato. Municípios pertencentes à mesma microrregião geográfica foram agrupados em estratos com aproximadamente o mesmo tamanho. Os dados de população provenientes do Censo Demográfico foram as medidas de tamanho usadas para os procedimentos de estratificação e de seleção dos municípios. Os municípios grandes em termos populacionais e aqueles pertencentes às regiões metropolitanas foram tratados cada um como um estrato e, portanto, incluídos na amostra com certeza, e denominados auto-representativos. Os demais municípios selecionados em cada estrato são denominados não auto-representativos e em cada um foram selecionados 5 setores.

O plano amostral é caracterizado por fração amostral fixa para cada região metropolitana² e para o restante da unidade da federação. Os municípios e os setores selecionados são mantidos na amostra até que estejam disponíveis os novos dados do Censo Demográfico, quando então são selecionadas novas unidades para a amostra. No momento em que foi feita a seleção de setores, o número de domicílios por setor para a amostra foi fixado e constante para todos os municípios. Quando a seleção da amostra foi atualizada com as novas definições de setores e com as medidas de tamanho baseadas nos dados do Censo Demográfico de 1991, o número de domicílios por setor na amostra foi fixado em 13. Com a realização do Censo Demográfico de 2000, além de novas informações sobre a população e os domicílios, houve uma redefinição da base de setores censitários. O território brasileiro foi dividido em um número bem maior de setores

¹ Em um plano amostral autoponderado, a probabilidade de uma unidade elementar qualquer pertencer à amostra é constante e igual à fração de amostragem geral, para um determinado nível geográfico.

² A PNAD considera nove regiões metropolitanas, a saber: Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre.

censitários do que havia em 1991. Assim, foi feita nova seleção da amostra, inclusive unidades primárias, e foi feito também um ajuste nos tamanhos das amostras de setores para dar conta dessa nova partição. O número de domicílios por setor na amostra foi fixado em 16.

A cada ano, em cada setor selecionado para a amostra, é preparada (ou atualizada) no campo uma listagem de domicílios, produzindo um cadastro atualizado para a seleção dos domicílios. Uma característica importante dessa operação de listagem refere-se ao Cadastro de Novas Construções, que é preparado de forma a conter os projetos responsáveis por alterações sérias nos tamanhos dos setores. O levantamento das novas construções é feito nos municípios da amostra, tanto nos setores da amostra como naqueles não selecionados para a amostra. Uma área de novas construções é excluída da área do setor original e é tratada em separado no momento da seleção de domicílios, que nesse caso, é feita de acordo com a fração amostral da área.

Como a seleção de domicílios em cada setor selecionado para a amostra é feita sistematicamente, para garantir a autoponderação da amostra, o intervalo de seleção de domicílios permanece fixo de ano para ano. Esse procedimento acarreta um aumento anual no número de domicílios na amostra. Em um dado setor da amostra, o número de domicílios selecionados depende do tamanho atualizado do setor, em número de domicílios, atualização essa dada pela operação de listagem. Esse fato não é observado nas PNADs de 2011 em diante porque, a partir de 2010, não houve atualização do Cadastro de Nacional de Endereços para Fins Estatísticos - CNEFE³ em todos os setores da amostra da PNAD.

Analogamente ao ocorrido após o Censo 2000, com a realização do Censo Demográfico de 2010, além de novas informações sobre a população e os domicílios, houve uma redefinição da base de setores censitários, resultando em um número bem maior de setores censitários do que havia em 2000. Assim, foi feita nova seleção da amostra para ser utilizada durante a década, a partir de 2011.

O procedimento de estimação adotado na PNAD é baseado em estimação de razão com ajuste de população. Inicialmente os pesos associados aos domicílios são obtidos considerando o plano amostral adotado sem tratamento para não resposta. Então, esses pesos são ajustados (multiplicados) por um fator que é calculado independentemente para cada região metropolitana e para o restante de cada unidade da federação. Esse fator é a razão entre a estimativa independente de população⁴ para o período de referência da pesquisa e a estimativa de população proveniente da amostra. Os pesos associados às pessoas são aqueles calculados para os domicílios onde moram. Até 2004, foi utilizada a projeção de população residente urbana como variável independente para a expansão da amostra das seis unidades da Federação da Região Norte, quando a pesquisa não cobria a área rural. Em 2004, apenas para essas seis Unidades da Federação, foi adotada a projeção da população residente separadamente por situação do domicílio, urbana ou rural,

³ A Operação de Listagem de domicílios nos setores da amostra foi substituída pela atualização do CNEFE, com a mesma finalidade.

⁴ As estimativas independentes de população são preparadas pelo próprio IBGE, anualmente, para diversos níveis geográficos.

como variável independente para a expansão da amostra. Em 2011, com a nova seleção da amostra para a década de 2010, a expansão passou a utilizar a estimativa da população total para essas Unidades da Federação.

A precisão das estimativas provenientes da amostra é medida pelo coeficiente de variação (CV) calculado para um conjunto selecionado de estimativas de totais de variáveis categóricas. A variância do estimador de razão adotado é calculada pelo método denominado *ultimate cluster* (Hansen et al., 1953). As estimativas dos coeficientes de variação para variáveis categóricas são divulgadas através do ajuste de um modelo de regressão do tipo *Generalized variance functions* (Wolter, 1985). Ou seja, para cada domínio de publicação dos resultados, são ajustados modelos de regressão para explicar os CVs das estimativas de total como função das próprias estimativas. Os coeficientes dos modelos de regressão ajustados são publicados juntamente com os resultados da pesquisa a cada ano.

Desde 1984, os resultados são publicados nos níveis nacional, de grandes regiões, unidades da federação e regiões metropolitanas abrangidas pela pesquisa.

Pesquisa Mensal de Emprego (PME)

De 1980 a 2001

A Pesquisa Mensal de Emprego foi iniciada em 1980. É uma pesquisa por amostra domiciliar realizada para fornecer estimativas do nível e das variações no emprego, desemprego e de outras características da força de trabalho. A população objetivo da pesquisa é constituída pelas pessoas de 10 anos ou mais de idade residentes na área urbana de cada uma das seis regiões metropolitanas em que é realizada, a saber: Rio de Janeiro, São Paulo, Belo Horizonte, Recife, Salvador e Porto Alegre. Os resultados são produzidos para cada uma dessas regiões, bem como para o agregado das seis regiões, agregado esse que em 1996 representava cerca de 25% da população brasileira. Para cada uma dessas regiões, a amostra foi selecionada em 2 estágios de forma independente. Em cada região metropolitana, os municípios constituem estratos de setores, os quais por sua vez são as unidades primárias de seleção dessa pesquisa. Os setores são selecionados sistematicamente com probabilidade proporcional a uma medida de tamanho (número de domicílios). Em cada região metropolitana, foi usada alocação proporcional dos setores ao longo dos estratos (municípios). Tal como na PNAD, após a atualização da listagem de domicílios em cada setor da amostra, o número de domicílios selecionados por setor aumenta de forma a manter a autoponderação. O Cadastro de Novas Construções preparado para a PNAD também é utilizado na PME.

Como a pesquisa é repetida a cada mês, foi estabelecido um esquema de rotação para a amostra a fim de evitar problemas de falta de cooperação dos entrevistados, que normalmente ocorrem em pesquisas por painéis fixos de domicílios. O procedimento de rotação foi definido de tal forma que cada domicílio permanece na amostra por 4 meses consecutivos, é retirado da amostra por 8 meses, retorna para mais 4 meses de pesquisa e,

então, é retirado definitivamente da amostra. Esse procedimento resulta em uma superposição de 75% da amostra a cada mês. Há também um processo de substituição dos setores que, por força do sistema de rotação da amostra, tiveram todos os seus domicílios incluídos na amostra.

O procedimento de estimação é semelhante ao da PNAD. É baseado no estimador de razão com ajuste para a população estimada para a data de referência. Para todas as estimativas de indicadores ou de totais calculados e divulgados a cada mês, são levadas em consideração apenas as informações da amostra daquele mês, ou seja, o aspecto longitudinal da amostra não é aproveitado na definição do estimador usado. Os coeficientes de variação são calculados para um conjunto selecionado de estimativas, usando o método do *ultimate cluster*.

Revisão 2002

A pesquisa foi submetida a uma revisão completa em 1982 e duas parciais, de vulto, em 1988 e 1993, por meio das quais foram realizados ajustamentos restritos somente ao plano de amostragem, principalmente no tamanho da amostra. Em 2001, passou por um amplo processo de revisão metodológica visando não só à captação mais abrangente das características de trabalho e das formas de inserção da mão-de-obra no mercado produtivo, como também à atualização da cobertura temática da pesquisa e sua adequação às mais recentes recomendações da Organização Internacional do Trabalho – OIT.

A revisão da PME incluiu o aprofundamento da investigação de temas já pesquisados, a adoção de instrumento eletrônico para coleta das informações, ajustes no plano de amostragem, seleção da amostra considerando a malha setorial do Censo Demográfico de 2000, alteração na cobertura geográfica, o uso de nova classificação de ocupação e de atividade e a estruturação de conceitos, definições e indicadores estabelecidos com base nas recomendações internacionais.

Na seleção das unidades primárias e secundárias de amostragem da PME revisada, implantada em 2001, foram adotadas a divisão territorial e a malha setorial vigentes em 1o de agosto de 2000 e utilizadas para a realização do Censo Demográfico 2000. Em síntese, os principais objetivos da revisão de 2002 da PME foram:

- Implementação de algumas mudanças conceituais no tema trabalho, seguindo recomendações internacionais;
- Ampliação da investigação para se ter melhor conhecimento da população economicamente ativa e da população disponível para o mercado de trabalho;
- Ajustamento no processo de rotação da amostra para dar consistência aos resultados de variação no tempo;
- Melhor operacionalização dos quesitos para captação das informações de forma a aprimorar a mensuração dos fenômenos; e,

- Introdução do uso do coletor eletrônico para a realização das operações de coleta visando a aprimorar o sistema operacional da pesquisa e a agilizar a apuração dos resultados.

A pesquisa revisada foi implantada em caráter experimental, em setembro de 2001, para viabilizar os ajustes que ainda eram necessários no novo sistema de apuração. De setembro de 2001 a dezembro de 2002, a pesquisa revisada foi realizada em paralelo à antiga com o objetivo de possibilitar a avaliação das alterações feitas e seus impactos. Em dezembro de 2002, o IBGE divulgou os resultados do período de outubro de 2001 a outubro de 2002 da nova pesquisa, acompanhada de um relatório de avaliação do impacto da nova metodologia nos resultados. Os resultados da PME revisada foram, então, disponibilizados a partir da pesquisa de março de 2002.

A Pesquisa Mensal de Emprego possui a priori aspectos de um plano amostral autoponderado dentro de cada região metropolitana. Isto implica que, em função do crescimento ou decréscimo natural do setor, verificado a cada realização da listagem, a quantidade de unidades domiciliares a serem selecionadas pode aumentar ou diminuir.

Na amostra original de setores, foi feita uma redução na probabilidade final de seleção de unidades domiciliares, nos casos em que se constatou baixa proporção de unidades domiciliares ocupadas em relação ao total de unidades domiciliares no setor. Além disso, como nesta pesquisa há tratamento para não resposta, a característica da autoponderação é perdida no momento da estimação de quantidades de interesse.

No momento em que a amostra foi dimensionada, em março de 2001, a sua composição foi a seguinte: 145 municípios, 2 029 setores, 37 212 domicílios.

O cálculo dos pesos para a expansão da amostra foi feito usando os pesos de desenho, ajustes para não resposta, ajustes para a inclusão do cadastro de novas construções e calibração para as estimativas independentes de população em cada região metropolitana, por meio de estimador de razão.

As estimativas dos coeficientes de variação são calculadas todo mês para todos os indicadores e são divulgadas na Internet. Além disso, são também obtidos resultados sobre a precisão da diferença de alguns indicadores em dois instantes distintos de tempo, baseados na metodologia que consiste em estimar intervalos de confiança para as diferenças temporais de um determinado conjunto de indicadores provenientes da pesquisa, para cada região metropolitana isoladamente e para o conjunto das seis. Detalhes sobre esta metodologia podem ser verificados em Lila e Freitas (2007).

Pesquisa de Orçamentos Familiares (POF)

Abrangência 11 áreas

O objetivo principal da pesquisa é obter informações sobre os padrões de gastos das famílias para a construção dos novos pesos para os itens que compõem os diversos índices

de preços que o IBGE calcula a cada mês. A primeira pesquisa sobre despesas familiares e renda no Brasil foi o ENDEF - Estudo Nacional da Despesa Familiar, realizada em 1974-75, que investigou também aspectos relativos à nutrição e antropometria, com plano amostral nos moldes do da PNAD. A Pesquisa de Orçamentos Familiares realizada em 1995-1996 foi planejada para representar a população urbana de 9 regiões metropolitanas, do Distrito Federal e do município de Goiânia. Essa pesquisa foi basicamente uma repetição da pesquisa realizada em 1987-88, com alguma atualização metodológica.

A coleta dos dados foi realizada no período de outubro/95 a setembro/96 de forma a captar os padrões sazonais de renda e despesas. As despesas foram obtidas através de entrevistas, usando diversos períodos de referência tais como: semana, mês, trimestre e semestre. Porém, as despesas com itens menos prováveis de serem lembrados exatamente pelos entrevistados (pequenas compras, despesas com alimentação e outras despesas de uso coletivo) foram obtidas através do registro diário em uma caderneta, feito preferencialmente pelo próprio informante, durante 7 dias, no caso da POF 1995-96 e de 14 dias no caso da POF 1987-88.

Cabe destacar vários aspectos inovadores que foram incorporados no planejamento da amostra da POF, já na sua primeira aplicação. A esse respeito, ver IBGE (1992), sobre a metodologia da POF 1987-88, e Bianchini e Vieira (1998), sobre a metodologia da POF 1995-96.

Em cada área da pesquisa, o plano amostral considera uma amostra em dois estágios de seleção com estratificação da unidade primária. A unidade primária de seleção, o setor censitário, foi estratificada em duas etapas: estratos geográficos (núcleo e periferia) e renda média do chefe do domicílio no setor. Os setores foram selecionados sistematicamente com probabilidade proporcional ao tamanho (medido em número de domicílios particulares ocupados). Em cada setor da amostra, os domicílios foram selecionados através de amostragem aleatória simples sem reposição. Pela primeira vez, em pesquisas do IBGE, a seleção de domicílios a pesquisar em cada setor foi feita sem empregar amostragem sistemática, com o sorteio aleatório efetuado por computador.

Para a POF 1995/96, o tamanho da amostra em cada área foi determinado a partir de uma precisão especificada para estimar a renda total do chefe ($CV=0,05$), com o número de domicílios a serem selecionados em cada setor da amostra fixado em 10, usando as informações investigadas no questionário básico do Censo Demográfico de 1991. A amostra de setores foi dividida em quatro subamostras, uma para cada trimestre de coleta. A alocação dos setores nas subamostras foi aleatória, preservando a estratificação adotada, de tal forma que todos os estratos estão representados em todos os trimestres.

Antes do início da coleta, em cada setor selecionado para a amostra, foi preparada uma listagem completa dos domicílios de forma a construir um cadastro atualizado para a seleção das unidades de segundo estágio. Entretanto, o número de domicílios selecionados em cada setor foi aumentado para 13 para compensar a seleção de domicílios vagos, fechados e possíveis recusas. Para setores com taxas de crescimento acima de certos níveis, o número de domicílios selecionados foi novamente aumentado, de acordo com

patamares de crescimento, de forma a reduzir a variância dos pesos, mas o número máximo de domicílios selecionados por setor foi 28. Uma novidade em relação à tradição das demais pesquisas domiciliares foi a eliminação do requisito da manutenção da autoponderação. A amostra da POF 1995-96 ficou com 19.816 domicílios selecionados.

O procedimento de estimação é baseado no estimador de razão com calibração⁵ na população residente em domicílios particulares urbanos, dada pela Contagem de População de 1996. Os pesos associados a cada domicílio da amostra foram obtidos usando o estimador natural derivado do plano amostral, com tratamento para não resposta, multiplicado por um fator que foi calculado independentemente para cada área da pesquisa. Esse fator é a razão entre a população residente em domicílios particulares permanentes urbanos, dada pela Contagem de População de 1996 e a estimativa da população correspondente proveniente da amostra. A data de referência da Contagem de População foi 01.08.96, que é próxima da data de referência definida para a POF, 15.09.96. A data de referência da POF foi usada para a correção da inflação, tornando todos os valores das despesas e receitas a preços constantes. Para o município da capital de cada Região Metropolitana, exceto Belém, foi calculado outro ajuste para calibrar a população do município, gerando um conjunto de pesos à parte para ser usado apenas na obtenção de estimativas no nível do município.

A precisão das estimativas amostrais foi medida pelo coeficiente de variação calculado para um conjunto de estimativas de total, usando o método do *ultimate cluster*. A publicação dos coeficientes de variação para variáveis categóricas foi feita usando a função generalizada de variância (*generalized variance functions*, Wolter, 1985). Ou seja, para cada domínio de publicação, foi ajustado um modelo de regressão para explicar os CVs das estimativas de total como função do valor das próprias estimativas. Os coeficientes do modelo de regressão ajustado foram publicados juntamente com as estimativas da pesquisa, bem como os próprios CVs para uma seleção de estimativas relativas a variáveis contínuas relacionadas com valores de despesas e de rendimentos.

Abrangência Nacional

A coleta dos dados da pesquisa foi realizada nas áreas urbanas e rurais em todo o território brasileiro no período de julho de 2002 a junho de 2003. A concepção do plano amostral adotado na pesquisa é basicamente a mesma que foi empregada na Pesquisa de Orçamentos Familiares 1995-1996. Como a POF 2002-2003 teve sua abrangência territorial ampliada para todo o território nacional⁶, o planejamento da amostra foi distinto no tocante a esta condição, além de utilizar outra variável para estratificação e dimensionamento da amostra.

⁵ Calibração é o processo pelo qual os fatores de expansão para a amostra são determinados de forma a buscar a consistência das estimativas a partir da amostra com os totais conhecidos da população.

⁶ Foram excluídas as áreas definidas pelo IBGE como sendo quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos e hospitais.

Assim, na POF 2002-2003, adotou-se um plano amostral conglomerado em dois estágios, com estratificação geográfica e estatística (a partir de variável que caracteriza os estratos socioeconômicos) das unidades primárias de amostragem que correspondem aos setores da base geográfica do Censo Demográfico 2000. As unidades secundárias de amostragem são os domicílios particulares permanentes. Os setores foram selecionados por amostragem sistemática com probabilidade proporcional ao número de domicílios no setor, ao passo que os domicílios foram selecionados por amostragem aleatória simples sem reposição, dentro dos setores selecionados.

Um dos aspectos específicos desta POF refere-se principalmente à estratificação, tanto geográfica como estatística. A estratificação geográfica teve por intuito espalhar geograficamente a amostra, garantindo a participação na amostra das diferentes partes do território brasileiro. Para a área urbana de cada Unidade da Federação, foram definidos os seguintes estratos geográficos: município da capital, região metropolitana sem o município da capital e restante da área urbana. Com o objetivo de permitir comparação com as pesquisas anteriores, foram consideradas as regiões metropolitanas pesquisadas na POF 1995-1996: Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre. Nas áreas rurais, em função dos altos custos de coleta, principalmente devidos a grandes deslocamentos, a estratificação não foi definida em cada Unidade da Federação. Assim, foram definidos cinco estratos rurais, um para cada Grande Região (Norte, Nordeste, Sudeste, Sul e Centro-Oeste). Ainda para efeito de comparação, nas regiões metropolitanas consideradas, foi mantida a estratificação geográfica das POFs passadas, de núcleo e periferia. O município da capital constituiu o núcleo, enquanto o restante da região metropolitana foi chamado de periferia. Para o caso do Distrito Federal, foram criadas duas subdivisões: uma contendo apenas a região administrativa de Brasília e outra contendo as demais regiões administrativas.

Diferentemente das POFs anteriores, os estratos estatísticos (socioeconômicos) da pesquisa foram definidos com base na variável "média de anos de estudo do responsável pelo domicílio no setor", investigada na Contagem da População 1996. Isso porque os dados sobre rendimento investigado no Censo Demográfico 2000 não estavam liberados à época do planejamento da POF 2002-2003.

O tamanho da amostra de setores foi determinado em função do tipo de estimador utilizado, do nível de precisão fixado para a estimativa da média de anos de estudo dos responsáveis pelo domicílio, obtido a partir da Contagem da População 1996 e do número esperado de domicílios com entrevista realizada em cada setor, em cada nível geográfico de controle da estimação. Foram identificados dois níveis geográficos de controle: área urbana de cada Unidade da Federação e área rural de cada Grande Região.

Foram fixados coeficientes de variação (CVs) de 3% para a estimativa da média de anos de estudo dos responsáveis pelos domicílios, para a área urbana de cada Unidade da Federação da Região Nordeste, Amazonas, Pará e Rondônia. O CV de 2% foi fixado para cada Unidade da Federação das Regiões Sudeste, Sul e Centro-Oeste. Acre, Amapá e Tocantins, da Região Norte, tiveram CV fixado em 4%. Nas áreas rurais das Regiões Norte

e Nordeste, foi fixado CV de 4% e, para cada uma das demais Grandes Regiões, o CV fixado foi de 3%.

A alocação do total de setores selecionados em cada estrato foi proporcional ao número total de domicílios particulares permanentes no estrato, com a condição de haver pelo menos dois setores na amostra de cada estrato. O número fixado de domicílios com entrevista por setor foi estabelecido de acordo com a área da pesquisa: 10 domicílios nos setores urbanos, 16 nos setores rurais das Regiões Norte e Nordeste, e 20 nos setores rurais das Regiões Sudeste, Sul e Centro-Oeste. O tamanho efetivo da amostra foi de 3 984 setores, correspondendo a um número esperado de 44 248 domicílios com entrevista.

A seleção dos setores foi feita independentemente em cada estrato, sistematicamente, e proporcional ao número de domicílios (ocupados e fechados) do setor da malha setorial de 2000. Selecionados os 3 984 setores da amostra, foi feita uma listagem, em campo, de todos os domicílios pertencentes a esses setores, com o objetivo de se obter um cadastro atualizado para proceder a seleção dos domicílios. Prevendo a perda de domicílios por entrevista não realizada na etapa de coleta de dados, optou-se por selecionar em cada setor um número maior de domicílios do que aquele estipulado durante o dimensionamento da amostra. Foi definido acréscimo baseado numa proporção esperada de entrevistas não realizadas, em vez de substituir domicílios. Estipulou-se em 25% essa proporção para compensar a não resposta, acarretando em 13 o número de domicílios selecionados por setor urbano na expectativa de se obter 10 entrevistas realizadas. Foi definido em 20 o número de domicílios selecionados por setor rural das Regiões Norte e Nordeste na expectativa de se obter 16 entrevistas realizadas. Nos setores rurais das Regiões Centro-Oeste, Sudeste e Sul, foi estabelecido em 25 na expectativa de 20 realizações de entrevistas.

Com base nas informações das listagens dos domicílios, foram identificados aqueles setores com altas taxas de crescimento em relação às informações do Censo Demográfico 2000 com elevadas taxas de domicílios vagos e ainda aqueles com grande ocorrência de domicílios de uso ocasional. Nesses setores foram feitos acréscimos maiores, no momento da seleção, até o limite de 28 domicílios para os setores das áreas urbanas e de 30 a 34 para os setores das áreas rurais, com a finalidade de compensar eventuais perdas na precisão do plano amostral. De posse do total de domicílios listados e do número de domicílios a serem efetivamente selecionados por setor, realizou-se a seleção aleatória sem reposição dos domicílios, independente em cada setor.

Os pesos amostrais foram inicialmente calculados com base no plano amostral efetivamente utilizado na seleção da amostra, incorporando ajustes para compensar a não resposta das unidades investigadas. Posteriormente, os pesos sofreram ajustes de calibração, procedimento que consistiu em obter, para cada Unidade da Federação (domínios de calibração), estimativas para o total de pessoas em determinados recortes iguais às respectivas projeções populacionais obtidas para 15 de janeiro de 2003. As variáveis utilizadas para estimar estes totais são denominadas variáveis de calibração. Foram definidas 20 variáveis de calibração, considerando o cruzamento de população por

sexo e classes de idade ou por situação do domicílio (urbana ou rural) e o procedimento foi desenvolvido utilizando-se o software estatístico *Generalized Estimation System* – GES.

Os estimadores de variância foram obtidos através de Linearização de Taylor e do Método do Conglomerado Primário (ver Hansen et al., 1953). Os erros amostrais foram avaliados através das estimativas dos coeficientes de variação (CVs), obtidos dividindo-se a variância estimada pela estimativa da quantidade de interesse. Na POF 2002-2003, optou-se por estimar os CVs de todas as estimativas do plano tabular de divulgação.

Em 2008-2009 o IBGE realizou a quarta POF, que além de manter o objetivo principal - orçamentos familiares -, que permite obter as estruturas de consumo das famílias e o valor total da conta família, também incluiu os temas nutrição e condições de vida, tendo um maior aprofundamento no tema nutrição. O modelo metodológico da pesquisa também atendeu novas demandas, sendo incluídas diversas variáveis relacionadas a Meio Ambiente, Turismo, Assistência à Saúde e Fecundidade. Nesta pesquisa, também foi incluída uma primeira experiência metodológica de investigação do consumo efetivo pessoal. Através de um novo questionário - Bloco de consumo pessoal individual (POF7), o consumo efetivo de alimentos e bebidas no domicílio e fora do domicílio foi investigado para uma subamostra de domicílios.

A POF 2008-2009 teve duração também de 12 meses e foi realizada de maio de 2008 a maio de 2009.

Em linhas gerais, o plano amostral adotado para a POF 2008-2009 é basicamente o mesmo que o implementado na POF 2002-2003, com destaque para a existência da Amostra Mestra de setores censitários, definida para as pesquisas domiciliares do IBGE, segundo metodologia descrita em Freitas et al. (2007).

A amostra mestra foi definida para permitir que várias subamostras possam ser selecionadas a partir de seu conjunto de setores censitários, selecionado a partir de um cadastro inicial contendo todos os setores censitários disponíveis à época do Censo Demográfico 2000. A amostra de setores censitários da POF 2008-2009 é uma das possíveis subamostras da amostra mestra.

Um dos aspectos específicos para a definição da amostra da POF refere-se principalmente à questão da estratificação dos setores censitários, não somente com relação à estratificação geográfica, mas também com relação à estratificação estatística. Para atender a essa característica das amostras das POFs, a metodologia de estratificação definida para a seleção da Amostra Mestra levou em consideração esses aspectos em sua construção. Mais detalhes sobre a definição da Amostra Mestra estão descritos nos itens específicos mais adiante neste documento.

Em resumo, a POF 2008-2009 adotou um plano amostral denominado como conglomerado em dois estágios, com estratificações geográfica e estatística das unidades primárias de amostragem que correspondem aos setores da base geográfica do Censo Demográfico 2000, a partir da estrutura oferecida pela Amostra Mestra desenhada pelo IBGE para o Sistema Integrado de Pesquisas Domiciliares. Os setores foram selecionados

por amostragem com probabilidade proporcional ao número de domicílios existentes no setor, dentro de cada estrato final da amostra mestra. A subamostra de setores para a POF 2008-2009 foi selecionada por amostragem aleatória simples em cada estrato. No plano adotado, as unidades secundárias de amostragem foram os domicílios particulares permanentes, que foram selecionados por amostragem aleatória simples sem reposição, dentro de cada um dos setores selecionados. Em seguida ao processo de seleção de setores e domicílios, os setores são distribuídos ao longo dos quatro trimestres da pesquisa, garantindo que em todos os trimestres, os estratos geográfico e socioeconômico estejam representados através dos domicílios selecionados.

O tamanho da amostra de setores foi determinado em função do tipo de estimador utilizado e do nível de precisão fixado para estimar o total dos rendimentos das pessoas moradoras responsáveis pelos domicílios, obtidos a partir dos dados do Censo Demográfico 2000, e, ainda, levando em consideração o número esperado de domicílios com entrevistas realizadas em cada setor, segundo cada domínio de estimação considerado. Foram identificados dois níveis geográficos de controle para o cálculo do tamanho da amostra, a saber: área urbana de cada Unidade da Federação e área rural de cada Grande Região.

Foram fixados diferentes coeficientes de variação para estimar com a precisão desejada o total da renda dos responsáveis pelos domicílios, segundo os diferentes domínios de estimação. Os níveis de precisão fixados foram estabelecidos a partir das análises realizadas com as precisões para a mesma variável e calculados a partir dos dados da POF 2002-2003. Para estimar o total nas Unidades da Federação da Região Norte, foram fixados coeficientes de variação que variaram de 10% a 15%. Para a Região Nordeste, os coeficientes fixados ficaram entre 5% e 10%. No caso das Regiões Sudeste e Sul, a variação dos coeficientes ficou entre 3% e 7%. Por último, na Região Centro-Oeste, os coeficientes de variação variaram entre 8% e 10%.

A alocação da amostra total de setores selecionados em cada estrato foi proporcional ao número total de domicílios particulares permanentes no estrato, com a condição de haver pelo menos três setores na amostra de cada estrato. Foi fixado o número de domicílios com entrevistas por setor de acordo com a área da pesquisa: 12 domicílios nos setores urbanos, 16 nos setores rurais. O tamanho efetivo da amostra foi de 4 696 setores, correspondendo a um número esperado de 59 548 domicílios com entrevista.

Os setores da amostra da POF 2008-2009 foram selecionados a partir dos setores da Amostra Mestra, por meio de amostragem aleatória simples. Foram então selecionados 4 696 setores para a amostra da pesquisa, de um total de 12 800 setores pertencentes à Amostra Mestra. Após o procedimento de seleção dos setores e a alocação desses setores nos quatro trimestres da pesquisa, deu-se início ao processo de atualização (operação de listagem) dos cadastros de endereços dos domicílios (cadastro de seleção). Nesta pesquisa, tal processo foi realizado em quatro etapas, contemplando cada um dos trimestres da pesquisa. A decisão por essa periodicidade trimestral para o processo de atualização foi tornar o cadastro o mais atualizado possível para a realização de cada etapa de seleção dos domicílios para compor a amostra a ser entrevistada a cada trimestre. O objetivo da atualização dos cadastros de seleção é o de minimizar a perda de entrevistas por motivos,

tais como: domicílio que já deixou de existir, domicílio em ruínas, etc. Mesmo com todos os cuidados com a atualização do cadastro de seleção, avaliou-se pela pertinência de se ampliar o número de entrevistas, prevendo eventuais perdas ao longo da fase de coleta das informações, por recusa do morador ou por não conseguir abrir o domicílio, por exemplo. Sendo assim, para esta pesquisa, estimou-se uma perda média de 15% das entrevistas e acréscimo de igual proporção foi atribuído ao total de domicílios a serem selecionados por setor da amostra. Em termos práticos, para compensar as perdas futuras com entrevistas não realizadas, foi selecionado, em cada setor urbano, um total de 13 domicílios, enquanto para os setores de situação rural foram selecionados 18 domicílios.

Com base nas informações das listagens dos domicílios, foram identificados aqueles setores com altas taxas de crescimento em relação às informações do Censo Demográfico 2000 e com elevadas taxas de domicílios fechados. Nesses setores foram feitos acréscimos maiores, no momento da seleção, até o limite de 28 domicílios para os setores das áreas urbana e rural, com a finalidade de compensar eventuais perdas na precisão das estimativas. De posse do total de domicílios listados e do número de domicílios a serem efetivamente selecionados por setor, realizou-se a seleção aleatória sem reposição dos domicílios, independente em cada setor.

Os setores de cada estrato foram aleatoriamente alocados por trimestre e seus domicílios espalhados ao longo do mesmo, para garantir a distribuição dos estratos da amostra ao longo dos 12 meses de duração da pesquisa.

Os pesos foram calculados, inicialmente, com base no plano de seleção efetivamente utilizado, incorporando ajustes para compensar a não resposta das unidades investigadas. Posteriormente, os pesos sofreram ajustes de pós-estratificação, procedimento que consistiu em obter para cada um dos pós-estratos definidos e segundo cada Unidade da Federação, estimativas para o total de pessoas que fossem equivalentes às respectivas projeções populacionais obtidas para 15 de janeiro de 2009.

Tal como na POF anterior, os estimadores de variância de totais e razões, quantidades estimadas nesta pesquisa, foram obtidos através de Linearização de Taylor e do Método do Conglomerado Primário (ver Hansen et al., 1953). Os erros amostrais foram avaliados através das estimativas dos coeficientes de variação, calculados para tabelas selecionadas.

A POF 2015-2016 está prevista para ser selecionada a partir da Amostra Mestra 2010, considerando as mesmas definições usadas na definição do plano amostral da POF 2008-2009, com pequenas modificações no tamanho da amostra.

Pesquisa sobre Padrões de Vida (PPV)

A Pesquisa sobre Padrões de Vida realizada em 1996-97 foi uma pesquisa piloto baseada no *Living Standard Measurement Study* (LSMS) estabelecido pelo Banco Mundial em 1980 para desenvolver métodos para a coleta e análise de dados sobre padrões de vida em países em desenvolvimento. As pesquisas do tipo LSMS foram realizadas em vários

países, com vários objetivos analíticos, tais como medir a distribuição do bem estar e o nível de pobreza dos domicílios, para entender como os domicílios reagem a programas econômicos e ambientais de governo, e permitir análises complexas da relação entre vários aspectos do bem estar do domicílio (ver Grosh e Muñoz, 1996 e Caillaux, 1998).

Essa pesquisa caracteriza-se por: “inclusão de temas socioeconômicos, estudados de forma integrada em um mesmo domicílio, um ano de permanência no campo (março de 1996 a março de 1997) de forma a captar fenômenos sazonais e, manutenção de um controle rigoroso tanto na aplicação do questionário como na entrada de dados e no processo de crítica das informações” (Caillaux, 1998).

Para atingir esses objetivos, foram considerados vários aspectos inovadores no planejamento da pesquisa. Os domicílios foram entrevistados duas vezes, com intervalo de duas semanas entre as entrevistas. O questionário foi quase todo pré-codificado de forma que os dados puderam ser digitados diretamente em microcomputadores imediatamente após a entrevista, de forma descentralizada no campo. O programa de entrada de dados realiza críticas de consistência e de validade das respostas. Os dados inconsistentes ou errados eram marcados pelo programa de tal forma que as respostas de certas questões puderam ser verificadas durante a segunda entrevista. Os informantes foram instruídos para registrarem suas despesas durante as duas semanas que antecederam a segunda visita.

A pesquisa foi planejada para investigar uma diversidade de temas sociais e econômicos, a saber: características do domicílio, características básicas demográficas, migração, saúde, mão-de-obra, fecundidade, rendimentos, investimentos e créditos, despesas com bens duráveis, despesas com alimentação, empreendimentos domiciliares, agricultura, avaliação do padrão de vida e antropometria.

O plano amostral inclui estratificação dos setores censitários (as UPAs - unidades primárias de amostragem) em 10 estratos geográficos: de Fortaleza, Recife, Salvador, o restante da área urbana da Região Nordeste, o restante da área rural da região Nordeste, as regiões metropolitanas de São Paulo, Rio de Janeiro, Belo Horizonte, o restante da área urbana da Região Sudeste, o restante da área rural da região Sudeste. Em cada estrato geográfico, foram definidos outros três estratos com base no rendimento médio domiciliar mensal do chefe do domicílio, usando os dados do Censo Demográfico de 1991. O tamanho da amostra em número de setores em cada estrato foi determinado por alocação proporcional. Os setores foram selecionados com probabilidade proporcional ao número de domicílios, com reposição, em cada um dos trinta estratos definidos. A amostra de setores foi dividida em quatro subamostras, uma para cada trimestre de coleta. A alocação da amostra nas subamostras foi aleatória e de tal forma a preservar a estratificação, ou seja, todos os estratos estão representados em todos os trimestres. Em cada setor selecionado para a amostra, foi realizada uma operação denominada listagem de domicílios, com o objetivo de obter um cadastro atualizado para a seleção das unidades de segundo estágio. A partir dessa listagem, os domicílios foram selecionados através de amostragem aleatória simples sem reposição. A operação de listagem foi realizada por partes, o mais próximo possível do início da coleta de cada trimestre. Nos setores pertencentes aos estratos de

região metropolitana ou de área urbana, foram selecionados 8 domicílios por setor. Nos estratos de setores rurais esse número foi fixado em 16.

O tamanho da amostra foi de 554 setores, sendo 278 na região Nordeste e 276 na região Sudeste, correspondendo a um tamanho total de amostra de 4 944 domicílios. Com o objetivo de compensar a não resposta por motivos de recusa, domicílios vagos ou fechados, foi selecionada uma segunda amostra de domicílios nos mesmos moldes da primeira. Os domicílios com não resposta foram então substituídos por domicílios provenientes dessa segunda amostra. Esse procedimento foi adotado em função do reduzido tamanho amostra, que tornava indesejável perder informação. Apesar disso, foram perdidos quatro questionários do primeiro trimestre de pesquisa. Essa perda foi compensada por reponderação dos domicílios informantes pelo inverso da taxa de resposta. Com o objetivo de evitar um aumento na taxa de não entrevistas devidas à recusa dos informantes, no processo de seleção da amostra de setores foram adotados procedimentos para evitar a seleção de setores já selecionados para a PNAD, a PME e a POF, uma vez que essas pesquisas foram realizadas na mesma época da PPV e possuem planos amostrais semelhantes.

A estimação de totais e dos erros amostrais associados foi realizada usando o estimador natural derivado do plano amostral adotado. A precisão das estimativas foi medida pelo coeficiente de variação calculado para todos os indicadores usados para a análise e disseminação dos resultados, usando o método do *ultimate cluster* (ver Albiéri e Bianchini, 1997).

Pesquisa de Economia Informal Urbana (ECINF)

Em 1997, o IBGE realizou sua primeira pesquisa domiciliar por amostragem em nível nacional para identificar atividades econômicas desenvolvidas nos domicílios ou em pequenas unidades produtivas, de forma a medir o papel e a dimensão dessas atividades na economia brasileira, através da identificação dos proprietários de negócios informais e da investigação das características de funcionamento das unidades produtivas (ver Jorge, 1995).

A população objetivo inclui as pessoas residentes na área urbana que trabalhavam por conta própria ou como empregadores com até cinco empregados, em pelo menos uma situação de trabalho de atividades não agrícolas. Os trabalhadores domésticos foram excluídos da população objetivo.

O objetivo da pesquisa foi produzir estimativas para cada um dos 26 estados, o Distrito Federal, cada uma das 10 regiões metropolitanas e o município de Goiânia.

Foram usados dois tipos de questionários na coleta dos dados: o primeiro para obter informação sobre as características dos domicílios e das pessoas moradoras, com o objetivo de identificar as pessoas engajadas em unidades produtivas do setor informal, através das características do trabalho; o segundo questionário foi usado para investigar as características das unidades produtivas do setor informal e seus proprietários.

O plano amostral considera dois estágios de seleção e estratificação das unidades de seleção. As unidades primárias de amostragem (UPAs) foram os setores urbanos estratificados primeiramente pela localização geográfica. Em cada estado, foram definidos dois ou três estratos geográficos dependendo se o estado possui ou não região metropolitana, o município da capital do estado, os demais municípios da região metropolitana, o restante do estado. Foi definido um segundo nível de estratificação das UPAs, dentro de cada estrato geográfico, de acordo com a renda média domiciliar do setor, obtida dos dados do Censo Demográfico de 1991. A seleção das unidades primárias em cada estrato foi feita através de amostragem sistemática com probabilidade proporcional ao tamanho (medida em número de domicílios ocupados). As unidades de segundo estágio foram os domicílios com moradores classificados como conta própria ou empregador com até cinco empregados. Os domicílios foram posteriormente classificados de acordo com o grupo de atividades e selecionados sistematicamente a partir da lista atualizada de domicílios realizada nos setores selecionados para a amostra.

Com o objetivo de evitar um aumento na taxa de não entrevistas devidas à recusa, no processo de seleção dos setores foi adotado um procedimento para eliminar as coincidências com os setores selecionados para a PNAD e PME, uma vez que essas duas pesquisas seriam realizadas na mesma época que a de economia informal.

O tamanho da amostra em cada área foi determinado a partir de um coeficiente de variação especificado para estimar o número de proprietários (conta próprias e empregadores) na economia informal (CV=5%, excepcionalmente, por motivos de custo, CV=6% para cada área da Região Norte). O número de domicílios a serem selecionados por setor foi fixado em 16 (ver Almeida e Bianchini, 1998).

No que se refere a aspectos de planejamento amostral, a pesquisa de economia informal difere das pesquisas domiciliares tradicionais do IBGE. Isto porque é preciso lidar com uma população rara, heterogênea e mais difícil de ser detectada. Todos esses fatores contribuem para aumentar a complexidade do plano amostral, da seleção da amostra, dos procedimentos de estimação e principalmente da preparação do cadastro de unidades amostrais de interesse, ou seja, da listagem de domicílios, a qual requer a realização de uma pesquisa em todos os domicílios do setor para identificar as atividades desenvolvidas pelos moradores de cada domicílio (ver Kalton e Anderson, 1986).

A listagem dos domicílios da ECINF foi uma operação com custo elevado, pois além de produzir uma lista completa de endereços das unidades domiciliares, envolveu a realização de entrevista para obter as informações necessárias para identificar a população objetivo, para obter as informações para a segunda estratificação por grupo de atividades que constituíram o objetivo da pesquisa.

Em 2003, as atividades foram classificadas utilizando a Classificação Nacional de Atividades Econômicas-CNAE-Domiciliar, que é uma adaptação da Classificação Nacional de Atividades-CNAE para as pesquisas domiciliares e foram agregadas nos seguintes grupos: Indústria de transformação e extrativa; construção civil; comércio e reparação; serviços de alojamento e alimentação; transporte, armazenagem e comunicações;

atividades imobiliárias, aluguéis e serviços prestados às empresas; educação, saúde e serviços sociais; outros serviços coletivos, sociais e pessoais; e, outras atividades.

Em cada setor da amostra, a alocação dos 16 domicílios foi feita proporcionalmente ao número de domicílios existentes em cada grupo de atividades no setor. Além disso, foram feitos alguns ajustes que ocasionaram um aumento médio de 30% no número de domicílios a serem selecionados por setor.

A estimação foi feita com base no estimador natural que deriva do plano amostral adotado, com tratamento para não resposta. Os domicílios listados como tendo proprietários do setor informal e que na entrevista não tinham mais essa característica foram excluídos na estimação de totais, mas receberam o valor zero para cada variável de interesse da investigação. Além dos pesos para estimação das características de proprietários de unidades produtivas do setor informal, foi também associado um peso à unidade produtiva do setor informal, que leva em conta o inverso do número de sócios da unidade produtiva. A precisão das estimativas foi medida pelos coeficientes de variação calculados para um conjunto selecionado de estimativas de total. O método do *ultimate cluster* foi usado para estimar a variância de cada variável.

A concepção do plano amostral adotado na pesquisa de 2003 é basicamente a mesma empregada na ECINF 1997. A estratificação por renda e as medidas de tamanho usadas durante a seleção foram baseadas nos dados do Censo 2000. Em 2003, a amostra da pesquisa teve 54 595 domicílios selecionados, em 2 499 setores censitários, enquanto que em 1997 foram 48 934 domicílios selecionados, em 2 340 setores.

Pesquisa Especial sobre Tabagismo (PETAB)

A Pesquisa de Tabagismo – PETAB é uma pesquisa suplementar que foi a campo junto com a PNAD 2008, utilizando, portanto, toda a estrutura amostral dessa pesquisa. Uma característica importante dessa investigação foi a necessidade de entrevistar a própria pessoa selecionada para a pesquisa, em contraposição à forma usual da PNAD, que admite que as informações sejam fornecidas por outra pessoa moradora no domicílio.

Em função disso, a PETAB foi realizada em uma subamostra da amostra de domicílios da PNAD e, em cada domicílio dessa subamostra, foi selecionada uma amostra de pessoas moradoras de 15 anos ou mais para responder às questões sobre tabagismo. Dessa forma, a amostra da PETAB é uma amostra probabilística de pessoas de 15 anos ou mais, obtida em quatro estágios de seleção, que toma por base a amostra da PNAD nos três primeiros estágios.

A Pesquisa Nacional por Amostra de Domicílios - PNAD é realizada por meio de uma amostra probabilística de domicílios obtida em três estágios de seleção: unidades primárias - municípios; unidades secundárias - setores censitários; e unidades terciárias - unidades domiciliares (domicílios particulares e unidades de habitação em domicílios coletivos). Na seleção das unidades primárias e secundárias (municípios e setores

censitários) da PNAD da primeira década deste século, foram adotadas a divisão territorial e a malha setorial vigentes em 1º de agosto de 2000 e utilizadas para a realização do Censo Demográfico 2000.

No terceiro estágio, a amostra da PETAB corresponde a 1/3 dos domicílios e das unidades de habitação em domicílios coletivos selecionados para a PNAD. A definição do tamanho da amostra foi feita considerando o objetivo de obter estimativas de proporção de pessoas com determinadas características relacionadas com o consumo de tabaco em nível nacional e em cada uma das Grandes Regiões.

No quarto estágio, em cada domicílio da amostra da PETAB, foi selecionado, com equiprobabilidade, um morador com 15 anos ou mais de idade para a investigação das características relacionadas ao fumo e ao tabaco.

O procedimento definido para a inclusão das áreas rurais de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá na PNAD, em 2004, foi mantido para a PETAB.

Na definição dos pesos de expansão da PETAB, também foram consideradas as seleções proporcionais de um terço no terceiro estágio de seleção e a seleção de um morador de 15 anos ou mais no quarto estágio de seleção, além de uma correção para os casos de não resposta deste morador. Adicionalmente, os pesos da PETAB foram ajustados para que as estimativas de população por sexo correspondessem aos totais populacionais por sexo da PNAD estimados a partir da amostra inteira em cada divisão geográfica definida acima.

A expansão da amostra da PNAD utiliza estimadores de razão cuja variável independente é a projeção da população residente de cada Unidade da Federação, segundo o tipo de área (região metropolitana de divulgação da pesquisa). Essas projeções consideram a evolução populacional ocorrida entre os Censos Demográficos sob hipóteses de crescimento associadas a taxas de fecundidade, mortalidade e migração. Os resultados da PETAB foram produzidos considerando os dados da revisão 2008 da projeção da população do Brasil como variável independente para a expansão da amostra. Além disso, para Rondônia, Acre, Amazonas, Roraima, Pará e Amapá, adotou-se a projeção da população residente, segundo a situação do domicílio (urbana e rural), como variável independente para expansão da amostra, em função da inclusão das áreas rurais na pesquisa em 2004.

A utilização do plano de amostragem da PNAD e da PETAB para estimar populações pequenas em números absolutos ou concentradas geograficamente, como pode ser o caso dos temas em questão, pode gerar estimativas com erros de amostragem elevados. Nesse sentido, visando a facilitar a avaliação da precisão das estimativas divulgadas, foram calculados os erros de amostragem, expressos pelos coeficientes de variação, para todas as estimativas (células) constantes do plano tabular de divulgação. Assim, para cada tabela de resultados apresentada na divulgação, foi preparada outra com os correspondentes coeficientes de variação estimados.

Pesquisa das Características Étnico-raciais da População (PCERP)

Realizada uma única vez, em 2008, a Pesquisa das Características Étnico-raciais da População - PCERP 2008 é uma pesquisa de objetivos múltiplos que visava compreender melhor o sistema de classificação da cor ou raça utilizado nas pesquisas domiciliares do IBGE, contribuindo para seu aprimoramento. Os objetivos específicos foram assim definidos:

- Ampliar o espectro de compreensão das categorias nas estatísticas oficiais em relação às questões étnico-raciais;
- Fornecer novos elementos de interpretação para possíveis alternativas de aprimoramento do sistema de classificação étnico-racial;
- Construir uma base empírica para subsidiar estudos e análises sobre o tema;
- Levantar as denominações correntes de cor, raça, etnia e origem de forma mais abrangente e completa, tanto do ponto de vista da composição étnica da população como das diversidades regionais;
- Identificar as dimensões que definem a construção e o uso dessa terminologia; e,
- Correlacionar os níveis de instrução e a posição na ocupação da população entrevistada com os dos pais, segundo os grupos de cor ou raça.

A população-alvo da PCERP 2008 é constituída pelos moradores de 15 ou mais anos de idade residentes em domicílios particulares permanentes pertencentes à área de abrangência geográfica da pesquisa. Esta área foi definida como sendo o recorte geográfico formado pelas seguintes Unidades da Federação: Amazonas, Paraíba, São Paulo, Rio Grande do Sul, Mato Grosso e Distrito Federal.

A principal finalidade do plano amostral para esta pesquisa foi permitir a obtenção de estimativas das proporções de pessoas em cada uma das cinco categorias de cor ou raça no conjunto das seis Unidades da Federação selecionadas. A pesquisa foi domiciliar e o plano amostral empregado foi amostragem conglomerada em três estágios, com estratificação das unidades primárias de amostragem. Os setores censitários formaram as unidades primárias de amostragem; os domicílios foram as unidades de segundo estágio; e os moradores de 15 anos ou mais de idade definiram as unidades de terceiro estágio.

Como parte integrante do SIPD, as unidades primárias de amostragem da pesquisa foram obtidas da Amostra Mestra, que é a estrutura amostral do sistema, composta por setores censitários selecionados com um método probabilístico e da qual é possível selecionar uma subamostra para as pesquisas que integram o sistema.

Portanto, a estratificação dos setores censitários (unidades primárias de amostragem) nesta pesquisa foi a mesma adotada para a Amostra Mestra da década de 2000, que está descrita de forma geral em item específico mais adiante neste documento e em detalhes em Freitas et al. (2007).

No segundo estágio, foi selecionado por amostragem aleatória simples um número fixo de domicílios particulares permanentes em cada setor selecionado no primeiro estágio.

Por fim, no terceiro estágio, em cada domicílio selecionado, um morador com 15 anos ou mais de idade foi selecionado para responder ao questionário, também por amostragem aleatória simples.

O tamanho da amostra foi determinado considerando o nível de precisão desejado para as estimativas das proporções de pessoas em cada uma das cinco categorias de cor ou raça. Vários cálculos baseados em amostragem aleatória simples foram efetuados a fim de definir o tamanho da amostra, fixando-se vários níveis de precisão para as estimativas das proporções e para os diversos níveis geográficos.

O plano amostral adotado na pesquisa não foi amostragem aleatória simples de pessoas, por isso foi feito um ajuste no tamanho da amostra obtido, considerando os valores do efeito de plano amostral (EPA). Esta medida, EPA, indica o quanto o plano amostral por conglomerados é menos eficiente (maior variância) que a amostragem aleatória simples, razão pela qual o tamanho da amostra obtido para amostragem aleatória simples é aumentado para que se alcance a mesma precisão. O EPA foi calculado levando-se em conta o número de domicílios/pessoas selecionado em cada setor, a variância do estimador por amostragem aleatória simples, e o coeficiente de correlação intraclasse, que mede a heterogeneidade dentro dos setores em relação à variável de interesse. Para maiores detalhes sobre a definição teórica do efeito de plano amostral, ver Lila e Freitas (2006).

Os dados utilizados nos cálculos foram obtidos no Censo 2000, tanto os referentes aos números de pessoas, total e por cor ou raça, quanto os referentes aos efeitos de plano amostral. Em resumo, o tamanho da amostra resultou em 1 511 setores, 15 110 domicílios e pessoas.

Os pesos amostrais ou fatores de expansão, utilizados para a obtenção de estimativas dos parâmetros relacionados com as características investigadas na PCERP 2008, foram calculados considerando as probabilidades de seleção, os ajustes pela não resposta, e os ajustes para calibração dos totais estimados. Para cada domicílio da amostra, foram calculados dois pesos: um para obtenção das estimativas das características gerais de todos os moradores e do domicílio, associado ao domicílio e a cada um de seus moradores; e outro para as características específicas sobre o tema da pesquisa das pessoas de 15 anos ou mais de idade, associado à pessoa selecionada no domicílio.

Observou-se que a não resposta da pessoa selecionada foi diferenciada por sexo e por faixa etária, levando à calibração das estimativas das pessoas de 15 anos ou mais de idade nestes domínios, obtidas com os dados da pessoa selecionada da PCERP 2008 pelos totais estimados pela POF 2008-2009.

As faixas etárias consideradas foram as seguintes:

- Para o Distrito Federal e o Estado da Paraíba: 15 a 24 anos; 25 a 34 anos; 35 a 44 anos; 45 a 54 anos; e 55 anos ou mais de idade; e,

- Para os Estados do Amazonas, São Paulo, Rio Grande do Sul e Mato Grosso: 15 a 24 anos; 25 a 34 anos; 35 a 44 anos; 45 a 54 anos; 55 a 64 anos; e 65 anos ou mais de idade.

A precisão das estimativas produzidas com os dados da PCERP 2008 é expressa pelo coeficiente de variação (CV) estimado, utilizando-se o Método do Conglomerado Primário (*ultimate cluster*), por meio do software SUDAAN (*Survey Data Analysis*).

O Sistema Integrado de Pesquisas Domiciliares - SIPD

Amostra Mestra para o SIPD - Base 2000

Com o objetivo de facilitar a integração entre as diferentes pesquisas domiciliares, de posse dos resultados do Censo Demográfico 2000, o IBGE iniciou o planejamento do Sistema Integrado de Pesquisas Domiciliares - SIPD, um amplo projeto que busca harmonizar conceitos e definições de variáveis comuns, procedimentos de coleta e listagem de setores censitários, dentre outros procedimentos fundamentais para a qualidade das pesquisas. Por serem de múltiplos propósitos e regulares, a PNAD Contínua e o sistema de Pesquisas de Orçamentos Familiares - POF foram definidas como as duas fontes de dados centrais do SIPD. Naturalmente, outros levantamentos irão integrar o sistema, de forma que as pesquisas nucleares serão veículos importantes para conduzir a investigação de temas suplementares aos seus questionários básicos. No entanto, algumas demandas poderão levar, também, à realização de pesquisas independentes, considerando, porém, a mesma infraestrutura do sistema.

Um ponto central dessa integração é a utilização de uma infraestrutura amostral comum, cadastro e amostra, definidos especialmente para o sistema. A amostra comum, denominada Amostra Mestra, é um conjunto de setores censitários, que são considerados unidades primárias de amostragem do plano amostral de cada uma das pesquisas do SIPD.

A descrição detalhada do plano amostral definido para a seleção da Amostra Mestra encontra-se em Freitas et al. 2007, de onde foram extraídas as descrições apresentadas de forma resumida a seguir.

A população alvo da Amostra Mestra inclui toda a população a ser investigada em todas as pesquisas e é constituída pelos moradores residentes em todos os domicílios dos setores censitários da Base Operacional Geográfica de 2000, que foi compatibilizada com a malha municipal de 2001, de todo o território nacional, totalizando 214 835 setores. Foram excluídas as áreas definidas pelo IBGE como sendo quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos e hospitais.

O plano amostral escolhido foi um desenho estratificado, com seleção das unidades primárias com probabilidade proporcional a uma medida de tamanho. A opção por este plano se justifica pelo fato das metodologias de todas as pesquisas domiciliares do IBGE

incluïrem planos amostrais que empregam algum tipo de estratificaço das UPAs e seleço destas unidades com probabilidade proporcional ao tamanho.

A estratificaço das UPAs, ou seja, dos setores censitrios, da Amostra Mestra foi definida levando-se em consideraço os objetivos das diversas pesquisas que sero contempladas por esta amostra e tambm as questes operacionais e domnios de divulgaço.

Alm da definiço de um tamanho mnimo para os estratos, foram consideradas as seguintes etapas de estratificaço: estratificaço por diviso administrativa; estratificaço geogrfica e espacial; estratificaço por situaço do setor; e, estratificaço estatstica, considerando a varivel "renda total dos responsveis pelos domiclios".

O tamanho da amostra foi determinado considerando a preciso desejada para a estimativa de um parmetro de interesse. O parmetro escolhido o "total de pessoas desocupadas no trimestre", um dos principais indicadores da PNAD Contnua. Para calcular o tamanho da amostra foi necessrio definir o esquema de rotaço da amostra de domiclios para a pesquisa contnua, trimestral, pois este esquema define o tamanho mnimo de amostra em cada estrato.

Aps diversos estudos, o esquema de rotaço definido para a PNAD Contnua foi o 1-2(5), que corresponde a 5 trimestres seguidos de entrevista de um mesmo domiclio. Como a pesquisa  trimestral, neste esquema o domiclio  entrevistado 1 ms e sai da amostra por 2 meses seguidos, sendo esta sequncia repetida 5 vezes. Com esse esquema,  possvel comparar as informaçes de uma mesma pessoa em anos consecutivos, pois h a sobreposiço esperada de 25% da amostra de um trimestre para o mesmo trimestre do ano seguinte.

A alocaço da amostra de setores nos estratos foi feita proporcionalmente ao nmero de domiclios. Ao final, o tamanho da Amostra Mestra 2000 ficou em 12.800 setores, para a obtenço de estimativas para um trimestre.

O mtodo de seleço da amostra escolhido foi Amostragem de Pareto PPT (ver Costa, 2007) combinado com a tcnica de nmeros aleatrios permanentes, que associa a cada setor um nmero aleatrio entre zero e um, que permanece atrelado a ele permanentemente, e o tamanho relativo do setor no grupo de rotaço, em nmero de domiclios particulares permanentes.

Amostra Mestra para o SIPD - Base 2010

Para a realizaço do Censo Demogrfico 2010, foi definida uma nova diviso do territrio nacional em setores censitrios. Assim, com a malha de setores atualizada e os dados do Censo Demogrfico 2010, a Amostra Mestra pode ser atualizada para a implantaço da Pesquisa Nacional por Amostra de Domiclios Contnua - PNAD Contnua, em 2011.

A definição do plano amostral da nova Amostra Mestra sofreu pequenas alterações em relação ao plano da amostra anterior, decorrentes de avaliação realizada com base nos testes da PNAD Contínua, realizados em 2009, e na realização da Pesquisa de Orçamentos Familiares de 2008-2009, cujas amostras foram selecionadas considerando a Amostra Mestra 2000.

A descrição detalhada do plano amostral definido para a seleção da Amostra Mestra 2010 encontra-se em Freitas e Antonaci (2014), de onde foram extraídas as descrições apresentadas de forma resumida a seguir.

A abrangência geográfica da amostra foi definida como todo o território nacional, dividido nos setores censitários da Base Operacional Geográfica de 2010, excluídas áreas com características especiais classificadas pelo IBGE como setores censitários de: aldeias indígenas, quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos, hospitais e agrovilas de projetos de assentamentos rurais. Para esta nova Amostra Mestra, foram excluídos, também, os setores censitários localizados em Terras Indígenas.

Para a definição das UPAs, foi estipulado um tamanho mínimo, em número de domicílios, para permitir a seleção de diversas amostras para as pesquisas do SIPD. Assim, setores censitários muito pequenos foram agregados com vizinhos para atender à restrição de tamanho mínimo, que foi definido em 60 domicílios particulares permanentes, incluindo os ocupados, os ocupados sem entrevista realizada e os vagos, dados provenientes do Censo 2010. A agregação foi feita com o objetivo de maximizar o número de grupos, juntando o mínimo possível, tendo como restrições a contiguidade, o tamanho mínimo e algumas características dos setores: tipo, situação e divisão administrativa.

Na área de abrangência definida para a Amostra Mestra 2010, o total de UPAs ficou em 292.067, correspondendo a 310.329 setores censitários em todo o país.

A estratificação das UPAs da nova Amostra Mestra foi definida com pequenas alterações em relação à amostra anterior, levando-se em consideração os objetivos das diversas pesquisas do Sistema, questões operacionais e domínios de divulgação. Entretanto, foram mantidas as etapas de estratificação, a saber: estratificação por divisão administrativa; estratificação geográfica e espacial; estratificação por situação do setor; e, estratificação estatística, considerando a variável "renda total dos domicílios".

Tal como na Amostra Mestra anterior, o tamanho da amostra foi calculado como o tamanho necessário para estimar o "total de pessoas desocupadas de 14 anos ou mais de idade", que é um dos principais indicadores da PNAD Contínua, com um nível de precisão pré-determinado, e o mesmo esquema de rotação de domicílios nas UPAs 1-2(5).

Foram mantidos, também, os esquemas de alocação da amostra de setores nos estratos (proporcionalmente ao número de domicílios) e de seleção da amostra (amostragem de Pareto PPT combinada com a técnica de números aleatórios permanentes).

Esta Amostra Mestra foi definida para ser usada ao longo da década de 2010, até que estejam disponíveis novas informações censitárias que permitam uma atualização completa e nova seleção de UPAs. Considerando isso, foi definido um esquema de rotação da amostra de UPAs.

O método escolhido de seleção da amostra permite a incorporação de atualizações no cadastro de seleção, acompanhando a evolução do crescimento das UPAs e mudanças na base operacional geográfica, além de permitir a renovação controlada da amostra. O esquema de rotação de setores da Amostra Mestra ao longo da década está descrito no item da PNAD Contínua, a seguir.

Assim, o tamanho da Amostra Mestra 2010 ficou em 15.096 UPAs (setores ou agregados de setores), para a obtenção de estimativas para um trimestre.

Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua)

A PNAD Contínua visa produzir informações básicas para o estudo do desenvolvimento socioeconômico do País e permitir a investigação contínua de indicadores sobre trabalho e rendimento. Produz resultados trimestralmente para o tema trabalho e rendimento e anualmente para outros temas sociodemográficos. Como pesquisa integrante do SIPD, utiliza a infraestrutura amostral comum a todas as pesquisas, a Amostra Mestra 2010, que corresponde a um plano amostral conglomerado em dois estágios de seleção com estratificação das unidades primárias de amostragem (UPAs). Por ser a pesquisa do sistema que demanda a maior amostra, a definição do tamanho da amostra de UPAs da pesquisa corresponde à definição do tamanho da Amostra Mestra para um trimestre.

No segundo estágio é selecionado um número fixo de domicílios particulares permanentes ocupados dentro de cada UPA da amostra, por amostragem aleatória simples do Cadastro Nacional de Endereços para Fins Estatísticos - CNEFE atualizado.

O tamanho da amostra de domicílios dentro de cada UPA foi fixado em 14, pois em estudos realizados, foi observado que o aumento neste número não acarretaria uma redução considerável no tamanho da amostra de UPAs, pelo fato de haver grande homogeneidade entre os domicílios de uma mesma UPA (ver Freitas e Lila, 2004).

A PNAD Contínua foi planejada para ter periodicidade de coleta trimestral, ou seja, a amostra total de domicílios é coletada em um período de 3 meses, para ao final deste ciclo serem produzidas as estimativas dos indicadores desejados. Um dos principais interesses em pesquisas contínuas que acompanham mercado de trabalho é a inferência a respeito de mudanças no comportamento dos indicadores, considerando o período de divulgação definido. Nestas situações a amostra é planejada de tal forma que haja rotação dos domicílios selecionados, mantendo uma parcela sobreposta entre dois períodos de divulgação subsequentes, tal como descrito nos itens sobre a Amostra Mestra.

Conforme citado no item anterior sobre a Amostra Mestra para o SIPD - Base 2010, além da rotação da amostra de domicílio, foi definido que a cada trimestre serão trocadas

2,5% das UPAs, o que resultará em uma substituição quase completa da amostra em 10 anos. Não havendo mudanças no cadastro de seleção ao longo do tempo, seria possível saber quais UPAS estariam entrando e saindo da amostra, em todos os trimestres, no momento da primeira seleção. Como é comum mudanças, especialmente nos tamanhos dos setores, acréscimo ou decréscimo no número de domicílios (medida de tamanho adotada na seleção), a incorporação destas mudanças a cada seleção proporciona melhorias nas estimativas, porém o conhecimento das UPAs na amostra seria obtido com pouca antecedência. A ideia é incorporar as atualizações uma vez por ano, incorporando-as ao cadastro para selecionar a Amostra Mestra correspondente aos 4 trimestres do ano seguinte.

Em relação à seleção dos domicílios para as diversas pesquisas do SIPD, foram definidos procedimentos para evitar que um mesmo domicílio participe de duas pesquisas num curto espaço de tempo, causando o cansaço do informante. Assim, um domicílio só será selecionado para a amostra de uma pesquisa se ele estiver há pelo menos 1 ano sem participar da amostra de qualquer pesquisa. E também foi definido que um domicílio selecionado para a amostra da PNAD Contínua só poderá voltar para esta mesma pesquisa quando for selecionada nova Amostra Mestra, o que está previsto para ocorrer após 10 anos.

Para auxiliar no controle da seleção das amostras, está sendo implantado um mecanismo que associa a cada domicílio um código único que será sua identidade, que não será alterado mesmo quando houver alterações e atualizações no CNEFE. Essa atualização do cadastro é feita a cada trimestre, nas UPAs que participam da rotação da amostra de domicílios da PNAD Contínua, cerca de 20% da amostra.

Na PNAD Contínua, a seleção de domicílios para a amostra é feita bem próxima da atualização do cadastro. Por isso, participam da seleção apenas as unidades classificadas como domicílios particulares permanentes ocupados (abertos ou fechados).

Na PNAD Contínua, que faz uso da Amostra Mestra 2010, o tamanho da amostra é de 211.344 domicílios, nas 15.096 UPAs (setores ou agregados de setores).

O cálculo dos pesos para a expansão da amostra e obtenção das estimativas na PNAD Contínua foi feito considerando as probabilidades de seleção de cada unidade (UPA e domicílio), os ajustes por não resposta e por calibração, tanto para as estimativas que serão produzidas trimestralmente, quanto para aquelas que serão calculadas com base na amostra anual e para as quais são definidos pesos específicos. A calibração usada até o momento considera a variável total da população em determinados níveis geográficos, provenientes das estimativas de população elaboradas pela Coordenação de População e Indicadores Sociais (COPIS) da Diretoria de Pesquisas, tendo como referência o ponto central do trimestre ou do ano. Entretanto, estuda-se a possibilidade de incluir as variáveis grupos de idade e sexo na calibração, para melhorar a qualidade das estimativas.

Para avaliar a precisão das estimativas, as estimativas de variância são obtidas usando o método do *ultimate cluster*, também conhecido como método do Conglomerado Primário (Cochran 1977, p.307). Como as quantidades de interesse na PNAD Contínua são

obtidas através de estimadores de razão, não existe uma fórmula exata para as variâncias destas quantidades. Para estimá-las é utilizada uma aproximação da variância de uma razão através de Linearização de Taylor. Este método consiste em definir uma variável linearizada z e aproximar a variância da estatística amostral de interesse pela variância dos valores z sob o plano amostral utilizado para a seleção da amostra. Para maiores detalhes, ver Lila e Freitas (2007).

Pesquisa Nacional de Saúde 2013 (PNS)

A Pesquisa Nacional de Saúde (PNS) é uma pesquisa domiciliar, de âmbito nacional, realizada em parceria com o Ministério da Saúde em 2013. Tem como objetivo caracterizar a situação de saúde e os estilos de vida da população, bem como a atenção à sua saúde, quanto ao acesso e uso dos serviços, às ações preventivas, à continuidade dos cuidados e ao financiamento da assistência. Antes disso, houve outras investigações sobre o tema Saúde, como Suplemento da PNAD, em 2003 e em 2008, que seguiram integralmente o plano amostral da PNAD correspondente.

A PNS é uma pesquisa isolada no âmbito do SIPD - Sistema Integrado de Pesquisas Domiciliares e a amostra da PNS constitui uma subamostra da Amostra Mestra atualizada com dados de 2010, definida para o Sistema, conforme descrito anteriormente.

Tal como a abrangência da Amostra Mestra, a população pesquisada na PNS corresponde aos moradores de domicílios particulares do Brasil, exceto os setores censitários localizados em Terras Indígenas e os setores censitários especiais (quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos e hospitais).

O questionário tem três partes, sendo que as duas primeiras são respondidas por um residente do domicílio e abrangem perguntas sobre as características desse domicílio e a situação socioeconômica e de saúde de todos os moradores. A terceira parte é, de fato, um questionário individual respondido por um morador adulto de 18 anos ou mais, sobre morbidade e estilos de vida. Além disso, para este indivíduo foram feitas aferições de peso, altura, circunferência da cintura e pressão arterial e exames laboratoriais para caracterizar o perfil lipídico, o nível de glicemia no sangue e determinar o teor de sódio na urina. Outra característica específica da PNS é que os exames laboratoriais foram feitos em uma subamostra de 25% dos setores censitários selecionados para a amostra da PNS.

Resumindo, a amostra da PNS foi selecionada em três estágios: 1º- seleção da subamostra de UPA em cada estrato da Amostra Mestra com probabilidade proporcional ao tamanho, medido pelo número de domicílios particulares permanentes - DPP em cada unidade; 2º - seleção por amostragem aleatória simples de domicílios em cada UPA selecionada no primeiro estágio; 3º - seleção por amostragem aleatória simples de um adulto com 18 ou mais anos de idade entre todos os moradores adultos do domicílio, para responder a um questionário individual.

Para calcular o tamanho de amostra da PNS necessário para a estimação de parâmetros de interesse em níveis diferentes de desagregação geográfica, foram considerados os seguintes aspectos: estimação de proporções com nível de precisão desejado em intervalos de 95% de confiança; efeito do plano de amostragem (EPA), por se tratar de amostragem por conglomeração em múltiplos estágios; número de domicílios selecionados por UPA; proporção de domicílios com pessoas na faixa etária de interesse. Diferentemente do previsto para as pesquisas do Sistema, na Pesquisa Nacional de Saúde, as unidades classificadas como domicílios particulares permanentes vagos também entraram na seleção de domicílios, em função da grande distância da atualização do cadastro de algumas UPAs, permitindo, assim, que unidades que viessem a ser ocupadas tivessem chances de seleção para a pesquisa.

A amostra da PNS permite a estimação dos principais indicadores no nível de UF e capital, mas alguns indicadores de interesse poderão ser também divulgados em menores níveis de desagregação geográfica: capital, restante da região metropolitana, e restante da UF. Em cada desagregação geográfica de divulgação de indicadores, o tamanho de amostra é de, pelo menos, 900 domicílios.

O tamanho total da amostra para a PNS foi definido em 81.254 domicílios, considerando uma taxa de não resposta de 20%. De fato, foram obtidas entrevistas para 64.348 domicílios, distribuídas em 6.069 UPAs.

Para os exames laboratoriais, foi selecionada uma subamostra de cerca de 25% das UPAs selecionadas e 10 domicílios por UPA, uma pessoa por domicílio, o que resultou em 1.727 UPAs e 17.902 pessoas selecionadas. Para facilitar a logística da coleta de material biológico, os setores censitários foram selecionados com probabilidade inversamente proporcional à dificuldade de coleta. Para o estabelecimento da dificuldade de coleta por meio de parâmetro numérico, foram identificados todos os municípios com 80 mil habitantes ou mais em todas as Unidades da Federação (UF). Em cada UF, foram calculadas as distâncias entre todos os municípios com menos de 80 mil habitantes selecionados na amostra e os municípios de grande porte populacional (80 mil ou mais habitantes) daquela UF, por meio das coordenadas geográficas dos centroides dos municípios. Entretanto, a coleta de material para os exames laboratoriais enfrentou muitas dificuldades em campo e não estava encerrada até o momento do fechamento deste documento.

Pesquisa por Amostragem na Coleta de Dados do Censo Demográfico

O Censo Demográfico de 1960 foi o primeiro a utilizar amostragem na coleta de dados relativos a um conjunto selecionado de características de pessoas, famílias e domicílios. Foram usados dois tipos de questionários: um questionário pequeno aplicado a todos os domicílios e seus moradores, não selecionados para a amostra (chamado questionário básico); e um questionário longo (chamado questionário da amostra) aplicado a todos os domicílios selecionados para a amostra, bem como seus moradores. Nos Censos de 1960, 1970 e 1980, foi utilizada uma única fração amostral de 25% dos domicílios. Em

1991, uma revisão amostral com grande impacto foi o emprego de duas frações amostrais diferentes de acordo com o tamanho do município, medido em função da projeção de população para a data de referência do Censo: 20% para os municípios com até 15.000 habitantes e 10% para os demais municípios. No Censo 2000, foram usadas as mesmas frações amostrais que em 1991.

No Censo 2010, para a parte do levantamento pesquisada por amostragem, foram aplicadas cinco frações de amostragem, considerando os tamanhos dos municípios em termos da população estimada em 1º de julho de 2009, tal como apresentada a seguir: 50% para os municípios com até 2 500 habitantes; 33% para os municípios com mais de 2 500 e até 8 000 habitantes; 20%, para os municípios com mais de 8.000 e até 20 000 habitantes; 10% para os municípios com mais de 20 000 e até 500.000 habitantes; e 5% para os municípios com população superior a 500.000 habitantes.

Em 2000 e em 2010, o procedimento usado para a determinação dos pesos para a expansão da amostra foi o mesmo do de 1991, porém simplificado, passando o ajuste do modelo de regressão a ser realizado em apenas uma etapa na área de ponderação.

Para a realização dos censos, o território nacional é dividido em partições geográficas denominadas setores censitários, de tal forma que seus limites respeitam as divisões internas dos municípios em zonas urbanas e rurais, e em distritos e subdistritos, caso existam. O setor censitário foi planejado de forma a que um entrevistador consiga realizar a operação de coleta no período de realização do Censo. Em 1991, foram definidos 163.266 setores censitários (que, no plano amostral, foram considerados como estratos). Em 2000, eram 215 811 setores e em 2010 foram definidos 310 120 setores com população.

Até o Censo Demográfico 2000, os domicílios particulares foram selecionados por amostragem sistemática em cada setor censitário. As famílias ou pessoas sós moradoras em domicílios coletivos (alojamentos estudantis, quartéis, prisões, hospitais, orfanatos, conventos, etc.) foram selecionadas, também de forma sistemática, independentemente da seleção de domicílios particulares, usando a mesma fração amostral definida para o setor a que pertence cada domicílio coletivo.

O procedimento de estimação de totais foi aplicado em cada área de ponderação separadamente. Uma área de ponderação é um conjunto de setores censitários que constituem a menor área geográfica para a qual as estimativas provenientes da amostra são avaliadas em termos de precisão. Em 1991, o método de Mínimos Quadrados Generalizados em duas etapas (*Generalized Least Squares Estimation Procedure* - GLSEP, Bankier et al., 1992), denominado no IBGE por MQG2, foi usado na determinação dos pesos e as variáveis auxiliares utilizadas foram definidas dentre aquelas investigadas para 100% da população, no próprio Censo Demográfico. Esse procedimento de estimação de regressão atribui um único peso fracionário a cada domicílio e a cada um de seus moradores, sendo importante destacar essas duas situações novas em relação aos censos anteriores: o peso fracionário e único para domicílios, famílias e pessoas.

Outra alteração foi introduzida na definição das áreas de ponderação. Para os Censos 2000 e 2010, foram usados métodos e sistemas automáticos de formação de áreas

de ponderação que conjugam critérios tais como: tamanho, para permitir estimativas com qualidade estatística em áreas pequenas; contiguidade, no sentido de serem constituídas por conjuntos de setores limítrofes com sentido geográfico; e homogeneidade em relação a um conjunto de características populacionais e de infraestrutura conhecidas. Além disso, alguns municípios grandes definiram as áreas de ponderação em função de suas áreas de planejamento municipal.

No Censo Demográfico 2010, a coleta de dados foi realizada usando computadores de mão, o que facilitou a implementação de novo esquema de seleção de domicílios para a amostra, por meio de um programa de seleção aleatória por grupo de domicílios, de acordo com a fração amostral definida para o setor.

No Censo 2000, o tamanho da amostra de domicílios foi 5.304.711, correspondendo a 20.274.412 pessoas. Em 2010, esses tamanhos foram: 6.192.332 domicílios, correspondendo a 20.635.472 pessoas.

Assim, por exemplo, em um setor com fração amostral de 20%, em vez de selecionar 1 em cada 5 domicílios, o algoritmo prevê a seleção de 2 em cada 10, aleatoriamente no grupo formado. Dessa forma, não existe um salto constante entre os números de ordem de unidades selecionadas em sequência, dificultando a introdução de vícios operacionais por parte dos entrevistadores.

Outra novidade introduzida no Censo 2010 foi o tratamento dados aos domicílios ocupados sem entrevista realizada, equivalente aos domicílios fechados nas pesquisas domiciliares. Os domicílios classificados como fechados são aqueles que sabidamente possuíam moradores na data de referência, mas que não tiveram entrevista realizada para o preenchimento das informações do questionário, independentemente do motivo da não realização da entrevista. A classificação de um domicílio na categoria de fechado é equivalente a considerá-lo como uma não resposta, que é um dos erros não amostrais mais comuns na realização de uma pesquisa, seja ela censitária ou por amostragem.

Visando estimar a parcela da população moradora nos domicílios fechados em cada um dos municípios brasileiros, foi aplicado um procedimento de imputação de domicílios e seus moradores. O pressuposto básico do procedimento de imputação é que a perda de dados seja aleatória, e se não for, que o padrão de não resposta seja conhecido ou pelo menos estimado, para ser considerado durante o tratamento da não resposta por imputação. Assim, no Censo Demográfico 2010, para estimar o número de moradores em domicílios fechados para cada município, definiu-se cada domicílio fechado como uma não resposta cujo atributo necessário é o número de moradores. O tratamento adotado para essa não resposta foi um procedimento de imputação por meio de seleção aleatória de um domicílio doador entre um conjunto de possíveis doadores em estratos específicos.

Em termos operacionais, o procedimento consistiu em imputar, para cada município, tantos domicílios quantos os classificados como fechados, com número de moradores de acordo com a distribuição obtida pelo conjunto de domicílios do estrato correspondente. E o total de moradores estimados no conjunto de domicílios fechados de cada município foi obtido pela soma dos moradores nos domicílios imputados.

Esse procedimento foi adotado para o conjunto de dados investigados censitariamente, chamado Conjunto Universo do Censo Demográfico. No caso dos resultados estimados provenientes da investigação por amostragem, o tratamento desse tipo de não resposta foi feito por meio do peso de expansão, uma vez que em cada área de ponderação foi considerado o tamanho efetivo da amostra de domicílios no ajuste do método MQG.

Resumo e conclusões

As tabelas do anexo apresentam de forma resumida algumas características gerais dos planos amostrais, os tamanhos das amostras, as distribuições da amostra por tipo de entrevista e as taxas de resposta para as várias pesquisas domiciliares, excetuando o Censo Demográfico.

As taxas de resposta representam a qualidade das pesquisas, pelo menos sob a ótica da aceitação por parte dos informantes, e são consideradas satisfatórias, em função da complexidade de cada pesquisa.

O IBGE vem realizando várias pesquisas domiciliares com diferentes níveis de complexidade em função das diferenças nos objetivos, na população-alvo, na abrangência geográfica, no nível de precisão e na periodicidade definidas para cada uma. Apesar das muitas semelhanças entre os planos amostrais das pesquisas citadas, as diferenças apontadas também revelam mudanças e aperfeiçoamentos metodológicos que foram possíveis de serem implantados ao tempo em que cada pesquisa foi planejada ou realizada.

Essas modificações (aperfeiçoamentos) incluem: a incorporação de estratificação de acordo com o nível de renda do setor; a previsão de perda de unidades da amostra por não resposta; o tratamento da não resposta; a redução do número de domicílios selecionados por setor; a redução do número de estágios de seleção; a eliminação do requisito de autoponderação; o esquema de seleção aleatória e não mais sistemática; o momento da atualização de setores (antiga operação de listagem para atualização do cadastro de domicílios), que passou mais próximo do momento da seleção de domicílios para a pesquisa; a utilização de subamostras de setores, de domicílios ou de pessoas para investigações específicas em uma dada pesquisa; a divulgação de erros amostrais associados a cada estimativa do plano tabular de divulgação; a adoção da amostra mestra para o conjunto de pesquisas domiciliares do SIPD.

O trabalho de revisão que está documentado na publicação de 2002 serviu de base para as discussões realizadas com vistas à reformulação da PME - Pesquisa Mensal de Emprego e da PNAD - Pesquisa Nacional por Amostra de Domicílios. Na ocasião, o desafio apontado era propor uma reformulação dos planos amostrais das pesquisas de forma a reduzir custos através de amostras menores conjugado com novos e aperfeiçoados procedimentos de estimação.

As discussões culminaram com a definição da Amostra Mestra para o SIPD - Sistema Integrado de Pesquisas Domiciliares, que começou a ser implantado com a

realização da PNAD Contínua, a partir do quarto trimestre de 2011. Além disso, outras pesquisas já tiveram sua amostra baseada nas Amostras Mestras: a Pesquisa das Características Étnico-raciais da População 2008 e a Pesquisa de Orçamentos Familiares 2008-2009, que foram definidas a partir da Amostra Mestra de 2000; a Pesquisa Nacional de Saúde 2013 e a Pesquisa de Orçamentos Familiares 2015-2016, em fase de planejamento, ambas com amostra baseada na Amostra Mestra de 2010.

Como aperfeiçoamentos possíveis, ainda objeto de estudos, vale destacar: avaliação do uso da variável população por grupo de idade e sexo no método de calibração dos pesos para expansão; avaliação da possibilidade de incorporar efeitos espaciais ou de vizinhança na estratificação dos setores; e, introdução de estimadores que se beneficiem da estrutura longitudinal da amostra, no caso de pesquisas contínuas.

Referências

- ALBIERI, S. e BIANCHINI, Z.M. *Aspectos de amostragem relativos à pesquisa domiciliar sobre padrões de vida*. Rio de Janeiro: IBGE, Departamento de Metodologia, 1997. 14p.
- ALMEIDA, R.A.P.; BIANCHINI, Z.M. *Sampling aspects of the 1997 Brazilian survey of the urban informal sector*. Proceedings of the Joint IASS/IAOS Conference. 1998.
- ALMEIDA, R.A.P.; BIANCHINI, Z.M. *Aspectos de amostragem da Pesquisa de Economia Informal Urbana 97*. Rio de Janeiro: IBGE, 1998. 32p. (Texto para Discussão, nº89).
- BANKIER, M.D.; RATHWELL, S.; MAJKOWSKI, M. *Two step generalized least squares estimations in 1991 Canadian Census*. Ottawa: Statistics Canada, Methodology Branch Working Paper. 1992.
- BIANCHINI, Z.M. e ALBIERI, S. *Principais Aspectos de Amostragem das Pesquisas Domiciliares do IBGE - revisão 2002*. Rio de Janeiro: IBGE, 27p. (Textos para Discussão, nº 8).
- BIANCHINI, Z.M. e ALBIERI, S. Uma Revisão dos Principais Aspectos dos Planos Amostrais das Pesquisas Domiciliares Realizadas pelo IBGE, *Revista Brasileira de Estatística*, v.60, n. 213, p.7-23, jan./jun. 1999.
- BIANCHINI, Z.M. e ALBIERI, S. Principales Aspectos de Muestreo de las Encuestas de Hogares del IBGE - revisión 2002. 17p. 2002. [Artigo apresentado no MECOVI - 10º Taller Regional sobre La Practica del Muestreo para el Diseño de las Encuestas de Hogares, 27 a 29 de noviembre, Buenos Aires, Argentina].
- BIANCHINI, Z.M.; VIEIRA, M. *Aspectos de amostragem da Pesquisa de Orçamentos Familiares 1995-1996*. Rio de Janeiro: IBGE, 1998. 106p. (Texto para Discussão, nº93).
- CAILLAUX, E.L. *Living standard survey 1996-1997*. Rio de Janeiro: IBGE. 1998. 10p. [Apresentado no Meeting of the Expert Group on Poverty Statistics (Rio Group), 13-15 Maio, 1998].
- COCHRAN, W.G. *Sampling Techniques* (3rd ed.). New York: Wiley. 1977.
- CARACTERÍSTICAS Étnico-raciais da População. Um estudo das categorias de classificação de cor ou raça. 2008. Brasil. Rio de Janeiro: IBGE, 95p. 2011.
- CORRÊA, S.T. Modelos Lineares Hierárquicos em Pesquisas por Amostragem Relacionando o Índice de Massa Corporal às Variáveis da Pesquisa sobre Padrões de Vida do IBGE. Rio de Janeiro: IBGE/ENCE. 2001. [Dissertação de Mestrado].

- CORRÊA, S.T.; SILVA, P.L.N. *Estimação de Variância para o Estimador da Diferença entre duas Taxas na Pesquisa Mensal de Emprego*. Rio de Janeiro: IBGE, Departamento de Metodologia, 2001. 26p.
- COSTA, G.T.L. *Coordenação de amostras PPT em Pesquisas Repetidas, utilizando o método de amostragem de Pareto*. Rio de Janeiro: IBGE/ENCE. 2007. [Dissertação de Mestrado].
- CRUZ, M.M. *Estimação de variâncias para séries dessazonalizadas pelo método X-12-Arima, considerando o desenho amostral*. Rio de Janeiro: IBGE/ENCE. 2002. [Dissertação de Mestrado].
- ECONOMIA Informal Urbana. Rio de Janeiro: IBGE. (Série Relatórios Metodológicos, vol.35). 87p. 2006.
- FREITAS, M.P.S. *Estratificação para a amostra de uma pesquisa domiciliar sobre mercado de trabalho*. Rio de Janeiro: IBGE/ENCE. 2002. [Dissertação de Mestrado].
- FREITAS, M.P.S.; LILA, M. F.; AZEVEDO, R.V. e ANTONACI, G.A. *Amostra Mestra para o Sistema Integrado de Pesquisas Domiciliares - Rio de Janeiro: IBGE, 2007*. 67p. (Texto para Discussão, nº23).
- FREITAS, M.P.S. e ANTONACI, G.A. *Sistema Integrado de Pesquisas Domiciliares. Amostra Mestra 2010 e Amostra da PNAD Contínua*. Rio de Janeiro: IBGE, 2014. 36p. (Texto para Discussão, nº50).
- GROSH, M. E.; MUÑOZ, J. *Manual for planning and implementing the LSMS Survey*. Poverty e Human Resources Division, Policy Research Department, The World Bank. 1996.
- HANSEN, M.H., HURWITZ, W.N. e MADOW, W.G. *Sample Survey Methods and Theory*, Vol. I e II. New York: Wiley. 1953.
- JORGE, A. The survey of the urban informal economy in Brazil. *Proceedings of the International Seminar on Informal Sector Employment Statistics*, 1995. pp. 239-256.
- JORGE, A.F. *Pesquisa de Economia Informal Urbana*. Rio de Janeiro: IBGE, 1996. 17p. [Artigo apresentado no Encontro Nacional de Produtores e Usuários de Informações Sociais, Econômicas e Territoriais].
- KALTON, G. e ANDERSON, D.W. Sampling rare populations. *The Journal of the Royal Statistical Society A*, 149, part 1, pp 65-82. 1986.
- KISH, L. *Survey sampling*. New York: Wiley. 1965.
- LILA, M.F. e FREITAS, M.P.S. *Estimação de Intervalos de Confiança para Estimadores de Diferenças Temporais na Pesquisa Mensal de Emprego*. Rio de Janeiro: IBGE, 2007. 101p. (Texto para Discussão, nº22).
- METODOLOGIA da Pesquisa Nacional por Amostra de Domicílios na década de 80. Rio de Janeiro: IBGE (Série Relatórios Metodológicos, vol.1). 1983.
- METODOLOGIA da Pesquisa Mensal de Emprego 1980. Rio de Janeiro: IBGE. (Série Relatórios Metodológicos, vol.2). 1983.
- METODOLOGIA do Censo Demográfico 2010. Rio de Janeiro: IBGE (Série Relatórios Metodológicos, vol.41). 703p. 2013.
- PESQUISA Mensal de Emprego. Rio de Janeiro: IBGE (Série Relatórios Metodológicos, vol.23). 72p. 2002.
- PESQUISA Mensal de Emprego. Rio de Janeiro: IBGE (Série Relatórios Metodológicos, vol.23. 2ª Edição). 89p. 2007.
- PESQUISA Mensal de Emprego. Outubro 2001 - Dezembro 2002. Rio de Janeiro: IBGE (Estatísticas básicas: séries retrospectivas, número 10). 299p. 2003.
- PESQUISA Nacional por Amostra de Domicílios. Volume. 25. Brasil. Rio de Janeiro: IBGE, 117p. 2004.

- PESQUISA Nacional por Amostra de Domicílios: Tabagismo 2008. Brasil. Rio de Janeiro: IBGE, 128p. 2009.
- PESQUISA Orçamentos Familiares. Volume 3. Aspectos de Amostragem. Rio de Janeiro: IBGE, 218p. (Série Relatórios Metodológicos, vol. 10). 1992.
- PESQUISA Orçamentos Familiares 2002-2003. Primeiros Resultados. Brasil e Grandes Regiões. 2ª Edição. Rio de Janeiro: IBGE, 270p. 2004.
- PESQUISA Orçamentos Familiares 2008-2009. Despesas, rendimentos e condições de vida. Rio de Janeiro: IBGE, 222p. 2010.
- PESSOA, D.G.C.; SILVA, P.L.N. *Análise de dados amostrais complexos*. In: Simpósio Nacional de Probabilidade e Estatística, 13.,1998, Caxambu. *Anais...* São Paulo: Associação Brasileira de Estatística, 1998.
- SILVA, P.L.N. *Planejamento, estimação e análise de dados em pesquisas por amostragem: desvendando a realidade brasileira com o "telescópio da estatística"*. Rio de Janeiro: IBGE, 1996. 28p. [Artigo apresentado no Encontro Nacional de Produtores e Usuários de Informações Sociais, Econômicas e Territoriais].
- SISTEMA Integrado de Pesquisas Domiciliares - SIPD. Rio de Janeiro: IBGE, 2007, 80p. (Texto para Discussão, nº 24).
- SZWARCWALD, C.L. et al. Pesquisa Nacional de Saúde no Brasil: concepção e metodologia de aplicação. *Ciência & Saúde Coletiva*, 19(2):333-342, 2014.
- WOLTER, K. M. *Introduction to Variance Estimation*. New York: Springer-Verlag. 1985.

Tabela 1 - Características gerais dos planos amostrais das várias pesquisas

(continua)			
<i>Pesquisa</i>	<i>Abrangência Geográfica</i>	<i>Número e definição dos Estágios de seleção</i>	<i>Tipo de estratificação das UPAs</i>
PNAD Até 2003	Nacional (exceto o Norte rural)	3 (municípios, setores e domicílios)	Geográfica
2004 em diante	Nacional		
PME	6 regiões metropolitanas	2 (setores e domicílios)	Geográfica
POF 1995/96	11 áreas urbanas	2 (setores e domicílios)	Geográfica e classes de renda do chefe/responsável
2002/03	Nacional		
2008/09	Nacional		
PPV	Regiões Nordeste e Sudeste	2 (setores e domicílios)	Geográfica e classes de renda do chefe
ECINF	Nacional, somente áreas urbanas	2 (setores e domicílios)	Geográfica e classes de renda domiciliar
PETAB	Nacional	4 (municípios, setores, domicílios, pessoas de 15 anos ou mais de idade)	Geográfica (PNAD 2008)
PCERP	6 Unidades da Federação	3 (setores, domicílios, pessoas de 15 anos ou mais de idade)	Geográfica e Classes de renda total dos responsáveis pelo domicílio (Amostra Mestra 2000)
Amostra Mestra 2000	Nacional	1 (setores)	Geográfica e Classes de renda total dos responsáveis pelo domicílio
2010	Nacional	1 (UPAs=setores ou agregados)	Geográfica e Classes de renda total dos domicílios
PNAD Contínua	Nacional	2 (UPAs e domicílios)	Geográfica e Classes de renda total dos domicílios (Amostra Mestra 2010)
PNS	Nacional	3 (UPAs, domicílios, pessoas de 18 anos ou mais de idade)	Geográfica e Classes de renda total dos domicílios (Amostra Mestra 2010)

Tabela 1 - Características gerais dos planos amostrais das várias pesquisas

(continuação)

<i>Pesquisa</i>	<i>Auto-ponderação</i>	<i>Seleção de setores</i>	<i>Variável usada como medida de tamanho</i>	<i>Cadastro de seleção de domicílios</i>	<i>Seleção de domicílios</i>
PNAD	Sim	Sistemática com PPT	Número de domicílios	Listagem + Novas construções	Sistemática simples
PME	Sim	Sistemática com PPT	número de domicílios particulares ocupados	Listagem + Novas construções	Sistemática simples
POF	Não	Sistemática com PPT	Número de domicílios particulares ocupados	Listagem	Aleatória simples sem reposição
PPV	Não	PPT com reposição	Número de domicílios particulares	Listagem	Aleatória simples sem reposição
ECINF	Não	Sistemática com PPT	Número de domicílios ocupados	Listagem com entrevista para identificar população objetivo	Sistemática simples
PETAB	Não	Sistemática com PPT	Número de domicílios	Listagem + Novas construções	Sistemática simples
PCERP	Não	Pareto PPT	Número de domicílios particulares permanentes	Atualização de setores via CNEFE	Aleatória simples
Amostra Mestra	---	Pareto PPT	Número de domicílios particulares permanentes	Atualização de setores via CNEFE	---
PNAD Contínua	Não	Pareto PPT	Número de domicílios particulares permanentes	Atualização de setores via CNEFE	Aleatória simples
PNS	Não	Pareto PPT	Número de domicílios particulares permanentes	Atualização de setores via CNEFE	Aleatória simples

Tabela 1 - Características gerais dos planos amostrais das várias pesquisas

<i>Pesquisa</i>	<i>Periodicidade</i>	<i>Período de coleta</i>	(conclusão) <i>Aspectos longitudinais</i>
PNAD	Anual (desde 1967, exceto em anos de Censo Demográfico)	Três meses	Nova seleção da amostra de domicílios (3º estágio)
PME	Mensal (desde 1980)	Um mês	Rotação da amostra de setores e domicílios
POF	1974/75 (ENDEF) 1987/88 1995/96 2002/03 2008/09	Um ano	Nova seleção da amostra a cada execução da pesquisa
PPV	1996/97 (pesquisa piloto)	Um ano	---
ECINF	1997 2003	Três meses	Nova seleção da amostra a cada execução da pesquisa
PETAB	---	Três meses	---
PCERP	---	Cinco meses	---
Amostra Mestra	Decenal (desde 2010)	---	Rotação da amostra de UPAs
PNAD Contínua	Trimestral (desde 2012)	Três meses	Rotação da amostra de UPAs e de domicílios
PNS	Quinquenal (a partir de 2013)	Três meses	Nova seleção da amostra a cada execução da pesquisa

Tabela 2 – Tamanho da amostra e distribuição por tipo de entrevista das várias pesquisas

<i>Pesquisa</i>	<i>Número de</i>	<i>Número de domicílios na amostra</i>			
	<i>setores selecionados</i>	<i>Selecionados</i>	<i>Média por setor</i>	<i>Eleitos (**)</i>	<i>Entrevistados</i>
PNAD 2013	9 166	148 697	16,2	124 031	116 543
PME Agosto 2014	2 851	44 512	15,6	37 390	33 458
POF 2008/09	4 696	68 373	13 (urbanos) e 18 (rurais)	62 238	55 970
PPV 96/97 (*)	554	4 944	8 (urbanos) e 16 (rurais)	4 944	4 940
ECINF 2003	2 499	54 595	21,8	42 506	41 333
PETAB 2008	7 818	51 011	6,5	39 847	39 425
PCERP 2008	1 499	14 990	10	14 150	12 860
Amostra Mestra 2000	12 800	--	--	--	--
2010 (***)	15 096	--	--	--	--
PNAD Contínua 1ºTrim/2015 (***)	15 095	211 170	14	194 664	186 355
PNS 2013(***)	6 069	81 254	13,4	69 994	64 348

(*) Foi adotado o procedimento de substituição controlada de domicílios durante a coleta.

(**) Os domicílios eleitos foram definidos de acordo com a população alvo de cada pesquisa.

(***) O dado refere-se ao número de UPAs em vez de número de setores.

-- Não se aplica dado numérico.

Tabela 3 - Taxas de resposta das várias pesquisas

<i>Pesquisa</i>	<i>Taxas</i>		
	<i>Entrevistados / selecionados (%)</i>	<i>Eleitos / selecionados (%)</i>	<i>Entrevistados / eleitos (%)</i>
PNAD 2013	78,4	83,4	94,0
PME Agosto 2014	75,2	84,0	89,5
POF 2008/09	81,9	91,0	89,9
PPV 1996/97 ^(*)	99,9	100,0	99,9
ECINF 2003	75,7	77,9	97,2
PETAB 2008	77,3	78,1	98,9
PCERP 2008	85,8	94,4	90,9
PNAD Contínua	88,2	92,2	95,7
PNS 2013	79,2	86,1	91,9

^(*) Foi adotado o procedimento de substituição controlada de domicílios durante a coleta.

Tabela 4 - Procedimentos de estimação das pesquisas por amostra

<i>Pesquisa</i>	<i>Estimador</i>	<i>Variável de calibração</i>	<i>Tratamento de não resposta</i>	<i>Publicação de erros amostrais</i>
PNAD 2013	Razão	Projeção de População	Não	Ajuste de modelo de regressão para variáveis categóricas
PME	Razão	Projeção de População	Sim	Todas as estimativas
POF 1995/96	Razão	Contagem de População de 1996	Sim	Seleção de variáveis contínuas e ajuste para variáveis categóricas
2002/03	Regressão	Projeção de População para 20 variáveis categóricas	Sim	Todas as estimativas do plano tabular
2008/09	Razão	Projeção de População	Sim	Tabelas selecionadas
PPV 1996/97	Natural do desenho	Não	Sim	Calculados para todos os indicadores divulgados
ECINF 1997 e 2003	Natural do desenho	Não	Sim	Seleção de variáveis
PETAB 2008	Razão	População por sexo estimada pela PNAD e Projeção de População	Sim	Todas as estimativas do plano tabular
PCERP 2008	Razão	População estimada pela POF 2008/09	Sim	Todas as estimativas do plano tabular
PNAD Contínua 2012 em diante	Razão	Projeção de População	Sim	Todas as estimativas do plano tabular
PNS 2013	Razão	Projeção de População	Sim	Todas as estimativas do plano tabular

Textos para Discussão já publicados

Antiga série

- 📖 Pesquisas Contínuas da Indústria - Vol. 1, nº 1, janeiro 1988
- 📖 Pesquisas Agropecuárias Contínuas: Metodologia - Vol. I, nº 2, 1988
- 📖 Uma Filosofia de Trabalho: As experiências com o SNIPC e com o SINAPI - Vol. I, nº 3, março 1988
- 📖 O Sigilo das Informações Estatísticas: Idéias para reflexão - Vol. I, nº 4, abril 1988
- 📖 Projeções da População Residente e do Número de Domicílios Particulares Ocupados: 1985-2020 - Vol. I, nº 5, maio 1988
- 📖 Classificação de Atividades e Produtos, Matérias-Primas e Serviços Industriais: Indústria Extrativa Mineral e de Transformação - Vol. 1, nº 6, agosto 1988
- 📖 A Mortalidade Infantil no Brasil nos Anos 80 - Vol. I, nº 7, setembro 1988
- 📖 Principais Características das Pesquisas Econômicas, Sociais e Demográficas - Vol. I, número especial, outubro 1988
- 📖 Ensaio sobre o Produto Real da Agropecuária - Vol. I, nº 9, setembro 1988
- 📖 Novo Sistema de Contas Nacionais, Ano Base 1980 - Resultados Provisórios - Vol. I, nº 10, dezembro 1988
- 📖 Pesquisa de Orçamentos Familiares - Metodologia para Obtenção das Informações de Campo - nº 11, janeiro 1989
- 📖 De Camponesa a Bóia-fria: Transformações do trabalho feminino - nº 12, fevereiro 1989
- 📖 Pesquisas Especiais do Departamento de Agropecuária - Metodologia e Resultados - nº 13, fevereiro 1989
- 📖 Brasil - Matriz de Insumo-Produto - 1980 - nº 14, maio 1989
- 📖 As Informações sobre Fecundidade, Mortalidade e Anticoncepção nas PNADs - nº 15, maio 1989
- 📖 As Estatísticas Agropecuárias e a III Conferência Nacional de Estatística - nº 16, junho 1989
- 📖 Brasil - Sistema de Contas Nacionais Consolidadas - nº 17, agosto 1989
- 📖 Brasil - Produto Interno Bruto Real Trimestral - Metodologia - nº 18, agosto 1989
- 📖 Estatísticas e Indicadores Sociais para a Década de 90 - nº 19, setembro 1989
- 📖 Uma Análise do Cotidiano da Pesquisa no DEREN (As Estatísticas do Trabalho) - nº 20, outubro 1989
- 📖 Coordenação Estatística Nacional - Reflexões sobre o caso Brasileiro - nº 21, novembro 1989
- 📖 Pesquisa Industrial Anual 1982/84 - Análise dos Resultados - nº 22, novembro 1989
- 📖 O Departamento de Comércio e Serviços e a III Conferência Nacional de Estatística - nº 23, dezembro 1989
- 📖 Um projeto de Integração para as Estatísticas Industriais - nº 24, dezembro 1989
- 📖 Cadastro de Informantes de Pesquisas Econômicas - nº 25, janeiro 1990
- 📖 Ensaio sobre a Produção de Estatística - nº 26, janeiro 1990
- 📖 O Espaço das Pequenas Unidades Produtivas: Uma tentativa de delimitação - nº 27, fevereiro 1990
- 📖 Uma Nova Metodologia para Correção Automática no Censo Demográfico Brasileiro: Experimentação e primeiros resultados - nº 28, fevereiro 1990
- 📖 Notas Técnicas sobre o Planejamento de Testes e Pesquisas Experimentais - nº 29, março 1990

- 📖 Estatísticas, Estudos e Análises Demográficas - Uma visão do Departamento de População - nº 30, abril 1990
- 📖 Crítica de Equações de Fechamento de Empresas no Censo Econômico de 1985 - nº 31, maio 1990
- 📖 Efeito de Conglomeração da Malha Setorial do Censo Demográfico de 1980 - nº 32, maio 1990
- 📖 A Redução da Amostra e a Utilização de Duas Frações Amostrais no Censo Demográfico de 1990 - nº 33, junho 1990
- 📖 Estudos e Pesquisas de Avaliação de Censos Demográficos - 1970 a 1990 - nº 34, julho 1990
- 📖 A Influência da Migração no Mercado de Trabalho das Capitais do Centro-Oeste - 1980 - nº 35, agosto 1990
- 📖 Pesquisas de Conjuntura: Discussão sobre Variáveis a Investigar - nº 36, setembro 1990
- 📖 Um Modelo para Estimar o Nível e o Padrão da Fecundidade por Idade com Base em Parturições Observadas - nº 37, outubro 1990
- 📖 A Estrutura Operacional de Uma Pesquisa por Amostra - nº 38, novembro 1990
- 📖 Produção Agrícola, Agroindustrial e de Máquinas e Insumos Agrícolas no Anos 80: Novas Evidências Estatísticas - nº 39, dezembro 1990
- 📖 A Inflação Medida pelo Índice de Preços ao Consumidor - nº 40, janeiro 1991
- 📖 A Participação Política Eleitoral no Brasil - 1988, Análise Preliminar - nº 41, fevereiro 1991
- 📖 Ensaios sobre Estatísticas do Setor Produtivo - nº 42, março 1991
- 📖 A Produção Integrada de Estatística e as Contas Nacionais: Agenda para Formulação de um Novo Plano Geral de Informações Estatísticas e Geográficas - nº 43, março 1991
- 📖 Matriz de Fluxos Migratórios Intermunicipais - Brasil - 1980 - nº 44, abril 1991
- 📖 Fluxos Migratórios Intrametropolitanos - Brasil - 1970-1980 - nº 45, abril 1991
- 📖 A Revisão da PNAD, A Questão Conceitual e Relatório das Contribuições - nº 46, maio 1991
- 📖 A Dimensão Ambiental no Sistema de Contas Nacionais - nº 47, maio 1991
- 📖 Estrutura das Contas Nacionais Brasileiras - nº 48, junho 1991
- 📖 Mercado do Couro e Resultados da Pesquisa Anual do Couro - nº 49, junho 1991
- 📖 As Estatísticas e o Meio Ambiente - nº 50, julho 1991
- 📖 Novo Sistema de Contas Nacionais Séries Correntes: 1981-85 Metodologia, Resultados Provisórios e Avaliação do Projeto - nº 51, julho 1991 (2 Volumes: Volume 1 - Metodologia, Resultados Provisórios e Avaliação do Projeto; Volume 2-Tabelas)
- 📖 O Censo Industrial de 1985 - Balanço da Experiência - nº 52, agosto 1991
- 📖 Análise da Inflação Medida Pelo INPC 1989 - nº 53, agosto 1991
- 📖 Revisão da PNAD: A Questão Amostral: Módulo II do Anteprojeto - nº 54, setembro 1991
- 📖 A Força de Trabalho e os Setores de Atividade - Uma Análise da Região Metropolitana de São Paulo - 1986-1990 - nº 55, outubro 1991
- 📖 Revisão da PNAD: Apuração das Informações: Módulo III do Anteprojeto - nº 56, novembro 1991
- 📖 Novos Usos para Pesquisa Industrial Mensal: A Evolução dos Salários Industriais, O Desempenho da Pecuária - nº 57, novembro 1991
- 📖 Revisão da PNAD: A Disseminação das Informações Módulo IV do Anteprojeto - nº 58, dezembro 1991
- 📖 Estatísticas Agropecuárias: Sugestões para o Novo Plano Geral de Informações - nº 59, dezembro 1991
- 📖 Análise Conjuntural e Pesquisa Industrial - nº 60, janeiro 1992
- 📖 Exploração dos Dados da Pesquisa Industrial Mensal de Dados Gerais - nº 61, fevereiro 1992

- 📖 Uma Proposta de Metodologia para a Expansão da Amostra do Censo Demográfico de 1991 - **nº 62**, outubro 1993
- 📖 Expansão da Fronteira e Progresso Técnico no Crescimento Agrícola Recente - **nº 63**, novembro 1993
- 📖 Avaliação das Condições de Habitação com Base nos Dados da PNAD - **nº 64**, setembro 1993
- 📖 Análise da Taxa de Desemprego Feminino no Brasil – **nº 65**, dezembro 1993
- 📖 Aspectos da Metropolização Brasileira: Comentários sobre os Resultados Preliminares do Censo Demográfico de 1991- **nº 66**, janeiro 1994
- 📖 Estimativas Preliminares de Fecundidade Considerando os Censos Demográficos, Pesquisas por amostragem e o Registro Civil - **nº 67**, janeiro 1994
- 📖 Apuração de Dados no IBGE: Problemas e Perspectivas - **nº 68**, fevereiro 1994
- 📖 Limeira - SP: Estimativas de Fecundidade e Mortalidade 1980/1988 - **nº 69**, março 1994
- 📖 Desemprego - Uma Abordagem Conceitual - **nº 70**, abril 1994
- 📖 Apuração dos Dados Investigados no Questionário Básico (CD 1.01) do Censo Demográfico de 1991 - **nº 71**, outubro de 1994
- 📖 Deslocamento Populacional e Segregação Sócio-Espacial – Migrantes Originários do Rio de Janeiro - **nº 72**, novembro de 1994
- 📖 Projeção Preliminar da População do Brasil para o Período 1980-2020 - **nº 73**, dezembro de 1994
- 📖 Considerações Preliminares Sobre a Migração Internacional no Brasil - **nº 74**, janeiro de 1995
- 📖 Estatísticas Agropecuárias Censitárias no Âmbito do Mercosul - Brasil, Argentina e Uruguai - **nº 75**, julho de 1995
- 📖 Projeções Preliminares das Populações das Grandes Regiões para o Período 1991-2010 - **nº 76**, agosto de 1995
- 📖 Dinâmica da Estrutura Familiar no Sudeste Metropolitano, Chefia Feminina e Indicadores Sócio-Demográficos: Um exercício exploratório utilizando modelo da regressão múltipla - **nº 77**, setembro de 1995
- 📖 O Uso das Matrizes de Insumo-Produto e Matrizes de Inovação para Medir Mudanças Técnicas - **nº 78**, outubro de 1995
- 📖 Estimativas dos Fatores de Correção para o Registro de Nascimentos Utilizando Registros tardios a nível de Brasil, Grandes Regiões, Unidades da Federação e Regiões Metropolitanas 1974/1994 - **nº 79**, abril de 1996
- 📖 Aspectos de Amostragem Relativos ao Censo Cadastro de 1995 - **nº 80**, junho de 1996
- 📖 Tendências Populacionais no Brasil e Pressão Sobre o Mercado de Trabalho Futuro - **nº 81**, setembro de 1996
- 📖 Transformações Estruturais e Sistemas Estatísticos Nacionais - **nº 82**, setembro de 1996
- 📖 Metodologias para o Cálculo de Coeficientes Técnicos Diretos em um Modelo de Insumo-Produto - **nº 83**, outubro de 1996
- 📖 Avaliação da Cobertura da Coleta do Censo Demográfico de 1991 - **nº 84**, outubro de 1996
- 📖 Componentes da Dinâmica Demográfica Brasileira: Textos Selecionados - **nº 85**, novembro de 1996
- 📖 Apuração dos Dados Investigados pelo Questionário da Amostra - CD 1.02 do Censo Demográfico de 1991 - **nº 86**, dezembro de 1996
- 📖 Estudo Preliminar da Evolução dos Nascimentos, Casamentos e Óbitos 1974-1990 - **nº 87**, janeiro de 1997

- 📖 Sistema de Contas Nacionais - Tabelas de Recursos e Usos - Metodologia - nº 88, dezembro de 1997
- 📖 Aspectos de Amostragem da Pesquisa de Economia Informal Urbana 97 - nº 89, junho de 1998
- 📖 Comparações da Renda Investigada nos Questionários do Censo Demográfico de 1991 - nº 90, julho de 1998
- 📖 Uma Revisão dos Principais Aspectos dos Planos Amostrais das Pesquisas Domiciliares Realizadas pelo IBGE - nº 91, setembro de 1998
- 📖 Planejamento Amostral para as Pesquisas Anuais da Indústria e do Comércio - nº 92, outubro de 1998
- 📖 Aspectos de Amostragem da Pesquisa de Orçamentos Familiares 1995-1996 - nº 93, dezembro de 1998
- 📖 Reflexões sobre um Programa de Estatísticas Ambientais - nº 94, abril de 1999
- 📖 O Comportamento das Importações e Exportações Brasileiras com Base no Sistema de Contas Nacionais 1980 - 1997 (versão preliminar) - nº 95, maio de 1999
- 📖 Meio Ambiente: sua integração nos sistemas de informações estatísticas - nº 96, maio de 1999
- 📖 Conta da Terra: considerações sobre sua realização no Brasil - nº 97, dezembro de 1999

Textos para discussão - nova série

- 📖 **Número 1** - Sistema integrado de contas econômico-ambientais - SICEA: síntese e reflexões / Sandra De Carlo. - Rio de Janeiro : IBGE, Departamento de Contas Nacionais, 2000.
- 📖 **Número 2** - Aspectos da produção de informação estatística oficial no contexto da sociedade atual: algumas questões teórico-metodológicas / Rosa Maria Porcaro - Rio de Janeiro : IBGE, Departamento de Metodologia, 2000
- 📖 **Número 3** - A Cor denominada: um estudo do suplemento da Pesquisa Mensal de Emprego de julho/98 / José Luis Petruccelli. - Rio de Janeiro : IBGE, Departamento de População e Indicadores Sociais, 2000.
- 📖 **Número 4** - Indicadores para a agropecuária - Rio de Janeiro : IBGE, Departamento de Agropecuária, 2001.
- 📖 **Número 5** - Estudos para definição da amostra da Pesquisa Industrial Mensal de Emprego e Salário / Ana Maria Lima de Farias. - Rio de Janeiro : IBGE, Departamento de Indústria, 2001.
- 📖 **Número 6** - A declaração de cor/raça no censo 2000: um estudo comparativo / José Luis Petruccelli. - Rio de Janeiro : IBGE, Departamento de População e Indicadores Sociais, 2002..
- 📖 **Número 7** - Dimensões preliminares da responsabilidade feminina pelos domicílios: um estudo do fenômeno a partir dos censos demográficos 1991 e 2000 / Sonia Oliveira, Ana Lucia Sabóia, Bárbara Cobo - Rio de Janeiro : IBGE, Departamento de População e Indicadores Sociais, 2002.
- 📖 **Número 8** - Principais Aspectos de Amostragem das Pesquisas Domiciliares do IBGE - revisão 2002 / Zélia Magalhães Bianchini e Sônia Albieri - Rio de Janeiro : IBGE, Departamento de Metodologia, 2003.
- 📖 **Número 9** - Censo Demográfico 2000 - Resultados da Pesquisa de Avaliação da Cobertura da Coleta / Luís Carlos de Souza Oliveira, Marcos Paulo Soares de Freitas, Márcia Regina Martins Lima Dias, Cláudia Maria Ferreira Nascimento, Edie da Silva Mattos e João José Amado Ramalho Júnior - Rio de Janeiro : IBGE, Coordenação Técnica do Censo Demográfico, 2003.
- 📖 **Número 10** - Sistema de informação estatística e a sociedade da informação / Rosa Maria Porcaro - Rio de Janeiro : IBGE, Departamento de Metodologia, 2003.

- 📖 **Número 11** - Indicadores para a agropecuária - 1996 a 2001 /Julio César Perruso, Marcelo de Moraes, Duriez, Roberto Augusto Soares P. Duarte e Carlos Alfredo Barreto Guedes - Rio de Janeiro : IBGE, Coordenação de Agropecuária, 2003.
- 📖 **Número 12** - A Unidade de Metodologia e a Evolução do Uso de Amostragem no IBGE, 2003 / Sônia Albieri - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2003.
- 📖 **Número 13** - Estimando a Precisão das Estimativas das Taxas de Mortalidade Obtidas a Partir da PNAD / Pedro Luis do Nascimento Silva e Djalma Galvão Carneiro Pessoa. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2004.
- 📖 **Número 14** - A Qualidade na Produção de Estatísticas no IBGE / Zélia Magalhães Bianchini. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2004
- 📖 **Número 15** - Calibration Estimation: When and Why, How Much and How / Pedro Luis do Nascimento Silva. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2004
- 📖 **Número 16** - Um panorama recente da desigualdade no Brasil a partir dos dados da PNAD 2002 / Ana Lucia Saboia e Barbara Cobo. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2004
- 📖 **Número 17** – Processamento das Áreas de Expansão e Disseminação da Amostra no Censo Demográfico 2000 / Ari Nascimento Silva, Luiz Alberto Matzenbacher e Bruno Freitas Cortez. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2004
- 📖 **Número 18** – Fatores de correção para o registro de nascimentos utilizando registros tardios segundo os grupos de idades das mulheres - Brasil e Unidades da Federação - 1984-2001 / Fernando Roberto Pires de Carvalho e Albuquerque e Selma Regina dos Santos. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2004
- 📖 **Número 19** – O processo de Imputação dos quesitos de migração no Censo Demográfico 2000 / Fernando Roberto P. de C. e Albuquerque, Janaína Reis Xavier Senna e Antonio Roberto Pereira Garcez - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2004
- 📖 **Número 20** – Tábuas de Mortalidade por sexo e grupos de idade - Grandes Regiões e Unidades da Federação - 1980, 1991 e 2000 / Fernando Roberto P. de C. e Albuquerque e Janaína Reis Xavier Senna - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2005
- 📖 **Número 21** – Tempo, trabalho e afazeres domésticos: um estudo com base nos dados da Pesquisa Nacional por Amostra de Domicílios - 2001 e 2005/ Cristiane Soares e Ana Lucia Saboia - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 22** – Estimação de Intervalos de Confiança para Estimadores de Diferenças Temporais na Pesquisa Mensal de Emprego / Mauricio Franca Lila e Marcos Paulo soares de Freitas - Rio de Janeiro: IBGE, Coordenação de Trabalho e Rendimento e Coordenação de Métodos e Qualidade, 2007
- 📖 **Número 23** – Amostra Mestra para o Sistema Integrado de Pesquisas Domiciliares / Marcos Paulo Soares de Freitas, Mauricio Franca Lila, Rosemary Vallejo de Azevedo e Giuseppe de Abreu Antonaci - Rio de Janeiro: IBGE, Coordenação de Métodos e Qualidade, 2007
- 📖 **Número 24** – Sistema Integrado de Pesquisas Domiciliares - SIPD / Coordenação de Trabalho e Rendimento - Rio de Janeiro: IBGE, 2007
- 📖 **Número 25** – Pesquisas Agropecuárias por Amostragem Probabilística no IBGE: Histórico e Perspectivas Futuras / Coordenação de Agropecuária - Rio de Janeiro: IBGE, 2007

- 📖 **Número 26** – Migração Pendular Intrametropolitana no Rio de Janeiro: Reflexões sobre o seu estudo, a partir dos Censos Demográficos de 1980 e 2000 / Antonio de Ponte Jardim e Leila Ervatti - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 27** – Características da fecundidade e da mortalidade segundo a condição migratória das mulheres, com base no quesito de "data fixa" / Fernando Roberto Pires de Carvalho e Albuquerque, Isabel Cristina Maria da Costa e Antonio Roberto Pereira Garcez - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 28** – Utilização de Modelos para Estimar a Mortalidade Brasileira nas Idades Avançadas / Jorcely Victório Franco, Juarez de Castro Oliveira e Fernando Roberto Pires de C. e Albuquerque - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 29** – Influência da mortalidade nos níveis de fecundidade da população brasileira e o intervalo médio entre duas gerações sucessivas - 1980, 1991, 2000 e 2005/ Fernando Roberto Pires de C. e Albuquerque e Maria Lúcia Pereira do Nascimento - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2008
- 📖 **Número 30** - Família nas pesquisas domiciliares: questões e propostas alternativas / Rosa Ribeiro, Ana Lúcia Sabóia - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2008
- 📖 **Número 31** – Setor e Emprego Informal no Brasil - Análise dos resultados da nova série do Sistema de Contas Nacionais / João Hallak Neto, Katia Namir, Luciene Kozovitz, Sandra Rosa Pereira - Rio de Janeiro : IBGE, Coordenação de Contas Nacionais, 2008
- 📖 **Número 32** - Diferenciais de idade entre os casais nas famílias brasileiras / Cristiane Soares. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2008
- 📖 **Número 33** – Estudos de modalidades alternativas de censos demográficos: aspectos de amostragem / IBGE, Diretoria de Pesquisas, Grupo de Trabalho de Amostragem, Estimção e Acumulação de Informações. - Rio de Janeiro : IBGE, 2009.
- 📖 **Número 34** – O Acompanhamento Estatístico da Fabricação de Medicamentos na Indústria Farmacêutica Brasileira/ Marcus José de Oliveira Campos e Luiz Antônio Casemiro dos Santos. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2009.
- 📖 **Número 35** – Áreas mínimas de Comparação / Weuber da Silva Carvalho, Gilson Flaeschen. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2010.
- 📖 **Número 36** – Contabilizando a Sustentabilidade: principais abordagens / Frederico Barcellos, Paulo Gonzaga M. de Carvalho e Sandra De Carlo. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2010.
- 📖 **Número 37** – Indicadores sobre Trabalho Decente: Uma contribuição para o debate da desigualdade de gênero / Cíntia Simões Agostinho e Ana Lucia Saboia. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2011.
- 📖 **Número 38** – Reflexões sobre pesquisas longitudinais: uma contribuição à implementação do Sistema Integrado de Pesquisas Domiciliares / Leonardo Athias. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2011.
- 📖 **Número 39** – Desafios e possibilidades sobre os novos arranjos familiares e a metodologia para identificação de família no Censo / Ana Lucia Saboia, Bárbara Cobo e Gilson Gonçalves Matos. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2012.

- 📖 **Número 40** – Metodologia Estatística da Pesca: Pesca embarcada / Aristides Pereira Lima Green e Guilherme Guimarães Moreira. - Rio de Janeiro : IBGE, Coordenação de Agropecuária e Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2012.
- 📖 **Número 41** – Pareamento Automático na Pesquisa de Avaliação da Cobertura da Coleta do Censo Demográfico / Djalma Galvão Carneiro Pessoa, Fábio Figueiredo Farias e Vinícius Layter Xavier. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2012.
- 📖 **Número 42** – Seminários IBGE – 15 anos disseminando conhecimento / Sonia Albieri. – Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2012.
- 📖 **Número 43** – Estimadores de Diferenças Temporais e suas Variâncias: Uma Abordagem Aplicada ao Estudo de Indicadores Sociais a partir dos Dados da PNAD/ Gilson Gonçalves de Matos, Ana Lucia Saboia, Leonardo Athias. – Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2013.
- 📖 **Número 44** – Disponibilização de Acesso a Microdados em Institutos Nacionais de Estatísticas: Experiência de países selecionados e Eurostat/ Priscila Koeller, Fernanda Vilhena e Maria Luiza Barcellos Zacharias. – Rio de Janeiro : IBGE, Coordenação das Estatísticas Econômicas e Classificações, Coordenação de Indústria e Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2013.
- 📖 **Número 45** – Capacidade Funcional dos Idosos: Uma análise dos suplementos Saúde da PNAD com a teoria da resposta ao item/ Kaizô Iwakami Beltrão, Moema De Poli Teixeira, Maria Isabel Coelho Alves Parahyba e Philip Fletcher. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2013.
- 📖 **Número 46** – Recomendações internacionais sobre estatísticas sociais e como importantes institutos de estatísticas divulgam seus temas, com destaque para a área social/ Ana Lucia Sabóia, Leonardo Athias. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2013.
- 📖 **Número 47** – Uma contribuição para a produção de indicadores educacionais no IBGE: Panorama nacional e experiências internacionais/ Betina Fresneda, Ana Lucia Sabóia. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2013.
- 📖 **Número 48** – Regionalização e Alto Crescimento: uma análise sobre o crescimento de empresas nas Regiões Metropolitanas brasileiras/ Cristiano Santos...[et al.]. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2013.
- 📖 **Número 49** – Indicadores de pobreza nos municípios de Minas Gerais: comparação de métodos de estimação em pequenas áreas/ Débora F. Souza [et al.]. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2014.
- 📖 **Número 50** – Sistema integrado de pesquisas domiciliares: amostra mestra 2010 e amostra da PNAD contínua / Marcos Paulo Soares de Freitas, Giuseppe de Abreu Antonaci. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2014.
- 📖 **Número 51** – O Sistema de Contas Nacionais evolução, principais conceitos e sua implantação no Brasil / João Hallak Neto. - Rio de Janeiro : IBGE, Coordenação de Contas Nacionais, Diretoria de Pesquisas, 2014.
- 📖 **Número 52** – Conjunto mínimo de indicadores padrão de qualidade a ser aplicado no Mercosul / Coordenação de Métodos e Qualidade . – Rio de Janeiro : IBGE, 2014

- 📖 **Número 53** – Codificar para contar – Um retrospecto das classificações econômicas usadas para fins estatísticos / Marcus José de Oliveira Campos. – Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2014.
- 📖 **Número 54** – Sistema de Estatísticas Econômicas Integradas SEEI – O projeto de Reformulação das Estatísticas Econômicas / Comitê de Reformulação das Estatísticas Econômicas – Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2014.