

**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 49

**Indicadores de pobreza nos municípios de Minas Gerais:
comparação de métodos de estimação em pequenas áreas**

Debora F. Souza
Nícia C. H. Brendolin
Viviane C. C. Quintaes
Solange C. Onel
Djalma G. C. Pessoa

Rio de Janeiro

2014

Instituto Brasileiro de Geografia e Estatística - IBGE
Av. Franklin Roosevelt, 166 - Centro - 20021-120 - Rio de Janeiro, RJ - Brasil

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 978-85-240-4312-3

© IBGE. 2014

Impressão

Gráfica Digital/Centro de Documentação e Disseminação de Informações - CDDI/IBGE, em 2014.

Capa

Gerência de Criação/CDDI

Indicadores de pobreza nos municípios de Minas Gerais : comparação de métodos de estimação em pequenas áreas / Debora F. Souza [et al.]. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2014.
p.53 - (Textos para discussão. Diretoria de Pesquisas, ISSN 1518-675X ; n. 49)

Inclui bibliografia.
ISBN 978-85-240-4312-3

1. Pobreza - Minas Gerais. 2. Pobreza - Métodos estatísticos. 3. Minas Gerais - Condições sociais. 4. Brasil - Condições sociais. 5. Pesquisa social - Minas Gerais. 6. Indicadores sociais - Minas Gerais. I. Souza, Debora F. II. Brendolin, Nícia C. H. III. Quintaes, Viviane C. C. IV. Onel, Solange Correa. V. Pessoa, Djalma G. C., 1940-VI. IBGE. Coordenação de Métodos e Qualidade. VII. Série.

Gerência de Biblioteca e Acervos Especiais CDU330.564.055.3(815.1)
RJ/IBGE/2014-03 DEM

Impresso no Brasil / Printed in Brazil

Sumário

Apresentação	3
Introdução	7
Indicadores de pobreza FGT para pequenas áreas	9
Estimadores diretos para indicadores de pobreza	10
Metodologias para estimação de indicadores de pobreza em pequenas áreas	11
Metodologia proposta por Fay e Herriot (1979)	11
Metodologia proposta por Elbers et al. (2002)	12
Metodologia proposta por Molina e Rao (2010)	14
Seleção das subamostras no censo	17
Resultados	20
Erro quadrático médio e vício relativos	22
Ordenação de áreas	40
Conclusões	45
Referências bibliográficas	47
Apêndice A - Modelos selecionados para aplicação do método FH	48
Apêndice B - Resultados da ordenação de áreas para o hiato de pobreza	52

Apresentação

No âmbito dos estudos sobre pobreza, desde o início da década de 2000, técnicos da Coordenação de Métodos e Qualidade, vêm participando de um Grupo de Trabalho da Diretoria de Pesquisas, para a construção do Mapa de Pobreza usando uma metodologia desenvolvida por técnicos do Banco Mundial e descrita em método proposto por Elbers et al. (2002). Como resultado da aplicação dessa metodologia, em 2008, foi publicado o Mapa de Pobreza e Desigualdade: municípios brasileiros - 2003, no formato DVD, composto por três conjuntos básicos de informações: os textos metodológicos que deram suporte ao tratamento das informações; as estimativas municipais apresentadas através de mapas de pobreza e desigualdade; e um conjunto de mapas com temas complementares que ajudam a contextualizar os determinantes da pobreza e da desigualdade.

Dando continuidade aos estudos, vêm sendo avaliados novos métodos de estimação em pequenas áreas que, em geral, combinam informações em nível geográfico agregado, coletadas em pesquisas domiciliares amostrais de múltiplos propósitos, com informações em níveis geográficos mais detalhados, como é o caso das obtidas nos censos demográficos.

Este documento contém a descrição desses estudos e seus resultados envolvendo a comparação de três métodos de estimação em pequenas áreas: o método desenvolvido em Fay e Herriot (1979), o descrito em Elbers et al. (2002) e o apresentado em Molina e Rao (2010).

Sonia Albieri

Coordenação de Métodos e Qualidade

Introdução

Institutos de estatísticas oficiais cada vez mais se deparam com grande demanda por informações detalhadas e precisas, e por outro lado sofrem com a constante restrição financeira na produção das pesquisas amostrais. Neste sentido, pesquisadores de vários órgãos produtores de estatísticas oficiais, atualmente, estudam metodologias de estimação em pequenos domínios a fim de fornecer estimativas para áreas geográficas ou domínios menores com precisão controlada, sem aumento de custos.

Para estimar pobreza em pequenas áreas, são utilizadas metodologias que combinam as informações coletadas em pesquisas domiciliares amostrais de múltiplos propósitos com a ampla cobertura geográfica dos censos demográficos.

Um método de estimação em pequenas áreas amplamente utilizado por institutos de estatística é o desenvolvido em Fay e Herriot (1979). Tal metodologia é baseada em modelo de regressão linear com efeitos aleatórios de área. Já, o Mapa de Pobreza publicado em 2008 pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2008) adotou o método proposto por Elbers et al. (2002) e teve por finalidade fornecer uma descrição detalhada da distribuição espacial da pobreza no país.

Recentemente, Molina e Rao (2010) desenvolveram uma metodologia na qual utilizam o máximo possível da informação amostral para estimação dos indicadores de pobreza nas pequenas áreas. Segundo os autores, esta proposta resulta em menores erros quadráticos médios quando comparados aos obtidos a partir do método apresentado em Elbers et al. (2002).

O objetivo principal deste trabalho é comparar três métodos de estimação em pequenas áreas: o método desenvolvido em Fay e Herriot (1979), o descrito em Elbers et al. (2002) e o apresentado em Molina e Rao (2010). Primeiro, comparam-se os métodos em termos de erro quadrático médio relativo e vício relativo. Para o estudo, foram extraídas subamostras de uma população a partir do desenho amostral utilizado na Pesquisa de Orçamentos Familiares de 2008-2009 (POF) e implementaram-se as três metodologias nestas subamostras. A amostra do Censo Demográfico de 2010 do Estado de Minas Gerais, excluindo-se os domicílios com rendimento total igual a zero, foi considerada a população. Ao final, compararam-se as estimativas da incidência de pobreza obtidas pelas metodologias a partir das subamostras com aquelas calculadas na população. Os resultados foram avaliados tanto para municípios quanto para microrregiões. Posteriormente, utilizou-se um critério de comparação baseado na ordenação de áreas, definida a partir das estimativas obtidas pelos métodos. Tal ordenação foi feita utilizando-se o coeficiente de correlação de postos de Spearman.

Os resultados desse estudo podem fornecer subsídios para escolher uma metodologia adequada para estimação de indicadores necessários para a produção de Mapa de Pobreza no país.

Este trabalho está organizado em 7 seções. As Seções 2 e 3 apresentam, respectivamente, os indicadores usuais de pobreza e seus estimadores diretos. A Seção 4 está subdividida em três partes. Na primeira, é apresentado o modelo descrito em Fay e Herriot (1979). Na segunda parte, definem-se as especificações do método desenvolvido por Elbers et al. (2002) e na terceira parte, descreve-se a metodologia proposta pelos autores Molina e Rao (2010). A quinta seção detalha a forma de obtenção das subamostras usadas no experimento. Na Seção 6, apresentam-se os resultados obtidos e as possíveis comparações entre os métodos. Por fim, na Seção 7, são apresentadas as conclusões.

Indicadores de pobreza FGT para pequenas áreas

Considere uma população finita de tamanho N particionada em D áreas de tamanhos N_1, \dots, N_D . Suponha que E_{dj} seja uma medida de bem-estar adequada do indivíduo j na pequena área d , tal como renda ou despesa, e seja z uma dada linha de pobreza; ou seja, o limiar para E_{dj} abaixo do qual uma pessoa é considerada pobre. A família de indicadores de pobreza, chamada de FGT, introduzida por Foster et al. (1984), para cada pequena área d , é definida como:

$$FGT_{cd} = \frac{1}{N_d} \sum_{j=1}^{N_d} FGT_{cdj} = \frac{1}{N_d} \sum_{j=1}^{N_d} \left(\frac{z - E_{dj}}{z} \right)^c I(E_{dj} < z), \quad c = 0, 1, 2, \quad d = 1, \dots, D,$$

em que $I(E_{dj} < z) = 1$ se $E_{dj} < z$ (pessoa em situação de pobreza) e $I(E_{dj} < z) = 0$ se $E_{dj} \geq z$ (pessoa fora de situação de pobreza). Para $c = 0$ tem-se a proporção de indivíduos pobres na área d , também chamada de incidência de pobreza. A medida FGT quando $c = 1$ é chamada de hiato de pobreza, e mede a média da distância relativa para a não-pobreza de cada indivíduo da área de interesse. Quando $c = 2$ a medida é chamada de severidade da pobreza. Este indicador eleva ao quadrado os hiatos de pobreza enfatizando a pobreza extrema.

Estimadores diretos para indicadores de pobreza

No processo de inferência, uma amostra aleatória de tamanho $n < N$ é retirada da população de acordo com um desenho amostral específico. Suponha que Ω seja o conjunto de índices das unidades contidas na população, s o conjunto de índices das unidades selecionadas na amostra e r o conjunto de índices das unidades não selecionadas (com tamanho $N - n$). Os subconjuntos de Ω , s , N e n contendo os índices das unidades da área d são denotados por Ω_d , s_d , N_d , e n_d respectivamente, em que $n = n_1 + \dots + n_D$. Os indicadores de pobreza FGT não ponderados da amostra são dados por:

$$fgt_{cd} = \frac{1}{n_d} \sum_{j \in s_d} fgt_{cdj} = \frac{1}{n_d} \sum_{j \in s_d} \left(\frac{z - E_{dj}}{z} \right)^c I(E_{dj} < z), \quad c = 0, 1, 2, \quad d = 1, \dots, D. \quad (1)$$

Um estimador direto de uma pequena área utiliza apenas os dados da área de interesse contidos na amostra. Sendo w_{dj} o peso amostral (inverso da probabilidade de inclusão) do indivíduo j da área d , os estimadores diretos das medidas de pobreza FGT são dados por:

$$fgt_{cd}^w = \frac{1}{\hat{N}_d} \sum_{j \in s_d} fgt_{cdj}^w = \frac{1}{\hat{N}_d} \sum_{j \in s_d} w_{dj} \left(\frac{z - E_{dj}}{z} \right)^c I(E_{dj} < z), \quad c = 0, 1, 2, \quad d = 1, \dots, D. \quad (2)$$

em que $\hat{N}_d = \sum_{j \in s_d} w_{dj}$ é o estimador direto do tamanho populacional da área d , N_d . Se os pesos amostrais não dependem da unidade j , por exemplo $w_{dj} = N_d/n_d$ quando tem-se amostragem aleatória simples, então (2) reduz-se à média não ponderada (1).

Tamanhos amostrais limitados em algumas áreas impedem o uso de estimadores como (1) ou (2). Para obtenção de estimadores confiáveis para estas áreas torna-se necessária a utilização de técnicas de estimação em pequenas áreas (Rao, 2003). Tais técnicas melhoram os procedimentos de estimação através de modelos que estabelecem alguma relação entre as áreas, baseados em informações auxiliares (providas do censo demográfico e/ou registros administrativos) relacionadas às variáveis de bem estar de interesse. Estes modelos fornecem estimadores “indiretos” que fazem uso de dados de outras áreas correlatas, os quais podem reduzir drasticamente os erros de estimação.

Metodologias para estimação de indicadores de pobreza em pequenas áreas

Metodologia proposta por Fay e Herriot (1979)

Esta subsecção detalha o modelo proposto em Fay e Herriot (1979). É um modelo linear com efeito aleatório de área e foi desenvolvido pelos autores para prever a renda *per capita* em áreas de até 1000 habitantes nos Estados Unidos. Ele é útil nos casos em que as variáveis auxiliares estão disponíveis no nível de área ou quando não é possível ligar as informações das unidades amostrais com os microdados do censo e com os registros administrativos.

Suponha que $\theta_d = g(\bar{Y}_d)$ seja uma função da medida \bar{Y}_d (por exemplo, proporção de pessoas pobres na área d) para a qual se deseja obter uma estimativa e que esteja relacionada às variáveis auxiliares de área $\mathbf{x}_d = (x_{1d}, \dots, x_{pd})'$ através do seguinte modelo linear:

$$\theta_d = \mathbf{x}'_d \boldsymbol{\beta} + v_d, \quad d = 1, \dots, D^*, \quad (3)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ é o vetor dos coeficientes da regressão. Além disso, os v_d 's são efeitos aleatórios de área independentes e identicamente distribuídos com $E(v_d) = 0$ e $Var(v_d) = \sigma_v^2$. Frequentemente, assume-se a normalidade destes efeitos aleatórios, mas é possível relaxar esta hipótese a fim de se realizar inferências “robustas”.

Em algumas aplicações, nem todas as áreas são selecionadas na amostra. Suponha que das D áreas existentes na população apenas D^* são selecionadas. Neste caso, assume-se o modelo da forma (3) para a população. Assume-se, também, que as áreas selecionadas obedecem ao modelo populacional.

A fim de se fazer inferências sobre as médias das pequenas áreas, \bar{Y}_d , sob o modelo (3), assume-se que os estimadores diretos $\hat{\bar{Y}}_d$ estejam disponíveis. Como no método James-Stein (ver James e Stein (1961)), temos que:

$$\hat{\theta}_d = g(\hat{\bar{Y}}_d) = \theta_d + e_d, \quad d = 1, \dots, D^*, \quad (4)$$

no qual os erros amostrais e_d são independentes com $E(e_d|\theta_d) = 0$ e variância $Var(e_d|\theta_d) = \sigma_{e,d}^2$ conhecida. Tais suposições podem ser restritivas em algumas aplicações. Por exemplo, o estimador

direto $\hat{\theta}_d$ pode ser viciado se a função $g(\cdot)$ for não-linear e o tamanho amostral da área pequeno. A hipótese da variância $\sigma_{e,d}^2$ ser conhecida pode ser relaxada ao estimá-la através de dados amostrais e então suavizar as variâncias estimadas para se ter uma estimativa mais estável de $\sigma_{e,d}^2$.

Combinando (3) com (4) obtem-se o modelo:

$$\hat{\theta}_d = \mathbf{x}'_d \boldsymbol{\beta} + v_d + e_d, \quad d = 1, \dots, D^*, \quad (5)$$

em que v_d e e_d são independentes.

Este modelo é um caso especial do modelo linear misto e é estimado utilizando-se as informações das áreas amostradas e, para cada domínio, é obtido um estimador composto, θ_d^{EBLUP} , dado pela média ponderada entre o estimador direto $\hat{\theta}_d$ e o estimador sintético $\tilde{\theta}_d = \mathbf{x}'_d \hat{\boldsymbol{\beta}}$:

$$\theta_d^{EBLUP} = \hat{\gamma}_d \hat{\theta}_d + (1 - \hat{\gamma}_d) \tilde{\theta}_d, \quad d = 1, \dots, D^*,$$

em que $\hat{\gamma}_d = \frac{\hat{\sigma}_v^2}{\hat{\sigma}_{e,d}^2 + \hat{\sigma}_v^2}$. Para as áreas nas quais não há amostra é utilizado apenas o estimador sintético.

O estimador para σ_v^2 pode ser obtido através de:

$$E \left[\sum_d (\hat{\theta}_d - \mathbf{x}'_d \tilde{\boldsymbol{\beta}})^2 / (\sigma_{e,d}^2 + \sigma_v^2) \right] = E [h(\sigma_v^2)] = D^* - p,$$

em que $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}^{(\sigma_v^2)}$. O estimador $\hat{\sigma}_v^2$ é obtido resolvendo-se a equação $h(\sigma_v^2) = D^* - p$ de modo iterativo e supondo que $\hat{\sigma}_v^2 = 0$ quando não existe solução positiva. Fay e Herriot (1979) sugeriram a seguinte solução iterativa: começando com o valor inicial $\sigma_v^{2(0)} = 0$, defina

$$\sigma_v^{2(l+1)} = \sigma_v^{2(l)} + \frac{1}{h'_*(\sigma_v^{2(l)})} \left[D^* - p - h(\sigma_v^{2(l)}) \right]$$

restringindo $\sigma_v^{2(l+1)} \geq 0$ em que

$$h'_*(\sigma_v^2) = - \sum_d (\hat{\theta}_d - \mathbf{x}'_d \tilde{\boldsymbol{\beta}})^2 / (\sigma_{e,d}^2 + \sigma_v^2)^2$$

é uma aproximação da derivada de $h(\sigma_v^2)$. A convergência é rápida, geralmente requer menos de 10 iterações.

Metodologia proposta por Elbers et al. (2002)

Nesta subseção, descreve-se a metodologia proposta por Elbers et al. (2002) para construção de indicadores de pobreza para pequenas áreas. O método consiste em estimar, na mesma amplitude territorial do censo demográfico, uma informação que resuma o bem-estar social obtida diretamente,

ou derivada, a partir de uma pesquisa por amostragem. Dessa forma, é possível utilizar a ampla base geográfica do censo para produzir estimativas dos principais índices de pobreza para áreas em níveis geográficos menores, assim como calcular a precisão destas estimativas.

Define-se y_{dj} como sendo função de uma variável de bem-estar do domicílio j pertencente à área d . O modelo pode ser escrito da seguinte forma:

$$y_{dj} = \mathbf{x}'_{dj}\boldsymbol{\beta} + \eta_{dj}, \quad \eta_{dj} \sim F(0, \Sigma), \quad (6)$$

em que F é uma distribuição com vetor de médias 0 e matriz de variâncias-covariâncias Σ e \mathbf{x}_{dj} é um vetor contendo os valores observados das variáveis comuns no censo e na pesquisa amostral, para o domicílio j da área d , $j = 1, \dots, n_d$ e $d = 1, \dots, D^*$. É possível introduzir indicadores de níveis geográficos mais agregados para controlar o efeito de localização quando este não é totalmente explicado pelas variáveis regressoras. Tais indicadores podem ser obtidos do próprio censo, bem como de outras bases de dados.

O componente de erro pode ser decomposto em duas partes: (i) uma associada ao domicílio e; (ii) outra associada ao nível geográfico (área) onde se localiza o domicílio (por exemplo, município, distrito, etc.). Pode-se dizer que o componente de erro que descreve os dados é decomposto segundo uma estrutura particular, possibilitando explicar grande parte da variabilidade total dos dados. Então, tem-se:

$$\eta_{dj} = u_d + e_{dj},$$

em que u_d e e_{dj} são independentes, com $u_d \sim N(0, \sigma_u^2)$ e $e_{dj} \sim N(0, \sigma_{e,dj}^2)$.

A estimativa da variância de u_d , $\hat{\sigma}_u^2$, é obtida regredindo-se os resíduos do modelo (6) nas variáveis indicadoras dos níveis de locação e então os parâmetros do modelo original são estimados através de mínimos quadrados generalizados.

Admitindo-se que os erros do nível domiciliar e_{dj} sejam heterocedásticos, Elbers et al. (2002) sugerem estimar a regressão logística:

$$\ln \left(\frac{e_{dj}^2}{A - e_{dj}^2} \right) = \mathbf{z}'_{dj}\boldsymbol{\alpha} + r_{dj}, \quad (7)$$

e estimam a variância no nível de domicílio segundo a fórmula:

$$\hat{\sigma}_{e,dj}^2 \approx \left[\frac{AB}{1+B} \right] + \frac{1}{2} \text{var}(r) \left[\frac{AB(1-B)}{(1+B)^3} \right],$$

onde $A = 1,05 \max \{e_{dj}^2\}$, $B = \exp \{ \mathbf{z}'_{dj}\hat{\boldsymbol{\alpha}} \}$, $\text{var}(r)$ é o erro quadrático dos resíduos da regressão logística estimada e \mathbf{z}'_{dj} é um vetor de variáveis explicativas.

Supondo que u_d e e_{dj} possuem distribuição normal, Elbers et al. (2002) derivaram uma estimativa

da variância do efeito de área u_d :

$$var(\hat{\sigma}_u^2) \approx \sum_d 2 \left\{ a_d^2 [(\hat{\sigma}_u^2)^2 + (\hat{\tau}_d^2)^2 + 2\hat{\sigma}_u^2 \hat{\tau}_d^2] + b_d^2 \frac{(\hat{\tau}_d^2)^2}{n_d - 1} \right\}.$$

em que $\hat{\tau}_d^2 = \frac{1}{n_d(n_d-1)} \sum_j (e_{dj} - e_d)^2$, $e_d = 1/n_d \sum_j e_{dj}$, $a_d = w_d / \sum_j w_j(1 - w_j)$, $b_d = w_d(1 - w_d) / \sum_j w_j(1 - w_j)$, $\hat{\sigma}_u^2 = \sum_d a_d \eta_d^2 - \sum_d b_d \hat{\tau}_d^2$, $\eta_d^2 = u_d + e_d$, $w_d = \sum_j w_{dj} / n_d$, n_d é o total de pessoas associado ao nível geográfico d e w_j é o fator de expansão da pesquisa do domicílio j .

Para obter a variável de bem-estar para todos os domicílios no censo, é necessário combinar os resultados estimados anteriormente pelos ajustes de modelos na pesquisa com as informações do censo, vetor de características x_{dj}^{censo} . As predições da variável de bem-estar são realizadas e, subsequentemente, estimativas de medidas de pobreza e seus respectivos erros padrões no nível de área podem ser gerados. Neste procedimento, uma série de simulações é conduzida para cada simulação l , $l = 1, \dots, L$, um conjunto de parâmetros é gerado a partir de distribuições derivadas das estimativas encontradas anteriormente. Os coeficientes $\tilde{\beta}^{(l)}$ são obtidos da distribuição normal multivariada cujos parâmetros são os coeficientes e matriz de covariâncias estimados pelo modelo (6). De forma análoga, os coeficientes $\tilde{\alpha}^{(l)}$ são gerados a partir do modelo (7). Adicionalmente, $\tilde{\sigma}_u^{2(l)}$, um valor simulado para a variância do componente de efeito de área é gerado a partir de uma distribuição gama com média $\hat{\sigma}_u^2$ e variância $var(\hat{\sigma}_u^2)$. Combinando os coeficientes $\tilde{\alpha}^{(l)}$ com os dados do censo, obtém-se um valor para a variância do componente de erro do domicílio. Para cada domicílio dos dados do censo, são extraídos $\tilde{u}_d^{(l)}$ e $\tilde{e}_{dj}^{(l)}$ de suas distribuições correspondentes e então a variável dependente é gerada através da equação $\tilde{y}_{dj}^{(l)} = x_{dj}^{censo} \tilde{\beta}^{(l)} + \tilde{u}_d^{(l)} + \tilde{e}_{dj}^{(l)}$, $j = 1, \dots, N_d$ e $d = 1, \dots, D$.

Assim, a partir da variável de bem-estar simulada, calculam-se os principais indicadores de pobreza e desigualdade L vezes. As estimativas pontuais dessas medidas são dadas pelas médias de todas as L iterações. Este procedimento permite que sejam estimadas as variâncias e, portanto, a precisão das estimativas.

Metodologia proposta por Molina e Rao (2010)

Esta subseção apresenta uma breve descrição da proposta metodológica dos autores Molina e Rao (2010) também para obtenção de indicadores de pobreza em pequenas áreas. A ideia do método é combinar os dados de domicílio da pesquisa aos dados do censo e, através de modelos hierárquicos de dois níveis com efeitos aleatórios no intercepto, prever a variável de bem-estar.

Seja $\mathbf{y}_d = (Y_{d1}, \dots, Y_{dN_d})' = (\mathbf{y}'_{ds}, \mathbf{y}'_{dr})'$ o vetor cujos componentes são funções da variável de

bem-estar do domínio d com N_d unidades da população, onde \mathbf{y}_{ds} é um subvetor de elementos da pesquisa e \mathbf{y}_{dr} é um subvetor de elementos não existentes na amostra. O modelo pode ser escrito da seguinte forma:

$$\begin{aligned} Y_{dj} &= \mathbf{x}'_{dj}\boldsymbol{\beta} + u_d + e_{dj}, \quad u_d \sim iid \quad N(0, \sigma_u^2), \\ e_{dj} &\sim iid \quad N(0, \sigma_e^2), \quad j = 1, \dots, N_d \quad e \quad d = 1, \dots, D, \end{aligned} \quad (8)$$

em que o efeito de área u_d e os erros e_{dj} são independentes. Os vetores \mathbf{y}_d , $d = 1, \dots, D$, sob o modelo (8) são independentes com $\mathbf{y}_d \sim N(\boldsymbol{\mu}_d, V_d)$, no qual $\boldsymbol{\mu}_d = X_d\boldsymbol{\beta}$ e $V_d = \sigma_u^2 \mathbf{1}_{N_d} \mathbf{1}'_{N_d} + \sigma_e^2 I_{N_d}$.

Os autores decompõem $\boldsymbol{\mu}_d = E(\mathbf{y}_d)$ e $V_d = Var(\mathbf{y}_d)$ de acordo com os elementos da amostra e fora da amostra. A distribuição dos elementos fora da amostra, \mathbf{y}_{dr} , dados os elementos pertencentes à amostra, \mathbf{y}_{ds} , é dada por:

$$\mathbf{y}_{dr} | \mathbf{y}_{ds} \sim N(\boldsymbol{\mu}_{dr|s}, V_{dr|s}), \quad (9)$$

em que o vetor de médias condicionais e a matriz de covariâncias são dados por:

$$\begin{aligned} \boldsymbol{\mu}_{dr|s} &= X_{dr}\boldsymbol{\beta} + \sigma_u^2 \mathbf{1}_{N_d-n_d} \mathbf{1}'_{n_d} V_{ds}^{-1} (\mathbf{y}_s - X_s\boldsymbol{\beta}), \\ V_{dr|s} &= \sigma_u^2 (1 - \gamma_d) \mathbf{1}_{N_d-n_d} \mathbf{1}'_{N_d-n_d} + \sigma_e^2 I_{N_d-n_d}, \end{aligned}$$

com $\gamma_d = \sigma_u^2 / (\sigma_u^2 + \sigma_e^2 / n_d)$. Note que $\mathbf{y}_{dr} | \mathbf{y}_{ds}$ e $\mathbf{y}_{dr} | \mathbf{y}_s$ seguem a mesma distribuição (9) devido à independência dos vetores \mathbf{y}_d .

Observe que a matriz $V_{dr|s}$ corresponde à matriz de covariâncias de um vetor \mathbf{y}_{dr} gerado pelo modelo

$$\mathbf{y}_{dr} = \boldsymbol{\mu}_{dr|s} + v_d \mathbf{1}_{N_d-n_d} + \boldsymbol{\epsilon}_{dr}, \quad (10)$$

com novos efeitos aleatórios v_d e $\boldsymbol{\epsilon}_{dr}$ independentes satisfazendo

$$v_d \sim N(0, \sigma_u^2 (1 - \gamma_d)) \quad e \quad \boldsymbol{\epsilon}_{dr} \sim N(\mathbf{0}_{N_d-n_d}, \sigma_e^2 I_{N_d-n_d}).$$

Utilizam-se as expressões anteriores para gerar variáveis normais univariadas, $v_d \sim N(0, \sigma_u^2 (1 - \gamma_d))$ e $\boldsymbol{\epsilon}_{dj} \sim N(0, \sigma_e^2)$, independentemente, para $j \in r_d$ e então obter as respostas Y_{dj} usando a equação (10) através do valor conhecido $\boldsymbol{\mu}_{dr|s}$. Na prática, todos os parâmetros desconhecidos do modelo $\boldsymbol{\beta}$, σ_u^2 e σ_e^2 são substituídos por estimadores adequados, assim, as variáveis Y_{dj} são geradas a partir das correspondentes distribuições normais estimadas de \hat{v}_d e $\hat{\epsilon}_{dj}$ e as medidas de pobreza correspondentes à área d , $d = 1, \dots, D$, são calculadas através da seguinte expressão:

$$fgt_{cd} = \frac{1}{N_d} \left[\sum_{j \in s_d} fgt_{cdj}^w + \sum_{j \in r_d} fgt_{cdj} \right].$$

Este processo é repetido L vezes. A metodologia proposta para estimar indicadores de pobreza pode ser resumida da seguinte forma: (a) ajuste do modelo (8) aos dados transformados \mathbf{y}_s ; (b) geração de L vetores $\mathbf{y}_r^{(l)}$, $l = 1, \dots, L$ a partir de (9) ou, equivalentemente, a partir de (10), substituindo os parâmetros desconhecidos por estimadores obtidos em (a); (c) cálculo das medidas

de pobreza utilizando aproximação de Monte Carlo (Robert e Casella, 2004), os L vetores gerados $\mathbf{y}_r^{(l)}$, $l = 1, \dots, L$, e os dados da amostra \mathbf{y}_s .

No caso em que não é possível associar as unidades do censo às da pesquisa e nem todas as áreas de interesse estão na pesquisa, os autores sugerem uma adaptação que consiste em calcular a média de Y_{dj} e a variância do erro aleatório da seguinte forma: $\hat{\mu}_{dj} = \mathbf{x}'_{dj}{}^{censo} \hat{\beta} + \hat{u}_d$ e $\hat{\sigma}_v^2 = \hat{\sigma}_u^2(1 - \hat{\gamma}_d)$, para as áreas pertencentes à amostra; $\hat{\mu}_{dj} = \mathbf{x}'_{dj}{}^{censo} \hat{\beta}$ e $\hat{\sigma}_v^2 = \hat{\sigma}_u^2$, para as áreas fora da amostra. Assim, L vetores de Y_{dj} são gerados e as medidas de pobreza são calculadas através da fórmula: $fgt_{cd} = (1/L) \sum_{l=1}^L (1/N_d) \sum_{j=1}^{N_d} fgt_{cdj}^l$ por aproximação de Monte Carlo.

Seleção das subamostras no censo

Para avaliação das metodologias apresentadas na seção anterior, foram selecionadas 400 subamostras de uma população, seguindo o desenho amostral da POF. O objetivo foi comparar, em termos de vício, erro quadrático médio relativos (VR e EQMR, respectivamente) e ordenação de áreas, as estimativas dos indicadores incidência e hiato de pobreza obtidas pelas três metodologias com os respectivos valores diretamente estimados a partir dos dados populacionais.

A amostra do Censo Demográfico de 2010 do Estado de Minas Gerais de domicílios particulares permanentes dos setores dos tipos normal e aglomerado subnormal, excluindo-se os domicílios localizados em terras indígenas foi considerada a população desse estudo, totalizando 763.505 domicílios. A escolha de Minas Gerais ocorreu em virtude deste ser um estado que se assemelha e reflete as condições socioeconômicas do país. Desta população, excluíram-se os domicílios com rendimento total igual a zero, cerca de 3,11% dos domicílios da população, devido ao fato dos processos de modelagem adotados pelos métodos não possuírem tratamento específico para concentração de zeros na variável resposta. Ao final, restaram 739.762 domicílios na população.

A edição de 2008-2009 da POF utilizou como base de seleção de setores a amostra mestra que foi implementada levando-se em consideração o seguinte esquema de estratificação:

1. Divisão administrativa: municípios das capitais, regiões metropolitanas e Regiões Integradas de Desenvolvimento - RIDES;
2. Espacial/geográfica: áreas de ponderação, municípios;
3. Situação dos setores censitários: urbana ou rural; e
4. Estatística: a partir da variável renda do responsável, obtida no Censo Demográfico 2000.

Com relação ao último nível do esquema de estratificação da amostra mestra, dentro de cada estrato geográfico definido, foi então calculado um quantitativo de estratos estatísticos, cujo número total foi diferente para cada Unidade da Federação, considerando as respectivas particularidades. Para a construção dos estratos estatísticos (socioeconômicos), a variável renda total do responsável pelo domicílio, obtida a partir dos resultados do Censo Demográfico 2000, foi utilizada para efeito dos cálculos.

A seleção de setores para compor a amostra mestra ocorreu de forma independente em cada estrato proporcional ao número de domicílios do setor da malha setorial do Censo Demográfico

2000. Os setores da amostra da POF 2008-2009, por sua vez, foram selecionados a partir dos setores da amostra mestra, através de um procedimento de seleção definido como amostra aleatória simples. 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. 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. 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: 13 domicílios nos setores urbanos, 18 nos setores rurais. Para mais detalhes, ver IBGE (2010).

As subamostras de setores usadas neste experimento foram selecionadas a partir da amostra mestra baseada na malha setorial do Censo Demográfico de 2010. Em cada setor de cada subamostra, os domicílios foram selecionados por amostragem aleatória simples sem reposição, mantendo o número de 13 domicílios por setor urbano e 18 por setor rural.

A área de interesse neste estudo é o município por ser uma unidade administrativa importante para implementação de políticas públicas. Vale ressaltar que, de acordo com o plano amostral adotado e exposto nesta seção, não se tem garantia de que os 853 municípios de Minas Gerais estejam contemplados nas 400 subamostras de domicílios. Conforme pode ser observado na Tabela 1, 449 municípios não apareceram em nenhuma das subamostras. Apenas 14 municípios tinham domicílios selecionados em todas as subamostras e 96 apareciam, pelo menos, em 301 subamostras.

Tabela 1: Distribuição do número de municípios pelo número de subamostras em que eles apareceram.

Número subamostras	Número municípios
0	449
153 a 200	191
201 a 300	103
301 a 399	96
400	14

A ausência de áreas de interesse na amostra, bem como o pequeno número de unidades amostradas em áreas selecionadas podem prejudicar o desempenho dos métodos, fazendo com que estes forneçam estimativas viciadas e/ou com baixa precisão. A agregação de áreas pode minimizar esse problema devido ao aumento do tamanho de amostra nos domínios. A fim de verificar esta hipótese, as comparações entre os métodos também foram realizadas para microrregiões as quais são grupos de municípios. Como pode ser visto na Tabela 2, das 66 microrregiões de Minas Gerais, 40 possuíam observações em todas as subamostras. As demais foram encontradas em mais de 155 subamostras.

Tabela 2: Distribuição do número de microrregiões pelo número de subamostras em que elas apareceram.

Número subamostras	Número microrregiões
155	1
270	1
337 a 399	24
400	40

Resultados

Nesta seção, apresentam-se os resultados da aplicação das metodologias FH (Fay e Herriot, 1979), ELL (Elbers et al., 2002) e MR (Molina e Rao, 2010), para municípios e microrregiões, nas 400 subamostras da população. No primeiro método, objetivou-se estimar a proporção de pessoas pobres e o hiato de pobreza em cada domínio, portanto utilizaram-se como variáveis respostas dos modelos as estimativas diretas destas medidas obtidas através dos dados das subamostras. A variável de bem-estar considerada, nos métodos ELL e MR, foi a renda domiciliar *per capita*, pois esta estava disponível nos dados da população e os indicadores derivados puderam ser comparados com aqueles simulados pelos métodos. Para cada subamostra, um modelo de regressão específico foi selecionado para os métodos ELL e MR. Para o método FH, foi escolhido um modelo diferente para cada medida de pobreza e tipo de área. Algumas das variáveis utilizadas nos modelos foram: anos de estudo, cor/raça e idade, referentes ao responsável pelo domicílio; tipo de escoamento sanitário e existência de bens duráveis, referentes ao domicílio. Tais variáveis representam características domiciliares, demográficas e educacionais, incluindo informações agregadas no nível de setor censitário. A Tabela 3 contém as variáveis do modelo de regressão selecionado para os dados populacionais (considerando nível de domicílio), os coeficientes estimados para esse modelo, número de vezes em que essas variáveis apareceram nos modelos selecionados para os métodos ELL e MR para as 400 subamostras e a média dos coeficientes estimados nas subamostras. A maioria das variáveis do modelo populacional apareceu nos modelos das subamostras e, em geral, os valores estimados dos coeficientes na população ficaram próximos das médias dos coeficientes nas subamostras. Este fato é um indício de que os modelos ajustados nas subamostras são válidos para a população, podendo ser utilizados para fazer previsões no nível estadual. Os modelos populacionais para a proporção de pobres e o hiato de pobreza ajustados nos níveis de município e microrregião encontram-se no Apêndice A.

Tabela 3: Variáveis escolhidas para o modelo populacional (considerando nível de domicílio), seus coeficientes estimados e limites inferior (LI) e superior (LS) dos intervalos de confiança de 95%, total de modelos das subamostras em que estas apareceram, médias dos coeficientes estimados nestes modelos e seus respectivos limites inferior e superior dos intervalos de confiança de 95%.

Descrição das variáveis	Subamostras				População		
	Total	LI	Média	LS	LI	Coeficiente	LS
Intercepto	400	4,3859	4,7965	5,6181	4,6861	4,6954	4,7047
Existência de bens intermediários	400	0,1619	0,1990	0,2350	0,2012	0,2047	0,2082
Existência de internet	400	0,2344	0,2846	0,3467	0,2169	0,2212	0,2255
Existência de telefone celular	399	0,1550	0,1926	0,2352	0,1674	0,1714	0,1754
Existência de telefone fixo	351	0,1294	0,1701	0,2090	0,1260	0,1298	0,1335
Domicílio com dois banheiros	292	0,1162	0,1546	0,1953	0,1423	0,1465	0,1506
Domicílio com três banheiros	392	0,3607	0,4572	0,5694	0,3257	0,3340	0,3422
Tipo de família - unipessoal	400	0,4036	0,5237	0,5950	0,4941	0,4992	0,5043
Tipo de família - casal sem filhos	390	0,2401	0,2842	0,3299	0,2447	0,2490	0,2534
Densidade de moradores por cômodo	393	-0,3562	-0,2652	-0,1765	-0,3112	-0,3053	-0,2994
Responsável com emprego formal	369	0,2665	0,4939	0,6951	0,6014	0,6085	0,6155
Responsável empregador ou conta própria	370	0,3226	0,5411	0,7520	0,6276	0,6349	0,6421
Responsável sem carteira de trabalho assinada	368	-0,1117	0,3432	0,5454	0,4678	0,4752	0,4826
Quadrado da idade do responsável	400	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
Responsável inativo	225	-0,2907	0,1990	0,3636	0,2575	0,2644	0,2713
Total de crianças com menos de 7 anos de idade	400	-0,2113	-0,1805	-0,1482	-0,1829	-0,1801	-0,1774
Total de crianças de 7 a 14 anos de idade	400	-0,2415	-0,2124	-0,1847	-0,2114	-0,2092	-0,2070
Proporção de moradores com nível superior completo	400	0,0086	0,0096	0,0105	0,0077	0,0078	0,0079
Total de moradores com emprego formal no setor	395	0,6983	1,2688	1,6944	0,9917	1,0050	1,0183
Total de empregadores ou conta própria no setor	251	0,7592	1,0122	1,3108	0,4687	0,4872	0,5056

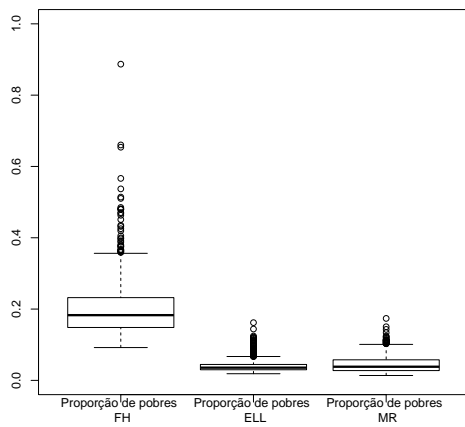
Para cada subamostra e método (ELL e MR), partindo-se de seu modelo ajustado, foram geradas 1000 replicações da renda domiciliar *per capita* para todos os domicílios da população. A linha de pobreza utilizada foi de R\$ 255,00, metade do salário mínimo vigente em 2010. Na aplicação do método MR, utilizou-se a adaptação sugerida pelos autores e descrita na Subseção “Metodologia proposta por Molina e Rao (2010)” para o caso de ausência de áreas de interesse na amostra. Os resultados dos métodos FH, ELL e MR foram obtidos utilizando rotina desenvolvida em SAS 9.2, pacote PovMap 1.2 e código desenvolvido em R 2.14.1, respectivamente. A seguir são apresentadas as comparações dos resultados obtidos pelos três métodos com os valores obtidos na população em termos de erro quadrático médio relativo, vício relativo e ordenação de áreas.

Erro quadrático médio e vício relativos

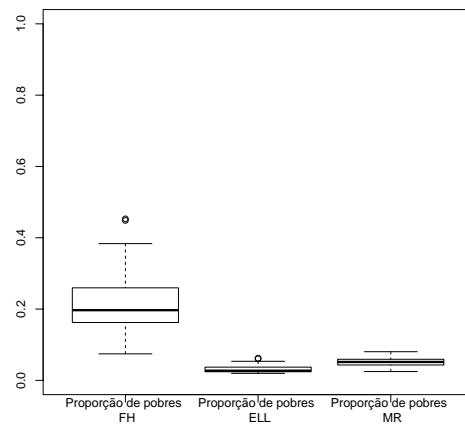
Nesta subseção são apresentadas as comparações dos métodos em termos de erro quadrático médio e vício relativos. Tais comparações foram realizadas tanto para municípios quanto para microrregiões.

Para cada método e área foram calculadas as seguintes medidas: (a) média das estimativas de pobreza: $fgt_{cd}^m = (1/400) \sum_{i=1}^{400} fgt_{icd}^m$; (b) vício relativo: $VR_{cd}^m = (1/400) \sum_{i=1}^{400} (fgt_{icd}^m - FGT_{cd}^{pop}) / FGT_{cd}^{pop}$; (c) erro quadrático médio relativo: $EQMR_{cd}^m = (1/400) \sum_{i=1}^{400} [(fgt_{icd}^m - FGT_{cd}^{pop}) / FGT_{cd}^{pop}]^2$. A quantidade fgt_{icd}^m é a estimativa do indicador c obtida pelo método m na área d para a subamostra i e FGT_{cd}^{pop} é o indicador c obtido diretamente com base na população para a área d .

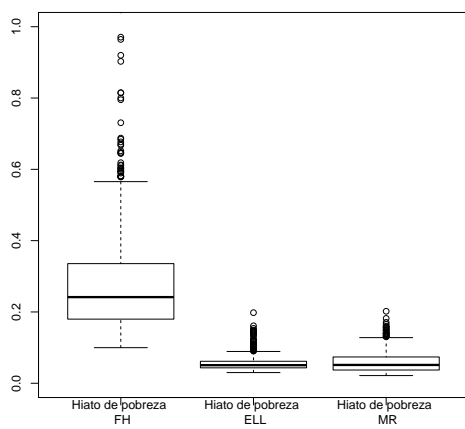
Para avaliar a variação das estimativas de proporção de pobres e hiato de pobreza fornecidas pelos métodos nas 400 subamostras, foram estimados os coeficientes de variação (CV) de tais indicadores para cada município e microrregião. Valor baixo do CV de um indicador para uma área específica indica que suas estimativas variam pouco entre as subamostras. A Figura 1 mostra que os coeficientes obtidos a partir dos métodos ELL e MR variam mais quando as áreas de interesse são os municípios, ou seja, os métodos tendem a produzir estimativas mais próximas entre si quando o tamanho de amostra aumenta e o número de áreas sem amostra diminui. Já, o método FH apresentou grande variação de estimativas tanto para municípios quanto para microrregiões.



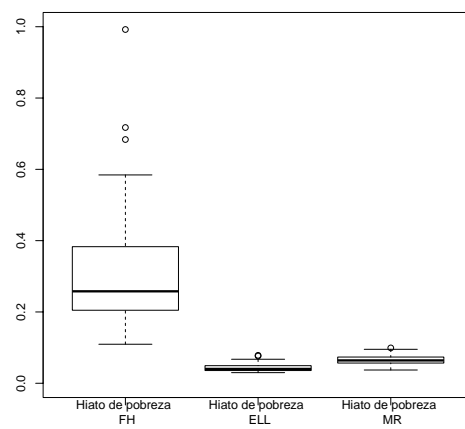
(a) Municípios



(b) Microrregiões



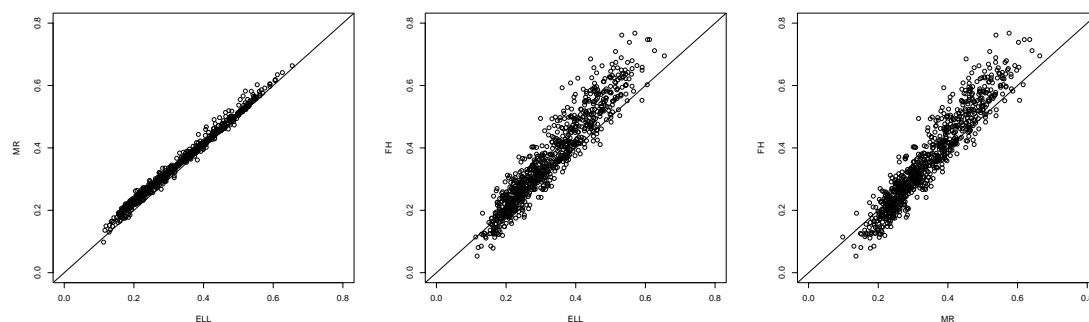
(c) Municípios



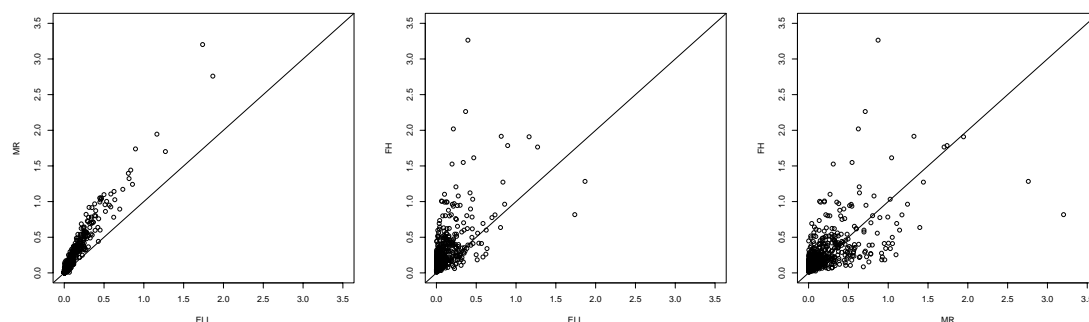
(d) Microrregiões

Figura 1: Distribuições dos coeficientes de variação estimados para os municípios e microrregiões de Minas Gerais considerando as estimativas de proporção de pobres e hiato de pobreza dadas pelos métodos FH, ELL e MR nas 400 subamostras.

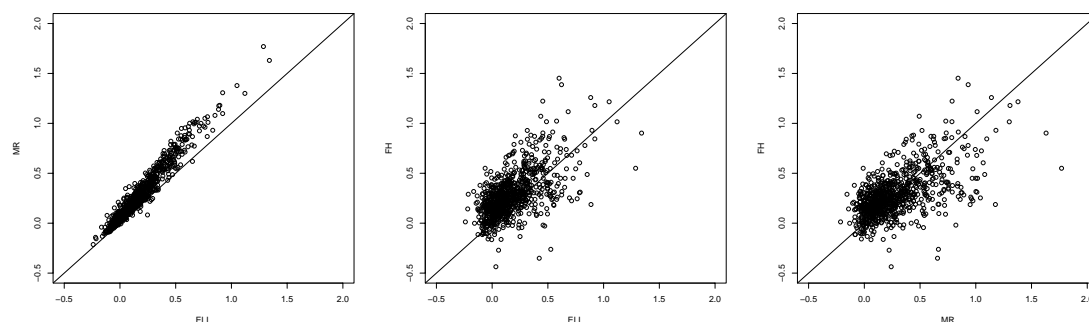
A Figura 2(a) apresenta a comparação entre as proporções de pobres (incidência de pobreza) estimadas pelos métodos FH, ELL e MR para cada um dos 853 municípios de Minas Gerais. A Figura 3 mostra as razões das proporções de pobres obtidas pelos métodos de estimação em pequenas áreas com aquelas calculadas diretamente na população para os municípios de Minas Gerais. Observa-se que os três métodos superestimaram as proporções de pobres resultantes da população. As estimativas fornecidas pelos métodos FH e MR foram, em geral, maiores que aquelas dadas pelo método ELL.



(a) Estimativas



(b) EQMs Relativos



(c) Vícios Relativos

Figura 2: Comparação entre as estimativas, EQMs relativos e vícios relativos dos indicadores de proporção de pobres obtidos pelos métodos FH, ELL e MR para os municípios de Minas Gerais.

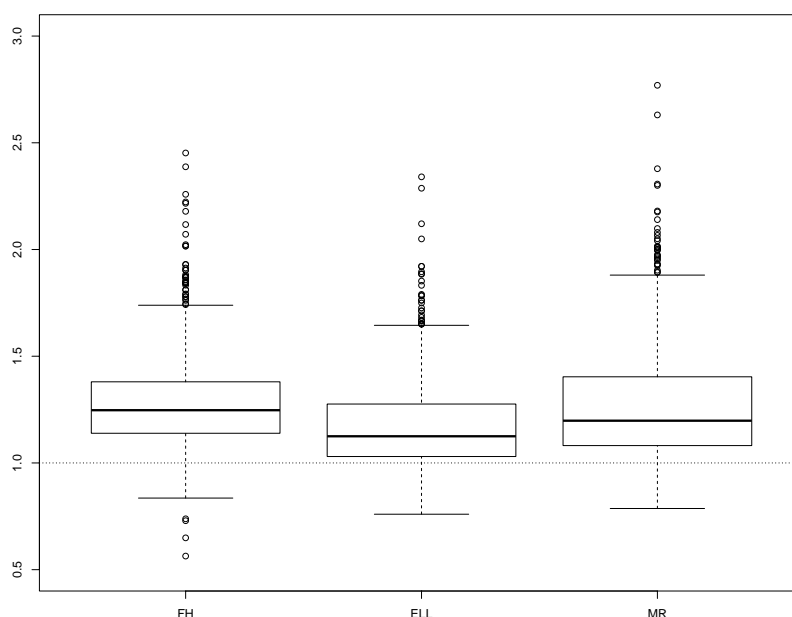


Figura 3: *Box plots* das razões entre as proporções de pobres obtidas pelos métodos de estimação em pequenas áreas e aquelas calculadas diretamente na população para os municípios de Minas Gerais.

Os métodos FH, ELL e MR apresentaram vícios relativos médios estimados de 28,11%, 17,42% e 27,49%, respectivamente, e EQMs relativos médios estimados de 21,13%, 7,71% e 15,40%, respectivamente, mostrando, portanto, que os métodos testados apresentam desempenhos diferentes com relação a vício e EQM relativos, mas todos superestimaram os valores da população. Os métodos FH e MR mostraram-se um pouco piores com relação a vício relativo, quando comparados ao método ELL. Em relação ao erro médio quadrático relativo, o método FH apresentou valores bem mais altos do que os demais métodos. Estes resultados podem ser observados na Tabela 4 e são ilustrados nas Figuras 2(b) e 2(c).

Tabela 4: Estatísticas descritivas dos EQMs relativos e dos vícios relativos dos indicadores de proporção de pobres obtidos pelos métodos de estimação em pequenas áreas para os municípios de Minas Gerais.

Resumo	Estimativa (em %)				EQM Relativo (em %)			Vício Relativo (em %)		
	População	FH	ELL	MR	FH	ELL	MR	FH	ELL	MR
Mínimo	7,32	5,33	11,29	9,81	1,20	0,04	0,05	-43,61	-24,02	-21,34
1° Quartil	18,02	24,37	22,66	25,68	6,30	0,46	1,05	13,91	3,01	8,14
Mediana	26,99	33,20	29,94	32,18	12,12	1,85	4,27	24,71	12,50	19,79
Média	29,07	35,97	32,16	34,30	21,13	7,71	15,40	28,11	17,42	27,49
3° Quartil	39,15	46,90	40,51	41,92	23,21	8,12	16,99	38,01	27,60	40,36
Máximo	65,25	76,79	65,45	66,36	326,40	186,70	320,30	145,20	134,10	177,00

O comportamento observado na Figura 2 pôde ser visto mesmo nos municípios que apareceram em todas as subamostras, como mostrado na Figura 4. Esta contém as proporções de pobres calculadas diretamente na população e as estimativas obtidas pelos métodos com seus respectivos intervalos de confiança de 95% para os municípios de Minas Gerais que apareceram nas 400 subamostras. Os extremos dos intervalos de confiança de cada município foram determinados a partir dos quantis 0,025 e 0,975 das distribuições das estimativas encontradas pelos métodos nas subamostras. Na maioria dos municípios, os intervalos dos métodos ELL e MR não contém os indicadores populacionais. O método FH apresentou estimativas muito diferentes por subamostra para estes municípios e, conseqüentemente, intervalos de confiança muito amplos que abrangeram os valores populacionais. Pode ser destacado que as estimativas dos métodos para um dado município não variam muito entre as subamostras. A estimativa obtida para o município de Belo Horizonte, representado na quarta posição desta figura, pelo método MR encontra-se mais próxima à proporção calculada na população em relação às estimativas obtidas pelos métodos FH e ELL. De maneira geral, os valores populacionais de 30,7% dos 853 municípios estavam contidos nos intervalos de confiança do método ELL e, considerando o método MR, apenas 21,2% dos intervalos de confiança municipais continham os valores populacionais. Já, considerando o método FH, cerca de 89,3% dos valores populacionais dos municípios de Minas Gerais estavam contidos nos intervalos de confiança.

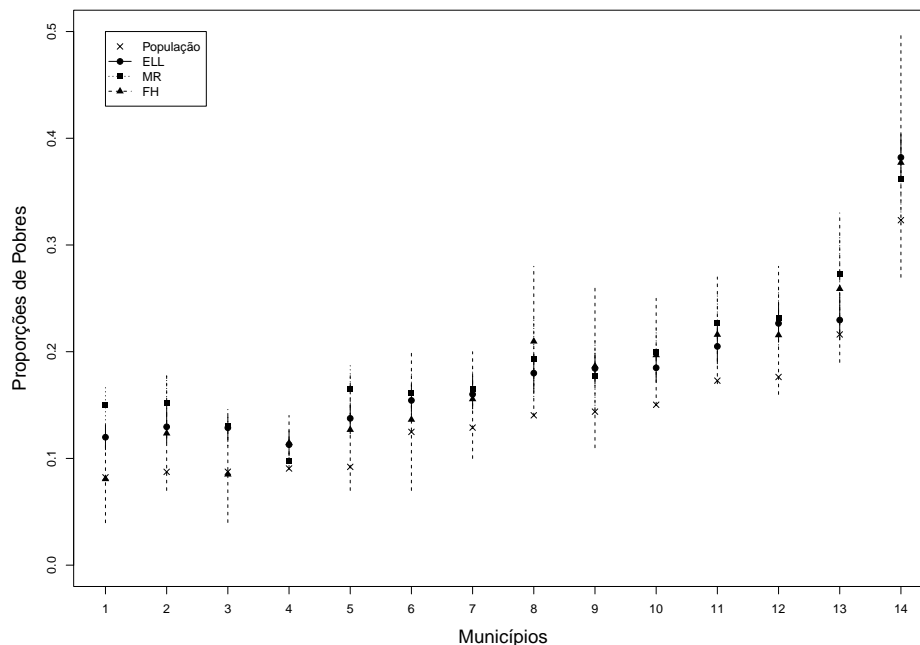
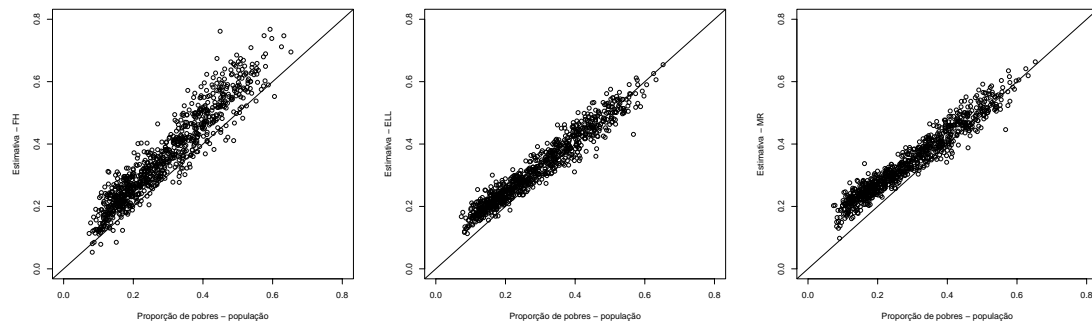
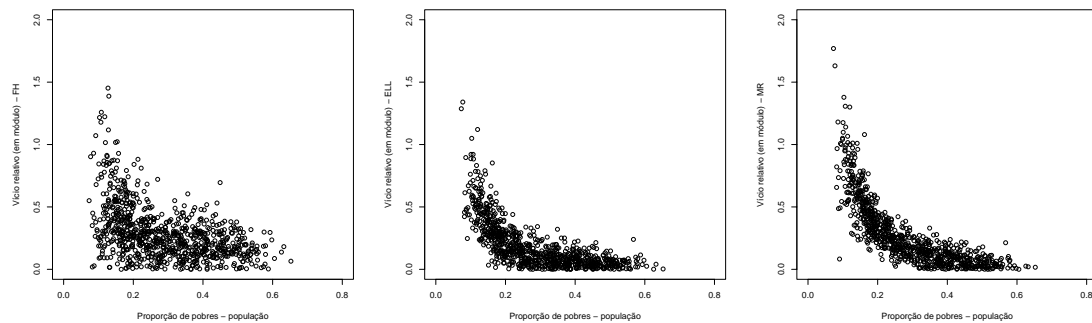


Figura 4: Proporções de pobres calculadas diretamente na população e estimativas obtidas pelos métodos FH, ELL e MR com seus respectivos intervalos de confiança de 95% para os municípios de Minas Gerais que apareceram nas 400 subamostras.

Na Figura 5, pode-se observar que as estimativas dadas pelos métodos ELL e MR tendem a ficar mais próximas dos indicadores populacionais à medida em que seus valores crescem. O método FH superestimou os valores populacionais da maioria dos municípios.



(a) Estimativa \times Proporção de pobres



(b) VR (em módulo) \times Proporção de pobres

Figura 5: Estimativas de proporções de pobres e vícios relativos (em módulo) obtidos pelos métodos FH, ELL e MR *versus* indicadores populacionais para os municípios de Minas Gerais.

No caso do indicador de hiato de pobreza, os métodos FH, ELL e MR superestimaram os valores populacionais na maioria dos municípios, como mostrado na Figura 6. Vale ressaltar que, para o hiato de pobreza, as estimativas dadas pelo método MR ficaram ainda mais distantes dos valores populacionais. A Tabela 5 evidencia este fato.

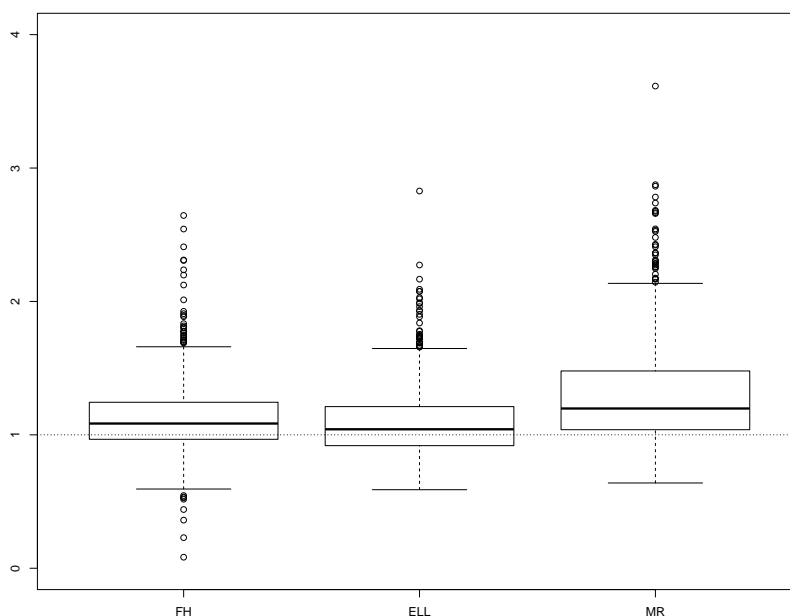


Figura 6: *Box plots* das razões entre os hiatos de pobreza obtidos pelos métodos de estimação em pequenas áreas e aqueles calculados diretamente na população para os municípios de Minas Gerais.

Tabela 5: Estatísticas descritivas dos EQMs relativos e dos vícios relativos dos hiatos de pobreza obtidos pelos métodos de estimação em pequenas áreas para os municípios de Minas Gerais.

Resumo	Estimativa (em %)				EQM Relativo (em %)			Vício Relativo (em %)		
	População	FH	ELL	MR	FH	ELL	MR	FH	ELL	MR
Mínimo	1,93	0,32	3,46	3,22	1,03	0,12	0,07	-91,68	-41,13	-36,07
1º Quartil	6,21	7,02	7,39	9,18	5,38	0,68	1,13	-3,34	-8,10	3,86
Mediana	10,35	11,12	10,54	12,38	10,45	2,03	4,63	8,49	4,18	19,76
Média	12,41	13,62	12,32	14,05	20,89	7,86	23,96	11,86	9,49	30,01
3º Quartil	17,25	19,77	16,32	17,92	21,83	5,82	23,86	24,43	21,17	47,89
Máximo	40,26	40,22	35,78	36,82	561,30	350,10	712,40	164,50	182,80	261,50

A Figura 7 apresenta a comparação dos três métodos em relação a vício relativo, EQM relativo, bem como as estimativas do indicador de hiato de pobreza para os municípios de Minas Gerais.

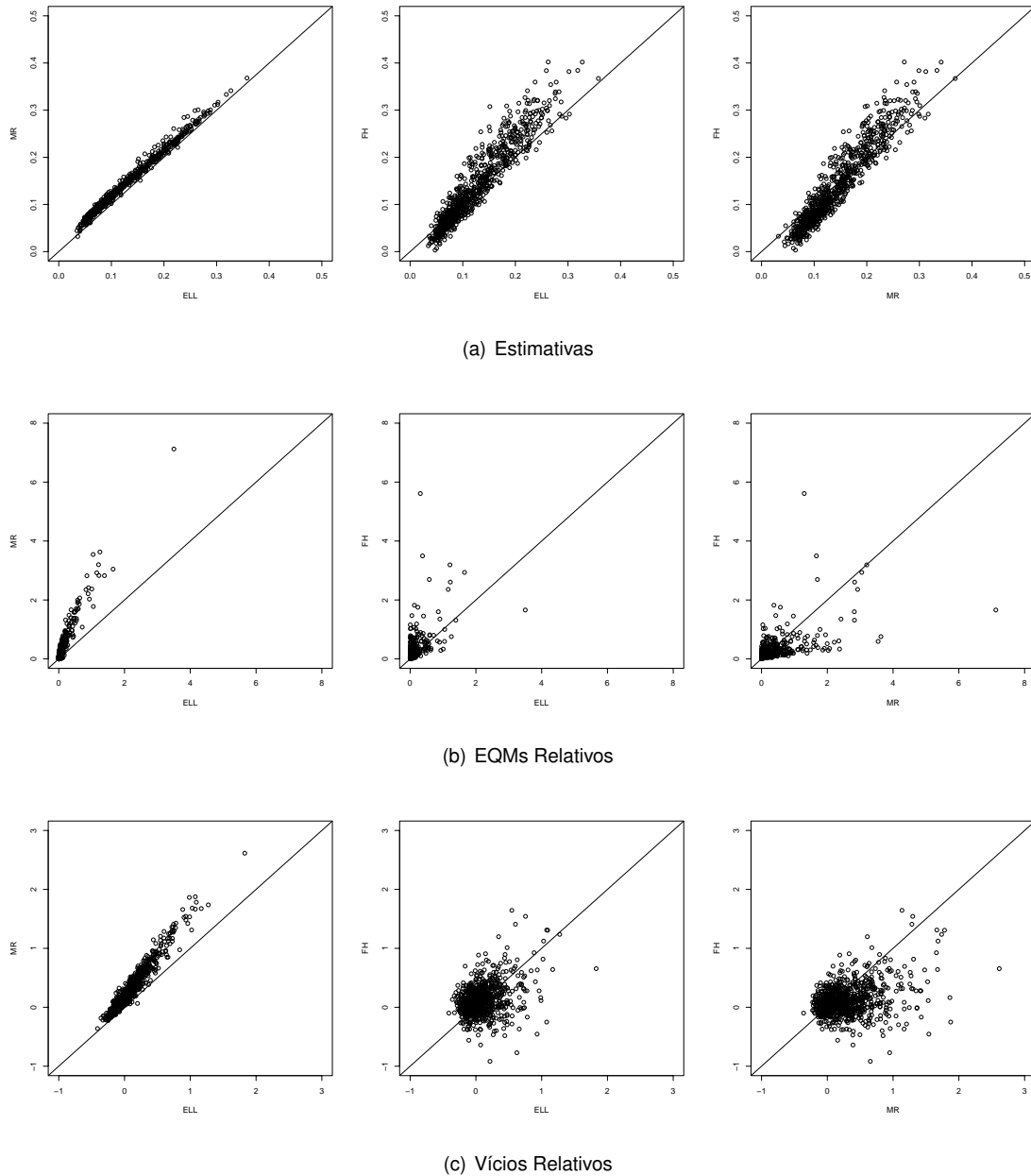


Figura 7: Comparação entre as estimativas, EQMs relativos e vícios relativos dos indicadores de hiato de pobreza obtidos pelos métodos FH, ELL e MR para os municípios de Minas Gerais.

Na Figura 8, pode-se observar os hiatos de pobreza populacionais, as estimativas obtidas pelos métodos FH, ELL e MR e seus respectivos intervalos de confiança de 95% dos municípios que apareceram em todas as subamostras. Como no caso das proporções de pobres, a maioria dos intervalos dos métodos ELL e MR não conteve os valores populacionais e o método FH apresentou amplos intervalos que contiveram estes valores. Para o município de Belo Horizonte,

representado na quarta posição do gráfico, os métodos apresentaram estimativas muito próximas ao valor encontrado na população. Considerando todos os 853 municípios de Minas Gerais, os hiatos de pobreza populacionais estavam contidos em 97,5%, 37,8% e 29,7% dos intervalos de confiança municipais dos métodos FH, ELL e MR, respectivamente.

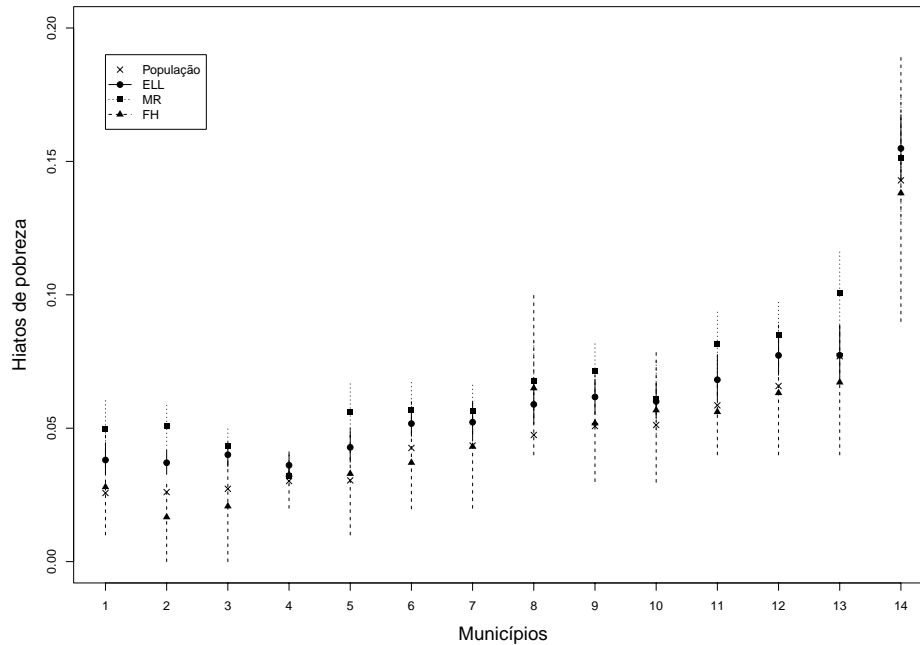
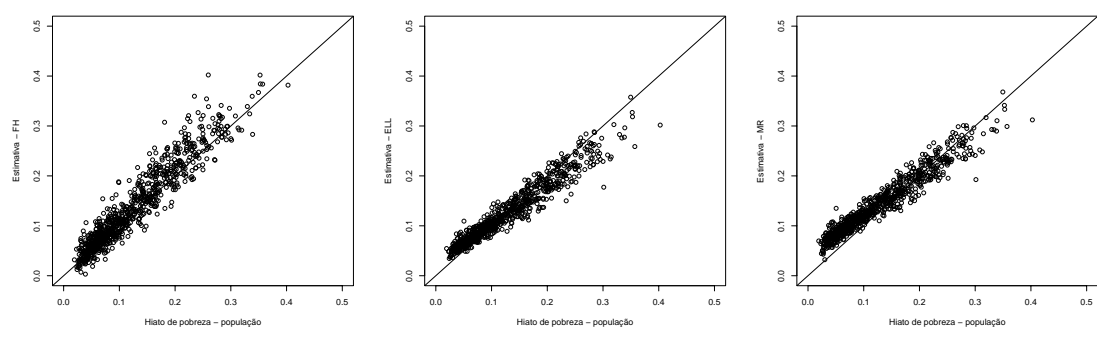
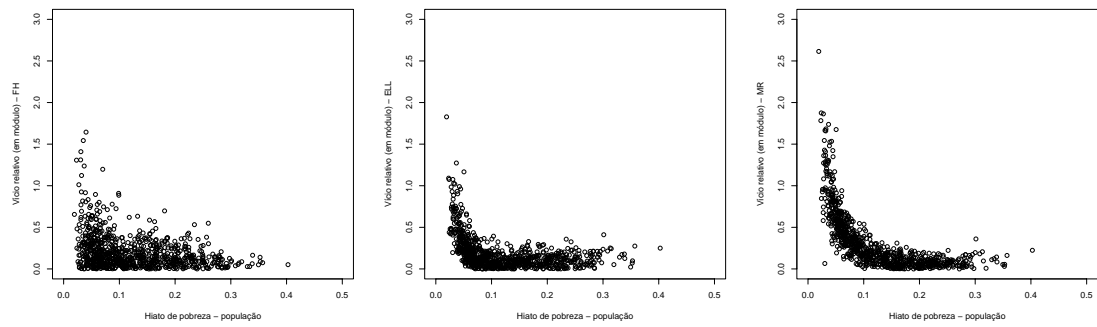


Figura 8: Hiato de pobreza calculados diretamente na população e estimativas obtidas pelos métodos FH, ELL e MR com seus respectivos intervalos de confiança de 95% para os municípios de Minas Gerais que apareceram nas 400 subamostras.

Através da Figura 9(b), observa-se que os vícios relativos são maiores para os menores valores do hiato de pobreza. Conforme a Figura 9(a), os métodos ELL e MR tendem a superestimar o hiato de pobreza nos municípios que apresentam valores mais baixos desse indicador e a subestimá-lo quando esses são altos. Pode ser observado, na Figura 9(a), que as estimativas de hiato de pobreza obtidas pelo método FH são maiores que os valores populacionais para a maioria dos municípios.



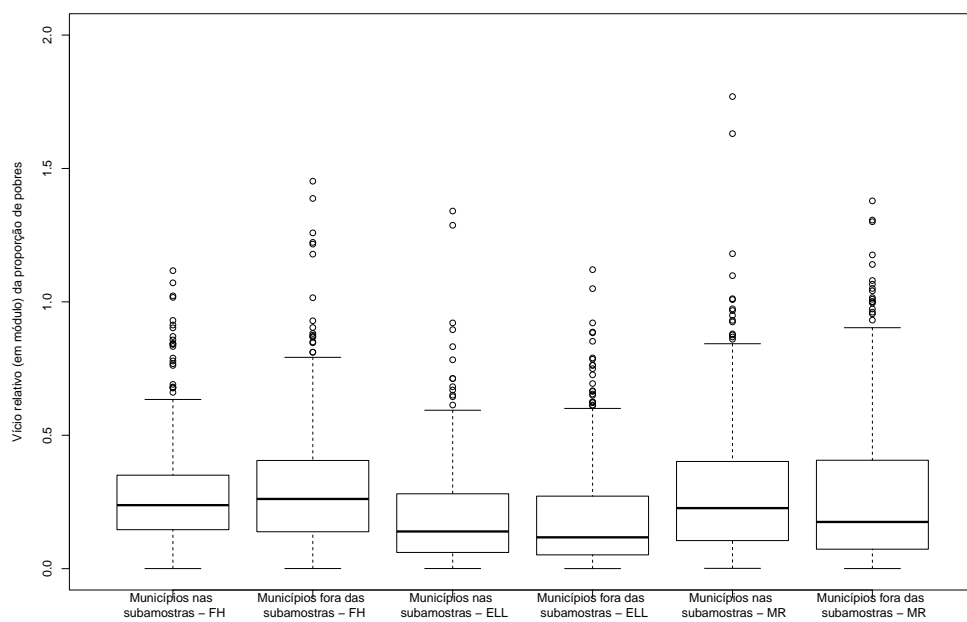
(a) Estimativa \times Hiato de pobreza



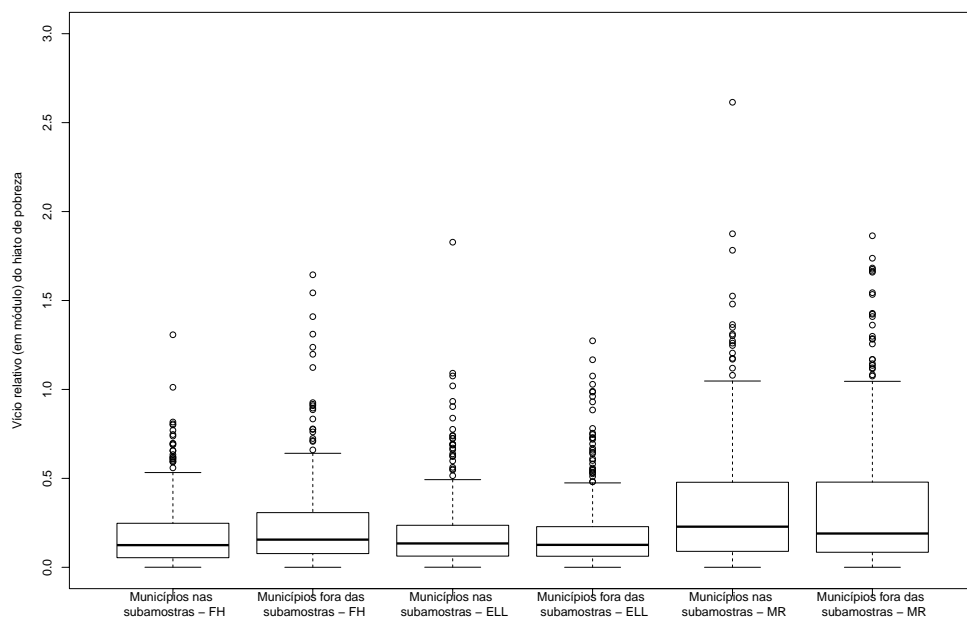
(b) VR (em módulo) \times Hiato de pobreza

Figura 9: Estimativas de hiatos de pobreza e vícios relativos (em módulo) obtidos pelos métodos FH, ELL e MR *versus* indicadores populacionais para os municípios de Minas Gerais.

A Figura 10 mostra as distribuições dos vícios relativos (em módulo) das proporções de pobres e hiatos de pobreza dados pelos métodos FH, ELL e MR nos municípios, separando-os em dois grupos: os que apareceram em pelo menos uma subamostra e aqueles não encontrados nas 400 subamostras. Os *box plots* indicam que não houve diferença entre as distribuições dos vícios relativos estimados das medidas de pobreza nos dois grupos de municípios para os métodos ELL e MR. No caso do método FH, os vícios relativos estimados do grupo dos municípios que apareceram em pelo menos uma subamostra são um pouco menores em relação ao outro grupo.



(a) Proporção de pobres

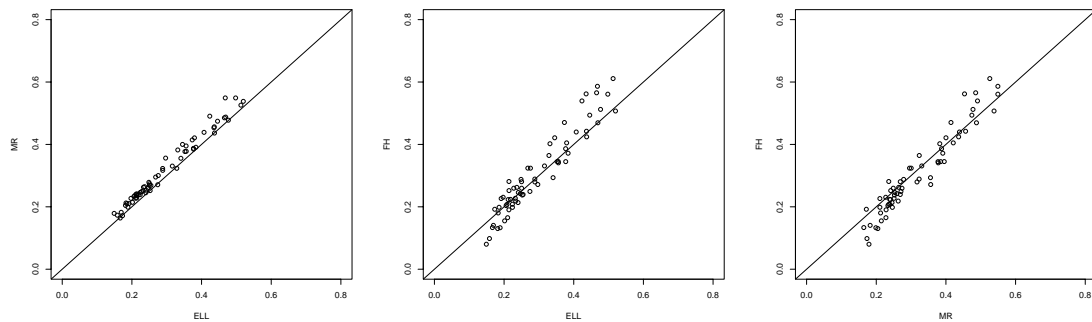


(b) Hiato de pobreza

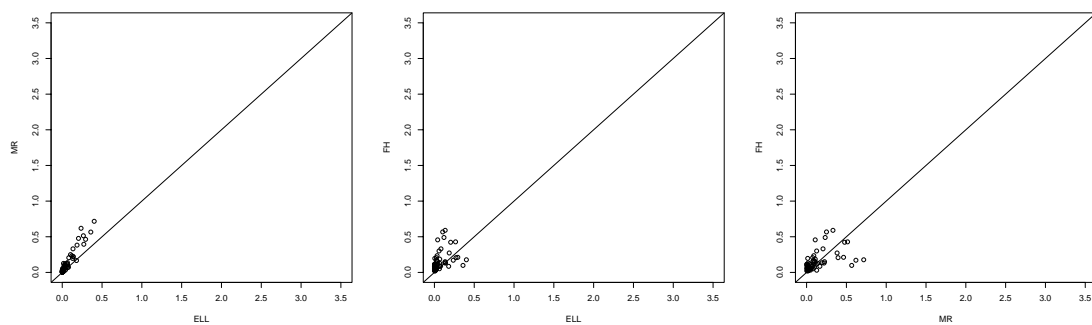
Figura 10: Distribuições dos vícios relativos (em módulo) das proporções de pobres e hiatos de pobreza dados pelos métodos FH, ELL e MR nos municípios dentro e fora das subamostras.

O pequeno número de domicílios amostrados, ou até mesmo a ausência de áreas na amostra, pode ter ocasionado o viés verificado nas estimativas municipais. Para verificar essa suposição, compararam-se os métodos FH, ELL e MR considerando o agrupamento de municípios em microrregiões.

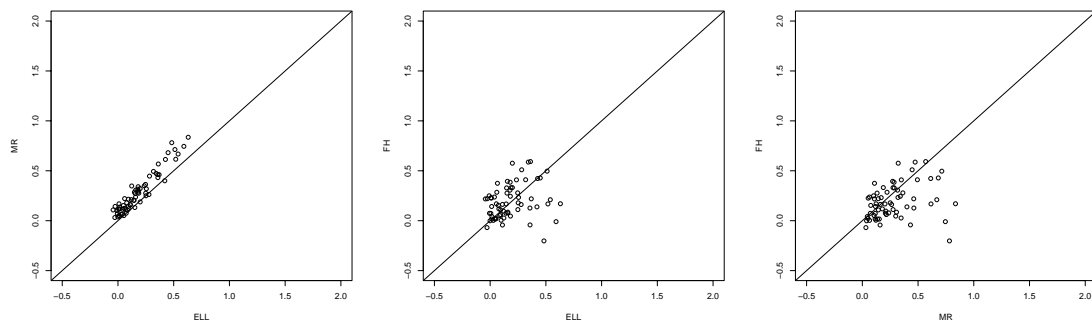
As Figuras 11(a), 11(b) e 11(c) mostram, respectivamente, as proporções de pobres estimadas, os EQMs e vícios relativos obtidos pelos métodos FH, ELL e MR para cada uma das 66 microrregiões de Minas Gerais.



(a) Estimativas



(b) EQMs Relativos



(c) Vícios Relativos

Figura 11: Comparação entre as estimativas, EQMs relativos e vícios relativos dos indicadores de proporção de pobres obtidos pelos métodos FH, ELL e MR para as microrregiões de Minas Gerais.

A Figura 12 apresenta as razões entre as proporções de pobres obtidas pelos três métodos e aquelas calculadas diretamente na população. Assim como no caso dos municípios, os métodos superestimaram as incidências de pobreza. No entanto, para o caso das microrregiões, houve redução nos valores dos EQMs e vícios relativos. O método FH apresentou, em média, os valores mais altos de EQMs relativos e o método MR apresentou o vício relativo médio mais alto. A Tabela 6 ilustra esses resultados.

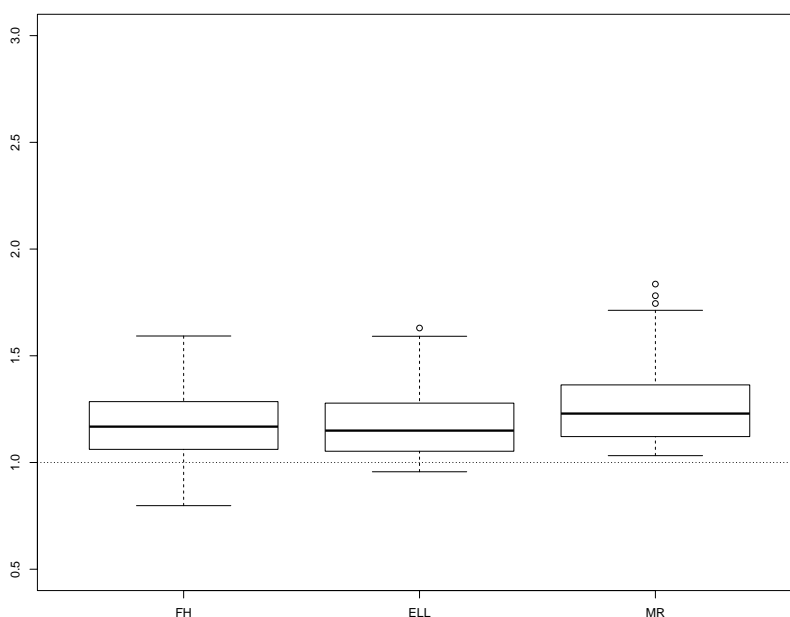


Figura 12: *Box plots* das razões entre as proporções de pobres obtidas pelos métodos de estimação em pequenas áreas e aquelas calculadas diretamente na população para as microrregiões de Minas Gerais.

Tabela 6: Estatísticas descritivas dos EQMs relativos e dos vícios relativos dos indicadores de proporção de pobres obtidos pelos métodos de estimação em pequenas áreas para as microrregiões de Minas Gerais.

Resumo	Estimativa (em %)				EQM Relativo (em %)			Vício Relativo (em %)		
	População	FH	ELL	MR	FH	ELL	MR	FH	ELL	MR
Mínimo	9,93	8,02	14,91	16,42	2,18	0,04	0,20	-20,22	-4,34	3,21
1º Quartil	16,90	21,49	21,39	23,66	5,75	0,37	1,81	6,46	5,36	12,20
Mediana	21,98	27,57	25,35	27,24	9,40	2,33	5,71	16,81	14,95	22,93
Média	26,35	30,77	29,56	31,84	14,29	6,38	12,61	18,99	18,45	28,25
3º Quartil	36,31	39,84	37,59	39,42	17,12	7,54	13,46	28,37	27,24	36,02
Máximo	50,54	61,12	51,99	54,92	59,08	40,16	71,74	59,27	63,04	83,61

A partir da Figura 13, observam-se as estimativas das proporções de pobres obtidas pelos métodos FH, ELL e MR com seus respectivos intervalos de confiança de 95% e as proporções de pobres populacionais para as microrregiões de Minas Gerais. Os indicadores populacionais das microrregiões estão contidos em 97,0% dos intervalos de confiança do método FH, por serem mais amplos, em 22,7% dos intervalos de confiança do método ELL e em 13,6% dos intervalos do método MR. Isso ocorre, principalmente, quando as proporções de pobres são mais elevadas. Em algumas microrregiões, nota-se, também, maior proximidade entre as estimativas dos métodos, como por exemplo na 13ª microrregião do gráfico a qual inclui o município de Belo Horizonte. As microrregiões que menos apareceram nas subamostras, 270 e 155 vezes, correspondem às posições 50 e 66, respectivamente.

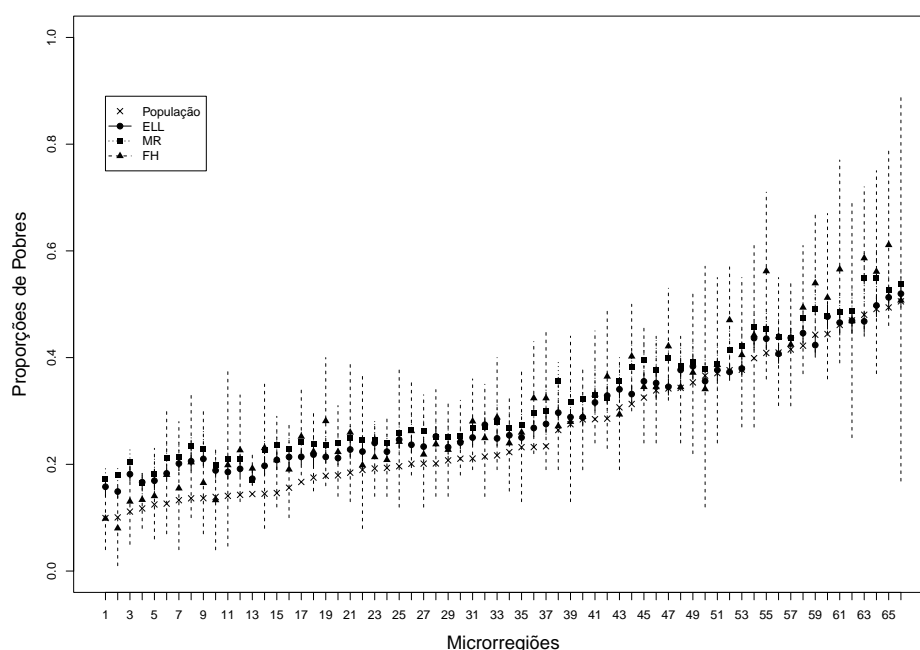
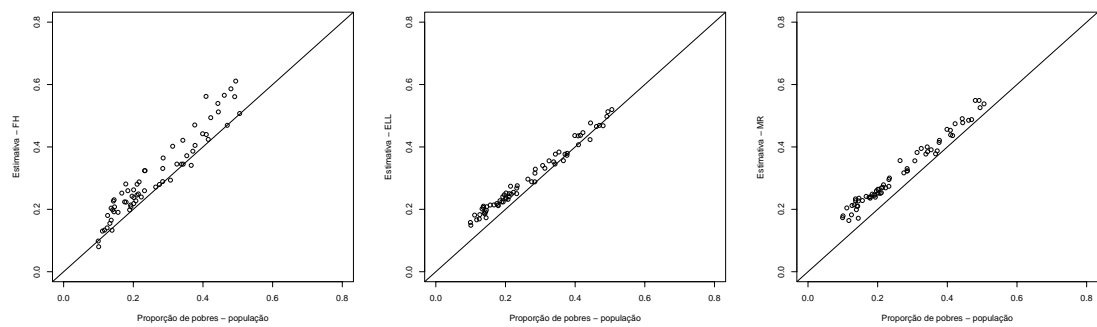
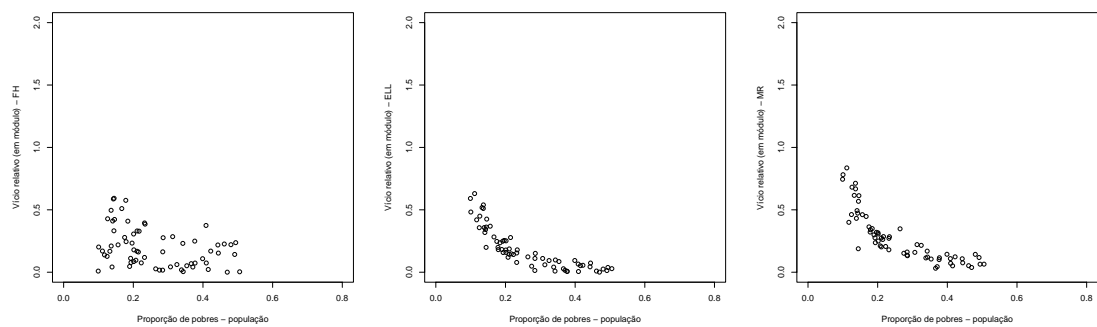


Figura 13: Proporções de pobres calculadas diretamente na população e estimativas obtidas pelos métodos FH, ELL e MR com seus respectivos intervalos de confiança de 95% para as microrregiões de Minas Gerais.

Nas microrregiões que apresentaram maior incidência de pobreza, as estimativas dadas pelos métodos ELL e MR se aproximaram mais dos valores populacionais como pode ser visto na Figura 14(a). As Figuras 14(a) e 14(b) mostram que as estimativas do método FH não seguiram esse mesmo padrão, os comportamentos dos módulos dos vícios relativos das microrregiões mais pobres e das menos pobres são parecidos.



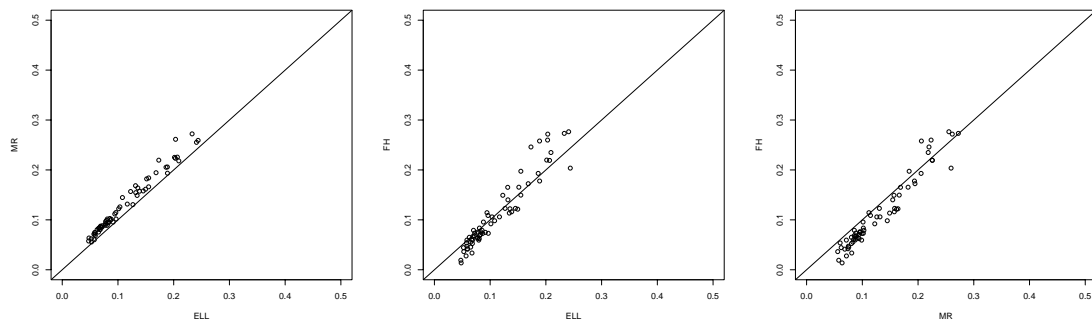
(a) Estimativa \times Proporção de pobres



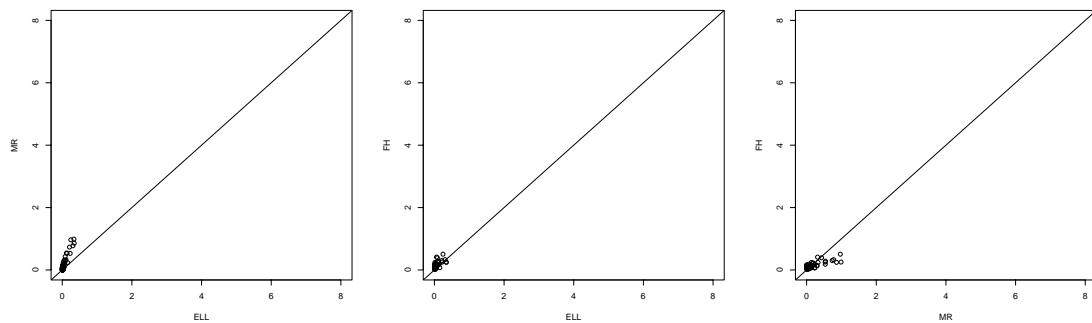
(b) VR (em módulo) \times Proporção de pobres

Figura 14: Estimativas de proporções de pobres e vícios relativos (em módulo) obtidos pelos métodos FH, ELL e MR *versus* indicadores populacionais para as microrregiões de Minas Gerais.

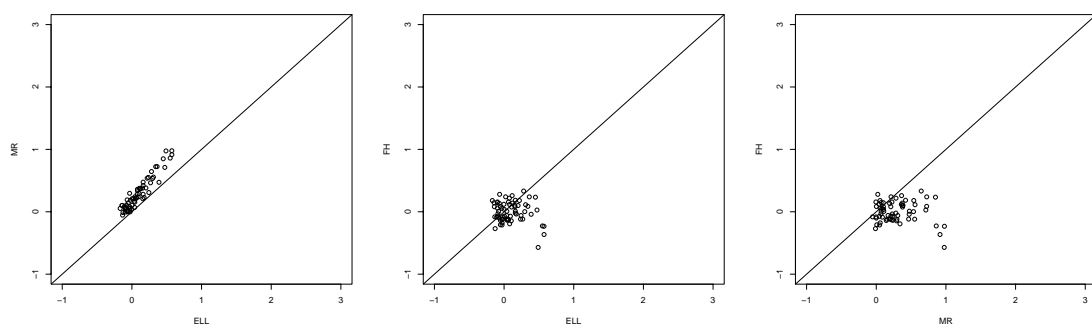
As Figuras 15(b), 15(c) mostram que houve redução nos valores dos EQMs e vícios relativos obtidos pelos métodos em relação ao que foi verificado para o hiato de pobreza nos municípios de Minas Gerais. Em média, o método MR foi o que forneceu maiores valores de EQM e vício relativos e o método ELL o que forneceu os menores valores dessas medidas, como pode ser visto na Figura 15(a) e na Tabela 7.



(a) Estimativas



(b) EQMs Relativos



(c) Vícios Relativos

Figura 15: Comparação entre as estimativas, EQMs relativos e vícios relativos dos indicadores de hiato de pobreza obtidos pelos métodos FH, ELL e MR para as microrregiões de Minas Gerais.

Tabela 7: Estatísticas descritivas dos EQMs relativos e dos vícios relativos dos hiatos de pobreza obtidos pelos métodos de estimação em pequenas áreas para as microrregiões de Minas Gerais.

Resumo	Estimativa (em %)				EQM Relativo (em %)			Vício Relativo (em %)		
	População	FH	ELL	MR	FH	ELL	MR	FH	ELL	MR
Mínimo	3,01	1,38	4,74	5,59	2,06	0,12	0,15	-57,06	-16,85	-5,32
1º Quartil	5,73	6,11	6,89	8,55	5,21	0,41	1,03	-11,68	-5,26	7,34
Mediana	8,27	8,14	8,97	10,18	7,81	1,18	5,68	0,48	5,64	22,48
Média	11,12	11,19	11,23	12,99	12,34	4,83	16,43	-0,15	9,81	29,30
3º Quartil	15,78	14,95	14,85	16,57	16,22	4,13	18,15	11,97	19,47	41,13
Máximo	25,79	27,66	24,38	27,23	50,67	34,39	98,97	33,30	57,40	97,86

As estimativas dadas pelos métodos ficaram mais próximas dos valores populacionais para as microrregiões do que para os municípios, como pode ser visto na Figura 16. Tal fato pôde ser observado também para a incidência de pobreza.

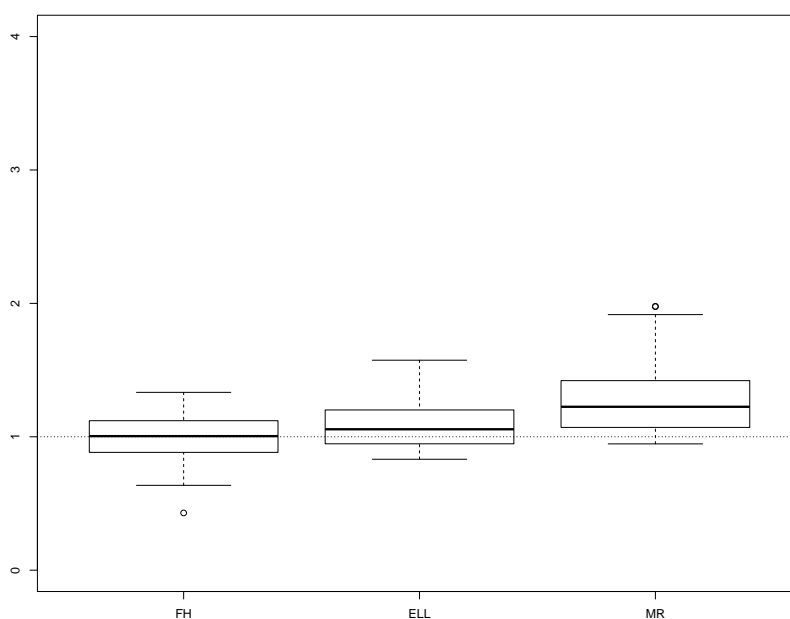
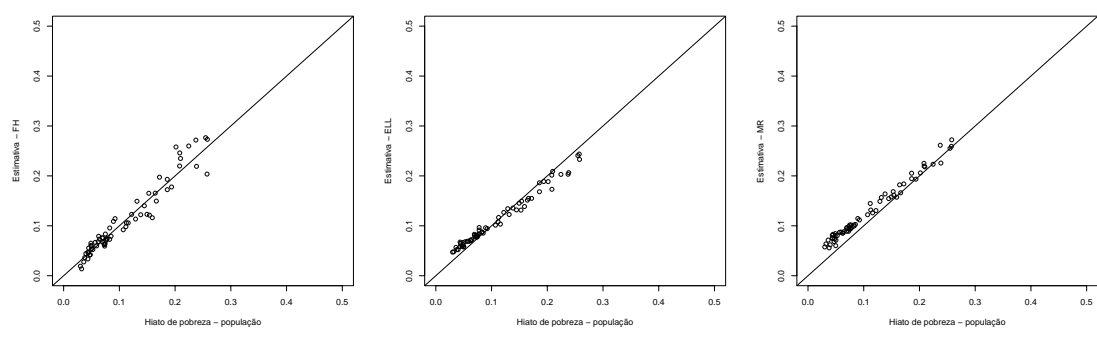
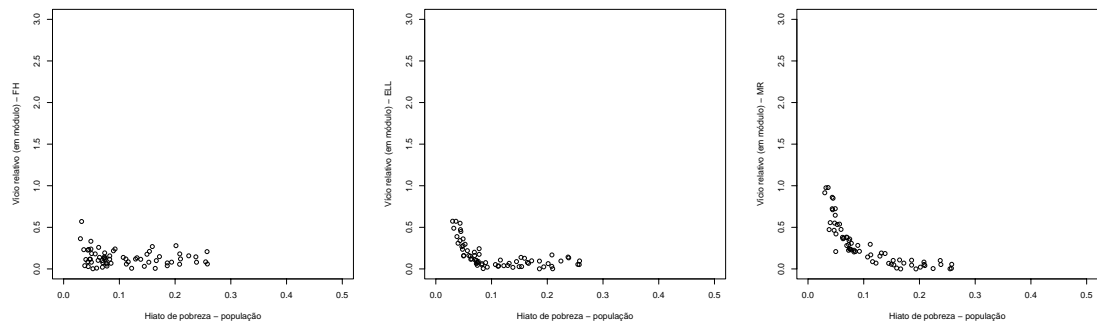


Figura 16: *Box plots* das razões entre os hiatos de pobreza obtidos pelos métodos de estimação em pequenas áreas e aqueles calculados diretamente na população para as microrregiões de Minas Gerais.

O método ELL subestimou o hiato de pobreza nas microrregiões que têm valores altos desse indicador. Já o método MR superestimou os hiatos de pobreza na maioria das áreas. Isto pode ser visto nas Figuras 17(a) e 18. Considerando estes dois métodos, os vícios relativos são maiores para os menores valores do indicador, conforme a Figura 17(b). Os módulos dos vícios relativos obtidos pelo método FH não diferem muito em relação às microrregiões que têm menores ou maiores hiatos de pobreza.



(a) Estimativa × Hiato de pobreza



(b) VR (em módulo) × Hiato de pobreza

Figura 17: Estimativas dos hiatos de pobreza e vícios relativos (em módulo) obtidos pelos métodos FH, ELL e MR *versus* indicadores populacionais para as microrregiões de Minas Gerais.

No caso do hiato de pobreza, as proporções de indicadores populacionais contidos nos intervalos de confiança de 95% foram iguais a 37,9% para o ELL e 36,4% para o MR. Os intervalos de confiança obtidos pelo método FH são muito amplos, assim os hiatos de pobreza de todas as microrregiões estão contidos nestes intervalos. Nota-se que na microrregião que inclui Belo Horizonte, identificada na 14ª posição da Figura 18, as estimativas dos métodos ficaram bastante próximas e seus intervalos de confiança tiveram pouca amplitude. As microrregiões que menos apareceram nas subamostras, 270 e 155 vezes, correspondem às posições 50 e 65, respectivamente. Nestas duas áreas, o método MR se aproximou mais do hiato de pobreza da população do que o método FH e o ELL.

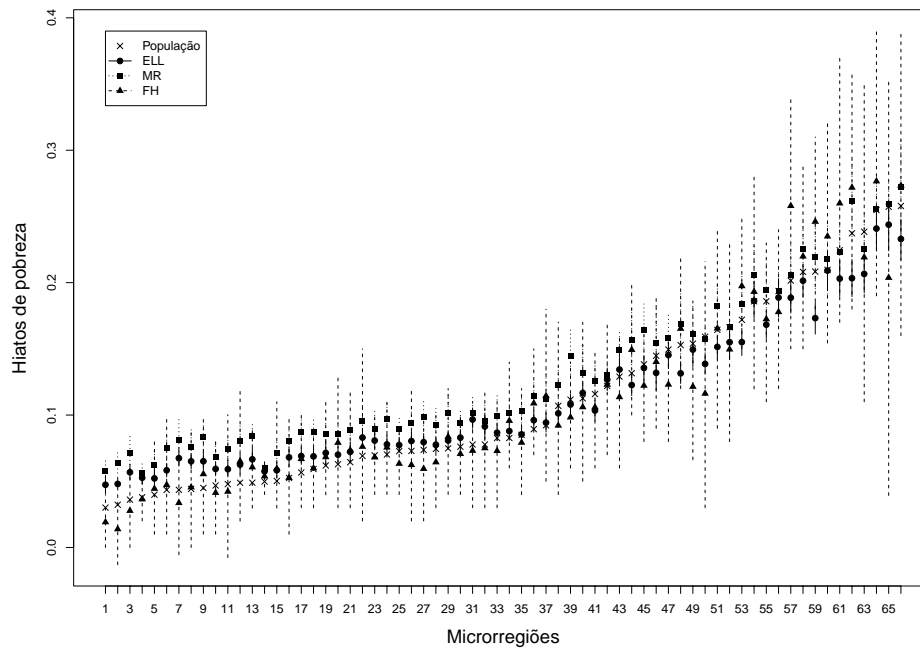


Figura 18: Hiatos de pobreza calculados diretamente na população e estimativas obtidas pelos métodos FH, ELL e MR com seus respectivos intervalos de confiança de 95% para as microrregiões de Minas Gerais.

Ordenação de áreas

Nesta subseção, comparam-se as ordenações dos municípios e microrregiões a partir das proporções de pobres estimadas por cada um dos três métodos com as ordenações obtidas a partir dos valores calculados na população.

Tais ordenações podem ser comparadas por meio do coeficiente de correlação de Spearman. Mais detalhes sobre o assunto são apresentados em Lehmann e D’Abrera (1975).

Sejam os postos das características associadas denotados por $(R_1, S_1), (R_2, S_2), \dots, (R_N, S_N)$. O coeficiente de correlação de postos de Spearman é definido por:

$$r_s = \frac{\sum (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum (R_i - \bar{R})^2 \sum (S_i - \bar{S})^2}},$$

em que $\bar{R} = \sum R_i/N$ e $\bar{S} = \sum S_i/N$.

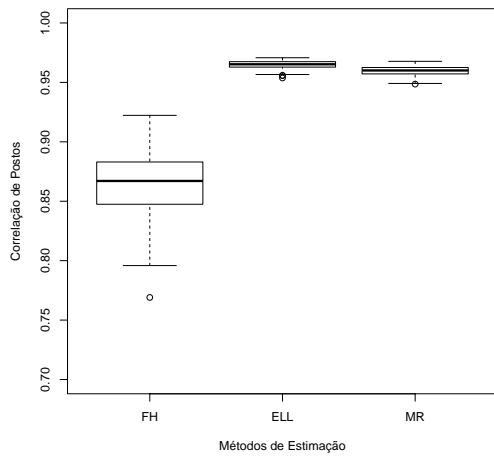
Este coeficiente pode ser usado como uma medida de concordância de duas ordenações de objetos segundo dois critérios diferentes. Observe que quando as duas ordenações coincidem exatamente $r_s = 1$ e que quando as ordenações são exatamente contrárias $r_s = -1$.

Para cada subamostra, foi medida a concordância entre a ordenação dos municípios ou microrregiões obtida com base nas estimativas FH, ELL e MR com a da população por meio do coeficiente de correlação de Postos de Spearman. Sejam, para cada subamostra i , $i = 1, \dots, 400$, os coeficientes de correlação de postos entre

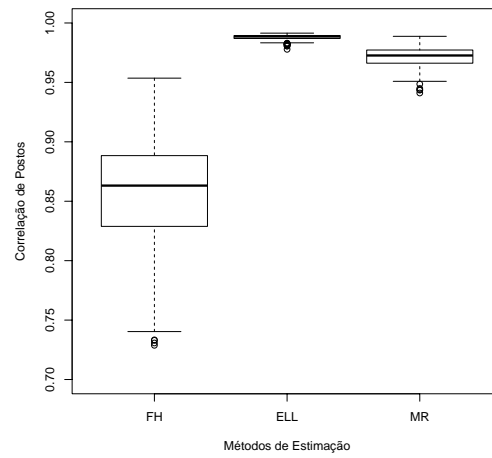
- $r_i(FH, POP)$ - as estimativas da proporção de pobres estimadas pelo método FH e os valores correspondentes calculados na população;
- $r_i(ELL, POP)$ - as estimativas da proporção de pobres estimadas pelo método ELL e os valores correspondentes calculados na população;
- $r_i(MR, POP)$ - as estimativas da proporção de pobres estimadas pelo método MR e os valores correspondentes calculados na população;
- $r_i(FH, ELL)$ - as estimativas da proporção de pobres estimadas pelos métodos FH e ELL.
- $r_i(FH, MR)$ - as estimativas da proporção de pobres estimadas pelos métodos FH e MR.
- $r_i(ELL, MR)$ - as estimativas da proporção de pobres estimadas pelos métodos ELL e MR.

A Figura 19 mostra os *box plots* das correlações de postos $r_i(FH, POP)$, $r_i(ELL, POP)$ e $r_i(MR, POP)$ para as 400 subamostras por (a) municípios e (b) microrregiões. É possível ver que há bastante concordância entre as ordenações definidas a partir dos valores estimados para os municípios de Minas Gerais pelos métodos ELL e MR e dos valores populacionais, sendo quase todos os valores das correlações de postos maiores que 0,95. O mesmo comportamento não é observado em relação ao método FH cujas correlações de postos variam de 0,76 a 0,93. Em geral, há maior concordância entre a ordenação dos municípios obtida a partir dos valores populacionais com a obtida a partir do método ELL do que com a obtida pelo método MR. Considerando as microrregiões, há uma maior diferença entre as correlações de postos obtidas para o método ELL em relação às obtidas considerando o método MR. Nota-se que, para o primeiro método, as correlações variam entre 0,98 e 0,99 enquanto que, para o segundo método, estes valores variam entre 0,94 e 0,99.

Os pares $r_i(FH, POP)$, $r_i(ELL, POP)$ e $r_i(MR, POP)$ obtidos para as 400 subamostras são apresentados na Figura 20. Nos gráficos (a) e (b), a concentração de pontos abaixo da bissetriz indica que há maior concordância entre a ordenação dos municípios induzida pelo método ELL com a obtida na população.

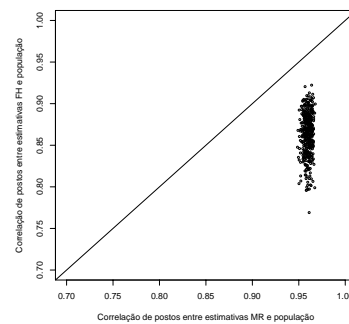
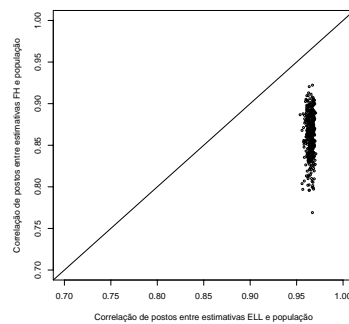
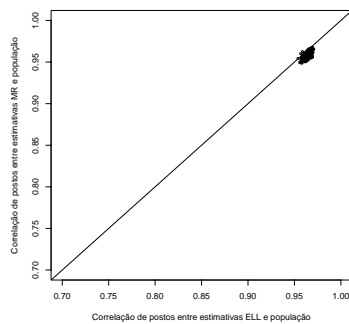


(a) Municípios

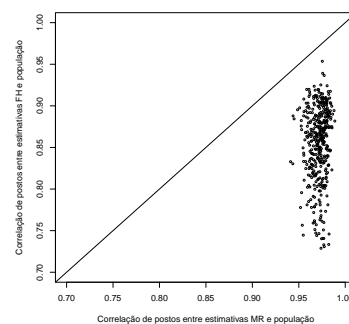
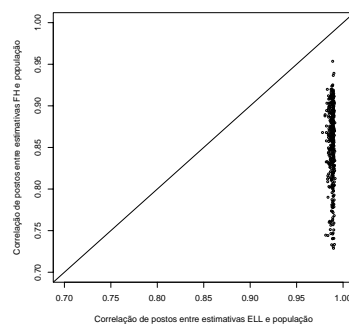
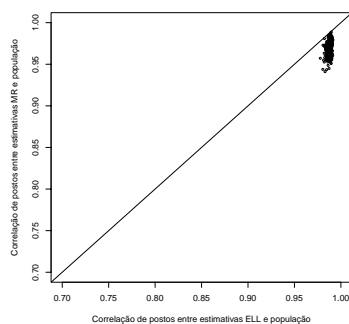


(b) Microrregiões

Figura 19: *Box plots* das correlações de postos entre estimativas de incidência de pobreza obtidas pelos métodos FH, ELL e MR e os valores populacionais considerando municípios e microrregiões.



(a) Municípios



(b) Microrregiões

Figura 20: Gráficos de dispersão das correlações de postos entre estimativas de incidência de pobreza obtidas pelos métodos FH, ELL e MR e os valores populacionais considerando municípios e microrregiões.

Designando-se por $r(FH, POP)$, $r(ELL, POP)$ e $r(MR, POP)$ os valores dos parâmetros que definem a correlação entre os postos das estimativas e dos valores populacionais, foram executados testes de Wilcoxon emparelhados com base nos pares observados $(r_i(ELL, POP), r_i(MR, POP))$, $(r_i(FH, POP), r_i(ELL, POP))$ e $(r_i(FH, POP), r_i(MR, POP))$, $i = 1, \dots, 400$ para testar as hipóteses $H_0 : r(ELL, POP) = r(MR, POP)$, $H_0 : r(FH, POP) = r(ELL, POP)$ e $H_0 : r(FH, POP) = r(MR, POP)$. Tanto para os municípios quanto para as microrregiões, as probabilidades de significância dos testes foram praticamente 0, mostrando que há diferenças significantes entre as correlações de postos dos resultados populacionais com os estimados a partir dos métodos FH, ELL e MR. Os intervalos de confiança obtidos para a diferença $r(ELL, POP) - r(MR, POP)$, $r(FH, POP) - r(ELL, POP)$ e $r(FH, POP) - r(MR, POP)$ nos municípios foram, respectivamente, (0,0045; 0,0058), (0,0957; 0,1016) e (0,0905; 0,0966), mostrando que a concordância de ordenação pelo método ELL com a ordenação obtida na população é significativamente maior que a obtida pelo método MR. No caso das microrregiões, os intervalos de confiança obtidos foram, respectivamente, (0,0146; 0,0167), (0,1199; 0,1311) e (0,1034; 0,1151). Deve-se ressaltar que há uma diferença significativamente maior entre os métodos ELL e MR quando consideram-se as microrregiões. No caso dos municípios, essa diferença, apesar de ser significativa, não é relevante.

A Tabela 8 apresenta resumos estatísticos das 400 replicações das correlações de postos $r_i(FH, POP)$, $r_i(ELL, POP)$, $r_i(MR, POP)$, $r_i(ELL, MR)$, $r_i(FH, ELL)$ e $r_i(FH, MR)$ entre as estimativas FH, ELL e MR com os valores populacionais e entre as estimativas FH, ELL e MR para municípios e microrregiões.

Tabela 8: Distribuição da correlação de postos de Spearman por municípios e microrregiões.

Áreas	Correlação de Postos	Min.	1ºQu.	Mediana	Média	3ºQu.	Max.
Municípios	$r_i(FH, POP)$	0,7691	0,8475	0,8671	0,8635	0,8830	0,9223
	$r_i(ELL, POP)$	0,9538	0,9629	0,9652	0,9649	0,9674	0,9707
	$r_i(MR, POP)$	0,9486	0,9572	0,9600	0,9597	0,9625	0,9677
	$r_i(ELL, MR)$	0,9744	0,9860	0,9889	0,9884	0,9910	0,9964
	$r_i(FH, ELL)$	0,7773	0,8490	0,8663	0,8638	0,8817	0,9174
	$r_i(FH, MR)$	0,7792	0,8493	0,8655	0,8635	0,8807	0,9172
Microrregiões	$r_i(FH, POP)$	0,7288	0,8289	0,8632	0,8547	0,8883	0,9536
	$r_i(ELL, POP)$	0,9779	0,9870	0,9886	0,9880	0,9894	0,9914
	$r_i(MR, POP)$	0,9411	0,9662	0,9727	0,9713	0,9772	0,9888
	$r_i(ELL, MR)$	0,9240	0,9686	0,9747	0,9738	0,9800	0,9932
	$r_i(FH, ELL)$	0,7182	0,8176	0,8530	0,8453	0,8770	0,9423
	$r_i(FH, MR)$	0,7433	0,8358	0,8655	0,8590	0,8904	0,9465

Na Tabela 9 observam-se os intervalos de confiança empíricos de 95% para as correlações de postos $r(FH, POP)$, $r(ELL, POP)$, $r(MR, POP)$, $r(ELL, MR)$, $r(FH, ELL)$ e $r(FH, MR)$ para municípios e microrregiões.

Tabela 9: Intervalos de confiança de 95% para os coeficientes de correlação de postos por municípios e microrregiões.

Áreas	Correlação de Postos	2,5%	97,5%
Municípios	$r(FH, POP)$	0,8038	0,9052
	$r(ELL, POP)$	0,9582	0,9693
	$r(MR, POP)$	0,9517	0,9658
	$r(ELL, MR)$	0,9790	0,9944
	$r(FH, ELL)$	0,8063	0,9070
	$r(FH, MR)$	0,8092	0,9032
Microrregiões	$r(FH, POP)$	0,7465	0,9189
	$r(ELL, POP)$	0,9829	0,9906
	$r(MR, POP)$	0,9530	0,9841
	$r(ELL, MR)$	0,9540	0,9894
	$r(FH, ELL)$	0,7425	0,9088
	$r(FH, MR)$	0,7659	0,9185

Com relação à ordenação de municípios e microrregiões, os valores dos coeficientes de correlação de postos mostraram que, no caso de municípios, há maior concordância entre os métodos ELL e MR do que entre cada um dos três métodos e a população. A partir dos intervalos de confiança de 95%, verificou-se que há pouca diferença entre os métodos de estimação ELL e MR com respeito ao poder de ordenar corretamente os municípios, baseado nas estimativas de proporção de pobres. Considerando as microrregiões, há maior concordância entre o método ELL e a população. Os mesmos resultados foram obtidos para as estimativas de hiato de pobreza, ver Apêndice B.

Conclusões

A principal proposta deste artigo foi comparar abordagens para estimação em pequenas áreas disponíveis na literatura atual. Para tanto, foram selecionadas, utilizando o desenho amostral da POF, 400 subamostras de uma população baseada nos dados do Estado de Minas Gerais do Censo Demográfico de 2010. As três metodologias foram aplicadas aos dados das subamostras e as estimativas das medidas de pobreza obtidas pelos métodos foram comparadas com os valores calculados diretamente na população em termos de vício e EQM relativos e ordenação de áreas.

Vale ressaltar que o método descrito em Molina e Rao (2010) foi adaptado para ser aplicado às condições das pesquisas domiciliares brasileiras, em que não se consegue identificar, no banco de microdados do censo demográfico, o domicílio que foi selecionado na amostra da pesquisa. Além disso, no método original, supõe-se que todas as áreas possuem unidades na amostra, o que não se observa na prática. Dessa forma, adotou-se, na implementação do método MR, a sugestão dos autores descrita no final da Subseção “Metodologia proposta por Molina e Rao (2010)”.

Cada método possui sua dificuldade de aplicação a dados reais. No método FH é preciso dispor de variáveis explicativas que descrevam cada área de interesse, advindas de censos demográficos e registros administrativos. Na prática, a existência dessas variáveis em níveis menores é restrita. Outra desvantagem do método é a avaliação dos erros das estimativas ser possível apenas para as áreas contidas na amostra. Quanto aos métodos ELL e MR, são necessárias regressoras que existam na pesquisa amostral e no censo demográfico no nível domiciliar. Na realidade, a tarefa de reunir tais variáveis não é simples, uma vez que pode não ser possível a compatibilização de conceitos de variáveis importantes para medir o bem-estar do domicílio. Além disso, muitas vezes são afetadas por não-resposta, inviabilizando seu uso nos modelos. O método ELL calcula as variâncias das estimativas diretamente a partir das simulações, pelo método de Monte Carlo. No entanto, esta forma de obtenção dos erros das estimativas é questionada em Molina e Rao (2010). Em relação aos erros do método MR, os autores propõem um método *bootstrap* paramétrico que demanda grande tempo computacional, mas suposto eficaz na redução do erro quadrático médio. Este estudo não avaliou o desempenho dos métodos em relação aos erros quadráticos médios das estimativas propostos por seus respectivos autores.

A área de interesse, neste estudo, foi o município. O desenho amostral utilizado não seleciona diretamente municípios, não garantindo que todos fossem contemplados nas subamostras. Assim, 449 municípios não apareceram em nenhuma das subamostras selecionadas. A fim de verificar se a ausência de áreas de interesse nas subamostras prejudicou o desempenho dos métodos na análise municipal, obteve-se, também, resultados dos indicadores de pobreza para as 66 microrregiões de Minas Gerais. Assim, 40 dessas áreas possuíam observações em todas as subamostras. As demais

microrregiões foram encontradas em mais de 155 subamostras.

Este estudo indicou que os valores de vício e EQM relativos foram maiores para os municípios que apresentaram valores mais baixos de incidência de pobreza. Os métodos de estimação em pequenas áreas superestimaram os valores de proporção de pobres resultantes da população. Observou-se que os EQMs e os vícios relativos médios desses indicadores foram maiores para o método FH. No caso das microrregiões, o método MR foi o que apresentou vício relativo médio maior e o método FH apresentou o maior EQM relativo médio. Como esperado, os valores de EQMs e vícios relativos foram menores para as microrregiões que os valores obtidos para os municípios. A estimativa obtida pelo método MR para Belo Horizonte, município com mais domicílios nas subamostras, encontra-se mais próxima à proporção calculada na população do que as estimativas obtidas pelos demais métodos. Já, a microrregião que contém esse município apresentou maior proximidade entre as estimativas dos métodos ELL e MR, mas ainda distante do valor populacional.

Assim como observado no caso do indicador de proporção de pobres, os vícios relativos foram maiores para os municípios com menores valores do hiato de pobreza. As estimativas dadas pelo método MR ficaram ainda mais distantes dos valores populacionais. Os métodos ELL e MR tenderam a superestimar o hiato de pobreza nos municípios que apresentaram valores mais baixos desse indicador e a subestimá-lo quando esses foram altos. Já, o método FH superestimou o hiato de pobreza da maioria dos municípios. O mesmo não foi verificado no caso das microrregiões, nas quais o método MR superestimou os hiatos de pobreza na maioria das áreas.

Foi verificado que os coeficientes de variação estimados das estimativas de proporção de pobres e hiato de pobreza fornecidas pelos métodos nas 400 subamostras variaram mais quando as áreas de interesse foram os municípios. Os métodos produziram estimativas mais próximas entre si nas microrregiões. Na maior parte das áreas, os intervalos de confiança de 95% dos métodos ELL e MR não contiveram os indicadores populacionais. O método FH forneceu intervalos de confiança com grandes amplitudes que contiveram os valores populacionais na maioria das áreas.

As comparações realizadas em termos de ordenação de municípios, utilizando as estimativas de proporção de pobres, mostraram que as diferenças encontradas no desempenho dos métodos ELL e MR foram menores que as diferenças encontradas nas comparações de vício e EQM relativos. O método que apresentou ordenações de municípios mais diferentes da ordenação da população foi o FH, com correlação média de 0,86. No caso das microrregiões, notou-se uma maior diferença entre os métodos de estimação em pequenas áreas. O método ELL apresentou uma concordância maior com os valores populacionais. Resta decidir se a ordenação dos pares seria um critério razoável para a escolha de um método de estimação. Apesar de haver interesse em focar políticas públicas de combate à pobreza nos municípios mais pobres, é também necessário identificar, dentro de cada área, a parcela de domicílios mais pobres.

Referências bibliográficas

- Elbers, C., Lanjouw, J. O. e Lanjouw, P. (2002) Micro-level estimation of welfare. *Relatório técnico*, The World Bank.
- Fay, R. E. e Herriot, R. A. (1979) Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association*, **74**, 269–277.
- Foster, J., Greer, J. e Thorbecke, E. (1984) A class of decomposable poverty measures. *Econometrica*, **52**, 761–766.
- IBGE (2005) *Aspectos de Amostragem Especial Microdados POF 2002-2003*. Rio de Janeiro: IBGE.
- (2008) *Mapa de pobreza e desigualdade: municípios brasileiros 2003*. Rio de Janeiro: IBGE.
- (2010) *Pesquisa de Orçamentos Familiares 2008-2009 Despesas, Rendimentos e Condições de Vida*. Rio de Janeiro: IBGE.
- James, W. e Stein, C. (1961) Estimation with quadratic loss. Em *Fourth Berkeley Symposium of Mathematical Statistics and Probability*, vol. 1, 361–379. Berkeley: University of California Press.
- Lehmann, E. e D'Abrera, H. (1975) *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco: Holden-Day.
- Molina, I. e Rao, J. N. K. (2010) Small area estimation of poverty indicators. *Canadian Journal of Statistics*, **38(3)**, 369–385.
- Moura, F. A. S. (2008) *Estimação em pequenos domínios*. São Pedro: Associação Brasileira de Estatística - 18^o SINAPE, 1 edn.
- Pessoa, D. G. C. (2013) Comparação entre estimadores sintéticos obtidos por regressão baseada em dados dos municípios e em dados da UF de Minas Gerais. *Relatório técnico*, IBGE.
- Quintaes, V. C. C., Souza, D. F. D., Machado, D. C., Assis, J. M., Duarte, C. B. e Santos, I. C. M. (2006) Análise do bem-estar a partir de um indicador de consumo agregado com base nas informações da POF 2002-2003. Em *XV Encontro Nacional de Estudos Populacionais*. ABEP.
- Rao, J. N. K. (2003) *Small Area Estimation*. London: Wiley.
- Robert, C. e Casella, G. (2004) *Monte Carlo Statistical Methods*. New York: Springer, 2 edn.

Apêndice A - Modelos selecionados para aplicação do método FH

Tabela 10: Variáveis escolhidas para o modelo populacional, seus coeficientes estimados e limites inferior (LI) e superior (LS) dos intervalos de confiança de 95%, total de modelos das subamostras em que estas apareceram, médias dos coeficientes estimados nestes modelos e seus respectivos limites inferior e superior dos intervalos de confiança de 95%. Modelos ajustados para o indicador de incidência de pobreza no nível municipal.

Descrição das variáveis	Subamostras				População		
	Total	LI	Média	LS	LI	Coeficiente	LS
Intercepto	400	-3,4453	0,8992	6,1456	0,4335	0,5202	0,6068
Proporção de domicílios com automóvel	163	-0,9318	-0,5316	-0,2648	-0,3189	-0,2758	-0,2327
Proporção de domicílios com computador	58	-1,7376	-0,5810	1,4868	-0,3278	-0,2703	-0,2127
Proporção de domicílios com geladeira	183	-1,5097	-0,8782	-0,3791	-0,3678	-0,3048	-0,2417
Proporção de responsáveis homens	93	-0,2166	0,5162	0,9398	0,1305	0,1797	0,2288
Proporção de domicílios com televisão	33	-2,3361	0,5989	2,2214	0,0622	0,1491	0,2360
Proporção de domicílios alugados	148	-1,0686	-0,6439	-0,3562	-0,2327	-0,1775	-0,1223
Proporção de responsáveis com nível superior completo	39	-1,2937	1,1050	2,3779	0,2078	0,3495	0,4911
Proporção de domicílios com famílias unipessoais	289	-4,3392	-1,9931	-0,7291	-0,4038	-0,3080	-0,2123
Proporção de domicílios com mulher chefe com filhos	94	-4,1343	-2,2151	-0,7246	0,1204	0,2289	0,3375
Proporção de responsáveis desocupados	99	1,4335	2,6771	4,6452	0,5024	0,7506	0,9989
Logaritmo da proporção de responsáveis com emprego formal	166	-0,2132	-0,1244	-0,0631	-0,0531	-0,0448	-0,0364
Logaritmo do número médio de idosos nos domicílios	44	-0,2562	-0,0976	0,1319	-0,0505	-0,0346	-0,0188
Proporção de responsáveis inativos	81	0,0930	0,1704	0,2989	0,0538	0,0702	0,0865
Logaritmo do número médio de crianças de 7 a 14 anos nos domicílios	56	-0,2392	0,1098	0,3526	0,0840	0,1041	0,1241
Proporção de pessoas que receberam bolsa família em 2010	97	-1,5472	1,5459	3,3115	0,3062	0,4590	0,6117

Tabela 11: Variáveis escolhidas para o modelo populacional, seus coeficientes estimados e limites inferior (LI) e superior (LS) dos intervalos de confiança de 95%, total de modelos das subamostras em que estas apareceram, médias dos coeficientes estimados nestes modelos e seus respectivos limites inferior e superior dos intervalos de confiança de 95%. Modelos ajustados para o indicador de incidência de pobreza no nível de microrregião.

Descrição das variáveis	Subamostras				População		
	Total	LI	Média	LS	LI	Coeficiente	LS
Intercepto	400	-22,6554	2,9968	21,9255	0,3124	1,4531	2,5938
Proporção de domicílios com automóvel	62	-1,6489	0,1905	1,8331	-0,3628	-0,2569	-0,1510
Proporção de domicílios com bens básicos	90	-26,7898	-9,5671	16,5493	-3,7726	-1,9713	-0,1701
Proporção de domicílios com água canalizada em pelo menos um cômodo	69	-4,6934	-1,5766	3,8152	0,4517	0,6777	0,9037
Proporção de domicílios com geladeira	73	-3,2744	-0,0508	2,6761	-0,7385	-0,5329	-0,3272
Proporção de domicílios com iluminação fornecida por companhia distribuidora	73	-11,7293	0,2082	9,0824	0,7429	1,4626	2,1822
Proporção de responsáveis que sabem ler e escrever	62	-3,5497	-0,9529	2,3517	-0,5219	-0,3477	-0,1735
Proporção de domicílios alugados	250	-4,1667	-2,0131	-0,8035	-0,5055	-0,3543	-0,2031
Proporção de domicílios cujo tipo de escoadouro é rede geral	59	-0,8066	0,1794	0,6769	-0,1653	-0,1212	-0,0770
Proporção de domicílios cujo lixo é jogado em rio, lago ou mar	95	-130,7323	-5,6718	130,0546	1,0458	9,0462	17,0466
Proporção de responsáveis com nível médio completo ou superior incompleto	83	-2,4977	2,0635	5,5461	-0,0673	0,1362	0,3397
Proporção de domicílios com parede de madeira para construção	99	-58,1609	12,4854	62,6894	-8,4095	-4,4761	-0,5426
Logaritmo da proporção de responsáveis desocupados	53	-0,5415	0,1752	0,8640	0,0169	0,0365	0,0562
Logaritmo da proporção de responsáveis com emprego formal	76	-0,7750	-0,1584	0,4519	-0,0616	-0,0300	0,0015
Logaritmo do número médio de idosos nos domicílios	91	-0,5419	0,3160	0,8462	-0,0117	0,0447	0,1011
Logaritmo do número médio de crianças com menos de 7 anos nos domicílios	65	-1,1968	-0,1674	0,9358	0,0627	0,1240	0,1853

Tabela 12: Variáveis escolhidas para o modelo populacional, seus coeficientes estimados e limites inferior (LI) e superior (LS) dos intervalos de confiança de 95%, total de modelos das subamostras em que estas apareceram, médias dos coeficientes estimados nestes modelos e seus respectivos limites inferior e superior dos intervalos de confiança de 95%. Modelos ajustados para o indicador hiato de pobreza no nível municipal.

Descrição das variáveis	Subamostras				População		
	Total	LI	Média	LS	LI	Coeficiente	LS
Intercepto	400	-2,3193	-0,0742	2,4329	0,0566	0,1345	0,2125
Proporção de domicílios com automóvel	74	-0,4620	-0,2422	-0,1265	-0,1267	-0,1067	-0,0867
Proporção de domicílios com geladeira	280	-1,0055	-0,5810	-0,2501	-0,2461	-0,2127	-0,1793
Proporção de domicílios com iluminação fornecida por companhia distribuidora	150	0,5644	1,2545	2,2618	0,0817	0,1627	0,2438
Proporção de responsáveis homens	103	-0,1898	0,2872	0,5183	0,0645	0,0871	0,1097
Proporção de domicílios em que o lixo é queimado, enterrado, jogado em terreno baldio	31	-0,4097	0,0916	0,2941	0,0258	0,0374	0,0490
Proporção de domicílios sem banheiro	58	-1,0922	-0,4133	0,5439	0,0557	0,0866	0,1175
Proporção de domicílios constituídos por família do tipo unipessoal	242	-1,9779	-0,8396	-0,3964	-0,1749	-0,1240	-0,0731
Proporção de responsáveis desocupados	74	-1,2537	1,3842	2,3529	0,3499	0,4772	0,6046
Logaritmo da proporção de responsáveis com emprego formal	238	-0,1485	-0,0831	-0,0371	-0,0391	-0,0339	-0,0287
Logaritmo da proporção de responsáveis sem carteira de trabalho assinada	51	-0,0666	-0,0158	0,0605	-0,0109	-0,0066	-0,0022
Proporção de responsáveis inativos	94	-0,0705	0,0856	0,1569	0,0234	0,0316	0,0399
Logaritmo do número médio de crianças com menos de 7 anos nos domicílios	62	-0,2103	-0,0557	0,1518	0,0124	0,0221	0,0319
Logaritmo do número médio de crianças de 7 a 14 anos nos domicílios	62	-0,1235	0,1030	0,2505	0,0427	0,0539	0,0652
Proporção de pessoas que receberam bolsa família em 2010	112	0,4513	0,8631	1,5527	0,2333	0,3086	0,3839
Razão da renda média dos 10% mais ricos em relação aos 40% mais pobres em 2010	41	-0,0030	0,0008	0,0032	0,0006	0,0008	0,0011

Tabela 13: Variáveis escolhidas para o modelo populacional, seus coeficientes estimados e limites inferior (LI) e superior (LS) dos intervalos de confiança de 95%, total de modelos das subamostras em que estas apareceram, médias dos coeficientes estimados nestes modelos e seus respectivos limites inferior e superior dos intervalos de confiança de 95%. Modelos ajustados para o indicador hiato de pobreza no nível de microrregião.

Descrição das variáveis	Subamostras				População		
	Total	LI	Média	LS	LI	Coeficiente	LS
Intercepto	400	-5,3375	2,0310	11,4334	0,9741	1,3338	1,6936
Proporção de domicílios com automóvel	51	-0,6850	-0,0008	0,6986	-0,1233	-0,0730	-0,0226
Proporção de domicílios com computador	41	-1,5515	0,1041	1,9010	-0,2617	-0,1738	-0,0859
Proporção de responsáveis que sabem ler e escrever	74	-1,8893	-0,5387	1,0318	-0,2073	-0,1376	-0,0679
Proporção de domicílios em que o lixo é queimado, enterrado, jogado em terreno baldio	67	-1,1959	0,3622	4,1444	0,0086	0,0543	0,1000
Proporção de domicílios em que o lixo é jogado em rio, lago ou mar	128	-54,4986	-6,3565	51,3864	-0,2418	3,2779	6,7975
Proporção de responsáveis sem instrução ou com fundamental incompleto	49	-1,3691	0,1107	2,2971	-0,3291	-0,2324	-0,1356
Proporção de responsáveis com nível de instrução fundamental completo ou médio incompleto	189	-3,9532	-2,0845	-0,6208	-0,7253	-0,4882	-0,2510
Proporção de domicílios com dois cômodos servindo de dormitório	39	-1,0969	0,5433	2,0723	0,0452	0,1567	0,2683
Proporção de domicílios com parede de alvenaria e taipa revestida	135	-6,1680	0,6953	7,3875	-1,2771	-0,9083	-0,5395
Proporção de domicílios com parede de madeira aproveitada	168	-11,7544	16,0527	33,7887	-5,2155	-3,5579	-1,9003
Proporção de domicílios constituídos por família do tipo unipessoal	148	-4,1610	-1,8538	0,6464	-0,3566	-0,2179	-0,0791
Proporção de domicílios constituídos por família do tipo casal sem filhos	49	-1,4948	0,9474	3,1846	-0,5014	-0,3335	-0,1656
Logaritmo da proporção de responsáveis com emprego formal	112	-0,3289	-0,1345	0,1486	-0,0827	-0,0640	-0,0453
Logaritmo da proporção de responsáveis sem carteira de trabalho assinada	89	-0,1858	0,0705	0,2538	-0,0354	-0,0212	-0,0070
Logaritmo do número médio de crianças com menos de 7 anos nos domicílios	84	-0,5387	-0,1151	0,4104	-0,0007	0,0327	0,0660

Apêndice B - Resultados da ordenação de áreas para o hiato de pobreza

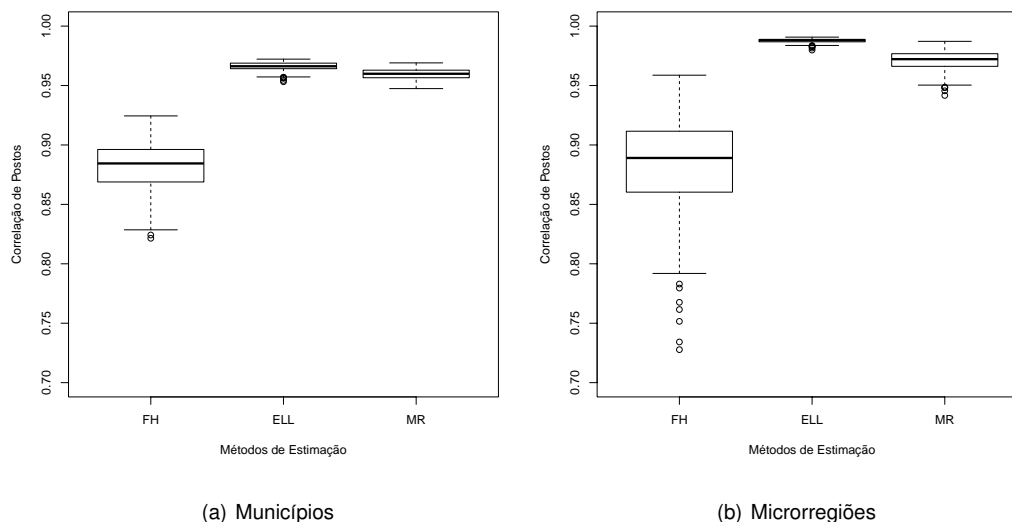
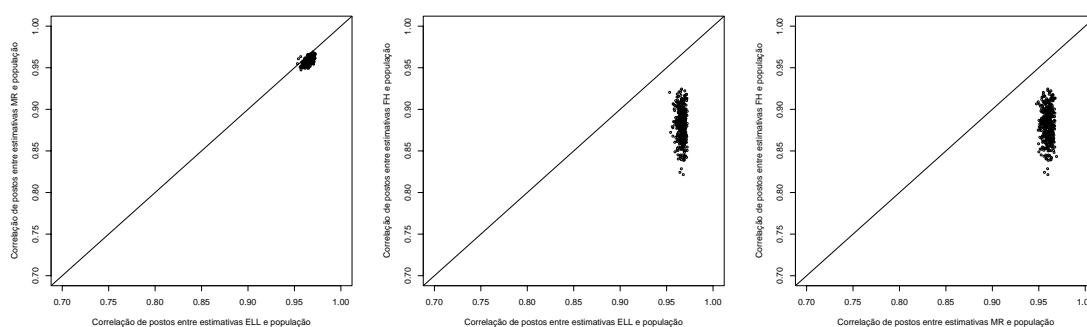


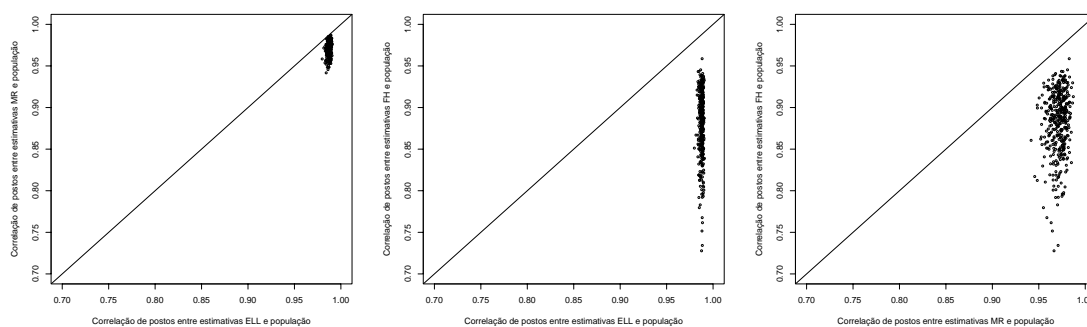
Figura 21: *Box plots* das correlações de postos entre estimativas de hiato de pobreza obtidas pelos métodos FH, ELL e MR e os valores populacionais considerando municípios e microrregiões.

Tabela 14: Distribuição da correlação de postos de Spearman por municípios e microrregiões.

Áreas	Correlação de Postos	Min.	1ºQu.	Mediana	Média	3ºQu.	Max.
Municípios	$r_i(FH, POP)$	0,8215	0,8689	0,8845	0,8825	0,8962	0,9244
	$r_i(ELL, POP)$	0,9532	0,9642	0,9662	0,9661	0,9688	0,9722
	$r_i(MR, POP)$	0,9474	0,9566	0,9599	0,9596	0,9629	0,9691
	$r_i(ELL, MR)$	0,9748	0,9855	0,9884	0,9878	0,9903	0,9959
	$r_i(FH, ELL)$	0,8124	0,8677	0,8823	0,8796	0,8948	0,9280
	$r_i(FH, MR)$	0,8130	0,8655	0,8796	0,8771	0,8914	0,9216
Microrregiões	$r_i(FH, POP)$	0,7278	0,8603	0,8891	0,8828	0,9115	0,9587
	$r_i(ELL, POP)$	0,9799	0,9868	0,9881	0,9877	0,9889	0,9907
	$r_i(MR, POP)$	0,9417	0,9662	0,9722	0,9711	0,9768	0,9872
	$r_i(ELL, MR)$	0,9290	0,9698	0,9754	0,9746	0,9803	0,9933
	$r_i(FH, ELL)$	0,7216	0,8522	0,8797	0,8738	0,9015	0,9531
	$r_i(FH, MR)$	0,7452	0,8635	0,8865	0,8809	0,9054	0,9504



(a) Municípios



(b) Microrregiões

Figura 22: Gráficos de dispersão das correlações de postos entre estimativas de hiato de pobreza obtidas pelos métodos FH, ELL e MR e os valores populacionais considerando municípios e microrregiões.

Tabela 15: Intervalos de confiança de 95% para os coeficientes de correlação de postos por municípios e microrregiões.

Áreas	Correlação de Postos	2,5%	97,5%
Municípios	$r(FH, POP)$	0,8423	0,9168
	$r(ELL, POP)$	0,9586	0,9715
	$r(MR, POP)$	0,9511	0,9667
	$r(ELL, MR)$	0,9782	0,9938
	$r(FH, ELL)$	0,8308	0,9158
	$r(FH, MR)$	0,8319	0,9130
Microrregiões	$r(FH, POP)$	0,7958	0,9377
	$r(ELL, POP)$	0,9842	0,9899
	$r(MR, POP)$	0,9534	0,9834
	$r(ELL, MR)$	0,9555	0,9888
	$r(FH, ELL)$	0,7863	0,9302
	$r(FH, MR)$	0,7926	0,9322

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/ Betina Fresneda, Cristiano Santos...[et al.]. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2013.