

Nota técnica 01/2007

IMPUTAÇÃO DE VALORES FALTANTES REFERENTES ÀS VARIÁVEIS DE RENDIMENTO DO TRABALHO

PESQUISA MENSAL DE EMPREGO

O IBGE, a partir da divulgação dos resultados da Pesquisa Mensal de Emprego de março de 2007, prevista para 26 de abril de 2007, passará a divulgar os indicadores relativos a rendimento do trabalho das pessoas com 10 anos ou mais de idade, considerando o tratamento de imputação de valores de rendimentos naqueles registros em que os rendimentos não tenham sido informados no processo de coleta, e que até então eram considerados ignorados.

Nessa ocasião, apresentará, também, a série histórica, que se inicia em março de 2002, com as estimativas de rendimentos médios recalculadas, ou seja, levando em conta este tratamento, cujas linhas gerais são descritas a seguir.

1. *Por que adotar o procedimento de imputação das variáveis de rendimentos?*

Entre os erros não-amostrais a que uma pesquisa está sujeita, a não-resposta é um dos que sempre se verificam, em maior ou menor grau, seja por dificuldade de estabelecer contato com o informante, seja por dificuldade do respondente em oferecer as informações solicitadas.

Tais dificuldades são mais percebidas em pesquisas repetidas mensalmente, como é o caso da Pesquisa Mensal de Emprego, cujo cronograma de coleta requer entrevistas num curto período de tempo. Ademais, o rendimento, que é visto como uma informação confidencial sensível, é um dos quesitos mais vulneráveis à não-resposta, mesmo quando o restante da pesquisa é realizado com sucesso. Esse comportamento também é observado em pesquisas censitárias e por amostragem, tanto no Brasil como internacionalmente.

Além disso, a não-resposta de rendimentos tende a ocorrer com mais frequência em todas as investigações quando esses são mais altos, o que, também, se verifica na Pesquisa Mensal de Emprego. Esse comportamento é identificado como não-resposta diferencial, e um tratamento estatístico é requerido para correção desse vício das estimativas produzidas. Optou-se, dessa forma, por adotar o procedimento de imputação.

2. *As variáveis que serão imputadas*

- ✓ Rendimento bruto mensal *habitual* do trabalho principal dos empregados e trabalhadores domésticos;
- ✓ Rendimento bruto mensal *efetivo* do trabalho principal dos empregados e trabalhadores domésticos, recebido no mês de referência;

- ✓ Retirada mensal *habitual* do trabalho principal dos trabalhadores por conta própria e empregadores;
- ✓ Retirada mensal *efetiva* do trabalho principal dos trabalhadores por conta própria e empregadores, no mês de referência;
- ✓ Rendimento mensal *habitual* dos trabalhadores no(s) outro(s) trabalho(s) que tinha(m) na semana de referência; e
- ✓ Rendimento mensal *efetivo* dos trabalhadores no(s) outro(s) trabalho(s) que tinha(m) na semana de referência, recebido no mês de referência.

3. A taxa de não-resposta dos rendimentos na Pesquisa Mensal de Emprego

As taxas médias de não-resposta dos quesitos de rendimentos da pesquisa, desde o início da produção da nova série, vêm decrescendo a partir do aprimoramento constante dos treinamentos, controles e procedimentos de coleta. Em 2006, a taxa média anual de não-resposta dos quesitos de rendimentos do trabalho principal atingiu 2,3%. Observe que, na tabela a seguir, essa mesma média calculada de março de 2002 a dezembro de 2006 correspondia a 5,6%.

Tabela 1: Taxa de não-resposta de pelo menos um dos quesitos de rendimento¹ do trabalho principal da PME (em %)

Região Metropolitana	Média histórica ²	Média de 2002	Média de 2003	Média de 2004	Média de 2005	Média de 2006
Total	5,6	8,0	9,0	6,2	2,9	2,3
Recife	7,7	10,1	12,9	11,9	2,4	1,5
Salvador	6,2	7,6	10,5	8,2	3,1	1,8
Belo Horizonte	3,9	8,0	7,1	2,6	1,2	1,2
Rio de Janeiro	8,3	12,9	10,6	8,3	5,7	4,9
São Paulo	3,7	4,7	6,1	4,0	2,1	1,7
Porto Alegre	4,7	5,6	8,8	4,7	2,4	2,3

Fonte: IBGE, Coordenação de Trabalho e Rendimento, Pesquisa Mensal de Emprego - PME.

(1) Incluindo habitual e efetivo.

(2) Média calculada de março de 2002 a dezembro de 2006, sobre o total de remunerados.

4. A metodologia de imputação

Adotou-se a metodologia que combina Árvores de Regressão com seleção probabilística de doadores em cada estrato construído através da árvore (Breiman et al, 1984) ¹. Em linhas gerais, a técnica de árvores de regressão consiste em um método de estratificação que utiliza um conjunto de características das pessoas respondentes da PME para classificar os

¹ BREIMAN, L., FRIEDMAN, J.H., OLSHEN R.H. and STONE, C.J. *Classification and Regression Trees*, 1984, Monterrey:Wadsworth and Brooks/Cole.

registros em grupos homogêneos. Essas características se expressam segundo as seguintes variáveis explicativas, associadas tanto para as pessoas que informaram rendimentos, quanto para aquelas que incorreram em não-resposta.

- ✓ Sexo;
- ✓ Idade;
- ✓ Condição do morador no domicílio;
- ✓ Anos de estudo;
- ✓ Posição na ocupação no trabalho principal;
- ✓ Tamanho do empreendimento do trabalho principal; e
- ✓ Horas habitualmente trabalhadas no trabalho principal.

Em seguida, cada registro que requer imputação (receptor) recebe o valor de rendimento de um registro (doador) selecionado aleatoriamente dentro do grupo homogêneo a que ambos pertencem.

O objetivo da utilização dessa técnica foi o de tratar a não-resposta diferencial nos quesitos de rendimento. Essa metodologia foi utilizada anteriormente pelo IBGE para a imputação de rendimentos no Censo Demográfico de 2000.

Cabe destacar que existe um conjunto de situações especiais, que, assim como o procedimento detalhado, serão apresentadas em nota metodológica, a ser disseminada com os resultados da PME de março de 2007.

5. A divulgação

Na data da divulgação da PME de março, **26 de abril de 2007**, o IBGE disseminará:

- ✓ Nota metodológica;
- ✓ Séries históricas dos indicadores de rendimentos, traçando um comparativo entre os indicadores divulgados e os indicadores recalculados em função do tratamento de imputação;
- ✓ Os indicadores de rendimentos divulgados como parte dos resultados da PME de março incorporando o novo procedimento;
- ✓ A partir de março de 2007, passa a ser divulgada a massa de rendimentos com base na PME, o que se torna possível a partir do tratamento da não resposta das variáveis de rendimentos;
- ✓ A partir de março de 2007, os arquivos de microdados da PME vão conter, para todos os registros, a informação indicativa se houve ou não imputação, propiciando aos usuários a alternativa de, para estudos e análises específicos, adotar ou não o rendimento imputado.

Rio de Janeiro, 26 de março de 2007.