

AGRICULTURA EM SÃO PAULO
Boletim Técnico do Instituto de Economia Agrícola

Ano 37

Tomo 1

1990

DIFERENÇAS DE CUSTOS DOS LEVANTAMENTOS POR AMOSTRAGEM ENTRE REGIÕES E TAMANHOS DE IMÓVEIS NO ESTADO DE SÃO PAULO⁽¹⁾

Denise Viani Caser⁽²⁾
José Roberto Vicente⁽²⁾
Luiz H.O. Piva⁽²⁾
Ana M.M.P. Camargo⁽²⁾
Mário P.A. Olivetti⁽³⁾

RESUMO

O objetivo deste trabalho foi o de testar a hipótese de que os custos dos levantamentos por amostragem para previsões e estimativas das safras agrícolas do Estado de São Paulo não diferiram entre si nas regiões e diferentes tamanhos de imóvel. Análise de variância, testes de Tukey e de χ^2 foram utilizados. Os resultados mostraram que existiam diferenças significativas nos custos, tanto entre regiões (DIRAS) como entre estratos de tamanho do imóvel.

Palavras-chave: Custos dos levantamentos amostrais; levantamentos agrícolas.

DIFFERENCES OF COSTS OF THE SAMPLE SURVEYS AMONG REGIONS AND FARM SIZES IN THE STATE OF SÃO PAULO, BRAZIL

SUMMARY

The intention of this paper was to check the hypothesis that the costs of the sample surveys used for the forecast and estimations of the crops in the State of São Paulo do not differ statistically among the geographic strata and among the area strata. The results showed that significant differences exist among DIRAs and also among area strata.

Key-words: costs of the sample surveys; agricultural surveys.

⁽¹⁾ Trabalho referente ao projeto SPTC 16-006/90. Recebido em 31/01/90. Liberado para publicação em 22/03/90. Os autores agradecem as sugestões do Prof. Rodolfo Hoffmann, do Departamento de Economia e Sociologia Rural da ESALQ/USP.

⁽²⁾ Pesquisador Científico do Instituto de Economia Agrícola (IEA).

⁽³⁾ Bacharel em Geografia, funcionário do Instituto de Economia Agrícola (IEA).

1 - INTRODUÇÃO

A Secretaria de Agricultura e Abastecimento do Estado de São Paulo elabora e publica regularmente, desde 1942, estatísticas de área e produção dos principais produtos agrícolas.

Inicialmente, a metodologia empregada era a que se rotula de levantamento subjetivo, no qual os engenheiros agrônomos alocados nas zonas produtoras avaliavam as safras baseados em seus conhecimentos pessoais e informações de terceiros. Entretanto, essa técnica impossibilitava obter estimativas da precisão dessas estatísticas e de elementos para a análise de problemas sócio-econômicos mais complexos do setor agrícola. Assim, em 1950, SCHATTAN (9) e STEVENS (10) introduziram o uso de amostragem nos levantamentos para previsão e estimativas de safras. Esse processo foi sendo alterado e aperfeiçoado com o tempo e, em 1954, iniciou-se a publicação regular das estimativas obtidas por esse método ao nível estadual. Concomitantemente, os levantamentos subjetivos continuaram a ser efetuados, tanto para efeito de comparação como também para estimar safras de produtos com áreas menores e/ou localizadas e obter dados ao nível de município. Os dois levantamentos (objetivo e subjetivo) são efetuados cinco vezes por ano, em fevereiro, abril, junho, setembro e novembro.

O interesse crescente por estatísticas ao nível de região levou CAMPOS & PIVA (2) a reformular os critérios de precisão utilizados e procurar estimar as produções para as Divisões Regionais Agrícolas (DIRAs) em que o Estado encontrava-se dividido. Essa amostra foi dimensionada pela partilha de Neyman, utilizando um esquema de dupla estratificação, por DIRA e por tamanho de imóvel, considerando as oito culturas mais importantes do Estado (algodão, amendoim, arroz, café, cana, feijão, milho e soja); o erro de amostragem foi limitado a 20% da média de cada DIRA (4). Em 1975 e 1977, as amostras foram dimensionadas utilizando esse mesmo esquema.

Uma nova estrutura foi definida por CAMARGO (1), dividindo cada DIRA em dois grupos de municípios de acordo com o valor da produção, cada grupo em doze estratos de área

e cada estrato em vários substratos ou blocos, de modo que cada elemento representasse, no máximo, quinhentos imóveis.

A partir de 1986, nova mudança na estrutura eliminou os grupos e blocos da amostra, com sorteios aleatórios dentro dos onze estratos de área em cada DIRA, conforme o esquema de CAMPOS & PIVA(2).

As amostras levadas a campo têm que ser substituídas periodicamente para evitar a exaustão do informante e o aumento do viés por falta de respostas, PINO & CASER(7). Até o momento, nenhum dos redimensionamentos foi efetuado considerando os custos, que são tomados como iguais entre estratos geográficos e de tamanho de imóvel; entretanto, essa premissa pode não ser verdadeira.

NEGRI NETO et alii(6) calcularam os benefícios e custos marginais das reduções dos erros amostrais dos levantamentos objetivos desenvolvidos pelo Instituto de Economia Agrícola (IEA) e pela Coordenadoria de Assistência Técnica Integral (CATI) e obtiveram uma relação variando entre 1.725,12 e 7,74. Para o cálculo dos custos foram aproveitados itens introduzidos no questionário de novembro de 1987 referentes a salário dos enumeradores, tempo total gasto com o levantamento de cada questionário, tempo de viagem, tempo de entrevista, distância percorrida e número de entrevistas realizadas em cada viagem; além desses tópicos, no cálculo dos custos, foram inseridas despesas de confecção, impressão e remessa dos questionários, digitação, processamento e depuração dos dados, salários e outros custos relacionados à análise dos resultados.

As informações geradas por aquela pesquisa referentes aos custos de campo foram utilizadas no presente estudo para indicar a existência ou não de diferenças entre as DIRAs e os estratos de tamanho de imóvel.

2 - OBJETIVOS

O presente trabalho tem como objetivo testar a hipótese de que os custos dos levantamentos por amostragem para as previsões e estimativas das safras agrícolas do Estado de São

(4) As DIRAs e os estratos de área, que constituem a amostra, estão nos quadros 1 e 2. *Agricultura em São Paulo, SP, 37(1):79-87, 1990.*

Paulo não diferem estatisticamente entre DIRAs e estratos de área; a mesma hipótese será testada para componentes dos custos totais.

Caso sejam encontradas diferenças estatisticamente significativas, procurar-se-á identificar as DIRAs e os estratos que diferem entre si.

3 - METODOLOGIA

Os dados básicos utilizados nesta pesquisa foram os provenientes das questões inseridas nos questionários do levantamento objetivo de novembro de 1987. A amostra utilizada era constituída por 3.662 elementos e foram aproveitados 3.027 questionários, que retornaram do campo preenchidos. Para o cálculo dos custos foram eliminados aqueles cujos itens referentes a tempo total gasto com a viagem e/ou salário médio do enumerador não estavam preenchidos, o que reduziu o número total de questionários a 2.708, preenchidos em 1.992 viagens efetuadas com essa finalidade.

Para testar as hipóteses de igualdade entre DIRAs e estratos de área, utilizou-se análise de variância, com as DIRAs e estratos constituindo os tratamentos. As pressuposições necessárias parecem ser satisfeitas no caso presente, uma vez que os dados analisados provêm de uma população que pode ser considerada infinita e foram obtidos de forma aleatória, provavelmente sem erros sistemáticos de medida. Todavia, existia a possibilidade de que a variância dos erros dentro do tratamento não fosse constante (heterocedasticia). Esse problema pode, às vezes, ser resolvido com uma transformação nas variáveis utilizando-se, por exemplo, logaritmos, se a variância for proporcional ao quadrado da média e raiz quadrada, seja variância for proporcional à média, HOFFMANN(4) e SIQUEIRA (8). Para verificar qual transformação seria a mais adequada ao presente, optou-se por tomar a que propiciou o menor valor de χ^2 (qui-quadrado) oriundo da relação U/G, HOFFMANN & VIEIRA (5), onde:

$$U = (\sum g_h) \ln \frac{\sum g_h s^2_h}{\sum g_h} - \sum g_h \cdot \ln s^2_h \quad (1)$$

Agricultura em São Paulo, SP, 37(1):79-87, 1990.

e

$$G = 1 + \frac{1}{3(H-1)} \left(\sum \frac{1}{g_h} - \frac{1}{\sum g_h} \right) \quad (2)$$

onde:

$g_h = n_h - 1$; H = número de tratamentos; e s^2_h = estimativa das variâncias dentro de cada tratamento.

Após efetuadas as análises de variância, utilizou-se o teste de Tukey, GOMES(3), para comparar todos os contrastes entre duas médias dos tratamentos:

$$\Delta = q \sqrt{1/2 \hat{V}(Y)} \quad (3)$$

onde:

q = valor da amplitude total studentizada; e $\hat{V}(Y)$ = estimativa da variância.

Assim, todo contraste entre duas médias que excedeu o valor Δ calculado foi significativo ao nível de probabilidade correspondente. Como o número de repetições nos tratamentos não era o mesmo, os resultados do teste são aproximados.

Para os testes de hipóteses, adotou-se o nível de 5% de probabilidade.

4 - RESULTADOS E DISCUSSÃO

Nas viagens efetuadas em novembro de 1987, os enumeradores da CATI levantaram, em média, 1,36 questionário, número que variou de 1,21 na DIRA de São Paulo a 1,48 na DIRA de São José do Rio Preto. Os enumeradores ocupavam diversos cargos, sendo também diversos os níveis de instrução; a maioria deles ocupava o cargo de auxiliar agropecuário, com as médias salariais variando entre Cz\$13.204,24 (175,78 BTN) na DIRA do Vale do Paraíba e Cz\$7.638,02 (101,68 BTN) na DIRA de Araçatuba.

Em média, a distância percorrida por questionário levantado foi de 31,2km, variando entre 36,9km na DIRA do Vale do Paraíba e 24,9km na DIRA de Sorocaba. O tempo médio despendido desde a saída até o retorno à sede foi de 56 minutos, ficando os valores extremos com as DIRAS do Vale do Paraíba (86 minutos) e de Bau-

ru (45 minutos). Para a entrevista foram necessários em média, 64 minutos, variando de 44 minutos em Sorocaba a 75 minutos em Campinas (quadro 1).

Entre os estratos de área, observa-se que os imóveis com áreas maiores custam, em geral, mais para serem levantados, em função do tempo de entrevista e as distâncias percorridas serem maiores (quadro 2).

As análises de variância foram efetuadas, inicialmente, com as variáveis na forma original, transformadas em logaritmos decimais e em raízes quadradas; esse procedimento possibilitou os cálculos dos valores de χ^2 . Pelos valores obtidos, não se rejeitam as hipóteses de homocedasticidade para os dados ao nível de DIRA, transformados em logaritmos, referentes às variáveis tempo de viagem, tempo total gasto, número de entrevistas e distância percorrida (quadro 3). Com os dados ao nível de estrato, os mesmos resultados foram obtidos para as variáveis custo do levantamento e tempo total de viagem, com a mesma transformação.

Nas demais variáveis, optou-se por utilizar, como descrito anteriormente, a transformação que possibilitou o menor valor de χ^2 , apesar dos testes indicarem a existência de heterocedasticidade. Isso foi feito para violar o menos possível essa pressuposição da análise de variância, embora a não existência de homogeneidade das variâncias, considerando-se experimentos balanceados com fatores fixos, não traga conseqüências sérias e a análise de variância possa ser utilizada, pois as conclusões não serão, conforme SIQUEIRA(8), significativamente alteradas.

Portanto, nas análises de variância consideradas, utilizou-se dados transformados em raízes quadradas para o custo do levantamento (nas DIRAs), tempo de entrevista (estratos de área) e em logaritmos para o tempo de entrevista (nas DIRAs), tempo total gasto, número de entrevistas, distância percorrida (estratos de área) e salários dos enumeradores (nas DIRAs e nos estratos).

Das análises de variância efetuadas para o tratamento DIRA, concluiu-se haver diferenças significativas nas variáveis custo total do levantamento e salários dos enumeradores (quadro 4).

O teste de Tukey mostrou que o custo do
Agricultura em São Paulo, SP, 37(1):79-87, 1990.

levantamento na DIRA do Vale do Paraíba era significativamente maior (a 5% de probabilidade) do que na DIRA de Sorocaba. Essa última DIRA apresentou custos também inferiores aos de Campinas, Marília, Ribeirão Preto, Araçatuba e São José do Rio Preto; a DIRA de Ribeirão Preto apresentou custos inferiores aos de São José do Rio Preto, enquanto as demais DIRAs não diferiram entre si. Na DIRA do Vale do Paraíba, os enumeradores tinham os maiores salários médios e os tempos de viagem eram também os maiores, em média, enquanto que Sorocaba apresentava os menores tempos de viagem e de entrevista e distância percorrida. Com relação aos salários dos enumeradores, a DIRA de Campinas apresentava o maior valor médio para os dados transformados em logaritmos, significativamente diferentes de todas as demais DIRAs, com exceção da do Vale do Paraíba. Como não há diferenças nos salários pagos pelo Estado para o mesmo cargo, entre regiões, pode-se concluir que no Vale do Paraíba e em Campinas, o pessoal alocado para efetuar o levantamento de campo ocupava cargos mais elevados e tinha, provavelmente, melhor qualificação. As demais variáveis não diferiram entre si.

Conclui-se, também, que existem diferenças significativas entre os estratos de área, para as variáveis custo total do levantamento, tempo gasto com a viagem e número de entrevistas realizadas em uma mesma viagem (quadro 4).

Pelo teste de Tukey, obteve-se como resultado que os questionários aplicados aos imóveis acima de 500 hectares (estratos 12 e 13) apresentaram custos totais para serem levantados, significativamente maiores do que os demais. O número de entrevistas realizadas numa mesma viagem apresentou-se significativamente maior no estrato 13, comparado aos estratos 4 e 8, enquanto que o tempo de viagem gasto com imóveis do estrato 4 foi significativamente maior do que o empregado com imóveis do estrato 12; esses resultados, todavia, devido à inversão das posições relativas das médias nos dados transformados, devem ser vistos com cautela.

5 - CONCLUSÕES E SUGESTÕES

Os resultados obtidos mostraram que, via-de-regra, existem diferenças no custo por ques-

QUADRO 1. – Valores Médios de Variáveis Componentes dos Custos do Levantamento de Questionários para Previsões de Safra por Amostragem, por Divisão Regional Agrícola (DIRA), Estado de São Paulo, Novembro de 1987

DIRA	Custo total (Cz\$/u.)	Tempo gasto (minuto)			Salário do enumerador (Cz\$/mês)	Distância percorrida (km)	Entrevista por viagem (nº)
		Total	Viagem	Entrevista			
São Paulo	312,44	125,0	77,7	47,1	8.112,03	34,8	1,21
Vale do Paraíba	423,37	138,6	85,9	52,6	13.204,24	36,9	1,37
Sorocaba	239,65	94,0	49,7	44,3	10.388,39	24,9	1,31
Campinas	359,07	128,5	53,9	74,6	12.936,84	30,8	1,30
Ribeirão Preto	327,94	124,2	56,0	68,3	9.675,88	31,4	1,36
Bauru	278,23	97,8	45,0	53,0	9.015,64	33,9	1,33
São José do Rio Preto	293,79	117,5	50,5	66,6	8.390,53	31,4	1,48
Araçatuba	320,00	128,3	67,2	61,1	7.638,02	36,3	1,35
Presidente Prudente	296,18	119,4	53,1	66,7	7.936,87	31,9	1,41
Marília	326,16	135,3	63,1	72,2	8.465,11	32,7	1,37
Estado	310,01	119,7	56,1	63,6	9.660,70	31,2	1,36

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola (IEA) e da Coordenadoria de Assistência Técnica Integral (CATI).

QUADRO 2. - Valores Médios de Variáveis Componentes dos Custos do Levantamento de Questionários para Previsões de Safra por Amostragem, por Estrato de Área, Estado de São Paulo, Novembro de 1987

Nº de ordem	Estrato de área		Custo total (Cz\$/u.)	Tempo gasto (minuto)			Salário do enumerador (Cz\$/mês)	Distância percorrida (km)	Entrevista por viagem (nº)
	Intervalo (ha)			Total	Viagem	Entrevista			
03	3,1 a	5,0	209,88	94,9	47,7	47,0	11.718,39	20,0	1,49
04	5,1 a	10,0	286,16	117,7	58,5	60,0	8.778,75	29,7	1,27
05	10,1 a	20,0	264,68	112,5	54,1	58,4	9.440,19	26,2	1,25
06	20,1 a	30,0	256,95	109,7	49,0	60,7	9.326,61	26,6	1,24
07	30,1 a	50,0	262,73	111,0	50,5	60,5	8.910,92	27,6	1,40
08	50,1 a	100,0	285,68	116,6	53,8	62,7	9.708,98	29,2	1,26
09	100,1 a	200,0	320,27	121,5	56,1	65,4	10.080,43	30,5	1,40
10	200,1 a	300,0	347,26	123,2	59,8	63,6	9.964,66	36,3	1,32
11	300,1 a	500,0	319,14	118,8	56,1	62,2	9.552,82	31,3	1,43
12	500,1 a	1.000,0	359,25	121,1	55,2	65,9	9.812,63	35,2	1,58
13	>	a 1.000,1	406,82	147,6	71,0	76,5	9.854,58	40,7	1,35
Estado			310,01	119,7	56,1	63,6	9.660,70	31,2	1,36

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola (IEA) e da Coordenadoria de Assistência Técnica Integral (CATI).

QUADRO 3. – Resultados dos Testes de χ^2 , Provenientes das Análises de Variância Efetuadas para Variáveis Componentes dos Custos do Levantamento de Questionários para Previsões de Safra por Amostragem, Estado de São Paulo, Novembro de 1987⁽¹⁾

Variável e forma dos dados	DIRA	Estrato
Custo total do questionário levantado		
Linear	137,34*	80,53*
Logarítmica	44,56*	11,78
Raiz quadrada	39,58*	25,89*
Tempo total gasto		
Linear	98,32*	64,16*
Logarítmica	14,28	31,96*
Raiz quadrada	37,91*	35,38*
Tempo de viagem		
Linear	154,50*	101,93*
Logarítmica	9,79	11,43
Raiz quadrada	28,90*	139,51*
Tempo de entrevista		
Linear	257,88*	35,20*
Logarítmica	22,77*	28,88*
Raiz quadrada	77,84*	27,66*
Salário do enumerador		
Linear	1.124,32*	93,44*
Logarítmica	453,31*	41,71*
Raiz quadrada	736,42*	64,93*
Distância percorrida		
Linear	62,64*	123,74*
Logarítmica	10,46	31,96*
Raiz quadrada	19,61*	52,97*
Número de entrevistas por viagem		
Linear	29,47*	149,87*
Logarítmica	12,33	56,69*
Raiz quadrada	12,52	82,90*

(¹) O asterisco indica significância a 5%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola (IEA) e da Coordenação de Assistência Técnica Integral (CATI).
Agricultura em São Paulo, SP, 37(1):79-87, 1990.

QUADRO 4. - Resultados dos Testes F Provenientes das Análises de Variância Efetuadas para Variáveis Componentes dos Custos do Levantamento de Questionários para Previsões de Safra por Amostragem, Estado de São Paulo, Novembro de 1987

Variável e tratamento	Valor de F ⁽¹⁾
Custo total do questionário levantado	
DIRA	6,87*
Estrato de área	7,36*
Tempo total gasto	
DIRA	1,05
Estrato de área	1,79
Tempo de viagem	
DIRA	1,29
Estrato de área	2,39*
Tempo de entrevista	
DIRA	1,26
Estrato de área	1,11
Salário do enumerador	
DIRA	16,93*
Estrato de área	0,90*
Distância percorrida	
DIRA	1,09
Estrato de área	1,55
Número de entrevistas por viagem	
DIRA	1,77
Estrato de área	2,57*

(¹) O asterisco indica significância a 5%.

tionário levantado por amostragem nos imóveis rurais do Estado de São Paulo, tanto entre as Divisões Regionais Agrícolas (DIRAs), como entre os estratos de tamanho de imóvel.

Essa conclusão sugere que amostras economicamente mais eficientes deveriam levar em conta esse fato, tomando o custo como uma das variáveis dimensionadoras.

LITERATURA CITADA

1. CAMARGO, Milton N. de. **Amostra para previsão e estimativa das safras agrícolas do Estado de São Paulo em vigor a partir de junho de 1981.** São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1988. 75p. (Relatório de Pesquisa, 27/88)
2. CAMPOS, Humberto & PIVA, Luiz H. de O. **Dimensionamento de amostra para estimativa e previsão de safra no Estado de São Paulo.** *Agricultura em São Paulo*, SP, 21(3):65-88, 1974.
3. GOMES, Frederico P. **Curso de estatística experimental.** 7.ed. São Paulo, Nobel, 1977.
4. HOFFMANN, Rodolfo. **Estatística para economistas.** São Paulo, Pioneira, 1980. 379p.
5. _____ & VIEIRA, S. **Análise de regressão: uma introdução à econometria.** São Paulo, Hucitec/EDUSP, 1977. 339p. (Obras didáticas - - Coleção Economia e Planejamento)
6. NEGRI NETO, Afonso et alii. **Custo e benefício social de previsões e estimativas de produção agrícola: o valor da informação.** *Agricultura em São Paulo*, SP, 35(1):37-49, 1988.
7. PINO, Francisco A. & CASER, Denise V. **Análise de erros não-amostrais em levantamentos para previsão e estimativa de safras no Estado de São Paulo.** São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1984. 25p. (Relatório de Pesquisa, 10/84)
8. SIQUEIRA, Arminda L. **Uso de transformação em análise de variância e análise de regressão.** São Paulo, IME/USP, 1983. 154p. (Tese - Mestrado)
9. SCHATTAN, Salomão. **Obtenção de estatísticas agrícolas pelo método de amostragem.** São Paulo, Departamento de Produção Vegetal, 1953. (Estudos de Economia Rural, 17)
10. STEVENS, W.L. **Estimativa e previsão de safras através de um levantamento por amostragem.** São Paulo, Divisão de Economia Rural, 1951.