

PROPOSTA DE REESTRUTURAÇÃO DO SISTEMA AMOSTRAL DE COLETA DE DADOS NO MERCADO VAREJISTA DE ALIMENTOS NA CIDADE DE SÃO PAULO: uma aplicação de amostragem estratificada probabilística em dois estágios¹

Vagner Azarias Martins²
Carlos Roberto Ferreira Bueno³
Raquel Castellucci Caruso Sachs⁴

1 - INTRODUÇÃO

A demanda por informações vem ganhando cada vez mais espaço em todos os ramos de atividade, tornando-se fonte principal para a tomada de decisões nos setores público, privado e acadêmico. No primeiro, ela é essencial para prover políticas governamentais; no segundo é a base para tomada de decisões, especialmente sobre investimentos; e no setor acadêmico, ela é imprescindível em trabalhos de pesquisas. Para obtê-la são utilizadas diversas ferramentas, estatísticas ou não, conforme a especificidade de cada caso.

Em geral, a informação é de natureza estatística e, por isso, deve ser apoiada em seis pilares definidos como utilidade, pertinência, validade, visibilidade, continuidade e integridade (SENRA, 1999). Segundo este mesmo autor, os dois primeiros pilares definem a demanda da pesquisa, os três seguintes definem a oferta e o último une o processo.

Nos pilares utilidade e pertinência existem as noções de relevância e abrangência da pesquisa, em que se define a área de estudo da qual serão geradas as informações e a importância da pesquisa. Também é verificada a necessi-

dade ou não da mesma. Nestes pilares objetiva-se que as informações sejam comparáveis e combináveis.

A oferta da pesquisa está assentada em três pilares. No pilar validade encontra-se a fase da produção, em que se define a metodologia, sua precisão e acurácia. O pilar visibilidade diz respeito à aceitabilidade da informação, e nele está a fase de disseminação de resultados. O pilar continuidade é a união das duas fases anteriores, procurando garantir a sistematização da informação.

O sexto pilar, integridade, funciona como elo entre a demanda e a oferta. Neste pilar assentam-se os critérios credibilidade e legitimidade.

A informação estatística nasce da demanda em se estudar certa característica de uma população. Pino (1999) entende que a informação estatística não deve ser tratada simplesmente como um conjunto de números obtidos estatisticamente, e sim, dados estatísticos analisados.

O Instituto de Economia Agrícola (IEA) coleta e calcula mensalmente os preços referentes ao mercado varejista de mais de 150 itens de alimentos dentro do domicílio para a cidade de São Paulo, desde o início da década de 1970 (JUNQUEIRA; CANTO, 1971). Este trabalho se consolidou em importante fonte de consulta de dados. Essas informações são amplamente utilizadas em diversos trabalhos técnico-científicos referentes a comercialização, análise setorial e estrutural, entre outros. Suas estimativas de preços estruturadas em formato de cesta mercado são base de indicadores de variação de preços que podem ser utilizados para a composição de Índices de Preços ao Consumidor (IPCs) e de Indicadores de Custo de Vida (ICVs).

Segundo Carmo (2004), as estatísticas e, sobretudo os indicadores de variação de preços, passaram a ter, a partir dos anos 1980,

¹Os autores agradecem as sugestões e observações em todas as fases do trabalho do Pesquisador Científico Dr. Francisco Alberto Pino e a participação das estagiárias Luciana Barboza Carnevalli, Mariana Xaubet e Izabelle Tomáz Felício na depuração dos dados da Pesquisa de Orçamentos Familiares. Cadastrado no SIGA NRP 1942 e registrado no CCTC, IE-27/2009.

²Estatístico, Pesquisador Científico do Instituto de Economia Agrícola (e-mail: vagneram@iea.sp.gov.br).

³Médico Veterinário, Pesquisador Científico do Instituto de Economia Agrícola (e-mail: crfbueno@iea.sp.gov.br).

⁴Engenheira Agrônoma, Mestre, Pesquisadora Científica do Instituto de Economia Agrícola (e-mail: raquelsachs@iea.sp.gov.br).

grande veiculação na mídia devido ao noticiário econômico da época. Essa grande exposição acabou popularizando os IPCs.

No município de São Paulo, o levantamento de dados no mercado varejista de alimentos também é realizado por outras instituições de pesquisa. A parceria da Fundação de Proteção e Defesa do Consumidor (Fundação PROCON) e o Departamento Intersindical de Estudos Socioeconômicos (DIEESE) calculam a cesta básica paulistana de 32 produtos, sendo 22 de alimentação. Essa pesquisa é realizada no setor supermercadista. A Fundação Instituto de Pesquisas Econômicas (FIPE) e o Instituto Brasileiro de Geografia e Estatística (IBGE) coletam preços de alimentos na cidade de São Paulo para compor mensalmente seus IPCs.

Coletando preços desde o início dos anos 1970, o IEA se consolidou como uma das principais fontes de informações do mercado varejista da cidade de São Paulo. Este trabalho propõe uma nova metodologia amostral probabilística considerando o novo cenário da cidade, alicerçado em sólidos pilares de informação estatística.

2 - JUSTIFICATIVA

Analisando a cidade de São Paulo nos dias de hoje, pode-se verificar claramente as várias transformações sofridas nestes últimos 30 anos. Entre elas se destaca o aumento da população, a nova distribuição territorial, a crescente heterogeneidade em termos socioeconômicos e mudanças no perfil do consumidor.

Segundo dados do último censo populacional (IBGE, 2000), sua população é de 10,4 milhões de pessoas, superando em mais de 70% a população registrada em dezembro de 1969, que era de 6,1 milhões/hab. (JUNQUEIRA; CANTO, 1971). Esse aumento na população ocorreu em grande parte nas regiões periféricas⁵. A população residente nessas regiões supera a marca de 6 milhões de habitantes. Como o modelo amostral vigente é anterior a este crescimento demográfico, ele hoje não contempla a totalidade da população da cidade de São Paulo.

Na época da implementação do sistema de coleta de dados no mercado varejista pelo IEA, a cidade de São Paulo era dividida em oito

distritos, entre eles o de São Paulo, que correspondia a 93,9% do total. O distrito de São Paulo era dividido em 48 subdistritos e foi nesse universo que se estruturou o levantamento sistemático de preços no varejo na cidade de São Paulo. A partir da Lei de Territorialização n. 11.220/92 (SÃO PAULO, 1992), a cidade passou a ser dividida geograficamente em 96 distritos (Figura 1). Essa nova divisão territorial procurou refletir mais claramente a nova realidade da cidade, em que as suas três grandes regiões (leste, norte e sul) foram divididas cada uma em área I e II. A área I de cada região se refere aos distritos mais próximos do centro da capital e a II compreende os distritos periféricos da cidade.

Diversos estudos no campo socioeconômico comentam sobre a heterogeneidade da cidade de São Paulo em seus 96 distritos. Eles vão desde indicadores de desenvolvimento humano comparáveis aos de países do primeiro mundo como é o caso do distrito Jardim Paulista, até o seu inverso, observado por exemplo no distrito de Marsilac, que apresenta péssimos indicadores de qualidade de vida (SPOSATI, 2001; TAKIYA, 2002). Fazendo uma análise distrital, Sposati (2001) concluiu que cada um dos distritos do município é uma cidade, dadas as dimensões territoriais e populacionais e as características socioeconômicas de cada um deles. Também estudando a cidade de São Paulo através de seus distritos, Takiya (2002) reuniu um elenco de variáveis socioeconômicas e através de análise multivariada criou um indicador de desenvolvimento humano chamado de fator principal.

Em ambos os estudos é possível perceber claramente que a população paulistana possui características socioeconômicas extremamente heterogêneas, exigindo que um modelo amostral possa refletir tal realidade.

Em relação ao perfil do consumidor, percebem-se importantes mudanças nos hábitos de compra de produtos alimentícios na cidade de São Paulo (MARTINS; MARGARIDO; BUENO, 2007). Para fins de delineamento amostral, a mudança de local de compra de produtos é a mais importante. Comparando os resultados das Pesquisas de Orçamentos Familiares (POF) de 1981/82 e 1998/99 da FIPE (1981/82; 1998/99), verifica-se que, em relação ao local de compra de frutas, legumes e verduras (FLV), as feiras livres perderam muito espaço para os supermercados (Tabela 1). A crescente participação dos supermercados na distribuição de FLV na cidade de

⁵Entende-se por regiões periféricas aquelas que se localizam nos extremos da cidade distantes da região central.

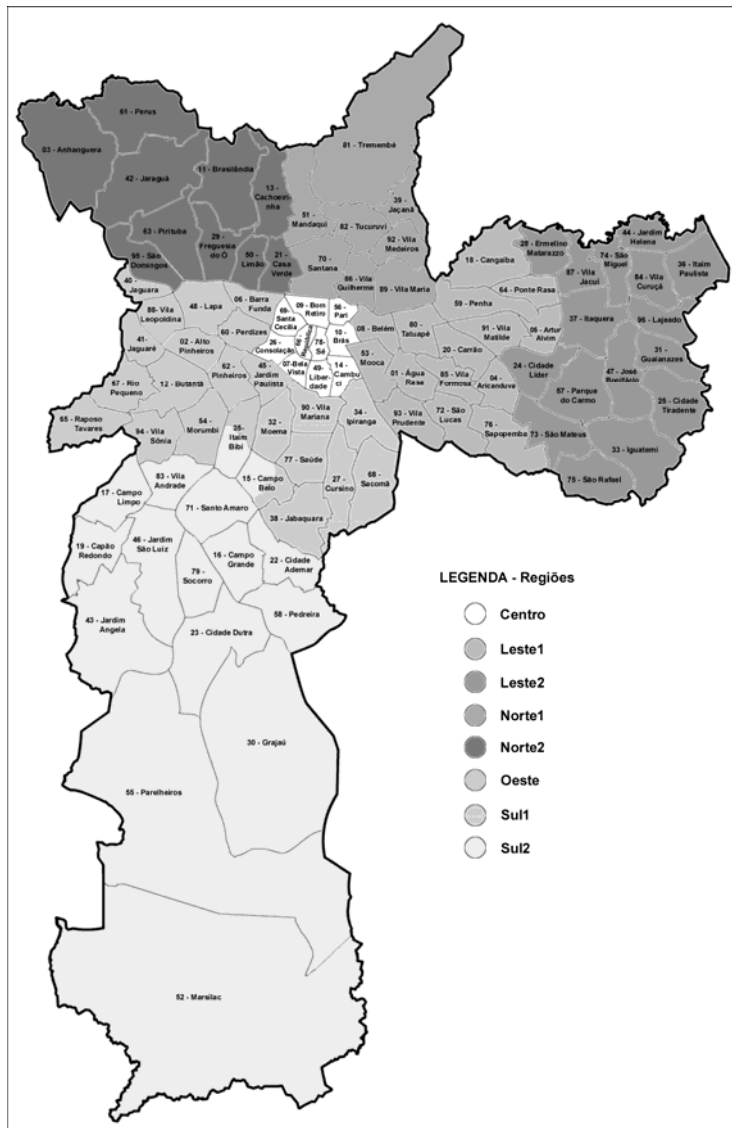


Figura 1 - Divisão Distrital da Cidade de São Paulo.
Fonte: Meyer; Grostein; Biderman (2003).

TABELA 1 - Participação e Variação Percentual dos Equipamentos Feiras Livres e Supermercado na Venda de FLV, Cidade de São Paulo, 1981/82 e 1998/99

Produto	1981/82		1998/99		Variação entre períodos (%)	
	Feira livre	Supermercado	Feira livre	Supermercado	Feira livre	Supermercado
Frutas	0,92	0,05	0,54	0,33	-41,14	577,11
Legumes	0,88	0,07	0,56	0,33	-35,86	344,44
Verduras	0,89	0,06	0,62	0,28	-30,28	374,47
Média	0,89	0,06	0,57	0,31	-35,76	432,01

Fonte: Elaborada pelos autores a partir de FIPE (1981/82; 1998/99).

São Paulo está intimamente ligada ao desenvolvimento do setor supermercadista, principalmente a área logística e o oferecimento de diversos serviços em um mesmo local, e em especial com a chegada de grandes redes varejistas no final da década de 1970 (MARTINS; MARGARIDO; BUENO, 2007; SESSO FILHO, 2003). Também é fato importante para a explicação desta mudança a crescente participação da mulher no mercado de trabalho. Segundo dados da Fundação Sistema Estadual de Análise de Dados (SEADE), a população economicamente ativa feminina⁶ em 2005 na região metropolitana de São Paulo foi de 55,5% (SEADE, 2007).

3 - MATERIAL

Neste trabalho foram utilizados dados da POF (FIPE, 1998/99) referentes à alimentação dentro do domicílio na cidade de São Paulo, indicadores de desenvolvimento humano na cidade de São Paulo, conforme Takiya (2002), e distritais do município, conforme Meyer; Grostein; Biderman (2003), e informes populacionais censitários do IBGE referentes à cidade de São Paulo no ano 2000.

Em relação ao sistema de referências dos equipamentos, foram utilizadas duas fontes. Para feiras livres e sacolões municipais, utilizaram-se cadastros oriundos da Prefeitura do Município de São Paulo, através do site no ano de 2006, e os referentes a sacolões, açougues, padarias e supermercados particulares foram coletados no guia OESP no site correspondente, também no ano de 2006 (SÃO PAULO, 2006; OESP, 2006).

4 - METODOLOGIA

Na primeira fase realizou-se depuração no conteúdo da POF (FIPE, 1998/99) para identificação dos itens alimentares⁷, que seriam objeto de estudo deste trabalho, através das variáveis tipo, variedade, unidade e equipamento de compra. Na etapa seguinte, procedeu-se a constru-

ção de estratos de forma a obter sub-populações para que a variabilidade fosse pequena dentro deles. De acordo com estudos de Sposati (2001) e Takiya (2002), optou-se por estrato de tipo de equipamento e classe socioeconômica.

O tamanho da amostra foi calculado a partir de estimativas de média e variância dos itens alimentares por meio de levantamento piloto. A distribuição do número de elementos pelos estratos seguiu a alocação de Neyman (BOLFA-RINE; BUSSAB, 2005). Para complementar o estudo, foram realizadas comparações do delineamento amostral baseadas no método estratificado com o de amostragem aleatória simples através do efeito do planejamento. Para tal, foi calculada a razão entre a estimativa da variância determinada pelo plano amostral e a estimativa da variância obtida pela amostra aleatória simples de mesmo tamanho (KISH, 1965).

4.1 - Depuração da POF/FIPE 1998/99

Os dados da POF 1998/99 foram depurados por faixa de renda⁸: dois a seis salários mínimos (SMs)⁹; seis a dez SMs; dez ou mais SMs; e faixa média ponderada¹⁰. Para todos os produtos foram estudadas as seguintes características: tipo/variedade, marca, embalagem, local de compra e quantidade consumida.

Foram estudados 434 itens em cinco grupos: produtos de origem animal semielaborados, produtos de origem animal industrializados, produtos de origem vegetal *in-natura*, produtos de origem vegetal semielaborados e produtos de origem vegetal industrializados. Dentro de cada grupo, os itens foram separados em subgrupos e selecionados conforme a sua importância referente aos critérios de quantidade consumida e de dispêndio familiar¹¹. O Anexo 1 apresenta a relação dos produtos sem as suas diversas variações em relação a tipo e ou variedade e unidade.

⁶Corresponde à parcela da população em idade ativa que está ocupada ou empregada.

⁷Os produtos em suas diferentes formas de apresentação em relação a unidades e variedades são nomeados itens alimentares neste trabalho.

⁸Não houve depuração para a renda inferior a dois salários mínimos porque nenhum dos 96 distritos do município de São Paulo possui renda menor do que essa faixa.

⁹Salário mínimo de R\$200,00 em dezembro de 1999.

¹⁰A ponderação é referente à população residente em cada faixa de renda.

¹¹Segundo dados da Fundação SEADE (2000), uma família de tamanho médio em São Paulo é composta por 3,78 pessoas.

4.2 - Amostragem Probabilística Estratificada em Dois Estágios

Basicamente, a amostragem probabilística estratificada objetiva reduzir as distorções verificadas na população, através de sua divisão em estratos homogêneos. Visa também elaborar uma amostra consistente a um custo menor (BOLFARINE; BUSSAB, 2005).

De acordo com as características da população no presente estudo, optou-se por utilizar o recurso desse tipo de estratificação para a elaboração do delineamento amostral. No primeiro, determinaram-se estratos constituídos de distritos com características socioeconômicas semelhantes, de forma a reduzir as grandes distorções verificadas nos estudos socioeconômicos considerados. Peceguini (1979) enumerou alguns motivos para estratificar por equipamento varejista, sendo o principal a variabilidade dos preços médios e variabilidades diferentes dos produtos. Portanto, o segundo nível refere-se à estratificação por equipamento varejista.

4.3 - Levantamento Piloto

As estimativas de média e variância podem ser obtidas por levantamento piloto, resultados de amostragens anteriores da mesma população ou semelhante e por conjecturas sobre a estrutura da população (COCHRAN, 1965). Neste caso utilizou-se o recurso do levantamento piloto de dados que *“por definição é uma amostra pequena, que antecede a principal e que é utilizada para calcular o tamanho desta última”* (SANTIANO et al., 2000).

Com base na POF 1998/99, foram selecionados os 434 itens mais consumidos pelas famílias paulistanas e os cinco equipamentos com maior frequência de venda: supermercados, feiras livres, açougues, sacolões e padarias. Com essa segmentação foram aplicados questionários específicos para cada equipamento.

A coleta dos dados¹² foi realizada entre junho e dezembro de 2004 em 78 dos 96 distritos da cidade de São Paulo dentro de todos os estratos socioeconômicos. Essa abrangência corres-

¹²Os autores agradecem os estagiários de nível médio Thiago Soares Velença e Alex Shimazaki de Menezes pelo trabalho de coleta das informações.

ponde a mais de 80% do total de distritos da cidade.

4.4 - Cálculo do Tamanho da Amostra¹³

Com base no levantamento piloto, foram calculadas a média e a variância dentro de cada estrato (equações 1 e 2).

$$\bar{y}_{KJi} = \sum_{KJi=1}^{n_{KJi}} \frac{y_{KJi}}{n_{KJi}} \quad (1)$$

sendo:

\bar{y}_{KJi} = média do item alimentício i dentro do equipamento J e classe socioeconômica K ;

y_{KJi} = preço do item alimentício i dentro do equipamento J e classe socioeconômica K ;

n_{KJi} = amostra piloto do item alimentício i dentro do estrato de equipamento J no estrato socioeconômico K .

$$S_{KJi}^2 = \sum_{KJi=1}^{n_{KJi}} \frac{(y_{KJi} - \bar{y}_{KJi})^2}{n_{KJi} - 1} \quad (2)$$

sendo:

S_{KJi}^2 = variância do item alimentício i dentro do estrato de equipamento J no estrato socioeconômico K ;

\bar{y}_{KJi} = média do item alimentício i dentro do estrato de equipamento J no estrato socioeconômico K ;

y_{KJi} = preço do item alimentício i dentro do estrato de equipamento J no estrato socioeconômico K ;

n_{KJi} = amostra piloto do item alimentício i dentro do estrato de equipamento J no estrato socioeconômico K .

¹³O cálculo da amostra teve como referências Cochran (1965) e Bolfarine e Bussab (2005).

Obtidas as estimativas do item alimentício i , seguiu-se ao processo de obtenção da média e da variância no equipamento J dentro da classe socioeconômica K ponderada por meio do peso da unidade da amostra nesse estágio definido da fração amostral (equações 3, 4 e 5).

$$W_{JK} = \frac{N_{JK}}{N_J} \quad (3)$$

sendo:

W_{JK} = fração amostral do equipamento J no estrato socioeconômico K ;

N_{JK} = número de equipamentos J dentro da classe socioeconômica K ;

N_J = número total de equipamentos J .

$$\bar{y}_{est_{Ji}} = \sum_{Ji=1}^{n_{Ji}} W_J * \bar{y}_{KJi} \quad (4)$$

sendo:

$\bar{y}_{est_{Ji}}$ = média do item alimentício i no equipamento J .

$$S_{est_{Ji}}^2 = \sum_{Ji=1}^{n_{Ji}} W_J * S_{KJi}^2 \quad (5)$$

sendo:

$S_{est_{Ji}}^2$ = variância do item alimentício i no equipamento J .

A variância fixada será calculada conforme Junqueira e Canto (1971), que determinaram como aceita a de 5% da média de cada item alimentício (equação 6).

$$L = \bar{y}_{est_{Ji}} * 0,05 \quad (6)$$

A partir das estimativas nos estratos e da variância fixada, o tamanho da amostra para cada item alimentício foi calculado da seguinte forma (equação 7):

$$n^0 = t_{0,05}^2 * \frac{\sum_{Ji=1}^{n_{Ji}} W_J * S_{est_{Ji}}^2}{L} \quad (7)$$

sendo:

n^0 = número de amostras para cada item alimentício em cada estrato de equipamento a um nível de 5%;

$t_{0,05}$ = valor da tabela t de *student* a 5%.

Como se trata de uma população finita, fez-se a correção para populações finitas (equação 8).

$$n = \frac{n^0}{\left(1 + \frac{n^0}{N}\right)} \quad (8)$$

sendo:

n = amostra corrigida.

A amostra corrigida foi dividida com base nas variâncias de correspondentes a classe socioeconômica K e equipamento J em cada item alimentício i . Essa partição seguiu a metodologia de alocação de Neyman (equação 9).

$$n_h = \frac{n * W_J * CV_{est_{ji}}}{\sum W_J * CV_{est_{ji}}} \quad (9)$$

sendo:

n_h = número de elementos na amostra por estrato socioeconômico K e equipamento J ;

$CV_{est_{ji}}$ = coeficiente de variação estratificado do item alimentício i e equipamento J .

O efeito de planejamento foi calculado da seguinte forma (equação 10):

$$e_p = \left(\frac{S_{AE2}^2}{S_{AAS}^2} \right) * 100 \quad (10)$$

sendo:

e_p = efeito do planejamento;

S_{AE2}^2 = variância média do delineamento amostral estratificado em 2 estágios;

S_{AAS}^2 = variância média do delineamento amostral simples.

O intervalo de confiança em um nível de significância de 5% foi calculado para cada item alimentício em cada estrato por equipamento, através da equação 11.

$$IC = \bar{y}_{est_{ji}} \pm t_{0,05} * \sqrt{\sum_{KJi=1}^{n_{KJi}} W_J^2 * \frac{S_{KJi}^2}{n_{KJi}}} \quad (11)$$

sendo:

IC = intervalo de confiança para cada item alimentício i em cada estrato K por equipamento J ;

$\bar{y}_{est_{ji}}$ = média estratificada de cada item alimentício i em cada estrato de equipamento J ;

$t_{0,05}$ = valor da tabela t de *student* para um nível de significância de 5%;

W_J^2 = fração amostral do equipamento J ;

S_{KJi}^2 = variância de cada item alimentício i dentro de cada estrato de equipamento J em cada estrato socioeconômico K ;

n_{KJi} = amostra de cada item alimentício i dentro de cada estrato de equipamento J em cada estrato socioeconômico K .

5 - ANÁLISE DOS RESULTADOS

Baseado nos estudos de Sposati (2001) sobre as distorções socioeconômicas dos distritos paulistanos e do indicador construído por Takiya (2002), optou-se por um delineamento amostral estratificado em dois estágios: distritais e principais locais de compra do paulistano.

No primeiro nível de estratificação, a segmentação distrital da cidade de São Paulo procedeu por meio de indicador de desenvolvimento humano, baseado em parâmetros socioeconômicos, conforme Takiya (2002), isto é, semelhantes em relação às suas características socioeconômicas. Esse autor analisou as variáveis distribuição territorial, perfil etário, nível de renda, escolaridade, condições de saúde, condições de habitação e segurança. Através da análise fatorial (MINGOTI, 2005) das variáveis, a autora obteve um indicador denominado de "fator principal". A partir deste resultado os 96 distritos foram

divididos em 6 estratos (Tabela 2 e Figura 2).

TABELA 2 - População Residente por Estrato Regional, Participação Percentual e Acumulada, Cidade de São Paulo, 2000

Estrato	População	Participação (%)	Acumulado (%)
I	3.241.165	0,31	0,31
II	2.083.658	0,20	0,51
III	1.881.445	0,18	0,69
IV	1.319.309	0,13	0,82
V	826.221	0,08	0,90
VI	1.054.069	0,10	1,00
Total	10.405.069	1,00	-

Fonte: IBGE (2000).

Segundo dados do último censo populacional (IBGE, 2000), a população residente na cidade de São Paulo é de 10,4 milhões de habitantes. A população residente em cada estrato está disposta na tabela 2.

Analisando a espacialização dessa estratificação, verifica-se que o estrato I contempla os distritos com os piores indicadores socioeconômicos da cidade, com concentração nas regiões leste II e sul II. Com exceção do distrito de Brasilândia, que se localiza na região norte II, esses distritos se encontram em áreas periféricas da cidade (Figura 2 e Tabela 3).

O estrato II se compõe de distritos basicamente das regiões sul, norte e leste da cidade, sendo os das regiões norte e sul não tão distantes da região central e os de Jaraguá e Perus (norte) localizados bem ao extremo da cidade.

Estes dois estratos são compostos por 33 distritos que, somados, são responsáveis pela residência de pouco mais da metade da população paulistana, 51%.

Os distritos do estrato III estão distribuídos praticamente em todas as regiões, mantendo pouca relação de vizinhança.

O estrato IV é formado em sua maioria por distritos das regiões leste I, norte I e central e é neste que o indicador socioeconômico passa a ser positivo.

O estrato V é formado por distritos da região oeste, relativamente afastados da região central, como o de Butantã e alguns da região central e da região leste da cidade. Este estrato concentra a menor parcela da população, perfazendo 8% do total.

No estrato VI, formado em sua totalidade

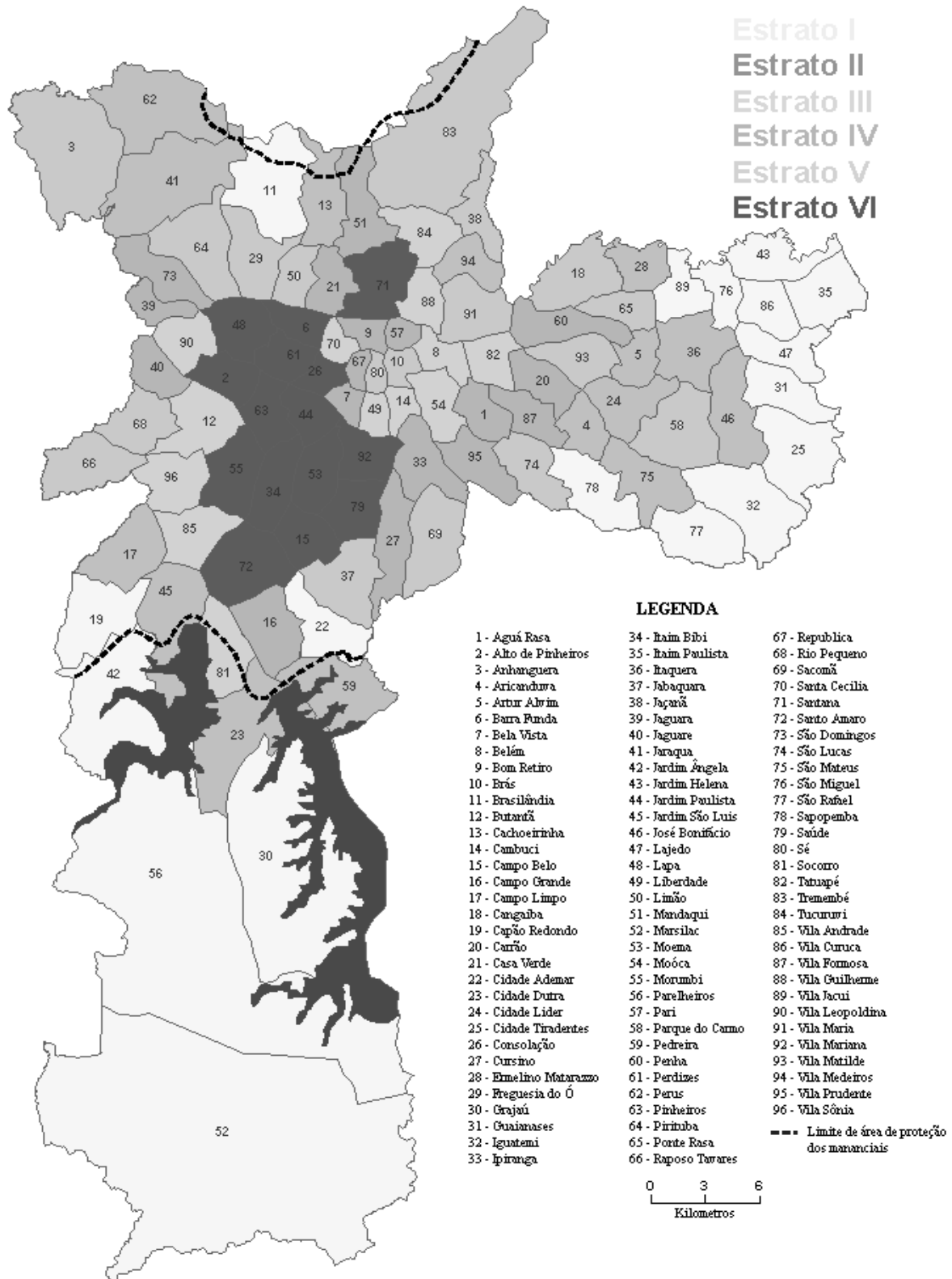


Figura 2 - Estratificação Socioeconômica da Cidade de São Paulo.
 Fonte: Takiya (2002) e dados da pesquisa.

TABELA 3 - Estratificação Socioeconômica dos Distritos da Cidade de São Paulo por Intervalos de Fator Principal, Cidade de São Paulo, 2006

Distrito	Fator principal	Estrato	Intervalo	Distrito	Fator principal	Estrato	Intervalo
Marsilac	-1,86	I	<(-1,00)	Penha	0,09	IV	>(0,00)<=(0,49)
Cid. Tiradentes	-1,70	I	<(-1,00)	Casa Verde	0,10	IV	>(0,00)<=(0,49)
Lajeado	-1,55	I	<(-1,00)	S. Domingos	0,10	IV	>(0,00)<=(0,49)
Brasilândia	-1,44	I	<(-1,00)	Bom Retiro	0,11	IV	>(0,00)<=(0,49)
Itaim Paulista	-1,41	I	<(-1,00)	V. Formosa	0,12	IV	>(0,00)<=(0,49)
Guaianases	-1,40	I	<(-1,00)	Pari	0,13	IV	>(0,00)<=(0,49)
Vila Curuçá	-1,38	I	<(-1,00)	V. Prudente	0,21	IV	>(0,00)<=(0,49)
Jardim Ângela	-1,31	I	<(-1,00)	Jaguará	0,23	IV	>(0,00)<=(0,49)
Iguatemi	-1,28	I	<(-1,00)	Ipiranga	0,24	IV	>(0,00)<=(0,49)
Jd. Helena	-1,25	I	<(-1,00)	República	0,24	IV	>(0,00)<=(0,49)
Sapopemba	-1,17	I	<(-1,00)	Mandaqui	0,29	IV	>(0,00)<=(0,49)
V. Jacuí	-1,10	I	<(-1,00)	Cursino	0,35	IV	>(0,00)<=(0,49)
Capão Redondo	-1,08	I	<(-1,00)	Bela Vista	0,36	IV	>(0,00)<=(0,49)
São Miguel	-1,08	I	<(-1,00)	Carrão	0,36	IV	>(0,00)<=(0,49)
Grajaú	-1,07	I	<(-1,00)	Água Rasa	0,39	IV	>(0,00)<=(0,49)
Cid. Ademar	-1,04	I	<(-1,00)	Cpo. Grande	0,43	IV	>(0,00)<=(0,49)
Parelheiros	-1,03	I	<(-1,00)	Jaguaré	0,49	IV	>(0,00)<=(0,49)
São Rafael	-1,01	I	<(-1,00)	Belém	0,52	V	>(0,50)<=(1,00)
Jd. São Luís	-0,93	II	>(-0,99)<=(0,51)	V. Guilherme	0,52	V	>(0,50)<=(1,00)
Itaquera	-0,86	II	>(-0,99)<=(0,51)	Santa Cecília	0,53	V	>(0,50)<=(1,00)
Perus	-0,86	II	>(-0,99)<=(0,51)	V. Andrade	0,57	V	>(0,50)<=(1,00)
Cachoeirinha	-0,83	II	>(-0,99)<=(0,51)	V. Sônia	0,57	V	>(0,50)<=(1,00)
Cidade Líder	-0,77	II	>(-0,99)<=(0,51)	Liberdade	0,61	V	>(0,50)<=(1,00)
Jaraguá	-0,76	II	>(-0,99)<=(0,51)	Tucuruvi	0,69	V	>(0,50)<=(1,00)
São Mateus	-0,74	II	>(-0,99)<=(0,51)	Tatuapé	0,73	V	>(0,50)<=(1,00)
Artur Alvim	-0,73	II	>(-0,99)<=(0,51)	Moóca	0,86	V	>(0,50)<=(1,00)
Erm. Matarazzo	-0,73	II	>(-0,99)<=(0,51)	Cambuci	0,92	V	>(0,50)<=(1,00)
Cpo. Limpo	-0,72	II	>(-0,99)<=(0,51)	V. Leopoldina	0,93	V	>(0,50)<=(1,00)
V. Medeiros	-0,69	II	>(-0,99)<=(0,51)	Socorro	0,94	V	>(0,50)<=(1,00)
José Bonifácio	-0,64	II	>(-0,99)<=(0,51)	Butantã	0,97	V	>(0,50)<=(1,00)
Aricanduva	-0,57	II	>(-0,99)<=(0,51)	Santana	1,02	VI	>(1,01)
Cid. Dutra	-0,54	II	>(-0,99)<=(0,51)	Saúde	1,05	VI	>(1,01)
Pedreira	-0,53	II	>(-0,99)<=(0,51)	Barra Funda	1,23	VI	>(1,01)
Sé	-0,48	III	>(-0,50)<=(0,00)	Lapa	1,30	VI	>(1,01)
Anhanguera	-0,44	III	>(-0,50)<=(0,00)	Cpo. Belo	1,45	VI	>(1,01)
Pq. do Carmo	-0,42	III	>(-0,50)<=(0,00)	Consolação	1,46	VI	>(1,01)
São Lucas	-0,41	III	>(-0,50)<=(0,00)	Perdizes	1,46	VI	>(1,01)
Ponte Rasa	-0,40	III	>(-0,50)<=(0,00)	V. Mariana	1,51	VI	>(1,01)
Sacomã	-0,38	III	>(-0,50)<=(0,00)	Itaim Bibi	1,69	VI	>(1,01)
Raposo Tavares	-0,36	III	>(-0,50)<=(0,00)	Pinheiros	1,83	VI	>(1,01)
Jaçanã	-0,35	III	>(-0,50)<=(0,00)	Morumbi	2,14	VI	>(1,01)
Freg. do Ó	-0,28	III	>(-0,50)<=(0,00)	Alto Pinheiros	2,15	VI	>(1,01)
Pirituba	-0,27	III	>(-0,50)<=(0,00)	Santo Amaro	2,25	VI	>(1,01)
V. Maria	-0,22	III	>(-0,50)<=(0,00)	Moema	2,30	VI	>(1,01)
Rio Pequeno	-0,21	III	>(-0,50)<=(0,00)	Jd. Paulista	2,48	VI	>(1,01)
Cangaíba	-0,15	III	>(-0,50)<=(0,00)				
Jabaquara	-0,15	III	>(-0,50)<=(0,00)				
Limão	-0,13	III	>(-0,50)<=(0,00)				
V. Matilde	-0,13	III	>(-0,50)<=(0,00)				
Tremembé	-0,07	III	>(-0,50)<=(0,00)				
Brás	-0,03	III	>(-0,50)<=(0,00)				

Fonte: Takiya (2002).

de por distritos das regiões oeste, sul I e central, pode-se verificar a vizinhança dos distritos de sua composição, ficando a exceção para o de Santana, localizado na região norte I.

O segundo nível de estratificação, equipamento varejista, baseou-se na POF 1998/99, em que foram analisados todos os equipamentos citados na pesquisa. A determinação dos equipamentos foi realizada com base na frequência média de compra pelos consumidores paulistanos e este critério foi o que determinou a escolha dos cinco equipamentos: açougue, feira livre, padaria, sacolões e supermercados. Portanto, os cinco primeiros cobriam 100% dos produtos levantados. O equipamento de maior frequência média foi o supermercado com 97%, ou seja, na média dos grupos, em apenas 3% dos itens pesquisados o supermercado não é um dos principais locais de compra. A feira livre, que segundo Martins; Margarido; Bueno (2007) perdeu aproximadamente 35% de sua participação na venda de FLV para os supermercados, se comparada com dados da POF 1981/82, é o segundo equipamento na venda de produtos alimentícios na cidade de São Paulo com 31%, seguida do açougue com 23%, do sacolão com 19% e da padaria com 18% (Tabela 4).

A partir dos níveis de estratificação, prosseguiu-se com a determinação da amostra piloto. Foram selecionados 223 equipamentos, sendo 80 supermercados, 54 feiras livres, 30 sacolões, 32 açougues e 27 padarias distribuídos nas classes socioeconômicas (Tabela 5). O tamanho da amostra piloto por definição é subjetiva e sua partição pelos estratos foi definida pela importância de cada equipamento.

Quanto aos resultados da amostra piloto, alguns coeficientes de variação para os equipamentos utilizados neste estudo apresentaram grandes amplitudes (Tabela 6). Uma causa provável é a extensa lista de itens alimentícios pesquisados aqui e a pequena quantidade de mão-de-obra especializada na coleta de preços que ocasionou um prolongado período de coleta (sete meses, de junho a dezembro de 2004). Esse fato pode ter afetado as estimativas de alguns itens, em especial frutas, que são fortemente influenciadas por efeitos sazonais. Apesar dessa dificuldade, o coeficiente de variação médio ficou próximo de 20%, considerado aceitável por Gomes e Pimentel (1984).

Obtidas as estimativas das variâncias oriundas do levantamento piloto e aplicando-as

no delineamento amostral sugerido, e considerando o maior número de itens alimentícios e erro de amostragem admissível de 5%, o tamanho da amostra resultou em 576 equipamentos. Os elementos foram distribuídos pelos estratos por meio da alocação ótima de Neyman, supondo custo fixo (Tabela 7).

Quanto à composição da amostra calculada, verifica-se que contempla 98% dos itens alimentícios em estudo neste trabalho. Os itens mexerica, limão e laranja necessitam de um tamanho de amostra de 284 supermercados. No equipamento feira, os itens maçã, limão, goiaba e tangerina necessitam de 215 locais de coleta para que todos os produtos estejam dentro do limite de 5% de erro. Para os equipamentos sacolão, açougue e padaria, o tamanho da amostra é suficiente para que todos os itens não ultrapassem o erro admitido.

A limitação no tamanho da amostra em 576 equipamentos foi definida por ser uma amostra factível de implementação sistemática no sistema de coleta de informações no mercado varejista pelo IEA e, não se justifica um aumento significativo na amostra para contemplar poucos itens.

Estabelecida a amostra, procedeu-se o cálculo do efeito do planejamento amostral (Tabela 8). No equipamento padaria, foi verificada a maior redução de variância média (79%). Já nos equipamentos sacolão e feira livre, a variância foi reduzida a menos da metade, 53% e 58%, respectivamente. Nota-se que, nestes três equipamentos, a localização geográfica é um fator determinante para a determinação da valoração dos itens alimentícios. Por outro lado, nos equipamentos supermercado e açougue houve menor redução, 29% e 9% respectivamente. Em relação ao supermercado, a redução menor pode ser explicada devido à ampla concorrência de grandes redes varejistas. O grupo de carnes *in natura* que predomina no equipamento açougue apresentou pouca variação nos estratos socioeconômicos. Neste equipamento, o grande número de pontos de venda em todas as regiões da cidade limita a variação de preços.

6 - CONSIDERAÇÕES FINAIS

A cidade de São Paulo nos dias de hoje é resultado de várias transformações que ocorreram nestes últimos 30 anos. Entre elas se des-

TABELA 4 - Frequência de Compra por Equipamento Varejista, Cidade de São Paulo, 2006

Equipamentos ¹	Frequência de venda em cada grupo (%)				Média
	Animal	Animal indústria	Vegetal	Vegetal indústria	
Supermercado	0,96	1,00	0,97	0,95	0,97
Feira livre	0,26	0,05	0,90	0,02	0,31
Açougue	0,70	0,18	0,00	0,05	0,23
Sacolão	0,09	0,05	0,55	0,05	0,19
Padaria	0,09	0,45	0,01	0,17	0,18
Vendedor ambulante	0,07	0,05	0,12	0,02	0,07
Avícola	0,22	0,00	0,03	0,00	0,06
Casa de frios	0,00	0,14	0,00	0,00	0,04
Bar	0,02	0,00	0,01	0,10	0,03
Loja de conveniência	0,00	0,05	0,02	0,05	0,03
Peixaria	0,09	0,00	0,00	0,00	0,02
Quitanda	0,00	0,00	0,09	0,00	0,02

¹Equipamentos selecionados em destaque.

Fonte: FIPE (1998/99).

TABELA 5 - Distribuição da Amostra Piloto por Estrato Socioeconômico e por Equipamento, Cidade de São Paulo, 2006

Estrato socioeconômico	Equipamento					Total
	Supermercados	Feiras livres	Sacolão	Padaria	Açougue	
I	13	8	7	4	4	36
II	10	11	3	3	5	32
III	19	12	4	4	8	47
IV	18	10	4	7	7	46
V	8	5	6	5	4	28
VI	12	8	6	4	4	34
Total	80	54	30	27	32	223

Fonte: Dados da pesquisa.

TABELA 6 - Coeficiente de Variação Máximo, Mínimo e Médio por Equipamento e Total, Cidade de São Paulo, 2006

Equipamento	Cv _{máx}	Cv _{méd}	Cv _{mín}
Supermercado	53,94	22,12	2,46
Açougue	39,82	14,95	0,00
Sacolão	48,52	27,07	0,00
Feira livre	42,58	24,22	0,00
Padaria	17,01	7,40	0,00
Total	53,94	21,99	0,00

Fonte: Dados da pesquisa.

tacam o aumento de sua população, que é de 10,4 milhões de pessoas (IBGE, 2000), sua nova distribuição territorial, sua crescente heterogeneidade em termos socioeconômicos e mudanças no perfil do consumidor.

Esse fato foi um dos indicadores para a mudança no modelo amostral de levantamento de preços no varejo vigente na instituição, devido a ser anterior a este crescimento demográfico e

não contemplar essa totalidade da população paulistana.

O estabelecimento de uma amostra com o intuito de estimar preços médios de produtos alimentícios através do recurso da estratificação em dois estágios, considerando as grandes distorções socioeconômicas verificadas na cidade de São Paulo e mantendo-se segmentação por equipamento varejista sugerido por Peceguini

TABELA 7 - Número de Elementos na Amostra por Classe Socioeconômica e por Equipamento, Cidade de São Paulo, 2006

Estrato	Equipamento	Amostra
I	Supermercado	35
II	Supermercado	37
III	Supermercado	29
IV	Supermercado	44
V	Supermercado	27
VI	Supermercado	48
Total	Supermercado	220
I	Açougue	14
II	Açougue	16
III	Açougue	22
IV	Açougue	15
V	Açougue	13
VI	Açougue	21
Total	Açougue	100
I	Sacolão	15
II	Sacolão	6
III	Sacolão	20
IV	Sacolão	11
V	Sacolão	16
VI	Sacolão	16
Total	Sacolão	83
I	Feira	24
II	Feira	35
III	Feira	22
IV	Feira	21
V	Feira	17
VI	Feira	11
Total	Feira	130
I	Padaria	4
II	Padaria	3
III	Padaria	11
IV	Padaria	9
V	Padaria	5
VI	Padaria	11
Total	Padaria	43
-	Geral	576

Fonte: Dados da pesquisa.

TABELA 8 - Efeito do Planejamento entre os Métodos de Amostragem Aleatória Simples e Amostragem Estratificada em Dois Estágios, Cidade de São Paulo, 2006

Equipamento	AAS	AE 2	Efeito do planejamento
Supermercado	1,48	1,05	0,29
Açougue	1,83	1,66	0,09
Sacolão	1,16	0,55	0,53
Feira livre	2,27	0,95	0,58
Padaria	3,07	0,64	0,79

Fonte: Dados da pesquisa.

(1979), é viável, segundo a apresentação neste estudo.

Tanto a estratificação através de parâmetros socioeconômicos como também de equipamentos varejistas se mostraram eficientes como instrumentos para a redução do tamanho

amostral e minimização da heterogeneidade. Comparando o delineamento amostral sugerido àquele de amostragem simples semelhante ao vigente, este método apresentou melhores resultados, estando em condições de uso imediato.

LITERATURA CITADA

BOLFARINE, H.; BUSSAB, W. O. **Elementos de amostragem**. São Paulo: Edgard Blucher, 2005. 274 p.

CARMO, H. C. E. do. **A teoria dos índices de preços e o sistema de metas de inflação no Brasil**. São Paulo: FEA/USP, 2004. 74 p. (Programa de Seminários Acadêmicos, n. 23).

COCHRAN, W. G. **Técnicas de amostragem**: John Wiley & Sons, Inc., New York. Rio de Janeiro: USAID, 1965. 428 p.

FUNDAÇÃO DE INSTITUTO DE PESQUISAS ECONÔMICAS - FIPE. **Pesquisa de Orçamentos Familiares - POF**. São Paulo: FIPE, 1981/82. Disponível em: <<http://www.fipe.org.br>>. Acesso em: out. 2006.

_____. **Pesquisa de Orçamentos Familiares - POF**. São Paulo: FIPE, 1998/99. Disponível em: <<http://www.fipe.org.br>>. Acesso em: out. 2006.

FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS - SEADE. **Mulher e trabalho**: a mulher no mercado de trabalho. Disponível em: <<http://www.seade.gov.br/produtos/mulher/index.php>>. Acesso em: ago. 2007.

GOMES, F.; PIMENTEL, A. **Estatística moderna na pesquisa agropecuária**. Piracicaba (SP): POTAFOS, 1984. 160 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Censo Demográfico 2000**. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/default_censo_2000.shtml>. Acesso em: fev. 2008.

JUNQUEIRA, P. C.; CANTO, W. L. Margens totais de comercialização. **Cesta de Mercado**, São Paulo, v. 18, n. 9/10, set./out. 1971.

KISH, L. **Survey sampling**. New York: John Wiley Professio, 1965. 664 p.

MARTINS, V. A.; MARGARIDO, M. A.; BUENO, C. R. F. Alteração no perfil de compra de frutas, legumes e verduras nos supermercados e feiras-livres na cidade de São Paulo. **Informações Econômicas**, São Paulo, v. 37, n. 2, p. 30-37, fev. 2007.

MEYER, R. M. P.; GROSTEIN, M. D.; BIDERMAN, C. **São Paulo metrópole**. São Paulo: EDUSP/Imprensa Oficial, 2003. 296 p.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte: Editora UFMG, 2005. 297 p.

OESP. **Cadastro de empresas da cidade de São Paulo**. Disponível em: <<http://www.listaoesp.com.br>>. Acesso em: jul. 2005.

PECEGUINI, E. E. **Redimensionamento da amostra para levantamento de preços a nível de varejo na cidade de São Paulo**. São Paulo: IEA, 1979.

PINO, F. A. Estatísticas agrícolas para o século XXI. **Agricultura em São Paulo**, v. 46, n. 2, p. 71-105, 1999.

Informações Econômicas, SP, v.39, n.8, ago. 2009.

SANTIAGO, M. M. D. et al. Mercado atacadista de produtos agrícolas: pesquisa piloto na cidade de São Paulo. **Informações Econômicas**, São Paulo, v. 30, n. 4, p. 36-52, abr. 2000.

SÃO PAULO (Cidade). Prefeitura do Município de São Paulo. **Legislação**: feiras-livres; sacolões e mercados municipais. 2006. <<http://portal.prefeitura.sp.gov.br/secretarias/abastecimento/legislacao/0001>>. Acesso em: 2006.

_____. Lei nº. 11.220, de 20 de maio de 1992. Institui a divisão geográfica da área do Município em Distritos, revoga a Lei 10.932, de 15 de janeiro de 1991, e da outras providências. **Câmara Municipal de São Paulo**. Disponível em: <<http://camaramunicipalsp.qaplaweb.com.br/cgi-bin/wxis.bin/iah/>>. Acesso em: 2006.

SENRA, N. C. Informação estatística: política, regulação, coordenação. **Ciência da Informação**, Brasília, v. 28, n. 2, maio/ago. 1999.

SESSO FILHO, U. A. **O setor supermercadista no Brasil nos anos 90**. Piracicaba (SP): ESALQ/USP, 2003.

SPOSATI, A. **Cidade em pedaços**. São Paulo: Brasiliense, 2001. 173 p.

TAKIYA, H. **Atlas ambiental do município de São Paulo**. São Paulo: Prefeitura do Município de São Paulo, 2002. (Relatório Final, v. 1).

**PROPOSTA DE REESTRUTURAÇÃO DO SISTEMA AMOSTRAL DE
COLETA DE DADOS NO MERCADO VAREJISTA DE ALIMENTOS NA
CIDADE DE SÃO PAULO: uma aplicação de amostragem
estratificada probabilística em dois estágios**

RESUMO: O Instituto de Economia Agrícola (IEA) coleta e calcula mensalmente preços de alimentos referentes ao mercado varejista para a cidade de São Paulo desde o início da década de 1970, tendo se consolidado como importante fonte de consulta de dados. Essas informações são amplamente utilizadas em diversos trabalhos técnico-científicos referentes à comercialização, análise setorial e estrutural, entre outros. Este trabalho propõe nova metodologia amostral considerando o novo cenário da cidade. Foram utilizados os dados da última Pesquisa de Orçamentos Familiares (POF), divulgada pela Fundação Instituto de Pesquisas Econômicas (FIPE) do ano de 1998/99, referentes à alimentação dentro dos domicílios na cidade de São Paulo; indicadores de desenvolvimento humano, dados populacionais censitários e cadastros de feiras livres, sacolões, açougues, padarias e supermercados da cidade de São Paulo. Foi realizada a estratificação distrital da cidade de São Paulo com base em parâmetros socioeconômicos em um primeiro nível de estratificação. O segundo nível de estratificação foi realizado nos principais locais de compra de alimentos dos paulistanos. A amostra foi calculada a partir de estimativas de média e variância dos itens alimentares através de um levantamento piloto. A distribuição das amostras pelos estratos seguiu a alocação de Neyman. Também foram realizadas comparações do método estratificado com o de amostragem aleatória simples através do efeito do planejamento.

Palavras-chave: preços, levantamento por amostragem, comércio varejista.

**PROPOSAL FOR RESTRUCTURING OF SAO PAULO CITY'S RETAIL FOOD
MARKET SAMPLING SYSTEM - an application of a
two-stage probabilistic stratified sampling**

ABSTRACT: The Agricultural Economics Institute (IEA) has tracked retail food prices on a monthly basis in the city of Sao Paulo since the early 1970's, having become a comprehensive reference center for technical and scientific research on trade and sectorial and structural analysis, among other tasks. The research at hand proposes a new sampling methodology taking into account a new city sce-

nario. The data used came from the 1998-99 Brazilian Household Budgets Research (POF) made public by the Economic Research Institute (FIPE), including food items within households, human development indicators, census population and data records on street markets, wholesale markets, butcheries, bakeries and supermarkets in Sao Paulo city. The first stratification level of Sao Paulo's districts was based on socioeconomic parameters. The second stratification level was carried out at the city's main food purchasing places. The sample was calculated from mean and variance estimates of food items through a pilot survey. Sample distribution across strata was determined by Neyman's allocation. Finally, comparisons were conducted between the stratified and the simple random sampling methods through the effect-based planning approach.

Key-words: prices, sampling survey, retail, Sao Paulo city.

Recebido em 25/03/2009. Liberado para publicação em 29/07/2009.

**PROPOSTA DE REESTRUTURAÇÃO DO SISTEMA AMOSTRAL DE
COLETA DE DADOS NO MERCADO VAREJISTA DE ALIMENTOS NA
CIDADE DE SÃO PAULO: uma aplicação de amostragem
estratificada probabilística em dois estágios**

Anexo 1

PRODUTOS DE ORIGEM VEGETAL *IN-NATURA*

Bulbos, Raízes e Tubérculos

Alho
Batata
Batata-doce
Beterraba
Cebola
Cenoura
Mandioca
Mandioquinha
Rabanete

Frutos e Flores Comestíveis

Abóbora
Abobrinha
Berinjela
Brócolis
Chuchu
Pepino
Pimentão
Quiabo
Tomate
Vagem
Folhas e Caules
Acelga
Agrião
Alface
Almeirão
Couve
Couve-flor
Escarola
Espinafre
Repolho
Rúcula
Salsa e cebolinha

Frutas

Abacate
Abacaxi
Ameixa
Banana
Caqui
Figo
Goiaba
Kiwi
Laranja
Limão
Maçã

Mamão
Manga
Maracujá
Melão
Mexerica
Morango
Melancia
Pêra
Pêssego
Tangerina
Uva

Cereais e Grãos

Amendoim cru
Arroz
Canjica
Feijão
Lentilha
Milho

PRODUTOS DE ORIGEM VEGETAL INDUSTRIALIZADOS

Derivados de Cereais e Grãos

Achocolatado
Amido de milho
Biscoito
Café
Ervilha
Farinha de milho
Farinha de mandioca
Farinha de trigo
Fubá
Gordura vegetal
Macarrão
Margarina
Milho
Óleo vegetal
Pão

Derivados de Frutas

Azeite
Azeitonas
Coco
Extrato de tomate
Goiabada
Molho de tomate
Pêssego em calda
Polpa de tomate

Derivados de Folhas e Caules

Açúcar
Chá
Palmito

PRODUTOS DE ORIGEM ANIMAL *IN-NATURA***Carne Bovina**

Acém
Alcatra
Braço
Capa de filé
Contrafilé
Costela de vaca
Coxão duro
Coxão mole
Cupim
Fígado
Filé mignon
Fraldinha
Lagarto
Maminha
Carne moída
Músculo
Patinho
Peito
Picanha

Carne de Frango

Asa
Coxa
Sobrecoxa
Filé
Frango
Peito

Carne suína

Costela de porco
Lombo
Pernil

Toucinho

Pescados

Cação
Camarão
Corvina
Merluza
Pescada
Sardinha

Leites

Leite

Ovos

Ovo

PRODUTOS DE ORIGEM ANIMAL INDUSTRIALIZADOS

Derivados de Carne Bovina

Hambúrguer
Almôndegas
Carne seca
Salame

Derivados de Carne Suína

Apresentado
Bacon
Linguiça
Mortadela
Presunto
Salsicha

Derivados do Leite

Iogurte
Leite condensado
Leite em pó
Manteiga
Queijo

Diversos

Caldo de carne
Caldo de galinha
Sardinha