

Gastos Domiciliares nas Regiões Metropolitanas: Uma Análise não Paramétrica Kernel usando Dados de Pseudo-Painel

Alexandre Nunes Almeida¹  | Domingos Isaias Maia Amorim²  |
Carlos Roberto Azzoni³  | Francisco José Silva Tabosa⁴  |
André Luiz Squarize Chagas⁵ 

¹ Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo. E-mail: alex.almeida@usp.br

² Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo. E-mail: domingos_isaias@usp.br

³ Departamento de Economia, Universidade de São Paulo. E-mail: cazzoni@usp.br

⁴ Departamento de Economia Agrícola, Universidade Federal do Ceará. E-mail: franzetabosa@ufc.br

⁵ Departamento de Economia, Universidade de São Paulo. E-mail: achagas@usp.br

RESUMO

O objetivo deste trabalho é analisar a relação entre a parcela despendida entre itens alimentícios (e não alimentícios) e o gasto domiciliar total através de modelos de regressão kernel não paramétricos. Para atender este objetivo, foram utilizados dados de um pseudo-painel construído a partir das Pesquisas de Orçamentos Familiares do IBGE entre 1987 e 2009. Observou-se que os resultados dos modelos paramétricos não se mostraram significativamente diferentes das estimativas médias não paramétricas entre os diferentes grupos de consumo. No entanto, ao analisar as estimativas não paramétricas na forma de percentis, pode-se constatar um impacto substancialmente diferente do gasto total sobre a parcela gasta nos grupos alimentação e saúde e despesas pessoais, mas não nas outras despesas. Ademais, também não se corroborou a hipótese de que o uso de métodos que levem em consideração as características não observadas e invariantes no tempo dos chefes dos domicílios convergiriam para resultados diferentes em relação aos de uma estrutura de dados de corte transversal.

PALAVRAS-CHAVE

Dispêndio familiar, Curva de Engel, Modelos Kernel não paramétricos, Dados de painel

Household Spending in Metropolitan Regions: A Non-Parametric Kernel Analysis Using Pseudo-Panel Data

ABSTRACT

The objective of this paper is to analyze the relationship between the share spent on food (and non-food) items and total household expenditures from non-parametric kernel regression models. To achieve this objective, we used data from a pseudo-panel constructed from Household Budget Surveys from 1987 to 2009. Our main findings suggest that results from parametric models were not significantly different from the average non-parametric estimates for the different consumption groups analyzed. However, when the nonparametric estimates in the terms of percentiles are analyzed, we found an impact substantially different from the total expenditure on the share spent mainly on food and health and personal expenses, but not in other expenditures. Moreover, the hypothesis that the use of panel data models, that takes into account the unobserved and time-invariant characteristics of the heads of households would lead to different results from pooled models, was not corroborated.

KEYWORDS

Household expenditures, Engel curve, Non-parametric Kernel models, Panel data

CLASSIFICAÇÃO JEL

D12, C14

1. Introdução

A importância de retratar a demanda e o padrão de gastos das famílias, principalmente de baixa renda, está em auxiliar no diagnóstico das suas condições de vida por meio da identificação do perfil de consumo coletivo e individual (Galvão et al., 2016). Esse retrato fornece também subsídios para o direcionamento de políticas sociais como aquelas de renda mínima, segurança alimentar, educação financeira etc. (Almeida e Azzoni, 2016).

Para atingir esse escopo, é fundamental fazer uso de pesquisas de orçamentos domiciliares, visto que tais pesquisas demonstram o comportamento econômico real dos agentes demandantes. Embora pouco frequentes, em razão do seu alto custo de execução, esse tipo de pesquisa aproxima de forma bastante eficaz as informações sobre o padrão de vida dos indivíduos, o que permite investigar aspectos do bem-estar e do custo de vida das famílias, tanto a nível nacional como regional (Azzoni et al., 2003; Menezes et al., 2008; Almeida e Azzoni, 2016).

Em domicílios em que a renda é menor, a parcela gasta com alimentação corresponde à parte mais significativa dos gastos totais, como já postula a Lei de Engel (Deaton, 1997). Segundo dados do IBGE, no Brasil, em 2009 e 2018, a parcela de gastos com itens alimentícios correspondia a 19,8% e 17,5% de todas as despesas de consumo (IBGE, 2010, 2019).

Isto posto, dada a crucial importância da alimentação para garantir um bem-estar mínimo às famílias — assim como sua expressiva participação nos gastos das famílias mais pobres (que são mais numerosas) —, é fundamental concentrar atenção nesse item. No entanto, considerando que a estrutura de consumo envolve também outros itens, igualmente relevantes, principalmente de indivíduos vivendo nas grandes metrópoles, não é razoável restringir esta análise unicamente aos alimentos.

No Brasil, seguindo a literatura internacional, o número de estudos sobre demanda é bastante extenso conforme permitem atestar os dois volumes editados por (Silveira et al., 2006, 2007), compostos por trabalhos já publicados e inéditos. Incluem-se também nesse menu os trabalhos de: (Resende e Oliveira, 2008; Coelho et al., 2009; Pintos-Payeras, 2009; Coelho et al., 2010; Hoffmann, 2010; Silva et al., 2012; Pereda e Alves, 2012; Barbosa et al., 2013; Silva e Coelho, 2014; Oliveira e Hoffmann, 2015; Travassos e Coelho, 2014; Almeida et al., 2016; Ferreira et al., 2016).¹

Em geral, a maior parte dos trabalhos supracitados: 1) parte do uso extensivo das pesquisas de orçamentos familiares do IBGE, e 2) está fundamentada nos trabalhos seminais de (Working, 1943) e (Leser, 1963). Isto é, as especificações da forma funcional da função de demanda baseiam-se em métodos de regressão lineares que associam as despesas de um bem (ou grupo de bens) pela família em questão com o logaritmo do gasto total familiar. Essencialmente, esses modelos seguem a seguinte

¹Os trabalhos de (Rossi, 1982) e (Medeiros, 1978) foram pioneiros no Brasil ao estimar curvas de Engel na cidade do Rio de Janeiro.

estrutura:

$$w_{ji} = \alpha + \beta \ln(x_j) + e_{ji} \quad (1)$$

em que w_{ji} é a parcela despendida com o bem i pela família j , $\ln(x_j)$ é o log do gasto (ou renda) total da família j , e o componente não observado e_{ij} satisfaz a seguinte condição de ortogonalidade $E(e_{ji}/x_j) = 0$ (Blundell e Duncan, 1998).

Essa forma mais comum de estimação de um modelo de demanda por bens foi desenvolvida por (Deaton e Muellbauer, 1980) e ficou conhecida como "*Price Independent Generalized Logarithmic-PIGLOG*" (Santana e Menezes, 2009). Não obstante, a Equação (1), ao adicionar um vetor de preços, pode ser estimada também simultaneamente através de um sistema de demanda conhecido por *Almost Ideal Demand System* (AIDS), que é atualmente muito popular dentro da análise empírica. Um sistema de demanda com n Equações (1) tem a vantagem de capturar consistentemente a interdependência mútua de muitos bens, de acordo com as escolhas do consumidor e sua restrição orçamentária (Sadoulet e De Janvry, 1995). Ademais, o modelo AIDS, além de satisfazer todos os pressupostos da teoria do consumidor, também apresenta como principal vantagem seu formato linear. Esse formato facilita a inclusão de restrições inerentes aos modelos de demanda, tais como: aditividade, homogeneidade e simetria (Barbosa et al., 2013).

Ao longo dos anos, observou-se também que havia um comportamento de consumo não linear entre indivíduos mais pobres (itens necessários) e mais ricos (itens de luxo) e que a Equação (1) sugerida acima não seria mais apropriada (Banks et al., 1997). Assim sendo, Banks et al. (1997) desenvolveram um sistema de demanda derivado do modelo AIDS, denominado *Quadratic Almost Ideal Demand System* (QUAIDS). Esse adiciona em (1) um termo quadrático do logaritmo do gasto total permitindo mensurar quais bens consumidos pelos domicílios apresentariam um caminho de expansão da renda não linear dentro da curva de Engel (Pereda e Alves, 2012).

Similar ao modelo AIDS, o QUAIDS tornou-se bastante popular, porque é simples de estimar e também por permitir testes e imposição das restrições teóricas inerentes à economia do consumidor (Coelho e Aguiar, 2010). Entretanto, mesmo diante desse avanço, parece ainda que nunca houve um consenso sobre o verdadeiro formato da curva de Engel que permita refletir de forma bem-comportada as preferências do consumidor e que esteja de acordo com a natureza dos dados (Blundell et al., 2003). E, em geral, a teoria econômica não fornece muita orientação sobre a escolha da forma funcional para fins de estimação de qualquer modelo econométrico (Schmalensee e Stoker, 1999; Blundell et al., 1998; Yatchew, 2003; Henderson e Parmeter, 2015). Nota-se também que poucos trabalhos vão além ao procurar métodos alternativos que possam ajustar uma forma funcional de demanda, do que simplesmente se basear na especificação piglog, como sugerido por Blundell et al. (1998), Bhalotra e Attfield (1998), Blundell e Duncan (1998) e Gong et al. (2005).

De acordo com Blundell et al. (1998), existe uma obsessão por linearidade por parte dos economistas empíricos. O uso de modelos não paramétricos ou semiparamétricos constitui um avanço importante, já que permite verificar se a especificação da função é adequada ou se existe uma melhor forma de representar as preferências da teoria do consumidor empiricamente (Blundell et al., 1998, 2003; Santana e Menezes, 2009).

Até a data presente, no Brasil, Santana e Menezes (2009) é o único trabalho que estima padrões de gastos com educação entre raças através de métodos semiparamétricos e levando em consideração a flexibilização de formas funcionais do modelo estimado. Sendo assim, o objetivo deste trabalho é analisar como as famílias brasileiras alocam seu dispêndio usando as Pesquisas de Orçamentos Familiares (POFs) do IBGE de 1987/88, 1995/1996, 2002/2003, 2008/2009.² Especificamente, esta pesquisa estimará a relação entre renda e dispêndio alimentar e também não alimentar, a partir do uso de técnicas de estimação econométricas não paramétricas desenvolvidas exclusivamente para dados empilhados e de painel. Serão considerados na análise seis grupos de bens: 1) alimentação; 2) habitação; 3) transporte; 4) educação; 5) saúde e despesas pessoais; e 6) vestuário.

A relação de interesse entre a parcela despendida (w_{ij}) com o bem i pela família j e o log do gasto (ou renda) total da família j ($\ln x_j$) estimada não parametricamente é dada por:

$$w_{ji} = \varphi \ln(x_j) + e_{ji} \quad (2)$$

em que $\varphi(\cdot)$ é uma função de regressão da média condicional de w_{ij} dado x_j estimada não parametricamente.

O uso desta condição como estratégia empírica se justifica porque tem a grande vantagem de *"let the data speak for itself concerning the actual form of the regression curve"* Eubank (1999). Dessa maneira, o uso da econometria clássica, por impor linearidade nos parâmetros a serem estimados (como é o caso do modelo AIDS e QUAIDS), constitui uma limitação que pode ser bastante restritiva, dependendo da natureza dos dados, mas que pode ser equacionada através do uso de métodos não paramétricos através da expressão (2) (Li e Racine, 2007; Henderson e Parmeter, 2015).

Portanto, espera-se, através das estimativas não paramétricas obtidas pelo presente trabalho, estabelecer comparação entre estas e os resultados dos métodos econométricos tradicionais, que são largamente usados para subsidiar a maior parte de recomendações de políticas econômicas e sociais. A hipótese central que norteia esta pesquisa reside principalmente em diagnosticar o formato da curva de Engel não parametricamente para diversos grupos de consumo, permitindo ao gestor fornecer eventualmente ajustes mais específicos nas políticas públicas. A análise se dará a

²Ressalta-se que, até a data de conclusão deste estudo, os microdados da POF 2017-2018 ainda não estavam disponíveis e a análise com dados mais recentes é deixada para estudos futuros.

partir de uma base de dados do tipo painel, em que indivíduos ou domicílios idênticos são acompanhados ao longo do tempo.

O restante deste artigo é organizado da seguinte forma: a próxima seção revisa brevemente a literatura relacionada ao escopo do trabalho e, na sequência, a seção 3 descreve os dados utilizados e apresenta as abordagens econométricas e não paramétricas utilizadas. A seção 4 apresenta os resultados, seguida das conclusões finais na seção 5.

2. Literatura

Como citado na introdução, a literatura sobre estudos relacionados à demanda por bens é bastante rica tanto no Brasil com no exterior. No entanto, análises que se concentram mais especificamente em métodos de econometria não paramétrica objetivando determinar verdadeiro o formato da curva de Engel ainda são bastante exíguas.

Bhalotra e Attfield (1998) estimaram semiparametricamente a curva de Engel para o meio rural no Paquistão, encontrando que o formato da curva assume uma especificação logarítmica quadrática, portanto, não linear. Neste caso, o uso dos tradicionais modelos *Almost Ideal Demand System*, *Log-Translog* e *Linear Expenditure System* seria inapropriado para uma análise da demanda por alimentos.

Conclusão similar foi encontrada por Blundell et al. (1998) ao analisarem os gastos com bebidas e alimentos no Reino Unido, por meio dos dados de pesquisas domiciliares entre 1980 e 1982. Usando diversas metodologias de regressão não paramétrica e controlando para componentes demográficos, os autores mostraram que o formato da curva de Engel para esses dois grupos difere substancialmente entre um e outro. Enquanto a parcela despendida com alimentos cai quando o gasto total domiciliar aumenta, para o caso de bebidas alcoólicas a curva de Engel se mostra como um U invertido (com relação à parcela gasta com esse item e ao gasto familiar total e ao logaritmo do gasto total).

Blundell et al. (1998), usando também dados de pesquisa domiciliares do Reino Unido entre 1980 e 1982, estimaram não parametricamente curvas de Engel para os grupos: alimentação, combustível doméstico, vestuário, roupas, transportes e outros bens para domicílios que possuíam uma ou duas crianças. Entre os resultados mais importantes encontrados, tem-se que a especificação linear piglog foi rejeitada em favor de uma função estimada não parametricamente para vários grupos de consumo, entre eles: bebidas alcoólicas, vestuário, transporte e outros bens. Gong et al. (2005) usaram um modelo parcialmente linear para identificar diferenças na alocação de gastos com alimentos, educação e bebidas alcoólicas nas famílias rurais da China. Averiguou-se no estudo se a estrutura de consumo difere de acordo com o gênero dos filhos. Com suas estimações, rejeitou-se a hipótese de uma relação linear (inversa) entre o item alimentação e o gasto domiciliar total, indicando que também neste caso

o uso de modelos lineares seria inapropriado.

Santana e Menezes (2009) estimaram a curva de Engel para o Brasil, focando nos gastos com educação. Ajustaram um modelo semiparamétrico aos dados da POF 2002-2003, utilizando como variáveis de controle as características do chefe da família, raça, composição do domicílio e localização. Entretanto, as autoras ignoraram qualquer análise referente à parte não paramétrica (gastos totais) sobre a variável dependente de interesse, focando apenas nas variáveis parametrizadas.

Os poucos estudos mencionados fornecem subsídios importantes para concluir que a flexibilização da função para estimação da Curva de Engel deve estar sendo sempre revisitada. Outra limitação presente em todos os citados estudos consiste no fato de suas análises basearem-se em dados de corte transversal. Esse obstáculo implica que ao analista é permitido fazer inferências estatísticas apenas sobre os determinantes observados que explicam o dispêndio familiar, enquanto que outras características igualmente importantes referentes às tomadas de decisões por parte das famílias, mas que não são observadas (por exemplo inteligência emocional dos chefes), são ignoradas (Cameron e Trivedi, 2005; Baltagi, 2008).

Dentro deste contexto, o uso de dados em painel sobrepõe o uso de dados de corte seccional com sucesso, uma vez que estes têm como principal vantagem analítica reduzir o viés de variáveis omitidas, ao permitirem que outras características invariantes no tempo e que não são observadas (referentes à unidade de análise) possam ser controladas (Cameron e Trivedi, 2005; Baltagi, 2008).

3. Material e Metodologia

As bases de dados utilizada neste trabalho são as Pesquisas de Orçamentos Familiares (POF), realizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) entre março de 1987 e fevereiro de 1988, entre outubro de 1995 e setembro de 1996, entre julho de 2002 e junho de 2003 e novamente entre maio de 2008 e maio de 2009. As POF diagnosticam a qualidade de vida das famílias brasileiras com base, principalmente, no seu orçamento doméstico (IBGE, 2010). De particular interesse para este estudo, há informações detalhadas sobre gastos mensais com alimentação, habitação, transporte, educação, saúde, higiene etc., além de informações demográficas e rendimentos salariais e não salariais.

As POF de 87/88, 95/96, 02/03 e 08/09 entrevistaram 13.611, 16.013, 48.470 e 55.970 unidades domiciliares, respectivamente. Para a sua realização, partiu-se dos resultados dos Censos Demográficos e com fatores de expansão da amostra construídos também a partir do Censos (IBGE, 2010, 2019). Para manter consistência na análise geográfica entre as POF, apenas observações dos residentes das 9 regiões metropolitanas (Belém, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre) mais os do Distrito Federal e de Goiânia foram utilizadas na análise. Ademais, todos os itens referentes aos gastos e rendimentos foram deflacio-

nados pelo IPCA/IBGE com a base para junho de 2019.

Os dados das POF são muito importantes para avaliar o custo de vida das pessoas, no entanto sua execução é onerosa pelo grau de detalhamento das informações referentes às despesas e às rendas das famílias investigadas. Ademais, os dados possuem como limitação não serem estruturados na forma de painel, ou seja, indivíduos e domicílios idênticos não são seguidos ao longo do tempo, o que diminui a precisão das estimativas econométricas (Cameron e Trivedi, 2005). Uma boa alternativa seria fazer uso de dados de cortes transversais repetidos como uma aproximação de dados longitudinais, técnica conhecida por “pseudo-painel” (Deaton, 1997). Browning et al. (1985) e Deaton (1997) foram os primeiros a fazer uso de pseudopainéis com vistas a analisar a oferta de trabalho no Reino Unido e em Taiwan, respectivamente (Cameron e Trivedi, 2005)).

A técnica para construir um pseudopainel é bastante simples e consiste em converter uma base de dados de corte transversal contemplando domicílios ou indivíduos observados em um determinado período t para o nível de cortes. Uma coorte pode ser interpretada como um grupo de indivíduos que compartilham – em determinado tempo – uma experiência em comum, definida a partir de características semelhantes médias observadas ao longo do tempo, como: idade, sexo, renda, localização geográfica, escolaridade dos pais etc. (Deaton, 1997)

Genericamente, para a montagem de uma coorte, considera-se que, se “ t ” é o momento da ocorrência, por exemplo, o ano de realização da pesquisa, e “ a ” é a idade do indivíduo nesse período (tempo t), então pode-se dizer que esta se aplica à coorte nascida no ano “ $t-a$ ”, assim como para as observações para as idades $a-1$ no tempo $t-1$ e assim sucessivamente (Ryder Norman, 1965). No total, no presente estudo, foram construídas 3.961 coortes formando um pseudopainel desbalanceado entre 1988 e 2009.

A Tabela 1 ilustra o gasto médio deflacionado em R\$ (Base=Jun/2019), de acordo com cada classe de renda, para os anos disponíveis da POF. O gasto médio considerando todas as classes de renda para as nove regiões metropolitanas, Distrito Federal e Goiânia orbitava entre R\$ 2.584 reais em 1988 e R\$ 2.181 reais em 2009. Nessas regiões, pôde-se observar também que o índice de Gini-consumo aumentou entre 1988 (0,51) e 1996 (0,54), reduziu-se em 2003 (0,49) e teve um leve aumento em 2009 (0,51). Quando se compara o gasto total mensal familiar com a renda total mensal dessa mesma família, a desigualdade é muito superior, e situa-se acima de 0,60, para qualquer ano da pesquisa.

Tabela 1. Média do gasto total mensal em R\$ (Base: Jun/19) por classe de renda entre 1988 e 2009 nas regiões metropolitanas

Decis	1988	1996	2003	2009
	Gasto Mensal	Gasto Mensal	Gasto Mensal	Gasto Mensal
1	271,65	262,16	284,43	293,61
2	566,29	559,59	567,01	560,62
3	791,13	790,62	791,75	791,96
4	1.025,67	1.026,60	1.026,60	1.024,43
5	1.298,87	1.293,61	1.296,08	1.296,49
6	1.639,69	1.629,79	1.629,69	1.633,09
7	2.077,01	2.084,54	2.096,60	2.073,71
8	2.747,63	2.749,69	2.753,81	2.749,07
9	3.942,06	3.972,16	3.934,64	3.897,63
10	9.454,54	10.600,21	9.296,80	10.392,68
Total	2.584,74	2.504,02	2.458,25	2.181,65
Gini - Consumo	0,5112	0,5456	0,4983	0,5191
Gini - Renda	0,637	0,654	0,6404	0,6225

Fonte: Dados da pesquisa.

A aplicação de modelos não paramétricos ainda é limitada, mas vem aumentando em muitos campos da pesquisa econômica (Henderson e Parmeter, 2015). Henderson e Parmeter (2015) também sugerem que, para fins de análise de política, um modelo completamente paramétrico, como colocado pela Equação 1 durante a introdução, é sempre preferível a um modelo não paramétrico da Equação 2, pois é de mais fácil interpretação.

No caso do presente estudo, o principal interesse é investigar, através de métodos não paramétricos, o quão robusto é o impacto de gasto familiar sobre a parcela despendida (em termos percentuais) em diversos bens e serviços com relação ao orçamento familiar total. Ademais, assumindo que existe uma heterogeneidade entre as diferentes coortes dos chefes afetando suas decisões de consumo, o uso de dados em painel permite controlar características que normalmente são consideradas como não observadas e invariantes no tempo (por exemplo: habilidades cognitivas).

Essa heterogeneidade que existe entre as unidades de análise, quando incorporadas ao modelo, além de resultar em estimativas não viesadas, oferecem ao pesquisador mais informação, mais variabilidade, menos colinearidade, mais graus de liberdade e, conseqüentemente, resultados mais eficientes (Baltagi, 2008). Entretanto, uma das grandes limitações em se trabalhar com dados de painel é o atrito das amostras, o que

indica que nem sempre é fácil acompanhar os mesmos indivíduos ao longo do tempo. O uso do pseudopainel é capaz de preencher essa lacuna, pois permite acompanhar pseudosindivíduos durante a montagem das diferentes coortes (Deaton, 1997).

A razão principal de usar o método não paramétrico é também evitar restrições ou erros de especificação, que podem resultar em errôneas formas funcionais específicas produzindo, por sua vez, estimativas tendenciosas (Yatchew, 2003; Li e Racine, 2007; Henderson e Parmeter, 2015). Nesse caso, considera-se primeiramente a seguinte especificação totalmente paramétrica da coorte c para o período $t = 1988, 1996, 2003$ e 2009 :

$$\bar{Y}_{c,t,i} = \bar{L}_{c,t}\beta + \bar{X}'_{c,t}\theta + \bar{C}_c + \bar{\varepsilon}_{c,t} \quad (3)$$

em que $\bar{Y}_{c,t}$ representa a média da parcela de gastos com um determinado bem i para a coorte c ; $\bar{L}_{c,t}$ é o gasto total mensal médio de cada coorte; $\bar{X}_{c,t}$ é um vetor com valores médios pertencentes a cada coorte com características domiciliares e individuais; e o vetor \bar{C}_c representa a média de componentes individuais não observados para cada coorte que são assumidos como constantes no tempo (portanto, sem o subscrito t), mas possivelmente correlacionados com $\bar{L}_{c,t}$ e $\bar{X}_{c,t}$. O termo erro é representado por $\bar{\varepsilon}_{c,t}$, e β e θ são parâmetros a serem estimados. O modelo proposto em (3) é conhecido como modelo de efeitos fixos (Cameron e Trivedi, 2005; Baltagi, 2008).

Usando a técnica conhecida como “*within transformation*”³ para eliminar o componente invariante no tempo \bar{C}_c (efeitos fixos) e produzir estimativas consistentes e não viesadas a Equação (3) torna-se:

$$\ddot{Y}_{c,t} = \ddot{L}_{c,t}\beta + \ddot{X}'_{c,t}\theta + \ddot{\varepsilon}_{c,t} \quad (4)$$

em que (4) é estimado usando o método de MQO tradicional.

Em seguida, considere a mesma especificação da Equação (4), mas agora sem a presença de parâmetros lineares (β e θ) no modelo:

$$\ddot{Y}_{c,t} = \varphi\left(\ddot{L}_{c,t}, \ddot{X}_{c,t}\right) + \ddot{\varepsilon}_{c,t} \quad (5)$$

onde $\varphi(\cdot)$ é uma função desconhecida do gasto total $\ddot{L}_{c,t}$, e o vetor $\ddot{X}_{c,t}$ que segue a mesma especificação da Equação (4).

Uma das principais limitações de estimar a Equação (5) usando abordagens não paramétricas é a “maldição da dimensionalidade” ou mais conhecida em inglês como “*curse of dimensionality*” (Li e Racine, 2007). Em outras palavras, se o número de

³A técnica “*within transformation*” consiste em calcular a diferença entre o modelo da Equação (3) e o modelo baseado nas médias (Cameron e Trivedi, 2005).

variáveis independentes é alto, a taxa de convergência do estimador não paramétrico para seu valor verdadeiro ($\hat{\varphi}_p \rightarrow \varphi$) diminui, produzindo dessa maneira estimativas inconsistentes (Li e Racine, 2007; Henderson e Parmeter, 2015). Recomenda-se para trabalhos aplicados, então, usar no máximo três variáveis na parte não paramétrica (Yatchew, 2003). Outra razão se deve também ao fato de que estimações não paramétricas baseadas principalmente em funções kernel, como no presente caso, são bastante demandantes do ponto de vista computacional, dependendo do número de observações e do número de interações durante a otimização do componente não paramétrico (Härdle, 1990). O objetivo é estimar $\varphi(\cdot)$ e a inclinação (gradientes) de $\varphi(\cdot)$ que pode ser derivada como $\beta(\cdot) = \partial\varphi(\cdot)/\partial(\cdot)$ (Henderson e Parmeter, 2015).

Em regressões não paramétricas que fazem uso da função kernel, como é o caso do presente estudo, a escolha da largura da janela (*bandwidth*) é sempre mais importante do que a escolha da função kernel (Li e Racine, 2007)⁴. Assim, técnicas conhecidas como o método de validação cruzada (*Cross-Validation*) são adotados para a escolha do ótimo bandwidth (Li e Racine, 2007; Henderson e Parmeter, 2015).

Tendo como objetivo central verificar a performance de métodos não paramétricos, desenvolvidos especificamente para dados de painel, é interessante comparar essas estimativas seguindo a mesma forma funcional dos tradicionais métodos paramétricos usados pela maioria de estudos empíricos. Ademais, é comum também reportar estimativas dos dados de painel de forma empilhada (*pooled*) como se eles fossem um único recorte transversal no tempo (*cross-section*). Desta forma, no entanto, exclui-se a possibilidade de controlar efeitos fixos que meçam a heterogeneidade não observada e invariante no tempo entre as coortes (Cameron e Trivedi, 2005; Baltagi, 2008).

No Quadro 1 são apresentadas diversas especificações paramétricas e não paramétricas propostas no trabalho. É de interesse principal na análise, as variáveis gasto total mensal e número de pessoas no domicílio, enquanto binárias que identificam a região de residência da coorte, o ano e cada coorte (para os modelos de efeitos fixos) são adotadas apenas como controles nos modelos estimados.

As especificações (6) e (7) serão estimadas usando o tradicional método de Mínimos Quadrados Ordinários (MQO) com variável instrumental, objetivando contornar vieses das estimativas provenientes de variáveis omitidas entre o termo de erro, gasto total e da variável dependente (Cameron e Trivedi, 2005). Isso foi colocado por Blundell e Duncan (1998); Bhalotra e Attfield (1998) e (Gong et al., 2005), que sugerem instrumentalizar o gasto total domiciliar usando a renda total do domicílio. Antecipando parte dos resultados, de fato, o teste Wu-Durbin-Hausman, de fácil implementação no STATA, e aplicado aos modelos (6) e (7), com os dados empilhados, rejeitou ao nível de significância de 1% a hipótese de que o gasto total seria exógeno para todos os grupos de consumo com exceção do grupo vestuário. Portanto, tanto nas especificações paramétricas como nas não paramétricas, o gasto total mensal, i.e., a principal

⁴O uso de um bandwidth ótimo na função kernel ajuda a minimizar o erro do quadrado médio, componente que regula o trade-off entre o viés e a variância (Li e Racine, 2007).

variável de interesse, foi incluído nos modelos utilizando-se a renda mensal como instrumento. Esse procedimento é realizado em dois estágios (Cameron e Trivedi, 2005). Primeiramente, o logaritmo neperiano do gasto total mensal é regredido através do método de MQO contra o logaritmo neperiano da renda total mensal acrescido do logaritmo neperiano do número de pessoas no domicílio. No segundo momento, os valores esperados do gasto total instrumentalizados do primeiro estágio entram na equação principal como variável explicativa de interesse.

Quadro 1. Modelos paramétricos e não paramétricos adotados no estudo

Modelos paramétricos:

Dados Empilhados

$$Y_c = L_c\beta + M_c'\gamma + D_c^R Regiao + D_c^A Ano + \varepsilon_c \quad (6)$$

Dados Painel

$$\ddot{Y}_{c,t} = \ddot{L}_{c,t}\beta + \ddot{X}'_{c,t}\gamma + \ddot{\varepsilon}_{c,t} \quad (7)$$

Modelos não paramétricos:

Dados Empilhados

$$Y_c = \varphi(L_c, M_c, Regiao_c, Ano_c) + \varepsilon_c \quad (8)$$

Dados Painel

$$Y_{c,t} = \varphi(L_{c,t}, M_{c,t}, Regiao_{c,t}, Ano_{c,t}) + \ddot{\varepsilon}_{c,t} \quad (9)$$

onde para cada coorte tem-se:

Y_c = Ln da parcela do bem em análise.

L_c = Ln da parcela do gasto total mensal instrumentalizado pela renda total.

M_c = Ln do número de pessoas no domicílio.

$Regiao$ = Binárias que identificam a região de residência da coorte.

Ano = Binárias que identificam ano da pesquisa.

Ainda com base no proposto pelo Quadro 1, também serão apresentados resultados para: a) uma versão totalmente não paramétrica (Equação 8) de um estimador projetado especificamente para dados empilhados denominados por “*Local Constant Linear Square-LCLS*” ; e d) um estimador não paramétrico de efeitos fixos (Equação 9) desenvolvido por (Li e Racine, 2004), no qual a identificação de cada coorte repetida ao longo dos anos mais o ano da pesquisa entram como variáveis discretas no modelo. Todas as propriedades teóricas para cada um desses quatro estimadores estão apresentadas em Henderson e Parmeter (2015) e estão omitidas aqui. Henderson e Parmeter (2015) também fornecem códigos de programação do *software* R (utilizado neste estudo) para replicar todos os exemplos práticos do livro, sendo esses códigos facilmente adaptáveis pelos usuários para seus respectivos problemas de pesquisa.⁵

Antes de adentrar na análise dos resultados, os estimadores kernel não paramétricos atualmente mais conhecidos partiram do estimador desenvolvido por Nadaraya (1964) e Watson (1964). O estimador Nadaraya-Watson ou “*Local Constant Linear Square*”-LCLS baseado na expressão $w_{ji} = \varphi(\ln(x_j)) + e_{ji}$ pode ser expresso por:

⁵Os códigos podem ser encontrados em: <http://www.the-smooth-operators.com/code>

$$\hat{\varphi}(x_j) = \frac{\frac{1}{nh} \sum_{i=1}^n k\left(\frac{x_n - x_j}{h}\right) w_{ji}}{\frac{1}{nh} \sum_{i=1}^n k\left(\frac{x_n - x_j}{h}\right)} \quad (10)$$

onde $k(\cdot)$ é uma função kernel que atende as propriedades estatísticas de simetria, integração para 1 e, continuamente diferenciáveis (Li e Racine, 2007; Henderson e Parmeter, 2015). A especificação das funções kernel mais utilizadas – uniforme, gaussiana e epanechnikov – estão listadas abaixo como:

$$\text{Kernel uniforme} \quad k(u) = \frac{1}{2} l(|u| \leq 1) \quad (11)$$

$$\text{Kernel gaussiana} \quad k(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \quad (12)$$

$$\text{Kernel epanechnikov} \quad k(u) = \frac{3}{4} (1 - u^2) l(|u| \leq 1) \quad (13)$$

Por fim, para qualquer estimativa não paramétrica utilizando-se a expressão (10), a escolha do parâmetro de largura da janela h é mais importante do que a escolha da função kernel $k(\cdot)$, já que é h que regula o *trade-off* entre variância e viés das estimativas (Li e Racine, 2007; Henderson e Parmeter, 2015). Existem diversos métodos que levam em consideração a natureza dos dados para calcular essa largura da janela ótima h , entre eles estão: a regra de bolso Silverman (1986), mínimos quadrados de validação cruzada e a máxima verossimilhança de validação cruzada (Henderson e Parmeter, 2015).

4. Resultados

Durante qualquer estimação econométrica que faça uso de pesquisas domiciliares, dois obstáculos devem ser cuidadosamente investigados. O primeiro diz respeito ao número de zeros (solução de canto) nos domicílios que declararam não ter consumido o bem durante a semana de realização da pesquisa, embora isso não queira dizer que o bem não tenha sido consumido, pois pode o ter sido a partir de estoques feitos anteriormente. Deve-se considerar, também, que o bem pode simplesmente não fazer parte da cesta de consumo do domicílio já que diferentes famílias podem ter os mais distintos hábitos de consumo, bem como renda, etc. Os procedimentos Tobit ou Heckman de dois estágios são mais comumente usados nesses casos e de fácil implementação quando os dados são censurados ou truncados. Outra solução seria agrupar as famílias por classes de renda diminuindo o número de observações nulas, uma prática bastante comum em estudos de demanda (Menezes et al., 2008). Contudo, ambos os procedimentos descritos acima não foram necessários, porque durante a construção das coortes, seguido da montagem do pseudopainel, foram utilizados valores médios

para: número de componentes das famílias, renda e gastos por grupos de dispêndio mensal de cada coorte. Como resultado, poucas observações apresentaram valores nulos e foram descartadas.

A Tabela 2 apresenta a média e o desvio padrão de cada variável de interesse. Em resumo, pode-se observar que, nas 9 regiões metropolitanas (Belém, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre), Distrito Federal e Goiânia, gastos com alimentação que incluem despesas realizadas dentro e fora do domicílio representaram em média cerca de 30% do gasto familiar total, seguidas de habitação (21%), outros bens (15%), transporte (12%), vestuário (9%), saúde (8%) e educação (4%). O grupo de outros bens e serviços (por exemplo: higiene, lazer, despesas pessoais, outras despesas do lar, outros imóveis, outras despesas, encargos, comunicação e fumo) não foi considerado na análise e os bens e serviços que compõem cada grupo citado estão apresentados no Apêndice.

Tabela 2. Média e desvio padrão das variáveis.

Descrição das variáveis	Média	D.P.
Parcela do gasta com alimentação por mês (%)	0,31	0,15
Parcela do gasta com habitação por mês (%)	0,21	0,09
Parcela do gasta com transporte por mês (%)	0,12	0,06
Parcela do gasta com vestuário por mês (%)	0,09	0,05
Parcela do gasta com saúde e despesas pessoais por mês (%)	0,08	0,06
Parcela do gasta com educação por mês (%)	0,04	0,03
Parcela do gasta com outros bens por mês (%)	0,15	0,1
Número de componentes	3,69	1,09
Gasto total mensal em R\$*	2747,84	3030,52
Renda total mensal em R\$*	4258,04	5725,15
=1 se a família está localizada na região Norte	0,10	0,3
=1 se a família está localizada na região Nordeste	0,31	0,46
=1 se a família está localizada na região Sudeste	0,28	0,45
=1 se a família está localizada na região Sul	0,16	0,37
=1 se a família está localizada na região Centro-Oeste	0,15	0,35
= 1 se observação é referente ao ano de 1988	0,23	0,42
= 1 se observação é referente ao ano de 1996	0,29	0,45
= 1 se observação é referente ao ano de 2003	0,24	0,43
= 1 se observação é referente ao ano de 2009	0,23	0,42

Fonte: Dados da pesquisa.

Nota: * Base: Jun/2019.

O número médio de membros por domicílio dentro de cada coorte está em torno de 3,69, enquanto o gasto total médio orbita por volta de R\$ 2.747 reais. A renda total média se situa por volta de R\$ 4.258 reais. Todos os valores de Junho/2019. Ademais, cerca de 60% das famílias residiam, entre 1998 e 2009, nas três grandes capitais da região Nordeste e Sudeste.

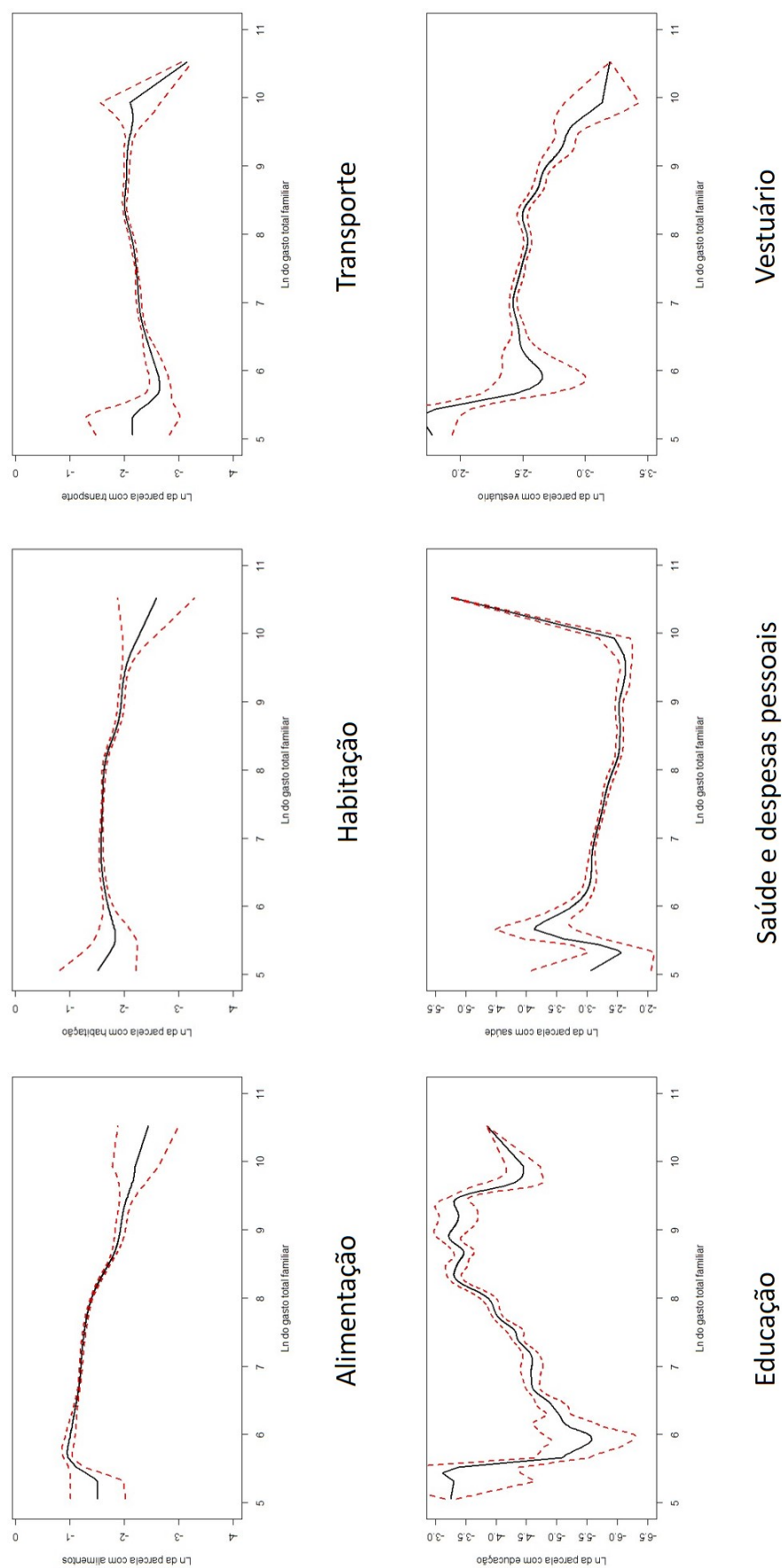
Dentro da literatura aplicada a modelos não paramétricos, é muito comum também mostrar a distribuição das estimativas de interesse, na forma de tabelas, ao nível de percentis ou através de gráficos com duas dimensões (Henderson e Parmeter, 2015). À título de ilustração, na Figura 1, estão plotados seis gráficos com estimativas das regressões não paramétricas (kernel gaussiana) a partir da largura da janela (*bandwidth*) estimado pelo método de mínimos quadrados de validação cruzada (LSCV) (Henderson e Parmeter, 2015) para grupos de bens e serviços selecionados. Essas foram feitas com base no método de “*Local Constant Linear Squares-LCLS*” da Equação (8) do logaritmo da parcela de cada item despendido contra o logaritmo do gasto total e o logaritmo do número de componentes por domicílio fixado ao seu valor mediano, portanto, constante. Para cada gráfico, são também apresentados os intervalos de confiança de 95% , com erros-padrão estimados por bootstrap seguindo o esboço proposto por Henderson e Parmeter (2015). O uso de bootstrap se tornou muito popular nos últimos anos por produzir estimativas de erros-padrão robustos para heterocedasticidade (Cameron e Trivedi, 2005).

Em resumo, os formatos das figuras das estimações não paramétricas parecem convergir para uma boa aproximação da especificação linear *piglog*, largamente utilizada pela pesquisa empírica para alguns grupos de consumo. Foi observado um comportamento linear por parte da curva estimada principalmente para os itens: alimentação, habitação e transporte.

É interessante observar também o formato de um N invertido para o grupo educação. Por exemplo, ao tirar o antilog do ln do gasto total familiar, nota-se que nesse grupo o gasto mensal com esse bem se reduz para famílias que têm gasto total de até R\$ 403 reais, contraposto por um aumento nas famílias com gasto mensal aproximado de até R\$ 13.500, mesmo ocorrendo depois uma leve queda.

Essa participação dos gastos com educação vem a refletir o crescimento do nível de escolaridade da população, entretanto, esse dispêndio representaria o crescimento das famílias que recebiam mais de 15 salários mínimos (Castro e Magalhães, 1998). Quanto aos indivíduos na parte inferior, os gastos com educação são influenciados pelas transferências de renda do Governo Federal, uma vez que a educação é uma das condicionalidades imposta pelo programa (Resende e Oliveira, 2008).

Figura 1. Curva de Engel



Nota: Curva de Engel não paramétrica (kernel) do logaritmo da parcela gasta (eixo Y) dos grupos de consumos selecionados contra o logaritmo do gasto total (eixo X) e o logaritmo do número de componentes por domicílio fixado ao seu valor mediano (constante). As linhas tracejadas representam 95% do intervalo de confiança com erros-padrão estimados através do método bootstrap usando 200 replicações.

Destaca-se o fato de que os gastos com saúde parecem declinar, mesmo que a taxas menores do que a do aumento no gasto total, até o nível de R\$ 20 mil/mês. Ademais, a partir desse nível de renda, o bem “saúde” parece assumir um comportamento de “bens de luxo” . Ao confrontar os resultados dos gastos com saúde e das cestas de alimentos, espera-se uma maior diversificação e melhora da dieta, que pode gerar benefícios à saúde (Resende e Oliveira, 2008), ou podem ser atribuídas a uma cobertura maior por parte do SUS, principalmente entre os segmentos mais pobres da população, e a intervenções na área de medicamentos (Diniz et al., 2007).

A curva de Engel que representa as despesas com vestuário, as quais incluem roupas e calçados, apresentou um formato similar ao do grupo educação, ou seja, o formato de N-invertido, mas com similaridade também ao grupo alimentação e habitação. Isso evidencia que a importância dada a esse item pelas famílias decresce no orçamento familiar até o nível de gasto total aproximado de R\$ 430 por mês, seguido por uma leve alta de até aproximadamente R\$ 1.100 por mês e voltando a reduzir-se nas classes de renda mais altas.

A estimação de modelos econométricos na sua forma logarítmica, adotada pelo presente trabalho, proporciona duas vantagens que se destacam: 1) melhor ajustamento dos dados; e 2) estimativas que podem ser consideradas elasticidades por definição (Gujarati e Porter, 2003). No caso da interpretação das elasticidades, por exemplo, se alterarmos a variável independente (por exemplo gasto total familiar) por 1% , esperamos observar mudanças percentuais na variável dependente (por exemplo parcela do dispêndio alimentar) equivalentes ao valor do coeficiente estimado nos modelos paramétricos ou ao valor das estimativas observadas nos modelos não paramétricos. Dessa forma, as Tabelas 3 a 4 reportam resultados paramétricos e não paramétricos para os seis grupos de consumo na forma de elasticidades, respectivamente. Como já discutido, para cada método, foram consideradas uma estimação com os dados empilhados e uma na forma de pseudopainel. Nos modelos não paramétricos, os erros-padrão foram estimados pelo método bootstrap utilizando-se 200 replicações. São mostrados também efeitos parciais de cada regressor estimado não parametricamente nos percentis P25, P50 e P75 e na média dos percentis. Os *bandwidths* foram estimados via *Least-Squares Cross Validation* (LSCV) utilizando a função kernel gaussiana.

Em linhas gerais, pode-se observar cada grupo de bens dos quais os resultados dos valores médios das estimativas não paramétricas não se mostraram substancialmente diferentes dos resultados obtidos a partir dos modelos totalmente paramétricos. A princípio, poderia sugerir-se que: 1) baseado no principal pressuposto da regressão não paramétrica de não impor nenhuma restrição sobre a função estimada relativa ao termo de erro e de linearidade dos parâmetros, nessa presente aplicação, a econometria tradicional forneceu uma boa aproximação referente ao impacto do gasto total e do número de membros sobre as parcelas de dispêndios dos grupos de consumo analisados; e 2) portanto, não haveria necessidade do uso de técnicas econométricas

mais avançadas de modelagem dos dados.

Contudo, foram observadas diferenças pontuais, mais igualmente relevantes entre os dois métodos, a serem descritas abaixo. Por exemplo, ao comparar os dois métodos para o grupo alimentação, considerando primeiramente o modelo paramétrico de dados empilhados (Tabela 3), observa-se uma elasticidade de -0,35 do gasto total sobre a parcela gasta com alimentação, enquanto, no modelo não paramétrico LCLS (Tabela 4), pode-se observar que as elasticidades do gasto total aumentam ao longo dos percentis, entre -0,47 no P25 para -0,25 no P75, tal como postula a lei de Engel.

Resultados semelhantes foram apresentados por Menezes et al. (2006) e Coelho et al. (2009), em que se constatou que, em determinadas regiões do país e para essas famílias de menor renda, o gasto com alimentação é ainda o de maior peso no orçamento familiar. Os dados da elasticidade-renda total encontrados nesses trabalhos indicam que os produtos são, em sua maioria, bens normais. Entre os bens normais, os produtos de elasticidade positiva mais baixa são açúcar, arroz, feijão, sal e massa de tomate (Menezes et al., 2006).

Como mostrado na Tabela 3, para o grupo alimentação, o modelo de painel produziu uma estimativa do gasto total sobre a parcela despendida com alimentos pouco inferior, de -0,37, em relação ao valor de -0,35 do modelo de dados empilhados. As estimativas não paramétricas do modelo Li e Racine de efeitos fixos na Tabela 4 também mostram uma elasticidade média de -0,37. Contudo, nota-se que as elasticidades dessa variável entre os diferentes percentis oscilam de -0,46 a -0,20, com resultados muito semelhantes aos efeitos parciais do estimador LCLS de dados empilhados discutidos acima.

De forma semelhante aos resultados paramétricos, as estimativas não paramétricas também corroboram a hipótese de que ao aumentar o número de membros do domicílio vê-se aumentada também a parcela gasta com alimentos, conforme o esperado. Nota-se também que o efeito parcial médio do número de membros no método Li e Racine (elasticidade de 0,29) não coincidiu em valor com o resultado retornado pelo modelo paramétrico MQO de efeitos fixos (elasticidade de 0,50).

Os gastos com habitação também se mostraram com comportamento de consumo análogos aos observados nos itens alimentícios. Entre os resultados paramétricos descritos na Tabela 3, as elasticidades estatisticamente significativas ficaram no intervalo que vai de -0,13 (modelo de dados empilhados) até -0,09 (modelo de efeitos fixos), enquanto os efeitos parciais dessa variável dos dois métodos não paramétricos, em nenhum deles, mostraram-se estatisticamente significativos aos níveis de 1% , 5% ou 10% . Além disso, enquanto no modelo paramétrico apenas o modelo de efeitos fixos o número de moradores se mostrou estatisticamente significativo (elasticidade de -0,14); nos modelos não paramétricos, apenas o estimador de dados empilhados LCLS mostrou efeitos parciais estatisticamente significativos na média (-0,31), no percentil 25 (-0,53) e no percentil mediano (-0,37).

Tabela 3. Regressões paramétricas com dados das coortes (empilhados e efeitos fixos) para os grupos de alimentação, habitação e transporte.

Variável Dependente: Ln da parcela do orçamento familiar

	Alimentação		Habitação		Transporte		Educação		Saúde e despesas pessoais		Vestuário	
	Empilh	Efeito Fixo	Empilh	Efeito Fixo	Empilh	Efeito Fixo	Empilh	Efeito Fixo	Empilh	Efeito Fixo	Empilh	Efeito Fixo
Ln (gasto total mensal)	0,35***	0,37***	0,13***	0,09***	0,16***	0,19***	0,66***	0,63***	0,28***	0,28***	0,04***	0,06***
	-0,01	0	-0,01	-0,01	-0,01	-0,01	-0,03	-0,02	-0,01	-0,01	-0,01	-0,01
Ln (número de membros)	0,45***	0,50***	-0,03	-	0,08**	0,01	0,28***	0,36***	-	-	0,10**	0,16***
	-0,02	-0,01	-0,03	0,14***	-0,04	-0,03	-0,08	-0,07	0,25***	0,27***	-0,04	-0,03
Região Norte	0,08***	-	-	-0,02	-0,04	-0,04	-0,04	-	0,10***	-0,05	0,11***	-
	-0,01	-	0,07***	-	-0,03	-	-0,06	-	-0,03	-0,03	-0,03	-
Região Sudeste	-	-	-0,02	-	0,18***	-	-0,10**	-	0,08***	0,08***	-	-
	0,09***	-	0,29***	-	-0,02	-	-0,05	-	-0,02	-0,02	0,18***	-
Região Sul	-0,01	-	-0,01	0,11***	0,05*	0,05*	-	-	0	0	-0,02	0,03
	0,06***	-	-0,02	-	-0,03	-	0,36***	-	-0,035	-0,035	-0,02	-
Região Centro-Oeste	-0,01	-	0,29***	-	0,22***	0,22***	-0,15**	-	0	0	-	-
	0,14***	-	-0,02	-	-0,02	-	-0,06	-	-0,03	-0,03	0,15***	-
1995	-0,01	-	-0,02	-	0,59***	0,59***	0,60***	-	-0,03	-0,03	-0,03	-0,03
	0,64***	-	-	0,13***	-0,02	-0,02	-0,06	-	0,98***	0,98***	0,54***	-
2003	-0,01	-	-0,01	-	-0,02	-	-0,05	-	-0,03	-0,03	-0,02	-
	-	-	0,03**	0,03**	0,45***	0,45***	0,58***	-	0,77***	0,77***	0,62***	-
2009	0,81***	-	-0,01	-	-0,03	-	-0,06	-	-0,03	-0,03	-0,03	-
	-0,01	-	0,08***	0,08***	0,65***	0,65***	0,55***	-	0,83***	0,83***	0,58***	-
Constante	1,40***	0,85***	-0,01	-	-0,02	-	-0,06	-	-0,03	-0,03	-0,03	-
	-0,06	-0,05	0,76***	0,80***	4,17***	3,761***	9,98***	9,53***	5,23***	4,52***	2,76***	2,28***
R2	0,62	0,25	-0,08	-0,06	-0,1	-0,096	-0,22	-0,2	-0,12	-0,11	-0,11	-0,09
Durbin-Wu-Hausman	9,09***	-	0,22	0,09	0,19	0,031	0,16	0,13	0,25	0,03	0,187	0,01
Hausman	-	-	27,41***	11,79***	28,57***	11,55***	50,32***	2,13	45,28***	4,84*	1,58	1,43*
n	3961	3961	3961	3961	3961	3961	3961	3961	3961	3961	3961	3961

Notas: 1) Erros padrões robustos para heterocedasticidade estão abaixo dos coeficientes entre parêntesis. 2) * p < 0.10, ** p < 0.05, *** p < 0.01.

Tabela 4. Elasticidades para estimadores não paramétricos selecionados. Variável Dependente: Ln da parcela do orçamento familiar gasta com alimentação, habitação, transporte, educação, saúde e despesas pessoais e transporte

Grupos de Despesas/Estimadores	Ln (gasto total mensal)				Ln (numero de membros do domicílio)				R2
	Média	P25	P50	P75	Média	P25	P50	P75	
Alimentação									
LCLS-Local Constant Linear Squares*	-0,38***	-0,47***	-0,33***	-0,25***	0,21***	0,12	0,2	0,30***	0,62
Li e Racine (2004)**	-0,05	-0,13	-0,03	-0,02	-0,06	-0,09	-0,36	-0,03	0,7
	-0,37***	-0,46***	-0,31*	-0,20***	0,29***	0,05	0,27	0,52***	
	-0,13	-0,06	-0,03	-0,1	-0,07	-0,35	-0,19	-0,25	
Habitação									
LCLS-Local Constant Linear Squares*	-0,09	-0,29	-0,07	0,13	-0,31**	-0,53***	-0,37***	-0,09	0,34
Li e Racine (2004)**	-0,21	-0,16	-0,17	-0,08	-0,15	-0,08	-0,08	-0,16	
	-0,03	-0,22	-0,03	0,13	-0,35*	-0,77	-0,44	-0,05	0,4
	-0,25	-0,3	-0,13	-0,11	-0,2	-0,56	-0,4	-0,31	
Transporte									
LCLS-Local Constant Linear Squares*	0,31**	0,04	0,26	0,57**	0,09	-0,16	0,07	0,28***	0,33
Li e Racine (2004)**	-0,13	-0,15	-0,41	-0,21	-0,08	-0,17	-0,21	-0,06	
	0,27*	0,008	0,22	0,48***	0,07	-0,35	0,04	0,4	0,43
	-0,16	-0,12	-0,19	-0,12	-0,36	-0,56	-0,35	-0,36	
Educação									
LCLS-Local Constant Linear Squares*	0,71***	0,38	0,80**	1,12***	1,03**	0,2	0,93	1,82**	0,28
Li e Racine (2004)**	-0,2	-0,23	-0,34	-0,09	-0,5	-0,73	-0,92	-0,71	
	0,57	0,18	0,69*	1,03***	0,82	-0,21	0,78	1,77**	0,38
	-0,57	-0,52	-0,34	-0,22	-1,01	-1,09	-1,19	-0,89	
Saúde e despesas pessoais									
LCLS-Local Constant Linear Squares*	0,33***	0,23**	0,36***	0,48***	-0,1	-0,30*	-0,06	0,06	0,33
Li e Racine (2004)**	-0,08	-0,11	-0,09	-0,06	-0,08	-0,16	-0,13	-0,16	
	0,19	0,14	0,32	0,52**	-0,09	-0,57	-0,1	0,32	0,47
	-0,26	-0,19	-0,21	-0,24	-0,7	-1,84	-0,51	-0,97	
Vestuário									
LCLS-Local Constant Linear Squares*	0,02	-0,29	-0,01	0,27***	0,07	-0,14	0,15	0,34***	0,33
Li e Racine (2004)**	-0,17	-0,27	-0,25	-0,12	-0,1	-0,25	-0,18	-0,12	
	0,03	-0,22	-0,02	0,22	0,12	-0,28	0,2	0,56**	0,43
	-0,24	-0,27	-0,15	-0,24	-0,31	-0,29	-0,74	-0,26	

Notas: 1) Tabela reporta as elasticidades para 25o(P25), 50o(P50), e 75o(P75) percentis junto com os erros padrões estimados pelo método bootstrap com 200 replicações. 2) * Regressões contém binárias para regiões e anos; **Regressões contém binárias para coortes e 4) Número de observações: 3.961.

O impacto do gasto total familiar sobre a parcela despendida com o grupo transporte ficou entre 0,16 (modelos MQO de dados empilhados) e 0,19 (modelo de efeitos fixos) nas estimativas paramétricas e entre 0,31 (modelo LCLS de dados empilhado) e 0,27 (modelo Li e Racine de efeitos fixos) nas estimativas não paramétricas. Nota-se ainda na Tabela 4 que, nos dois modelos, LCLS e Li e Racine, apenas o P50 produziu elasticidades estatisticamente significativas de magnitudes entre 0,57 e 0,48, respectivamente. Dentre os resultados encontrados na literatura, Diniz et al. (2007) apresentam que o gasto com transporte é mais alto quanto mais rica se torna a população. Sendo esse um reflexo do fato de que nos estratos superiores de renda é maior a participação dos residentes nas cidades e metrópoles (Pinheiro e N, 2007). Ainda nesse grupo de consumo, transporte, uma elasticidade de 0,08 foi observada para o número de moradores no modelo paramétrico, enquanto no modelo não paramétrico LCLS uma elasticidade de 0,28 foi encontrada no percentil 75. Curiosamente, nenhum dos modelos de efeitos fixos paramétricos ou não paramétricos, resultados estatisticamente significativos foram encontrados para esta variável.

Outra análise consiste em inspecionar o impacto do gasto total familiar sobre a parcela do orçamento familiar gasta com educação. Se as elasticidades do gasto entre os dois modelos paramétricos escolhidos (Tabela 3) variam pouco, entre 0,66 e 0,63; nos modelos não paramétricos (Tabela 4), os efeitos parciais oscilaram bastante. Por exemplo, elasticidades variaram de 0,80 (P50) até 1,12 (P75) no modelo LCLS e entre 0,69 (P50) e 1,03 (P75) no modelo Li e Racine. Para a variável número de membros de domicílio, elasticidades de 1,82 e 1,77 no percentil superior P75 foram estatisticamente significativas nos modelos LCLS e Li e Racine, respectivamente. Os gastos com educação estão sensivelmente ligados à renda *per capita* familiar (Bertasso et al., 2006). Nos estratos mais baixos, os gastos com educação tendem a ser influenciados pelo efeito das transferências governamentais (Bolsa-Escola), o que pode ser atribuído à condicionalidade imposta pelo programa, enquanto, por outro lado, os indivíduos que compõem o extrato mais rico se sobressaem com o peso destacado aos gastos com os cursos regulares de 1º, 2º e 3º graus e pré-escolar, decorrência da combinação do nível de renda e uma alta proporção de adolescentes nas famílias (Resende e Oliveira, 2008; Hoffmann, 2010).

No grupo de saúde, as elasticidades do gasto total dos dois modelos paramétricos produziram estimativas praticamente iguais, valor de 0,28; enquanto elasticidades gravitando entre 0,23 (P25) e 0,48 (P75) foram observadas no modelo não paramétrico LCLS que considera, lembrando, a natureza dos dados como empilhados. Esses resultados sugerem uma relativa inércia composta por medicamentos e pagamentos de mensalidades de planos de saúde nos estratos progressivamente mais ricos da população, enquanto as famílias mais pobres têm quase todo o seu gasto com saúde composto por despesas com medicamentos (Andrade e Lisboa, 2006; Silveira et al., 2006). No modelo Li e Racine de efeitos fixos não paramétricos, apenas o P75 mostrou uma elasticidade estatisticamente significativa de 0,52. As elasticidades dos dois modelos paramétricos, em relação aos membros do domicílio para a parcela despen-

didada com saúde, situaram em um intervalo de -0,25 a -0,27; enquanto nos métodos não paramétricos apenas o P25 do modelo LCLS produziu uma elasticidade significativa de -0,30.

Por fim, no grupo de vestuários também foram observados comportamentos de consumo similares ao do grupo alimentação e habitação em relação ao gasto total. Nos métodos paramétricos, elasticidades negativas do gasto ficaram em torno de -0,04 (dados empilhados) e -0,06 (efeitos fixos), enquanto nos métodos não paramétricos, apenas o modelo Li e Racine, uma elasticidade positiva e estatisticamente significativa de 0,27 no P75 foi observada. Em relação ao número de componentes, apenas o percentil 75 nos modelos LCLS e Li e Racine apresentou estimativas estatisticamente significativas com valores positivos de 0,34 e 0,56, respectivamente.

5. Conclusão

O objetivo deste trabalho foi realizar uma análise econométrica a respeito da curva de Engel. De interesse principal, foi comparar estimações paramétricas do modelo *piglog*, tradicionalmente adotadas pelos trabalhos empíricos no Brasil, com os modelos não paramétricos, esses últimos ainda pouco usados. Como mencionado por Henderson e Parmeter (2015), a recente disponibilidade de grandes conjuntos de dados somada aos avanços na capacidade computacional têm promovido a difusão de trabalhos mais aplicados no campo da econometria não paramétrica.

Como também mencionado pelos autores supracitados, do ponto de vista econômico relacionado ao dimensionamento de políticas públicas, os resultados de um modelo totalmente paramétrico, por serem mais fáceis de interpretação, são mais favoráveis do que os resultados de métodos não paramétricos. No entanto, eles também destacam que estimativas não paramétricas podem ser bastante úteis já que procuram complementar os resultados paramétricos no caso de amostras muito heterogêneas e fornecer respostas (no nosso caso) à questão de quanto as famílias em cada nível de vida se beneficiariam a partir de um aumento do gasto total domiciliar (reportado na forma de percentis) para cada grupo de despesa, mantendo o resto constante. No caso do presente trabalho, as elasticidades médias da variável gasto total geradas pelos diferentes estimadores não paramétricos coincidiram relativamente com os coeficientes dos modelos paramétricos para quase todos os grupos de consumo. No entanto, ao abrir a média e analisar os percentis nas estimativas não paramétricas, podem-se constatar diferenças estatisticamente mais robustas apenas para os grupos alimentação e saúde e despesas pessoais. Mais especificamente, nos dois casos, nossos resultados poderiam subsidiar, por exemplo, ajustes em políticas de segurança alimentar e na distribuição de medicamentos pelo SUS para famílias que estão entre estratos sociais muito distintos. Essas hipóteses são interessantes, merecendo atenção dos trabalhos futuros a partir de abordagens kernel não paramétricas semelhantes à nossa.

Ademais, não se corroborou também a hipótese de que o uso de métodos, paramétricos ou não paramétricos, desenvolvidos especificamente para controlar as características não observadas e invariantes no tempo dos chefes de domicílios através de uma estrutura de dados de painel, não convergiriam para resultados similares ao de uma estrutura de dados de corte transversal aumentando a precisão das estimativas. Em praticamente todos os grupos de bens e serviços, resultados dos primeiros foram muito próximos que os dos segundos em relação ao gasto total e ao número de componentes na família.

Uma limitação importante do presente trabalho é a de que se deve levar em consideração que todas as decisões de adquirir o bem são feitas com base na restrição orçamentaria das famílias, tal como postula a teoria do consumidor. Nesse caso, como exposto durante a introdução, o uso de modelos AIDS ou QUAIDS de estimação simultânea juntamente com as equações *piglog* para cada grupo de consumo seria mais apropriado. No entanto, o uso desse ferramental dentro da econometria kernel não paramétrica que flexibilize a função *piglog* dos modelos lineares até a presente data ainda não foi desenvolvido e representa uma avenida promissora na área de métodos quantitativos.

Referências

- Almeida, A. N. e Azzoni, C. R. (2016). Custo de vida comparativo das regiões metropolitanas brasileiras: 1996-2014. *Estudos Econômicos*, 46(1):253–276.
- Almeida, A. T. C. d., Mesquita, S. P. d., e Silva, M. V. B. d. (2016). Impactos do Programa Bolsa Família sobre a diversificação do consumo de alimentos no Brasil. *Pesquisa e Planejamento Econômico*, 46(1):7–39.
- Andrade, M. V. e Lisboa, M. d. B. (2006). Determinantes dos gastos pessoais privados com saúde no Brasil. In: *Gasto e Consumo das Famílias Brasileiras Contemporâneas*, volume 1. Instituto de Pesquisa Econômica Aplicada. Brasília.
- Azzoni, C. R., Carmo, H. E., e Menezes, T. A. d. (2003). Comparações da paridade do poder de compra entre cidades: aspectos metodológicos e aplicação ao caso brasileiro. *Pesquisa e Planejamento Econômico*, 33(1):91–126.
- Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.
- Banks, J., Blundell, R., e Lewbel, A. (1997). Curvas quadráticas de Engel e demanda do consumidor. *Review of Economics and Statistics*, 79(4):527–539.
- Barbosa, A. L. N. H., Menezes, T. A., e Andrade, B. C. (2013). Demanda por produtos alimentares nas áreas rurais e urbanas do Brasil. Texto para Discussão, n. 1829. Instituto de Pesquisa Econômica Aplicada.
- Bertasso, B. F., Silveira, F. G., e Magalães, L. C. G. (2006). Tipologia socioeconômica das famílias das grandes regiões urbanas brasileiras e seu perfil de gastos. In:

- Gasto e Consumo das Famílias Brasileiras Contemporâneas*, volume 1. Instituto de Pesquisa Econômica Aplicada. Brasília.
- Bhalotra, S. e Attfield, C. (1998). Intrahousehold resource allocation in rural Pakistan: a semiparametric analysis. *Journal of Applied Econometrics*, 13(5):463–480.
- Blundell, R. e Duncan, A. (1998). Kernel regression in empirical microeconomics. *Journal of Human Resources*, 34(1):62–87.
- Blundell, R., Duncan, A., e Pendakur, K. (1998). Semiparametric estimation and consumer demand. *Journal of Applied Econometrics*, 13(5):435–461.
- Blundell, R. W., Browning, M., e Crawford, I. A. (2003). Nonparametric Engel curves and revealed preference. *Econometrica*, 71(1):205–240.
- Browning, M., Deaton, A., e Irish, M. (1985). A profitable approach to labor supply and commodity demands over the life-cycle. *Econometrica*, 53(3):503–543.
- Cameron, A. C. e Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge University Press.
- Castro, P. F. d. e Magalhães, L. C. G. d. (1998). Recebimento e dispêndio das famílias brasileiras: evidências recentes da Pesquisa de Orçamentos Familiares (POF)-1995/1996. Texto para Discussão, n. 614. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Coelho, A. e Aguiar, D. (2010). O modelo Quadrático Almost Ideal Demand System (Quaids): uma aplicação para o Brasil. Texto para Discussão. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Coelho, A. B., Aguiar, D. R. D. d., e Eales, J. S. (2010). Food demand in Brazil: an application of Shonkwiler & Yen two-step estimation method. *Estudos Econômicos (São Paulo)*, 40(1):186–211.
- Coelho, A. B., Aguiar, D. R. D. d., e Fernandes, E. A. (2009). Padrão de consumo de alimentos no Brasil. *Revista de Economia e Sociologia Rural*, 47(2):335–362.
- Deaton, A. (1997). *The analysis of household surveys: a microeconomic approach to development policy*. The World Bank.
- Deaton, A. e Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge University Press.
- Diniz, B. P. C., Servo, L. M. S., Piola, S. F., e Eirado, M. (2007). Gasto das famílias com saúde no Brasil: evolução e debate sobre gasto catastrófico. Teto para Discussão. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Eubank, R. L. (1999). *Nonparametric regression and spline smoothing*. CRC Press.

- Ferreira, A. S., Coelho, A. B., et al. (2016). Substituição entre Alimentos Orgânicos e Convencionais no Brasil: O Papel dos Preços e do Dispêndio. *Anais do XLIV Encontro Nacional de Economia*.
- Galvão, M. C., Almeida, A. N., Spolador, H. F. S., e Azzoni, C. R. (2016). Custo de vida, amenidades e salários nas regiões metropolitanas brasileiras. *Revista Brasileira de Estudos Regionais e Urbanos*, 10(2):199–216.
- Gong, X., Van Soest, A., e Zhang, P. (2005). The effects of the gender of children on expenditure patterns in rural China: A semiparametric analysis. *Journal of Applied Econometrics*, 20(4):509–527.
- Gujarati, D. N. e Porter, D. C. (2003). *Econometria básica*. McGraw-Hill.
- Härdle, W. (1990). *Applied nonparametric regression*. Cambridge University Press.
- Henderson, D. J. e Parmeter, C. F. (2015). *Applied nonparametric econometrics*. Cambridge University Press.
- Hoffmann, R. (2010). Estimativas das elasticidades-renda de várias categorias de despesa e de consumo, especialmente alimentos, no Brasil, com base na POF de 2008-20091. *Revista de Economia Agrícola*, 2(2):49–62.
- IBGE (2010). *Pesquisa de Orçamentos Familiares 2008-2009: despesas, rendimentos e condições de vida*. Instituto Brasileiro de Geografia e Estatística.
- IBGE (2019). *Pesquisa de Orçamentos Familiares 2017–2018: primeiros resultados*. Instituto Brasileiro de Geografia e Estatística.
- Leser, C. E. V. (1963). Forms of Engel functions. *Econometrica: Journal of the Econometric Society*, 12(3):694–703.
- Li, Q. e Racine, J. (2004). Cross-validated local linear nonparametric regression. *Statistica Sinica*, 42(1):485–512.
- Li, Q. e Racine, J. S. (2007). *Nonparametric econometrics: theory and practice*. Princeton University Press.
- Medeiros, J. A. S. (1978). Curvas de Engel e transformação de Box-Cox: uma aplicação aos dispêndios em alimentação e educação na cidade de São Paulo. *Pesquisa e Planejamento Econômico*, 8(3):795–828.
- Menezes, T. A., Azzoni, C. R., e Silveira, F. G. (2008). Demand elasticities for food products in Brazil: a two-stage budgeting system. *Applied Economics*, 40(19):2557–2572.
- Menezes, T. A. d., Silveira, F. G., Magalhães, L. C. G. d., Tomich, F. A., e Vianna, S. T. W. (2006). Gastos alimentares nas grandes regiões urbanas do Brasil: aplicação do modelo AID aos microdados da POF 1995/1996 IBGE. Texto para Discussão, n. 896. Instituto de Pesquisa Econômica Aplicada (Ipea).

- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, 9(1):157–159.
- Oliveira, F. C. R. e Hoffmann, R. (2015). Consumo de alimentos orgânicos e de produtos light ou diet no Brasil: fatores condicionantes e elasticidades-renda. *Segurança Alimentar e Nutricional*, 22(1):541–557.
- Pereda, P. C. e Alves, D. C. d. O. (2012). Qualidade alimentar dos brasileiros: teoria e evidência usando demanda por nutrientes. *Pesquisa e Planejamento Econômico*, 42(2):239–260.
- Pinheiro, L. S. e N, F. (2007). Perfil das despesas e dos rendimentos das famílias brasileiras sob a perspectiva de gênero. In: *Gasto e Consumo das Famílias Brasileiras Contemporâneas*, volume 2. Instituto de Pesquisa Econômica Aplicada. Brasília.
- Pintos-Payeras, J. A. (2009). Estimação do sistema quase ideal de demanda para uma cesta ampliada de produtos empregando dados da POF de 2002-2003. *Economia Aplicada*, 13(2):231–255.
- Resende, A. C. C. e Oliveira, A. M. H. C. d. (2008). Avaliando resultados de um programa de transferência de renda: o impacto do Bolsa-Escola sobre os gastos das famílias brasileiras. *Estudos Econômicos*, 38(2):235–265.
- Rossi, J. W. (1982). Elasticidades de Engel para dispêndios familiares na cidade do Rio de Janeiro. *Pesquisa e Planejamento Econômico*, 12(2):579–606.
- Ryder Norman, B. (1965). The cohort as a concept in the study of social change. *American Sociological Review*, 30(6):843–861.
- Sadoulet, E. e De Janvry, A. (1995). *Quantitative development policy analysis*, volume 5. Johns Hopkins University Press Baltimore.
- Santana, P. J. e Menezes, T. A. d. (2009). Diferenças raciais no padrão de gastos com educação: uma abordagem semiparamétrica. *Nova Economia*, 19(3):383–405.
- Schmalensee, R. e Stoker, T. M. (1999). Household gasoline demand in the United States. *Econometrica*, 67(3):645–662.
- Silva, A. S. P., Scherer, C. E. M., e Porsse, A. (2012). A nova classe média: alterações de consumo e seus efeitos regionais. *Revista Brasileira de Estudos Regionais e Urbanos*, 6(1):1–15.
- Silva, M. M. d. C. e Coelho, A. B. (2014). Demanda por frutas e hortaliças no Brasil: uma análise da influência dos hábitos de vida, localização e composição domiciliar. Texto para Discussão. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Silveira, F. G. O., Servo, L. M. O., Menezes, T. O., e Piola, S. F. O. (2006). *Gasto e consumo das famílias brasileiras contemporâneas*. Instituto de Pesquisa Econômica Aplicada. Brasília.

- Silveira, F. G. O., Servo, L. M. O., Menezes, T. O., e Piola, S. F. O. (2007). *Gasto e consumo das famílias brasileiras contemporâneas*, volume 2. Instituto de Pesquisa Econômica Aplicada. Brasília.
- Travassos, G. F. e Coelho, A. B. (2014). A questão da separabilidade fraca na estimação de sistemas de demanda: Uma aplicação para a demanda de carnes no Brasil. *Pesquisa e Planejamento Econômico*, 44(3):545–578.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 26(4):359–372.
- Working, H. (1943). Statistical laws of family expenditure. *Journal of the American Statistical Association*, 38(221):43–56.
- Yatchew, A. (2003). *Semiparametric regression for the applied econometrician*. Cambridge University Press.

Agradecimentos

O primeiro autor agradece a Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (Auxílio n. 2016 /17566-5) pelo apoio financeiro desta pesquisa.

 Este artigo está licenciado com uma *CC BY 4.0 license*.

Apêndice

Composição dos grupos

Grupos			
Habituação	Transporte	Vestuário	Higiene
Energia Elétrica	Urbano (apenas ônibus)	Roupas de Homem	Frtfume
Água e esgoto	Gasolina - veículo próprio	Roupas de Mulher	Produtos para cabelo
Gás doméstico	Álcool - veículo próprio	Roupas de Criança	Sabonete
Condomínio	Manutenção e acessórios do veículo	Calçados e apetrechos	Instrumentos e produtos de uso pessoal
Telefone fixo Aluguel monetário	Viagens esporádicas	Joias e Bijuterias	
Saúde	Educação	Lazer	Despesas Pessoais
Remédios	Cursos pagos	Brinquedos e jogos	Cabeleireiro
Plano/Seguro saúde	Outros cursos e atividades	Periódicos, livros e revistas não didáticos	Pedicura
Consultas e tratamento dentário	Livros didáticos e revistas técnicas	Recreação e esportes	Consertos de artigos pessoais
Consulta médica	Artigos escolares	Jogos e apostas	
Serviços de Cirurgia			
Hospitalização			
Exames diversos			
Outras Despesas do Lar	Outro Imóveis	Outras Despesas	Encargos
Artigos de Limpeza	Imóveis de uso ocasional	Cerimônias e festas	Contribuições trabalhistas
Mobiliários e Artigos do Lar	Imóvel (aquisição)	Serviços profissionais	Previdência Privada
Manutenção do lar			Impostos
Imóvel (reforma)			
Têxteis e armarinhos			
Comunicação	Fumo		
Correios	Cigarros, charutos etc.		
Alimentação			
Arroz	Tomate	Pescados frescos	Café Moído
Feijão	Cebola	Frango	Refrigerantes
Alimentos preparados	Alface	Ovos	Cervejas e Chope
Macarrão	Banana	Leite de Vaca	Outras bebidas alcoólicas
Farinha de trigo	Laranja	Leite em pó	Enlatados e Conservas
Farinha de mandioca	Maça	Queijos	Massa de Tomate
Batata	Carne de boi primeira	Pão francês	Maionese
Cenoura	Carne de boi segunda	Biscoito	Sal Refinado
Mandioca	Carne suína	Óleo de Soja	Alimentação fora do domicílio
Açúcar refinado		Azeite de Oliva	
Açúcar cristal			