Semiparametric Estimation of Consumer Demand Systems in Real Expenditure

by

Krishna Pendakur and Stefan Sperlich[*]

Department of Economics, Simon Fraser University, Burnaby, BC, CANADA, V5A
1S6, e-mail: pendakur@sfu.ca; and Georg-August Universität Göttingen, Institut für
Statistik und Ökonometrie, Platz der Göttinger Sieben 5, D - 37073 Göttingen, e-mail:
stefan.sperlich@uni-goettingen.de

Abstract

Microdata concerning consumer demand typically show considerable variation in real
expenditures, but very little variation in prices. We propose a semiparametric strategy
for the consumer demand problem in which expenditure share equations are estimated

1

nonparametrically in the real expenditure direction and estimated parametrically (with fixed- or varying-coefficients) in price directions. In our model, Engel curves are unrestricted: demands may have any rank. Because the demand model is derived from a cost function, it may be restricted to satisfy integrability and used for consumer surplus calculations. Since real expenditure is unobserved, but rather estimated under the model, we face a semiparametric model with a nonparametrically generated regressor. We show efficient convergence rates for parametric and nonparametric components. We illustrate the feasibility of our proposed strategy using Canadian expenditure and price data: Engel curves display curvature which cannot be encompassed by standard parametric models. We also find that the rationality restriction of Slutsky symmetry is rejected in the fixed-coefficients model, but not in the varying-coefficients model.

Keywords: Consumer Demand, Engel Curves, Semiparametric Econometrics, Generated Regressors.

1 Introduction

Microdata concerning consumer demand typically exhibit considerable variation in real expenditure, but very little variation in prices. We propose a semiparametric strategy for the consumer-demand problem in which expenditure share equations are estimated nonparametrically in the real expenditure direction and parametrically in price directions. Our approach puts flexibility where the data can actually provide a lot of information — the Engel curve — and puts parametric structure on the price effects, where the data often have little variation.

Parametric approaches to the estimation of consumer demand systems are attractive because they allow the researcher to estimate a cost function, which makes consumer surplus calculations clear and easy. However, they typically impose strict limits on the complexity of Engel curves. Nonparametric approaches to the estimation of these systems have the advantage of letting the data determine the shape of Engel curves, but do not typically allow the researcher to estimate a cost function. Our semiparametric model combines the advantages of both by allowing Engel curves to be arbitrarily complex and yielding a cost function suitable for consumer surplus analysis.

We proceed by specifying a cost function which we call a *utility-dependent translog*. This is essentially a translog cost function, except that some or all of its parameters depend on utility in a nonparametric way. In our base model, the compensated demand system is a partially-linear model which is nonparametric in the utility direction and parametric

with fixed coefficients in the M price directions. We also extend the model to allow the price effects to depend on utility, which yields a varying coefficients econometric structure.

In our model, Engel curves (expenditure-share equations over real expenditure at a particular price vector) are unrestricted, so the demand system can have any rank up to $(M - 1)$; see Lewbel (1991) for a detailed discussion. This contrasts sharply with parametric approaches, such as the popular quadratic almost-ideal demand system in which the demand system is rank 3 and Engel curves are quadratic in the logarithm of expenditure. Although our semiparametric approach is more restrictive than a fully nonparametric approach — such as Haag, Hoderlein, and Pendakur (2005) — at least two important advantages exist. First, our model is composed of functions which have intuitive economic interpretations. Second, our approach avoids the curse of dimensionality faced in multivariate nonparametric regression. A fully nonparametric approach to estimating a consumer demand system has $(M + 1)$ dimensions, whereas our approach has only one-dimensional nonparametric functions.

Our cost function cannot be inverted analytically, so we cannot substitute indirect utility into the compensated demand system to generate an uncompensated demand system. Instead, following the lead of McLaren, Rossiter and Powell (2000) and Wong and McLaren (2005), we numerically invert the cost function into real expenditure — a convenient cardinalisation of utility — and substitute that into the compensated demand system. Thus, our empirical model uses a two-step estimator: in the first step,

we construct a consistent predictor for real expenditure, while in the second step, we estimate the fixed- or varying-coefficient semiparametric model with this constructed regressor. We provide asymptotics for both the fixed- and varying-coefficients consumer demand models with a constructed regressor. In particular, we show $\sqrt{n}-$convergence for the parametric price effects and efficient convergence rates for the nonparametric components. We also provide several consistent predictors for real expenditure, each of which satisfies the requirements for our constructed regressor model.

Lewbel and Pendakur (2008) as well as Sarmiento (2005) have considered problems related to ours. Lewbel and Pendakur (2008) considered a parametric model similar to ours, but they forced price effects to enter linearly. The nonparametric dependence of price effects on real expenditure is a strength of our varying-coefficients approach. Whereas Sarmiento (2005) used a varying-coefficients structure to approximate a full demand system, we use a similar strategy to identify an exact cost function. To our knowledge, no semiparametric consumer demand models have been proposed in which dimensionality is reduced, flexibility in the Engel curve and in price effects is retained, and integrability to an exact cost function is possible. Our model fills this gap.

We analyse Canadian price and expenditure data and find that some expenditure-share equations are 'S-shaped' or even more complex. In addition, we find evidence that compensated price effects are different for rich and poor households. In a cost-of-living experiment, we find that both the complexity of Engel curves and the utility-dependence of price effects affect welfare analyses. Considering the rationality restriction of Slutsky

symmetry, we reject symmetry in the fixed-coefficients version of our model, but not in the varying-coefficients version. This suggests that rejection of Slustky symmetry in previous parametric work may be due to the unduly restrictive incorporation of price effects in typical parametric models. This explanation may supplement other explanations having to do with the unobserved heterogeneity; see Lewbel (2001).

The remainder of our paper is organized as follows: in section 2, we introduce our model of cost and demand, while in section 3, we provide the econometric theory for the estimation of our demand system. In section 4, we discuss inference and the testing and maintaining of rationality conditions. In section 5, we present an empirical example using Canadian microdata.

2   Model

Denote log-prices as $\boldsymbol{p} = [p^1, ..., p^M]$ and log total-expenditure as x. Define indirect utility $V(\boldsymbol{p}, x)$ to be the utility attained when facing log-prices $\boldsymbol{p}$ with log total-expenditure (log-budget) x. Define log-cost, $\log C(\boldsymbol{p}, u)$, as the inverse of $V(\boldsymbol{p}, x)$ over x giving the minimum log total-expenditure required to attain the utility level u when facing log-prices $\boldsymbol{p}$. Denote expenditure shares yielding the share of total expenditure commanded by the jth good as w^j and let $\boldsymbol{w} = [w^1, ..., w^M]$ be the expenditure-share vector. Note that since expenditure shares sum to 1, $w^M = 1 - \sum_{j=1}^{M-1} w^j$. Let $\{W_i^1, ..., W_i^M, P_i^1, ..., P_i^M, X_i\}_{i=1}^n$ be a random $(2M + 1)$ vector summarizing the expenditure shares, log-prices and log-expenditures of a sample of n individuals.

2.1 Utility-Dependent Translog

Consider the "utility-dependent translog" (UTL) log-cost function

$$\log C(\boldsymbol{p}, u) = u + \boldsymbol{p}^\top \overline{\boldsymbol{\beta}}(u) + \frac{1}{2} \boldsymbol{p}^\top \overline{\boldsymbol{A}}(u) \boldsymbol{p}, \tag{1}$$

where u is an ordinal index of utility, so u can always be replaced by $\phi(u)$ where ϕ is an unknown increasing monotonic transformation.[1] The restrictions $\boldsymbol{\iota}^\top \overline{\boldsymbol{\beta}}(u) = 1$ and $\overline{\boldsymbol{A}}(u)^\top \boldsymbol{\iota} = \boldsymbol{0}_M$ are sufficient for homogeneity. The overbars on $\overline{\boldsymbol{\beta}}$ and $\overline{\boldsymbol{A}}$ emphasize that they are arbitrary functions of utility, rather than of an observable variable. The word 'translog' is used because if $\overline{\boldsymbol{\beta}}$ and $\overline{\boldsymbol{A}}$ are independent of utility, the model collapses to the (homothetic) translog model of Christensen, Jorgenson, and Lau (1971). If $\overline{\boldsymbol{A}}$ is independent of utility, then prices come in via 'fixed coefficients' and we shall apply semiparametric partially-linear modelling. If $\overline{\boldsymbol{A}}$ depends arbitrarily on utility, then prices come in via a 'varying coefficients' structure, and we shall use semiparametric varying coefficients modelling. Note that when $\boldsymbol{\beta}$ is linear and $\overline{\boldsymbol{A}}$ is independent of utility, the

---

[1] The presence of u as the leading term in the cost function is not restrictive. Rather, it helps clarify that indirect utility can be (log) money metric at a base price vector $\overline{\boldsymbol{p}}$ as discussed below.

At this point, there are no demographic effects. They can be incorporated into the cost function as follows. Denote a vector of demographic characteristics $\boldsymbol{z}$ where $\boldsymbol{z} = \boldsymbol{0}_T$ for some reference household type. Write the log-cost function as

$$\log C(\boldsymbol{p}, u) = u + \boldsymbol{p}^\top \overline{\boldsymbol{\beta}}(u) + \frac{1}{2} \boldsymbol{p}^\top \overline{\boldsymbol{A}}(u) \boldsymbol{p} + \boldsymbol{p}^\top \boldsymbol{\Gamma} \boldsymbol{z}.$$

Note that this formulation does not allow $\boldsymbol{z}$ to affect cost independently of prices. All the methods proposed below may be adapted to this model.

model collapses to the almost-ideal case.

The dual indirect utility function V is defined by

$$u = V(\boldsymbol{p}, x) \equiv x - \boldsymbol{p}^\top \overline{\boldsymbol{\beta}}(u) - \frac{1}{2}\boldsymbol{p}^\top \overline{\boldsymbol{A}}(u)\,\boldsymbol{p}, \tag{2}$$

which can be solved analytically only in special cases, such as the almost-ideal case.

Applying Shephard's lemma, $\boldsymbol{w} = \nabla_{\boldsymbol{p}} \log C(\boldsymbol{p}, u)$, to the log-cost function (1) yields a vector of compensated (Hicksian) expenditure share equations $\boldsymbol{\omega}(\boldsymbol{p}, u)$ given by

$$\boldsymbol{\omega}(\boldsymbol{p}, u) = \overline{\boldsymbol{\beta}}(u) + \overline{\boldsymbol{A}}(u)\,\boldsymbol{p}\,, \tag{3}$$

where $\overline{\boldsymbol{A}}(u) = \overline{\boldsymbol{A}}(u)^\top$. This compensated expenditure-share system is very simple: were utility u observed, it would be estimable by semiparametric methods for partially-linear or varying-coefficients models, depending on whether or not $\overline{\boldsymbol{A}}$ varies with u.

Set a reference vector of prices to unity, so that log-reference prices are $\overline{\boldsymbol{p}} = \boldsymbol{0}_M$, and note that indirect utility satisfies

$$V(\overline{\boldsymbol{p}}, x) = x. \tag{4}$$

The choice of reference price vector is innocuous because we can always make utility (log-) money metric at one price vector. Define 'log real-expenditure', $x^R = R(\boldsymbol{p}, x)$, as the level of expenditure at $\overline{\boldsymbol{p}}$ which yields the same level of utility as x at $\boldsymbol{p}$. It is implicitly defined by

$$V(\boldsymbol{p}, x) = V(\overline{\boldsymbol{p}}, x^R), \tag{5}$$

and given by

$$x^R = R(\boldsymbol{p}, x) = \log\left[C\overline{\boldsymbol{p}}, V(\boldsymbol{p}, x)\right].$$

Here, R and V are ordinally equivalent representations of preferences because R is a monotonically increasing function of V. However, R is cardinalised — its value is measured in base-price log-money units. Substituting $x^R = R(\boldsymbol{p}, x)$ into the compensated demand system yields a demand system which depends only on observables. Define $\boldsymbol{\beta}\left(x^R\right) = \overline{\boldsymbol{\beta}}\left[V(\overline{\boldsymbol{p}}, x^R)\right]$ and $\boldsymbol{A}\left(x^R\right) = \overline{\boldsymbol{A}}\left[V(\overline{\boldsymbol{p}}, x^R)\right]$. Here, the overbars are removed to indicate that these functions depend on the particular cardinalisation of utility, x^R.

Of course, the analytical form for R may be difficult to recover if $\log C$ is not analytically invertible. However, in our model, R can be estimated numerically as follows. Define $x^N = N(\boldsymbol{p}, x)$ as the log *nominal* expenditure function which gives the log-expenditure necessary at $\boldsymbol{p}$ to give the same utility as x at $\overline{\boldsymbol{p}}$. Since $R(\overline{\boldsymbol{p}}, x) = x$, we have

$$
\begin{aligned}
V(\boldsymbol{p}, x^N) &= V(\overline{\boldsymbol{p}}, x) \Leftrightarrow & (6) \\
x^N &= N(\boldsymbol{p}, x) = x + \boldsymbol{p}^\top \boldsymbol{\beta}\left(x\right) + \frac{1}{2}\boldsymbol{p}^\top \boldsymbol{A}\left(x\right)\boldsymbol{p}.
\end{aligned}
$$

Log real-expenditure is then given by the inverse of N with respect to x at each $\boldsymbol{p}$; i.e.,

$$x^R = R(\cdot, x) = N^{-1}(\cdot, x). \tag{7}$$

Here, R can be found at each $\boldsymbol{p}$ by numerical inversion of N. If log-cost is increasing in utility at $\boldsymbol{p}$, then N is monotonically increasing in x, and is easily inverted numerically.[2]

---

[2]Cost is locally weakly increasing in $\boldsymbol{p}$ if and only if expenditure shares are weakly greater than zero. The parametric component $\boldsymbol{Ap}$ guarantees violations of positivity with $\boldsymbol{p}$ large (small) enough

Note that log real-expenditure x^R is an interesting tool for welfare economics: the cost-of-living index I for a person facing prices $\boldsymbol{p}$ compared to $\overline{\boldsymbol{p}}$ may be defined by

$$\log I(\boldsymbol{p}, x) = x - R(\boldsymbol{p}, x) = x - x^R.$$

Given $x^R = R(\boldsymbol{p}, x)$, uncompensated shares over x^R are given by substituting log real-expenditure x^R (which holds utility constant) for utility u in the compensated demand system and replacing $\overline{\boldsymbol{\beta}}$ and $\overline{\boldsymbol{A}}$ with $\boldsymbol{\beta}$ and $\boldsymbol{A}$:

$$\boldsymbol{w}(\boldsymbol{p}, x^R) = \boldsymbol{\beta}\left(x^R\right) + \boldsymbol{A}(x^R)\boldsymbol{p}. \tag{8}$$

If $\overline{\boldsymbol{A}}$ is independent of utility, then $\boldsymbol{A}$ is independent of log real-expenditure, x^R, and we have the following partially-linear semiparametric structure:

$$\boldsymbol{w}(\boldsymbol{p}, x^R) = \boldsymbol{\beta}\left(x^R\right) + \boldsymbol{A}\boldsymbol{p} \tag{9}$$

where $\boldsymbol{A}$ is a matrix of parameters rather than of functions.

A nice feature is the clear interpretability that obtains here: we shall call the $\boldsymbol{\beta}(\cdot)$ 'Engel curve functions' because they yield the Engel curves at the reference price vector. We shall further call the elements of $\boldsymbol{A}$ 'compensated price effects' because they give the effect of price changes on demand holding utility constant. Thus, our structure is fairly simple: the demand system is characterized by a set of Engel curve functions and a matrix of compensated price effects.

when $\boldsymbol{A} \neq \boldsymbol{0}$. Thus, we cannot restrict the UTL to satisfy increasingness. However, no commonly-used parametric demand system is globally increasing either. Since C cannot be restricted to global increasingness in u, one might be cautious about evaluating N at price vectors very far from the observed price vectors.

3 Estimation

We refer to the model given by equation (9) as a 'fixed-coefficients' model because the coefficients governing compensated price effects are fixed with respect to all other variables. In contrast, the model given by equation (8) has compensated price effects which depend on x^R, so that substitution responses may be different for rich and poor persons. We refer to this as a 'varying-coefficients' model. The econometrics involved in estimating these two types of models are rather different from each other, so we treat them separately, beginning with the simpler fixed-coefficients model.

Consider estimation of the UTL demand system where $\boldsymbol{A}$ $(\overline{\boldsymbol{A}})$ is independent of log real-expenditure (utility) given by equation (9). If x^R were observed, then this would be a standard partially-linear model. Since it is unobserved, we instead use a generated regressor approach: we first construct a consistent predictor for x^R for each observation and then estimate the partially-linear model (9) by semiparametric methods. In this section, we provide two consistent initial predictors for x^R which satisfy the conditions required for consistency of the estimated demand system and for efficiency of the estimator for $\boldsymbol{A}$.

3.1 Estimation of Fixed-Coefficients Price Effects

To estimate the matrix $\boldsymbol{A}$ $\sqrt{n}$- consistently and efficiently, we discuss two approaches: the kernel-smoothing estimator of Rodríguez-Poó, Sperlich, and Fernández (2005) and

the differencing estimator of Yatchew (1997).[3] Both approaches are based on the same idea — differencing out the nonparametric vector-function $\boldsymbol{\beta}$.

It is convenient to describe the estimation method equation-by-equation. We face the data generating process

$$W_i^j = \beta^j \left[R(\boldsymbol{P}_i, X_i) \right] + \boldsymbol{a}^j \boldsymbol{P}_i + \epsilon_i^j \quad , \; j = 1, \ldots, M \tag{10}$$

where $\boldsymbol{a}^j$ are the rows of $\boldsymbol{A}$, so that $\boldsymbol{A} = \left[\boldsymbol{a}^1 | \; ... | \; \boldsymbol{a}^M \right]$, and the disturbances ϵ_i^j, $i = 1, \ldots, n$ are independently and identically distributed with mean zero and variance function $\sigma_{jj}(x, \boldsymbol{p})$ for all j, being bounded from above. Define the compact set $\mathcal{X}^R$ as the region of interest of x^R. For any approach, we shall use the following assumptions:

[**A1**] The vector of error terms $\boldsymbol{\epsilon} = (\epsilon^1, \ldots, \epsilon^M)^\top$ has the covariance function $Cov[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top | x, \boldsymbol{p}]$ $= \{\sigma_{lk}(x, \boldsymbol{p})\}_{l,k=1}^M = \boldsymbol{\Sigma}_\epsilon(x, \boldsymbol{p})$ is bounded and Lipschitz continuous in each argument.

[**A2**] The (marginal) density $f(\cdot)$ of X^R is uniformly bounded away from both zero and infinity, and has a continuous second derivative on $\mathcal{X}^R$.

[**A3**] The functions $\beta^j(\cdot)$ have bounded and continuous second derivatives on $\mathcal{X}^R$.

The econometric strategy is to difference out the contribution of β^j smoothly using kernel weights on the distances, thus leaving only the contribution of $\boldsymbol{a}^j \boldsymbol{P}_i$. Consider for each j the sample

$$w_i^j - w_k^j = \beta^j(x_i^R) - \beta^j(x_k^R) + \boldsymbol{a}^j(\boldsymbol{p}_i - \boldsymbol{p}_k) + \epsilon_i^j - \epsilon_k^j \; , \forall \; i \neq k \; .$$

---

[3]Alternatively, one might use the approaches as proposed by Speckman (1988).

Weighting inversely to the distance $|x_i^R - x_k^R|$ will cancel the contribution of β^j due to its smoothness. Our estimator is given by

$$
\begin{aligned}
\hat{A}_{RSF} &= \hat{H}_{PP}^{-1}\hat{H}_{PW} &\qquad (11)\\
\hat{H}_{PP} &= \binom{n}{2}^{-1} \sum_{i=1}^{n}\sum_{k=i+1}^{n-1} (\boldsymbol{p}_i - \boldsymbol{p}_k)(\boldsymbol{p}_i - \boldsymbol{p}_k)^{T\top}\hat{v}_{ik}\\
\hat{H}_{PW} &= \binom{n}{2}^{-1} \sum_{i=1}^{n}\sum_{k=i+1}^{n-1} (\boldsymbol{p}_i - \boldsymbol{p}_k)(\boldsymbol{w}_i - \boldsymbol{w}_k)^{\top}\hat{v}_{ik}\\
\text{where} &\qquad \hat{v}_{ik} = K_h(\hat{x}_i^R - \hat{x}_k^R)
\end{aligned}
$$

Here, $K_h(v) = \frac{1}{h}K(vh^{-1})$ is a kernel function. We further assume:

[**R1**] The kernel K is bounded, symmetric, compactly supported, and integrates to one, having first moment equal to zero and a continuous second derivative.

[**R2**] The prices $\boldsymbol{P}^j$ have nondegenerate conditional distributions given X^R and X, respectively, with $E[\boldsymbol{P}^j|X^R = x^R] = g_j(x^R)$, and $E[Cov(\boldsymbol{P}|X^R)] = \boldsymbol{\Sigma}_{P|X^R}$. The functions $g_j(\cdot)$ have bounded first derivatives on $\mathcal{X}^R$.

[**R3**] For bandwidth h we need $nh^6 \to \infty$ and $nh^8 \to 0$.

We shall need a consistent (nonparametric) predictor for all x^R from the range $\mathcal{X}^R$:

[**X1**] There exists a consistent predictor for each $x^R \in \mathcal{X}^R$ so that

$$
\hat{x}_i^R = R(\boldsymbol{p}_i, x_i) + B_X(x_i, \boldsymbol{p}_i) + u_i , \quad u_i := u(x_i, \boldsymbol{p}_i) ,
$$

where $B_X(\cdot)$ is the bias, $u(x, \boldsymbol{p})$ the stochastic error with $E[u(x, \boldsymbol{p})] = 0$ and variance

function $\sigma_X^2(x, \boldsymbol{p})$. Both, B_X and σ_X are Lipschitz continuous in x and converge to zero as $n \to \infty$. Further, $E[u_i u_k \sigma_X(x_i, \boldsymbol{p}_i) \sigma_X(x_k, \boldsymbol{p}_k)] = O(\frac{1}{n})$ uniformly for all $i \neq k$.

These assumptions are fulfilled by any simple kernel estimator. As both x^R and the conditional mean function of W come from related data generating processes, we have to allow for correlation between the stochastic errors u_i and ϵ_i from equation (10).

[X2] For all $j = 1, \ldots, M$ we have $E[u_i^r (\epsilon_k^j)^s] \leq O(\frac{1}{n^r}) \; \forall \; i, k, \; s = 1, 2, \; r = 1, 2$; and $E[u_x^\gamma] = \sigma_X^\gamma(x, \boldsymbol{p}) = o(\frac{1}{n})$ for $\gamma > 2$. Finally, the third cumulant $E[\epsilon_i^j \epsilon_k^j u_i]$ is assumed to be of order $o(\frac{1}{n})$. All these rates hold uniformly for all i, k.

In addition, for the efficient estimation of the parametric part, we need

[X3] For all x_i^R, $i = 1, \ldots, n$ there exist predictors $\hat{x}_i^R$ such that $\sup_i |\hat{x}_i^R - x_i^R| h^{-2} = o_p(1)$ for the bandwidths given in [R3].

We can now establish the asymptotics of our model as a corollary of results from Rodríguez-Poó, Sperlich, and Fernández (2005):

Corollary 1 *Under assumptions [A1]–[A3], [R1]–[R3], [X1]–[X3], we have for each row $\boldsymbol{a}^j$ of $\boldsymbol{A}$ as n goes to ∞*

$$\sqrt{n} \left(\hat{\boldsymbol{a}}_{RSF}^j - \boldsymbol{a}^j \right) \longrightarrow \mathcal{N} \left(\boldsymbol{0}_M, E[\boldsymbol{\Sigma}_{P|X^R}^{-1}] E[P_X \sigma_{jj}(X, \boldsymbol{P}) P_X^\top] E[\boldsymbol{\Sigma}_{P|X^R}^{-1}] \right) \; ,$$

where $P_X := \boldsymbol{P} - E[\boldsymbol{P}|X^R]$. The covariance matrix between vector $\boldsymbol{a}^j$ and $\boldsymbol{a}^k$ is

$$E \left[\boldsymbol{\Sigma}_{P|X^R}^{-1} \right] E[P_X \sigma_{jk}(X, \boldsymbol{P}) P_X^\top] E[\boldsymbol{\Sigma}_{P|X^R}^{-1}] \; ,$$

for all $k, j = 1, \ldots, M$.

14

Yatchew (1997) proposed a different approach which does not employ smoothing. Consequently, this approach is simpler in practise though less efficient asymptotically. First, we order the observations $\{\boldsymbol{w}_i, \hat{x}_i^R, \boldsymbol{p}_i\}_{i=1}^n$ by $\hat{x}_i^R$. Consider for each equation $j = 1, \ldots, M$ the differences

$$\sum_{i=0}^m d_i w_{k-i}^j = \sum_{i=0}^m d_i \beta^j(\hat{x}_{k-i}^R) + \sum_{j=1}^M \sum_{i=0}^m d_i p_{k-i}^j a_j + \sum_{i=0}^m d_i \epsilon_{k-i}^j , \quad k = m+1, \ldots, n$$

with differencing coefficients $d_0, d_1, \ldots, d_m$ fulfilling

$$\sum_{i=0}^m d_i = 0, \text{ and } \sum_{i=0}^m d_i^2 = 1 . \tag{12}$$

Again, the contribution of β^j is canceled out due to its assumed smoothness and the assumption:

[**Y1**] The data are ordered such that $\frac{1}{n} \sum_{i=1}^n |\hat{x}_i^R - \hat{x}_{i-1}^R|^2 = O(n^{-2(1-\delta)})$ for δ positive and arbitrarily close to zero.

Define optimal differencing weights by minimizing $\sum_{k=1}^m (\sum_l d_l d_{l+k})^2$ under the constraints (12); see Hall, Kay, and Titterington (1990). Then, define the 'differenced' vectors and matrices, $\Delta y = \sum_{k=0}^m d_k y_{i-k}$ with y being $\boldsymbol{W}$ or $\boldsymbol{P}$. Finally, apply ordinary least squares to the differenced data:

$$\hat{\boldsymbol{A}}_Y = [\Delta \boldsymbol{P} \Delta \boldsymbol{P}^\top]^{-1} \Delta \boldsymbol{P}^\top \Delta \boldsymbol{W} \tag{13}$$

The asymptotics follow mainly from results in Yatchew (1997):

Corollary 2 *Under assumptions [A1]–[A3], [X1]–[X3], [Y1], and [R2] it holds that for each row $\boldsymbol{a}^j$ of $\boldsymbol{A}$ we have for n going to ∞*

$$\sqrt{n}\left(\hat{\boldsymbol{a}}_Y^j - \boldsymbol{a}^j\right) \longrightarrow \mathcal{N}\left(\boldsymbol{0}_M, (1 + \frac{1}{2m})E[\boldsymbol{\Sigma}_{P|X^R}^{-1}]E[P_X\sigma_{jj}(X, \boldsymbol{P})P_X^\top]E[\boldsymbol{\Sigma}_{P|X^R}^{-1}]\right),$$

where $P_X := \boldsymbol{P} - E[\boldsymbol{P}|X^R]$. The covariance matrix between vector $\boldsymbol{a}^j$ and $\boldsymbol{a}^k$ is

$$(1 + \frac{1}{2m})E[\boldsymbol{\Sigma}_{P|X^R}^{-1}]E[P_X\sigma_{jk}(X, \boldsymbol{P})P_X^\top]E[\boldsymbol{\Sigma}_{P|X^R}^{-1}].$$

The two approaches are similar: in Yatchew's approach the kernel weights were substituted by the so-called 'differencing weights', and the parameter m corresponds to bandwidth h in the first approach. In the empirical work below, we implement Yatchew's estimator for the fixed-coefficient model.

3.2 Estimation of Engel Curve Functions

An efficient nonparametric estimator for $\boldsymbol{\beta}(x^R)$ may be obtained by standard methods if x_i^R is known and we have a $\sqrt{n}$-consistent estimate of the matrix $\boldsymbol{A}$. In this case, we simply apply a local estimator on the M one-dimensional regression problems

$$\boldsymbol{W}^j - \hat{\boldsymbol{a}}^j\boldsymbol{P} = \beta^j(X^R) + \epsilon^j \qquad j = 1, \ldots, M.$$

However, in our case, we have only a consistent predictor x_i^R, and so we must take into account the bias and randomness of the $\hat{x}_i^R$.

Let $K_h(v) = \frac{1}{h}K(vh^{-1})$ again be our kernel function with bandwidth h. We denote estimators of β^j and their first derivatives by $\theta_1(x^R) = \beta^j(x^R)$, $\theta_2(x^R) = \frac{\partial}{\partial v}\beta^j(v)|_{v=x^R}$

16

and employ a local-linear estimator:

$$\hat{\theta}(x^R) = \mathrm{argmin}_\theta \sum_{i=1}^{n} \left[(w_i^j - \hat{\boldsymbol{a}}^j \boldsymbol{p}_i) - \theta_1 - \theta_2(\hat{x}_i^R - x^R) \right]^2 K_h(\hat{x}_i^R - x^R)$$

Having predictors in the kernel as well as inside the square brackets complicates the asymptotic theory. Since we are only interested in the levels of β^j, we concentrate here on the asymptotic distribution of $\hat{\beta}^j(v)$. The following theorem gives the asymptotic behavior of $\hat{\beta}^j$ with a constructed regressor. The proof is given in the appendix.

Theorem 3 *Assume that [A1]–[A3], [R1] and [X1]–[X2] hold. Assume that the following rates hold uniformly: $B_X = o(h)$, and $\sigma_X^2 = O(\frac{1}{ng_n})$ for a g_n such that $\frac{1}{ng_n}$ is $o(h^2)$. Assume that there exists $(x^0, \boldsymbol{p}^0)$ such that $x^R = R(\boldsymbol{p}^0, x^0)$ and that for $n \to \infty$, nh and h^{-1} go to ∞. Then, with $\boldsymbol{\beta}(x^R) := \{\beta^j(x^R)\}_{j=1}^{M}$ and $\hat{\boldsymbol{\beta}}(x^R) := \{\hat{\beta}^j(x^R)\}_{j=1}^{M}$, we have for x^R in $\mathcal{X}^R$:*

$$\sqrt{(nh \wedge ng_n)} \left[\hat{\boldsymbol{\beta}}(x^R) - \boldsymbol{\beta}(x^R) - B_\beta(x^R) \right] \longrightarrow \mathcal{N}\left(0, \boldsymbol{\Sigma}_\beta(x^R)\right)$$

with bias

$$B_\beta(x^R) = \frac{h^2}{2}\mu_2(K)\boldsymbol{\beta}''(x^R) - B_X(x^0, \boldsymbol{p}^0)\boldsymbol{\beta}'(x^R) ,$$

where $\boldsymbol{\beta}'(x^R)$, $\boldsymbol{\beta}''(x^R)$ are the vectors of the first and second derivatives, respectively, of $\boldsymbol{\beta}(x^R)$. Further, recalling that $f(\cdot)$ denotes the marginal density of X^R, then with $\mu_l(K) = \int v^l K(v)dv$ we have

$$\frac{1}{nh \wedge ng_n}\boldsymbol{\Sigma}_\beta(x^R) = \frac{1}{nh}f^{-1}(x^R)||K||_2^2\boldsymbol{\Sigma}_\epsilon(x^R) \oplus \sigma_X^2(x^0, \boldsymbol{p}^0)\boldsymbol{\beta}'^2(x^R) ,$$

where $\oplus$ denotes 'element-wise' summation.

The newly-introduced parameter g_n corresponds to a smoothing parameter in the prediction of x^R. In case of using kernel-smoothing methods and g_n as a bandwidth it is clear that, without bias-reducing methods, the bias B_X is of rate g_n^2, the variance σ_X^2 of rate $\frac{1}{ng_n}$. Thus, the assumptions $B_x = o(h)$ and $\sigma_X^2 = o(h^2)$ are trivially fulfilled. Theorem 3 shows that the estimator for $\boldsymbol{\beta}(x^R)$ with a generated regressor is consistent and that its asymptotic bias and variance are driven in part by the bias and variance, respectively, of the predictor.

We now consider improving the estimation of $\boldsymbol{\beta}$ by applying iteration. Recall that the value of log real expenditures x_i^R can be recalculated using estimates of $\boldsymbol{\beta}$ and $\boldsymbol{A}$. There are two purposes for iteration: first, one can use iteration to generate model-consistent results. Here, if $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{A}}$ are iterated with $\hat{x}_i^R$, the estimates of each may 'settle down' in such a way that $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{A}}$ imply $\hat{x}_i^R$, and $\hat{x}_i^R$ implies $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{A}}$. Such estimates may be called 'model-consistent', and have the advantage that either set of estimates completely characterises the model: that is, one could present empirical results on *either* $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{A}}$ *or* on $\hat{x}_i^R$ without any loss of information.

A second purpose of iteration is to try to reduce (or to eliminate) the influence of the pre-estimation of $\hat{x}^R$. The corollaries and theorem above hold for any predictor of log real-expenditure fulfilling fairly weak conditions, including those which result from iteration. Thus, iteration does not reduce the efficiency of the estimator. One could try to establish conditions on the model enabling us to apply a contraction result, which would show that iteration would give an asymptotically efficient estimate. However,

such an exercise is quite difficult and beyond the scope of the present paper. We are more interested in the practical question as to whether the initial prediction affects the final estimates in real data. ¿From the results given above, we see that after iteration the asymptotic distribution of $\hat{\boldsymbol{\beta}}$ only changes in the additive term containing the bias and the variance of the predictor. We may assess the contribution of these terms via a simple subsampling approach as follows:

(i) Predict the n log real-expenditures x_i^R for the full sample using initial estimates of $\boldsymbol{\beta}$ and $\boldsymbol{A}$ which come from a random subsample of the data.

(ii) Estimate with them and the full sample the function $\boldsymbol{\beta}$ and $\boldsymbol{A}$, recalculate with them the x_i^R, and iterate this until convergence.

(iii) Repeat steps (i) and (ii) B times to determine the distribution of the final estimates of $\boldsymbol{\beta}$, $\boldsymbol{A}$, and the log real-expenditures x_i^R.

If the final estimates do not vary over the different subsamples, then the initial prediction has little impact on the final estimates. In this case, iteration yields an efficient estimator. In the empirical work below, we shall show that the iterated estimates are roughly independent of the initial pre-estimates, and therefore we suggest using iteration in practice. The sampling distribution for the nonparametric estimates should be estimated by resampling (bootstrap) methods, which we discuss in Section 6.

3.3 Consistent Predictors of Log Real-Expenditure x_i^R

The results above are all based on the assumption that we have predictors of the x_i^R, $i = 1, \ldots, n$ fulfilling conditions given in [X1]–[X3]. We first show an initial estimate of x_i^R computed from initial estimates of $\boldsymbol{\beta}$ and $\boldsymbol{A}$. Then, we discuss different approaches to initial estimates of $\boldsymbol{\beta}$ and $\boldsymbol{A}$ which satisfy [X1]–[X3] and are not too burdensome.

Our initial estimators of x_i^R will use initial estimates of $\boldsymbol{\beta}$ and $\boldsymbol{A}$ that we shall call $\boldsymbol{\beta}_0$ and $\boldsymbol{A_0}$. Define N_0 as the log-nominal expenditure function using these initial estimates $\boldsymbol{\beta}_0$ and $\boldsymbol{A_0}$:

$$N_0(\boldsymbol{p}, x) = x + \boldsymbol{p}^\top \boldsymbol{\beta}_0(x) + \boldsymbol{p}^\top \boldsymbol{A_0} \boldsymbol{p}.$$

Then define R_0 as the inverse with respect to x of N_0, so that $R_0(\cdot, x) = N_0^{-1}(\cdot, x)$, and use $\hat{x}_i^R = R_0(\boldsymbol{p}, x_i)$ as the (initial) predictor of x_i^R.

Given monotonically increasing costs in utility, both $R(\boldsymbol{p}, x)$ and $N(\boldsymbol{p}, x)$ are increasing in x for each $\boldsymbol{p}$. Therefore, we may invert N and derive the convergence rate. For each $\boldsymbol{p}$ fixed, and $t = \hat{N}(\boldsymbol{p}, x)$, $R(\boldsymbol{p}, t) = N^{-1}(\boldsymbol{p}, t)$, $\hat{R}(\boldsymbol{p}, t) = \hat{N}^{-1}(\boldsymbol{p}, t)$

$$
\begin{aligned}
\sup_t |\hat{R}(\boldsymbol{p}, t) - R(\boldsymbol{p}, t)| &= \sup_u |\hat{R}\{\boldsymbol{p}, \hat{N}(\boldsymbol{p}, u)\} - R\{\boldsymbol{p}, \hat{N}(\boldsymbol{p}, u)\}| \\
\sup_u |u - R\{\boldsymbol{p}, \hat{N}(\boldsymbol{p}, u)\}| &= \sup_u |R\{\boldsymbol{p}, N(\boldsymbol{p}, u)\} - R\{\boldsymbol{p}, \hat{N}(\boldsymbol{p}, u)\}| \\
&\leq \sup_t |\frac{d}{dt} R(\boldsymbol{p}, t)| \ \sup_u |N(\boldsymbol{p}, u) - \hat{N}(\boldsymbol{p}, u)|.
\end{aligned}
$$

This implies that $\hat{x}_i^R \equiv \hat{R}(\boldsymbol{p}_i, x_i)$ inherits the convergence rate of $\hat{N}(\cdot, \cdot)$ which itself inherits the convergence rate of the initial estimates $\boldsymbol{\beta}_0(\cdot)$ and $\hat{\boldsymbol{A}}_0$.

Recall that $\overline{\boldsymbol{p}} = \boldsymbol{0}_M$, which implies that $R(\overline{\boldsymbol{p}}, x) = x$, and that for observations facing $\overline{\boldsymbol{p}}$, $x_i^R = x_i$ and $\boldsymbol{A}\boldsymbol{p} = \boldsymbol{0}_M$. A consistent initial estimator for $\boldsymbol{\beta}(x^R)$, denoted $\boldsymbol{\beta}_0(x^R)$, may be obtained by nonparametric estimation of expenditure shares on log-expenditure using only those observations facing $\overline{\boldsymbol{p}}$:

$$\boldsymbol{\beta}_0(x^R) = E[\boldsymbol{W}_i | X_i = x, \boldsymbol{P}_i = \overline{\boldsymbol{p}}].$$

This estimator may be constructed either by estimating the univariate nonparametric Engel curve at $\overline{\boldsymbol{p}}$, using only the observations which face $\overline{\boldsymbol{p}}$, or by estimating the nonparametric demand system using all n observations facing all price vectors, and evaluating this demand system at $\boldsymbol{p} = \overline{\boldsymbol{p}}$.

Note that the matrix $\boldsymbol{A}$ is the matrix of log-price derivatives of compensated expenditure share equations. In general, the matrix of compensated semi-elasticities, $\boldsymbol{\Upsilon}(\boldsymbol{p}, x)$, may be expressed in terms of observables as:

$$\boldsymbol{\Upsilon}(\boldsymbol{p}, x) = \boldsymbol{\nabla_p}\boldsymbol{w}(\boldsymbol{p}, x) + \boldsymbol{\nabla_x}\boldsymbol{w}(\boldsymbol{p}, x)\boldsymbol{w}(\boldsymbol{p}, x)^\top.$$

Given our fixed-coefficients model, $\boldsymbol{\Upsilon}(\boldsymbol{p}, x) = \boldsymbol{A}$ is a matrix of constants. Following Haag, Hoderlein, and Pendakur (2005), one may estimate $\boldsymbol{\Upsilon}(\boldsymbol{p}, x)$ via nonparametric methods using local polynomial modelling of the Marshallian expenditure-share system. Methods like these yield estimated compensated semi-elasticities which depend on $\boldsymbol{p}, x$, which we denote $\widehat{\boldsymbol{\Upsilon}}(\boldsymbol{p}, x)$. A consistent estimator for $\boldsymbol{A}$ is thus given by

$$\widehat{\boldsymbol{A}}_0 = \frac{1}{n}\sum_{i=1}^{n}\widehat{\boldsymbol{\Upsilon}}(\boldsymbol{P}_i, X_i). \tag{14}$$

Here, $\widehat{\boldsymbol{A}}_0$ may or may not be symmetric.

We note that Lewbel (2001) showed that $\widehat{\boldsymbol{\Upsilon}}(\boldsymbol{p}, x)$ is only a consistent estimate of $\boldsymbol{\Upsilon}(\boldsymbol{p}, x)$ when the disturbance terms on the right-hand side of the regression are not behavioral parameters or when they satisfy rather complex covariance conditions that are difficult to verify in practice. An alternative estimator that avoids this problem was proposed by Crossley and Pendakur (2006) as well as by Hoderlein (2005). They suggested an estimator of the compensated semi-elasticity matrix which exploits the Slutsky symmetry restrictions:

$$\widehat{\boldsymbol{\Upsilon}}(\boldsymbol{p}, x) = \frac{1}{2} \left[\widehat{\boldsymbol{\nabla}_{\boldsymbol{p}} \boldsymbol{w}}(\boldsymbol{p}, x) + \widehat{\boldsymbol{\nabla}_{\boldsymbol{p}} \boldsymbol{w}}(\boldsymbol{p}, x)^\top + \widehat{\boldsymbol{\nabla}_{\boldsymbol{x}}} \left\{ \boldsymbol{w}(\boldsymbol{p}, x) \boldsymbol{w}(\boldsymbol{p}, x)^\top \right\} \right],$$

where $\widehat{\boldsymbol{\nabla}_{\boldsymbol{p}} \boldsymbol{w}}(\boldsymbol{p}, x)$ is an estimator of the derivative of the budget-share vector with respect to the log-price vector and $\widehat{\boldsymbol{\nabla}_{\boldsymbol{x}}} \left\{ \boldsymbol{w}(\boldsymbol{p}, x) \boldsymbol{w}(\boldsymbol{p}, x)^\top \right\}$ is an estimator of the derivative of the outer-product of budget-shares with respect to log-expenditure. Because $\boldsymbol{\nabla}_{\boldsymbol{x}} \left\{ \boldsymbol{w}(\boldsymbol{p}, x) \boldsymbol{w}(\boldsymbol{p}, x)^\top \right\}$ can be consistently estimated regardless of the structure of unobserved heterogeneity, this symmetry-restricted version of $\widehat{\boldsymbol{\Upsilon}}$ is not subject to Lewbel's (2001) critique. The matrix $\widehat{\boldsymbol{A}}_0$ is again taken as the sample average of $\widehat{\boldsymbol{\Upsilon}}(\boldsymbol{p}_i, x_i)$.

Note that for either estimator of $\boldsymbol{\Upsilon}(\boldsymbol{p}, x)$, we use high-dimensional nonparametric pre-estimators that have very slow convergence rates. Fortunately, the averaging in equation (14) reduces the variance of $\widehat{\boldsymbol{A}}_0$ such that with an undersmoothed pre-estimate $\widehat{\boldsymbol{\Upsilon}}$ we end up with a rate that easily fulfills the rates necessary to satisfy [X1]—[X3], the assumptions used in Theorem 3. See Hengartner and Sperlich (2008) for similar uses of averaging to improve convergence rates.

3.4 Estimation with Varying Price Effects

We now turn to the case where $\boldsymbol{A}(x^R)$ $(\overline{\boldsymbol{A}}(x^R))$ depends on log real expenditure (utility). We again assume that a consistent predictor of log real-expenditures X_i^R exists (we provide one such predictor below). The model

$$E[\boldsymbol{W}|\boldsymbol{p}, x^R] \;=\; \boldsymbol{\beta}\left(x^R\right) + \boldsymbol{A}(x^R)\boldsymbol{p} \tag{15}$$

can be interpreted as a varying-coefficients model that is linear in $\boldsymbol{p}$ but with coefficients that vary with x^R. For observed $\{\boldsymbol{p}, x^R\}$ these models have been well-studied in the non- and semiparametric literature – see Cleveland, Grosse, and Shyu (1991) as well as Fan and Zhang (1999) – but not for constructed regressors and simultaneous equations.

We use a local-linear model to get estimates of the functions β^j and $\boldsymbol{a}^j = (a_1^j, a_2^j, \ldots, a_M^j)$ at a given point x_0^R. For all j, minimize over the scalars β_0^j and β_1^j and the vectors $\boldsymbol{a}_0^j$ and, $\boldsymbol{a}_1^j$ the kernel-weighted sum-of-squares

$$\sum_{i=1}^n \left[W_i^j - \beta_0^j - \beta_1^j(\hat{x}_i^R - x_0^R) - \left\{ \boldsymbol{a}_0^j + \boldsymbol{a}_1^j(\hat{x}_i^R - x_0^R)\right\}' \boldsymbol{P}_i \right]^2 K_h(\hat{x}_i^R - x_0^R) \tag{16}$$

and then set $\widehat{\beta^j}(x_0^R) = \beta_0^j$, $\widehat{\boldsymbol{a}^j}(x_0^R) = \boldsymbol{a}_0^j$ for all j, k. For the ease of notation, let us set log-prices $\boldsymbol{P}_i^0 \equiv \boldsymbol{0}_M$ for all i, and $a_0^j(x^R) = \beta^j(x^R)$. We need the assumptions

[V1] $E[(p^j)^{2s}] < \infty$ for some $s > 2$, $j = 0, \ldots, M$. Further, the second derivative of $r_{jk}(x^R) := E[p^j p^k | x^R]$ is continuous and bounded from zero on $\mathcal{X}^R$ the for all j, k.

and replace [A3] by

[V2] The second derivatives of $\boldsymbol{A}(x^R)$ are continuous and bounded on $\mathcal{X}^R$ for all j, k.

23

The following theorem describes the asymptotic behavior for the varying-coefficient model with constructed regressors. The proof is deferred to the appendix.

Theorem 4 *Assume the same conditions as in Theorem 3 without [A3], but adding [V1] and [V2]. Define the estimators as in (16), and set $\alpha_k = \left(a_0^k, a_1^k, \cdots, a_M^k\right)^\top (x^R)$ for $k = 1, \ldots, M$. Then we have*

$$\sqrt{(nh \wedge ng_n)} \left\{ \hat{\alpha}_k - \alpha_k - B_k(x^R) \right\} \longrightarrow \mathcal{N}\left(0, \boldsymbol{\Sigma}_{\alpha_k}(x^R)\right) .$$

with bias

$$B_k(x^R) = \frac{h^2}{2} \mu_2(K) \alpha_k'' - B_X(x^0, \boldsymbol{p}^0) \alpha_k' ,$$

where α_k', α_k'' are the vectors of the first and second derivatives, respectively, of $\alpha_k(x^R)$.

The covariance structure is given by

$$\frac{1}{nh \wedge ng_n} \boldsymbol{\Sigma}_{\alpha_k}(x^R) = \frac{1}{nh} f^{-1}(x^R) ||K||_2^2 \Omega \boldsymbol{\Sigma}_{\epsilon k,k}(x^R) \oplus \sigma_X^2(x^0, \boldsymbol{p}^0) \alpha_k'^2 ,$$

where Ω is the inverse of $E\left[(P^0, P^1, \ldots, P^M)^\top (P^0, P^1, \ldots, P^M)|x^R\right]$, and $\boldsymbol{\Sigma}_{\epsilon k,k}(x^R)$ is the $(k,k)^{\text{th}}$ element of $\boldsymbol{\Sigma}_\epsilon(x^R)$.

Theorem 4 is analogous to Theorem 3 (the nonparametric part in the fixed coefficients model). Theorem 4 shows that the varying-coefficients estimator for $\boldsymbol{A}(x^R)$ with a constructed regressor is consistent, and that its bias and variance are partly driven by the bias and variance, respectively, of the predictor of x^R.

3.4.1 Consistent Predictor for $\hat{x}_i^R$

As in the fixed-coefficients case, we proceed by plugging initial estimates of $\boldsymbol{\beta}$ and $\boldsymbol{A}$ and into the function $N(\boldsymbol{p}, x)$, and then inverting around x to get $R(\boldsymbol{p}, x)$ which defines our predictor for real expenditure. We may get consistent initial estimators for $\boldsymbol{\beta}(x^R)$ as in Section 3.3. However, for $\boldsymbol{A}$ we must use a different approach.

In the varying-coefficients model, the matrix $\boldsymbol{A}$ of compensated semi-elasticities depends only on x^R and not on prices. Although $x^R = R(\boldsymbol{p}, x)$ is a complicated function of prices and expenditure in general, recall that at the base price vector, R satisfies $R(\overline{\boldsymbol{p}}, x) = x$. Thus, we obtain a consistent estimator for the matrix-valued function by

$$\widehat{\boldsymbol{A}}_0(x^R) = \widehat{\boldsymbol{\Upsilon}}(\overline{\boldsymbol{p}}, x).$$

Here, there is no averaging, so this estimator inherits the slow convergence rate of the nonparametric pre-estimator $\widehat{\boldsymbol{\Upsilon}}$. We note, however, that this slow convergence rate is only for the initial estimator. As discussed above for the fixed coefficients model, iteration may help. Consequently, an iterated predictor for x^R will not (necessarily) inherit this suboptimal rate. Stronger statements would require a technical mathematical analysis of the convexity of the estimation problem. Finally, as in the fixed coefficients case, we can either calculate the symmetry-restricted or the unrestricted estimate for $\boldsymbol{\Upsilon}$ to get $\boldsymbol{A}_0(x^R)$. Then, as in section 3.3, we can derive the convergence rate of the predictors $\hat{x}_i^R$.

4 Inference and Rationality Conditions

4.1 Bootstrap Confidence Bands

Here, we describe how to construct uniform confidence bands around $\hat{\beta}^j$. In the varying-coefficients model, one may construct uniform confidence bands around $\widehat{a^j}$ in the same way. We first define the statistic

$$S_j = \sup_{x^R} |\hat{\beta}^j(x^R) - \beta^j(x^R)| \widehat{\Sigma}_{\beta\ jj}^{-0.5}(x^R)\ ,\ \ j = 1, \ldots, M,$$

where $\widehat{\Sigma}_{\beta\ jj}^{0.5}(x^R)$ is the estimated standard deviation of $\hat{\beta}^j$ at point x^R. Following Härdle, Huet, Mammen, and Sperlich (2004) we determine the distribution of S_j via the wild bootstrap. To this end, we calculate

$$S_j^* = \sup_{x^R} |\check{\beta}^j(x^R) - E^*\left[\check{\beta}^j(x^R)\right]| \check{\Sigma}_{\beta\ jj}^{-0.5}(x^R)\ ,$$

where the $\check{\ }$ indicates estimates from bootstrap samples, and E^* refers to the expectation over the bootstrap sample estimates. The confidence bands are defined by

$$\left[\ \hat{\beta}^j(x^R) - s_j^*\ \check{\Sigma}_{\beta\ jj}^{0.5}(x^R)\ ,\ \ \hat{\beta}^j(x^R) + s_j^*\ \check{\Sigma}_{\beta\ jj}^{0.5}(x^R)\ \right]$$

at each point x^R, where s_j^* is the $(1 - \alpha)$ quantile of S_j^* and $\alpha \in (0, 1)$.

In Section 3, we show how to account for the variance caused by the prediction of the $x_i^R,\ i = 1, \ldots, n$. We show below that, in our empirical example, the variance of the initial prediction does not influence the distribution of the final estimates when iteration is used. In this case, the bootstrap samples can be generated with estimates given the

sample $\{\hat{x}_i^R, \boldsymbol{p}_i\}_{i=1}^n$ for $j = 1, \ldots, M$ by

$$w_i^{j^*} = \hat{\beta}^j(\hat{x}_i^R) + \hat{\boldsymbol{a}}^j \boldsymbol{p}_i + \epsilon_i^{j^*},$$

where the $\epsilon_i^{j^*}$ reflects the covariance structure (between the M different equations) of

the original errors. For the wild bootstrap, we use the Radamacher distribution where

the bootstrap error vector $\boldsymbol{\epsilon}_i^* = \boldsymbol{e}_i \mu_i$ with $\boldsymbol{e}_i$ is the sample residual vector, and μ_i is a

scalar-valued independent random variable satisfying $P[\mu_i = 1] = P[\mu_i = -1] = 0.5$; see

Davidson and Flachaire (2005) for details and particular advantages of this method.

4.2   Maintaining or Testing Homogeneity

The UTL may easily be restricted to satisfy homogeneity by normalizing prices with

respect to p^M. We shall discuss the case for the fixed-coefficients version, but the pro-

cedure is analogous for the varying-coefficients version. In the fixed-coefficients UTL,

homogeneity is satisfied if and only if $\boldsymbol{\iota}^\top \overline{\boldsymbol{\beta}}(u) = 1$ and $\boldsymbol{A}\boldsymbol{\iota} = \boldsymbol{0}_M$. Since only $(M-1)$

independent expenditure share equations are estimated, the summation restriction on $\overline{\boldsymbol{\beta}}$

only affects the calculation of N and R, and does not restrict the estimation.

The linear restriction on the parameters $\boldsymbol{A}$ can be implemented in various ways, nor-

malizing prices being one of them. Denote the kth normalized price as $\widetilde{p}^k = p^k - p^M$ and

let $\widetilde{\boldsymbol{p}} = [\widetilde{p}^1, ..., \widetilde{p}^{M-1}]$, and denote $\widetilde{P}_i^k = P_i^k - P_i^M$ and $\widetilde{\boldsymbol{P}}_i = [\widetilde{P}_i^1, ..., \widetilde{P}_i^{M-1}]$. Denote the

$\widetilde{\boldsymbol{A}}$ as the $(M-1) \times (M-1)$ upper-left submatrix of $\boldsymbol{A}$, and denote $\widetilde{\boldsymbol{\beta}}$ as the $(M-1)$

vector function giving all but the last element of $\boldsymbol{\beta}$, and $\widetilde{\boldsymbol{W}}$ and $\widetilde{\boldsymbol{w}}$ defined analogously

as the first $(M-1)$ expenditure shares and share functions. Log nominal expenditure over normalized prices is

$$N(\widetilde{\boldsymbol{p}}, x) = x + \widetilde{\boldsymbol{p}}^\top \widetilde{\boldsymbol{\beta}}\,(x) + p^M + \frac{1}{2}\widetilde{\boldsymbol{p}}^\top \widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{p}}$$

and log real-expenditure is still its inverse with respect to x at each $\widetilde{\boldsymbol{p}}$. Then, the homogeneity-restricted model analogous to (9) is given by

$$E[\widetilde{\boldsymbol{W}}|X^R = x^R, \widetilde{\boldsymbol{P}} = \widetilde{\boldsymbol{p}}] = \widetilde{\boldsymbol{\beta}}(x^R) + \widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{p}}. \tag{17}$$

This model may be estimated via the techniques outlined above.

In our semiparametric partially-linear model, the Null Hypothesis of homogeneity reduces to

$$H_0: \ \boldsymbol{A}\boldsymbol{\iota} = \boldsymbol{0}_M$$

against the alternative that at least one of these equations does not hold. This null hypothesis can be checked by a Wald test at the parametric rate since the variance-covariance structure is explicitly given in Corollary 1 and Corollary 2. The variance-covariance matrix must be estimated nonparametrically. Although it can be shown that this is theoretically valid, for practical applications one should use the bootstrap or subsampling. If the real data sample is relatively small, these resampling methods should also be used to determine the critical value of the test statistic. For the varying-coefficients case, this test is evaluated at particular values of x^R.

4.3 Maintaining or Testing Symmetry

Symmetry is also easily imposed in our framework. Again, the procedure is analogous between the fixed- and varying-coefficients models, so we discuss the fixed-coefficients case only. The Slutsky matrix $\boldsymbol{S}$ is given by

$$\boldsymbol{S} = \boldsymbol{A} + \boldsymbol{w}\boldsymbol{w}^\top - diag\{\boldsymbol{w}\} \ . \tag{18}$$

Symmetry is satisfied if and only if $\boldsymbol{A}$ is symmetric. To impose symmetry, we apply a linear estimator under linear cross-equation restrictions in the partially model. Thus, for example, in equation (13), we would include the restriction $\boldsymbol{A} = \boldsymbol{A}^\top$ (or, $\widetilde{\boldsymbol{A}} = \widetilde{\boldsymbol{A}}^\top$ for the homogeneity-restricted version).

If we first estimate $\boldsymbol{A}$ without the symmetry restriction, then we may test symmetry via a parametric hypothesis test for

$$H_0: \ a_i^j = a_j^i \ \forall \ i, j = 1, \ldots, M \quad \text{vs} \quad H_1: \ a_i^j \neq a_j^i \text{ for at least one } i \neq j = 1, \ldots, M$$

with $\boldsymbol{A} = \{a_i^j\}_{i,j=1}^M$. Given theorems 1 and 2, we have the covariance matrix of the vector of all differences of interest,

$$\boldsymbol{\alpha} = \left(a_1^2 - a_2^1, a_1^3 - a_3^1, \cdots, a_1^M - a_M^1, a_2^3 - a_3^2, \cdots, a_{M-1}^M - a_M^{M-1}\right)^\top$$

which we denote $\boldsymbol{\Sigma}_\alpha$. Then, our test statistic is $\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_\alpha^{-1} \boldsymbol{\alpha}$ that under H_0 converges to a χ^2-distribution with $\{(M-1)(M-2)\}/2$ degrees of freedom, and under H_1 goes to infinity. This can be written as a Wald-type-test applying the same variance-covariance matrix as in the homogeneity test.

5 Empirical Application

The data used in this paper come from the following public use sources: (1) the Family Expenditure Surveys 1969, 1974, 1978, 1982, 1984, 1986, 1990, 1992 and 1996; (2) the Surveys of Household Spending 1997, 1998 and 1999; and (3) Browning and Thomas (1999), with updates and extensions to rental prices from Pendakur (2001, 2002). Price and expenditure data are available for 12 years in 5 regions (Atlantic, Quebec, Ontario, Prairies and British Columbia) yielding 60 distinct price vectors. Prices are normalized so that the price vector facing residents of Ontario in 1986 is $(1, ..., 1)$.

In Table 1, we present summary statistics for 6,952 observations of rental-tenure unattached individuals aged 25-64 with no dependents. Analysis is restricted to these households to minimize demographic variation in preferences. In the empirical analysis, we used annual expenditure in nine categories: food-in, food-out, rent, clothing, household operation, household furnishing & equipment, private transportation operation, public transportation and personal care. Personal care was the left-out equation, yielding eight expenditure share equations that depend on 9 log-prices and log-expenditure.

These expenditure categories accounted for about three quarters of the current consumption of the households in the sample. Because Canadian expenditure data are at the annual level, the infrequency problem which plagues other data sources (e.g., UK expenditure data) with endogeneity problems is not a major issue here.

All models estimated maintain the restriction of homogeneity, and models used in con-

sumer surplus exercises maintain the additional restriction of Slutsky symmetry. Thus, for the 9-good demand system, the $\boldsymbol{A}$ ($\boldsymbol{A}(x^R)$) matrix (function) is a (9×9) matrix of compensated semi-elasticities with row-sums of zero, and $\boldsymbol{\beta}(x^R)$ is a 9 element vector-function of log real-expenditure which everywhere sums to one. We implemented our models using a predictor of log real-expenditure that uses pre-estimates of $\boldsymbol{A}$ and $\boldsymbol{\beta}$. For the fixed-coefficient model, we computed $\widehat{\boldsymbol{A}}_0$, the pre-estimate of $\boldsymbol{A}$, as the average of the fully nonparametric estimate of symmetry-unrestricted compensated semi-elasticity matrix at each observation. For the varying-coefficients model, we computed $\widehat{\boldsymbol{A}}_0(x^R)$, the pre-estimate of $\boldsymbol{A}(x^R)$, as the nonparametric estimate of the symmetry-unrestricted compensated semi-elasticity matrix at the base price vector $\overline{\boldsymbol{p}} = 0_M$. For both models, we computed $\boldsymbol{\beta}_0$, the pre-estimate of $\boldsymbol{\beta}$, as equal to the Engel curve for observations facing base prices $\overline{\boldsymbol{p}} = 0_M$. Given these initial estimates, we created our generated regressor for each observation, x_i^R.

For the fixed-coefficients model, we used Yatchew's difference estimator of $\boldsymbol{A}$ with 100th order moving-average difference coefficients. Holding the order of the difference constant, our results were essentially unchanged if we use optimal or optimal symmetric difference coefficients instead. Estimates were essentially identical if 50th or 20th order moving average differencing coefficients are used instead. For the nonparametric part, we used a Gaussian kernel with the cross-validated bandwidth of 0.24 for all equations.

For the varying-coefficients model, we used the same bandwidth for the matrix-valued function $\boldsymbol{A}(\boldsymbol{x}^R)$ since symmetry restrictions are cross-equation restrictions. In addition,

symmetry-unrestricted estimates using different cross-validated bandwidths across equations yielded very similar estimates of the matrix $\boldsymbol{A}$ ($\boldsymbol{A}(\boldsymbol{x}^R)$). Our results, including all the results of all tests described below, do not qualitatively change if a bandwidth (common to all equations) 50 percent larger or 25 percent smaller is used instead.

We used subsampling to assess the influence of the pre-estimation step, and considered the homogeneity-restricted but symmetry-unrestricted fixed-coefficients model. We drew 200 subsamples containing 2,000 observations each from the 6,952 observations in the data described above. For each subsample, we created pre-estimates of $\boldsymbol{A}$ and $\boldsymbol{\beta}$ as described above, and used these pre-estimates to estimate the iterated model on the entire sample of 6,952 observations. We iterated the model 6 iterations past the pre-estimation for each of the 200 consistent pre-estimates. We shall discuss only the behaviour of the estimate of $\boldsymbol{A}$ across subsamples. The sum of the 64 variances of the elements of the pre-estimates of $\boldsymbol{A}$ across the subsamples is 0.0129. The sum of the 64 variances of the elements of iterated estimates of $\boldsymbol{A}$ across the subsamples is 0.0000009, which is smaller by a factor of about 15,000. Clearly, the variance of the constructed regressor hardly affects the variance of the iterated estimator. Thus, we conclude that the iterated estimator is efficient and we present only iterated estimates below.

The fixed- and varying-coefficients models differ only in their treatment of compensated price effects. In the varying-coefficients model, these effects may differ over log-real expenditure. The Slutsky symmetry restriction likewise only concerns compensated price effects. Thus, we begin with a discussion of estimated compensated price effects

and of symmetry tests, and then proceed to discuss estimated Engel curve functions.

5.1 Compensated Price Effects and Symmetry

In the following tables and figures, we present estimates for iterated models using 6 iterations past the pre-estimation step. In practise, the estimates numerically converged after about 3 iterations. Simulated standard errors are provided in italics below each estimate. In Table 2, we present the symmetry-unrestricted estimate of $\boldsymbol{A}$.

In Table 3, we present symmetry-restricted estimates of the matrix of compensated price effects. As one might expect, the simulated standard errors for off-diagonal terms are much smaller than those reported in Table 2 because when they are true, the symmetry restrictions are quite informative.

The estimate of $\boldsymbol{A}$ reported in Table 2 does not appear to satisfy symmetry. A Wald test of symmetry based on the joint hypothesis that all off-diagonal terms equal their symmetric partner rejects the hypothesis. The test statistic is $\tau^{SYM} = \boldsymbol{\alpha}^{\top} \boldsymbol{\Sigma}_{\alpha}^{-1} \boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is the sample estimate of the difference between off-diagonal terms equal under symmetry, and $\boldsymbol{\Sigma}_{\alpha}^{-1}$ is their covariance estimated via the bootstrap. The sample value of the test statistic is 424, which is larger than 48, the 1 percent critical value of the χ^2_{28}, so we reject the hypothesis of symmetry in the fixed coefficient model.[4] The failure of symme-

---

[4]One could alternatively bootstrap the entire test statistic to account for possible small-sample bias. However, given the fast convergence in the partially-linear model, it is not surprising that this alternative approach also yields a very strong rejection of symmetry.

try could be due to the presence of unobserved behavioural heterogeneity as noted by Lewbel (2001) and Matskin (2005). Alternatively, it could be due to the restriction that the matrix of compensated price effects $\boldsymbol{A}$ is independent of utility. Below, we argue that this latter possibility may be true.

We estimated the varying-coefficients version of the model as described above. This model can encompass the partially-linear model if $\boldsymbol{A}(x^R)$ is independent of x^R, so it is natural to test whether or not this additional flexibility is necessary. Using symmetry-unrestricted varying-coefficients estimates, we constructed a matrix-valued function of deviations $\widehat{\boldsymbol{A}}_D(x_t^R) = \widehat{\boldsymbol{A}}(x_t^R) - \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{A}}(x_t^R)$ over a grid of T equispaced points in the range of x^R. Since $\widehat{\boldsymbol{A}}$ is asymptotically normal, so too is its deviation from its mean over T points in log-real expenditure. Under the null hypothesis of a fixed-coefficient model, these deviations should be zero. Thus, we constructed the $(M \times M \times T)$-vector $\boldsymbol{\lambda}$ as the vectorisation of $\widehat{\boldsymbol{A}}_D(x_t^R)$ over all T points, and simulated its variance, denoted $\boldsymbol{\Sigma}_\lambda$, under the null that the partially-linear model is true. Our test statistic is then $\tau^{PLM} = \boldsymbol{\lambda}\boldsymbol{\Sigma}_\lambda^{-1}\boldsymbol{\lambda}$ which is distributed asymptotically as a $\chi^2_{M^2(T-1)}$. One could compare the sample value of τ^{PLM} to its asymptotic distribution. However, in practice, because the $\boldsymbol{A}$ matrix converges more slowly in the varying-coefficients model than in the fixed-coefficients model, we accounted for possible sample bias by bootstrapping the entire statistic. Using $T = 9$, the value of our test statistic is 876, and the 1 percent critical value of its simulated distribution under the null is 124. Thus, for these data, we reject the hypothesis that the matrix of compensated price effects $\boldsymbol{A}$ is independent of utility,

and may comfortably use the varying coefficients model.

As noted above, in the fixed-coefficients model, we reject symmetry. We constructed a test statistic for symmetry in the varying-coefficients model analogous to that used in the fixed-coefficients model. In the varying-coefficients model, the vector $\boldsymbol{\alpha}$ depends on x^R, so we denote this vector-function as $\boldsymbol{\alpha}(x^R)$, and evaluate it at a grid of T equispaced points in the range of x^R. Thus, we may construct pointwise tests of symmetry as $\tau_t^{SYM} = \boldsymbol{\alpha}^\top(x_t^R)\boldsymbol{\Sigma}_{\alpha(x_t^R)}^{-1}\boldsymbol{\alpha}(x_t^R)$ for $t = 1, ..., T$. Using sample values of $\boldsymbol{\alpha}^\top(x_t^R)$ and a bootstrap estimate of $\boldsymbol{\Sigma}_{\alpha(x_t^R)}^{-1}$, each of these tests is asymptotically χ_{28}^2 with a 1 percent critical value of 48. However, to account for possible small-sample bias, we bootstrap the entire statistic. Figure 1 shows 90 percent confidence bands for the test statistic under the null of symmetry on the grid of 9 points. For $T = 9$, we find no pointwise rejection of symmetry. That symmetry is rejected in the partially-linear model, but not rejected in the varying-coefficients model, could be due to one of two factors. First, the utility-independence restriction on $\boldsymbol{A}$ could be false, which leads to a false rejection of symmetry in the partially-linear model. Alternatively, it may be that symmetry is false, but the relative imprecision of the estimated $\boldsymbol{A}(x^R)$ in the varying-coefficients model yields a test with low power, so that symmetry is not rejected even though it is false. Below, we consider the latter possibility.

In Table 4, we present estimates of compensated price effects in the symmetry-unrestricted varying-coefficients model, evaluated at the median log real-expenditure level of 9.30. The data are densest near the median, so this is where the varying-coefficients model

is estimated most precisely. Simulated standard errors are given in italics below each estimate. Clearly, the precision of the estimates is lower than in the fixed-coefficients model (Table 2). Most elements of $\boldsymbol{A}(x^R)$ are estimated with about half the precision of the corresponding estimates in the partially-linear model, and some with much less precision. This suggests that the non-rejection of symmetry in the varying-coefficients model may be due to the imprecision of the estimated coefficients: we should treat this non-rejection of symmetry with caution.

Consumer surplus calculations are unique only if estimated demands satisfy symmetry. We take the non-rejection of symmetry in the varying-coefficients model as licence to use symmetry-restricted varying-coefficients estimates in a consumer-surplus exercise below. In Table 5, we present estimates of compensated price effects for the symmetry-restricted varying-coefficients model evaluated at the median of log-real expenditure.

The differences between the models are most easily seen graphically. In Figures 2 to 4, we depict estimated values of selected elements of $\widehat{\boldsymbol{A}}(x^R)$ at 39 equispaced points in the range of log-real expenditure. The displayed elements correspond to own-price effects for food-in and rent, and the cross-price effects of food-out on food-in and vice-versa. In each figure, black and grey lines indicate varying- and fixed-coefficients estimates, respectively. Quadratic almost-ideal (QAI) estimates (see, e.g., Banks, Blundell, and Lewbel, 1997) are presented with dark dotted lines. Simulated 90 percent uniform confidence bands for the symmetry-restricted varying-coefficients estimates are indicated with crosses at 9 equispaced points in the range of log real-expenditure.

The median and average of log real-expenditure are 9.30 and 9.27. In this part of the distribution, it is clear that the fixed-coefficients model does fairly well in capturing the compensated price effects. In addition, the fixed-coefficients model gives estimates of compensated price effects very similar to those of the quadratic almost-ideal model. This is because although the QAI model has compensated semi-elasticities that depend on expenditure, there is (typically) only one matrix of parameters governing price effects, so that QAI price effects are not flexible over expenditure.

Over much of the middle of the distribution, the compensated price effects shown in Figures 2 and 3 essentially overlap under the fixed- and varying-coefficients models. However, even in the middle of the distribution, one can see in Figure 4 that the estimated rent compensated own-price effect is poorly approximated by the fixed coefficients and QAI models. The fixed-coefficients model estimates are too low, and lie outside the uniform confidence band of the varying-coefficients model estimates.

The fixed-coefficients model performs poorly far from the middle of the distribution of log real-expenditure. For example, the food-in compensated own-price effect is large and positive at the bottom of the distribution, but small and negative throughout the middle of the distribution. This means that although middle-income individuals are able to substitute away from food when its price rises, poorer individuals are not able to do so. Thus, use of the fixed-coefficients model would bias welfare analysis in potentially important ways, as we shall show below.

5.2 Engel Curve Functions

In Figures 5 to 12, we depict estimated Engel curve functions for the estimated models Since $\boldsymbol{w}(\boldsymbol{p}, x^R) = \boldsymbol{\beta}\left(x^R\right)$ at the base price vector, the Engel curves $\boldsymbol{\beta}\left(x^R\right)$ give the estimated expenditure share at the base price vector. Expenditure shares are evaluated at 39 equally-spaced points over the middle 95 percent of the implied log real-expenditure distribution. In each figure, black and grey lines indicate varying- and fixed-coefficients estimates, respectively. Thick and thin lines indicate symmetry-restricted and unrestricted estimates, respectively. Simulated 90 percent uniform confidence bands for the symmetry-restricted varying-coefficients estimates at 9 equally-spaced points are shown with crosses. QAI estimates are shown with dark dotted lines.

Figures 5 to 12 show estimated shares for food at home, food out, rent, household operation, household furnishing/equipment, clothing, private transportation operation and public transportation, respectively. The left-out expenditure share is personal care.

The expenditure-share equations for food-in and food-out are roughly linear, as is found in nonparametric investigations of the shape of Engel curves; see, e.g., Banks, Blundell, and Lewbel (1997). Not surprisingly, all models have roughly the same estimated Engel curve functions for these almost linear expenditure-share equations. Rent shares are roughly 'U-shaped' as found in previous work, and there is some evidence of rank greater than 2 at the extremes of the expenditure distribution. For rent shares, the QAI model performs poorly. For example, the estimated rent share at the bottom decile cutoff of $x^R = 8.6$ is almost 2 percentage points higher given the QAI model than given the

38

varying-coefficients model. In this part of the distribution, the fixed-coefficients model also performs relatively poorly, driven in large measure by the falseness of the fixed-coefficients assumption at the bottom of the distribution, as seen in Figure 3.

Some expenditure share equations appear to be 'S-shaped' as noted in previous work on Engel curves; see Blundell, Chen, and Kristensen (2003). The curvature of the private transportation operation expenditure-share equation varies greatly over expenditure and suggests rank greater than 2. In particular, expenditure shares are nearly flat for the bottom quintile of the population, rising steeply through the middle of the distribution, and falling for the top quintile. The complexity of this Engel curve is difficult to capture in a quadratic specification. For this reason, the QAI estimate of the private transportation share is fairly distant from both the varying- and fixed-coefficients estimates throughout the distribution of expenditure.

5.3 Cost-of-Living Experiment

We assess the economic significance of our models with a cost-of-living experiment. In Canada, rent is not subject to sales taxes, which can amount to 15 percent for goods such as food-out and clothing. Consider the cost-of-living index associated with subjecting rent to a 15 percent sales tax for people facing the base price vector. The log cost-of-living index for this change is given by $\{N(\boldsymbol{p}, x) - x\}$, where N is the nominal expenditure function, and $\boldsymbol{p} = [0, 0, \log(1.15), 0, ..., 0]^{\top}$, the new price vector. Using estimated values,

this is

$$\boldsymbol{p}^{\top}\widehat{\boldsymbol{\beta}}\left(x\right) + \frac{1}{2}\boldsymbol{p}^{\top}\widehat{\boldsymbol{A}}\boldsymbol{p} = = \widehat{\beta}^{rent}(x)\log(1.15) + \frac{\widehat{a}^{rent,rent}}{2}\log(1.15)^2$$

for the fixed-coefficients model and

$$\boldsymbol{p}^{\top}\widehat{\boldsymbol{\beta}}\left(x\right) + \frac{1}{2}\boldsymbol{p}^{\top}\widehat{\boldsymbol{A}}(\boldsymbol{x})\boldsymbol{p} = \widehat{\beta}^{rent}(x)\log(1.15) + \frac{\widehat{a}^{rent,rent}(x)}{2}\log(1.15)^2$$

for the varying-coefficients model.

In Figure 13, we illustrate how the cost-of-living index varies over expenditure for this hypothetical price change given estimates from symmetry-restricted fixed- and varying-coefficients models, as well as estimates from the (symmetry-restricted) QAI model. Here, neither the fixed-coefficients model nor the QAI perform very well in approximating the estimated cost-of-living impact indicated by the varying-coefficients estimates. In the lower part of the distribution, the first two models overstate the cost-of-living impact. For example, at the bottom decile of the real expenditure distribution ($x^R = 8.6$), the QAI and fixed-coefficients estimates of the cost-of-living impact are 7.4 percent and 7.0 percent, respectively, but the varying-coefficients estimate is 6.4 percent. In the upper part of the distribution, the fixed coefficients model performs better, but the QAI again overstates the cost-of-living impact of the price increase. The reason for these patterns can be seen in Figures 3 and 6. Both the fixed-coefficients and the QAI models have inflexible compensated price effects, and Figure 3 suggests that this inflexibility is most costly at the bottom of the distribution of real-expenditure, which is where both models perform poorly. In addition, the QAI faces the restriction that Engel curves are quadratic, which results in a poor fit in comparison to the nonparametric Engel curve

functions at both ends of the distribution.

We conclude from this investigation that our approach yields insights about the shape of expenditure-share equations that may be hard to see in a 'traditional' Engel-curve by Engel-curve nonparametric regression approach. In particular, our approach allows the investigator to estimate a complete demand system wherein expenditure-share equations may be arbitrarily complex in their relationship with real-expenditure. Further, in the varying-coefficients version of our model, the investigator may include price effects that vary with real-expenditure. Our empirical work suggests that the varying-coefficients extension is both statistically significant and economically important.

6 Appendix

Proof of Theorem 3:

As the local-linear estimator has been well studied – see Fan and Gijbels (1996) – we show here only how the bias and variance terms change when a nonparametrically generated regressor is used. Because the matrix $\boldsymbol{A}$ is estimated with the parametric rate, it is clear that the randomness caused by its estimation can be neglected when looking at the asymptotics of our nonparametric estimator. Furthermore, it will be seen in the proof that it is sufficient to do the explicit calculations for only one of the M equations.

For notational parsimony, we let Y_i denote $W_i^j - \boldsymbol{a}^j \boldsymbol{P}_i$ and b denote β^j, $\epsilon(X_i, \boldsymbol{P}_i)$ denote ϵ_i^j with variance function $\sigma_\epsilon^2(X_i, \boldsymbol{P}_i)$ for an arbitrary $j = 1, \ldots, M$. Further, we write B_X and σ_X^2 as functions of x^R and recall assumption [X1]. Then, for ξ_i between x^R and

41

X_i^R we have

$$Y_i = b(x^R) + b'(x^R)(X_i^R - x^R) + \frac{b''(\xi_i)}{2}(X_i^R - x^R)^2 + \epsilon(x_i, \boldsymbol{p}_i) . \tag{19}$$

The estimator for b (and b') in x^R is defined by

$$\begin{pmatrix} \hat{b}(x^R) \\ \hat{b}'(x^R) \end{pmatrix} = \begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \tag{20}$$

where

$$\begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix} = \left\{ \begin{pmatrix} 1 & \hat{X}_1^R - x^R \\ \vdots & \vdots \\ 1 & \hat{X}_n^R - x^R \end{pmatrix}^\top diag\left\{ K_h(\hat{X}_i^R - x^R) \right\}_{i=1}^n \begin{pmatrix} 1 & \hat{X}_1^R - x^R \\ \vdots & \vdots \\ 1 & \hat{X}_n^R - x^R \end{pmatrix} \right\}^{-1}$$

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1 & \hat{X}_1^R - x^R \\ \vdots & \vdots \\ 1 & \hat{X}_n^R - x^R \end{pmatrix}^\top diag\left\{ K_h(\hat{X}_i^R - x^R) \right\}_{i=1}^n \boldsymbol{Y} .$$

Combining (19) with (20) we see that for calculating bias it is sufficient to consider

$$\begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} \tag{21}$$

with vector $(M_1, M_2) =$

$$\begin{pmatrix} 1 & \hat{X}_1^R - x^R \\ \vdots & \vdots \\ 1 & \hat{X}_n^R - x^R \end{pmatrix}^\top diag\left\{ K_h(\hat{X}_i^R - x^R) \right\}_{i=1}^n \begin{pmatrix} b'(x^R)(X_1^R - x^R) + b''(\xi_1)\frac{(X_1^R - x^R)^2}{2} \\ \vdots \\ b'(x^R)(X_n^R - x^R) + b''(\xi_n)\frac{(X_n^R - x^R)^2}{2} \end{pmatrix} .$$

We first calculate the inverse matrix in equation (20)

$$
\begin{aligned}
N_{11} &= \sum_{i=1}^{n} K_h(\hat{X}_i^R - x^R) = n \int \left\{ \frac{B_X(v)}{h^2} K' \left(\frac{v - x^R}{h} \right) + o(h) \right\} f(v)dv + nf(x^R) \\
&= nf(x^R) + o(n)
\end{aligned}
$$

due to the rate assumptions on B_X, σ_X^2 and because K' integrates to zero. Further,

$$
\begin{aligned}
N_{12} &= \sum_{i=1}^{n} (\hat{X}_i^R - x^R) \, K_h(\hat{X}_i^R - x^R) = n \int \left\{ hv - B_X(x^R + hv) \right\} K(v) f(x^R + hv) dv \\
&\quad + n \int \left\{ hv - B_X(x^R + hv) \right\} \frac{B_X(x^R vh)}{h} K'(v) f(x^R + hv) dv + o \left\{ nB_X(x^R) \right\} \\
&= nB_X(x^R) f(x^R) \{ \mu_1(K') - 1 \} + o \left\{ nB_X(x^R) \right\} \\
N_{22} &= \sum_{i=1}^{n} (\hat{X}_i^R - x^R)^2 \, K_h(\hat{X}_i^R - x^R) \\
&= n \int \left\{ hv - B_X(x^R) \right\}^2 \left[K(v) + \frac{B_X(x^R)}{h} K'(v) + o \left\{ \frac{B_X(x^R)}{h} \right\} \right] f(x^R + hv) dv \\
&= n\mu_2(K) h^2 f(x^R) + no(h^2)
\end{aligned}
$$

For the vector (M_1, M_2) we have basically to repeat calculations as we have done for N_{11} (when calculating M_1) and N_{12} (when considering M_2) and get.

$$
\begin{aligned}
M_1 &= nf(x^R) \left\{ -b'(x^R) B_X(x^R) + h^2 \mu_2(K) \frac{b''(x^R)}{2} \right\} + o \left(nh^2 \right) \\
M_2 &= nb'(x^R) B_X^2(x^R) f(x^R) \{ 1 - \mu_1(K') \} \, .
\end{aligned}
$$

Putting this into (21) yields the bias stated in the theorem.

For the variance one has to consider the expectation of

$$
\begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \begin{pmatrix} M_1' \\ M_2' \end{pmatrix} \begin{pmatrix} M_1' \\ M_2' \end{pmatrix}^{\top} \begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \tag{22}
$$

43

with vector

$$
(M'_1, M'_2) = \begin{pmatrix} 1 & \hat{X}^R_1 - x^R \\ \vdots & \vdots \\ 1 & \hat{X}^R_n - x^R \end{pmatrix}^\top \, diag\left\{K_h(\hat{X}^R_i - x^R)\right\}^n_{i=1} \times
$$

$$
\begin{pmatrix} b'(x^R)(X^R_1 - x^R) + b''(\xi_1)\frac{(X^R_1 - x^R)^2}{2} + \epsilon_1 \\ \vdots \\ b'(x^R)(X^R_n - x^R) + b''(\xi_n)\frac{(X^R_n - x^R)^2}{2} + \epsilon_n \end{pmatrix}.
$$

Looking at the last vector in the definition of (M'_1, M'_2) it is clear that the matrix $(M'_1, M'_2)^\top (M'_1, M'_2)$ can be decomposed additively in four symmetric matrices, one containing only the b' (denoted by C^1), one only containing the b'' (denoted by C^2), one with with both (denoted by C^3), and one with only the error terms ϵ_i (denoted by C^4). Under expectation, the other terms either vanish or are obviously of higher order.

We start with C^1. For some ξ_i between X^R_i and x^R we get:

$$
E[C^1_{11}] = E\left[\sum_{j=1}^n \sum_{i=1}^n b'^2(x^R)\{B_X(X^R_i) + u_i\}\{B_X(X^R_j) + u_j\}\epsilon_i\epsilon_j \right.
$$

$$
\left. \left\{K_h(X^R_i - x^R) + \frac{B_X(X^R_i) + u_i}{h}K'_h(\xi_i)\right\}\left\{K_h(X^R_j - x^R) + \frac{B_X(X^R_j) + u_j}{h}K'_h(\xi_i)\right\}\right]
$$

$$
= n^2 f^2(x^R)B^2_X(x^R)b'^2(X^R_0) + o\{n^2 B^2_X(x^R)\} \, ,
$$

where $\nu_k = \int v^k K^2(v)dv$. Further,

$$
\begin{aligned}
E[C_{12}^1] &= E\left[\sum_{j=1}^n \sum_{i=1}^n b'^2(x^R)\{B_X(X_i^R)+u_i\}\{B_X(X_j^R)+u_j\}\epsilon_i\epsilon_j \right. \\
&\quad \{X_i^R - x^R + B_X(X_i^R)+u_i\}\left\{K_h(X_i^R - x^R) + \frac{B_X(X_i^R)+u_i}{h}K_h'(\xi_i)\right\} \\
&\quad \left.\left\{K_h(X_j^R - x^R) + \frac{B_X(X_j^R)+u_j}{h}K_h'(\xi_i)\right\}\right] \\
&= n^2 f^2(x^R)B_X^3(x^R)b'^2(X_0^R) + o\{n^2 B_X^3(x^R)\} \\
E[C_{22}^1] &= n^2 f^2(x^R)B_X^4(x^R)b'^2(X_0^R) + o\{n^2 B_X^4(x^R)\} \ .
\end{aligned}
$$

Similarly, for some ζ_i between X_i^R and x^R we have:

$$
\begin{aligned}
E[C_{11}^2] &= E\left[\sum_{j=1}^n \sum_{i=1}^n \frac{b''(\zeta_i)}{2}(X_i^R - x^R)^2 \frac{b''(\zeta_j)}{2}(X_j^R - x^R)^2 \left\{K_h(X_i^R - x^R)+ \right.\right. \\
&\quad \left.\left. \frac{B_X(X_i^R)+u_i}{h}K_h'(\xi_i)\right\}\left\{K_h(X_j^R - x^R) + \frac{B_X(X_j^R)+u_j}{h}K_h'(\xi_i)\right\}\right] \\
&= \frac{n^2 h^4}{4}f^2(x^R)b''^2(X_0^R)\mu_2^2(K) + o(n^2 h^4) \\
E[C_{12}^2] &= \frac{n^2 h^5}{4}f^2(x^R)b''^2(x^R)\mu_2(K)\mu_3(K) + o(n^2 h^5) \\
E[C_{22}^2] &= \frac{n^2 h^6}{4}f^2(x^R)b''^2(x^R)\mu_3^2(K) + o(n^2 h^6) \ .
\end{aligned}
$$

Next, considering the mixture of b' and b'' we have

$$
\begin{aligned}
E[C_{11}^3] &= E\left[\sum_{j=1}^n \sum_{i=1}^n b'(x^R)\{B_X(X_i^R)+u_i\}\frac{b''(\zeta_j)}{2}(X_j^R - x^R)^2 \left\{K_h(X_i^R - x^R)+ \right.\right. \\
&\quad \left.\left. \frac{B_X(X_i^R)+u_i}{h}K_h'(\xi_i)\right\}\left\{K_h(X_j^R - x^R) + \frac{B_X(X_j^R)+u_j}{h}K_h'(\xi_i)\right\}\right] \\
&= \frac{n^2 h^2}{2}f^2(x^R)b''(X_0^R)b'(X_0^R)\mu_2(K)\{-B_X(x^R)\} + o\left\{n^2 h^2 B_X(x^R)\right\} \\
E[C_{12}^3] &= \frac{n^2 h^3}{2}f^2(x^R)b''(X_0^R)b'(X_0^R)\mu_3(K)\{-B_X(x^R)\} + o\left\{n^2 h^3 B_X(x^R)\right\} \\
E[C_{22}^3] &= \frac{n^2 h^4}{2}f^2(x^R)b''(X_0^R)b'(X_0^R)\mu_4(K)\{-B_X(x^R)\} + o\left\{n^2 h^4 B_X(x^R)\right\} \ .
\end{aligned}
$$

45

Finally, for C^4 we have

$$
\begin{aligned}
E[C_{11}^4] &= \frac{n}{h} f(x^R) \sigma_\epsilon^2(x^R) \nu_0 + o(\frac{n}{h}) \\
E[C_{12}^4] &= \frac{n}{h} f(x^R) \sigma_\epsilon^2(x^R) \nu_1 h + o(n) \\
E[C_{22}^4] &= \frac{n}{h} f(x^R) \sigma_\epsilon^2(x^R) \nu_2 h^2 + o(nh) .
\end{aligned}
$$

For more details of the calculations see Sperlich (2008).

Plugging this results in equation (22) yields the variance we have stated in the theorem.

□

Proof of Theorem 4:

Also, the local-linear varying-coefficients estimator has been well studied; see, for example, Cleveland, Grosse, and Shyu (1991) as well as Fan and Zhang (1999) or Cai, Fan, and Li (2000). The calculations to incorporate the additional bias and variance arising because of the use of a generated regressor are basically the same as for Theorem 3. Note that, now, skipping the index $j = 1, \ldots, M$ of W and of the functions a_k , $k = 0, \ldots, M$ for the ease of notation,

$$
W_i = \sum_{k=0}^{M} P_i^k \left\{ a_k(x^R) + a_k'(x^R)(X_i^R - x^R) + \frac{a_k''(\xi_i)}{2}(X_i^R - x^R)^2 \right\} + \epsilon(x_i, \boldsymbol{p}_i) .
$$

The estimator of a_k in x^R is defined then by the $(2k+1)^{\text{th}}$ element of

$$
\left(\boldsymbol{R}^\top \boldsymbol{K} \boldsymbol{R} \right)^{-1} \boldsymbol{R}^\top \boldsymbol{K} \boldsymbol{W}
$$

where

$$
\boldsymbol{R} = \begin{pmatrix} P_1^0 & P_1^0(\hat{X}_1^R - x^R) & \cdots & P_1^M & P_1^M(\hat{X}_1^R - x^R) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P_n^0 & P_n^0(\hat{X}_n^R - x^R) & \cdots & P_n^M & P_n^M(\hat{X}_n^R - x^R) \end{pmatrix}
$$

$$
\boldsymbol{K} = diag\left\{ K_h(\hat{X}_i^R - x^R) \right\}_{i=1}^n .
$$

It is clear that this leads to the same equations as in proof of Theorem 3 but now always with $P_i^k P_j^l$, $k, l = 0, \ldots, M$, $i, j = 1, \ldots, n$ inside the (double) sums. Taking the expectation with respect to $X^R = x^R$ this leads to the elements of matrix Ω^{-1} which cancel in the bias but not for the variance; see Theorem 3 of Fan and Zhang (1999). For more details of the matrix calculations we also refer to their paper. □

References

Banks J, Blundell R, Lewbel A. 1997. "Quadratic Engel Curves and Consumer Demand", *Review of Economics and Statistics* **79**: 527-539.

Blundell R, Chen X, Kristensen D. 2003. Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves. *Working Paper CWP15/03, Centre for Microdata Methods and Practice.*

Browning M, Thomas I. 1999. Prices for the FAMEX. *Working Paper, Department of Economics, McMaster University.*

Cai Z, Fan J, Li R. 2000. Efficient Estimation and Inference for Varying-Coefficient Models. *Journal of the American Statistical Association* **95**: 888-902.

Christensen LR, Jorgenson DW, Lau LJ. 1971. "Conjugate Duality and the Transcendental Logarithmic Production Function. *Econometrica* **39**: 255-256.

Cleveland WS, Grosse E, Shyu WM. 1991. Local Regression Models. In *Statistical Models in S*, Chambers JM, Hastie TJ (eds). Wadsworth/Brooks-Cole, Pacific Grove, CA: 309-376.

Crossley T, Pendakur K. 2006. The Social Cost-of-Living: Welfare Foundations and Estimation. *IFS Working Papers, W06/10.*

Davidson R, Flachaire E. 2005. The Wild Bootstrap, Tamed at Last. *Working Paper, www.econ.queensu.ca/working_papers/papers/qed_wp_1000.pdf*

Fan J, Gijbels I. 1996. *Local Polynomial Modelling and Its Applications*, Chapman and Hall: London.

Fan J, Zhang W. 1999. Statistical Estimation in Varying Coefficient Models. *Annals of Statistics* **27**: 1491-1518.

Haag B, Hoderlein S, Pendakur K. 2005. Testing and Imposing Rationality Restrictions in the Nonparametric Estimation of Consumer Demand. *Working Paper, Simon Fraser University, Department of Economics.*

Härdle W, Huet S, Mammen E, Sperlich S. 2004. Bootstrap Inference in Semiparametric Generalized Additive Models. *Econometric Theory* **20**: 265-300.

Hall P, Kay JW, Titterington DM. 1990. Asymptotically Optimal Difference-based Estimation of Variance in Nonparametric Regression. *Biometrika* **77**: 521-528.

Hengartner N, Sperlich S. 2005. Rate Optimal Estimation with the Integration Method in the Presence of Many Covariates. *Journal of Multivariate Analysis* **95**: 246-272.

Hoderlein S. 2005. Slutsky Symmetry and Heterogeneity. *Working Paper, University of Mannheim.*

Lewbel A. 1991. The Rank of Demand Systems: Theory and Nonparametric Estimation. *Econometrica* **59**: 711-730.

Lewbel A. 2001. Demand Systems With and Without Error Terms. *American Economic Review* **91**: 611-618.

Lewbel A, Pendakur K. 2008. Tricks with Hicks: The EASI Demand System. *American Economic Review*, forthcoming.

McLaren KR, Rossiter P, Powell AA. 2000. Using the Cost Function to Generate Flexible Marshallian Demand Systems. *Empirical Economics* **25**: 209-227.

Pendakur K. 2001. Consumption Poverty in Canada 1969-1998. *Canadian Public Policy* **27**: 1-34.

Pendakur K. 2002. Taking Prices Seriously in the Measurement of Inequality. *Journal of Public Economics* **86**: 47-69.

Rodríguez-Póo JM, Sperlich S, Fernández AI. 2005. Semiparametric Three Step Estimation Methods for Simultaneous Equation Systems, *Journal of Applied Econometrics* **20**: 699-721.

Sarmiento C. 2005. A Varying Coefficient Approach to Global Flexibility in Demand Analysis: A Semiparametric Approximation. *American Journal of Agricultural Economics* **87**: 38-47.

Speckman P. 1988. Kernel Smoothing in Partial Linear Models. *J.Royal Statistical Society, Ser.B.* **50**: 413-436.

Sperlich S. 2008. A Note on Nonparametric Estimation with Constructed Variables and Generated Regressors. *Preprints ZfS, Centre for Statistics, Georg August Universität Göttingen, Germany.*

Wong GKK, McLaren K. 2005. Specification and Estimation of Regular Inverse Demand Systems: A Distance Function Approach. *American Journal of Agricultural Economics*, **87**: 823-834.

Yatchew A. 1997. An Elementary Estimator of the Partial Linear Model. *Economics Letters* **57**: 135-143.

		Min	Max	Mean	Std Dev
Table 1: The Data					
expenditure shares	food-in	0.02	0.62	0.17	0.09
	food-out	0.00	0.64	0.08	0.08
	rent	0.01	0.95	0.40	0.13
	clothing	0.00	0.53	0.09	0.06
	operation	0.01	0.63	0.08	0.05
	furnish/equip	0.00	0.65	0.04	0.06
	private trans	0.00	0.59	0.08	0.09
	public trans	0.00	0.35	0.04	0.04
log-expenditure		6.68	10.95	9.16	0.60
log-prices	food-in	-1.41	0.34	0.13	0.45
	food-out	-1.46	0.53	0.26	0.51
	rent	-1.32	0.37	-0.03	0.42
	clothing	-0.87	0.43	0.23	0.33
	operation	-1.40	0.32	0.12	0.46
	furnish/equip	-0.94	0.20	0.13	0.32
	private trans	-1.53	0.53	0.01	0.52
	public trans	-1.14	0.69	0.14	0.63

Table 2: Estimated Compensated Price Effects, $\widehat{\boldsymbol{A}}$

	food-in	food-out	rent	clothing	hh oper	furn/equ	priv tr	pub tr
food-in	-0.035	-0.009	-0.088	-0.040	-0.037	0.032	0.058	-0.064
	0.026	*0.020*	*0.007*	*0.027*	*0.027*	*0.017*	*0.008*	*0.009*
food-out	0.054	-0.022	0.018	-0.008	0.061	-0.053	0.008	-0.026
	0.026	*0.020*	*0.007*	*0.028*	*0.025*	*0.016*	*0.008*	*0.009*
rent	-0.073	0.107	0.100	0.075	-0.104	0.036	-0.054	0.080
	0.038	*0.029*	*0.010*	*0.039*	*0.039*	*0.023*	*0.011*	*0.013*
clothing	0.001	0.013	-0.002	0.044	-0.023	0.010	-0.028	0.022
	0.015	*0.012*	*0.004*	*0.016*	*0.016*	*0.008*	*0.004*	*0.005*
hh oper	0.055	-0.047	0.002	-0.023	-0.048	0.007	-0.002	-0.018
	0.019	*0.014*	*0.005*	*0.019*	*0.020*	*0.012*	*0.006*	*0.007*
furn/equ	0.023	-0.085	-0.016	-0.054	0.086	-0.005	0.029	0.000
	0.020	*0.014*	*0.005*	*0.021*	*0.020*	*0.012*	*0.006*	*0.007*
priv tr	-0.018	0.075	-0.032	0.027	0.003	-0.003	-0.013	0.002
	0.029	*0.023*	*0.008*	*0.030*	*0.030*	*0.018*	*0.008*	*0.009*
pub tr	-0.026	-0.023	0.029	0.007	0.017	0.011	0.002	0.003
	0.015	*0.011*	*0.004*	*0.015*	*0.016*	*0.009*	*0.004*	*0.005*

Table 3: Estimated Symmetry-Restricted Compensated Price Effects

	food-in	food-out	rent	clothing	hh oper	furn/equ	priv tr	pub tr
food-in	-0.073	0.053	-0.088	0.002	0.057	0.019	0.039	-0.043
	0.020	*0.013*	*0.007*	*0.011*	*0.010*	*0.009*	*0.007*	*0.006*
food-out		-0.050	0.059	0.020	-0.011	-0.052	0.012	-0.031
		0.014	*0.006*	*0.008*	*0.011*	*0.008*	*0.005*	*0.005*
rent			0.071	-0.001	-0.028	-0.027	-0.027	0.054
			0.010	*0.004*	*0.005*	*0.005*	*0.006*	*0.003*
clothing				0.038	-0.020	0.001	-0.031	0.026
				0.012	*0.009*	*0.006*	*0.004*	*0.004*
hh oper					-0.051	0.035	-0.012	-0.012
					0.014	*0.009*	*0.005*	*0.005*
furn/equ						0.039	0.008	0.000
						0.007	*0.005*	*0.013*
priv tr							0.012	0.003
							0.008	*0.003*
pub tr								-0.001
								0.003

Table 4: Compensated Price Effects, VCM, median x^R

	food-in	food-out	rent	clothing	hh oper	furn/equ	priv tr	pub tr
food-in	-0.018	-0.082	-0.014	-0.071	0.049	0.032	-0.046	-0.018
	0.030	*0.011*	*0.042*	*0.041*	*0.025*	*0.012*	*0.015*	*0.030*
food-out	-0.013	0.007	0.003	0.059	-0.040	0.014	-0.031	-0.013
	0.027	*0.009*	*0.034*	*0.029*	*0.019*	*0.011*	*0.011*	*0.027*
rent	0.099	0.138	0.026	-0.058	-0.013	-0.057	0.090	0.099
	0.040	*0.013*	*0.049*	*0.049*	*0.029*	*0.016*	*0.016*	*0.040*
clothing	0.011	-0.007	0.032	-0.013	0.002	-0.026	0.022	0.011
	0.015	*0.005*	*0.016*	*0.018*	*0.010*	*0.005*	*0.006*	*0.015*
hh oper	-0.042	-0.003	-0.027	-0.042	0.002	-0.002	-0.023	-0.042
	0.021	*0.007*	*0.027*	*0.025*	*0.015*	*0.006*	*0.009*	*0.021*
furn/equ	-0.089	-0.022	-0.054	0.103	0.001	0.035	0.002	-0.089
	0.020	*0.007*	*0.028*	*0.027*	*0.017*	*0.008*	*0.009*	*0.020*
priv tr	0.076	-0.044	0.053	-0.057	0.015	-0.006	-0.014	0.076
	0.036	*0.013*	*0.041*	*0.045*	*0.026*	*0.012*	*0.014*	*0.036*
pub tr	-0.016	0.030	0.013	0.032	0.015	0.003	-0.001	-0.016
	0.015	*0.006*	*0.020*	*0.021*	*0.013*	*0.006*	*0.006*	*0.015*

Table 4: Compensated Price Effects, VCM, median x^R

	food-in	food-out	rent	clothing	hh oper	furn/equ	priv tr	pub tr
food-in	-0.018	-0.082	-0.014	-0.071	0.049	0.032	-0.046	-0.018
	0.030	*0.011*	*0.042*	*0.041*	*0.025*	*0.012*	*0.015*	*0.030*
food-out	-0.013	0.007	0.003	0.059	-0.040	0.014	-0.031	-0.013
	0.027	*0.009*	*0.034*	*0.029*	*0.019*	*0.011*	*0.011*	*0.027*
rent	0.099	0.138	0.026	-0.058	-0.013	-0.057	0.090	0.099
	0.040	*0.013*	*0.049*	*0.049*	*0.029*	*0.016*	*0.016*	*0.040*
clothing	0.011	-0.007	0.032	-0.013	0.002	-0.026	0.022	0.011
	0.015	*0.005*	*0.016*	*0.018*	*0.010*	*0.005*	*0.006*	*0.015*
hh oper	-0.042	-0.003	-0.027	-0.042	0.002	-0.002	-0.023	-0.042
	0.021	*0.007*	*0.027*	*0.025*	*0.015*	*0.006*	*0.009*	*0.021*
furn/equ	-0.089	-0.022	-0.054	0.103	0.001	0.035	0.002	-0.089
	0.020	*0.007*	*0.028*	*0.027*	*0.017*	*0.008*	*0.009*	*0.020*
priv tr	0.076	-0.044	0.053	-0.057	0.015	-0.006	-0.014	0.076
	0.036	*0.013*	*0.041*	*0.045*	*0.026*	*0.012*	*0.014*	*0.036*
pub tr	-0.016	0.030	0.013	0.032	0.015	0.003	-0.001	-0.016
	0.015	*0.006*	*0.020*	*0.021*	*0.013*	*0.006*	*0.006*	*0.015*

Table 5: Symmetry-Restricted Compensated Price Effects, VCM, median x^R

	food-in	food-out	rent	clothing	hh oper	furn/equ	priv tr	pub tr
food-in	-0.081	0.059	-0.088	0.018	0.040	0.024	0.023	-0.033
	0.018	*0.016*	*0.009*	*0.014*	*0.013*	*0.011*	*0.008*	*0.007*
food-out		-0.045	0.049	0.016	-0.002	-0.058	0.016	-0.035
		0.019	*0.008*	*0.010*	*0.011*	*0.010*	*0.008*	*0.007*
rent			0.106	-0.004	-0.028	-0.033	-0.032	0.055
			0.013	*0.004*	*0.005*	*0.005*	*0.009*	*0.005*
clothing				0.023	-0.010	-0.006	-0.027	0.026
				0.014	*0.012*	*0.008*	*0.005*	*0.004*
hh oper					-0.045	0.036	-0.015	-0.015
					0.017	*0.010*	*0.006*	*0.006*
furn/equ						0.043	0.012	0.002
						0.008	*0.005*	*0.005*
priv tr							0.021	0.003
							0.013	*0.005*
pub tr								-0.006
								0.004

Figure 1: Pointwise Tests of Symmetry in the Varying Coefficients Model, simulated 90% confidence intervals

- test statistic
- 95% crit
- 5% crit

value

log real-expenditure

Figure 2: Compensated Semi-elasticities: Food-ln own-price effect

value

log real-expenditure

— symmetry-restricted VCM — symmetry-restricted PLM
-- QAI × 90% uniform confidence bands

57

Figure 3: Compensated Semi-elasticities: Food-In, Food-out cross-price effect

Figure 4: Compensated Semi-elasticities: Rent own-price effect

Figure 5: Estimated Food-in Shares

Figure 6: Estimated Food-Out Shares

symmetry-restricted varying-coefficient   symmetry-restricted fixed-coefficient
unrestricted fixed-coefficient   ▪ ▪ ▪ QAI
×   90% uniform confidence bands

Figure 7: Estimated Rent Shares

Figure 8: Estimated Household Operation Shares

Figure 9: Estimated Household Furnish/Equip Shares

Figure 10: Estimated Clothing Shares

Figure 11: Estimated Private Transportation Shares

share

log real-expenditure

— symmetry-restricted varying-coefficient   — symmetry-restricted fixed-coefficient
— unrestricted fixed-coefficient   ···· car oper
× 90% uniform confidence bands

Figure 12: Estimated Public Transportation Shares

share

log real-expenditure

— symmetry-restricted varying-coefficient   — symmetry-restricted fixed-coefficient
— unrestricted fixed-coefficient   ···· pub trans
× 90% uniform confidence bands

Figure 13: Cost-of-Living Change: 15% Rent increase