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

Consumer demand microdata typically exhibit a great deal of expenditure variation but not very much price variation. In this paper, we propose a semiparametric approach to the consumer demand problem in which expenditure share equations are nonparametric in the real expenditure direction and parametric (with fixed coefficients) in price directions. Here, Engel curves are unrestricted so that demands may have any rank. We also consider a 'varying coefficients' extension in which price effects depend on real expenditure. Because the demand model is derived from a model of cost, it may be restricted to satisfy integrability and used for consumer surplus calculations. Since real expenditure is not observed, 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. The estimation procedures are introduced for both cases, under integrability restrictions and without. Further we give specification tests to check these integrability restrictions. An empirical illustration with Canadian price and expenditure data shows that Engel curves display curvature which cannot be encompassed by standard parametric models. In addition, we find that although the rationality restriction of Slutsky symmetry is rejected in our fixed coefficients model, it is not rejected in the varying coefficients extension.

Keywords: Consumer Demand, Engel curves, Semiparametric Econometrics, Generated regressors.

*Pendakur gratefully acknowledges the financial support of the Social Sciences and Humanities Research Council of Canada through its standard research grant program; Sperlich the one of the Spanish "Dirección General de Investigación del Ministerio de Ciencia y Tecnología", number SEJ2004-04583/ECON. Both authors thank Emmanuel Guerre and the participants of the 2004 CAM/Cemmap Workshop on "Nonparametric Methods under Behavioural Restrictions", in particular Arthur Lewbel and Enno Mammen for helpful discussion.

1 Introduction

Consumer demand microdata typically exhibit a great deal of expenditure variation but not very much price variation. In this paper, we propose a semiparametric approach to the consumer demand problem in which expenditure share equations are nonparametric in the real expenditure direction and parametric in price directions. This 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 are likely less rich.

Parametric approaches to the estimation of consumer demand systems have the nice feature of allowing the researcher to estimate a cost function, which makes consumer surplus calculations clear and easy. However, parametric approaches typically impose strict limits on the complexity of Engel curves. Nonparametric approaches to the estimation of consumer demand 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 gains 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 cost function 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 in the M price directions. In an extension, we allow price effects to depend on utility and come in via a 'varying coefficients' structure.

As a consequence, in our model Engel curves – expenditure-share equations over real expenditure at a particular price vector – are unrestricted, and so the demand system can have any rank up to $M-1$ (see Lewbel (1991) for a detailed discussion of the rank of demand systems). This contrasts sharply with parametric approaches, such as the popular quadratic almost ideal demand system wherein the demand system is rank 3 and Engel curves are quadratic in log-expenditure.

Since the demand system is derived from a cost function, it may be restricted to satisfy homogeneity, symmetry and concavity, which together comprise the integrability conditions. However, because the cost function cannot be inverted analytically, we cannot substitute indirect utility into the compensated demand system to generate an uncompensated demand system. Instead, we numerically invert the cost function into real expenditure – a convenient cardinalisation of utility – and substitute that into the compensated demand system.

As mentioned above, our approach is semiparametric – expenditure shares are nonparametric over real expenditure, but price effects come in parametrically. Although this is more restrictive than a fully nonparametric approach (such as Haag, Hoderlein and Pendakur 2005), there are at least two important advantages. First, our model is comprised of functions which have intuitive economic interpretations. Second, our approach eliminates the curse of dimensionality faced in multivariate nonparametric regression. In our context, a fully nonparametric approach to estimating a consumer demand system has $M+1$ dimensions, whereas our approach has only 1 nonparametric dimension.

If real expenditure were directly observed, then standard semiparametric methods for partially linear models could be applied. However, because it is not observed, we numerically invert our cost function to generate a consistent predictor for real expenditure. In the econometric theory we must account for the constructed regressor in both the estimation of the parametric and of the nonparametric part of the model. We provide asymptotics for the partially linear expenditure-share system with any consistent predictor of real expenditure. In particular, we are able to show $\sqrt{n}-$convergence for the parametric price effects and efficient convergence rates for the nonparametric components. We also provide consistent predictors for real expenditure, each of which satisfies the requirements for our constructed regressor model.

For several decades, economists have been searching for parametric models of consumer demand systems that have sufficient flexibility in Engel curves to accommodate actual behavior (see, for example, Banks, Blundell and Lewbel 1997). This search led to a plethora of non- and semiparametric investigations of price-invariant Engel curves (for example: Blundell, Duncan and Pendakur 1998; Blundell, Chen and Christensen 2004) which revealed substantial evidence of curvature which cannot be accommodated in existing parametric models. Haag, Hoderlein and Pendakur (2005) propose a fully nonparametric approach to modelling consumer demand over prices and expenditure, but do not attempt to reduce the dimensionality of the problem as one would in a semiparametric approach. Lewbel and Pendakur (2006) propose a fully parametric approach to modeling consumer demand over prices and expenditure where Engel curves are unrestricted, but price effects are restricted. 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 is possible. Our model fills this gap.

The econometric challenge is twofold: to handle semiparametric estimators with nonparametrically generated regressors, and to identify valid predictors for real expenditure. Semiparametric partial linear regression was introduced by Robinson (1988) and Speckman (1988), but without considering generated regressors. For the same regression problem Yatchew (1997,2003) introduced differencing methods with complete asymptotic theory. A similar approach has also been applied for the unpleasant case when nonparametrically generated regressors enter in the nonparametric part of the model, see Rodríguez-Póo, Sperlich, and Fernández (2005) who consider the problem of estimating a semiparametric triangular system. For the estimation of the nonparametric part we use local linear smoothers, first introduced by Lejeune (1985). See also Fan and Gijbels (1996) for a survey. However, to our knowledge the theory has not been done for the more sophisticated local linear model with nonparametrically generated regressors.

The varying-coefficient model seems to us a natural way to make price effects depend on utility. A varying coefficients model retains the unit dimensionality of the nonparametric problem, but allows price effects to be different for rich versus poor consumers. Lewbel and Pendakur (2006) consider a parametric model similar to this, but force price effects to come in linearly. Alternatively, one could consider an additive interaction model as in Sperlich, Tjøstheim, and Yang (2002). Cleveland, Grosse and Shyu (1991) introduced the idea of varying coefficients, Fan and Zhang (1999) and Cai, Fan and Li (2000) provide asymptotic theory for it. Again, they do not consider the case of generated regressors. Sperlich (2005) provides some asymptotic theory for kernel density and Nadaraya-Watson estimators when the explanatory variables are (nonparametrically) constructed, but these results cannot be applied in our context. In this paper, we therefore develop the theory for the more complex cases of local polynomial and varying-coefficient models. Actually, the asymptotic results we provide are of general interest because they are applicable to any semi- or nonparametric estimation problem with generated regressors in which local polynomial, partial linear or varying coefficient models are used.

Typical resampling methods to make valid inference on nonparametric estimates are the wild bootstrap and subsampling. Härdle, Huet, Mammen, and Sperlich (2004) provide wild bootstrap methods for a large set of semiparametric models. Davidson and Flachaire (2005) introduce a "tamed" version of the wild bootstrap that seems particulary interesting for our context. Subsampling is extensively discussed in the book of Politis, Romano, and Wolf (1999).

We note that one of the main criticisms of nonparametric methods is their lack of feasibility in practice. Our method is straightforward to implement, is numerically robust, can handle large data sets in reasonable time, and the results are easy to interpret. We use differencing methods, that is ordinary least squares on transformed data, to get parametric estimates of price effects. We use univariate local linear smoothers on transformed data to get estimates of the nonparametric components of the model. The simplicity of this procedure makes this methodology practical for empirical researchers. We implement the model with Canadian price and expenditure data. The estimated expenditure-share equations exhibit quite a lot of nonlinearity. We find that some expenditure-share equations are 'S-shaped' or even more complex.

In addition to uncovering complexity in the curvature and rank of Engel curves, our approach illuminates several aspects of how prices affect demand. First, although we reject symmetry in the partially linear, or fixed-coefficients, version of our model, we do not reject symmetry in the varying-coefficients version of the model. This suggests that rejection of the rationality restriction of Slustky symmetry may be due to unduly restrictive incorporation of price effects in typical parametric models. This explanation may supplement other recent explanations having to do with the unobserved heterogeneity, see Lewbel (2001). Second, the additional complexity of utility-varying price effects is statistically important. The fixed-coefficients restriction is rejected against a varying-coefficients alternative. This suggests that the standard practice of using a single matrix of parameters in a parametric model to capture substitution effects is not sufficient to capture actual behaviour. Third, the variation of substitution effects over utility may be economically important. We find that estimated consumer surplus measures over a hypothetical price change are quite different between the models.

The rest of the paper is organized as follows. In Section 2 we introduce our model of cost and demand. In Section 3 we provide the econometric theory for estimation of our demand system and discuss efficiency. In Section 4 we discuss how the integrability conditions can be imposed and tested, and the use of bootstrap. In Section 5 we allow the price effects to depend on utility and provide consistent estimators for the resulting semiparametric varying coefficient model. Again, the complete asymptotic theory is derived and integrability conditions are discussed. Section 6 presents an empirical example using Canadian microdata. Section 7 concludes and discusses some straight forward extensions such as the inclusion of demographic effects. All proofs are deferred to the Appendix.

2 The 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 give the utility attained when facing log-prices $\boldsymbol{p}$ with log total-expenditure (log-budget) x, and log-cost $\ln C(\boldsymbol{p}, u)$ as its inverse over x giving the minimum log total-expenditure required to attain the utility level u when facing log-prices $\boldsymbol{p}$. Denote expenditure-share functions $\mathbf{w} = [w^1, ..., w^M]$. 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 giving 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

$$\ln C(\boldsymbol{p}, u) = u + \boldsymbol{p}'\overline{\boldsymbol{\beta}}(u) + \frac{1}{2}\boldsymbol{p}'\mathbf{A}\boldsymbol{p} \tag{1}$$

where u is an ordinal index of utility (that is, u can always be replaced by $\phi(u)$ where ϕ is an unknown increasing monotonic tranformation).[1] The restrictions

$$\iota'\overline{\boldsymbol{\beta}}(u) = 1,$$

$$\mathbf{A}'\iota = \mathbf{0}_M$$

are sufficient for homogeneity. The overbar on $\overline{\boldsymbol{\beta}}$ is to emphasize that it is a function of utility, rather than of an observable variable. The phrase 'utility-dependent' is to emphasize that $\overline{\boldsymbol{\beta}}$ (and later $\mathbf{A}$) varies arbitrarily across utility. The word 'translog' is used because if $\overline{\boldsymbol{\beta}}$ is independent of utility, the model collapses to the (homothetic) translog model of Christensen, Jorgensen and Lau (1971). Note that when $\boldsymbol{\beta}$ is linear, it collapses to the almost ideal case.

The dual indirect utility function V is defined by

$$u = V(\boldsymbol{p}, x) \circeq x - \boldsymbol{p}'\overline{\boldsymbol{\beta}}(u) - \frac{1}{2}\boldsymbol{p}'\mathbf{A}\boldsymbol{p}, \tag{2}$$

but it cannot be solved for analytically except in special cases, such as the Almost Ideal case.

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

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

$$\ln C(\boldsymbol{p}, u) = u + \boldsymbol{p}'\overline{\boldsymbol{\beta}}(u) + \frac{1}{2}\boldsymbol{p}'\mathbf{A}\boldsymbol{p} + \boldsymbol{p}'\boldsymbol{\Gamma}\mathbf{z}.$$

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

Applying Shepphard's Lemma to the log-cost function (1) yields a vector of compensated expenditure share equations $\omega(p, u)$ given by

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

where

$$\mathbf{A} = \mathbf{A}'.$$

This compensated expenditure-share system is very simple, and if utility u were observed, it would be estimable by standard semiparametric methods for partially linear models.

If the log-cost function were analytically invertible, we would substitute V in for u to derive uncompensated share equations, but as noted above, the log-cost function (1) is not invertible except in special cases. An alternative way to replace u with something observable is to construct a real expenditure variable which holds utility constant, and substitute that into the compensated expenditure share system. Set a reference vector of prices to unity, so that log-reference prices are $\overline{\boldsymbol{p}} = \mathbf{0}_M$, and note that indirect utility satisfies

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

This latter point is innocuous since 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

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

and given by

$$x^R = R(\boldsymbol{p}, x) = \ln C(\overline{\boldsymbol{p}}, V(\boldsymbol{p}, x)).$$

Combining (5) with (4) yields

$$V(\boldsymbol{p}, x) \mathrel{\hat{=}} R(\boldsymbol{p}, x) \tag{6}$$

That is, V and R are ordinally equivalent representations of preferences. 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. Of course, since the solution for R uses analytical forms for both C and V, recovering R is easy only if C is analytically invertible. However, in our model R can be estimated numerically even if an analytical solution for V is not available.

2.2 UTL Almost Observable Demand System

The UTL does not have an analytical solution for log real-expenditure or for indirect utility, and so it does not have an analytical solution for expenditure shares in terms of observable variables. Instead, we use numerical methods to estimate log real-expenditure R at each price vector.

Since in equation (4) we have $V(\overline{\boldsymbol{p}}, x) \mathrel{\hat{=}} x$, uncompensated demands, $\mathbf{w}(\boldsymbol{p}, x)$, at base prices are given by

$$\mathbf{w}(\overline{\boldsymbol{p}}, x) = \omega(\overline{\boldsymbol{p}}, V(\overline{\boldsymbol{p}}, x)) = \overline{\boldsymbol{\beta}}(V(\overline{\boldsymbol{p}}, x)) + \mathbf{A}\overline{\mathbf{p}} = \overline{\boldsymbol{\beta}}(V(\overline{\boldsymbol{p}}, x)) + \mathbf{0}_M = \boldsymbol{\beta}(x)$$

where $\boldsymbol{\beta}(x) \equiv \overline{\boldsymbol{\beta}}(V(\overline{\boldsymbol{p}}, x))$ is a vector-function of (observable) log-expenditure rather than of utility.

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 $\overline{\boldsymbol{p}}$. It solves

$$\begin{aligned} V(\boldsymbol{p}, x^N) &= V(\overline{\boldsymbol{p}}, x) \Leftrightarrow \\ x^N &= N(\boldsymbol{p}, x) = x + \mathbf{p}'\boldsymbol{\beta}(x) + \frac{1}{2}\mathbf{p}'\mathbf{A}\boldsymbol{p}. \end{aligned} \tag{7}$$

If the functions $\boldsymbol{\beta}$ and the parameters $\mathbf{A}$ are known, then N is known and given by (7). Log real-expenditure is given by the inverse of N with respect to x at each $\boldsymbol{p}$, i.e.

$$x^R = R(\cdot, x) = N^{-1}(\cdot, x). \tag{8}$$

Thus, if the functions $\boldsymbol{\beta}$ and the parameters $\mathbf{A}$ are known, then N is given by (7), and 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]

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:

$$\begin{aligned} \mathbf{w}(\boldsymbol{p}, x) &= \omega(\boldsymbol{p}, V(\boldsymbol{p}, x)) = \omega(\boldsymbol{p}, V(\overline{\boldsymbol{p}}, R(\boldsymbol{p}, x))) \\ &= \overline{\boldsymbol{\beta}}(u) + \mathbf{A}\boldsymbol{p}, \qquad (9) \\ &= \overline{\boldsymbol{\beta}}(V(\overline{\boldsymbol{p}}, R(\boldsymbol{p}, x))) + \mathbf{A}\boldsymbol{p}. \qquad (10) \end{aligned}$$

Equivalently, we may write the shares as a function of $\boldsymbol{p}$, x^R :

$$\mathbf{w}(\boldsymbol{p}, x^R) = \boldsymbol{\beta}\left(x^R\right) + \mathbf{A}\boldsymbol{p}. \tag{11}$$

This uncompensated expenditure-share system is linear in prices and an unspecified function of log real expenditure.

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

[2] Cost is locally weakly increasing in $\boldsymbol{p}$ if and only if expenditure shares are weakly greater than zero. However, the parametric component $\mathbf{A}\boldsymbol{p}$ guarantees violations of positivity with $\boldsymbol{p}$ large (small) enough when $\mathbf{A} \neq \mathbf{0}$. 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.

Note finally 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 $\overline{\boldsymbol{p}}$ is defined by

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

It varies with x unless preferences are homothetic (which in this model requires $\overline{\boldsymbol{\beta}}$ independent of u).

3 Estimation

We begin by assuming the existence of a consistent predictor for the log real-expenditure of every individual in the sample. Then, we show how to estimate the utility-dependent translog demand system given by (11) with a constructed regressor in the nonparametric component. This yields an estimator for $\mathbf{A}$ with optimal convergence rate and an estimator for $\boldsymbol{\beta}$. Following this, we discuss different consistent initial predictors for log real-expenditure which satisfy the conditions required for consistency of the estimated demand system and for efficiency of the estimator for $\mathbf{A}$.

3.1 Estimation of the linear part

To estimate matrix $\mathbf{A}$ $\sqrt{n}$- 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 actually based on the same idea: differencing out the nonparametric vector-function $\boldsymbol{\beta}$. It is convenient to describe the estimation equation-by-equation. We face the data generating process

$$W_i^j = \beta^j \left(R(\boldsymbol{P}_i, X_i) \right) + \boldsymbol{a}^j \boldsymbol{P}_i + \epsilon_i^j \quad , \; j = 1, \ldots, M \tag{12}$$

where $\boldsymbol{a}^j$ are the rows of $\mathbf{A}$, so that $\mathbf{A} = \left[\boldsymbol{a}^1 | \; ... | \; \boldsymbol{a}^M \right]$, the disturbances ϵ_i^j, $i = 1, \ldots, n$ are i.i.d. with mean zero and variance function $\sigma_{jj}(x, \boldsymbol{p})$ for all j, being bounded from above.

Taking expectations, we may write

$$\mathbf{w}(\boldsymbol{p}, x) = E[W | R(\boldsymbol{p}, x), \boldsymbol{p}] = \boldsymbol{\beta} \left(R(\boldsymbol{p}, x) \right) + \boldsymbol{A} \boldsymbol{p}$$

with $\boldsymbol{\beta}$ nonparametric. If the structure of R is ignored we would get a regression problem

$$E[W | R(\boldsymbol{p}, x), \boldsymbol{p}] = \gamma \left(\boldsymbol{p}, x \right) + \mathbf{A} \boldsymbol{p}$$

where γ is an $M-$vector function of $(\boldsymbol{p}, x)$. Clearly, γ and $\mathbf{A}$ are not nonparametrically identified, so we cannot estimate $\mathbf{A}$ by applying a standard partial linear model that ignores the structure of R. Instead, we first replace $R(\boldsymbol{p}, x)$, i.e. x^R by a predictor $\hat{x}^R$, and then estimate $\mathbf{A}$ and $\boldsymbol{\beta}$ as in Rodríguez-Poó, Sperlich, and Fernández (2005) or Yatchew (1997). Define the a compact set $\mathcal{X}^R$ including the neighborhood of all x^R of interest. For any approach we will use the following assumptions:

[3] We do not suggest that the two approaches we outline are exhaustive. Given a consistent predictor for log real-expenditure, one might also use the methods proposed by Speckman (1988) and Robinson (1988).

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

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

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

The idea of Ahn and Powell (1993) or Rodríguez-Poó, Sperlich, and Fernández (2005) is to 'smoothly' difference out the contribution of β^j using kernel weights on the distances. Consider for each j the sample

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

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

$$\begin{aligned} \hat{\boldsymbol{A}}_{RSF} &= \hat{H}_{PP}^{-1}\hat{H}_{PW} \qquad (13) \\ \hat{H}_{PP} &= \begin{pmatrix} n \\ 2 \end{pmatrix}^{-1} \sum_{i=1}^{n} \sum_{k=i+1}^{n-1} (\boldsymbol{p}_i - \boldsymbol{p}_k)(\boldsymbol{p}_i - \boldsymbol{p}_k)^T \hat{v}_{ik} \\ \hat{H}_{PW} &= \begin{pmatrix} n \\ 2 \end{pmatrix}^{-1} \sum_{i=1}^{n} \sum_{k=i+1}^{n-1} (\boldsymbol{p}_i - \boldsymbol{p}_k)(\boldsymbol{w}_i - \boldsymbol{w}_k)^T \hat{v}_{ik} \\ \text{where} \qquad & \hat{v}_{ik} = K_h(\hat{x}_i^R - \hat{x}_k^R) \end{aligned}$$

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, integrates to one, with 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 $E[Cov(\boldsymbol{P}|X^R)] = \Sigma_{P|X^R}$. 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 will 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

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

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 for n going to infinity. 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 $i \neq k$.

Note that 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 (12). These assumptions are not restrictive. Actually, they even allow for the worst case were one estimates x_i nonparametrically from

$$x_i^R \ = \ E[x_i^R|y_i] + \epsilon_i^j$$

for some known instruments y_i and for any j. I.e. here, the errors in both regression problems are perfectly correlated. Therefore these although rather technical assumptions again are "trivial":

[X2] For all $j = 1, \ldots, M$ we have $E[u_i^r(\epsilon_k^j)^s] = O(\frac{1}{n^r})$ or smaller $\forall\ i, k,\ s = 1, 2,\ r = 1, 2$; 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 are supposed to 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].

Now, we can establish the asymptotics. As they follow from the theoretical results of Rodríguez-Poó, Sperlich, and Fernández (2005) we state them just in terms of a Corollary:

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

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

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

$$E\left[\Sigma_{P|X^R}^{-1}]E[P_X \sigma_{jk}(X, \boldsymbol{P})P_X^T]E[\Sigma_{P|X^R}^{-1}\right] ,$$

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

Note that in this corollary as well as in the following corollaries and theorems our assumptions are somewhat stronger than necessary for convenience and clarity.

The suggested estimator for $\boldsymbol{A}$ is asymptotically efficient. Nevertheless it is worth considering an alternative that may be in practice superior due to its simplicity. Yatchew (1997) proposes an approach which does not employ smoothing. First, we order the observations $\{\boldsymbol{w}_i, \hat{x}_i^R, \boldsymbol{p}_i\}_{i=1}^n$ by $\hat{x}_i^R$. Then 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

$$\sum_{i=0}^m d_i = 0, \text{ and } \sum_{i=0}^m d_i^2 = 1 . \tag{14}$$

Optimal differencing weights can be found e.g. in Hall, Kay, and Titteringen (1990) and in Yatchew (2003). Again, the contribution of β^j is cancelled out due to its assumed smoothness.

Assume that

[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 (14) (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:

$$\hat{\boldsymbol{A}}_Y = [\Delta\boldsymbol{P}\Delta\boldsymbol{P}^T]^{-1}\Delta\boldsymbol{P}^T\Delta\boldsymbol{W} \tag{15}$$

Again, as the asymptotics can be derived following mainly the lines of Yatchew (1997,2003), we state them as a Corollary:

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

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

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

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

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. However, in Yatchew's approach we can see clearly how the differencing affects the variance of our estimate.

3.2 Estimation of the nonparametric part

The utility-dependent translog demand system (11) is comprised of the functions $\boldsymbol{\beta}$ and the fixed matrix $\mathbf{A}$. Given consistent initial predictors for log real-expenditure, we show above how to estimate $\mathbf{A}$ consistently and efficiently. 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 $\mathbf{A}$. In this case, we 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 $\hat{x}_i^R$ when discussing the asymptotics of an estimator.

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

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

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

Theorem 3 *We assume that [A1] - [A3] and [R1] hold. Further we assume [X1] - [X2] and that $B_X = o(h)$, and $\sigma_X^2 = O(\frac{1}{ng_n})$ for a g_n such that $\frac{1}{ng_n}$ is $o(h^2)$ (all uniformly). We assume that there exists $x^0, \boldsymbol{p}^0$ such that $x^R = R(\boldsymbol{p}^0, x^0)$. Further, for $n \to \infty$, nh and h^{-1} go also to ∞. Then, with $\boldsymbol{\beta}(x^R) := \{\beta^j(x^R)\}_{j=1}^M$ and $\hat{\boldsymbol{\beta}}(x^R) := \{\hat{\beta}^j(x^R)\}_{j=1}^M$ it holds for x^R interior point of $\mathcal{X}^R$*

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

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, respectively second, derivatives. Further, recall that $f(\cdot)$ denotes the marginal density of X^R, then with $\mu_l(K) = \int v^l K(v)dv$ it is

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

where $\oplus$ means "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 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 and (as assumed in the theorem), the variance σ_X^2 is of rate $\frac{1}{ng_n}$. Thus, the assumptions $B_x = o(h)$ and $\sigma_X^2 = o(h^2)$ are trivially fulfilled.

From Sperlich (2005) we can also derive the asymptotic distribution of the local constant estimator for $\boldsymbol{\beta}$. However, since the local linear estimator is almost as easy to implement as the local constant one, but is known to be more efficient (see Lejeune (1985) and Fan and Gijbels (1996)), we use the local linear version.

We are also interested in whether or not we can improve the estimation of $\boldsymbol{\beta}$ by using iteration. As we have already seen in Subsection 2.2 and we will recall in the following subsections, log real-expenditures x_i^R can be recalculated having estimates of $\boldsymbol{\beta}$ and $\mathbf{A}$. There are two purposes for iteration: first, one can use iteration to generate model-consistent results. Here, if $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{A}}$ are iterated with $\hat{x}_i^R$, the estimates of each may 'settle down' in such a way that $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{A}}$ imply $\hat{x}_i^R$ and $\hat{x}_i^R$ implies $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{A}}$. 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 $\hat{\mathbf{A}}$ *or* on $\hat{x}_i^R$ without any loss of information.

A second purpose for iteration is to try to reduce or eliminate the influence of the pre-estimation of 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 actually more interested in the practical question as to whether or not 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 $\mathbf{A}$ which come from a random subsample of the data.
- (ii) Estimate with them and the full sample the function $\boldsymbol{\beta}$ and $\mathbf{A}$, 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 $\boldsymbol{\beta}$, $\mathbf{A}$, and the log real-expenditures x_i^R.

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

3.3 Consistent initial estimator for $\hat{x}_i^R$

The results above are all based on the assumption that we have predictors for the x_i^R, $i = 1, \ldots, n$ fulfilling conditions given in [X1]-[X3] in Theorem 3. Our results are rather general holding for any consistent predictor fulfilling some minimal conditions. In this paper we treat only the "worst case", i.e. that we do not have any additional information or data to predict the real expenditure than the information and data we have to estimate the demand system. We show that even then our method can be applied as will be seen also in our real data application.

To this aim we first show the behavior of an initial estimate of x_i^R computed from initial estimates of $\boldsymbol{\beta}$ and $\mathbf{A}$. Then, we discuss different approaches to initial estimates of $\boldsymbol{\beta}$ and $\mathbf{A}$ which satisfy [X1]-[X3] and are not too burdensome.

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

$$N_0(\boldsymbol{p}, x) = x + \boldsymbol{p}'\boldsymbol{\beta}_0(x) + \boldsymbol{p}'\mathbf{A}_0\boldsymbol{p}.$$

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 for x_i^R.

Given monotonic increasing costs in utility, we have increasingness of $R(\boldsymbol{p}, x)$ and $N(\boldsymbol{p}, x)$ in x for each $\boldsymbol{p}$, so that we can invert N and derive the convergence rate. For each $\boldsymbol{p}$ fixed, and

$t = \hat{N}(\boldsymbol{p}, x),\ R(\boldsymbol{p}, t) = N^{-1}(\boldsymbol{p}, t),\ \hat{R}(\boldsymbol{p}, t) = \hat{N}^{-1}(\boldsymbol{p}, t)$

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

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

3.3.1 Initial Estimator for $\boldsymbol{\beta}$

Recall that $\overline{\boldsymbol{p}} = \mathbf{0}_M$, which implies $R(\overline{\boldsymbol{p}}, x) = x$, and that for observations facing $\overline{\boldsymbol{p}}$, $x_i^R = x_i$ and $\mathbf{A}\boldsymbol{p} = \mathbf{0}_M$. 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 $\overline{\boldsymbol{p}}$. Since

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

we may construct the following consistent initial estimator for $\boldsymbol{\beta}$:

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

This estimator may be constructed either by estimating the univariate nonparametric Engel curve at $\overline{\boldsymbol{p}}$, using only the N_0 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 $\boldsymbol{p} = \overline{\boldsymbol{p}}$. While the latter approach works with any stochastic environment for the process generating $\{\mathbf{P}_i, X_i\}_{i=1}^n$, the former approach works only in an environment with many individuals facing the same price vector, and not in an environment with fully individual-level prices (such as unit values).

3.3.2 Initial Estimator for $\boldsymbol{A}$

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

$$\boldsymbol{\Upsilon}(\boldsymbol{p}, x) = \nabla_{\boldsymbol{p}} \mathbf{w}(\boldsymbol{p}, x) + \nabla_{\mathbf{x}} \mathbf{w}(\boldsymbol{p}, x) \mathbf{w}(\boldsymbol{p}, x)'.$$

Given our model, $\boldsymbol{\Upsilon}(\boldsymbol{p}, x) = \mathbf{A}$ is a matrix of constants. In the following we discuss three ways to proceed.

Unrestricted Estimator via averaging: Following Haag, Hoderlein and Pendakur (2005), one may easily 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 $\mathbf{A}$ is thus given by

$$\mathbf{A}_0 = \frac{1}{n}\sum_{i=1}^{n} \widehat{\boldsymbol{\Upsilon}}(\mathbf{P}_i, X_i). \tag{16}$$

Here $\mathbf{A}_0$ may or may not be symmetric.

We note that Lewbel (2001) shows that $\widehat{\boldsymbol{\Upsilon}}(\boldsymbol{p}, x)$ is only a consistent estimate of $\boldsymbol{\Upsilon}(\boldsymbol{p}, x)$ if the disturbance terms on the right-hand side of the regression are not important behavioral parameters or if they satisfy rather complex covariance conditions that are hard to verify in practice. The reason is that although $\nabla_{\boldsymbol{p}}\mathbf{w}(\boldsymbol{p}, x)$ may be consistently estimated via nonparametric methods in the presence of unobserved behavioral heterogeneity, $\nabla_{\mathbf{x}}\mathbf{w}(\boldsymbol{p}, x)\mathbf{w}(\boldsymbol{p}, x)'$ contains a cross-product of such heterogeneity which makes consistent estimation of this term difficult.

Symmetry-Restricted Estimator via averaging: Crossley and Pendakur (2005) and Hoderlein (2005) propose the following symmetry-restricted estimator of the average compensated semi-elasticity matrix.

$$\begin{aligned}
\boldsymbol{\Upsilon}(\boldsymbol{p}, x) + \boldsymbol{\Upsilon}(\boldsymbol{p}, x)' &= 2\boldsymbol{\Upsilon}(\boldsymbol{p}, x) \\
&= \nabla_{\boldsymbol{p}}\mathbf{w}(\boldsymbol{p}, x) + \nabla_{\boldsymbol{p}}\mathbf{w}(\boldsymbol{p}, x)' + \\
&\quad \nabla_{\mathbf{x}}\mathbf{w}(\boldsymbol{p}, x)\mathbf{w}(\boldsymbol{p}, x)' + \mathbf{w}(\boldsymbol{p}, x)\nabla_{\mathbf{x}}\mathbf{w}(\boldsymbol{p}, x)' \\
&= \nabla_{\boldsymbol{p}}\mathbf{w}(\boldsymbol{p}, x) + \nabla_{\boldsymbol{p}}\mathbf{w}(\boldsymbol{p}, x)' + \nabla_{\mathbf{x}}\left(\mathbf{w}(\boldsymbol{p}, x)\mathbf{w}(\boldsymbol{p}, x)'\right).
\end{aligned}$$

Given nonparametric estimates $\widehat{\nabla_{\boldsymbol{p}}}\ \mathbf{w}(\boldsymbol{p}, x)$ and $\widehat{\nabla_{\mathbf{x}}}(\mathbf{w}(\boldsymbol{p}, x)\mathbf{w}(\boldsymbol{p}, x)')$, we construct

$$\widehat{\boldsymbol{\Upsilon}}(\boldsymbol{p}, x) = \frac{1}{2}\left(\widehat{\nabla_{\boldsymbol{p}}}\mathbf{w}(\boldsymbol{p}, x) + \widehat{\nabla_{\boldsymbol{p}}}\mathbf{w}(\boldsymbol{p}, x)' + \widehat{\nabla_{\mathbf{x}}}\big(\mathbf{w}(\boldsymbol{p}, x)\mathbf{w}(\boldsymbol{p}, x)'\big)\right),$$

and estimate

$$\mathbf{A}_0 = \frac{1}{n}\sum_{i=1}^{n} \widehat{\boldsymbol{\Upsilon}}(\mathbf{P}_i, X_i)\ . \tag{17}$$

Because $\nabla_{\mathbf{x}}\left(\mathbf{w}(\boldsymbol{p}, x)\mathbf{w}(\boldsymbol{p}, x)'\right)$ 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. We will come back to the symmetry restriction in the next section when discussing the different aspects of integrability.

Note that for both estimators, (16) and (17), for the matrix of constants $\mathbf{A}$, we use high-dimensional nonparametric pre-estimators that have very slow convergence rates. Fortunately, the averaging in the estimates of $\boldsymbol{A}_0$ reduces the variance 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. This method of reducing the dimensionality to get efficient convergence rates is well known in the nonparametric literature for marginal integration estimators, see e.g. Newey (1994), or Hengartner and Sperlich (2005) for an improved version of this estimator.

Parametric rates for A with M-demands: An alternative approach to constructing an initial estimator for $\mathbf{A}$ is to use the 'M-demand' approach proposed by Browning (2005). Assume that at least one good has share equations which are monotonic in utility (x^R), and denote a good in this class as good 1. For example, food-at-home shares are known to be decreasing in log-expenditure. Then, we may express expenditure shares $w^2, ..., w^{M-1}$ as functions of $\boldsymbol{p}, w^1$, rather than as functions of $\boldsymbol{p}, x^R$. First, using the initial estimate of $\boldsymbol{\beta}$ for good 1, β_0^1, invert w^1 around x to get

$$x^R = (\beta_0^1)^{-1}\left(w^1 - \boldsymbol{a}^1\boldsymbol{p}\right)$$

where $\boldsymbol{a}_1$ is a row-vector comprised of the first row of $\boldsymbol{A}$. Second, substitute this expression for x^R into the expenditure shares $w^2, ..., w^{M-1}$ using initial estimates β_0^j, $j = 2, ..., M-1$, to get M-demands:

$$w^j(\boldsymbol{p}, w^1) = \beta_0^j\left((\beta_0^1)^{-1}\left(w^1 - \mathbf{a}^1\boldsymbol{p}\right)\right) + \mathbf{A}\boldsymbol{p}, \tag{18}$$

$j = 2, ..., M-1$. Taking $\boldsymbol{\beta}_0$ as given, the M-demand system (18) may be estimated via GMM to get a consistent initial estimate for $\mathbf{A}$, which we denote $\mathbf{A}_0$.

If w^1 is measured with error, or if it contains additive preference heterogeneity, then the M-demand system above suffers from an endogeneity problem. However, if we take β_0^j and $(\beta_0^1)^{-1}$ as given, the endogeneity is inside a known function.

4 Inference and Estimation under Integrability

4.1 Bootstrap Confidence Bands

Our aim is to construct symmetric confidence bands around the nonparametric function estimate β^j. For the construction of uniform confidence bands around $\hat{\beta}^j$ 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, compare Theorem 3. Following Härdle, Huet, Mammen, and Sperlich (2004) we determine the distribution of S_j via 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 estimates. Then, the confidence bands are given 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 $(\alpha \in (0,1))$ of S_j^*.

In Section 3 we have seen how to take into account the variance caused by the prediction of the x_i^R, $i = 1, \ldots, n$. In our application it actually does not influence the distribution of the final estimates. Therefore, the bootstrap samples can be generated with our 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 ϵ_i^{j*} reflects the covariance structure (between the M different equations) of the original errors. However, when it turns out that the prediction of the x_i^R affects the final estimate in the particular real data problem, one would start the bootstrap already from the prediction of the x_i^R (can also be done via subsampling, see Politis, Romano, and Wolf (1999).

For the wild bootstrap, we propose the Radamacher distribution where the bootstrap error vector $\epsilon_i^* = \mathbf{e}_i\mu_i$ with $\mathbf{e}_i$ 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. Other bootstrap methods are certainly thinkable, too.

4.2 Homogeneity

Homogeneity is easily satisfied in this context, by normalizing prices with respect to p^M. As noted in the discussion of the cost function (1), homogeneity is satisfied if and only if $\iota'\overline{\boldsymbol{\beta}}(u) = 1$ and $\mathbf{A}\iota = \mathbf{0}_M$. 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 restrict the estimation.

The linear restriction on the parameters $\mathbf{A}$ can be implemented in various ways, normalizing prices being one of them. Denote the k'th 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 $\widetilde{\mathbf{P}}_i = [\widetilde{P}_i^1, ..., \widetilde{P}_i^{M-1}]$. Denote the $\widetilde{\mathbf{A}}$ as the $M-1$ by $M-1$ upper-left submatrix of $\mathbf{A}$, and denote $\widetilde{\boldsymbol{\beta}}$ as the $M-1$ vector function giving all but the last element of $\boldsymbol{\beta}$, and $\widetilde{\mathbf{W}}$ and $\widetilde{\mathbf{w}}$ defined analogously as the first $M-1$ expenditure shares and share functions. Log nominal expenditure over normalized prices is

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

and log real expenditure is still its inverse with respect to x at each $\widetilde{\boldsymbol{p}}$. Then, the homogeneity-restricted model analogous to (11) is given by

$$E[\widetilde{\mathbf{W}}|X^R = x^R, \widetilde{\mathbf{P}} = \widetilde{\boldsymbol{p}}] = \widetilde{\boldsymbol{\beta}}(x^R) + \widetilde{\mathbf{A}}\widetilde{\boldsymbol{p}}. \tag{19}$$

This model may be estimated via the techniques outlined above.

A goodness-of-fit test for homogeneity could compare estimates from the unrestricted model (11) to the restricted model (19) following Aït-Sahalia, Bickell and Stoker (2001) or Haag, Hoderlein and Pendakur (2005). However, in our semiparametric partial linear model we can not only avoid the curse of dimensionality, the test on homogeneity does even reduce here to a simple parametric test. To see this just note that if the rows of $\boldsymbol{A}$ sum to zero, the $\boldsymbol{\beta}$ will automatically sum up to one by definition (being then simply the sum over all M expenditure shares). Therefore, it is sufficient to consider the null hypothesis

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

against the alternative that at least one of these equations does not hold. This null hypothesis can be checked by a Wald-type-test at a parametric rate since the variance-covariance structure is explicitly given in Corollary 1 and Corollary 2 respectively. Unfortunately, the variance-covariance matrix has to estimated nonparametrically. Although it can be shown that this is theoretically valid, for practical applications one should use 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. The estimation of our model under the null hypothesis to get bootstrap samples under H_0 – necessary to get the statistic's null distribution – has been discussed above. Further, we recommend undersmoothing the nonparametric part $\beta(\cdot)$ so as to avoid distorting the bootstrap with a nonparametric bias term.

4.3 Symmetry

Symmetry is also easily imposed in this context. The Slutsky Matrix $\mathbf{S}$ is given by

$$\mathbf{S} = \mathbf{A} + \mathbf{w}\mathbf{w}^T - diag\{\mathbf{w}\} \ . \tag{20}$$

Symmetry is satisfied if and only if $\mathbf{A}$ is symmetric. To get an estimate $\widetilde{\mathbf{A}}$ ($\mathbf{A}$) that is symmetric we just have to apply a linear estimator under linear cross-equation restrictions in the partially model. Thus, for example, in equation (15), we would include the restriction $\widetilde{\mathbf{A}} = \widetilde{\mathbf{A}}'$ ($\mathbf{A} = \mathbf{A}'$).

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

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

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

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

which we denote Σ_α. Then our test statistic is $\alpha^T \Sigma_\alpha^{-1} \alpha$ that under H_0 converges to a chi-square distribution with $(M-1)(M-2)/2$ degrees of freedom, and under H_1 goes to infinity. It is obvious that also this can be written as a Wald-type-test as the test for homogeneity, applying the same variance-covariance matrix. Thus, as above, the variance-covariance matrix has to estimated nonparametrically and we therefore again recommend using bootstrap or subsampling which can also be used to find the critical values of the test statistic.

4.4 Concavity

Concavity is satisfied if the Slutsky matrix $\mathbf{S}$ is negative semidefinite. Since $\mathbf{w}\mathbf{w}^T - diag\{\mathbf{w}\}$ is negative definite, negative semidefiniteness of $\mathbf{A}$ is sufficient for concavity. However, restricting $\mathbf{A}$ to be negative semidefinite is overly restrictive.

Alternatively one may search within any reasonable matrix-norm for the negative semidefinite matrix closest to the estimate of $\mathbf{S}$. As long as this matrix norm is consistent with the objective functions used for the estimation of $\mathbf{A}$ and $\mathbf{w}$, this does not lead to any efficiency loss. The degree of difficulty of implementation depends strongly on the matrix norm we choose.

A test could then be constructed on the difference between the unrestricted estimate and the final, negative semidefinite estimate. A reasonable weighting which accounts for the variance-covariance structure $\mathbf{S}$ is difficult to construct and would depend on many nonparametric auxiliary estimates, so the practical application of this would be difficult. Thus, it may be preferable to use a simple subsampling based test, which would not require us to generate bootstrap samples based on the projection of the unrestricted $\widehat{\mathbf{S}}$ onto its negative semidefinite analog.

5 Varying Price Effects

So far our log-cost function (1) has first-order price effects which depend on utility, but second-order price effects which are independent of utility. A natural extension is to let the second-order price effects depend on utility, as in the following:

$$\ln C(\boldsymbol{p}, u) = u + \boldsymbol{p}'\overline{\boldsymbol{\beta}}(u) + \frac{1}{2}\boldsymbol{p}'\overline{\mathbf{A}}(u)\boldsymbol{p}. \tag{21}$$

Here, $\overline{\mathbf{A}}$ is a matrix-valued function of utility u. Indirect utility is defined by

$$u = V(\boldsymbol{p}, x) \mathrel{\hat{=}} x - \boldsymbol{p}'\overline{\boldsymbol{\beta}}(u) - \frac{1}{2}\boldsymbol{p}'\overline{\mathbf{A}}(u)\boldsymbol{p}, \tag{22}$$

which is log-money-metric at base prices, because

$$V(\overline{\boldsymbol{p}}, x) \mathrel{\hat{=}} x.$$

The compensated expenditure-share system corresponding to (21) is given by Shepphard's Lemma as

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

and at a vector of base prices is equal to

$$\omega(\overline{\boldsymbol{p}}, u) = \overline{\boldsymbol{\beta}}(u).$$

Uncompensated expenditure-shares at base prices are given by

$$\mathbf{w}(\overline{\boldsymbol{p}}, x^R) = \boldsymbol{\beta}(x),$$

because if $\boldsymbol{p} = \overline{\boldsymbol{p}}$, then $x^R = x$.

Since $\overline{\mathbf{A}}$ is a an unrestricted function of u, we may create the matrix-valued function $\mathbf{A}\left(x^R\right)$ as

$$A\left(x^R\right) = \overline{\mathbf{A}}(u) = \overline{\mathbf{A}}(V(\boldsymbol{p}, x)) = \overline{\mathbf{A}}(V(\overline{\boldsymbol{p}}, x^R))$$

At price vectors other than $\boldsymbol{p} = \overline{\boldsymbol{p}}$, the observable expenditure-share functions $\mathbf{w}$ are given by

$$\mathbf{w}(\boldsymbol{p}, x^R) = \boldsymbol{\beta}\left(x^R\right) + \mathbf{A}(x^R)\boldsymbol{p}. \tag{23}$$

As in the simpler model, we may solve for N as

$$\begin{aligned} V(\boldsymbol{p}, x^N) &= V(\overline{\boldsymbol{p}}, x) \Leftrightarrow \\ N(\boldsymbol{p}, x) &= x + \boldsymbol{p}'\boldsymbol{\beta}(x) + \boldsymbol{p}'\mathbf{A}(x)\boldsymbol{p}, \end{aligned} \tag{24}$$

and if the functions $\boldsymbol{\beta}$ and $\mathbf{A}$ are known, then N is known, and $x^R = R(\boldsymbol{p}, x)$ solves $V(\boldsymbol{p}, x) = V(\overline{\boldsymbol{p}}, x^R)$ as in (5), and $R(\boldsymbol{p}, x)$ is given by the inverse of N with respect to x.

Note finally that the nice feature of having a clear interpretability is shared with the simpler partial linear model: the demand system is again characterized by a set of Engel curve functions, i.e. the $\boldsymbol{\beta}^j(\cdot)$ and a matrix of compensated price effects, $\boldsymbol{A}(\cdot)$.

5.1 Estimation of the Model with varying Price Effects

We start by assuming that we can get some consistent predictors for the log-real expenditures X_i^R, as we did in the partial linear model with constant $\mathbf{A}$. The estimators defined above do not help here. Instead, notice that the model

$$E[\mathbf{W}|\boldsymbol{p}, x^R] = \boldsymbol{\beta}\left(x^R\right) + \mathbf{A}(x^R)\boldsymbol{p} \tag{25}$$

can be interpreted as a varying coefficient model which is linear in $\boldsymbol{p}$ but with coefficients that vary with x^R. For $\boldsymbol{p}$, x^R observed these models are well studied in the non- and semiparametric literature, see e.g. Cleveland, Grosse and Shyu (1991), or Fan and Zhang (1999). We need to modify such models to allow for a constructed regressor, the predictor of x^R.

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

$$\sum_{i=1}^n \left[W_i^j - \beta_0^j - \beta_1^j(\hat{x}_i^R - x_0^R) - \left\{\mathbf{a}_0^j + \mathbf{a}_1^j(\hat{x}_i^R - x_0^R)\right\}'\mathbf{P}_i\right]^2 K_h(\hat{x}_i^R - x_0^R) \tag{26}$$

and then set $\widehat{\beta^j}(x_0^R) := \beta_0^j$, $\widehat{\mathbf{a}^j}(x_0^R) := \mathbf{a}_0^j$ for all j, k. Here, $K_h(\cdot)$ is a kernel function defined as before.

For the ease of notation let us set log-prices $\mathbf{P}_i^0 \equiv \mathbf{0}_M$ for all i, and $a_0^j(x^R) := \beta^j(x^R)$. In addition to the assumptions of Theorem 3 we need

[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.

Further, we replace [A3] now by

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

Then, we can state the following result for which the proof again is given in the appendix.

Theorem 4 *Assume the same conditions as in Theorem 3 without [A3], but adding [V1], [V2]. Define the estimators as in (26), and set* $\alpha_k := \left(a_0^k, a_1^k, \cdots, a_M^k\right)^T (x^R)$ *for* $k = 1, \ldots, M$. *Then it holds*

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

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, respectively second, derivatives in* x^R.

The covariance structure is given by

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

where Ω *is the inverse of* $\Omega^{-1} := E\left[(P^0, P^1, \ldots, P^M)^T(P^0, P^1, \ldots, P^M)|x^R\right]$ *and* $\Sigma_{\epsilon k,k}(x^R)$ *is the* (k,k)*'th element of* $\Sigma_\epsilon(x^R)$.

Certainly, the statement could have been formulated the same way for the vector $\gamma_j = (a_j^1, a_j^2, \ldots, a_j^M)$. For the unrestricted estimator the covariance structure is then given by

$$\frac{1}{nh}f^{-1}(x^R)||K||_2^2\Omega_{j,j}\Sigma_\epsilon(x^R) \oplus \sigma_X^2(x^0, \boldsymbol{p}^0)\gamma_j'^2$$

for $j = 0, \ldots, M$

Given reasonable predictors for x_i^R, one can estimate the varying coefficient version of the utility-dependent translog demand system. However, as we show in the next subsection, it is much harder to create good predictors for x_i^R given the structural model of cost in this case. An alternative to methods outlined above can be found in Cai, Fan, and Li (2000) who propose a Maximum Likelihood approach. Although such an approach requires specification of the conditional distribution of the observed shares, it offers a relatively easy model check which we will describe below.

Regarding the value of iteration, we refer to the discussion in the preceding section as the arguments do not change between the fixed- and varying-coefficients versions of our model.

5.1.1 Consistent initial estimator for $\hat{x}_i^R$

As in Subsection 3.3 we have to find consistent predictors $\hat{x}_i^R$ for all i. Again, we proceed by plugging initial estimates of $\boldsymbol{\beta}$ and $\mathbf{A}$ 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 $\boldsymbol{\beta}(x^R)$ we as in Subsection 3.3. However, for $\mathbf{A}$ we must use a different approach.

In the varying coefficients model, the matrix $\mathbf{A}$ 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 again can obtain a consistent estimator for the matrix-valued function $\mathbf{A}$ similar as in the simpler case when it was assumed to be constant:

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

Here, however, there is no averaging, so this estimator inherits the slow convergence rate of the nonparametric pre-estimator $\widehat{\boldsymbol{\Upsilon}}$. Again we can either calculate the symmetry-restricted or the unrestricted estimator for $\boldsymbol{\Upsilon}$ to get our $\mathbf{A}_0(x^R)$. Then, as in Section 3.3 we can derive the convergence rate of the predictors $\hat{x}_i^R$. Finally, notice that here, the initial estimator for the matrix function $\mathbf{A}$ cannot be obtained via the M-demand strategy.

5.1.2 Testing and Bootstrap Confidence Bands

For $\mathbf{A}$ independent of x^R, it is straight forward to impose restrictions coming from integrability, and inference can be derived directly from the asymptotic theory. In contrast, when we allow $\mathbf{A}$ to be nonparametric in the direction of real expenditure, imposing integrability and doing statistical inference is much more complicated. Therefore, a first step should be to check whether such an effort is justified. This means to test $\mathbf{A}$ for significant deviations from being constant. This can be done either by a bootstrap (or subsampling) based test similar to those proposed in Härdle, Huet, Mammen and Sperlich (2004) or, as mentioned above based on a Likelihood approach as suggested in Cai, Fan and Li (2000). Note that the latter mentioned article treats explicitly this testing problem. Since our problem is “simply” an extension to the case when generated regressors are included, we skip further discussion here.

When the coefficient matrix turns out to depend significantly on x^R, testing symmetry can be done the way it is suggested by Haag, Hoderlein and Pendakur (2005). The advantage of starting with our model is the strong reduction of dimensionality we yield, the disadvantage is the inclusion of a nonparametrically generated regressor. Asymptotic theory tells us that our approach is preferable but unfortunately, in practice, i.e. for finite samples we do not know.

Homogeneity and concavity, always understood as “local”, can be imposed the same way as before with the (computational) burden that $\mathbf{A}$ can differ at each x^R.

For model (25) the bootstrap procedure actually does not change, and therefore the construction of confidence bands for the functions $\boldsymbol{a}^j(x^R)$ works as before.

6 Empirical Example

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 $(1, ..., 1)$.

Table 1 gives summary statistics for 6952 observations of rental-tenure unattached individuals aged 25-64 with no dependents. Analysis is restricted to these households to minimize demographic variation in preferences. The empirical analysis uses annual expenditure in nine expenditure categories: food-in, food-out, rent, clothing, household operation, household furnishing & equipment, private transportation operation, public transportation and personal care. Personal care is the left-out equation, yielding eight expenditure share equations which depend on 9 log-prices and log-expenditure. These expenditure categories account for about three quarters of the current consumption of the households in the sample.

Table 1: The Data

		Min	Max	Mean	Std Dev
expenditure shares	food-in	0.02	0.62	0.17	0.09
	food-out	0.00	0.64	0.08	0.08
	rent	.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

All models estimated in the empirical work maintain the restriction of homogeneity, and models used in consumer surplus exercises maintain the additional restriction of Slutsky symmetry. Thus, for the 9-good demand system, the $\mathbf{A}$ ($\mathbf{A}(x^R)$) matrix (function) is a 9×9 matrix of compensated semi-elasticities with row-sums of zero, and $\beta(x^R)$ is a 9 element vector-function of log real-expenditure which everywhere sums to one. We implement our models using a predictor of log real-expenditure that uses pre-estimates of $\mathbf{A}$ and β. For the fixed-coefficient model, we compute $\mathbf{A}_0$, the pre-estimate of $\mathbf{A}$, as the average of the fully nonparametric estimate of symmetry-unrestricted compensated semi-elasticity matrix at each observation, see above. For the varying-coefficient model, we compute $\mathbf{A}_0(x^R)$, the pre-estimate of $\mathbf{A}(x^R)$, 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 compute β_0, the pre-estimate of β, as equal to the Engel curve for observations facing base prices $\overline{\boldsymbol{p}} = 0_M$. For the fixed-coefficients model, we compute a pre-estimate of log real-expenditure for each observation, and use Yatchew's difference estimator for $\mathbf{A}$ with 100'th order moving-average difference coefficients. Holding the order of the difference constant, results are essentially unchanged if we use optimal or optimal symmetric difference coefficients instead. Estimates are essentially identical if 50'th or 20'th order moving average differencing coefficients are used instead. For the nonparametric part, we use the cross-validated bandwidth.

For the varying-coefficients model, we use a bandwidth of 0.24 log real-expenditure units, which was found by cross-validation, for all equations. Results, including all the results of all tests described, do not qualitatively change if a bandwidth 50% larger or 25% smaller is used instead.

We use the subsampling approach described above to assess the influence of the pre-estimation step, and consider the homogeneity-restricted but symmetry-unrestricted fixed-coefficients model. We drew 200 subsamples containing 2000 observations each from the 6952 observations in the data described above. For each subsample, we created pre-estimates of $\mathbf{A}$ and β as described above, and used these pre-estimates to estimate the iterated model on the entire sample of 6952 observations. We iterated the model 6 iterations past the pre-estimation for each of the 200 consistent pre-estimates. If the variance of final iterated estimates across the subsamples is zero, then the pre-estimation step does not have an impact on the final estimate, see discussion above. Since for fixed $\mathbf{A}$, there is a unique β and log real-expenditure, we will discuss only the behaviour of the estimate of $\mathbf{A}_0$ across subsamples. The sum of the 64 variances of the elements of the pre-estimates of $\mathbf{A}$ across the subsamples is 0.0129. The sum of the 64 variances of the elements of iterated estimates of $\mathbf{A}$ across the subsamples is 0.0000009, which is smaller by a factor of about 15000. The variance of the iterated estimates across subsamples is numerically close to zero, and is greatly dwarfed by the sampling variance of the estimates, which sum to 0.0219 as we shall see below. Thus, we can conclude that the final iterated estimator is efficient and we only present these estimates in the results 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.

6.1 Compensated Price Effects and Symmetry

Table 2 gives the symmetry-unrestricted estimate of $\mathbf{A}$ resulting from iterating our model between estimates of $\mathbf{A}$ and β and their implied values for the constructed right-hand side variable x^R. In practise, the estimates 'settle down' after about 3 iterations. All tables present estimate values after 6 iterations. Simulated standard errors are given in italics below each estimate.

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

	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 presents 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 if they are true, the symmetry restrictions are quite informative.

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

The estimate of **A** reported in Table 2 does not 'appear' to satisfy symmetry, and indeed, a Wald test of symmetry based on the joint hypothesis that all off-diagonal terms equal their symmetric partner rejects the hypothesis. The symmetry test statistic is $\tau^{SYM} = \alpha' \Sigma_{\alpha}^{-1} \alpha$, where α is the sample estimate of the difference between off-diagonal terms equal under symmetry, and Σ_{α}^{-1} is their covariance estimated via the bootstrap. The sample value of the test statistic is 424, which is larger than 48, the 1% critical value of the χ^2_{28}, so we may reject the hypothesis of symmetry in the fixed coefficient model.[4] The failure of symmetry 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 $\mathbf{A}$ is independent of utility. Below, we argue that this latter possibility may be true.

We estimated the varying coefficient version of the model via a locally linear varying coefficients semiparametric model. We use a gaussian kernel for the local linear estimator, with a bandwidth of 0.24 selected by cross-validation. We use a single bandwidth for all 8 equations. For unrestricted models, one could use a different bandwidth in each equation. However, for the symmetry-restricted model, the model is a locally weighted SUR regression wherein all equations are stacked together. In this case, it is more natural to use a single bandwidth for all equations, so we employ this structure for all varying-coefficient models. In this model, $\mathbf{A}$ depends on u, or equivalently, on x^R. This model can encompass the partially linear model if $\mathbf{A}$ is independent of x^R, so it is natural to test whether or not this additional flexibility is necessary.

Using symmetry-unrestricted varying coefficient estimates, we construct a matrix-valued function of deviations $\widehat{\mathbf{A}}_D(x_t^R) = \widehat{\mathbf{A}}(x_t^R) - \frac{1}{T}\sum_{t=1}^{T} \widehat{\mathbf{A}}(x_t^R)$ over a grid of T equispaced points in the range of x^R. Since $\widehat{\mathbf{A}}$ is asymptotically normal, so 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 construct the $M \times M \times T-$vector λ as the vectorisation of $\widehat{\mathbf{A}}_D(x_t^R)$ over all T points, and simulate its variance, denoted Σ_λ, under the null that the partially linear model is true. Our test statistic is then $\tau^{PLM} = \lambda \Sigma_{\lambda}^{-1} \lambda$ which is distributed asymptotically as a $\chi^2_{M^2(T-1)}$. One could compare the the sample value of τ^{PLM} to its asymptotic distribution. However, in practise because the $\mathbf{A}$ matrix converges more slowly in the varying-coefficients model than in the fixed-coefficients model, we account for possible sample bias by bootstrapping the entire statistic. Using $T = 9$, the value of our test statistic is 876, and the 1% critical value its simulated distribution under the null is 124. Thus, for these data we may reject the hypothesis that the matrix of compensated price effects $\mathbf{A}$ is independent of utility, and may comfortably use the varying coefficients model.

As noted above, in the fixed coefficients model, we reject symmetry. We construct a test statistic for symmetry in the varying-coefficients model analagous to that used in the fixed-coefficients model. In the varying-coefficients model, the vector α depends on x^R, so we denote this vector-function as $\alpha(x^R)$, 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 $\tau_t^{SYM} = \alpha'(x_t^R)\Sigma^{-1}_{\alpha(x_t^R)}\alpha(x_t^R)$ for $t = 1, ..., T$. Using sample values of $\alpha'(x_t^R)$ and a bootstrap estimate of $\Sigma^{-1}_{\alpha(x_t^R)}$, each of these tests is asymptotically χ^2_{28} with a 1% critical value of 48. However, to account for possible

[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.

small-sample bias, we bootstrap the entire statistic. Figure 1 shows 90% 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. It may be that symmetry is true, but the utility-independence restriction on $\mathbf{A}$ is 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 $\mathbf{A}(x^R)$ in the varying-coefficient model yields a test with low power, so that symmetry is not rejected even though it is false. Below, we consider the latter possibility.

Table 4 gives 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 most precise. Simulated standard errors are given in italics below each estimate. Clearly, the precision of the estimates is lower than in the fixed coefficients model. Most elements of $\mathbf{A}(x^R)$ 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, and thus we should treat this non-rejection of symmetry with caution.

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*

Since our model of demand is generated from a model of cost and uses real expenditure – a dual of utility – as a dependent variable, symmetry must hold for the model to be sensible. In addition, consumer surplus calculations are only unique for estimated models satisfying symmetry. Thus, 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. Table 5 gives estimates of compensated price effects for the symmetry-restricted varying-coefficients model evaluated at the median of log real-expenditure.

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

The differences between the models are most easily seen graphically. Figures 2-4 give estimated values of selected elements of $\widehat{\mathbf{A}}(x^R)$ 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, eg, Banks, Blundell and Lewbel 1997) are presented with dark dotted lines. Simulated 90% 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 the 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 which 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-coefficient model performs worst 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 will show below).

6.2 Engel Curve Functions

Figures 5-12 show estimated functions β for the models we estimated. Since $\mathbf{w}(\boldsymbol{p}, x^R) = \beta\left(x^R\right)$ at the base price vector, the Engel curve functions $\beta\left(x^R\right)$ give the estimated expenditure share at the base price vector. Expenditure shares are evaluated at 39 equally spaced points over the middle 99% 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% 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 5-12 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 non-parametric investigations of the shape of Engel curves (eg, 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 does not do as good a job. 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 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 (shown in Figure 3).

Some expenditure share equations appear to be 'S-shaped' as noted in previous work on Engel curves (Blundell, Chen and Christensen (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, steeply rising through middle of the distribution, and falling for the top quintile. The complexity of this Engel curve is difficult to capture in a quadratic specification, and for this reason, the QAI estimate of the private transportation share is fairly distant from both the varying- and fixed-coefficient estimates throughout the distribution of expenditure.

The approach presented here, which allows for demand systems of any rank up to $M - 1$, does reveal features of the data that parametric investigations neglect. Further, given the structural model for cost and demand, it also reveals features that 'engel-curve by engel-curve' approaches would also miss due to their small sample sizes. Differences appear both in the shape of Engel curves and in the compensated price responses across the distribution of expenditure. Given the simulated uniform confidence bands shown in the Figures, it is clear that the restricted models are rejected in statistical terms. However, it remains to be shown that the restrictions embodied in these models are costly in terms of their economic significance.

6.3 A 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 typically amount to 15% for goods such as food-out and clothing. Consider the cost-of-living index associated with subjecting rent to a 15% 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 $\mathbf{p} = [0, 0, \ln(1.15), 0, ..., 0]'$, the new price vector. Using estimated values, this is

$$\boldsymbol{p}'\widehat{\beta}(x) + \frac{1}{2}\boldsymbol{p}'\widehat{\mathbf{A}}\boldsymbol{p} = \widehat{\beta}^{rent}(x)\ln(1.15) + \frac{\widehat{a}^{rent,rent}}{2}\ln(1.15)^2$$

for the fixed-coefficients model and

$$\boldsymbol{p}'\widehat{\beta}(x) + \frac{1}{2}\boldsymbol{p}'\widehat{\mathbf{A}}(\mathbf{x})\mathbf{p} = \widehat{\beta}^{rent}(x)\ln(1.15) + \frac{\widehat{a}^{rent,rent}(x)}{2}\ln(1.15)^2$$

for the varying coefficients model.

Figure 13 shows 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, both the former 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% and 7.0%, respectively, but the varying-coefficients estimate is 6.4%. 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-coefficient and 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 '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-coefficient version of our model, the investigator may include price effects that vary with real expenditure. Our empirical work suggests that the varying-coefficient extension is both statistically significant and economically important.

7 Conclusions

We propose a cost function whose implied consumer demand system has parametric price effects and nonparametric real expenditure effects. Because the demand system is nonparametric in a single dimension, real expenditure, we avoid the curse of dimensionality typically associated with the fully nonparametric estimation of consumer demand. Our demand system may have any rank, and may be restricted to satisfy homogeneity, symmetry and concavity, which together comprise the integrability restrictions. We show $\sqrt{n}$-convergence of the parametric components and convergence rates for the nonparametric components. An application with Canadian price and expenditure data shows our method's potential.

8 Appendix: Proofs

Proof of Theorem 3:
As the local linear estimator is well studied [see e.g. Lejeune (1985) or Fan and Gijbels (1996)], we show here only how the bias and variance terms change due to the use of a nonparametrically generated regressor. As 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. Further, it will be seen in the proof that it is sufficient to do the explicit calculations for only one of the M equations.

To ease the notation, we set $Y_i := W_i^j - \boldsymbol{a}^j \boldsymbol{P}_i$ and $b := \beta^j$, $\epsilon(X_i, \boldsymbol{P}_i) := \epsilon_i^j$ 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 X_i^R we have

$$Y_i = b(x^R) + b'(x^R)\{X_i^R - x^R\} + \frac{b''(\xi_i)}{2}\{X_i^R - x^R\}^2 + \epsilon(x_i, \boldsymbol{p}_i) \ . \tag{27}$$

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

$$\begin{pmatrix} \hat{b}(x^R) \\ \hat{b}'(x^R) \end{pmatrix} = \begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \tag{28}$$

where

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

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

Combining (27) with (28) we see that for calculating bias it is sufficient to consider

$$\begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} \tag{29}$$

with vector $(M_1, M_2) :=$

$$\begin{pmatrix} 1 & \hat{X}_1^R - x^R \\ \vdots & \vdots \\ 1 & \hat{X}_n^R - x^R \end{pmatrix}^T diag\left(K_h(\hat{X}_i^R - x^R)\right)_{i=1}^n \begin{pmatrix} b'(x^R)\{X_1^R - x^R\} + b''(\xi_1)\frac{\{X_1^R - x^R\}^2}{2} \\ \vdots \\ b'(x^R)\{X_n^R - x^R\} + b''(\xi_n)\frac{\{X_n^R - x^R\}^2}{2} \end{pmatrix}.$$

We first calculate the inverse matrix in equation (28):

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

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

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

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 (29) yields the bias stated in the theorem.

For the variance one has to consider the expectation of

$$\begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \begin{pmatrix} M_1' \\ M_2' \end{pmatrix} \begin{pmatrix} M_1' \\ M_2' \end{pmatrix}^T \begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}^{-1} \tag{30}$$

with vector $(M_1', M_2') :=$

$$\begin{pmatrix} 1 & \hat{X}_1^R - x^R \\ \vdots & \vdots \\ 1 & \hat{X}_n^R - x^R \end{pmatrix}^T diag\left(K_h(\hat{X}_i^R - x^R)\right)_{i=1}^n \begin{pmatrix} b'(x^R)\{X_1^R - x^R\} + b''(\xi_1)\frac{\{X_1^R - x^R\}^2}{2} + \epsilon_1 \\ \vdots \\ b'(x^R)\{X_n^R - x^R\} + b''(\xi_n)\frac{\{X_n^R - x^R\}^2}{2} + \epsilon_n \end{pmatrix}.$$

Looking at the last vector in the definition of (M_1', M_2') it is clear that the matrix $(M_1', M_2')^T$ (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_i^R and x^R we get:

$$\begin{aligned}
E[C_{11}^1] &= E\Bigg[\sum_{j=1}^n\sum_{i=1}^n b'^2(x^R)\{B_X(X_i^R)+u_i\}\{B_X(X_j^R)+u_j\}\epsilon_i\epsilon_j \\
&\quad \left\{K_h(X_i^R-x^R)+\frac{B_X(X_i^R)+u_i}{h}K_h'(\xi_i)\right\}\left\{K_h(X_j^R-x^R)+\frac{B_X(X_j^R)+u_j}{h}K_h'(\xi_i)\right\}\Bigg] \\
&= n^2f^2(x^R)B_X^2(x^R)b'^2(X_0^R)+o(n^2B_X^2(x^R))\ ,
\end{aligned}$$

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

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

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

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

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

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

Finally, for C^4 we have

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

For more details of the calculations compare Sperlich (2005).

Plugging now this results in (30) gives the variance we have stated in the Theorem. □

Proof of Theorem 4:
Also the local linear varying coefficient estimator is already well studied, see e.g. Cleveland, Grosse and Shyu (1991), Fan and Zhang (1999) or Cai, Fan and Li (2000). The calculations to incorporate the additional bias and variance coming in by the use of a generated regressor, are basically the same as for Theorem 3. Note that now, skipping the index $j=1,\dots,M$ of W and of the functions a_k , $k=0,\dots,M$ 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 $2k+1$'th element of

$$
\left(R^TKR\right)^{-1}R^TK\boldsymbol{W}\ ,
$$

where

$$
\begin{aligned}
R &= \begin{pmatrix} P_1^0 & P_1^0(\hat{X}_1^R-x^R) & \dots & P_1^M & P_1^M(\hat{X}_1^R-x^R)\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ P_n^0 & P_n^0(\hat{X}_n^R-x^R) & \dots & P_n^M & P_n^M(\hat{X}_n^R-x^R)\end{pmatrix}\\
K &= diag\left(K_h(\hat{X}_i^R-x^R)\right)_{i=1}^{n}\ .
\end{aligned}
$$

It is clear that this leads to the same equations as in proof of Theorem 3 but now always with $P_i^kP_j^l$, $k,l=0,\dots,M$, $i,j=1,\dots,n$ 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, compare Theorem 3 of Fan and Zhang (1999). For more details of the matrix calculations we also refer to their paper. □

References

Aït-Sahalia, Y., Bickel, P. and T. Stoker, (2001), "Goodness-of-fit tests for kernel regression with an application to option implied volatilities", *Journal of Econometrics*, **105**, 363-412.

Ahn, H. and J.L. Powell, (1993), "Semiparametric estimation of censored selection models with a nonparametric selection mechanism", *Journal of Econometrics*, **58**, 3-29.

Banks, J., R. Blundell and A. Lewbel, (1997), "Quadratic Engel Curves and Consumer Demand", *Review of Economics and Statistics*, **79(4)**, 527-539.

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

Blundell, R., A. Duncan and K. Pendakur, (1998), "Semiparametric Estimation of Consumer Demand", *Journal of Applied Econometrics*, **13**, 435-461.

Browning, M. and I. Thomas, (1999), "Prices for the FAMEX", unpublished working paper, Department of Economics, McMaster University.

Browning, M., (2005), "Modelling Commodity demands and Labour Supply with M-demands", working paper.

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

Christensen, L.R., D.W. Jorgenson and L.J. Lau, (1971), "Conjugate Duality and the Transcendental Logarithmic Production Function," *Econometrica* 39, 255-256.

Cleveland, W.S., E. Grosse and W.M. Shyu, (1991), "Local Regression Models", In *Statistical Models in S* (J.M. Chambers, and T.J. Hastie, eds.) 309-376. Wadsworth/Brooks-Cole, Pacific Grove, CA.

Crossley, T. and Pendakur, K., (2005), "The social cost-of-living: welfare foundations and estimation" IFS Working Papers, W06/10.

Davidson, R. and Flachaire, E., (2005), "The Wild Bootstrap, Tamed at Last", *unpublished working paper, available at* www.econ.queensu.ca/working_papers/papers/qed_wp_1000.pdf.

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

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

Haag, B., S. Hoderlein and K. Pendakur, (2005), "Testing and Imposing Rationality Restrictions in the Nonparametric Estimation of Consumer Demand", SFU working paper.

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

Hall, P., J.W. Kay and D.M. Titterington, (1990), "Asymptotically optimal difference-based estimation of variance in nonparametric regression", *Biometrika*, **77**, 521-528.

Hengartner, N. and S. Sperlich, (2005), "Rate Optimal Estimation with the Integration Method in the Presence of Many Covariates", *Journal of Multivariate Analysis*, in press.

Hoderlein, S., (2005), "Slutsky Symmetry and Heterogeneity", unpublished working paper, University of Mannheim.

Lejeune, M., (1985), "Estimation non-parametrique par noyaux: Regression Polynomiale Mobile", *Revue de Statistique Appliqueés* XXXIII, 43-67.

Lewbel, A., (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation", *Econometrica* 59(3), pp 711-730.

Lewbel, A., (2001), "Demand Systems With and Without Error Terms", *American Economic Review*, 91(3), pp 611-618.

Lewbel, A. and K. Pendakur, (2006), "Tricks with Hicks: The EASI Demand System", *unpublished working paper, available at* www.sfu.ca/ pendakur/hixtrix22.pdf.

Newey, W., (1994), "Kernel estimation of partial means", *Econometric Theory*, **10**, 233–53.

Pendakur, K., (2001), "Consumption Poverty in Canada 1969-1998", *Canadian Public Policy*, June 2001, pp 1-34.

Pendakur, K., (2002), "Taking Prices Seriously in the Measurement of Inequality", *Journal of Public Economics*, **86(1)**, 47-69.

Politis, D., Romano, J. and Wolf, M., (1999), *Subsampling,* Springer-Verlag, New York.

Robinson, P., (1988), "Root-N-Consistent Semiparametric Regression", *Econometrica*, **56**, 931 - 954.

Rodríguez-Póo, J.M., S. Sperlich, and A.I. Fernández, (2005), “Semiparametric Three Step Estimation Methods for Simultaneous Equation Systems", *Journal of Applied Econometrics*, **20**, 699 - 721.

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

Sperlich, S., (2005), "A Note on Nonparametric Estimation with Constructed Variables and Generated Regressors", working paper

Sperlich, S., D. Tjøstheim, and L. Yang, (2002), "Nonparametric Estimation and Testing of Interaction in Additive Models", *Econometric Theory*, **18**, 197 - 251.

Yatchew, A., (1997), "An elementary estimator of the partial linear model", *Economics Letters*, **57**, 135 - 143.

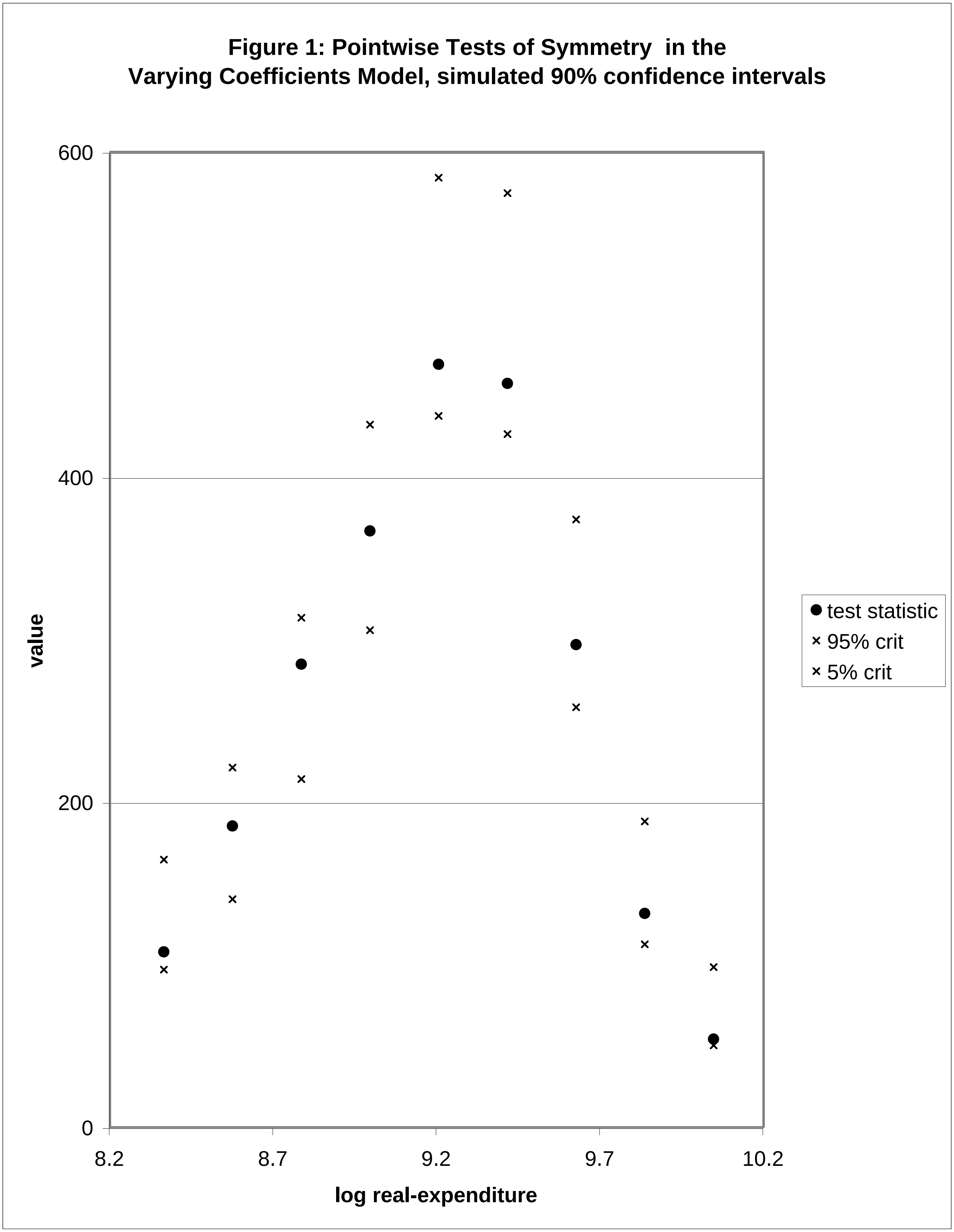


Figure 1: Pointwise Tests of Symmetry in the
Varying Coefficients Model, simulated 90% confidence intervals
600
400
200
0
value
8.2
8.7
9.2
9.7
10.2
log real-expenditure
test statistic
95% crit
5% crit

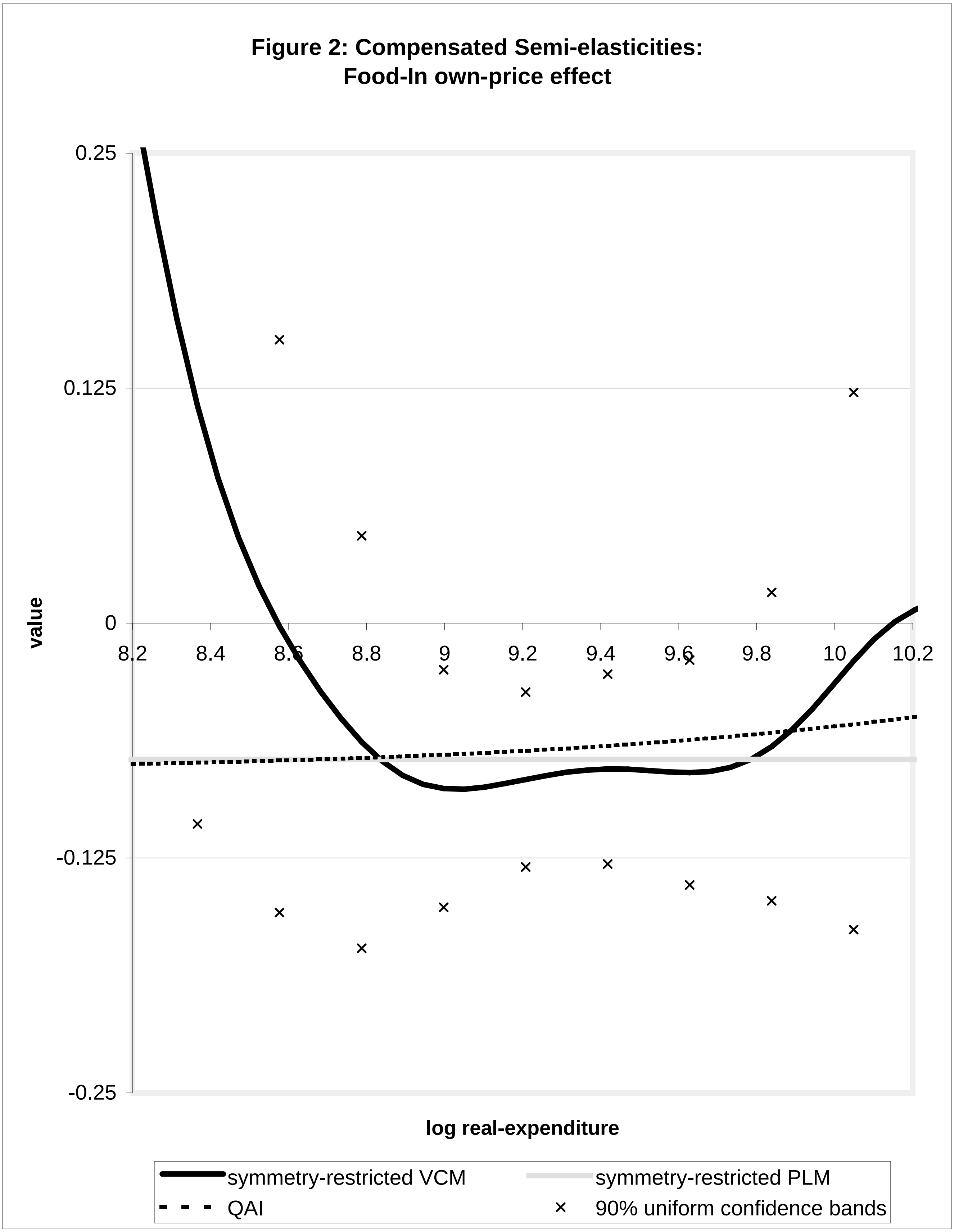


Figure 2: Compensated Semi-elasticities:
Food-In own-price effect
0.25
0.125
0
-0.125
-0.25
value
8.2
8.4
8.6
8.8
9
9.2
9.4
9.6
9.8
10
10.2
log real-expenditure
symmetry-restricted VCM
symmetry-restricted PLM
QAI
90% uniform confidence bands

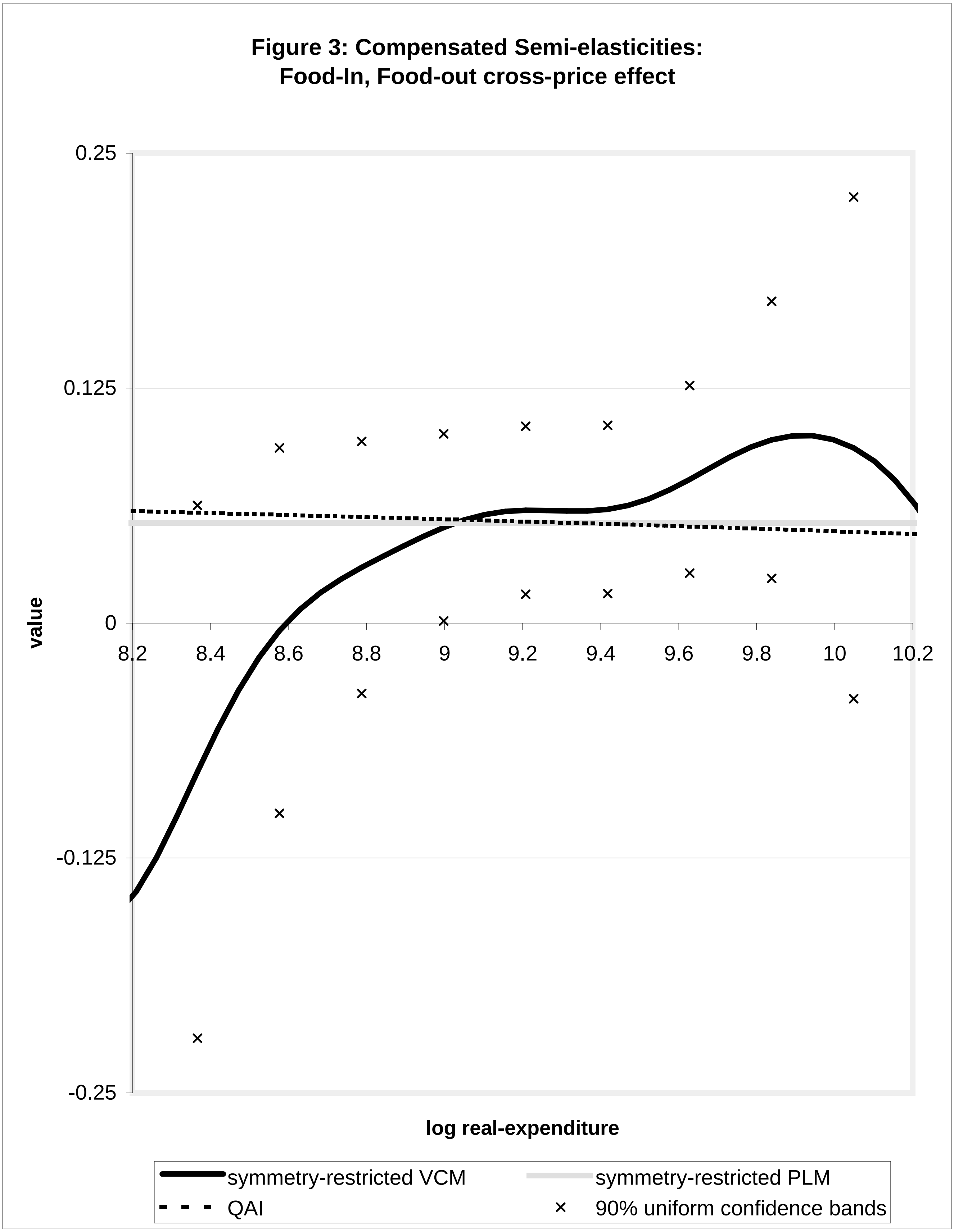


Figure 3: Compensated Semi-elasticities:
Food-In, Food-out cross-price effect
0.25
0.125
0
-0.125
-0.25
value
8.2
8.4
8.6
8.8
9
9.2
9.4
9.6
9.8
10
10.2
log real-expenditure
symmetry-restricted VCM
symmetry-restricted PLM
QAI
90% uniform confidence bands

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

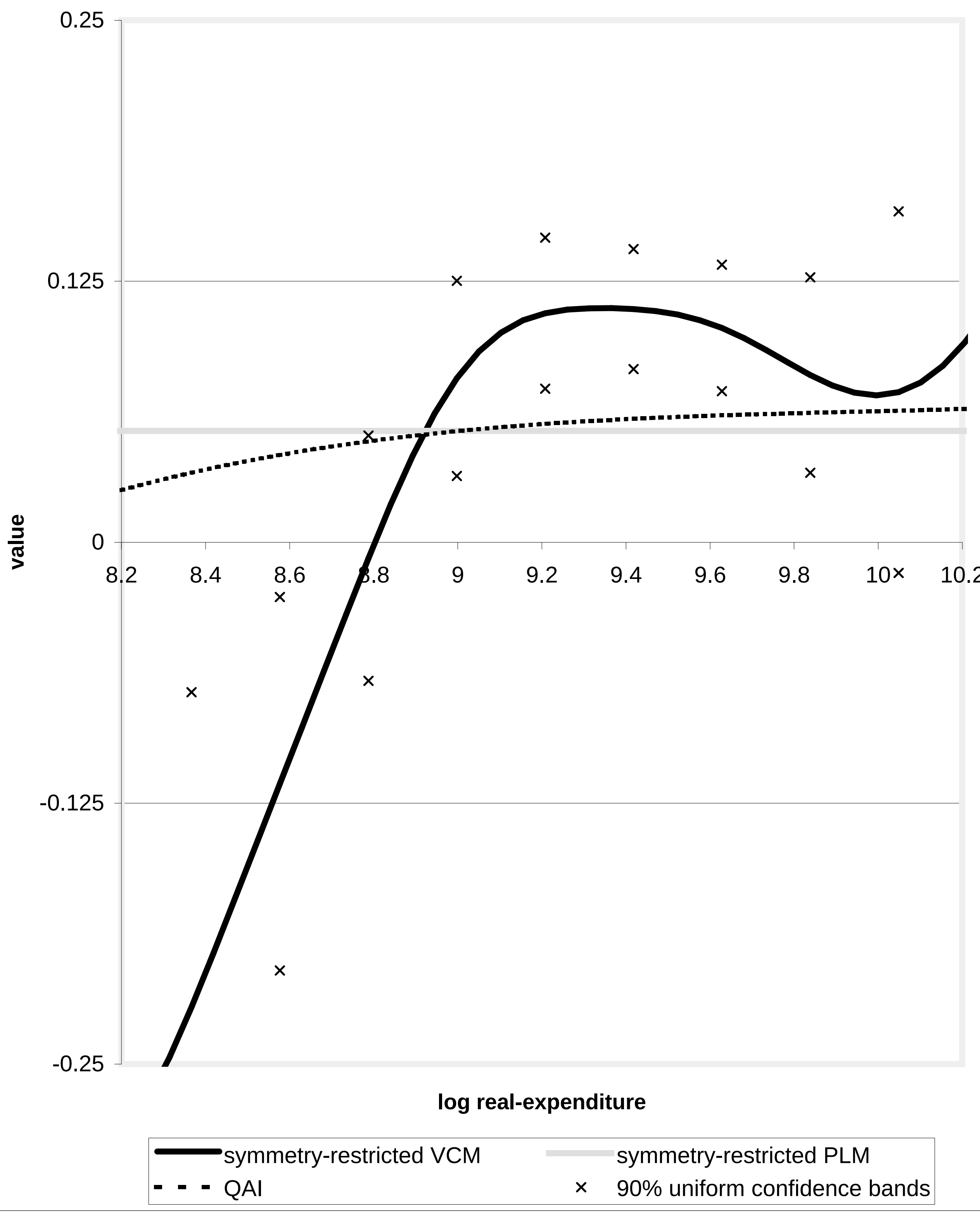


Figure 5: Estimated Food-in Shares

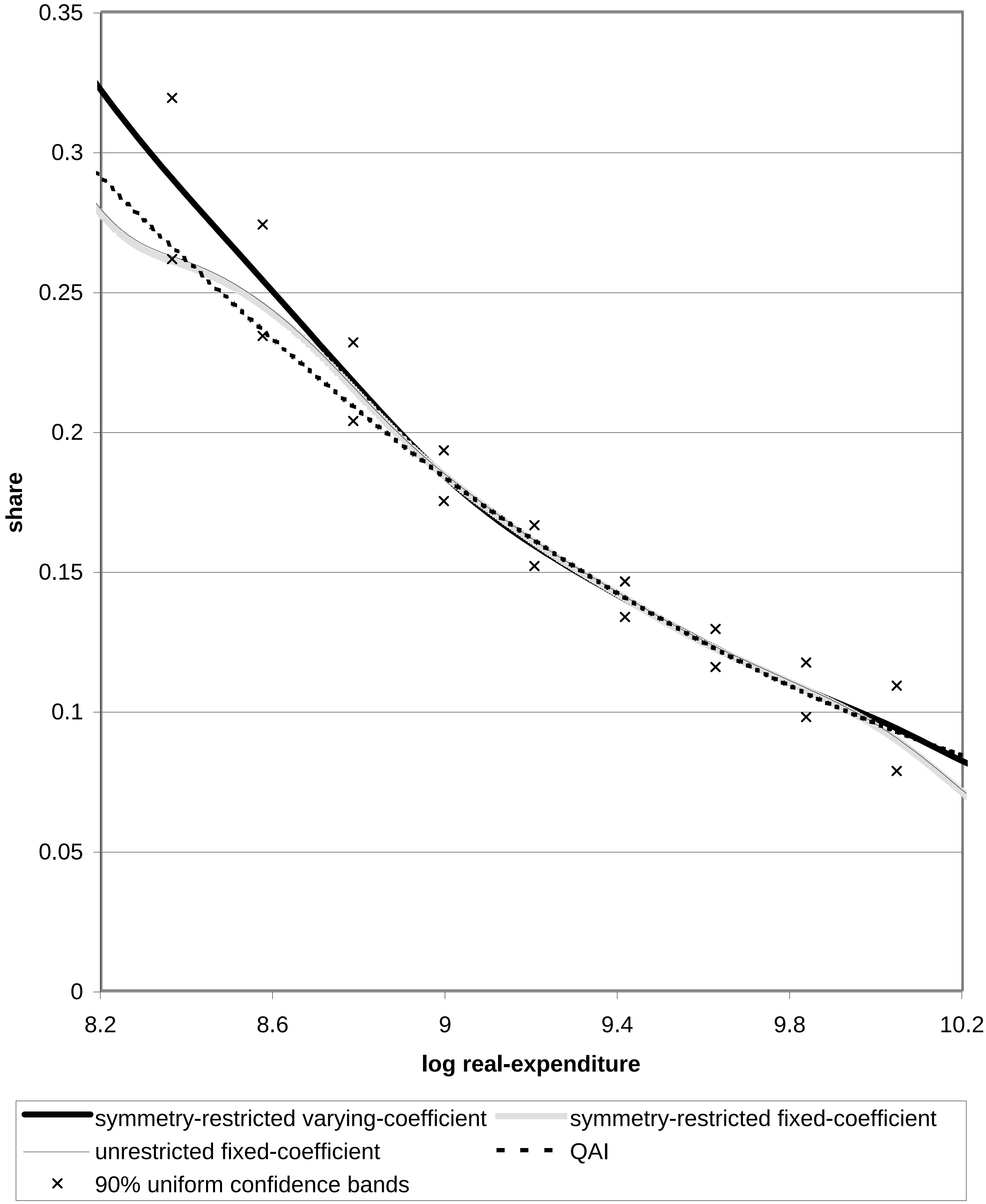
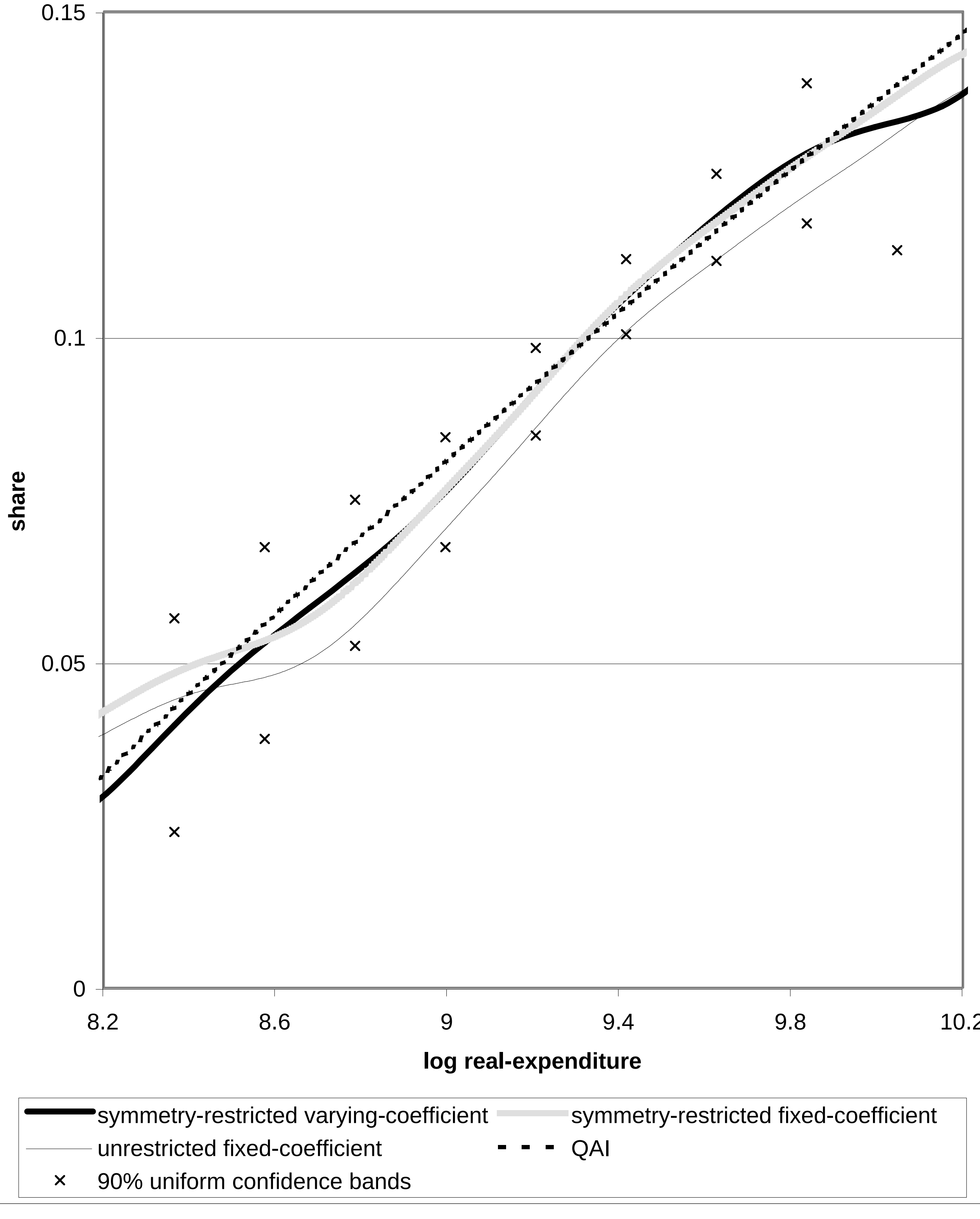


Figure 6: Estimated Food-Out Shares



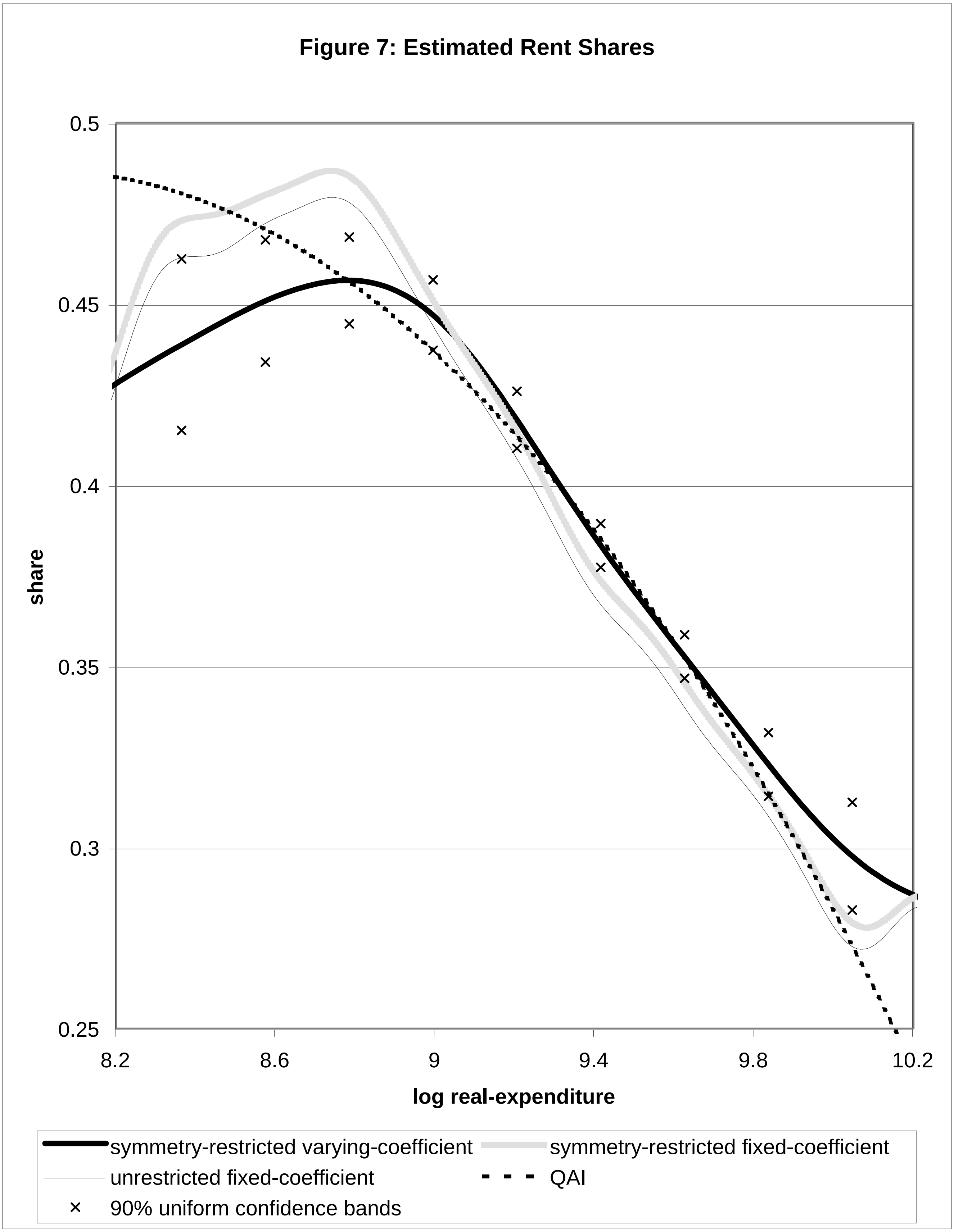


Figure 7: Estimated Rent Shares
0.5
0.45
0.4
0.35
0.3
0.25
share
8.2
8.6
9
9.4
9.8
10.2
log real-expenditure
symmetry-restricted varying-coefficient
symmetry-restricted fixed-coefficient
unrestricted fixed-coefficient
QAI
90% uniform confidence bands

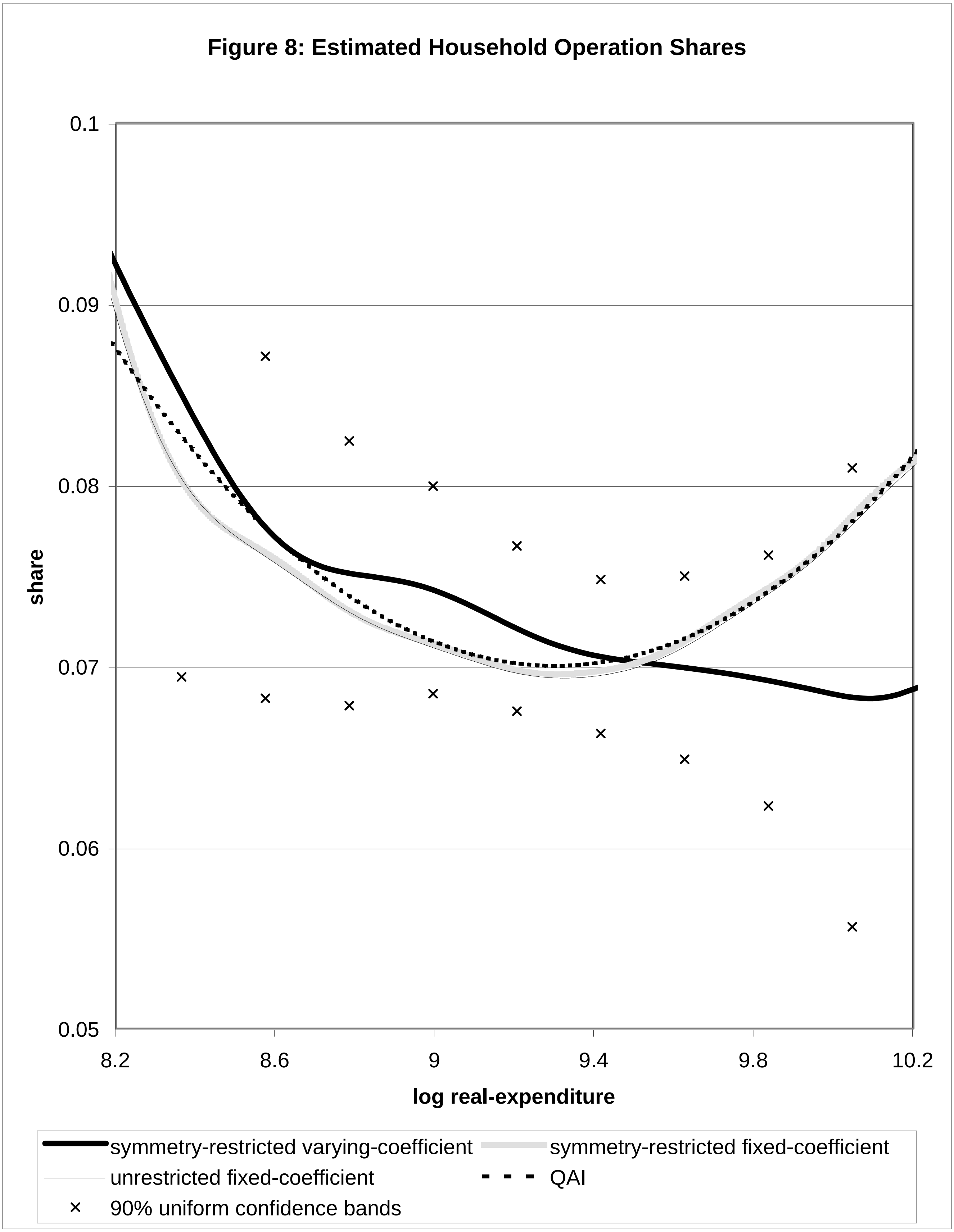


Figure 8: Estimated Household Operation Shares
0.1
0.09
0.08
0.07
0.06
0.05
share
8.2
8.6
9
9.4
9.8
10.2
log real-expenditure
symmetry-restricted varying-coefficient
symmetry-restricted fixed-coefficient
unrestricted fixed-coefficient
QAI
90% uniform confidence bands

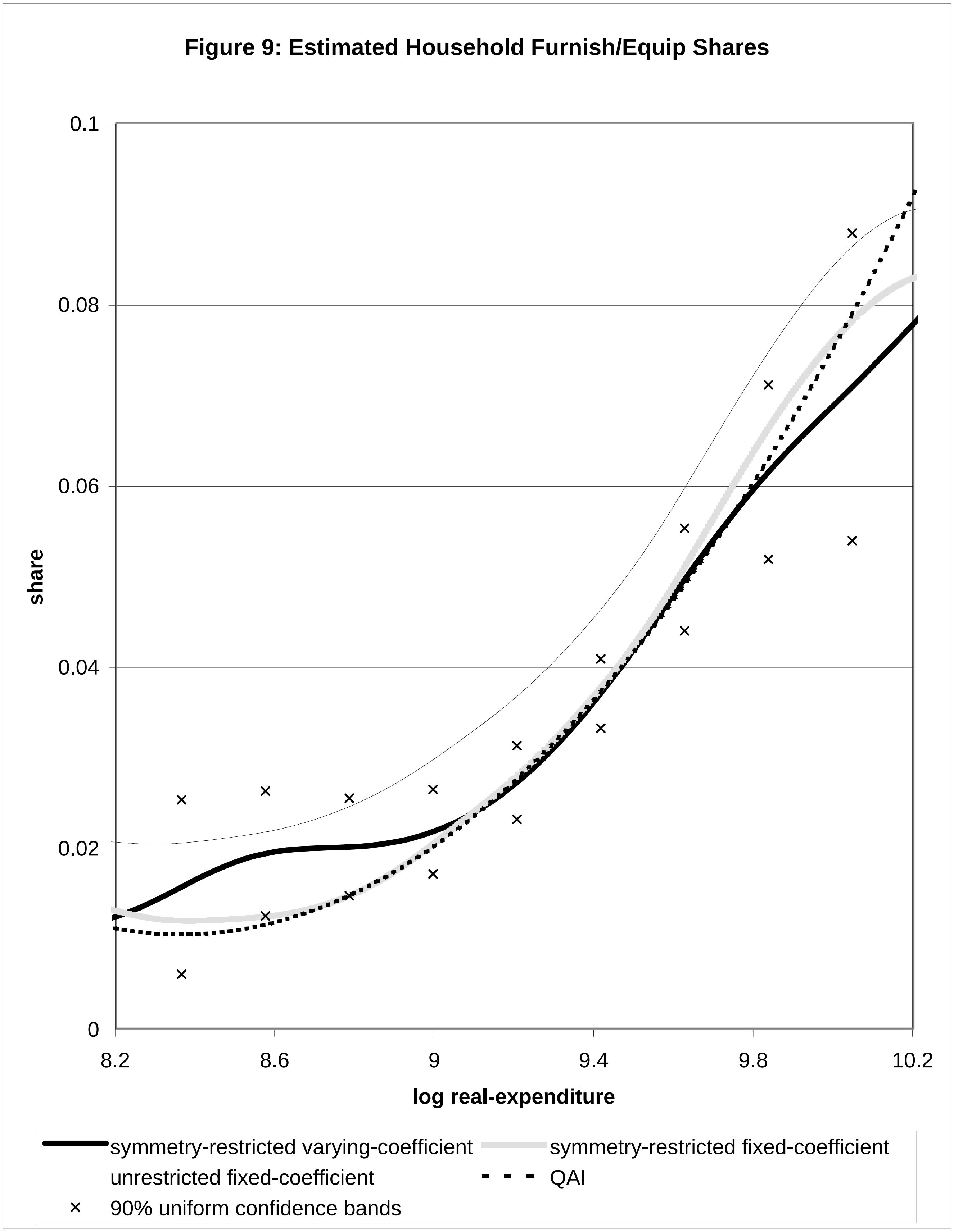


Figure 9: Estimated Household Furnish/Equip Shares
0.1
0.08
0.06
0.04
0.02
0
share
8.2
8.6
9
9.4
9.8
10.2
log real-expenditure
symmetry-restricted varying-coefficient
symmetry-restricted fixed-coefficient
unrestricted fixed-coefficient
QAI
90% uniform confidence bands

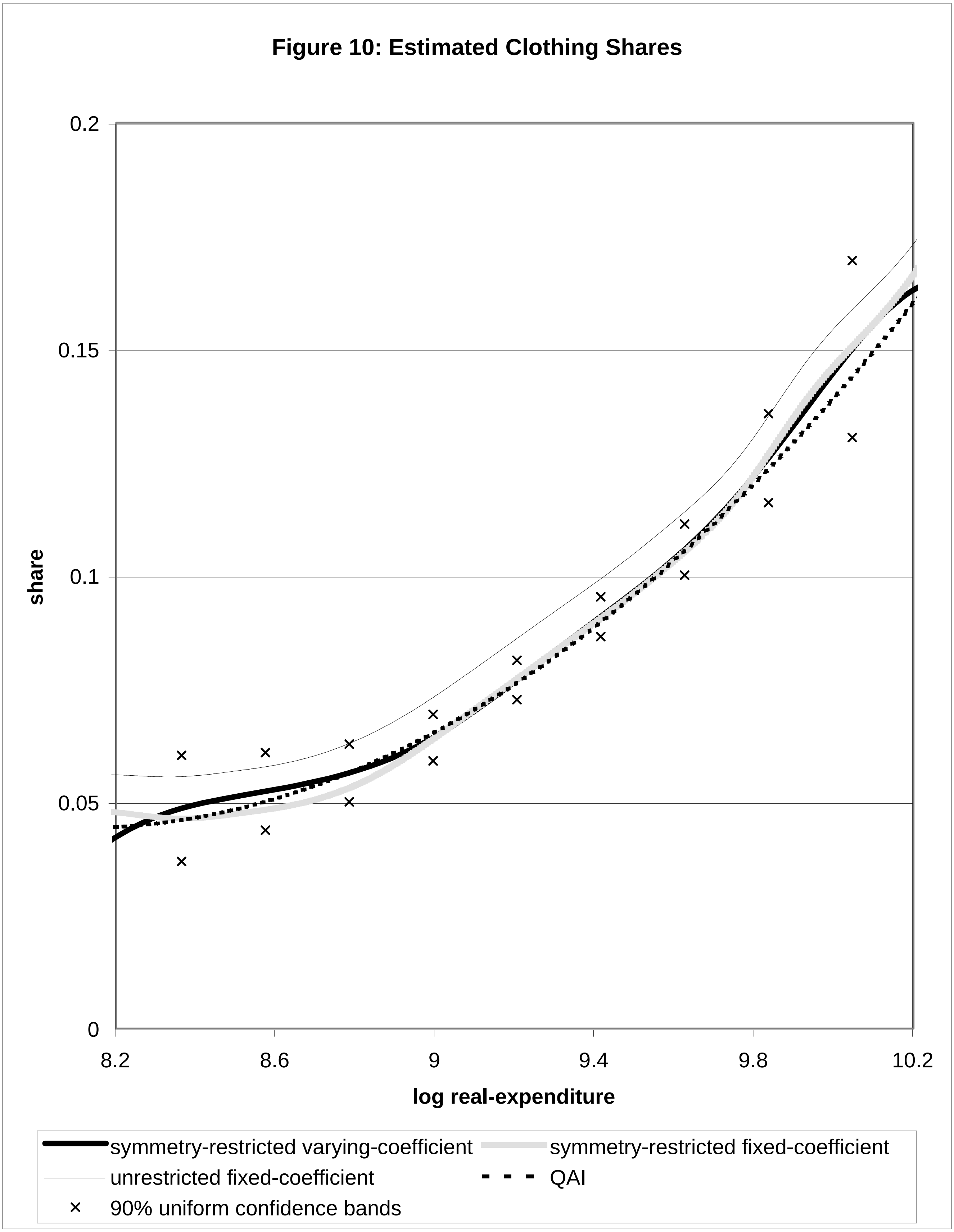
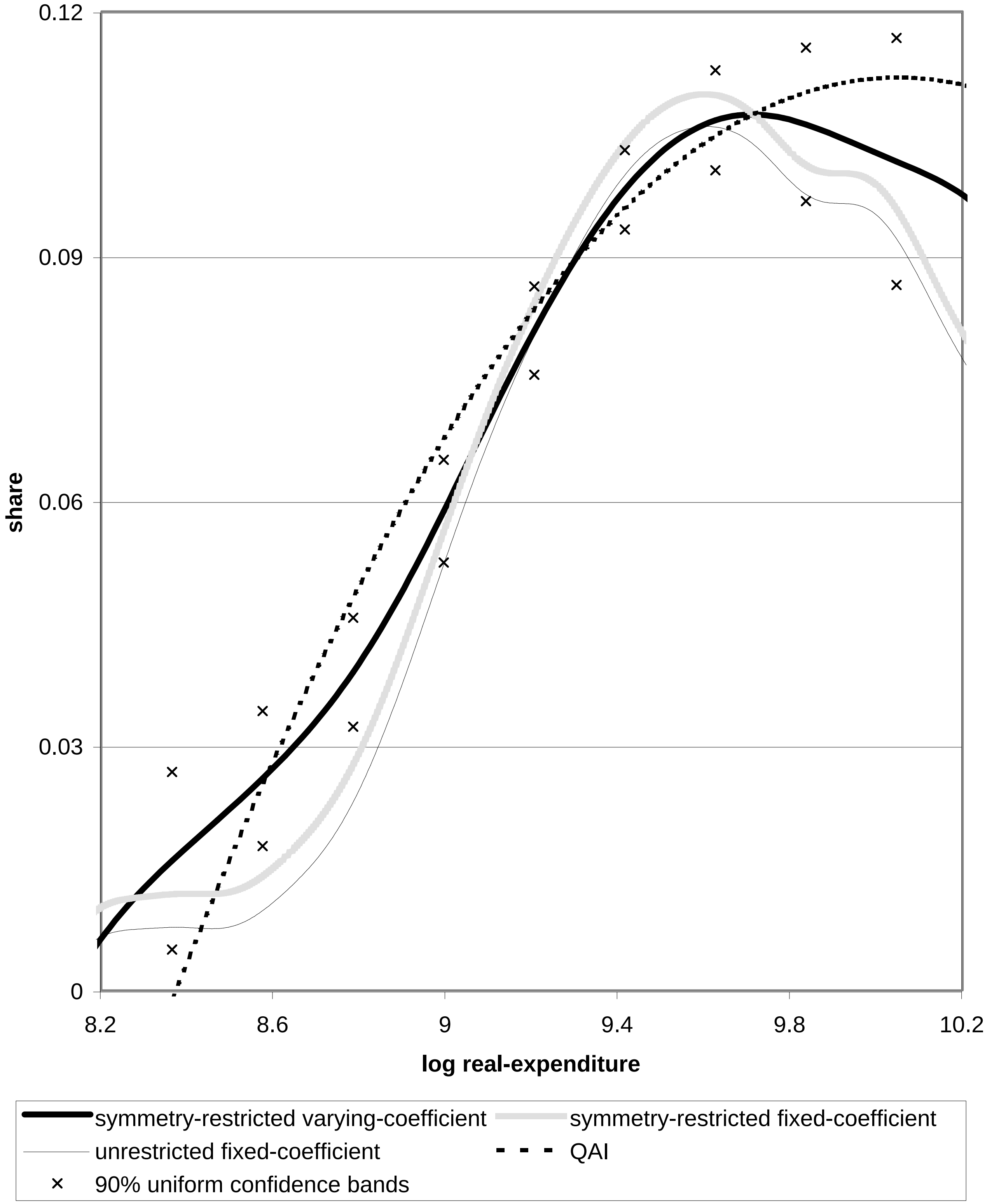


Figure 10: Estimated Clothing Shares
0.2
0.15
0.1
0.05
0
share
8.2
8.6
9
9.4
9.8
10.2
log real-expenditure
symmetry-restricted varying-coefficient
symmetry-restricted fixed-coefficient
unrestricted fixed-coefficient
QAI
90% uniform confidence bands

Figure 11: Estimated Private Transportation Shares



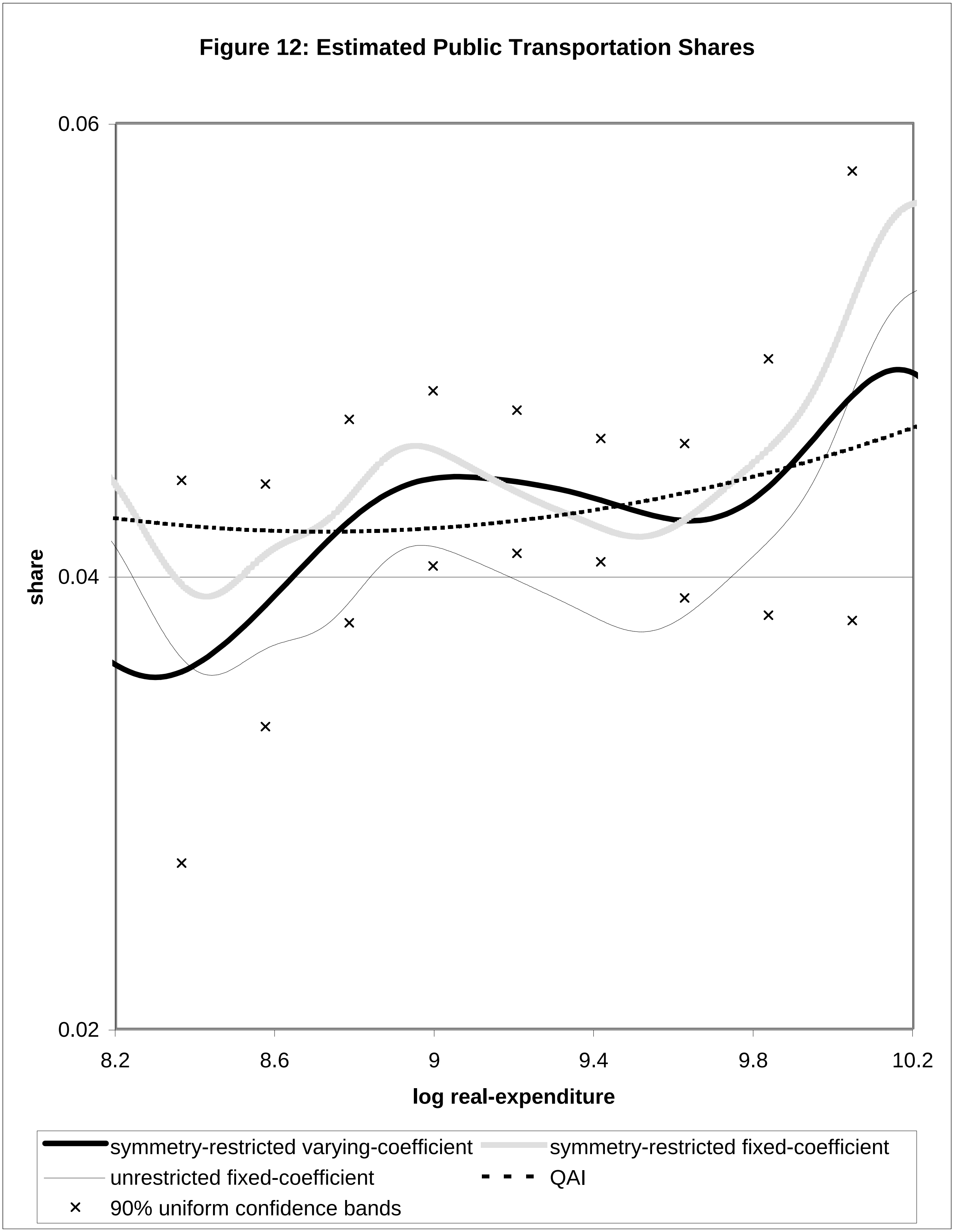


Figure 12: Estimated Public Transportation Shares
0.06
0.04
0.02
share
8.2
8.6
9
9.4
9.8
10.2
log real-expenditure
symmetry-restricted varying-coefficient
symmetry-restricted fixed-coefficient
unrestricted fixed-coefficient
QAI
90% uniform confidence bands

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