A Varying Coefficient Approach to Global Flexibility in Demand Analysis: A Semiparametric Approximation

Camilo Sarmiento

To reduce dimensionality issues, this article derives a globally flexible demand system that can be estimated non-parametrically with a specially devised temporal kernel. Statistical and economic results from a meat demand application underscores the usefulness of a temporal kernel in globally approximating an integrable demand system.

Key words: demand system, global flexibility, integrability, kernel regression, meat demand, time varying coefficient.

Much progress has been made in specifying globally flexible demand systems that model demand functions and their derivatives (elasticities) for the entire range of the data (Diewert; Gallant).[1] In some multivariate applications, the effectiveness of globally flexible approximations may, however, be limited by dimensionality problems (C.J. Stone) and possible inaccurate parameterizations of nonlinearities. The issue of dimensionality is particularly relevant in time series applications, and price variability necessary for identification of price elasticities tends to require time series rather than cross-sectional data (Jorgenson, Lau, and Stoker).

A recurrent challenge in econometric analysis is choosing a functional form. For example, Granger has recently raised the questions: "How much nonlinearity should a model contain? Should time-varying parameters be considered (perhaps as an alternative to nonlinearity)?" This article uses time varying coefficients and kernel regression in place of nonlinear parameterizations to achieve both integrability and global flexibility in demand analysis. To do so, the paper applies the Rank Theorem (see Rudin, p. 228) to model global nonlinearities of any continuously differentiable demand system in terms of time varying coefficients, which reduces the dimensionality requirement in non-parametric estimation of elasticities to the one-dimensional space.

The functional structure derived from using the Rank Theorem yields a globally flexible varying coefficient (GFVC) demand system that represents the level of the function and its derivatives (elasticities) for the entire range of the data. In this article, for estimation, a one-dimensional temporal kernel of the GFVC demand system, which captures both nonlinearities and changes in economic structure non-parametrically, is implemented. The temporal kernel of the GFVC model allows elasticities to be estimated for any differentiable demand system by modeling curves in a one-dimensional space. Global flexibility is thus achieved by using the Rank theorem and the temporal kernel.

This varying coefficient approach is applied to meat demand analysis, which previous work has shown the presence of structural breaks and nonlinearities.[2] Results from the estimated GFVC meat demand system indicate: (*a*) global flexibility in functional structure can be attained by means of time varying coefficients; and (*b*) mean squared error (MSE) with the temporal kernel estimator is lower than the MSE with traditional ways of modeling

---

Camilo Sarmiento is currently associate research scientist at the Pacific Institute for Research and Evaluation, Calverton Office Park, MD.

The author thanks Richard Just for stimulating his research in econometric modeling of functional structure, and Wade Brorsen, Rita Curtis, and three anonymous reviewers for constructive comments that significantly improved the contribution of the article.

This article was submitted and accepted for publication while the author was assistant professor at North Dakota State University.

[1] For a comprehensive analysis and empirical performance of globally flexible systems, see Piggott.

[2] For the range of notable efforts in meat demand analysis in both functional form and modeling structural breaks, see Alston and Chalfant, Moschini and Meilke, and Moschini and Moro.

varying coefficients (time trend, dummy variables). Therefore, modeling global flexibility with time varying coefficients increases model performance, and the best statistical fit under a MSE criterion occurs when the coefficient variation in the demand system is estimated non-parametrically.

Globally Flexible Demand System: A Time Varying Coefficient Approach

A Stone Index Approach

Assume Φ_{kt} represents the logarithm of consumption for commodity k; $\mathbf{P}_t = (\ln p_{1t}, \ldots, \ln p_{nt})$ where p_{it} represents the price of good i at time t and $Y_t = \ln y_t$ where y_t is income at time t. Following the variable definitions, a demand system with n-commodities and variables expressed in logarithms can be represented as:

$$(1) \qquad \mathbf{F}(\mathbf{X}_t) = [\Phi_{1t}(\mathbf{X}_t), \ldots, \Phi_{nt}(\mathbf{X}_t)]'$$

where $\mathbf{X}_t = (\mathbf{P}_t, Y_t)'$; and nonlinearities and integrability are defined in the n-dimensional space. To reduce the scope of the dimensionality requirement in the specification of integrability and global flexibility in (1), this article applies the Rank Theorem (Rudin, p. 228).

Under smoothness assumptions, the Rank Theorem defined in terms of a time dependent ω-ball generates a semi-log linear representation of the demand system in (1) using time varying coefficients. In particular, under the assumption that the demand equations in (1) are continuously differentiable, by the Rank Theorem, there exists an open set $\mathrm{E}_t \subset \Re^n$ such that (1) can be represented as:

$$(2) \qquad \mathbf{F}(\mathbf{X}_t) = \mathbf{A}_t\mathbf{X}_t + \boldsymbol{\psi}(\mathbf{A}_t\mathbf{X}_t)$$

where

$$\mathbf{A}_t = \begin{bmatrix} a_{11t} & a_{12t} & \ldots & a_{1nt} & a_{1n+1t} \\ & & & \vdots & \\ a_{n-1,1t} & a_{n-1,2t} & \ldots & a_{n-1,nt} & a_{n-1,n+1t} \\ a_{n1t} & a_{n2t} & \ldots & a_{nnt} & a_{nn+1t} \end{bmatrix};$$

$$a_{ijt} = \partial\Phi_i/\partial\ln p_{jt}, \quad \text{for } j \leq n;$$
$$a_{\text{in+1t}} = \partial\Phi_i/\partial\ln y_t;$$

$\boldsymbol{\psi}(\mathbf{A}_t\mathbf{X}_t) = 0$ only if the rank of $\mathbf{A}_t$ equals n (i.e., full rank);

and

$$\mathbf{X}_t = (\mathbf{P}_t, Y_t)' \in \omega_t \subset \mathrm{E}_t.$$

Therefore, there exists a time dependent ω_t-ball (i.e., time dependent neighborhood) such that any continuously differentiable demand system is exactly represented by the time varying coefficient model in (2);[3] and the system is globally flexible (Diewert, Gallant, Piggott) since it represents the function and its derivatives (elasticities) for the entire range of the data. Time dependent local neighborhoods (i.e., ω_t-balls) in the application of Rank Theorem then yield an exact globally flexible functional form in demand with time varying coefficients, as opposed to the nonlinear form with respect to all explanatory variables in (1).

The definition of the variables in logarithms in (2), furthermore, implies (from the Rank Theorem) that time varying coefficients in $\mathbf{A}_t$ correspond to elasticities and, hence, integrability conditions (symmetry, homogeneity, and summability) in $\mathbf{A}_t$ can be imposed following Stone's seminal work. Specifically, the property of symmetry of the Slutsky matrix implies that

$$(3a) \qquad \Phi^*_{it}\big(a_{ijt} + \Phi^*_{jt}a_{in+1t}\big) = \Phi^*_{jt}\big(a_{jit} + \Phi^*_{it}a_{jn+1t}\big)$$

where the share of commodity j in total consumption is defined as Φ^*_{jt}; and the Stone index $\sum_j \Phi^*_{jt} \ln p_{jt}$ emerges in each equation when (3a) is incorporated in (2). Furthermore, the property of homogeneity of degree zero of Marshallian demands implies that:

$$(3b) \qquad \sum_j a_{ijt} = 0$$

which can be imposed by choosing a numeraire p_{nh}, while the adding up property implies that:

$$(3c) \qquad a_{nn+1t} = 1/\Phi^*_{nt} - \sum_{i \neq n} \big(\Phi^*_{it}/\Phi^*_{nt}\big)a_{in+1t}.$$

Any differentiable and integrable demand system is then exactly represented by (2) with the coefficient restrictions in (3).

An Almost Ideal Demand Index Approach

The structure of the Stone index while convenient to incorporate integrability conditions in (2) has also been a subject of much criticism

[3] The Rank Theorem applies under the same assumption of the implicit function theorem, that is, the ω-ball in which demand equations are explicitly defined. This underscores the generality of the Rank Theorem.

in demand analysis. For example, Alston and Green demonstrate limitations of the Stone index to incorporate integrability conditions in terms of evaluation of derivatives because the differential of the function does not take into account the derivative of shares in the Stone index. To derive a theoretically consistent globally flexible demand system with time varying coefficients that does not depend on the Stone index for specification of integrability, consider a demand system that uses shares Φ^*_{kt} as the dependent variable and explanatory variables in logarithms:

$$(4) \qquad \mathbf{F}^*(\mathbf{X}_t) = \left[\Phi^*_{1t}(\mathrm{X}^*_t), \ldots, \Phi^*_{nt}(\mathbf{X}^*_t)\right]'$$

where $\mathrm{X}^*_t = (\mathbf{P}_t, \mathrm{Y}_t - \underline{P}_t)'$ and the logarithm of income Y_t is normalized by $\underline{P}_t$, an arbitrary quadratic function with respect to $\mathbf{P}_t$:

$$\underline{P}_t = \sum_i \varphi_{it} \ln p_{it} + 1/2 \sum_i \sum_j \beta^*_{ijt} \ln p_{it} \ln p_{jt}.$$

This normalization of the income variable facilitates imposition of integrability conditions, while it does not restrict generality of the demand system representation.

As in (2), application of the Rank Theorem implies that there exists an open set $E_t \subset \Re^n$ such that (4) can be represented as:

$$(5) \qquad \mathbf{F}^*(\mathbf{X}^*_t) = \mathbf{A}^*_t\mathbf{X}^*_t + \boldsymbol{\psi}(\mathbf{A}^*_t\mathbf{X}^*_t)$$

where

$$\mathbf{A}^*_t = \begin{bmatrix} a^*_{11t} & a^*_{12t} & \ldots & a^*_{1nt} & a^*_{1n+1t} \\ & & & & \vdots \\ a^*_{n-1,1t} & a^*_{n-1,2t} & \ldots & a^*_{n-1,nt} & a^*_{n-1,n+1t} \\ a^*_{n1t} & a^*_{n2t} & \ldots & a^*_{nnt} & a^*_{nn+1t} \end{bmatrix}$$

$$a^*_{ijt} = \partial\Phi^*_i/\partial \ln p_{jt}, \quad \text{for } j \leq n$$

$$a^*_{in+1t} = \partial\Phi^*_i/\partial \ln y_t = -\partial\Phi^*_i/\partial\underline{P}_t$$

and

$$\mathbf{X}^*_t \in \omega_t \subset \mathrm{E}_t.$$

Moreover, since $\underline{P}_t$ is a function of $\mathbf{P}_t$ it follows that:

$$a^*_{ijt} = \partial\Phi^{d*}_i/\partial \ln p_{jt} + \left(\partial\Phi^*_i/\partial\underline{P}_t\right) \times (\partial\underline{P}_t/\partial \ln p_{jt}), \quad \text{for } j \leq n;$$

$$\mathrm{a}^*_{\mathrm{in}+1t} = \partial\Phi^*_i/\partial \ln \mathrm{y}_t$$

which implies that

$$(6a) \qquad a^*_{ijt} = \beta_{ijt} - \gamma_{it}\left[\varphi_{jt} + \sum_i \beta^*_{ijt} \ln p_{it}\right] \quad \text{for} \quad j \leq n$$

$$(6b) \qquad a^*_{\mathrm{in}+1t} = \gamma_{it}$$

where $\beta_{ijt} = \partial\Phi^{d*}_i/\partial\ln p_{jt}$ and $\gamma_{it} = \partial\Phi^*_i/\partial \ln y_t$.

Therefore, there exists a time dependent ω_t-ball such that any continuously differentiable demand system is exactly represented by (5) and its derivatives in (6). Indeed, the structure in (6) exactly represents the derivatives of (4) at time t under each ω_t-ball. A global representation of derivatives is thus achieved by allowing coefficients in (6) to vary across time, while a time varying intercept can be used to represent the one-dimensional nonlinear function, $\boldsymbol{\psi}(\mathbf{A}^*_t\mathbf{X}^*_t)$, i.e.,

$$(7) \qquad \mathbf{H}_t = \boldsymbol{\psi}(\mathbf{A}^*_t\mathbf{X}^*_t).$$

The varying coefficients in (6) in (7) are, however, consistent with an integrable demand system (symmetry, summability, and homogeneity) in (4) only if

$$\beta_{it} = \beta^*_{it};$$

$$\beta_{ijt} = \beta_{jit}; \sum_i \varphi_{it} = 1; \sum_j \beta_{ijt} = 0;$$

$$\sum_i \gamma_{it} = 0$$

and

$$\mathbf{H}_t = (\varphi_{1t}, \ldots, \varphi_{n-1,t}, \varphi_{nt})'$$

where any differentiable and integrable demand system can be modeled using these appropriate coefficient restrictions; and system (5) with integrability conditions entails the structure of the almost ideal demand system (Deaton and Muellbauer).

An alternative representation of (5) that includes symmetry, summability, and homogeneity conditions is, therefore,

$$(8) \qquad \mathbf{F}(\mathbf{X}_t) = \mathbf{Z}_t \cdot \Gamma_t + (\mathrm{Y}_t - \underline{P}_t)\gamma_t + \boldsymbol{\psi}[\mathbf{Z}_t \cdot \Gamma_t + (\mathrm{Y}_t - \underline{P}_t)\gamma_t]$$

where

$$\mathbf{Z}_t = [\mathbf{Z}_{1t} \ldots \mathbf{Z}_{n-1t}]$$

with sub-matrices:

$$\mathbf{Z}_{1t} = \begin{bmatrix} \ln(p_{1t}/p_{nt}) & \ln(p_{2t}/p_{nt}) & \ldots & \ln(p_{n-1t}/p_{nt}) \\ 0 & \ln(p_{1t}/p_{nt}) & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \ldots & \ln(p_{1t}/p_{nt}) \\ -\ln(p_{1t}/p_{nt}) & -\ln\left(p_{2t}p_{1t}\big/p_{nt}^2\right) & \ldots & -\ln\left(p_{n-1t}p_{1t}\big/p_{nt}^2\right) \end{bmatrix};$$

and

$$\mathbf{Z}_{jt} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \vdots & & & \vdots \\ \ln(p_{jt}/p_{nt}) & \ln(p_{j+1t}/p_{nt}) & \ldots & \ln(p_{n-1t}/p_{nt}) \\ 0 & \ln(p_{jt}/p_{nt}) & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \ldots & \ln(p_{jt}/p_{nt}) \\ -\ln(p_{jt}/p_{nt}) & -\ln\left(p_{j+1t}p_{jt}\big/p_{nt}^2\right) & \ldots & -\ln\left(p_{n-1t}p_{jt}\big/p_{nt}^2\right) \end{bmatrix};$$

and, the price index in (4) is now structured as:

$$\underline{P}_t = \underline{P}_t(\mathbf{H}_t, \Gamma_t) = \sum_i \varphi_{it} \ln(p_{it}/p_{nt}) + 1/2 \sum_i \sum_j \beta_{ijt} \ln(p_{it}/p_{nt}) \ln(p_{jt}/p_{nt});$$

with parameter vectors defined by:

$$\Gamma_t = (\beta_{11t}, \ldots, \beta_{1n-1t}, \beta_{22t}, \ldots, \beta_{2n-1t}, \ldots, \beta_{j-1,n-1t}, \beta_{jjt}, \ldots, \beta_{jn-1t}, \ldots, \beta_{n-1,n-1t})'$$

$$\gamma_t = \left(\gamma_{1t}, \ldots, \gamma_{n-1,t}, -\sum_i \gamma_{it}\right)';$$

and

$$\boldsymbol{\psi} = \mathbf{H}_t = \left(\varphi_{1t}, \ldots, \varphi_{n-1,t}, 1 - \sum_i \varphi_{it}\right)'.$$

The derivatives and integrability condition in (4) can then be globally represented (for the entire range of the data) using the time varying parameter demand system:

$$\text{(9)} \quad \Lambda_t = \mathbf{Z}_t \cdot \Gamma_t + [\mathrm{Y}_t - \underline{P}_t(\mathbf{H}_t, \Gamma_t)]\gamma_t + \mathbf{H}_t$$

where the system in (9) is referred as a globally flexible varying coefficient (GFVC) demand system that only requires modeling curves in a one-dimensional space to capture global nonlinearities of the demand system. This dimensionality reduction in (9), however, implies that the GFVC is not best for separating changes in economic structure from changes in variables since they are jointly modeled in (9).

In the application of (9), the caveat is that the variation of coefficients (e.g., Γ_t) across observations needs to be properly captured.

A Kernel Regression of a Globally Flexible Demand System

Econometric Model

A statistical model of the GFVC demand system is obtained by adding a residual. In particular, the econometric model of the GFVC demand system, in which the conditional expectation component of the model stems from the Rank Theorem, is

$$\text{(10)} \quad \mathbf{Q}_t = [Q_{1t} \ldots Q_{nt}]' = \mathbf{F}_t + \mathbf{U}_t$$

where

$$Q_{itjt} = \Lambda_{jt} + \varepsilon_{jt}$$

and

$$\Lambda_{jt} = z_{jt}\Gamma_t + (Y_t - \underline{P}_t)\gamma_{jt} + h_{jt}$$

where $\boldsymbol{z}_{jt}$ and h_{jt} are the jth rows of $\mathbf{Z}_t$ and $\mathbf{H}_t$, respectively, and ε_{jt} is the jth row of the econometric residual $\mathbf{U}_t$. To avoid ad hoc parameterizations in the estimation of elasticities, however, a non-parametric estimator of (10) needs to be specified.

A Temporal Weighted Estimator: A Kernel Approach

To model nonparametrically the temporal variation of elasticities in the GFVC demand system in (10), an estimator is specified following C.J. Stone and Cleveland criterion of kernel regression defined in terms of minimizing a weighted sum of squares of the errors.[4] In particular, a kernel estimator of the GFVC demand system in (10) evaluated at coefficients in observation q is

$$\text{(11)}\quad \underset{\{\Gamma_q, \mathrm{H}_q, \gamma_q\}}{\mathrm{Min}} = \sum_{t=1}^{M} \left\{ k\left[\frac{d(t,q)}{h}\right] \times (\mathbf{Q}_t - \mathbf{Z}_t\Gamma_q - (\mathrm{Y}_t - \underline{P}_{tq})\gamma_q - \mathbf{H}_q)' (\mathbf{Q}_t - \mathbf{Z}_t\Gamma_q - (\mathrm{Y}_t - \underline{P}_{tq})\gamma_q - \mathbf{H}_q) \right\}$$

$$\text{for } \underline{P}_{tq} = \underline{P}_t(\mathbf{H}_q, \Gamma_q)$$

where the distance between observation t and q is $d(t, q)$ (e.g., the time length $|t - q|$); and the bandwidth length h determines the size of the temporal weights allotted to each observation, while the density function k of the kernel captures the shape of the weights. For example, a Gaussian weight function for the kernel in (11) evaluated at time q is:

$$\text{(12)}\quad k\left(\frac{d(t,q)}{h}\right) = (2\pi)^{-1/2} \times \exp\left[-\left(\frac{1}{2}\right)\left(\frac{d_{tq}}{h}\right)^2\right]$$

where $d_{tq} = d(t, q)$.

Now, re-defining $k(\frac{d(t,q)}{h}) = \beta_{tqh}^2$, a nonparametric estimator of time varying coefficients around a reference point q that minimizes (11) is, therefore, the nonlinear least squares estimator of the system:

$$\text{(13)}\quad \begin{aligned} &\beta_{1qh}\mathbf{Q}_1 = \beta_{1qh}(\mathbf{Z}_1\Gamma_q + (Y_1 - \underline{P}_{1q})\boldsymbol{\gamma}_q + \mathbf{H}_q) \\ &\beta_{2qh}\mathbf{Q}_2 = \beta_{2qh}(\mathbf{Z}_2\Gamma_q + (Y_2 - \underline{P}_{2q})\boldsymbol{\gamma}_q + \mathbf{H}_q) \\ &\beta_{Mqh}\mathbf{Q}_M = \beta_{Mqh}[\mathbf{Z}_M\Gamma_q + (Y_M - \underline{P}_{Mq})\boldsymbol{\gamma}_q + \mathbf{H}_q] \end{aligned}$$

where, from the definition of the nonlinear least squares estimator, the temporal kernel estimator of the GFVC demand system in (10) then solves:

$$\text{(14a)}\quad \mathbf{R}_q'\mathbf{E}_q = 0$$

$$\text{(14b)}\quad (\mathbf{E}_q'\mathbf{E}_q)/n\left(\sum_t \beta_{tqh}^2\right) = \sigma_q^2$$

where $\mathbf{R}_q$ is the matrix of the gradient of (13) for all observations with respect to all parameters in the model:[5]

$$\mathbf{R}_q = \begin{bmatrix} \beta_{1qh}\nabla_{1q} \\ \beta_{2qh}\nabla_{2q} \\ \vdots \\ \beta_{Mqh}\nabla_{Mq} \end{bmatrix}$$

where

$$\begin{aligned} \nabla_{tq} &= [\partial\mathbf{Q}_i/\partial\Gamma_q : \partial\mathbf{Q}_i/\partial\gamma_q : \partial\mathbf{Q}_i/\partial\mathbf{H}_q] \\ &= [\mathbf{Z}_i - \gamma_q\partial\underline{P}_{tq}/\partial\Gamma_q : (Y_t - \underline{P}_{tq})i : i - \gamma_q\partial\underline{P}_{tq}/\partial\mathbf{H}_q] \end{aligned}$$

where i is an $n \times 1$ vector of ones, and the vector E_q is the regression error:

$$\mathrm{E}_q = \begin{bmatrix} \beta_{1qh}\mathbf{Q}_1 - \beta_{1qh}(\mathbf{Z}_1\Gamma_q + \gamma_q(\mathrm{Y}_1 - \underline{P}_{1q}) + \mathbf{H}_q) \\ \beta_{2qh}\mathbf{Q}_2 - \beta_{2qh}(\mathbf{Z}_2\Gamma_q + \gamma_q(\mathrm{Y}_1 - \underline{P}_{1q}) + \mathbf{H}_q) \\ \vdots \\ \beta_{Mqh}\mathbf{Q}_M - \beta_{Mqh}(\mathbf{Z}_M\Gamma_q + \gamma_q(\mathrm{Y}_1 - \underline{P}_{1q}) + \mathbf{H}_q) \end{bmatrix}.$$

A global approximation of derivatives results from evaluating (14) at different time periods in the data set. In the case of nonspherical disturbances, however, the variance–covariance matrix of the estimator in (14) may not be adequate for inference testing. Traditionally, demand system estimation incorporates the presence of contemporaneous serial correlation across equations, $\mathrm{Cov}(\varepsilon_{jt}, \varepsilon_{kt}) = \sigma_{jk} \neq 0$, by using the seemingly unrelated regression (SUR) estimator. A kernel estimator of the GFVC demand system that filters out

[4] See also Fan (1992, 1993) and Ruppert and Wand (1994).

[5] In a linear in parameter model, for example, Stone model, R_q corresponds to the matrix of explanatory variables.

contemporaneous serial correlation is considered next.

Kernel with Contemporaneous Serial Correlation

From the temporal-weighted least-squares representation of the kernel estimator in (14), a filter of non-spherical disturbances can be incorporated similarly to the case of the non-linear least squares estimator. In particular, the kernel estimator of the GFVC demand system with a filter for contemporaneous serial correlation is the SUR estimator of (13), i.e.,

$$\text{(15a)}\quad \tilde{\mathbf{R}}_q'(\boldsymbol{\Omega}_q \otimes \mathrm{I}_m)^{-1}\tilde{\mathbf{E}}_q = 0$$

$$\text{(15b)}\quad \boldsymbol{\Omega}_q = \left(\sum\nolimits_t \beta_{tqh}^2\right)^{-1} \times \begin{bmatrix} \mathbf{E}_{1q}{}'\mathbf{E}_{1q} \ldots \mathbf{E}_{nq}'\mathbf{E}_{1q} \\ \vdots \\ \mathbf{E}_{nq}'\mathbf{E}_{1q} \ldots \mathbf{E}_{nq}'\mathbf{E}_{nq} \end{bmatrix}$$

where

$$\tilde{\mathbf{R}}_q = \begin{bmatrix} \mathbf{R}_{1q} \\ \mathbf{R}_{2q} \\ \vdots \\ \mathbf{R}_{nq} \end{bmatrix}, \quad \text{for} \quad \mathbf{R}_{jq} = \begin{bmatrix} \beta_{1qh}\nabla_{j1q} \\ \beta_{2qh}\nabla_{j2q} \\ \vdots \\ \beta_{Mqh}\nabla_{jMq} \end{bmatrix}$$

where ∇_{jtq} is the gradient of equation j at time t with respect to parameters of the model at reference point q:

$$\nabla_{jtq} = [\partial Q_{jt}/\partial \Gamma_q : \partial Q_{jt}/\partial \gamma_{jq} : \partial Q_{jt}/\partial h_{jq}] = [\mathbf{z}_{ji} - \gamma_{jq}\partial \underline{P}_{tq}/\partial \Gamma_q : Y_t - \underline{P}_{tq} : 1 - \gamma_{jq}\partial \underline{P}_{tq}/\partial h_{jq}]$$

and $\tilde{\mathbf{E}}_q$ is the regression error,

$$\tilde{\mathbf{E}}_q = \begin{bmatrix} \mathbf{E}_{1q} \\ \mathbf{E}_{2q} \\ \vdots \\ \mathbf{E}_{nq} \end{bmatrix} \quad \text{for} \quad \mathbf{E}_{jq} = \begin{bmatrix} \beta_{1qh}\mathrm{Q}_{j1} - \beta_{1qh}(\mathbf{z}_{j1}\Gamma_q + \gamma_{jq}(\mathrm{Y}_1 - \underline{P}_{1q}) + h_{jq}) \\ \beta_{2qh}\mathrm{Q}_{j2} - \beta_{2qh}(\mathbf{z}_{j2}\Gamma_q + \gamma_{jq}(\mathrm{Y}_2 - \underline{P}_{2q}) + h_{jq}) \\ \vdots \\ \beta_{Mqh}\mathrm{Q}_{jM} - \beta_{Mqh}(\mathbf{z}_{jM}\Gamma_q + \gamma_{jq}(\mathrm{Y}_M - \underline{P}_{Mq}) + h_{jq}) \end{bmatrix}.$$

And the variance–covariance matrix associated with the kernel in (15) is:

$$\mathrm{Cov}(\hat{\Gamma}_q, \hat{\gamma}_q, \hat{\mathbf{H}}_q) = \tilde{\mathbf{R}}_q'(\boldsymbol{\Omega}_q \otimes \mathbf{I})^{-1}\tilde{\mathbf{R}}_q$$

where the standard errors are conditional on the estimated bandwidth.[6]

Yet, the bandwidth length cannot be simultaneously estimated with the other coefficients in (15) since in that case $h \to 0$. Because of this difficulty, the criterion of *cross validation* is often used to estimate the bandwidth (e.g., see Engle et al. 1986, Schmalensee et al. 1999). Cross validation, a mean squared error criterion, is frequently implemented by minimizing the estimated prediction error:

$$\text{(16)}\quad \underset{h}{\mathrm{Min}}\, J(h) = (M \cdot n)^{-1} \times \sum_{q=1}^{M}(\mathbf{Q}_q - \hat{\hat{\mathbf{Q}}}_q)'(\mathbf{Q}_q - \hat{\hat{\mathbf{Q}}}_q)$$

where $\hat{\hat{\mathbf{Q}}}_q = \mathbf{Z}_q\hat{\hat{\boldsymbol{\Gamma}}}_q + (\mathrm{Y}_t - \hat{\hat{\underline{P}}}_{tq})\hat{\hat{\boldsymbol{\gamma}}}_q + \hat{\hat{\mathbf{H}}}_q$ is computed as the "leave-one-out" estimator, which deletes the qth observation in (13) when estimating $\hat{\hat{\Gamma}}_q$, $\hat{\hat{\gamma}}_q$, and $\hat{\hat{\mathbf{H}}}_q$.[7]

A simple algorithm that solves (15) based on (16) for bandwidth selection is

1. Select the bandwidth length by solving (16) and solve (14a) for elasticity estimates;
2. From step (1), find the variance–covariance estimates in (15b);
3. Use estimates in step (2) to solve (15a) for coefficient estimates;
4. Iterate steps (2)–(3),

where this step-based approach, in particular, corresponds to Zellner's estimator for the case of a parametric SUR estimator (i.e., when $h \to \infty$ in (15)). The presence of time varying coefficients may then be incorporated at any desired time using the temporal kernel, and automation in the estimation of the bandwidth in the estimator avoids imposing arbitrary discrete partitions of coefficients in the data. A GFVC meat demand model is next estimated non-parametrically.

[6] Again, in a linear in parameter model $\tilde{\mathbf{R}}_q$ corresponds to the matrix of explanatory variables.

[7] That is, the parameters $\hat{\hat{\Gamma}}_q$, $\hat{\hat{\gamma}}_q$, and $\hat{\hat{\mathbf{H}}}_q$ are estimated as in (14), but without the qth observation in (13).

Data and Kernel-Based Elasticities Estimates

The temporal kernel estimator of the GFVC demand system is applied to a meat demand system that includes the three main meat groups: beef, pork, and poultry, plus the numeraire captured by all other goods. The analysis uses annual USDA data from 1950 to 2000 on per-capita retail weights of beef, pork and poultry consumption, and Bureau of Labor Statistics data on price indexes for each meat group considered. Data on personal disposable income and a price index for the numeraire that captures all other goods complete the data needs for estimation of a weakly separable meat demand system.

For the distance between observations t and q, the temporal kernel in (15) with weights based only on time lengths, $d(t, q) = |t - q|$, is a non-parametric estimator of the GFVC model. A potential limitation of using only the time length to specify the kernel could be, however, its inability to capture closeness in variable values in terms of the explanatory variables. Alternatively, to model coefficient variation across time in a one-dimensional space, while incorporating closeness in variable values in terms of the explanatory variables, a distance function is

$$(17)\quad d(t,q) = \sum_{i=1}^{n} o_i |r_{it} - r_{qt}| + o_{n+1}|t - q|$$

where $o_i = \{0, 1\}$; and r_{it} is the ranking of variable i at observation t in terms of its value-size.[8] The structure in (17) avoids the dimensionality curse of unrestricted multivariate kernels (Stone), while the need to define the temporal kernel in terms of some or all the components in (17) can be determined based on statistical evidence.

A numerical approach based on Gauss–Newton algorithm solves the first-order conditions in (15), while the components of the distance function in (17) and the bandwidth length in (12) are chosen based on the lowest estimated prediction error (EPE) in (16). The selected distance function with the lowest EPE only uses the time length between observations and, for this distance function,[9] table 1 shows the estimated prediction error (EPE) in (16) for different bandwidth lengths applied to the meat demand system. The bandwidth determines the weights allotted to an observation for each point of approximation. The "leave-one-out" estimator indicates that the EPE is minimized at $h = 3.1$, and the temporal kernel EPE is 63% lower than the EPE associated with the nonlinear least squares estimator. After selecting the bandwidth that minimizes the EPE in table 1, coefficient estimates for any given reference point of the GFVC demand system can be obtained by solving (15). In the analysis of elasticities, however, the kernel tends to be ineffective (large standard errors) when the point of approximation is far-off from the mean value of the data.

Table 1. Estimated Prediction Error under Different Bandwidth Lengths

Bandwidth	EPE
∞	0.4460
15	0.2662
10	0.2412
8	0.2256
7	0.2125
6	0.1941
5	0.1697
4	0.1408
3.5	0.1278
3.4	0.1258
3.3	0.1241
3.2	0.1229
3.1	0.1221
3	0.1222
2	0.2078
1.5	0.3855
1	0.7276

Note: The estimated prediction error (EPE) is calculated using the "leave-one-out" estimator. The EPE in the table is multiplied by 10^3.

In selecting the reference points, upper and lower bounds of the explanatory variables in the data are thus avoided, and the selected reference points in this article are chosen so that they represent different periods in the data, but at the same time consist of observations within values of other observations. Accordingly, the kernel estimator of the GFVC demand system is evaluated at the reference points: 1955, 1975, and 1995, and elasticity estimates at each reference point are illustrated in table 2. Table 2 also reports a special case of the kernel estimator of the GFVC demand system in (15), which occurs when $h \to \infty$. (Table 3 reports estimates of the variance–covariance matrix for different reference points.)

[8] This use of ranks homogenizes the metric across explanatory variables and the metric of a time trend to assign weights in the distance function.

[9] The EPE, when using a distance function based only trend and the rank of income, is 0.128, and when using all explanatory variables in the distance function, the EPE is 0.284.

Table 2. Own Price and Income Elasticity Estimates Derived from the Globally Flexible Varying Coefficient Demand System

Variable	GLS	Kernel-1955	Kernel-1975	Kernel-1995
Beef price—BE	−0.32 (−1.81)	−0.66 (−6.91)	−0.62 (−3.18)	−0.48 (−0.68)
Pork price—PE	−1.25 (−5.72)	−1.38 (−4.42)	−1.08 (−3.52)	−0.83 (−3.24)
Poultry price–poultry Eq.	−0.63 (−3.29)	−0.73 (−3.30)	−0.74 (−2.27)	−0.81 (−3.04)
Income—BE	0.31 (5.23)	0.76 (4.91)	0.12 (0.50)	−0.25 (−0.47)
Income—PE	0.20 (2.21)	0.55 (2.07)	−0.01 (−0.30)	0.14 (0.51)
Income–poultry Eq.	0.95 (4.25)	0.07 (0.23)	0.86 (2.27)	0.65 (1.91)

Notes: The symbol BE refers to the beef equation, and the symbol PE refers to the pork equation. All values in parenthesis are *t*-values. GLS estimates assume time invariant coefficients.

Table 3. Selected Variance–Covariance Matrix Coefficient Estimates from the Temporal Kernel

Variable	1955	1975	1995
Variance BE	0.74	1.08	0.16
Covariance BE and PE	0.26	0.38	0.06
Covariance BE and poultry Eq.	0.07	−0.17	−0.09
Variance PE	1.37	0.59	0.04
Covariance PE and poultry Eq.	0.64	0.05	−0.03
Variance poultry Eq.	0.58	0.63	0.09

Notes: The symbol BE refers to the beef equation, and the symbol PE refers to the pork equation. All reported estimates of the covariance matrix in the table need to be multiplied by 10^{-5}.

Analysis of the GFVC demand system estimated with the temporal kernel in (15) shows some interesting patterns. For example, in table 2, the norm (absolute value) of the own price elasticity of beef and pork is lowest at the most recent reference point, 1995. In the year 1955, results indicate (with statistical significance) that beef and pork have positive income elasticities, while the income elasticity of poultry is not significant. In 1975, in contrast, the income elasticities for beef and pork are statistically insignificant, while the income elasticity for poultry is significant. (Interestingly, Moschini and Meilke show that around 1975 there was a structural break in meat demand.)

In 1995, estimates show that the income elasticity of beef is negative, but not statistically significant, while poultry has the largest income elasticity. The size of the price elasticity for beef and pork are lower in 1995 than in 1975 and 1955, while the own price elasticity of poultry is largest in 1995. Overtime, red meat products (beef and pork) have thus become less price and income elastic and, in particular, red meat products have become quite income inelastic in comparison to 1955.

MSE Performance of the Temporal Kernel

In meat demand, kernel estimation of time varying coefficients adds non-parametric flexibility, while significance of the elasticity estimates tends to be maintained at each reference point in table 2. Alternatively, for parsimony, the GFVC model can be approximated using simpler structures for modeling the coefficient variation, e.g., dummy variables and time trends. Yet, ad hoc parameterizations in the variation of coefficients across time and arbitrary partitions of the data may impose distortions in coefficient estimates. This section compares the temporal kernel to parameterizations of time varying coefficients with dummy variables and time trends.

For the model comparison, cross-validation evaluates goodness of fit through observations not included in the regression capturing the effect of the variance as well as the biases from the model specification (see Piggott). Specifically, the method of cross validation computed as the "leave-one-out" estimator (see [16]) provides a mean square error (MSE) criterion for model selection (see Schmalensee et al.). Table 4 shows the MSE resultant from the cross-validation analysis of the GFVC model with intercept shifters, as well as with dummy variables (time trends) for both intercept and interaction variables.

Clearly, results in table 4 show that the temporal kernel has lower MSE than traditional approaches for modeling varying coefficients. For example, the model with dummy variables that provides best MSE performance occurs for the case with one partition for both intercept and slope with an estimated prediction error (EPE) of 0.27, which represents a lower performance than the EPE of 0.122 under the temporal kernel. Overall, under the traditional method, flexibility attained with dummy

Table 4. EPE Using the "Leave-One-Out" Estimator of Different Varying Coefficient Forms

GFVC Model	EPE
With parameter stability	0.446
Intercept dummy: One partition	0.470
Intercept dummy: Two partitions	0.446
Intercept dummy: Three partitions	0.498
Intercept trend	0.375
Intercept quadratic trend	0.344
Intercept-slope dummy: One partition	0.278
Intercept-slope dummy: Two partitions	0.297
Intercept-slope dummy: Two partitions	3.254
Parametric model: Time trend as interaction variable	0.259
Parametric model: Quadratic trend as interaction variable	0.213
Parametric model: Cubic trend as interaction variable	0.284
Temporal Kernel	0.122

Note: The estimated prediction error (EPE) is calculated using the "leave-one-out" estimator. The globally flexible varying coefficient model (GFVS) model is estimated under the different structures for the coefficient variation described in the table.

variables quickly becomes very expensive in terms of degrees of freedom after two partitions reflected by drastic increases in MSE (see table 4).

A parametric modeling of the variation of the coefficients with a time trend is similarly limited relative to the temporal kernel. For parametric specification, the lowest MSE uses a quadratic trend polynomial entering in the intercept and as an interaction variable with EPE of 0.21, which—again—represents a lower performance than under the temporal kernel. Therefore, in addition to the properties of non-parametric flexibility and automation provided by the temporal kernel, cross-validation shows empirically that the temporal kernel has lower MSE than traditional ways of modeling varying coefficients (time trend, dummy variables).

Conclusion

This article applies the Rank Theorem to derive an exact representation of a demand system in terms of time varying coefficients. Restrictions are then imposed on the elements of the vector of parameters to impose integrability conditions, and the integrable demand system with time varying coefficients representing derivatives of the function for all observations in the data; and the system is then globally flexible. In estimation, this GFVC model only requires modeling curves in a one-dimensional space to capture global nonlinearities of the demand system.

To approximate the variation of coefficients in the GFVC demand system, the paper derived a temporal kernel estimator that allows elasticities to be estimated without any ad hoc parametric assumptions in the functional form, and elasticity estimates can be tailored to the observation of interest to the researcher. As a result, different from past work, flexibility was achieved by using the Rank Theorem and a temporally weighted estimator.

In the meat demand application, the temporal kernel maintains statistical significance of price elasticity estimates, and the estimator simultaneously captures nonlinearities and structural change with minimum parameterizations. The temporal kernel also yields better MSE fit than alternative forms of the GFVC demand system that includes traditional parametric modeling of the variation of time varying coefficients with dummy variables or time trends.

A temporal kernel approach may thus be considered as a viable approach in estimation, and the importance of price elasticity estimates based on time series data underscores a useful place of the GFVC among the variety of forms available for food demand analysis. Future work may explore the use of the temporal kernel in the context of other forms (e.g., see Barten and Bettendorf), and consider the use of the rank theorem to define new classes of globally flexible systems in production with reduced dimensionality.

[Received July 2003; accepted May 2004.]

References

Alston, J.L., and J.A. Chalfant. "The Silence of the Lambdas: A Test of the Almost Ideal and Rotterdam Models." *American Journal of Agricultural Economics* 75(1993):304–13.

Alston, J.L., and R. Green. "Elasticities in AIDS Models." *American Journal of Agricultural Economics* 72(1990):442–45.

Barten, A.P., and L.J. Bettendorf. "Price Formation of Fish: An Application of an Inverse Demand System." *European Economic Review* 33(1989):1509–26.

Cleveland, W.S. "Robust Locally Weighted Regression and Smoothing Scatterplots." *Journal of the American Statistical Association* 79(1979):829–36.

Deaton, A.S., and J. Muellbauer. "An Almost Ideal Demand System." *American Economic Review* 70(1980):312–26.

Diewert, W.E. "Application of Duality Theory." In M.D. Intrilligator and D.S. Kendtrick, eds. *In Frontiers of Quantitative Economics*. Amsterdam: North-Holland Publishing Co., 1974, pp. 106–71.

Engle, R.F., W.J. Granger, J. Rice, and A. Weiss. "Semiparametric Estimates of the Relation between Weather and Electricity Sales." *Journal of the American Statistical Association* 81(1986):310–20.

Fan, J. "Design Adaptive Nonparametric Regression." *Journal of the American Statistical Association* 87(1992):998–1004.

——. "Local Linear Regression Smoothers and their Minimax Efficiencies." *Annals of Statistics* 30(1993):196–216.

Gallant, A.R. "On the Bias in Flexible Functional Forms and Essentially Unbiased from the Fourier Flexible Form." *Journal of Econometrics* 15(1981):211–45.

Granger, C.W.J. "Macroeconomics: Past and Future." *Journal of Econometrics* 100(2001):17–19.

Jorgenson, D.W., L.J. Lau, and T.M. Stoker. "The Trascendental Logarithmic Model of Aggregate Consumer Behaviour." In R.L. Basmann and G. Rhodes, eds. *Advances in Econometrics*. Vol.1. 1982, pp. 97–238.

Moschini, G., and K.D. Meilke. "Modeling the Pattern of Structural Change in U.S. Meat Demand." *American Journal of Agricultural Economics* 71(1989):235–61.

Moschini, G., and D. Moro. "Structural Change and Demand Analysis: A Cursory Review." *European Review of Agricultural Economics* 23(1996):239–61.

Piggott, N.E. "The Nested PIGLOG Model: An Application to US Food Demand." *American Journal of Agricultural Economics* 84(2003):1–15.

Rudin, W. *Principles of Mathematical Analysis*, Third Edition. New York: McGraw-Hill, 1976.

Ruppert D., and M.P. Wand. "Multivariate Locally Weighted Least Squares Regression." *Annals of Statistics* 9(1994):65–78.

Schmalensee, R., and T. Stoker. "Household Gasoline Demand in the U.S." *Econometrica* 67(1999):645–62.

Silverman, B.W. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.

Stone, R. "*Measurement of Consumer's Expenditures and Behavior in the United Kingdom*." Vol. 1. Cambridge, UK: Cambridge University Press, 1954.

Stone, C.J. "Optimal Convergence Rates for Nonparametric Estimators." *Annals of Statistics* 8(1980):1348–60.

Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias." *Journal of the American Statistical Association* 57(1962):348–68.