Econometric Theory, **28**, 2012, 1087–1120.
doi:10.1017/S0266466612000412

REGRESSOR DIMENSION REDUCTION WITH ECONOMIC CONSTRAINTS: THE EXAMPLE OF DEMAND SYSTEMS WITH MANY GOODS

STEFAN HODERLEIN AND ARTHUR LEWBEL
Boston College

Microeconomic theory often yields models with multiple nonlinear equations, nonseparable unobservables, nonlinear cross equation restrictions, and many potentially multicolinear covariates. We show how statistical dimension reduction techniques can be applied in models with these features. In particular, we consider estimation of derivatives of average structural functions in large consumer demand systems, which depend nonlinearly on the prices of many goods. Utility maximization imposes nonlinear cross equation constraints including Slutsky symmetry, and preference heterogeneity yields demand functions that are nonseparable in unobservables. The standard method of achieving dimension reduction in demand systems is to impose strong, empirically questionable economic restrictions such as separability. In contrast, the validity of statistical methods of dimension reduction such as principal components has not hitherto been studied in contexts like these. We derive the restrictions implied by utility maximization on dimension-reduced demand systems and characterize the implications for identification and estimation of structural marginal effects. We illustrate the results by reporting estimates of the effects of gasoline prices on the demands for many goods, without imposing any economic separability assumptions.

1. INTRODUCTION

Consider a policy question such as, "What would be the effects of an increase in gasoline taxes or prices on the demands for every good consumers purchase?" By standard utility theory, Marshallian demand functions for each good depend (typically nonlinearly) on the prices of every consumption good. The number of different goods that consumers buy is large, and observed price variation across consumers in most data sets is limited (because, absent price discrimination, all consumers in a given time period and region pay similar if not identical prices).

Excellent research assistance by Sonya Mihaleva is gratefully acknowledged. Address correspondence to Arthur Lewbel, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA; e-mail: lewbel@bc.edu.

© Cambridge University Press 2012

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

As a result, available data sets do not have sufficient price variation to estimate demand functions containing the separate prices of all the dozens or hundreds of different goods that theory says must be considered.

This is a common problem in many microeconometric applications. Structural models often involve large numbers of covariates, yielding the curse of dimensionality in nonparametric models. Even in semiparametric and parametric models, having large numbers of closely related regressors, such as many prices, often leads to multicolinearity, flat objective and likelihood functions, and related problems. The difficulties associated with having many regressors are exacerbated in contexts such as demand systems both by the constraints of utility theory and by the complexity of empirically observed behavior, which together require models that are highly nonlinear, have complicated nonlinear cross equation restrictions, and possess nonseparable vectors of unobserved preference heterogeneity parameters. To deal with these problems, we show how statistical dimension reduction techniques such as principal components can be applied in this context of nonlinear, nonseparable, multiple equation systems.

The typical way of dealing with these difficulties is to impose strict constraints on behavior using, e.g., strong restrictions on preferences such as separability. In the demand system context, both traditional separability assumptions (Gorman, 1959; Blackorby, Primont, and Russell, 1978) or modern variants such as latent separability (Blundell and Robin, 2000) greatly restrict preferences by assuming that very many ad hoc equality constraints exist on the cross price effects of goods within each group.[1] The limitations of these grouping schemes become obvious when the goals of the analysis concern prices and demands of individual goods, because separability into groups forces common price effects on all goods within each group. Alternatives to separability such as the Hicks (1936) and Leontief (1936) composite commodity theorem or the newer generalized composite commodity theorems (Lewbel, 1996; Davis, 2003) permit estimation of group demands without separability but cannot provide estimates of the effects of prices and income on the demands for individual goods.

Analogous to separability assumptions are restrictions commonly imposed in the marketing and industrial organization literatures in which demand functions are taken over a small number of characteristics rather than over a large number of goods (as in hedonic pricing models) and typically restrict goods to discrete choices. A prominent example is Berry, Levinsohn, and Pakes (1995). These methods work well when the number of characteristics describing goods is small relative to the number of goods (as when the goods consist of many brands of a single type of product) and when quantities purchased tend to be small and hence well approximated by discrete counts, such as automobile purchases. In contrast, our proposed method can be applied when the relevant set of characteristics is large or difficult to enumerate, as when the goods are diverse such as food, clothing, gasoline, etc., and when purchases are better approximated by continuous quantities rather than discrete counts, as in gallons of gasoline per year.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

As an alternative to these strong (and in many contexts, empirically questionable) economic restrictions, we propose the use of statistical methods of dimension reduction such as principal components to identify the marginal effects (derivatives or elasticities) of continuous covariates of interest on average structural functions in nonlinear, nonseparable models possessing economic constraints. The validity of statistical methods of dimension reduction has not hitherto been studied in contexts like ours of nonseparable errors and nonlinear cross equation restrictions.

For concreteness, we illustrate our approach using demand systems. The specific complication for demand systems is that they possess as many prices as there are equations (goods), the prices of all goods appear (generally nonlinearly) in all equations, unobserved random utility parameters appear nonadditively and nonseparably in every equation, and utility maximization yields nonlinear functional restrictions such as homogeneity and Slutsky symmetry both within and across equations. However, given the preceding discussion it should be clear that our basic methodology can be generalized to other high-dimensional economic models with many continuously distributed covariates.

To describe our main contribution, consider the following setup. Let P_1 be a small vector of regressors of interest to the researcher, in our case the log price of gasoline, and let P_2 be the much larger vector of other regressors, in our case logged prices of all the other goods that consumers buy. In ordinary linear regression models, a standard solution to multicolinearity is to replace the large set of multicolinear nuisance regressors, in this case P_2, with a much smaller set of regressors R_2 constructed by applying a principal components, factor analysis, or other statistical dimension reducing algorithm to P_2 (see, e.g., Judge, Griffiths, Hill, Lütkepohl, and Lee, 1985, Ch. 22). We consider a similar solution to multicolinearity in the context of the nonlinear, nonseparable set of equations that comprise consumer demand systems with random utility parameters.

Note that the issue we are addressing is not lack of price variation per se but multicolinearity. If almost all of the variation in prices P_2 is driven by a small number of factors, then the elements of P_2 will be highly multicolinear, corresponding to insufficient relative price variation among all the prices. In contrast, the principal components will consist largely of these few factors, which can possess sufficient variation relative to each other over time to strongly identify their separate effects on demand.

Our first result establishes conditions under which the derivatives of dimension-reduced nonparametric regressions involving P_1 and R_2 (in place of P_1 and P_2) can be used to obtain derivatives of average structural functions in the sense of Blundell and Powell (2004), Altonji and Matzkin (2005), or Hoderlein (2005) and Hoderlein and Mammen (2007). Under suitable assumptions we show a direct link between average structural derivatives and the derivatives of constrained low-dimensional regressions, but in the most general case obtaining these derivatives requires an empirically estimable additional correction term.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

A second set of results takes this first result as a building block and considers implications and restrictions arising from economic theory in the dimension-reduced regression. We assume that individuals purchase quantities of many different goods to maximize utility functions (that may vary in unobserved ways across individuals) under a linear budget constraint, yielding Marshallian demand equations that are functions of P_1 and P_2 that possess certain testable restrictions (the so-called integrability constraints). We derive the restrictions on the dimension-reduced nonparametric demand functions involving P_1 and R_2 that arise from applying dimension reduction to the original integrability constrained Marshallian demands that are functions of P_1 and P_2.

For example, we show that the dimension-reduced demand functions do not themselves satisfy Slutsky symmetry (their Slutsky matrix is not even square) but under general conditions a certain transformation of them is symmetric. We emphasize that no form of separability or two-stage budgeting is assumed in this analysis. Dimension reduction based on replacing P_2 with R_2 is accomplished solely by exploiting mostly testable restrictions on relative price movements rather than difficult to test separability restrictions on preferences.

In our empirical applications, we apply these results to estimate the underlying structural derivatives of dependent variables (in our application, a large collection of budget shares) with respect to a regressor of interest (log gasoline price), and we estimate similar semielasticities with respect to other regressors such as total expenditures and demographic characteristics.

Our theorems employ elements of Brown and Walker (1989), Lewbel (2001), Christensen (2004), and Hoderlein (2005) regarding unobserved preference heterogeneity and Lewbel (1996) and Davis (2003) regarding price dimension reduction. However, by combining and extending these methods, we obtain far more general results. In particular these earlier papers can only identify aggregate effects across aggregates of goods, whereas we, e.g., identify the average effect of the price of gasoline on each of the many different goods that individuals buy, not just group price effects.

Although our modeling method and associated identification results are new, the resulting estimators take a standard two-step estimation form, involving parametric or nonparametric regressions (the demand system) that are estimated with some ordinary regressors such as P_1 and some generated regressors, R_2, so no new or nonstandard limiting distribution theory is required.

To show how our dimension reducing technique applies in different settings, we provide two empirical applications. The first is a simple parametric aggregate consumer model that illustrates the basic methodology and provides some macroeconomic estimates of gasoline own and cross price effects. The second application is a fully nonparametric individual consumer level study, which permits direct estimation and rigorous testing of the modeling methodology without household aggregation or functional form restrictions. Both applications show that our approach is easily implemented, obtains economically plausible results, and provides a convenient way of dealing with the problem

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

of the dimensionality of the vector of regressors in large structural econometric models.

2. THE MODEL

In this section we introduce the model formally, lay out the basic assumptions, and provide an in depth discussion of the economic content of the model in the case of consumer demand.

2.1. Notation

Let P_1 be a J_1 vector of regressors of interest, in our application logged prices, and let P_2 be a large J_2 vector of other regressors, in our case the logged prices of all other goods. Let P be the $J = J_1 + J_2$ vector of the logged prices of all goods, so $P' = (P_1' : P_2')$ where $'$ denotes transpose. Let R be the K vector (with K much smaller than J) defined by $R' = (P_1' : R_2')$, so R consists of the logged prices of interest, P_1, and some summary measures, R_2, of the remaining logged prices P_2. Let W be the J vector of budget shares of all goods consumed, so each element of W is the fraction of total expenditures spent on a good, and the corresponding element of P is the price of that good. The elements of W are nonnegative and sum to one. Define Y to be logged total expenditures and let Z denote a vector of other observable characteristics of consumers that affect, determine, or covary with their preferences. The vector Z might include variables such as household size, age, sex, region, and home ownership status. Finally, let A be a vector of all unobserved characteristics that affect preferences and hence demand.

The notation using capital letters reflects the fact that these variables are random, varying across the population and time. Realizations of these variables are denoted by small letters, when necessary with subscripts as y_{it}, w_{it}, z_{it}, a_{it}, and p_t, where i indexes the consumer and t indexes the price regime, where prices may vary by location and also by time. Let $||\cdot||$ denote euclidean norm; let vec denote the vectorization operator. Let F_{AU} denote the joint cumulative distribution function of two random vectors A and U and let f_{AU} denote the joint probability density function. Finally, let $\mathbb{E}$ denote expectation.

2.2. Assumptions

Using this notation, we are now able to state our assumptions formally. The first specifies the outcome equation of interest.

Assumption 1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which are defined the random vectors $A : \Omega \to \mathcal{A}$, $\mathcal{A} \subseteq R^\infty$, and $(W, P, Y, Z) : \Omega \to \mathcal{W} \times \mathcal{P} \times \mathcal{Y} \times \mathcal{Z}$, $\mathcal{W} \subseteq \mathbb{R}^J, \mathcal{P} \subseteq \mathbb{R}^J, \mathcal{Y} \subseteq \mathbb{R}, \mathcal{Z} \subseteq \mathbb{R}^L$, $t = 1, 2$, with J and L finite integers, such that

(a) $W = \phi(P, Y, Z, A)$, where $\phi : \mathcal{P} \times \mathcal{Y} \times \mathcal{Z} \times \mathcal{A} \to \mathcal{W}$ is a Borel measurable function; and
(b) realizations of W, P, Y, Z are observable, whereas those of A are not.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

Although this and later assumptions may apply to many general contexts, we will for concreteness discuss and interpret them in a consumer demand scenario: The function ϕ describes our large-scale structural model in a heterogeneous population, involving all regressors and also observed and unobserved heterogeneity parameters Z and A. If ϕ comes from utility maximization, then we may think of A as random utility parameters where the variation is either across individuals or across unobserved factors that determine preferences, whereas Z are observable characteristics that affect utility, so each consumer's utility function is completely determined by Z and A. When holding Z and A fixed, ϕ as a function of P and Y is the vector of Marshallian budget share demand functions. At this level of generality, each consumer can have a utility function and associated Marshallian demands that are completely different from everyone else's. Because $\mathcal{A} \subseteq R^{\infty}$, the vector $\mathcal{A}$ can have any dimension.

We assume that demand functions involving all prices cannot be estimated with any useful precision because the J vector P is too large and multicolinear. Therefore we instead consider demand equations $W = m(R, Y, Z) + V$, where m is defined as the conditional expectation of W conditioning on R, Y, and Z. The error vector V embodies both the effects of variation in unobserved taste attributes A across consumers and the effects of variation in P_2 that are not captured by variation in R.

To link the large set of regressors P to the dimension-reduced set R, we introduce the following assumption.

Assumption 2.

(a) There exist a J vector U and a J by K matrix of constants Γ such that

$$P = \Gamma R + U, \tag{1}$$

where Γ has rank K and all rows of Γ sum to one.

(b) Moreover, ϕ is continuously differentiable in the first $J+1$ components, and there exists a function η such that

$$|| \left[D_y \phi(\Gamma r + u, y, z, a)' : \operatorname{vec}\left[D_r \phi(\Gamma r + u, y, z, a) \right]' \right] || \leq \eta(u, a),$$

with $\int \eta(u, a) F_{AU}(da, du) < \infty$, uniformly in (r, y, z).

By itself, Assumption 2(a) can be interpreted as simply defining a vector U given a suitable matrix Γ. The restriction that rows of Γ sum to one is a convenient scale normalization on the elements of R and rules out trivial cases such as $\Gamma = 0$. The first J_1 elements of P and R are the same vector P_1, so the top left J_1 by J_1 corner of Γ is the identity matrix, the top right J_1 by $K - J_1$ corner of Γ consists of zeros, and the first J_1 elements of U are identically zero. Define U_2 to be the J_2 vector of remaining, nonzero elements of U. Note that because t indexes price regimes, realizations of R, U may be denoted r_t, u_t.

Assumption 2(a) generally holds if the elements of R are defined as linear homogeneous functions of P, i.e., if $R = \Pi P$ for some matrix Π, because if there

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

exists a nonsingular matrix S having first K rows equal to Π, then write S^{-1} as $(\Gamma : T)$ and equation (1) holds with U defined as T times the bottom $J-K$ rows of S times P. An example is defining R as K suitably scaled principal components of P, which is a standard method of obtaining dimension reduction in linear regression models. Having $R = \Pi P$ is also a common way to construct price indexes R; e.g., a Stone index has this form with rows of Π corresponding to budget share weights. We will therefore refer to R, and in particular R_2, as log price indexes.

We will later assume that R and U are uncorrelated or independent. There are many different models of price behavior that give rise to Assumption 2 with these additional restrictions. For example, equation (1) with R independent of U is the structure of a standard factor analysis model for log prices, and could also arise as a cointegration model for nonstationary log prices with R being $I(1)$ processes, U equaling independent stationary processes, and K being the number of cointegrating relationships that hold among the J logged prices.

A special case of Assumption 2(a) is the generalized composite commodity decomposition of prices proposed in Lewbel (1996). In that model, R is a vector of price indexes for distinct groups of goods, each row of the matrix Γ has one element equal to one and all other elements equal to zero, and U is a vector of differences between group prices and the prices of individual goods. Even when prices satisfy this particular decomposition, we still obtain results that differ from, and are more general than, those of Lewbel (1996).

Assumption 2(b) assumes differentiability of the structural demand functions and provides a mild boundedness condition on the derivatives required to interchange differentiation and integration.

We now provide our main identifying assumptions regarding the joint distribution of the observables and unobservables in our model.

Assumption 3. The joint distribution of R, Y, Z, A, and U obeys the following restrictions:

(a) (R, U) is independent of (Y, Z, A).
(b) A is conditionally independent of Y, conditioning on Z.
(c) (A, U) are jointly absolutely continuously distributed with respect to Lebesgue measure, conditional on (Y, Z, R).

Assumption 3(a) indicates that prices are determined independently of the variables that determine a consumer's demand functions. This price independence implies that we are performing a partial equilibrium analysis. Strictly, this assumption can only hold when prices are determined exclusively by supply (e.g., if goods are produced by constant returns to scale technologies and provided by perfectly competitive suppliers). However, for individual consumer data it is a reasonable assumption under much more general conditions because the contribution of any one consumer's demands to aggregate demand, and hence to price determination, is infinitesimal. However, this does rule out possible effects of common

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

demand shocks. This assumption also limits the possibility of including leisure as a good in the model, because it would then restrict permitted relationships between an individual's price of leisure (his wage rate) and other variables that affect demand.

In Appendix A we show that it is possible to relax Assumption 3(a) if necessary, at the expense of adding additional terms to the analysis that, although identified in theory, may be difficult to estimate in practice. In one of our empirical applications, which uses aggregate data, we consider instrumenting prices with supply side variables, which serves as a check on the sensitivity of our results to possible violations of this assumption.

Assumption 3(b) states that total expenditures are conditionally independent of random utility parameters. This assumption affects the variables that must be included in the conditioning set Z. For example, if a consumer's age affects his or her preferences, and if age is correlated with total expenditures, then age must be included in the vector of observed characteristics Z to satisfy this assumption. This assumption is implicit in most empirical demand system applications, either because such models assume no unobserved variation in preferences or because all unobserved variation is assumed to be contained in ordinary model errors.

Assumption 3(c) states that the unobservables are continuously distributed, which is stronger than necessary but simplifies some derivations. Taken together, the components of Assumption 3 immediately imply that $f_{AU|YZR} = f_{U|R} f_{A|Z}$.

Assumptions 1–3 comprise the core set of assumptions about our structural model, and we derive a general version of our results under these assumptions in the Appendixes. To obtain simpler (but more empirically convenient) results we invoke additionally the following assumption.

Assumption 4. R and U are independently distributed.

Though Assumption 4 is clearly restrictive, it holds, e.g., given a standard factor analysis model for prices. Also, Lewbel (1996) and Davis (2003) provide empirical support for this assumption for the special case of the generalized composite commodity model. This assumption may be more problematic for prices of goods of interest that are included in R, such as gasoline prices in our empirical application. This potential problem can be addressed in two ways. First, Assumption 4 can be easily tested empirically; e.g., one possible test could be based on regressing the estimated U residuals squared on a second-degree polynomial in R, analogous to White's test for heteroskedasticity. Second, in Appendix A we provide more general results that do not depend on Assumption 4. These extensions can be used in applications where R is constructed to be uncorrelated with U but not necessarily independent, e.g., when U is defined as the residual from regressing P on R. For Assumption 4 it is not necessary to assume that U has mean zero.

It is useful to compare these assumptions to those that are required for the standard methods of obtaining dimension reduction in demand systems, which are behavioral separability restrictions. First, starting again from Assumption 1,

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

standard demand system estimation does not allow for unobserved heterogeneity A, or at most permits restrictive separable forms that appear only in additive error terms. See Lewbel (2001) for a fuller discusion of this point. Next, for dimension reduction even weak separability requires that the cross derivatives of utility with respect to any two goods in different groups be proportional to the cross derivatives of any other pair of goods in those two groups (for details on this and alternative stronger forms of separability see, e.g., Blackorby et al., 1978). For estimation, one would also require that every consumer in the sample have the same grouping of goods and that the econometrician knows a priori which goods go into which groups. It is also required that the subutility functions for each group be known to the econometrician so that price indexes for each group can be constructed prior to estimation of the across group demand functions. Virtually all of these very strong behavioral assumptions required for traditional demand system estimation are difficult to test empirically, because doing so requires estimation of the non–dimension-reduced demand system, and the whole reason for dimension reduction is precisely the inability to estimate a non–dimension-reduced system with any precision. So, for empirical work, the restrictiveness of Assumptions 2–4 should be assessed relative to these alternatives.

2.3. Objects of Interest

The goal of our analysis is to relate properties of the unobservable vector of demand functions $W = \phi(P, Y, Z, A)$ to the observable, empirically tractable dimension-reduced regression $m(R, Y, Z)$, defined as

$$m(r, y, z) = \mathbb{E}[W \mid R = r, Y = y, Z = z] = \mathbb{E}[\phi(P, Y, Z, A) \mid R = r, Y = y, Z = z].$$

This conditional expectation always exists because budget shares W are bounded. We call $\phi(p, y, z, a)$ the theoretical demand function and $m(r, y, z)$ the empirical demand function because the latter is what we may feasibly estimate.

The empirical demand function $m(r, y, z)$ can be compared to the demand functions m^* that are usually estimated in practice, $W = m^*(P, Y, Z) + V^*$ where V^* is a vector of model errors. Traditionally, one would make sufficient separability assumptions to reduce P and W to manageable dimensions and would assume that $m^*(p, y, z)$ itself equals the demand functions arising from utility maximization, so rationality properties such as homogeneity and Slutsky symmetry would be imposed on (or tested with) m^*. This would be a valid assumption if, e.g., all consumers with the same value of z had identical preferences and if V^* corresponded to classical measurement errors. In that case, all of the theorems in this paper would hold replacing ϕ with m^*. However if V^* is instead the residual obtained from defining m^* as the conditional expectation of ϕ given p, y, z, then V^* would be the effect of variation in unobserved preference attributes A, and $m^*(p, y, z)$ need not satisfy rationality properties. See, e.g., Brown and Walker (1989), Lewbel (2001), Christensen (2004), and Hoderlein (2005), who derive and test properties of m^* and V^* in this context.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

In the present paper we do not deal with m^* because empirically feasible estimation of m^* requires additional restrictions such as separability. We instead directly condition the underlying theoretical demand functions ϕ on Y, Z, and a set of dimension reducing price indexes R to obtain the empirical demand function $m(r, y, z)$ that can be estimated.

It is important to recognize that there is no one "correct" definition of R and that the resulting demand functions we estimate depend on the choice of R. The underlying behavioral model $\phi(p, y, z, a)$ does not depend on R. As we show in the next section, the demand functions and associated elasticities we estimate are essentially averages across goods and prices. So, e.g., if one is primarily interested in own and cross price elasticities for food, one would construct R differently than if the focus is on energy, analogous to the way one can estimate different averages by grouping data in different ways, though the underlying data generating process does not depend on how the data are grouped.

3. MARGINAL EFFECTS OF INTEREST

We first consider how the marginal effects of interest, the income and price derivatives, $D_y m(r, y, z)$ and $D_r m(r, y, z)$ of the dimension-reduced regression, relate to the theoretical ones $D_y \phi(p, y, z, a)$ and $D_p \phi(p, y, z, a)$. Note that these derivatives are semielasticities; e.g., they are derivatives of budget shares with respect to log total expenditures y and with respect to log prices p or log price indexes r. More rigorous statements of the assumptions and theorems along with some generalizations and associated proofs are provided in the Appendixes.

THEOREM 1. *Let Assumptions 1–3 hold. Then*

$$D_y m(r, y, z) = \mathbb{E}\left[D_y \phi(P, Y, Z, A) \mid R = r, Y = y, Z = z\right]. \tag{2}$$

If Assumption 4 also holds then

$$D_{r_1} m(r, y, z) = \mathbb{E}\left[D_{p_1} \phi(P, Y, Z, A) \mid R = r, Y = y, Z = z\right] \tag{3}$$

and

$$D_r m(r, y, z) = \mathbb{E}\left[D_p \phi(P, Y, Z, A) \mid R = r, Y = y, Z = z\right] \Gamma. \tag{4}$$

Appendix B contains the proofs of this and more general versions of this theorem where we dispense in particular with the restrictive Assumption 4. Equation (2) shows that the income semielasticities of the empirical demands are a best approximation (in the sense of closest projection) to the theoretical income semielasticities because they equal the conditional expectations of the theoretical elasticities conditioning on the usable data R, Y, Z. Equivalently, the log total expenditure derivative of the empirical demands, $D_y m(r, y, z)$, equals the average of the log total expenditure derivatives over all individuals that have log total expenditures y and observable characteristics z and who are in price regimes that have log price indexes r.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

If we divide the population into subpopulations characterized by certain values $R = r, Y = y, Z = z$, then we can interpret $\mathbb{E}[D_y\phi(P, Y, Z, A) \mid R = r, Y = y, Z = z]$ as the average treatment effect of a marginal change in y for this subpopulation. This is as much as we can learn from mean regressions; the structural function ϕ or its derivatives are not identified, only functionals thereof. Given the generality of the preference heterogeneity involved, this result is in line with the literature (see Altonji and Matzkin, 2005; Imbens and Newey, 2009; Hoderlein, 2005). Given our results in Hoderlein and Mammen (2007), we conjecture that in the absence of further assumptions we would not obtain identification of ϕ even if we were to consider quantiles. As in Hoderlein (2005) or Imbens and Newey (2009), we can extend this result to cover endogenous regressors by simply adding control function residuals to the conditioning set. Note that we do not require Assumption 4 to hold for this result.

Equation (3) shows that a similar result is obtained for the derivatives of empirical demands with respect to the log prices of interest $p_1 = r_1$, whereas equation (4) shows that derivatives with respect to all log price indexes (including other elements of r) equal weighted averages of derivatives with respect to individual log prices, where the weights are given by Γ. The added aggregation (by averaging over the individual regressors) of course makes identification of ϕ more difficult. The full independence assumption, Assumption 4, however, ensures that this step does not add any distributional effect if we differentiate. We show in Appendix A that when this condition does not hold, then the empirical price derivatives equal the preceding averages of theoretical price derivatives plus a correction term that captures precisely such distributional effects. This term can be estimated, because it depends entirely on W and on the conditional distribution of U given R.

Note one key difference between equation (3) and equation (4): The first derivative is with respect to the regressor of interest, $r_1 = p_1$, which is retained and not reduced; the second derivative is with respect to a summary statistic. Although the former has exactly the same structure as the income effect, the latter is a weighted average of the original derivatives, with weights that depend on the influence of the individual regressor on the summary statistic R.

4. RATIONALITY RESTRICTIONS IN THE DIMENSION-REDUCED MODEL

In this section we consider the key restrictions utility maximization subject to a linear budget constraint imposes on individual demands. We start with the restrictions that arise from a linear budget constraint under nonsatiation (adding up and homogeneity of degree zero in the levels of prices and total expenditure) and then turn to the Slutsky conditions (i.e., Slutsky symmetry and negative semidefiniteness). We show that the dimension-reduced demand functions satisfy adding up and homogeneity but not the Slutsky conditions. Indeed, the Slutsky matrix is not even square. However, given some distribution assumptions, certain

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

linear combinations of the dimension-reduced demand functions do satisfy these Slutsky conditions.

Consider adding up and homogeneity first. The adding up constraint is that budget shares sum to one, i.e., the constraint that $1'_J \phi(p, y, z, a) = 1$ for theoretical demands and $1'_K m(r, y, z) = 1$ for empirical demands, where 1_L denotes the L vector of ones for any positive integer L. Homogeneity of theoretical demands means that $\phi(p, y, z, a)$ is homogeneous of degree zero in $\exp(p)$ and $\exp(y)$, whereas homogeneity of empirical demands is the property that $m(r, y, z)$ is homogeneous of degree zero in $\exp(r)$ and $\exp(y)$.

THEOREM 2. *If Assumption 1 holds and $\phi(p, y, z, a)$ satisfies adding up then $m(r, y, z)$ satisfies adding up. If Assumptions 1–4 hold and $\phi(p, y, z, a)$ is homogeneous then $m(r, y, z)$ is homogeneous.*

Theorem 2 shows that theoretical demand homogeneity and adding up implies empirical demand homogeneity and adding up, so in particular price homogeneity carries over to price indexes r. Like Theorem 1, Theorem 2 can also be extended to cases where U and R are not independent, in which case $m(r, y, z)$ will still satisfy adding up and $m(r, y, z)$ will be homogeneous if the conditional pdf of U conditioned on R is homogeneous of degree zero in R. See Appendix A for details.

In budget share form the theoretical Slutsky matrix is defined by

$$S(p, y, z, a) = D_p \phi(p, y, z, a) + \left[D_y \phi(p, y, z, a)\right] \phi(p, y, z, a)' + \phi(p, y, z, a) \phi(p, y, z, a)' - \operatorname{diag}\{\phi(p, y, z, a)\}, \quad (5)$$

where $\operatorname{diag}(\cdot)$ is a diagonal matrix with the vector $(\cdot)$ on the diagonal. Rationality implies that this matrix S is negative semidefinite and symmetric. A corresponding empirical Slutsky matrix based on $m(r, y, z)$ does not exist, because the number of goods is J but the number of prices (i.e., price indexes) comprising r is only K, so the price (index) derivatives matrix $D_r m(r, y, z)$ is J by K, which is not square. It is therefore not possible for $m(r, y, z)$ itself to satisfy Slutsky symmetry and negative semidefiniteness.

However, consider the K vector of functions

$$\widetilde{m}(r, y, z) = \Gamma' m(r, y, z),$$

so $\widetilde{m}(r, y, z) = \mathbb{E}[\Gamma' W \mid R = r, Y = y, Z = z]$. We may think of $\widetilde{W} = \Gamma' W = \widetilde{\phi}(P, Y, Z, A) = \Gamma' \phi(P, Y, Z, A)$ as a K vector of budget shares of constructed commodities, with each such commodity defined as a linear combination (defined by columns of Γ) of all the goods a consumer may buy. We may then interpret $\widetilde{m}(R, Y, Z)$ as a vector of demand functions for these constructed commodities and interpret R as the price indexes corresponding to these constructed commodities. These constructed commodities are analogous to groups of goods in the Blundell and Robin (2000) latent separability model and to the produced goods in the Gorman (1976) general linear technologies.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

We show subsequently in Theorem 3 that symmetry and negative semidefiniteness of the theoretical Slutsky matrix, and hence rationality of individual consumers, have observable (testable) implications for the empirical demands $m(r, y, z)$ through the constructed commodities $\widetilde{m}(r, y, z)$.

The matrix that corresponds to a Slutsky matrix for constructed commodities is, by analogy with equation (5), the K by K matrix

$$\widetilde{S}(r, y, z) = D_r\widetilde{m}(r, y, z) + \left[D_y\widetilde{m}(r, y, z)\right]\widetilde{m}(r, y, z)' + \widetilde{m}(r, y, z)\widetilde{m}(r, y, z)' - \operatorname{diag}\left\{\widetilde{m}(r, y, z)\right\}. \tag{6}$$

Also define the K by K matrices

$$H(r, y, z) = \operatorname{Cov}\left[D_y\widetilde{\phi}(P, Y, Z, A), \widetilde{W} \mid R = r, Y = y, Z = z\right],$$

$$G(r, y, z) = \operatorname{Var}\left[\widetilde{W} \mid R = r, Y = y, Z = z\right] + \operatorname{diag}\left\{\widetilde{m}(r, y, z)\right\} - \Gamma'\operatorname{diag}\left\{m(r, y, z)\right\}\Gamma.$$

The matrix H is just the conditional covariance between the derivative of $\widetilde{\phi}$ with respect to y and $\widetilde{\phi}$ itself, which we could also write as

$$H(r, y, z) = \mathbb{E}\left[\left[D_y\widetilde{\phi}(P, Y, Z, A)\right]\left[\widetilde{W}' - \widetilde{m}(R, Y, Z)'\right] \mid R = r, Y = y, Z = z\right].$$

Similarly, the first term in G is just the conditional variance of $\widetilde{\phi}$, and G is by construction symmetric. Define $m_2(r, y, z) = \mathbb{E}[WW' \mid R = r, Y = y, Z = z]$, $\widetilde{m}_2(r, y, z) = \mathbb{E}[\Gamma'W\left(\Gamma'W\right)' \mid R = r, Y = y, Z = z]$, and finally define the matrix L by

$$L(r, y, z) = D_r\widetilde{m}(r, y, z) + D_r\widetilde{m}(r, y, z)' + D_y\widetilde{m}_2(r, y, z) + 2\left[\left(\widetilde{m}_2(r, y, z)\right) - \Gamma'\operatorname{diag}\left\{m(r, y, z)\right\}\Gamma\right].$$

Using these definitions, we consider the following theorem.

THEOREM 3. *Let Assumptions 1–4 hold. Then*

$$\mathbb{E}\left[\Gamma'S(P, Y, Z, A)\Gamma \mid R = r, Y = y, Z = z\right] = \widetilde{S}(r, y, z) + H(r, y, z) + G(r, y, z). \tag{7}$$

If $S(p, y, z, a)$ is negative semidefinite then $L(r, y, z)$ is negative semidefinite. If $S(p, y, z, a)$ is symmetric then $\widetilde{S}(r, y, z) + H(r, y, z)$ is symmetric. If $S(p, y, z, a)$ is symmetric and negative semidefinite then $\widetilde{S}(r, y, z) + H(r, y, z) + G(r, y, z)$ is symmetric and negative semidefinite.

Like Theorems 1 and 2, Theorem 3 can be extended to eliminate Assumption 4 by adding additional terms related to the conditional distribution of U given R. Details of this extension and proofs are in the Appendixes. Together, Theorems 2

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

and 3 provide all of the restrictions on the empirical demand functions m that are implied by rationality of ϕ.

Theorem 3 first shows that negative semidefiniteness of the Slutsky matrix can be tested, or imposed upon estimation, through the matrix $L(r, y, z)$. This result is an extension of Hoderlein (2005). All the terms in $L(r, y, z)$ are functions of conditional expectations of W or $\widetilde{W}$, conditioning on r, y, and z, and so can be readily estimated.

Theorem 3 also gives the relationship between the theoretical Slutsky matrix S and the constructed commodity Slutsky matrix $\widetilde{S}$. This relationship depends on the matrices H and G. Like L, the matrices S and G depend only upon expectations conditioning on r, y, and z and so can be readily estimated. However, H depends upon $D_y\widetilde{\phi}(P, Y, Z, A)$, the total expenditure derivative of the theoretical demand functions, and so for testing or imposing Slutsky symmetry we will need to either assume H is symmetric or make modeling assumptions regarding $\widetilde{\phi}$ that suffice to either identify H or deduce its properties. For example, if

$$\Gamma'\phi(P, Y, Z, A) = \mu(P, Z, A) + \tau(Z)Y \tag{8}$$

holds for some functions μ and τ, then one can immediately check that H will be a zero matrix. Another example is that if $K = 2$ then the constructed commodity Slutsky matrix $\widetilde{S}$ is 2 by 2, and in any model with just two commodities, Slutsky symmetry always holds given homogeneity and the adding up constraint that budget shares sum to one. It then follows from rationality of ϕ that in this case $\widetilde{S}$ and H will be symmetric.

The matrix H is similar to the matrix defined in equation (7) of Lewbel (2001), so much of the analysis of H provided there extends to the present framework, with the main difference being that here the covariance defining H is conditioned upon price indexes R rather than individual good prices P. Even when H is not symmetric, if the derivative of $\phi(P, Y, Z, A)$ with respect to Y varies little with P and with A then the H matrix will be numerically small, so its effects might be safely ignored in a policy analysis. Lewbel (2001) gives an example like this in a model that is quadratic in Y. See also Christensen (2004), where the author uses panel data to estimate a matrix like H. Also, as in Hoderlein (2005), negative semidefiniteness may be more relevant than symmetry, and negative semidefiniteness can be directly tested or imposed using matrix L without modeling H.

In our empirical applications, adding up is imposed on estimation, and we test homogeneity, negative semidefiniteness, and symmetry. Only the test of symmetry requires assumptions regarding H, and for that test we assume H is symmetric.

In some applications, it is not necessary to estimate the entire system of J equations comprising m, and it would suffice only to estimate the much smaller system of K equations $\widetilde{m}$. By the construction of Γ, just as the first J_1 elements of R equal P_1, we have that the first J_1 elements of $\widetilde{m}$ are the same as the first J_1 elements of m, which are the empirical budget shares of the goods that have prices P_1. So, e.g., if P_1 is a vector of energy prices, and one is only interested in the effects of energy prices P_1 on the demand for energy goods, then those are given

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

just by the first J_1 elements of $\widetilde{m}$. The remaining elements of $\widetilde{m}$ could be estimated to test or impose Slutsky conditions, but that would still only require estimating K equations in total rather than the much larger J. Note that by Theorem 2 and the construction of Γ, $\widetilde{m}$ will satisfy homogeneity if ϕ satisfies homogeneity, so if ϕ is rational and H is symmetric then $\widetilde{m}$ will satisfy both homogeneity and Slutsky symmetry.

5. A PARAMETRIC EXAMPLE

In this section we illustrate Theorems 1–3 using the almost ideal demand (AID) system of Deaton and Muellbauer (1980). For this simple example we assume all consumers have the same overall functional form for demands, but all the parameter values are permitted to vary in unknown ways across consumers. To simplify the presentation, let all preference attributes be unobserved, so the model contains A but not Z. We could easily reintroduce observed demographic characteristics Z in this example by having some or all parameters that depend on A depend instead on both A and Z.

The parameters of the model are a scalar valued function $\delta(A)$, J vector valued functions $\alpha(A)$ and $\beta(A)$, and a $J \times J$ matrix valued function $\gamma(A)$. The model consists of theoretical demand functions of the form

$$W = \phi(P, Y, A) = \alpha(A) + \gamma(A)P + \beta(A)[Y - G(P, A)], \tag{9}$$

where $G(P, A)$, defined as

$$G(P, A) = \delta(A) + \alpha(A)'P + \frac{1}{2}P'\gamma(A)P, \tag{10}$$

is the log of a price index that deflates total expenditures. This is just the ordinary AID model, except that we allow all the parameters of the model to be unknown functions of A, so all the parameters of the model are allowed to vary across consumers. Equivalently, we assume a potentially different AID model for every consumer.

With these theoretical demand functions, adding up requires $1'_J\alpha(A) = 1$, $1'_J\beta(A) = 0$, and $1'_J\gamma(A) = 0'_J$. Homogeneity requires $\gamma(A)1_J = 0_J$, and Slutsky symmetry requires symmetry of $\gamma(A)$. A sufficient though stronger than necessary condition for Slutsky negative semidefinitenss is negative semidefiniteness of $\gamma(A)$. These restrictions together yield a rational AID system, with indirect utility function given by

$$F\left[\ln(Y - G(P, A)) - P'\beta(A), A\right], \tag{11}$$

where F is any function that is strictly monotonically increasing in its first element. The AID budget share functions (9) are obtained by applying Roy's identity to equation (11).

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

Substituting equation $P = \Gamma R + U$ into equation (9) gives

$$W = \alpha(A) + \gamma(A)U + \gamma(A)\Gamma R + \beta(A)[Y - G(\Gamma R + U, A)],$$
$$G(\Gamma R + U, A) = \delta(A) + \alpha(A)'U + \frac{1}{2}U'\gamma(A)U + \alpha(A)'\Gamma R + U'\gamma(A)\Gamma R + \frac{1}{2}R'\Gamma'\gamma(A)\Gamma R. \quad (12)$$

Let Assumptions 1–4 hold with these rational theoretical demand functions and assume also that the covariances of $\beta(A)$ with $\delta(A)$, $\alpha(A)$, and $\gamma(A)$ are all zero (more generally, if Z were present these would need to be zero conditional covariances conditioning on Z). This will make the H matrix in Theorem 3 be zero. A sufficient though not necessary condition for these covariances to equal zero is that $\beta(A)$ not depend upon A. In this case H will equal zero because equation (8) will hold.

With these assumptions, averaging equation (12) over U and A gives

$$W = m(R, Y) + V, \quad (13)$$

where $\mathbb{E}(V \mid R, Y) = 0_J$ and

$$m(R, Y) = \alpha + CR + b[Y - g(R)], \quad (14)$$

$$g(R) = \overline{d} + a'\Gamma R + \frac{1}{2}R'\Gamma'CR, \quad (15)$$

where $\overline{d} = \mathbb{E}[\delta(A)] + \mathbb{E}[\alpha(A)]\mathbb{E}(U) + \frac{1}{2}\mathbb{E}\left[U'\mathbb{E}[\gamma(A)]U\right]$, $C = \mathbb{E}[\gamma(A)]\Gamma$, $b = \mathbb{E}[\beta(A)]$, and $a = \mathbb{E}[\alpha(A)] + \mathbb{E}[\gamma(A)]\mathbb{E}(U)$. The demand functions $m(r, y)$, defined by substituting equation (15) into (14), could by equation (13) be estimated by, e.g., nonlinear least squares using observed budget shares w_{it}, log price indexes r_t, and log total expenditures y_{it} from a collection of (demographically homogeneous if we condition on Z) consumers i, each observed in one or more time periods t, because, $w_{it} = m(r_t, y_{it}) + V_{it}$. These demand functions $m(R, Y)$ look similar to the almost ideal (though with different parameter values from those of the individual consumers); however, they are not quite almost ideal because the number of prices (elements of R) is only K, whereas the number of budget shares (elements of m) is J, so C is $J \times K$, not square like $\gamma(A)$.

Still, in this particular parametric example we obtain the nearly linear AID-like model $W = a + CR + b\widetilde{Y} + V$ where $\widetilde{Y}$ equals log deflated total expenditures $Y - g(R)$. This looks just like an ordinary aggregate AID model, except that instead of regressing each element of W on a constant, $\widetilde{Y}$, and all J prices P, we obtain the coefficients a, b, and C by just regressing each element of W on a constant, $\widetilde{Y}$, and the K price indexes R, which include the prices of interest $P_1 = R_1$.

In the model of equations (13)–(15) adding up requires $1_J'a = 1$, $1_J'b = 0$, and $1_J'C = 0_J'$. This can be obtained in the usual way by omitting one good and using these adding up constraints to obtain the parameters corresponding

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

to the last good. Homogeneity requires in addition that $C1_K = 0_J$, and Slutsky symmetry requires symmetry not of C (which is impossible because C is not square) but of $\Gamma' C$. Also, if $\gamma(A)$ is negative semidefinite then $\Gamma' C$ will be semidefinite.

6. AN AGGREGATE EMPIRICAL EXAMPLE: GASOLINE PRICES

As a first empirical illustration of the methodology, we aggregate the parametric AID based model of the previous section across consumers in each time period and estimate the resulting model using total U.S. consumption data to assess the impact of gasoline prices on the demands for a large number of consumer goods.

From Tables 2.4.5 and 2.4.6 of the U.S. National Income and Product Accounts (NIPA) we obtain budget shares W_{jt} and associated logged prices (consumption deflators) P_{jt} for $J = 22$ different goods j for 45 years t from 1960 to 2004. The logged prices P_{jt} are demeaned over time (this corresponds to a particular choice of units for measuring quantities of each good), so principal components can be easily used for constructing price indexes. Specifically, we let $R_{1t} = P_{1t}$ be the logged price of gasoline in time t, and we then define R_2, the last J_2 elements of R, as follows. Let S_{it} be the ith principal component of the matrix of P_2 consisting of all logged prices except gasoline. Then by construction $S_{it} = P_{2t}\pi_i$ where π_i is a 21 element vector of weights. The corresponding logged price index $R_{i+1,t}$ for $i = 1, 2$, and 3 is then defined by $R_{i+1,t} = P_{2t}'\pi_i/(1_{21}'\pi_i)$ where 1_{21} denotes the vector of 21 ones. This scaling of the principal component is necessary to ensure that R satisfies the homogeneity property of price indexes (e.g., if all prices $\exp(P)$ double then the price indexes $\exp(R)$ must also double).

We then define Γ and U_t by $P_t = \Gamma R_t + U_t$ where the jth row of Γ consists of the coefficients from a constrained linear time series regression of P_{jt} on the four elements of R_t, and we define U_{jt} (the jth element of U_t for each t) as the time t residual from this regression. Assumption 2 requires the rows of Γ to sum to one, so this constraint is imposed on the regressions defining each row of Γ.

The AID model is quadratic in logged prices (through the function g), so for this model Assumption 4 can be relaxed to only require that the conditional mean and variance of U given R be independent of R. Because U is defined as a set of regression residuals, U is by construction uncorrelated with R. To check the variance condition, we apply a White (1980) test for heteroskedasticity to each of the preceding price regressions. Table 1 gives the estimates of Γ, along with the R^2's of these regressions and the White test statistics.

To apply the model of the previous section to total U.S. consumption data, we must aggregate this model across consumers in each time period. By equations (13)–(15), we may write the budget shares for individual consumers (households) h in time t as $W_{ht} = A_t + b \ln X_{ht} + V_{ht}$, where X_{ht} is total expenditures and $A_t = a + Cr_t - bg(R_t)$. Total expenditures for each good are therefore given by $X_{ht}W_{ht} = A_t X_{ht} + bX_{ht} \ln X_{ht} + V_{ht}X_{ht}$, and averaging this expression across all

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

TABLE 1. U.S. NIPA price-based Γ estimates and heteroskedasticity of residuals test

	Γ_1	Γ_2	Γ_3	Γ_4	R^2	White p-value
Gasoline	1.0000				1.0000	
Motor vehicles	−0.0590	0.9816	0.1456	−0.0682	0.6281	0.9864
Furniture	−0.3211	1.1438	0.2840	−0.1066	0.6705	0.9528
Appliances	−0.1590	0.8147	0.4143	−0.0700	0.5530	0.7999
Video audio	−0.3267	0.0076	1.2170	0.1021	0.4442	0.0058
Other durables	0.0413	0.7077	0.2823	−0.0312	0.5366	0.7903
Food off premise	0.0987	0.8619	0.1694	−0.1300	0.7493	0.8681
Purchased meals	0.0014	1.0451	0.0736	−0.1201	0.7794	0.8320
Women's clothing	−0.0605	0.5824	0.4817	−0.0036	0.6031	0.1610
Men's clothing	0.2149	0.3860	0.3338	0.0653	0.5952	0.2281
Fuel oil and coal	0.6417	0.5261	−0.1141	−0.0538	0.7770	0.9965
Tobacco	0.1965	1.2327	−0.4349	0.0058	0.7872	1.0000
Household	0.0290	0.7994	0.0306	0.1411	0.6697	0.7546
Other nondurable	0.0415	0.6664	0.1024	0.1898	0.6845	0.3158
Owner rent	0.0831	0.8850	−0.0887	0.1206	0.6791	0.6379
Tenant rent	−0.0586	1.0462	−0.0762	0.0886	0.6797	0.7334
Electricity gas	−0.0749	1.0533	−0.0578	0.0794	0.7347	0.9721
Telephone	−0.4258	1.1690	0.1323	0.1245	0.7011	0.3389
Car care	−0.1429	1.1905	−0.0880	0.0404	0.7265	0.9096
Transportation	−0.0840	1.0810	0.0117	−0.0087	0.7736	0.9543
Business	−0.1553	1.2586	−0.0633	−0.0399	0.7733	0.4342
Education	−0.0710	1.2792	−0.1566	−0.0516	0.7901	0.8685
Other services	−0.0115	1.0563	−0.0147	−0.0301	0.7051	0.9967

consumers in time t then gives aggregate (per capita) budget shares

$$W_t = \frac{\mathbb{E}_t\left(X_{ht} W_{ht}\right)}{X_t} = A_t + b\frac{\mathbb{E}_t\left(X_{ht} \ln X_{ht}\right)}{X_t} + V_t,$$

where $\mathbb{E}_t$ denotes averaging across households h in time t, with $X_t = \mathbb{E}_t(X_{ht})$. This notation means that X_{ht} equals an X_t that varies randomly over time, plus a random term that for any time period t is mean zero across households h. Here $X_t = \mathbb{E}_t(X_{ht})$ is per capita total expenditures in time period t, and $V_t = \mathbb{E}_t(V_{ht} X_{ht})/X_t$ is an aggregate error term.

Lewbel (1991) showed that in the United States. using income, $\mathbb{E}_t(X_{ht} \ln X_{ht}) \approx X_t\kappa + X_t \ln X_t$, where κ is a constant. In particular, he finds that a linear time series regression of $\mathbb{E}_t(X_{ht} \ln X_{ht})$ on X_t and $X_t \ln X_t$, imposing a coefficient of one on the latter term, has an R^2 over 0.998, so roughly the approximation error here is less than 0.2%. Substituting this approximation into the preceding expression for W_t then gives $W_t = (A_t + b\kappa) + b \ln X_t + V_t$, where V_t now includes this small approximation error resulting from aggregation across X_{ht}. The resulting

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

model we estimate is then

$$W_t = a + C R_t + b\left[\ln X_t - g\left(R_t\right)\right] + V_t, \tag{16}$$

$$g\left(R_t\right) = d + a'\Gamma R_t + \frac{1}{2} R_t'\Gamma' C R_t, \tag{17}$$

where $d = \overline{d} + \kappa$. As with ordinary AID model estimation, the constant d is underidentified and so is set to zero.

We estimate this aggregate, macro data model using budget shares W_t constructed as NIPA consumption expenditures on each good j divided by the sum of these expenditures across all goods, and per capita total expenditures X_t defined as this sum of NIPA expenditures across all goods, divided by U.S. population. To avoid separability assumptions, our list of goods includes durables and also nondurables and services. We do not attempt to construct flows of services derived from durables, noting that aggregation across consumers smooths out most of the lumpiness in durable consumption that results from infrequency of purchase.

By analogy with the common approximate AID model, we first linearly regress W_t on a constant, R_t, and on $\ln X_t - P_t' W_t$. This is not a consistent estimator but yields reasonable starting values $\widehat{a}$, $\widehat{b}$, and $\widehat{C}$ for our formal estimator. We estimate equation (16) using three-stage least squares (specifically, generlized method of moments[GMM] with a weighting matrix that is is efficient under homoskedasticity, which for large systems such as ours can have good finite-sample properties and fewer numerical convergence problems than general efficient GMM; see, e.g., Wooldridge, 2002, for details). Estimation is based on the moment conditions

$$\mathbb{E}\left[\left(W_t - a - C R_t - b\left[\ln X_t - a'\Gamma R_t - \frac{1}{2} R_t'\Gamma' C R_t\right]\right) Q_{st}\right] = 0_J, \qquad s = 1, \ldots, S,$$

where the set of S instruments Q_{1t},...,Q_{St} consists of a constant, $\ln X_t$, all $K = 4$ elements of R_t, a time trend t, and $R_t'\Gamma'\widehat{C} R_t$ using the preceding starting value of $\widehat{C}$. Note that estimation error in $\widehat{C}$ only affects efficiency but not consistency of the resulting three-stage least squares estimator, because $R_t'\Gamma'\widetilde{C} R_t$ is a valid instrument (in the sense of being uncorrelated with the model error) for any choice of the matrix $\widetilde{C}$. We could more generally use any functions of X_t and R_t as instruments, but this set works well numerically in our small sample, because the model is nearly linear in $Q_{1t}, \ldots, Q_{St}$.

We impose the adding up constraints by omitting one good from the system. We estimate the model without the homogeneity and Slutsky symmetry restrictions $C 1_K = 0_J$ and $\Gamma' C = C'\Gamma$, so these restrictions can be tested. Wald tests of these restrictions are reported in Table 2, and Table 3 reports results of testing homogeneity separately in each of the 22 separate demand equations. As is often the case with aggregate data, overall symmetry is rejected, and for most but not all goods homogeneity is also rejected (though see the caveats that follow regarding possible statistical invalidity of these tests). Far fewer rejections are obtained with individual consumer data in the next section.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

TABLE 2. Aggregate model systemwide tests

	Homogeneity	Symmetry
Wald test statistic	2,872.74	44.4299
p-value	0.00	0.00

Table 4 reports C_{j1} for each equation j, which is the estimated derivative of the budget share of good j with respect to the logged price of gasoline, holding deflated log total expenditures $\ln x_t - g(r_t)$ fixed. Table 4 also reports b_j for each good j, which is the estimated derivative of the budget share of good j with respect to log total expenditures. Standard errors are provided in parentheses.

Most empirical implementations of demand system models do not adjust standard errors to account for price index estimation. Errors in demand systems are typically rather large (because of substantial unobserved heterogeneity of preferences), so the additional contribution to the final standard errors of first-stage estimation of price indexes is likely to be relatively modest. Still, because our

TABLE 3. Aggregate model homogeneity tests for each equation

	Homogeneity test statistic	p-value
Gasoline	15.6499	0.00
Motor vehicles	1.6021	0.21
Furniture	27.2294	0.00
Appliances	1.8169	0.18
Video audio	10.9130	0.00
Other durables	0.0424	0.84
Food off premise	18.6336	0.00
Purchased meals	64.6694	0.00
Women's clothing	3.7925	0.05
Men's clothing	52.0058	0.00
Fuel oil and coal	0.1661	0.68
Tobacco	69.7992	0.00
Household	227.3928	0.00
Other nondurable	11.4715	0.00
Owner rent	10.5153	0.00
Tenant rent	21.9833	0.00
Electricity gas	174.7613	0.00
Telephone	24.8777	0.00
Car care	4.8632	0.03
Transportation	67.9379	0.00
Business	1.6500	0.20
Education	14.7159	0.00
Other services	22.9348	0.00

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

aggregate model is completely parametric, to account for estimation error in the construction of price indexes R and to account for any autocorrelation in the residuals, we also estimated confidence intervals for these parameters using an ordinary block bootstrap. In this bootstrap consecutive time blocks of the original data are drawn with replacement, which are then used to construct price indexes, estimate the model, and construct parameter t-statistics. We construct

TABLE 4. Aggregate model parameter estimates

	Gas-price derivatives C_{j1}	Log total expenditures derivatives b_j
Gasoline	0.0042817	0.0136943
	(0.0014) [−0.0001; 0.0061]	(0.0196) [−0.0099; 0.035]
Motor vehicles	−0.0015811	0.0291848
	(0.0019) [−.00535; 0.00137]	(0.0267) [−0.026; 0.077]
Furniture	0.0000058	0.0029748
	(0.0002) [−0.00039; 0.00018]	(0.0025) [−0.009,0.0071]
Appliances	0.0003378	0.0129098
	(0.0001) [0.00020; 0.00066]	(0.0020) [0.0026; 0.0170]
Video audio	−0.0002001	0.0119140
	(0.0003) [−0.00046; 0.00018]	(0.0046) [0.0025; 0.0167]
Other durables	0.0012377	0.0342129
	(0.0006) [0.00044; 0.0029]	(0.0078) [0.024; 0.066]
Food off premise	0.0017509	−0.0139890
	(0.0010) [−0.00049; 0.0027]	(0.0145) [−0.131; 0.0074]
Purchased meals	0.0020063	0.0316373
	(0.0006) [0.00047; 0.0033]	(0.0082) [−0.0025; 0.0564]
Women's clothing	0.0001581	0.0125307
	(0.0002) [−0.00077; 0.00025]	(0.0031) [−0.0159; 0.0321]
Men's clothing	−0.0000122	0.0108421
	(0.0001) [−0.00079; 0.00084]	(0.0020) [−0.081; 0.121]
Fuel oil and coal	0.0004234	−0.0069874
	(0.0002) [−0.00007; 0.00068]	(0.0029) [−0.021; 0.0010]
Tobacco	−0.0010198	−0.0256529
	(0.0003) [−0.0021; −0.00009]	(0.0045) [−0.041; −0.0023]
Household	−0.0010344	−0.0525102
	(0.0003) [−0.0029; −0.0001]	(0.0048) [−0.076; −0.015]
Other nondurable	0.0002252	0.0118459
	(0.0003) [−0.0002; 0.0013]	(0.0039) [0.0076; 0.022]
Owner rent	−0.0006820	−0.0449707
	(0.0005) [−0.0037; −0.00006]	(0.0078) [−0.075; -0.0059]
Tenant rent	−0.0007580	0.0032383
	(0.0004) [−0.00016; 0.000028]	(0.0052) [−0.0081; 0.0178]
Electricity gas	0.0015898	0.0132918
	(0.0006) [−0.00058; 0.0023]	(0.0081) [−0.018,0.035]

Table continues on overleaf

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

TABLE 4. *continued*

	Gas-price derivatives C_{j1}	Log total expenditures derivatives b_j
Telephone	−0.0010171	−0.0090374
	(0.0004) [−0.0023; 0.00056]	(0.0052) [−0.022; 0.00026]
Car care	−0.0013308	0.0030477
	(0.0006) [−0.0030; 0.00078]	(0.0079) [−0.018; 0.025]
Transportation	0.0006045	0.0214435
	(0.0002)[0.000067; 0.0013]	(0.0024) [−0.018; 0.025]
Education	−0.0008962	−0.0047274
	(0.0010) [−0.0028; 0.0032]	(0.0136) [−0.022; 0.020]
Business	0.0001459	0.0122341
	(0.0003) [0.000022; 0.00025]	(0.0038) [−0.0011; 0.029]
Other services	−0.0027634	−0.0402280
	(0.0008) [−0.0053; 0.0019]	(0.0108) [−0.077; 0.0081]

Note: Parametric standard errors are in parentheses; lower and upper bounds of block bootstrap based 95% confidence intervals are in square brackets.

95% confidence intervals across the bootstrap simulations for these t-statistics and convert the results back into confidence intervals for the underlying parameters. Because of the small sample size of $n = 45$, we picked rather small blocks of size 4. Still, we had to discard some bootstrap simulations that experienced severe colinearity issues resulting in a singular or ill-conditioned variance covariance matrix. Attempts to bootstrap with larger block sizes produced a higher percentage of numerical failures of this type. This points to the limits of attempting sophisticated asymptotic analyses in this specific small-sample application.

Table 5 provides the mean of each budget share W_j in the data, and reports the corresponding estimates of the elasticity (at the mean) of quantity demands

TABLE 5. Aggregate model parameter estimates

	Derivatives gas price	Derivatives total expenditures	Mean budget share	Elasticity gas price	Elasticity total expenditures
Gasoline	0.0042817	0.0136943	0.0320	−0.8664	1.4273
	(0.0014)	(0.0196)		(0.0430)	(0.6107)
Motor vehicles	−0.0015811	0.0291848	0.0584	−0.0271	1.4997
	(0.0019)	(0.0267)		(0.0322)	(0.4570)
Furniture	0.0000058	0.0029748	0.0117	0.0005	1.2540
	(0.0002)	(0.0025)		(0.0152)	(0.2160)
Appliances	0.0003378	0.0129098	0.0134	0.0253	1.9662
	(0.0001)	(0.0020)		(0.0104)	(0.1482)

Table continues

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

TABLE 5. *continued*

	Derivatives gas price	Derivatives total expenditures	Mean budget share	Elasticity gas price	Elasticity total expenditures
Video audio	−0.0002001	0.0119140	0.0135	−0.0148	1.8802
	(0.0003)	(0.0046)		(0.0237)	(0.3368)
Other durables	0.0012377	0.0342129	0.0334	0.0371	2.0259
	(0.0006)	(0.0078)		(0.0165)	(0.2347)
Food off premise	0.0017509	−0.0139890	0.1277	0.0137	0.8904
	(0.0010)	(0.0145)		(0.0080)	(0.1134)
Purchased meals	0.0020063	0.0316373	0.0591	0.0339	1.5353
	(0.0006)	(0.0082)		(0.0098)	(0.1392)
Women's clothing	0.0001581	0.0125307	0.0428	0.0037	1.2925
	(0.0002)	(0.0031)		(0.0050)	(0.0717)
Men's clothing	−0.0000122	0.0108421	0.0184	−0.0007	1.5907
	(0.0001)	(0.0020)		(0.0077)	(0.1089)
Fuel oil and coal	0.0004234	−0.0069874	0.0060	0.0708	−0.1684
	(0.0002)	(0.0029)		(0.0342)	(0.4855)
Tobacco	−0.0010198	−0.0256529	0.0135	−0.0755	−0.8983
	(0.0003)	(0.0045)		(0.0235)	(0.3339)
Household	−0.0010344	−0.0525102	0.0430	−0.0240	−0.2200
	(0.0003)	(0.0048)		(0.0078)	(0.1105)
Other nondurable	0.0002252	0.0118459	0.0201	0.0112	1.5884
	(0.0003)	(0.0039)		(0.0137)	(0.1951)
Owner rent	−0.0006820	−0.0449707	0.1016	−0.0067	0.5573
	(0.0005)	(0.0078)		(0.0054)	(0.0763)
Tenant rent	−0.0007580	0.0032383	0.0380	−0.0200	1.0853
	(0.0004)	(0.0052)		(0.0097)	(0.1376)
Electricity gas	0.0015898	0.0132918	0.0325	0.0488	1.4084
	(0.0006)	(0.0081)		(0.0174)	(0.2480)
Telephone	−0.0010171	−0.0090374	0.0284	−0.0358	0.6817
	(0.0004)	(0.0052)		(0.0128)	(0.1819)
Car care	−0.0013308	0.0030477	0.0280	−0.0476	1.1090
	(0.0006)	(0.0079)		(0.0199)	(0.2832)
Transportation	0.0006045	0.0214435	0.0098	0.0620	3.1975
	(0.0002)	(0.0024)		(0.0172)	(0.2444)
Education	−0.0008962	−0.0047274	0.0597	−0.0150	0.9208
	(0.0010)	(0.0136)		(0.0160)	(0.2281)
Business	0.0001459	0.0122341	0.0202	0.0072	1.6053
	(0.0003)	(0.0038)		(0.0132)	(0.1876)
Other services	−0.0027634	−0.0402280	0.0686	−0.0403	0.4133
	(0.0008)	(0.0108)		(0.0110)	(0.1568)

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

with respect to gasoline prices and with respect to total expenditures. These are immediately obtained from the derivative estimates and the budget share means. For comparative convenience this table includes the derivative estimates from Table 4 that are used to construct the reported elasticities.

The total expenditure elasticities are for the most part quite sensible. Durables generally have high expenditure elasticities as would be expected since consumers can adjust the timing of their purchases of durables to when their incomes are relatively high. The gasoline expenditure elasticity is unexpectedly high, though the corresponding total expenditure derivative is not significantly different from zero.

The own price elasticity for gasoline, −.87, is negative, inelastic, and statistically significant. Recall a novel feature of the methodology is that it provides detailed cross price effects, and in this application these are the separate effects of gasoline prices on every other good consumers purchase. Of course most of these cross effects are small or insignificant, but some are quite interesting. For example, we find when gasoline prices increase 1%, the demand for transportation other than personal car use (planes, trains, buses, etc.) goes up .06 percent, which is useful information for analyzing gasoline tax effects and responses to oil shocks.

To check for possible simultaneity with supply we experimented with reestimating the model replacing R_t in the construction of the instruments Z_{st} with some of the same supply side variables used for that purpose by Lewbel and Ng (2005) and Jorgenson, Lau, and Stoker (1982). We do not report these results because they did not change the overall conclusions in any substantial way, and to maintain greater consistency with our price aggregation theory (particularly Assumption 3(a)).

Overall, our aggregate data results illustrate the feasibility of using our methodology to estimate demand systems with large numbers of equations and goods. We consider this application to be an extreme test of the methodology because our data consists of only 45 time periods, so obtaining any kind of reasonable estimates in a model with 22 goods and no imposed separability seems like a success. (Details on the construction of these goods are provided in Table 6.) Still, these results come with numerous caveats. Estimation of such a large model with relatively few observations means that our asymptotic standard errors, hypothesis tests, and bootstrap confidence intervals likely suffer from relatively large biases associated with small samples.

We have not dealt with possible nonstationarity of prices. As noted earlier, nonstationarity helps to satisfy our price aggregation assumptions to the extent that it makes variation in U small relative to variation in R (as seen in the generally high R^2 statistics in Table 1). However, nonstationarity of regressors in nonlinear models invalidates the standard asymptotic limiting distribution theory underlying our hypothesis tests. Little is known about the distribution of nonlinear model parameter estimates with nonstationary data in general, though see Lewbel and Ng (2005) for a demand system functional form that can be estimated

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

TABLE 6. Groups of goods of the parametric model

Groups of goods	Description
Gasoline	Gasoline and oil
Motor vehicles	New autos, net purchases of used autos, other motor vehicles, tires, tubes, accessories, and other parts
Furniture	Furniture, including mattresses and bedsprings
Appliances	Kitchen and other household appliances, china, glassware, tableware and utensils
Video audio	Video and audio goods, including musical instruments, and computer goods
Other durable	Ophthalmic products and orthopedic appliances, wheel goods, sports and photographic equipment, boats and pleasure aircraft, jewelry and watches, books and maps, other durable house furnishings
Food off premise	Food purchased for off-premise consumption
Purchased meals	Purchased meals and beverages, food furnished to employees including military and food produced and consumed on farms
Women's clothing	Women's and children's clothing and accessories, shoes
Men's clothing	Men's and boys' clothing and accessories except shoes
Fuel oil and coal	Fuel oil and coal
Household	Toilet articles and preparations, semidurable house furnishings, cleaning and polishing preparations, and miscellaneous household supplies and paper products, drug preparations and sundries
Other nondurable	Nondurable toys and sport supplies, stationery and writing supplies, magazines, newspapers, and sheet music, flowers, seeds, and potted plants
Owner rent	Owner-occupied nonfarm dwellings—space rent
Tenant rent	Tenant-occupied nonfarm dwellings—rent
Electricity gas	Electricity, gas, water, and other sanitary services
Telephone	Telephone and telegraph, domestic service, other household operation
Car care	Repair, greasing, washing, parking, storage, rental, and leasing, other user-operated transportation
Transportation	Purchased local transportation (mass transit systems, taxicab), purchased intercity transportation (railway, bus, airline, other)
Education	Higher education, nursery, elementary, and secondary schools, other
Other services	Rental value of farm dwellings, other housing, admissions to specified spectator amusements and other recreation, religious and welfare activities, foreign travel by U.S. residents

with nonstationary price data. Finally, both our parametric functional form and the aggregation across consumers impose strong restrictions that may not be valid empirically. These we relax in our next application.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

7. NONPARAMETRIC ESTIMATION AND TESTING

The estimates presented in the previous section depended on many restrictive, parametric modeling assumptions. In this section we work with consumption data from a relatively homogeneous sample of individual households. The resulting elasticity estimates therefore have less general policy relevance than the previous macro estimates, but these data permit direct nonparametric estimation of elasticities and testing. In particular, we can directly estimate the functions $m(r, y, z)$ by nonparametric regressions of w on r, y, z and can nonparametrically test if they are homogeneous and if the corresponding Slutsky related matrix $\widetilde{S}(r, y, z)$ is symmetric and negative semidefinite.

The data are observations of expenditures, demographic composition, and other characteristics of a demographically relatively homogeneous set of households drawn from the years 1974 to 1993 of the British Family Expenditure Survey (FES). The sample consist of about 7,000 households, all consisting of childless couples with at least one employed member and with the head of the household being a white collar worker. To account for remaining observable demographic heterogeneity in a nonparametrically tractable way, we include a covariate Z constructed as the first principal component of the the remaining demographic household characteristics reported in the FES. The expenditures of all goods are grouped into 12 categories, which are Food, Catering, Alcohol, Tobacco, Housing, Household Goods, Household Services, Clothing, Motoring, Fares, Leisure Services, and Gasoline. We construct R to be three-dimensional, with first element being the log price of gasoline as before and the second two elements as the suitably scaled first two principal components of the log price indexes of the remaining 11 goods, using the same scaling procedure as with the aggregate data example. We also construct Γ using the same procedure as with the aggregate data. We experimented with including more price and demographic components, but these caused problems associated with the nonparametric curse of dimensionality

We nonparametrically estimate the demand functions $m(r, y, z)$ using kernel based local polynomial regressions and apply bootstrap methods to obtain critical values for hypothesis test statistics as in Hoderlein (2005) and Haag, Hoderlein, and Pendakur (2009). The reader may refer to those papers for associated econometric distribution theory and details. The principal components R are estimated at a parametric rate, and $m(r, y, z)$ is a smooth function of r, so by a standard Taylor expansion of the estimated m around the true R, one may readily show that (as usual for a parametric generated regressor in a nonparametric regression) estimation error in R is to first-order asymptotically irrelevant for estimation of m.

Endogeneity of total expenditure could be an issue, so we estimate the model a second time allowing for endogeneity of y by using labor income as an instrument and including corresponding control function residuals in the nonparametric regressions as in Newey, Powell, and Vella (1999). With this household level data we do not instrument prices, both for consistency with our formal

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

assumptions (particularly Assumption 3(a)) and because the contribution of any one consumer's demands (or even all consumers in the subset of the population we sample from) to aggregate demand, and hence to price determination, should be small.

Table 7 reports the mean elasticity of each budget share with respect to gasoline prices and with respect to total expenditures, along with bootstrapped

TABLE 7. Nonparametric model estimates

	Mean budget share	Elasticity gas price	Elasticity total expenditures
Gasoline	0.0413	−0.8359	0.8228
		(−1.1409; −0.2786)	(0.7104; 0.9268)
Food	0.1482	0.0626	0.2601
		(−0.0978; 0.2333)	(0.2150; 0.2927)
Catering	0.0473	0.1061	1.4109
		(−0.2752; 0.4708)	(1.3316; 1.5069)
Repair	0.0231	−0.2641	1.5104
		(−0.3162; 0.3696)	(1.2460; 1.6978)
Fuel oil and coal	0.0527	−0.0718	0.3390
		(−0.3215; 0.1663)	(0.2798; 0.3864)
Furniture	0.0526	0.1197	1.3370
		(−0.1954; 0.4948)	(1.1496; 1.4589)
Other household	0.0218	−0.0610	0.4932
		(−0.5344; 0.1103)	(0.4128; 0.5983)
HH services	0.0441	0.2334	1.2238
		(−0.0638; 0.6112)	(1.1236; 1.2998)
Men's clothing	0.0123	−0.0103	1.5378
		(−0.5107; 0.3470)	(1.2326; 1.8457)
Women's clothing	0.0463	0.0166	1.3670
		(−0.5695; 0.4504)	(1.2442; 1.5144)
Personal care	0.0267	−0.2039	1.0906
		(−0.4539; 0.1945)	(0.9866; 1.1860)
Motor vehicles	0.0479	0.1352	1.5124
		(−0.1644; 0.5643)	(1.2514; 1.6680)
Other motoring	0.0424	−0.4848	1.0603
		(−1.0108; 0.0793)	(0.9488; 1.1882)
Transportation	0.0211	0.0308	1.4781
		(−0.3975; 0.4541)	(1.2061; 1.6570)
Leisure goods	0.0230	0.1294	1.0129
		(−0.2638; 0.4659)	(0.9086; 1.0825)
Leisure services	0.0637	−0.2334	1.7671
		(−0.6906; 0.1731)	(1.6291; 1.8806)
Rent and mortgage	0.1165	−0.4687	1.3118
		(−0.8023; 0.0191)	(1.2112; 1.4044)

Note: Numbers in parentheses are lower and upper bounds of the 95% confidence interval.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

95% confidence intervals. The results show similar patterns to the aggregate data, such as high total expenditure elasticities for durables. Gasoline demand itself has almost the same own price elasticity as in the aggregate data and shows many similar cross price effects (again including a small positive cross effect on other transportation), though these nonparametric cross price effects are not estimated with enough precision to be statistically significant.

Table 8 lists results of various tests of rationality restrictions. We evaluate the estimated nonparametric demand functions and their derivatives at each data point, and for each observation we test whether the demand functions at that point satisfy homogeneity, negative semidefiniteness, and Slutsky symmetry. For each hypothesis, Table 8 lists the percentage of observations at which the hypothesis is not rejected at the 0.95 confidence level.

Table 8 shows that homogeneity (the absence of money illusion) is generally accepted, with a rejection rate of 11% of the data when we allow for endogenous regressors. We test symmetry and negative semidefiniteness of the composite commodity Slutsky matrix $\widetilde{S}(r, y, z)$ both with and without imposing homogeneity. Note that imposing homogeneity speeds the rate of convergence by reducing dimensionality (see, e.g., Tripathi and Kim, 2003). We reject symmetry and negative semidefiniteness in 13% to 26%, depending on whether or not we impose homogeneity or allow for endogeneity.

Overall, we find that our method of handling price multicolinearity through price dimension reduction yields reasonable nonparametric demand model estimates that are usually consistent with consumer rationality. The somewhat lower acceptance rates for symmetry versus homogeneity could be due to the fact that symmetry involves a greater number of restrictions, or some couples might employ bargaining or other non–social welfare maximizing expenditure allocation schemes that yield symmetry violations of the unitary model as in Browning and Chiappori (1998).

TABLE 8. Nonparametric model tests: Percentage of population not rejecting each hypothesis at the 95% confidence level

Hypothesis	% not rejected
Homogeneity under exogeneity	80
Homogeneity under endogeneity	89
Negative semidefiniteness under exogeneity	81
Negative semidefiniteness under exogeneity and homogeneity	82
Negative semidefiniteness under endogeneity and homogeneity	87
Symmetry under exogeneity	74
Symmetry under exogeneity and homogeneity	78
Symmetry under endogeneity and homogeneity	80

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

TABLE 9. Groups of goods of the nonparametric model

Groups of goods	Description
Gasoline	Gasoline and oil
Food	Bread, meat and fish, milk products, drinks, fruits and vegetables, other food
Catering	Restaurant meals, canteen meals, take-away meals and snacks
Repair	Payments to contractors, including those for house painting and decorating, purchase of materials and tools (other than electrical tools)
Fuel oil and coal	Fuel oil and coal
Furniture	New and secondhand furniture, floor coverings, household textiles, mattresses, pillows, cushions, etc., electrical appliances (excluding audiovisual equipment) and other household equipment
Other household	Soap and detergents, other cleaning materials, toilet paper, and other
Household services	Postage, telephone charges, telephones, telemessages, domestic services, fees and subscriptions
Men's clothing	Men's clothing, including shirts and sportswear
Women's clothing	Women's clothing, children's clothing, other clothing and footwear
Personal care	Personal articles, chemist's goods, personal services
Motor vehicles	Net purchases of new and secondhand cars, vans, motorcycles, scooters
Other motoring	Repair and servicing of motor vehicles, spares and accessories, vehicle tax and insurance
Transportation	Rail fares, bus fares, taxi fares, car hire charges and other travel costs
Leisure goods	Records, toys, photo and sports goods, personal articles, books and newspapers, gardening products
Leisure services	TV licenses and rentals, entertainment and recreation, foreign holidays and UK holidays
Rent and mortgage	Rent, excluding rates and other charges paid together with rent, interest payments on a representative mortgage

8. CONCLUSIONS

We have shown how demand systems can be estimated, with rationality imposed, while replacing the complete set of prices of all goods with a small set of prices of interest and a few summary price indexes. We illustrate the results by reporting estimates of the effects of gasoline prices on the demands for many goods, both with an aggregate data parametric model to obtain some macroeconomic elasticities and nonparametrically for a demographically homogeneous sample of individual households. The methodology is easy to implement either in parametric demand models or using ordinary nonparametric regressions.

In parametric models, this dimension reduction mitigates problems of multicollinearity of prices. In our parametric model, we replace J equals 22 separate prices

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

with K equals four price aggregates. The gains from reducing dimensionality of prices in the nonparametric context are more dramatic. In our nonparametric application we have J equals 12 prices that we reduce to K equals three price aggregates, along with a demographic and a total expenditures variable. Our methodology therefore reduces the dimension of these nonparametric demand function regressions from 14 to five regressors, which yields an enormous increase in rates of convergence without behavioral demand function restrictions.

More generally, we have shown how statistical dimension reduction techniques can be applied in optimizing models arising from microeconomic theory, which have multiple nonlinear equations, nonseparable unobservables, nonlinear cross equation restrictions, and many potentially multicolinear covariates.

NOTE

1. The undesirable result is that, to empirically implement modern flexible models that reduce or eliminate restrictions on second and higher order cross price elasticities between groups, empirical practioners have assumed a greater degree of separability (fewer groups), thereby implicitly imposing severe restrictions on the first-order cross price elasticities of individual goods within groups.

REFERENCES

Altonji, J.G. & R.L. Matzkin (2005) Cross section and panel data estimators for nonseparable models with endogenous regressors. *Econometrica* 73, 1053–1102.

Berry, S., J. Levinsohn, & A. Pakes (1995) Automobile prices in market equilibrium. *Econometrica* 63, 841–890.

Blackorby, C., D. Primont, & R.R. Russell (1978) *Duality, Separability, and Functional Structure: Theory and Economic Applications*. North-Holland.

Blundell, R.W. & J.L. Powell (2004) Endogeneity in semiparametric binary response models. *Review of Economic Studies* 71, 655–679.

Blundell, R. & J.M. Robin (2000) Latent separability: Grouping goods without weak separability. *Econometrica* 68, 53–84.

Brown, R. & M. Walker (1989) The random utility hypothesis and inference in demand systems. *Econometrica* 57, 815–829.

Browning, M. & P.A. Chiappori (1998) Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica* 66, 1241–1278.

Christensen, M. (2004) Integrability of Demand Accounting for Unobservable Heterogeneity: A Test on Panel Data. Manuscript. Centre of Applied Microeconomics, University of Copenhagen.

Davis, G.C. (2003) The generalized composite commodity theorem: Stronger support in the presence of data limitations. *Review of Economics and Statistics* 85, 476–480.

Deaton, A. & J. Muellbauer (1980) An almost ideal demand system. *American Economic Review* 70, 312–326.

Gorman, W.M. (1959) Separable utility and aggregation. *Econometrica* 27, 469–481.

Gorman, W.M. (1976) Tricks with utility functions. In M.J. Artis & A.R. Nobay (eds.), *Essays in Economic Analysis: Proceedings of the 1975 AUTE Conference, Sheffield*, pp. 211–243. Cambridge University Press.

Haag, B., S. Hoderlein, & K. Pendakur (2009) Testing and imposing Slutsky symmetry in nonparametric demand systems. *Journal of Econometrics* 153, 33–50.

Hicks, J.R. (1936) *Value and Capital*. Oxford University Press.

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

Hoderlein, S. (2005) Nonparametric Demand Systems, Instrumental Variables and a Heterogeneous Population. Manuscript, Brown University.
Hoderlein, S. & E. Mammen (2007) Identification of marginal effects in nonseparable models without monotonicity. *Econometrica* 75, 1513–1518.
Imbens, G.W. & W.K. Newey (2009) Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica* 77, 1481–1512.
Jorgenson, D.W., L.J. Lau, & T.M. Stoker (1982) The transcendental logarithmic model of aggregate consumer behavior. In R.L. Basmann & G. Rhodes (eds.), *Advances in Econometrics*, vol. 1, pp. 97–238. JAI Press.
Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl, & Lee, T. (1985) *The Theory and Practice of Econometrics*, 2nd ed. Wiley.
Leontief, W. (1936) Composite commodities and the problem of index numbers. *Econometrica* 4, 39–59.
Lewbel, A. (1991) The rank of demand systems: Theory and nonparametric estimation. *Econometrica* 59, 711–730.
Lewbel, A. (1996) Aggregation without separability: A generalized composite commodity theorem. *American Economic Review* 86, 524–543.
Lewbel, A. (2001) Demand systems with and without errors. *American Economic Review* 91, 611–618.
Lewbel, A. & S. Ng (2005) Demand systems with nonstationary prices. *Review of Economics and Statistics* 87, 479–494.
Newey, W.K., J.L. Powell, & F. Vella (1999) Nonparametric estimation of triangular simultaneous equations models. *Econometrica* 67, 567–603.
Tripathi, G. & W. Kim (2003) Nonparametric estimation of homogeneous function. *Econometric Theory* 19, 640–663.
White, H. (1980) Heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48, 817–836.
Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

APPENDIX A: Technical Assumptions and Extensions

To allow dependence of U on R we now consider replacing Assumption 4 with the following assumption.

Assumption 5. The conditional distribution of U_2 given $R = r$ is continuous, with conditional pdf $f_{U_2|R}(U_2 \mid r)$ that is differentiable in r.

We then obtain Theorems $1'$, $2'$, and $3'$, which follow, which extend Theorems 1, 2, and 3, respectively. Generally speaking, the effect of Assumption 5 in place of Assumption 4 is to add additional terms that depend upon the conditional density of U_2.

THEOREM $1'$. *Let Assumptions 1–3 and 5 hold. Then*

$$D_r m(r,y,z) = \mathbb{E}\left[D_p\phi(P,Y,Z,A)\Gamma + W D_r \log f_{U_2|R}(U_2 \mid R) \mid R = r, Y = y, Z = z\right]. \tag{A.1}$$

We introduce the notation $Cor(r,y,z) = -\Gamma'\mathbb{E}\left[W D_r \log f_{U_2|R}(U_2 \mid R) \mid R = r, Y = y, Z = z\right]$.

THEOREM $2'$. *Let Assumptions 1–3 and 5 hold. If $\phi(p,y,z,a)$ satisfies adding up then $m(r,y,z)$ satisfies adding up. If $\phi(p,y,z,a)$ is homogeneous and $f_{U_2|R}(U_2 \mid r) = f_{U_2|R}(U_2 \mid r+\lambda)$ for any scalar λ then $m(r,y,z)$ is homogeneous.*

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

Theorem 2′ requires that the conditional density of the errors U in the price dimension reduction equation be homogeneous of degree zero in the price indexes $\exp(R)$. This is empirically testable in a variety of ways; e.g., one could directly apply general nonparametric specification tests to test the hypothesis that $f_{U_2|R}(U_2 \mid R)$ equals the conditional density of U_2 just conditioning on $R - R_J$ where R_J is the last element of R. This test would use a parametric estimate of U_2 and kernel estimates of the required conditional densities.

An implication of this density homogeneity is that the conditional variance of U given R is homogeneous. A simple parametric test of the hypothesis would be to regress $U_2 U_2'$ on $R - R_J$ and R_J and test if the coefficients on R_J are insignificant.

THEOREM 3′. *Let Assumptions 1–3 and 5 hold. Assume $f_{U_2|R}(U_2 \mid r)$ is differentiable in r. Then*

$$\mathbb{E}\left[\Gamma' S(P, Y, Z, A)\Gamma \mid R = r, Y = y, Z = z\right] = \widetilde{S}(r, y, z) + H(r, y, z) + G(r, y, z) + Cor(r, y, z). \tag{A.2}$$

Therefore, if $S(p, y, z, a)$ is negative semidefinite then $\widetilde{S}(r, y, z) + H(r, y, z) + G(r, y, z) + Cor(r, y, z)$ is negative semidefinite and if $S(p, y, z, a)$ is symmetric then $\widetilde{S}(r, y, z) + H(r, y, z) + Cor(r, y, z)$ is symmetric. In addition, if S is negative semidefinite then

$$D_r \widetilde{m}(r, y, z) + \Gamma'\left\{D_y m_2(r, y, z) + 2(m_2(r, y, z))\right\}\Gamma - \operatorname{diag}\left\{\widetilde{m}(r, y, z)\right\} + Cor(r, y, z)$$

is negative semidefinite.

APPENDIX B: Proofs and Tables

Proof of Theorem 1 and 1′. By boundedness of W the expectation of W exists and therefore also the conditional expectation

$$\begin{aligned} m(r, y, z) &= \mathbb{E}[\phi(P, Y, Z, A) \mid R = r, Y = y, Z = z] \\ &= \int_{\mathcal{A}\times\mathcal{U}_2} \phi(\Gamma b + u, y, z, a) F_{U_2, A|R, Y, Z}(du_2, da \mid r, y, z). \end{aligned}$$

Note that the first K elements of u are zero by construction and the remaining elements are u_2. By Assumption 3(a), we may separate the integration over U and R from the (A, Y, Z), and we may use the conditional independence of A from Y to drop y from the conditional distribution of A. Summarizing, we obtain

$$m(r, y, z) = \int_{\mathcal{A}} \left(\int_{\mathcal{U}_2} \phi(\Gamma r + u, y, z, a) f_{U_2|R}(u_2 \mid r) du_2 \right) F_{A|Z}(da \mid z). \tag{A.1}$$

For Theorem 1, replace $f_{U_2|R}(u_2 \mid r)du_2$ with $F_{U_2}(du_2)$ in the preceding expression. Equation (2) follows immediately by taking the derivative of equation (A.1) with respect to y and interchanging differentiation and integration, given bounded convergence. Finally, taking the derivative with respect to r gives

$$D_r m(r, y, z) = \int_{\mathcal{A}} \left(\int_{\mathcal{U}_2} D_r \left[\phi(\Gamma r + u, y, z, a) f_{U_2|R}(u_2 \mid r)\right] du_2 \right) F_{A|Z}(da \mid z),$$

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

and

$$D_r\left[\phi(\Gamma r+u,y,z,a)f_{U_2|R}(u_2\mid r)\right]$$
$$=D_p\phi(p,y,z,a)\Gamma f_{U_2|R}(u_2\mid r)+\phi(p,y,z,a)D_r f_{U_2|R}(u_2\mid r)$$
$$=\left[D_p\phi(p,y,z,a)\Gamma+\phi(p,y,z,a)D_r f_{U_2|R}(u_2\mid r)\right]f_{U_2|R}(u_2\mid r),$$

which gives equation (A.1). The same derivation replacing $f_{U_2|R}(u_2\mid r)du_2$ with $F_{U_2}(du_2)$ makes the second term in this expression be zero, yielding equation (4). Equation (3) follows from equation (4) given $r_1=p_1$ and the structure of Γ. ■

Proof of Theorem 2 and 2′. Adding a constant λ to each element of r in equation (A.1) gives

$$m(r+\lambda,y,z)=\int_{\mathcal{A}}\left(\int_{\mathcal{U}_2}\phi(\Gamma r+u+\Gamma 1_K\lambda,y,z,a)f_{U_2|R}(u_2\mid r+\lambda)du_2\right)F_{A|Z}(da\mid z).$$

Now $\Gamma 1_K=1_J$ so $\phi(\Gamma r+u+\Gamma 1_K\lambda,y,z,a)=\phi(\Gamma r+u,y,z,a)$ if $\phi(p+\lambda,y,z,a)=\phi(p,y,z,a)$, and so if we also have either $f_{U_2|R}(u_2\mid r+\lambda)=f_{U_2|R}(u_2\mid r)$ or if $f_{U_2|R}(u_2\mid r)$ is replaced with $F_{U_2}(du_2)$, then $m(r+\lambda,y,z)=m(r,y,z)$. ■

Proof of Theorem 3 and Theorem 3′. For compactness of notation, let conditioning on r,y,z denote conditioning on the event $R=r,Y=y,Z=z$. Define $\widetilde{\phi}(P,Y,Z,A)=\Gamma'\phi(P,Y,Z,A)$. Dropping the argument (P,Y,Z,A) from ϕ and $\widetilde{\phi}$ for clarity we have $\widetilde{m}(r,y,z)=\mathbb{E}\left[\widetilde{\phi}\mid r,y,z\right]$. Multiplying equation (2) in Theorem 1 by Γ' gives $D_y\widetilde{m}(r,y,z)=\Gamma' D_y m(r,y,z)=\Gamma'\mathbb{E}\left[D_y\phi\mid r,y,z\right]=\mathbb{E}\left[D_y\left(\Gamma'\phi\right)\mid r,y,z\right]=\mathbb{E}\left[D_y\widetilde{\phi}\mid r,y,z\right]$. Similarly, multiplying equation (4) in Theorem 1 by Γ' gives $D_r\widetilde{m}(r,y,z)=\Gamma' D_r m(r,y,z)=\Gamma'\mathbb{E}\left[D_p\phi\mid R=r,Y=y,Z=z\right]\Gamma$. Substituting these into the definition of $\widetilde{S}(r,y,z)$ gives

$$\widetilde{S}(r,y,z)=\Gamma'\mathbb{E}\left[D_p\phi\mid r,y,z\right]\Gamma+\mathbb{E}\left[D_y\widetilde{\phi}\mid r,y,z\right]\mathbb{E}\left[\widetilde{\phi}\mid r,y,z\right]'$$
$$+\mathbb{E}\left[\widetilde{\phi}\mid r,y,z\right]\mathbb{E}\left[\widetilde{\phi}\mid r,y,z\right]'-\operatorname{diag}\left\{\mathbb{E}\left[\widetilde{\phi}\mid r,y,z\right]\right\}.$$

Now by the definition of the Slutsky matrix S we have

$$\mathbb{E}\left[\Gamma' S\Gamma\mid r,y,z\right]=\Gamma'\mathbb{E}\left[D_p\phi\mid r,y,z\right]\Gamma$$
$$+\mathbb{E}\left[\left(D_y\left(\Gamma'\phi\right)\right)\left(\phi'\Gamma\right)+\Gamma'\phi\phi'\Gamma-\Gamma'\operatorname{diag}\{\phi\}\Gamma\mid r,y,z\right]$$
$$=\Gamma'\mathbb{E}\left[D_p\phi\mid r,y,z\right]\Gamma+\mathbb{E}\left[\left(D_y\widetilde{\phi}\right)\widetilde{\phi}'\mid r,y,z\right]$$
$$+\mathbb{E}\left[\widetilde{\phi}\widetilde{\phi}'\mid r,y,z\right]-\Gamma'\operatorname{diag}\{\mathbb{E}[\phi\mid r,y,z]\}\Gamma.$$

So

$$\mathbb{E}\left[\Gamma' S\Gamma\mid r,y,z\right]-S(r,y,z)=\operatorname{Cov}\left[D_y\widetilde{\phi},\widetilde{\phi}'\mid r,y,z\right]+\operatorname{Var}\left[\widetilde{\phi}\mid r,y,z\right]$$
$$+\operatorname{diag}\left\{\mathbb{E}[\widetilde{\phi}\mid r,y,z]\right\}-\Gamma'\operatorname{diag}\{\mathbb{E}[\phi\mid r,y,z]\},\Gamma,$$

which gives equation (7). Semidefiniteness and symmetry of S imply semidefiniteness and symmetry of $\mathbb{E}\left[\Gamma' S\Gamma\mid r,y,z\right]$, which given (7) and symmetry of G proves the rest of the

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press

theorem. The proof of Theorem 3 follows exactly be the same arguments, including the correction term as in Theorem $1'$.

Finally, to see the last statement in Theorem $3'$ (regarding the matrix L in Theorem 3), note that semidefiniteness of S implies semidefiniteness of $\mathbb{E}\left[\Gamma'\left(S+S'\right)\Gamma \mid r, y, z\right]$. Next, note that

$$\mathbb{E}\left[\Gamma'\left(D_y\phi\phi'+\phi D_y\phi'\right)\Gamma \mid r, y, z\right] = \Gamma' D_y \mathbb{E}\left[WW' \mid r, y, z\right]\Gamma = D_y\widetilde{m}_2(r, y, z),$$

where we used again the conditional independence of A and Y for the first equality. The result follows by collecting terms and noticing that the last two terms are symmetric. ■

https://doi.org/10.1017/S0266466612000412 Published online by Cambridge University Press