Imposing and Testing for Shape Restrictions in Flexible Parametric Models[1]

PUBLISHED AS:

Wolff, H. (2016): *"Imposing and Testing for Shape Restrictions in Flexible Parametric Models", Econometric Reviews. Vol. 35(6), pp. 1013-1039.*

Abstract

In many economic models, theory restricts the shape of functions, such as monotonicity or curvature conditions. This paper reviews and presents a framework for constrained estimation and inference to test for shape conditions in parametric models. We show that 'regional' shape-restricting estimators have important advantages in terms of model fit and flexibility (as opposed to standard 'local' or 'global' shape-restricting estimators). In our empirical illustration, this is the first paper to impose and test for all shape restrictions required by economic theory simultaneously in the "Berndt and Wood" data. We find that this dataset is consistent with 'duality theory', whereas previous studies have found violations of economic theory. We discuss policy consequences for key parameters, such as whether energy and capital are complements or substitutes.

Keywords: Demand System, Cost Function, Shape Restrictions, Flexible Functional Forms, Hypothesis Test, Elasticity of Substitution, Energy

JEL Code: C51 - Model Construction and Estimation; D21 - Firm Behavior; C11 - Bayesian Analysis; Q43 - Energy and the Macroeconomy; C52 - Model Evaluation and Testing

1 I wish to thank Guido Imbens, Jeffrey LaFrance, Daniel Brent, Jen Brown, Yanay Farja and Muzhe Yang for motivation, and valuable discussions during the course of this research. Furthermore, thanks are due Ernst Berndt, for providing the data set. Financial support from the Center for Statistics and the Social Sciences at the University of Washington is greatly appreciated.

While a difficult literature, we believe that research on models permitting flexible imposition of true regularity should expand.

-- William A. Barnett & Meenakshi Pasupathy, 2003 --

1. INTRODUCTION

A lot of work in economics is characterized as follows: A researcher makes a set of behavioural assumptions, develops a theoretical model and derives functions of interest. An econometrician then estimates these functions. Finally, we use the resulting empirical model for policy analysis. The data provide the link between 'economic theory' and the 'empirical model'. Two fundamental questions motivate this paper: is the data in line with economic theory, and how can we test for this relationship?

One critical implication of economic theory is that behavioural assumptions manifest themselves in the form of uniquely defined 'shape conditions'. For example, if it is assumed that a firm shows the behaviour of a cost minimizer, then the dual cost function is concave and monotonically increasing in input prices.[2] The potential gap between a well-established economic theory, on one side, and the empirical model on the other, is of great concern as reflected in the large literature on regularity preserving estimation procedures of the past 30 years, see Gallant and Golub (1984), Diewert and Wales (1987) and Barnett and Binner (2004) for literature reviews on this in *parametric* functional forms. For a literature review on the more recently developing *non-parametric* shape-preserving methods, see Henderson and Parmeter (2009). Nonetheless, in empirical applications shape conditions are reported to be often violated and policy recommendations derived from such models are dubious at best; see the examples in Salvanes and TjØtta (1998), Griffiths, O'Donnell and Tan-Cruz (2000), Barnett (2002) or Blundell (2004).

This paper reviews and presents a framework for estimation and inference in flexible parametric models to test for shape conditions, hence to test whether the behavioural assumptions of the underlying economic theory is reflected in the data. In reviewing estimators, we differentiate methods in

- 'standard econometric techniques' which have been traditionally used to impose shape

[2] By the Shephard Lemma, a shape consistent dual cost function implies that the input demand system is downward sloping and the law of demand holds. Similar relationships between 'behaviour' and 'shape conditions' hold in many other contexts: if individuals are utility maximizers, then the indirect utility function is quasi-convex in prices. If firms maximize profits, then supply is upward sloping and the profit function is convex in output and input prices.

restrictions either *locally* (at one singular point of the regressor space) or *globally* (at the entire domain) and

- the class of *regional* shape-restricting estimators (imposing the shape restrictions on a subset of the domain).

Importantly, in many applications, economic theory has to hold over a continuous subset of the regressor space—in short: ‘region’. This region, first proposed by Gallant and Golub (1984), is defined as the *‘empirically relevant region’*, where the researcher is drawing inferences from—hence the area of the values of the regressor data itself and the area where policy simulations/forecasts are performed. With standard estimators, however, economic theory is formally tested either locally (at one singular point of the domain) or it is imposed globally (at the entire domain, typically leading to very restrictive functional forms). In contrast, regional testing has been ad hoc. For example, one could report the percentage of data points at which the estimated function violates shape conditions. This informal test statistic, however, is problematic. In fact, Figure 1 and 2 below will we show that a function violating shape conditions at all data points can be very “close” to economic theory. The reverse is also possible: one single violation can generate an empirical model completely out of sync with economic theory. This paper reviews and illustrates a formal testing procedure which outperforms such ad-hoc tests.

The setup of the testing procedure is to first estimate a flexible parametric function unrestrictedly. If it satisfies all required shape restrictions over the ‘region’, then the estimated empirical model does not reject the assumed underlying economic theory. If, however, the estimated model violates shape restrictions, a second procedure follows: Building upon work by Gallant and Golub (1984), Terrell (1996) and Wolff *et al*. (2010), we re-estimate the function, subject to the (infinite number of) constraints that all shape conditions are satisfied within the connected ‘region’. Finally, the comparison of the regionally restricted estimate to the unrestricted estimate provides us with the test statistic. The intuition of the proposed test is simple: if we statistically reject the shape properties, then we reject the behavioural assumptions of the underlying economic theory as well. Hence, the objective is to test the gap between the empirical model and the economic theory.

We apply the discussion to the "Berndt and Wood" dataset—that has been extensively used to test the performance of new estimators in the econometric literature—estimating a flexible input demand system of four production factors to the U.S. In fact, it may even be the most frequently used dataset to test new techniques, see e.g. Berndt and Wood (1975), Berndt and Khaled (1979), Galant and Golub (1984), Diewert and Wales (1987), Barnett, Geweke and Wolfe (1991), Friesen (1992) and Terrell (1996).

Here we take a fresh look at the Berndt and Wood data and make the following contributions. First, whereas previous applications only impose a subset of shape conditions, this is the first paper enforcing *all* shape conditions derived from conventional economic theory in the Berndt and Wood data[3]. Second, we discuss which estimators have the technical flexibility to impose all of the required shape conditions and find that 'regional' techniques outperform 'globally' regular estimators (as these latter can only impose a small subset of these conditions without overly restricting the flexibility of the functional form). Regional estimators also outperform local estimators as these do not guarantee regularity away from one singular point. Third, our empirical results demonstrate that estimator choice[4] can lead to significantly different implications. In particular, we find that estimates based on standard econometric techniques would have erroneously rejected economic theory, when evaluated over the '*empirically relevant region*'. In contrast, the newer class of regional regularity estimators provide strong support that the Berndt and Wood data is in fact consistent with economic theory, when evaluated over the same '*empirically relevant region*'. Fourth, we show how a new non-convex 'string' approach by Wolff et al. (2010) substantially (i) reduces computing time and (ii) increases model fit compared to the previous approach to define the region as one convex hypercube, as i.e. constructed by Gallant and Golub (1984) or Terrell (1996).

Finally, one of the central parameters in the Berndt and Wood data is the elasticity of substitution

---

[3] We are not the first to impose multiple restrictions simultaneously. Parmeter et al. (2013) recently presents an estimation framework to impose a large number of constrains simultaneously using a constraint weighted bootstrapping method. In their application, Parmeter et al. (2013) impose seven monotonicity constraints on a Translog of estimated input distance functions.

[4] *Estimator choice* means choosing among the three classes of estimators: local, regional and global regularity imposing estimators. And within the class of regional estimators, it is choosing between the string approach and the convex hypercube approach. In addition the unconstrained estimator will be used for testing purposes and to provide a baseline in the empirical section.

between energy and capital.[5] We show that estimated elasticities by standard econometric techniques can be dramatically different compared to the results produced by the regional approach. In particular, for the Berndt and Wood data we find that energy and capital are complements rather than substitutes, and we discuss the potential policy implications of this result.

The rest of the article is structured as follows. Section 2 provides background on the theory of shape conditions and reviews local, regional and global regularity-imposing techniques. Section 3 compares these approaches in terms of model fit and theoretical consistency. Using the famous Berndt and Wood data, we then test for shape conditions and present estimation results regarding the distribution of the elasticity of substitution matrix. Section 4 concludes.

2. BACKGROUND

2.1 Shape Conditions

This section first reviews the concept of shape conditions, with a specific example from duality theory, which will be then applied in Section 3 to estimate the Berndt and Wood data. An input factor demand system is estimated which is derived by Shephard's Lemma from the dual cost function $c^*(p,y,t) = \min\{p`x \mid f(x,t) \geq y\}$. Hence it is assumed that firms exhibit the behaviour of a cost minimizer. Inputs are denoted by $x \in \Re^N_+$ and are transformed into output $y \in \Re^1_+$ by a smooth production function $f(x,t)$, which depends upon technological change $t \in \Re^1_+$.

Duality theory implies that $c^*(z)$, where $z = [p,y,t] \in \Re_+^{2+N}$, is

HD1 : homogenous of degree one in input prices $p \in \Re^N_+$

M_p : monotone increasing in p

C_p : concave in p

Furthermore, if firms operate in a competitive environment, we have that

[5] Whether energy and capital are substitutes or complements is a fundamental question that has received much attention in economics, i.e. see Dasgupta & Heal (1979) or Apostolakis (1999).

C_y : convex in y, and

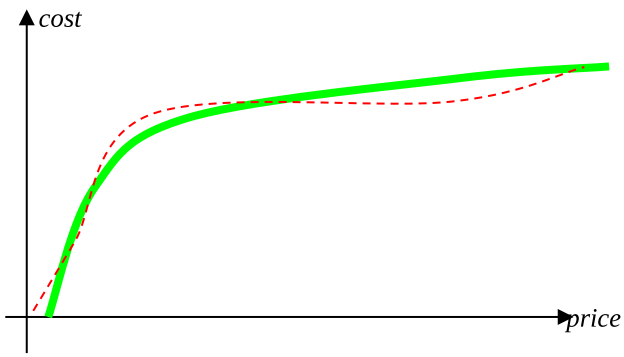
M_y : monotone increasing in y.

If a function satisfies all five properties HD1, $M_{p,y}$ and $C_{p,y}$, we say that the function is 'regular' or synonymously, 'well-behaved.' To impose these regularity conditions, for many years the literature concentrated on the estimation of factor demand systems from globally regular generating functions, such as the Cobb-Douglas and the Constant Elasticity of Substitution. These *first order flexible functional forms* satisfy the restrictions of homogeneity, monotonicity, and curvature by well-known parametric restrictions; at the same time, these forms restrict the potential values for the second order effects prior to the estimation. This implies that the elasticities of substitutions—which are often key parameters in policy analysis—cannot be estimated, but are fixed. Moreover, this does not allow formal testing of the underlying economic theory because these first order flexible functions are strictly a subset within the class of functions generated by the theory.

Due to the restrictiveness of first order flexible forms, the literature moved toward local approximation functions to the true data generating function itself, by series expansions. The result was the class of *second order flexible functional forms*, such as the popular Translog, Generalized McFadden, and the Generalized Leontief, providing the capability to attain arbitrary elasticities of substitution. Nevertheless, this increased flexibility came with a cost: sacrificing the guarantee of regularity.[6] In fact, a series of studies (for a review, see Barnett, Geweke and Wolfe (1991)) demonstrated that these forms often have very small regions of theoretical regularity.

[6] For linear-in-the-parameters functional forms, Lau (1986) proved that flexibility is incompatible with global regularity with the imposition of both concavity and monotonicity. For example, a globally consistent second order Translog reduces the feasible parameter values of its squared terms to be zero, thus restricting the functional form to its first order series expansion, the Cobb-Douglas, which has constant cross elasticities of value one.

Figure 1: True versus approximation function



Even if the true data generating cost function c^* is deterministic and regular (bold), the estimated approximation function $\hat{c}$ (dashed) can be irregular. The movement from c^* to $\hat{c}$ is not a simple parallel movement. Instead, (because realized (weighted) residuals sum up to zero) $\hat{c}$ oscillates around c^*. This phenomenon has been demonstrated with numerical examples using the Translog, the Generalized Leontief, and the AIM (Wolff *et al.* 2010).

Why do we need regularity preserving estimators? On one side Moschini (1999) raises the issue of whether one should really force the function to satisfy economic theory (or if one should rather let the data speak for itself and search for other (non-neoclassical) explanations). However, we argue, regularity-preserving techniques are indispensable for at least five reasons:

(1) Any finite order flexible functional form c represents an approximation to the true function c^*. If c^* is regular and *stochastic*, then $\hat{c}$, estimated with some non-regularity preserving estimator, can fit outliers produced by c^* and thus violate regularity (even though the determinist part of c^* is regular!).

(2) Even if c^* was regular and *deterministic*, $\hat{c}$ can oscillate around the true relationship. Because of its approximating nature, $\hat{c}$ has a different tracking behaviour over its domain, so it does not lie completely above c^*, but slightly next to it, as shown in Figure 1. This is perhaps the most important reason to use regularity-retaining techniques. Otherwise one risks erroneously concluding that data is ill behaved, whereas, in fact, the true data generation process is regular.

(3) A regularity preserving point estimate is required for correctly specifying hypothesis tests. More on this issue is outlined in the empirical section.

(4) Numerical iterations between parametric functions are often used to obtain non-analytic results (i.e.

iterations between a demand system and a supply system, until convergence).[7] If such estimated functions violate curvature conditions, this can lead to local equilibria or numerical unbounded results.

(5) In natural sciences, many technological relationships (i.e. acceleration of a rocket) are known to be shape restricted, but measurement error produces oscillating results. To differentiate between measurement error and functional form, shape preserving techniques are used.[8]

2.2 The Domain

Following the terminology as suggested in Wolff *et al*. (2010), regularity imposing methods can be categorized into three classes: (a) global, (b) local and (c) regional. Let us define $\boldsymbol{\psi}$ as a subset of the domain of the right hand side variable space $\Re^{N+2}{}_{+}$ (here spanned by z). Whereas global refers to the unbounded positive orthant $\Re^{N+2}{}_{+}$, local refers to one singular point, and regional refers to a connected subset of $\Re^{N+2}$. Note that conceptually the local, global and regional approaches only differ in the way $\boldsymbol{\psi}$ is defined. If $M_{p,y}$ and $C_{p,y}$. holds $\forall\ z \in \boldsymbol{\psi}$, we say that regularity is imposed (i) *globally* if $\boldsymbol{\psi} = \Re_{+}^{2+N}$, (ii) *locally* if $\boldsymbol{\psi}$ consists of one singular point, and (iii) *regionally* if $\boldsymbol{\psi}$ is some connected subset of $\Re_{+}^{2+N}$.

Global and local approaches are by far the most common standard econometric methods currently employed, because of their relative ease of computation. Instead the regional approach is more demanding both computationally and because the researcher needs to explicitly specify the connected set $\boldsymbol{\psi}$. This regional connected subset, first proposed by Gallant and Golub (1984) represents the '*empirically relevant region'*, and is defined by the model analyst. In particular it is assumed that it is known prior to the estimation at which ranges of the data the model shall generate forecasts.[9] Note, for the regional approach it is not particularly

[7] Note that these set of models (i.e. Computational General Equilibrium (CGE) models or partial equilibrium models) do require well-behaved *parametric* functions. Hence, while shape imposing techniques have recently garnered much attention in the nonparametric literature (Henderson and Parmeter, 2009), shape conditions in parametric functions will continue to be of interest in such settings.

[8] Repeatedly it had been argued that a 'regular model' may forecast better out of sample - although their 'in sample' fit statistics are inferior compared to an irregular model. This is an interesting point but not a general result. For a discussion on this, see Edwards and Terrell (2004).

[9] If the estimated function will be used in the context of a larger simulation model, numerical requirements may lead to a larger set $\boldsymbol{\psi}$, sometimes extending beyond the initial range of interest in economic analysis. For example, to guarantee the proper functioning of iterative algorithms used in partial equilibrium models, the set $\boldsymbol{\psi}$ has to

important if the function is irregular immediately outside the boundary of $\boldsymbol{\psi}$ because inferences will not be drawn from those regions. The *regional approach* can offer important advantages over the *local* approach because it imposes theoretical consistency not only locally, at a given singular evaluation point, but also over the entire empirically relevant region of the domain associated with the function being estimated. The method also provides benefits relative to the *global* approach, through higher flexibility derived from being less constraining, generally leading to a better model fit to the sample data. In the empirical Section 3 we test for these claims, comparing local, regional and global approaches.

2.3 Shape imposing estimators

2.3.1. The Inequality Constraint Function

Any estimator can be described as maximizing some statistical criterion function subject to a set of constraints. This subsection first defines the constraining function for shape conditions. Note that the two curvature conditions C_y and C_p (in short $C_{y,p}$) and the two monotonicity conditions $M_{y,p}$ must hold on a connected subset $\boldsymbol{\psi} \subset \Re_+^{2+N}$ of the price × output × time space. The restrictions $C_{y,p}$ and $M_{y,p}$ can hence be characterized by $H = 4$ vector-valued-functions $i_h(\mathbf{z};\boldsymbol{\beta})$, $h = 1,\ldots,H$, whereby the restrictions hold whenever, for a given $\boldsymbol{\beta}$, $\mathbf{i}()$ is nonnegative for all z in the relevant region $\boldsymbol{\psi}$,

$$\mathbf{i}(\mathbf{z};\boldsymbol{\beta}) \equiv [i_1,i_2,\ldots,i_H] \geq \mathbf{0} \quad \forall \ p,y,t \in \boldsymbol{\psi}.$$

Hence for monotonicity and curvature, define the following four sets of inequality constraints:

$$i_1 = \nabla_p c \geq 0, \qquad i_2 = \nabla_y c \geq 0, \qquad i_3 = -\,\mathrm{eig}[\nabla_{pp} c] \geq 0, \qquad \text{and} \qquad i_4 = \mathrm{eig}[\nabla_{yy} c] \geq 0,$$

whereby ∇_j indicates the derivative with respect to variable j and eig[] computes the eigenvalue of the Hessian $\nabla_{jj}c$, which checks whether the curvature of c is concave.

2.3.2. Local Regularity Preserving Estimator

An estimator of local regularity maximizes some statistical criterion function (i.e. minimizing the sum of squared error distances) subject to the constraint that $\mathbf{i}(\mathbf{z};\boldsymbol{\beta}) \equiv [i_1,i_2,\ldots,i_H] \geq \mathbf{0}$ at *one* point $\mathbf{z} = \mathbf{z}_o$ of the regressor

encompass the range of values occurring in the numerical procedure.

space. Using this approach, Ryan and Wales (1998) yield functional forms with a high degree of flexibility. However, this approach does not eliminate the risk that regularity is violated near the selected point. Finding the optimal set of local constraints for a given dataset can be time consuming and lead to complications for statistical testing and verification.

2.3.3. Global Regularity Preserving Estimator

Imposing global regularity requires that shape conditions hold at all values of the regressor space in $\Re^{N+2}$. This is typically done by now well-known parametric restrictions described e.g. in Diewert and Wales (1987). The guarantee that the whole domain is consistent with economic theory comes at the cost of sacrificing flexibility and model fit. For example, imposing global regularity on the Generalized Leontief cost function eliminates the possibility of complementary inputs. Further, monotonicity restrictions have traditionally been omitted because (as proven by Lau, 1986) estimators imposing both monotonicity and concavity *globally* restrict the valid parameter space of the functional form such that it loses the property of second order flexibility, and therefore offer no advantage to functions with only first order flexibility such as CES and Cobb-Douglas. Barnett (2002) and Barnett and Pasupathy (2003) show how omission of global monotonicity can lead to very misleading results in empirical applications.

2.3.4. Regional Regularity Preserving Estimator

The regional regularity approach was first proposed by Gallant and Golub (1984). Here a statistical criterion function is maximized subject to an infinite number of constraints $\mathbf{i}(\mathbf{z};\boldsymbol{\beta}) \equiv [i_1, i_2, \ldots, i_H] \geq \mathbf{0} \;\; \forall \;\; p,y,t \in \boldsymbol{\psi}$ which have to hold at any point within a convex hypercube $\boldsymbol{\psi}$. Imposing regional regularity generally leads to a better statistical fit of the data to the model, compared to the global regularity approach. However, Gallant and Golub (1984) did not demonstrate the tractability of their double inequality constrained optimization technique and it seems that empirical implementation can be formidable with optimization tools currently available. It was more than a decade until Terrell (1996) advanced ideas relating to the empirical application of regional regularity. Instead of explicitly using a constrained optimization algorithm as in Gallant and Golub (1984), Terrell (1996) decomposed the problem into a series of steps: First, the convex set $\boldsymbol{\psi}$ is approximated by a dense grid consisting of thousands of singular regressor values. Secondly, using a Bayesian framework, an unconstrained posterior distribution of the parameter vector $\boldsymbol{\beta}$, conditional on the endogenous variable $\mathbf{y}$, $p_u(\boldsymbol{\beta}|\mathbf{y})$, is derived that does not incorporate the regularity conditions. Thirdly, a Gibbs sampler is used to draw parameter vector outcomes from $p_u(\boldsymbol{\beta}|\mathbf{y})$, and an Accept-Reject algorithm is applied to assess regularity for each outcome at all grid points within $\boldsymbol{\psi}$. Finally, point estimates are derived and inferences are drawn based on the set of regular parameter vectors and its truncated posterior distribution. Further details of the method are provided in part I of the Appendix. This procedure has two challenges: (a) Due to the approximation of the relevant regressor space by the grid, the possibility cannot be eliminated that the function is irregular for some non-grid points. In this sense this technique does not compel regularity on a connected set but imposes local regularity at multiple singular points. (b) The Gibbs simulator requires sampling from the entire support $\boldsymbol{\Theta}$ of the unconstrained posterior $p_u(\boldsymbol{\beta}|\mathbf{y})$. This can be time consuming if, as is often the case in practice, the regular region for $\boldsymbol{\beta}$ is only a small subset of $\boldsymbol{\Theta}$ (Terrell 1996). Aiming to address theses challenges (a) and (b), Wolff *et al*. (2010) developed an alternative estimator that builds upon Terrell (1996) to impose regional regularity, see Box 1. The main differences are as follows: Their Metropolis-Hastings Accept-Reject Algorithm (MHARA) simulator decreases the computing time by reducing the amount of regularity checks because it does not need to sample over the whole unconstrained posterior (which includes potentially large regions of irregular parameter draws), but it samples directly from the regular support $\boldsymbol{\Theta}^R|\boldsymbol{\psi} = \{\boldsymbol{\beta}: \mathbf{i}(\mathbf{p};\boldsymbol{\beta}) \geq \mathbf{0}\ \forall\ \mathbf{z} \in \boldsymbol{\psi}\}$. Second, a mode estimator is proposed that is shown to increase the model fit (compared to the previous posterior mean point estimates). Third, Wolff *et al*. (2010) advocate to impose regularity on non-convex sets of the domain, $\boldsymbol{\psi}^{string}$, referred to as the string approach.[10] Wolff et al. (2010) show in simulations that this greatly enhances the model fit, particularly in cases where a higher dimensional point-cloud of regressor data appears as non-convex. In Wolff *et al*. (2010), however, no empirical application has been offered using real world

[10] The construction of the approximation grids of the hypercube and the string approach are detailed in Table 2 below and in Appendix II.

data to verify these improvements. This paper is the first to do so in Section 3.

Box 1: Step by step procedure to impose regional regularity–adopted from Wolff et al. (2010)

Step 1	Estimate $\mathbf{y} = \mathbf{f}(\mathbf{P};\beta) + \varepsilon$ *without* imposing inequality constraints to obtain the unconstrained estimate $\mathbf{b}_u$ of β as well as the estimated $L \times L$ covariance matrix $\mathbf{cov}(\mathbf{b}_u)$.
Step 2	Define $\mathbf{i}()$ that characterizes the regularity conditions for the function being estimated.
Step 3	Define ψ. If the proposed region is not convex (string approach), define a sequence of I convex subsets ψ_i such that $\psi = \bigcup_{i=1}^{I} \psi_i$.
Step 4	Initialize the Markov Chain with a regular parameter vector: If $\mathbf{b}_u \in \Theta^R$, set $\mathbf{b}^{(0)} = \mathbf{b}_u$ else $\mathbf{b}^{(0)} = \mathbf{0}$. Set $j = 0$.
Step 5	Generate a candidate $\mathbf{b}^{(*)}$ by the proposal distribution $p(\mathbf{b}^{(*)};\mathbf{b}^{(j)})$.
Step 6	If $\mathbf{b}^{(*)}$ is irregular at the vertices of ψ, set r=0.
Step 7	If $\mathbf{b}^{(*)}$ is regular in ψ_g, calculate $r = p(\mathbf{b}^{(*)}\|\mathbf{y},\psi)/p(\mathbf{b}^{(j)}\|\mathbf{y},\psi)$, else set r=0.
Step 8	if $r > 1$, $\mathbf{b}^{(j+1)} = \mathbf{b}^{(*)}$ else if Uniform(0,1) $\leq r$, $\mathbf{b}^{(j+1)} = \mathbf{b}^{(*)}$, else $\mathbf{b}^{(j+1)} = \mathbf{b}^{(j)}$.
Step 9	Increment j by $j = j+1$. Go to step 5, until $j = J+S$, whereby $\{\mathbf{b}^{(j)}\}_{j=1}^{S}$ are the burn-in draws to be discarded after the final loop such that $\{\mathbf{b}^{(j)}\}_{j=S+1}^{J+S}$ are the outcomes to be considered for constructing $p(\beta\|\mathbf{y},\psi)$.
Step 10	Analyze $p(\beta\|\mathbf{y},\psi)$, i.e. calculate point estimates and perform inferences.

3. EMPIRICAL ILLUSTRATION

Background to the Application of Estimating a Flexible Input Demand System

This section first briefly motivates the empirical importance of *estimating* cross-partials (such as the elasticity of substitution between energy and capital) using *second order* flexible functional forms.[11] Climate change concerns drive many countries to debate over imposing a tax on energy use to reduce CO_2 emissions. To assess the costs and benefits of such an energy tax policy, we estimate the own price elasticity of demand for energy, but also require information on cross-price elasticities that describe the effects on important markets that are potentially linked to energy. Firms facing a carbon tax could substitute away from energy

[11] Again, note, first order flexible forms do not allow to numerically estimate the cross partial, but would a-priori determine the value of the cross-partial. For example using the Cobb-Douglas a-prior determines that the elasticity of substitution is equal to one. Also, global imposition of all shape conditions leads to the same problem (see Lau,1986).

towards other inputs such as capital and labor—which may be less polluting but more costly. To this end, estimated demand systems have provided key ingredients to answer many important questions in production analysis (Chambers 1988, Griffiths, O'Donell and Tan Cruz 2000, Kumbhakar and Tsionas 2005), policy studies and welfare analysis (Evans and Heckman 1984, 1986, Koebel, Falk and Laisney 2003), as well as in the debate on the sources of economic growth (Mankiw, Romer and Weil 1992, Hsieh 2000, Antras 2004).

The empirical illustration contains 3 subsections: Section 3.1, re-estimates the demand system for four production inputs to the U.S. manufacturing sector using the Berndt and Wood (1975) data set for capital (K), labor (L), energy (E), and materials (M). This KLEM dataset has been reported to violate regularity conditions implied by economic theory. For that same reason, however, the KLEM data have been applied to a considerable number of studies imposing regularity techniques (i.e. Berndt and Wood 1975, Berndt and Khaled 1979, Galant and Golub 1984, Diewert and Wales 1987, Berndt, Geweke and Wolfe 1991, Friesen 1992, Terrell 1996), providing a substantial foundation on which to investigate and to compare the performance of alternative estimators.

Does estimator choice (described in Section 2) has any policy importance? For instance, does the local approach produce estimates that are significantly different from the global or regional approaches? To exemplify, Section 3.2 looks at the elasticity of substitution between capital and energy, a parameter that has attracted a great deal of attention in the last decades (see e.g. Apostolakis 1990). Finally, section 3.3, asks whether the KLEM data are consistent with 'duality theory' and we outline a procedure for testing this.

The Berndt and Wood data have been described in more detail in many places in the literature (e.g. Berndt and Wood (1975), Berndt and Khaled (1979), Gallant and Golub (1984)). Table 1 provides summary statistics of the annual data from 1947 to 1971. In particular the min/max values of the right hand side price variables will be considered below for the construction of the various ψ sets.

Table 1: Summary statistics of the Berndt and Wood data set from 1947 to 1971

	input quantities				input prices				output
	K	L	E	M	p_K	p_L	p_E	p_M	y
mean	20.45	106.07	16.78	237.56	1.18	1.77	1.35	1.30	313.80
std	7.77	43.59	5.54	85.14	0.19	0.46	0.12	0.14	87.67
min	8.58	45.10	7.76	112.35	0.74	1.00	1.00	1.00	182.83
max	34.11	190.26	29.48	407.71	1.50	2.76	1.65	1.55	466.83

Note: Variables are produced using index numbers and deflators. For details on the data construction see Berndt and Wood (1975) and Berndt and Khaled (1979).

3.1 Comparing Shape Imposing Techniques – An Illustration using the Berndt and Wood Data

The main purpose of the following eight sets of estimations is to assess potential (dis)advantages of the regional estimation versus the standard econometric techniques (the local and the global approaches) both in terms of model fit and the propensity for regularity violations. Performance statistics of various estimators as applied to the second order flexible Generalized Leontief cost function

$$c(\mathbf{z};\boldsymbol{\beta}) = \Sigma_{i=1}^{N}\Sigma_{j=1}^{N} b_{ij} p_i^{0.5} p_j^{0.5} + \Sigma_{i=1}^{N} b_i p_i + \Sigma_{i=1}^{N} b_{it} p_i t y + t\,\Sigma_{i=1}^{N} a_i p_i + y^2\,\Sigma_{i=1}^{N} \beta_i p_i + y t^2\,\Sigma_{i=1}^{N} \gamma_i p_i$$

with $b_{ij} = b_{ji}$, are displayed in Table 3. In order to be able to directly compare our results with previous studies, we use the exact same specification of the demand system as in Diewert and Wales (1987), and in Terrell, (1996).[12] Hence the N estimated equations are

$$\mathbf{x}/\mathbf{y} = \nabla_p c(\mathbf{z};\boldsymbol{\beta})/\mathbf{y} + \mathbf{u}. \tag{1}$$

It is assumed that the $T \times 1$ error vectors $\mathbf{u}_n$, $n = 1,\ldots,4$ are contemporaneously correlated, such that the estimating equations can be written in form of the Gaussian seemingly unrelated regression (SUR) system with $t = 1,2,\ldots T$. For details on the specification see Diewert and Wales, 1987 and Terrell, 1996.

Economic theory restricts the Generalized Leontief cost function to be M_p, M_y, C_p and C_y. In general, economists are well aware of these, imposing *all* four of these conditions when estimating *first* order flexible

[12] For a motivation of this particular specification see Diewert and Wales (1987) and Terrell (1996). In particular, all right hand side variables are assumed to be exogenous. This seems to be a standard approach in this literature: Except for the Berndt et al. articles in the seventies, none of the above papers estimating the KLEM input demand system uses instruments. For a justification of this, see the discussions by Diewert (2004), Barnett and Binner (2004), and Antras (2004). A possible extension is to estimate the system in the context of an error correction model (Friesen, 1992).

functional forms (such as a Cobb-Douglas). In contrast, the standard practice is that only a small subset of these conditions is enforced when using the Generalized Leontief or any other *second* (or higher) order flexible form. In particular, the three conditions M_p, M_y and C_y conditions have rarely been considered. A remarkable exception of explicitly imposing both C_p and M_p (but not M_y and C_y) is Terrell (1996), whose results are shown below for comparison.

The fact that the previous literature contemplated subsets of the regularity conditions cannot be justified from the perspective of economic theory. Why should a violation of monotonicity be less harmful than a concavity violation? The omission of regularity may be explained by the previous lack of standard estimators that have the ability to maintain regularity over all conditions. Such a gap between economic theory and the empirical model is problematic for the interpretability of the results (and is especially worrisome if one wishes to derive any policy conclusions for the U.S. manufacturing sector, which accounts for about 20% of the GDP). To our knowledge, this paper is the first study systematically taking into account *overall* regularity including M_y and C_y when using flexible functional forms.

Definitions for $\boldsymbol{\psi}$ are first provided in Table 2, separating in column (1) the approaches into 'standard econometric techniques' (for 'unconstrained, 'local' and 'global' and the 'regional' approach) versus the 'regional approach'. Column wise Table 3 displays the regression results, with the columns again indicating the various estimation approaches (unconstrained to global). In addition in the columns of Table 3, the symbols in parentheses indicate which shape conditions are imposed. To provide an example, column (4) and (5) of Table 3 display the sympbol $\boldsymbol{\psi}^{\square p}_{i}(M_p,C_p)$, which implies that concavity and monotonicity is imposed with respect to *p* on the convex hypercube. In Table 2, this construction corresponds to row (4) (for $i = 1$) and row (5) (for $i = 2$). Note that this particular set has a Lebesgue measure of zero because it does not expand into the dimension of *y* and *t*. Instead, if all four regularity conditions are imposed, a simplified notation uses the indicator 'all', which corresponds to row (6) of Table 2. In Table 3, for each estimation approach we display in the columns the results of both point estimates, the mean of the regularity posterior distribution, and the mode. As in Terrell, after each estimation, the regularity conditions are evaluated at the smaller set $\boldsymbol{\psi}^{\square}_{1}$ and larger set

$\boldsymbol{\psi}^{\square}{}_2$ as indicated in the respective rows of Table 3. In particular, the smaller hypercube $\boldsymbol{\Psi}^{\square p}{}_1 = \{\mathbf{p}: \mathbf{p} \in \times_{k=1}^{3}[1.0, 1.5]\}$ is defined such as it covers most but not all of the observations in the dataset, while the larger set $\boldsymbol{\Psi}^{\square p}{}_2 = \{\mathbf{p}: \mathbf{p} \in \times_{k=1}^{3}[0.5, 3.0]\}$ covers all data entirely plus areas of the regressor space that are intended for subsequent policy forecasts (i.e. to simulate the introduction of a tax on prices in the range from 1.5 to 3.0, or to simulate a subsidy on the range of prices from 0.5 to 1.0). Violations of regularity conditions are expressed as the percentage of grid points where violations occur, whereby the grids are constructed as described in Table 2, and for details please refer to Appendix II.

Table 2: Regularity imposing sets

Approach		Row	Definition of ψ	Comment
Standard Econometric Techniques	Unconstrained	(1)	$\boldsymbol{\psi} = \emptyset$:	$M_{p,y}$ and $C_{p,y}$ is not imposed. 'Unconstrained' refers to the 'inequality constraints' only: Symmetry and HD1 is imposed by parametric equality restrictions.
	Local	(2)	$\boldsymbol{\psi} = z_1$	z_1 is the first observation in the sample corresponding to the year 1947. z_1 is used because this matches the tables in Berndt and Wood (1979), Diewert and Wales (1984), Barnett et al., (1991) and Terrell (1996).[13]
	Global	(3)	$\boldsymbol{\psi} = \Re_+^{2+N}$	Nonnegative orthant of all right hand side variables of *z.*
Regional	Cube[14] approach	(4)	$\boldsymbol{\Psi}^{\square p}{}_1 = \{\mathbf{p}: \mathbf{p} \in \times_{k=1}^{3}[1.0, 1.5]\}$	This was chosen by Terrell (1996). It does not cover the entire empirical relevant data space. Some observed prices lie outside the [1.0, 1.5] interval, see min/max values in table 1.*
		(5)	$\boldsymbol{\Psi}^{\square p}{}_2 = \{\mathbf{p}: \mathbf{p} \in \times_{k=1}^{3}[0.5, 3.0]\}$	This set was chosen by Terrell (1996). It ensures that $\boldsymbol{\psi}^{\square p}{}_2$ covers all observed prices and beyond.*
		(6)	$\boldsymbol{\Psi}^{\square z}{}_i$ ('all')= $\{\mathbf{z} \in \boldsymbol{\psi}^{\square p}{}_i \times [y_{min}, y_{max}] \times [t_{min}, t_{max}]\}$	$\boldsymbol{\psi}^{\square z}{}_i$, $i \in \{1,2\}$: These sets expand $\boldsymbol{\psi}^{\square p}{}_i$ over the remaining dimensions *t* and *y*.
	String approach	(7)	$\boldsymbol{\psi}^{\text{string}} = \bigcup_{t=1}^{26} \boldsymbol{\psi}_i$	$\boldsymbol{\psi}^{\text{string}}$ covers 26 = T+1 points in $\Re_+^{2+N}$ by connecting straight lines $\boldsymbol{\psi}_i$ between the right hand side variables mean, $\mathbf{z}^M$, and each of the T observations. I approximate each line $\boldsymbol{\psi}_i$ by $\boldsymbol{\psi}_{ig}$ by taking 10 equidistant grid points between $\mathbf{p}^M$ and the i^{th} observation z_i, leading to 1+(F-1)T = 226 grid points.

* This set has Lebesgue measure of zero because the set does not expand into the dimension *y* and *t.*

[13] In case of the Berndt and Wood dataset, imposing the regularity conditions at other points in the sample space (z_2 to z_{25}) does not essentially change the empirical results. This conclusions, however, is not a general result but is entirely data-driven.

[14] In this paper all grid sets $\boldsymbol{\psi}^{\square ''}{}_{ig}$ are constructed with F = 10, as in Terrell, 1996. For details see Appendix II.

Table 3: Generalized Leontief Input Demand System, estimated with 8 different approaches

Model Performance Statistics: Regularity Violations/ Fit Statistics		**Estimation Approach**															
		Unrestricted	**Global**			**Local**		**Regional**									
		1	**2a**		**2b**	**3**		**4**		**5**		**6**		**7**		**8**	
Restricted domain with shape conditions imposed in parenthesis		Ø (‘none’)	$\Re_+^{2+N}$ (C_p)		$\Re_+^{2+N}$ (‘all’)	z_1(‘all’)		$\Psi^{\square p}_1(M_p,C_p)$		$\Psi^{\square p}_2(M_p,C_p)$		$\Psi^{\square z}_1$(‘all’)		$\Psi^{\square z}_2$(‘all’)		Ψ^{string}(‘all’)	
evaluated at	**% of shape violations at grid points**	**mode**	**mean**	**mode**	**mode(b^{gr})**	**mean**	**mode**	**mean**	**mode**	**mean**	**mode**	**mean**	**mode**	**mean**	**mode**	**mean**	**mode**
$\Psi^{\square}_1$	C_p violations	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	M_p violations	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	C_y violations	4.0	25.1	33.4	0.0	0.0	0.0	0.0	0.0	0.0	53.0	0.0	0.0	0.0	0.0	0.0	0.0
	M_y violations	0.0	2.0	11.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\Psi^{\square}_2$	C_p violations	100.0	0.0	0.0	0.0	27.1	27.1	16.1	27.0	0.0	0.0	15.2	27.0	0.0	0.0	27.4	27.4
	M_p violations	3.1	45.8	0.0	0.0	5.0	5.0	5.2	5.3	0.0	0.0	3.6	5.3	0.0	0.0	4.9	5.8
	C_y violations	31.2	33.4	25.1	0.0	32.9	32.9	34.4	33.3	29.7	50.5	1.6	33.3	0.0	0.0	0.0	32.8
	M_y violations	1.3	11.3	2.0	0.0	1.8	1.8	1.7	1.9	1.8	0.9	2.4	1.9	0.0	0.0	4.8	2.4
Generalized Variance of the Fit[1]		1.44	0.52	0.52	0.27	1.26	1.26	1.18	1.26	0.74	0.76	1.14	1.26	0.17	0.27	1.20	1.26

Note: The model fit is calculated by *Generalized Variance of the Fit*. See e.g Barnettt (1976). Here it is defined as $100\bullet|\Sigma|^{-1}$. The statistic is proportional to the ordinate value of the unconstrained posterior $p(\mathbf{b}|y,\varnothing)$ evaluated at the respective point estimate. Due to the choice of priors here it is proportional to the likelihood value of the unconstrained Maximum Normal Likelihood regression.

3.1.1 Unconstrained estimation

In the first column of Table 3, the demand system is estimated by iterated SUR unrestrictedly. Firstly, compared to any other columns, the unrestricted estimate, $\mathbf{b}^{u}$, provides the best model fit statistics but $c(z;\mathbf{b}^{u})$ violates the regularity conditions *everywhere*, both in $\boldsymbol{\psi}^{\square}_1$ land $\boldsymbol{\psi}^{\square}_2$ leading to, among other things, a failure of the fundamental law of demand.[15] Contemplating these poor regression results, a researcher could now pursue a multitude of directions, until something more consistent is obtained, i.e. trying other functional forms or applying the data to another economic theory. However, if goalposts are changed in an ad hoc manner, such procedures can be rife with statistical testing and verification problems. Hence other estimation approaches leading to a well-behaved economic model are required to test this hypothesis.

3.1.2. Global Approach

The global approach to impose concavity is probably one of the most common techniques, when estimating flexible input demand systems. In the case of the Generalized Leontief, this unfortunately allows the cost function to model substitutes only (Diewert and Wales 1987). Because in the KLEM data at least energy and capital seem to have a stark complementary relationship, maintaining global concavity reduces the model fit. Moreover, the global imposition of concavity is not sufficient for *overall* regularity. In fact, all remaining conditions are violated as can be seen in column 2a. Now, applying the global approach not only to C_p but to all regularity conditions (estimated by (1) with $b_{ij}^{u}=0$, $i\neq j$) a-priori fixes the elasticity of substitution estimates to 0. Here, as displayed in column 2b, such a procedure performs very poorly in terms of model fit because this globally regular estimate $\mathbf{b}^{gr}$ emerges from the 'small' parameter subset $\boldsymbol{\Theta}^{R}|\Re_{+}^{2+N} \subset \boldsymbol{\Theta}^{R}|\boldsymbol{\psi}$. In conclusion our empirical results indicate that the global approaches are extremely restrictive, a finding that agrees with the results in Diewert and Wales, 1987, Terrell, 1996 and the simulation study in Wolff et al. 2010.

[15] For comparison, this estimation exactly repeats the unrestricted estimation of Diewert and Wales (1987: table II) and Terrell's (table 3: 1996).

3.1.3 Local approach

The third column displays results from local imposition of overall regularity. As expected, here the model fit is inferior when compared to the unrestricted approach, but is superior to the global approach. Unfortunately, the local approach does not guarantee that regularity is satisfied in the relevant empirical area, leading to a high percentage of over 33% of violations of C_y at $\boldsymbol{\psi}^{\square}_{2}$. One has to be careful, however, with the interpretation of these percentage statistics: On one side, violating shape conditions at 100% of the data points does not necessarily imply that the estimated model is "very far" from a good model. And on the reverse, a single violation of one grid point could lead to extremely poor results. The latter point is illustrated in Figure 2.

Before proceeding to the regional approach, it is worth noting that the local approach does not lead to any violation in $\boldsymbol{\psi}^{\square}_{1}$. Hence, if one is solely interested in $\boldsymbol{\psi}^{\square}_{1}$, one could stop here, because the below regional approach of $\boldsymbol{\psi}^{\square}_{1}$ in the next Section does not improve the estimation (compare upper panel of column (6) to column (3)). Note, however, that $\boldsymbol{\psi}^{\square}_{1}$ does not cover all the relevant data points (compare the min/max values in Table 1 and the definition of $\boldsymbol{\psi}^{\square}_{1}$ in Table 2).

Since neither the unconstrained, local nor the global approach produces a well-behaved flexible demand system at all data points, we now turn to the regional approach.

Fig 2: Illustration of an irregular cost function violating the shape restriction at one grid point

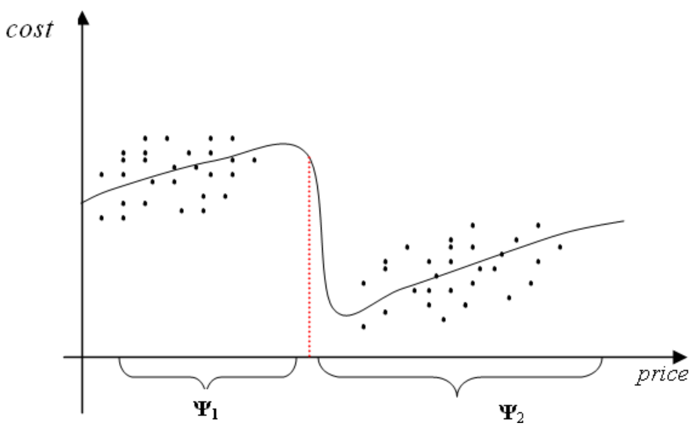


Illustration of a cost function regular within ψ_1 and within ψ_2. The overall model is irregular, however, on the domain $\psi = \psi_1 \cup \psi_2$ although only one grid point is violated (here indicated on the domain by the dotted line). Overall regularity is violated because cost must not decrease with rising input prices.

3.1.4 Regional Regularity

First Terrell's (1996) estimation is replicated, i.e. using the Gibbs accept-reject simulator and the same $\boldsymbol{\psi}$-sets as defined by Terrell. He applied this method to impose M_p and C_p. This successfully leads to regularity preserving results if interested in the function's domain $\boldsymbol{\psi}^{\square z}{}_1$. In contrast, on the domain $\boldsymbol{\psi}^{\square z}{}_2$, the constraints with respect to y would be violated (see column 4 and 5 in Table 3). This shows that there is a need to impose these other conditions as well.

We now turn towards the MHARA estimation method imposing all regularity conditions on any set $\boldsymbol{\psi}$ of interest. Hence, as shown in the columns (6) and (7) regularity holds in $\boldsymbol{\psi}^{\square z}{}_1$ and $\boldsymbol{\psi}^{\square z}{}_{2.}$ We display the results for both, the mean and the mode of the posterior. In case of column (6), the mode increases the model fit (as measured by the likelihood value) by over 10% (from 1.14×10^{-2} to 1.26×10^{-2}). The percentage increase is even more dramatic in the case of imposing the regularity conditions on the larger set $\boldsymbol{\psi}^{\square z}{}_2$, in column (7), achieving an increase in model fit by 43.6% from 0.17 to 0.27. These results are consistent with Lemma 1 of Wolff et al. (2010).

The String Approach

So far, we only have described the approaches based on a convex cube $\boldsymbol{\psi}^{\square}$. A motivation for the KLEM data set to further investigate in non-convex sets for $\boldsymbol{\psi}$ is probably best described by Gallant and Golub, 1984: '*The exogenous variable* [p_t and y_t] *for t* = 1947,…,1971 *lie in a five dimensional space and can be projected into a three dimensional space with a negligible loss of information…The projected point cloud has an irregular shape. It is a sort of a fat rope lying mostly on the ground in the shape of a tilde (~) with the two end-points and the middle elevated.*' This description indicates that constructing $\boldsymbol{\psi}^{\square}$ as a convex cube (which includes all the data points) may lead to an unnecessary voluminous set containing relatively little useful data ranges. A simple construction rule of a non-convex set containing all the KLEM data is described in Table 2 and labelled as the 'string approach'. Comparing the regular parameter support of the string approach versus the cube approach shows that the former is a parameter superset of the latter, $\boldsymbol{\Theta}^R|\boldsymbol{\psi}^{string} \supset \boldsymbol{\Theta}^R|\boldsymbol{\psi}^{\square z}{}_1$, benefiting the resulting *flexibility* property of the functional form and leading to a higher model fit, see

Lemma 1 of Wolff et. al (2010). In the application of column (8) (String approach), the model fit strikingly increases from 0.17 to 1.20, compared to the hypercube approach in column (7). Also, note that comparing the mean of the string approach to the mode leads to an additional improvement of the model fit of about 5.3%.

The string approach has potential important advantages because: (a) it represents the method which is to the largest possible extend 'data driven', (b) it leads to a well-behaved demand model and (c) with 226 regularity checks, it is computationally *much* faster than the regional regularity preserving cube method, that checks one million times.[16]

Summarizing Table 3, unconstrained and local regularity estimates increase the model fit in all specifications at the cost of violating regularity within $\boldsymbol{\psi}$. This produces estimation results that are problematic in terms of economic interpretation and further policy analysis. Imposing regional regularity solves this problem and significantly increases the model fit when compared to the global approach. Moreover, it is relevant for model fit to use the mode instead of the mean. Finally, the string technique proposed by Wolff *et al*. (2010) reduces computing time, making the regional regularity approach more tractable for empirical analysis. Instead of the full evaluation grid consisting of over one million points, only a maximum of 343900 points have to be evaluated for the cube approaches. Furthermore, for the string approach only 226 points have to be assessed. This significantly decreases the computational burden when compared to previous approaches.

Whether, these advances have any consequences for our estimation results will be investigated in the following subsection with respect to the U.S. matrix of the elasticity of substitution of input factors.

3.2 Elasticities

We examine the estimated marginal posterior distributions for input demand elasticities $\partial x_i/\partial p_j \cdot x_i/p_j$. Table 4 reports the means, modes and standard deviations of these density functions. For the purpose of analyzing the potential effects of an environmental tax on energy use, here we are interested in the capital-energy elasticity. The long-run growth potential of the manufacturing sector depends crucially on the magnitude of this

---

[16] With the Q^*-grid construction approach we need about 30% of the original Q-grid computing time. In comparison the string approach estimation is faster requiring less than 0.05% of the original computing time.

parameter (Dasgupta and Heal, 1979). In particular the question of whether capital and energy are complements or substitutes has received much attention (see e.g. Apostolakis 1990).[17]

[17] If they are substitutes, then an increase in energy taxes would lead, ceteris paribus, to an increase in the capital stock, potentially benefiting the sector in the long run. In this case, energy conservation policies promoting new energy-saving physical capital would be predicted to have the desired effect. However, if they are complements, then rising energy prices would adversely effect capital formation and, hence, such policies could be counterproductive.

Table 4: Price elasticities matrices at 1947, evaluated at the mode and the mean estimate

Unrestricted SUR

	K	L	E	M
		mode estimates		
K	-0.0974	0.4461	-0.1315	-0.2172
L	0.0921	-0.1774	0.0922	-0.0069
E	-0.1579	0.5362	-0.6167	0.2384
M	-0.0168	-0.0026	0.0154	0.0040
		mean estimates		
K	-0.0978	0.4384	-0.1320	-0.2086
L	0.0905	-0.1909	0.0926	0.0078
E	-0.1585	0.5384	-0.6136	0.2337
M	-0.0162	0.0029	0.0151	-0.0018
		std		
K	0.0625	0.1576	0.0422	0.1998
L	0.0325	0.2355	0.0402	0.2393
E	0.0507	0.2338	0.1284	0.2067
M	0.0155	0.0898	0.0133	0.0994

Local Approach

	K	L	E	M
		mode estimates		
K	-0.1488	0.2897	-0.1412	0.0004
L	0.0598	-0.4092	0.0888	0.2606
E	-0.1696	0.5163	-0.6659	0.3192
M	0.0000	0.0978	0.0206	-0.1184
		mean estimates		
K	-0.1485	0.2896	-0.1412	0.0002
L	0.0598	-0.4110	0.0888	0.2624
E	-0.1696	0.5160	-0.6655	0.3191
M	0.0000	0.0985	0.0206	-0.1191
		std		
K	0.0492	0.1228	0.0415	0.1422
L	0.0253	0.1751	0.0392	0.1689
E	0.0498	0.2280	0.1203	0.2006
M	0.0110	0.0634	0.0129	0.0665

Global concavity

	K	L	E	M
		mode estimates		
K	-0.1369	0.0685	0.0145	0.0539
L	0.0141	-0.3349	0.0569	0.2638
E	0.0174	0.3307	-0.5425	0.1943
M	0.0042	0.0990	0.0125	-0.1157
		mean estimates		
K	-0.1364	0.0682	0.0145	0.0537
L	0.0141	-0.3334	0.0570	0.2624
E	0.0174	0.3311	-0.5431	0.1945
M	0.0042	0.0984	0.0126	-0.1151
		std		
K	0.0502	0.0437	0.0141	0.0417
L	0.0090	0.1712	0.0289	0.1676
E	0.0169	0.1681	0.1307	0.1208
M	0.0032	0.0629	0.0078	0.0638

Regional approach on cube1

	K	L	E	M
		mode estimates		
K	-0.1488	0.2897	-0.1412	0.0004
L	0.0598	-0.4092	0.0888	0.2606
E	-0.1696	0.5163	-0.6659	0.3192
M	0.0000	0.0978	0.0206	-0.1184
		mean estimates		
K	-0.1672	0.2300	-0.1312	0.0683
L	0.0475	-0.3898	0.0831	0.2592
E	-0.1575	0.4830	-0.6566	0.3310
M	0.0053	0.0973	0.0214	-0.1239
		std		
K	0.0468	0.1046	0.0389	0.1239
L	0.0216	0.1721	0.0339	0.1720
E	0.0467	0.1971	0.1109	0.1864
M	0.0096	0.0645	0.0120	0.0691

Regional approach on cube2

	K	L	E	M
		mode estimates		
K	-0.1605	0.1938	-0.0585	0.0251
L	0.0400	-0.3584	0.0915	0.2268
E	-0.0703	0.5321	-0.6585	0.1966
M	0.0019	0.0851	0.0127	-0.0997
		mean estimates		
K	-0.1613	0.1237	-0.0398	0.0774
L	0.0255	-0.3661	0.0691	0.2715
E	-0.0477	0.4015	-0.5641	0.2103
M	0.0060	0.1019	0.0136	-0.1214
		std		
K	0.0483	0.0729	0.0228	0.0781
L	0.0151	0.1599	0.0330	0.1575
E	0.0274	0.1921	0.1100	0.1610
M	0.0061	0.0591	0.0104	0.0616

String approach

	K	L	E	M
		mode estimates		
K	-0.1488	0.2897	-0.1412	0.0004
L	0.0598	-0.4092	0.0888	0.2606
E	-0.1696	0.5163	-0.6659	0.3192
M	0.0000	0.0978	0.0206	-0.1184
		mean estimates		
K	-0.1474	0.2900	-0.1429	0.0004
L	0.0599	-0.4215	0.0886	0.2730
E	-0.1716	0.5153	-0.6714	0.3277
M	0.0000	0.1024	0.0211	-0.1236
		std		
K	0.0475	0.1246	0.0406	0.1443
L	0.0257	0.1748	0.0397	0.1725
E	0.0487	0.2308	0.1204	0.1992
M	0.0112	0.0647	0.0129	0.0698

Figure 3: Posterior distributions of the elasticity of substitution between energy and capital

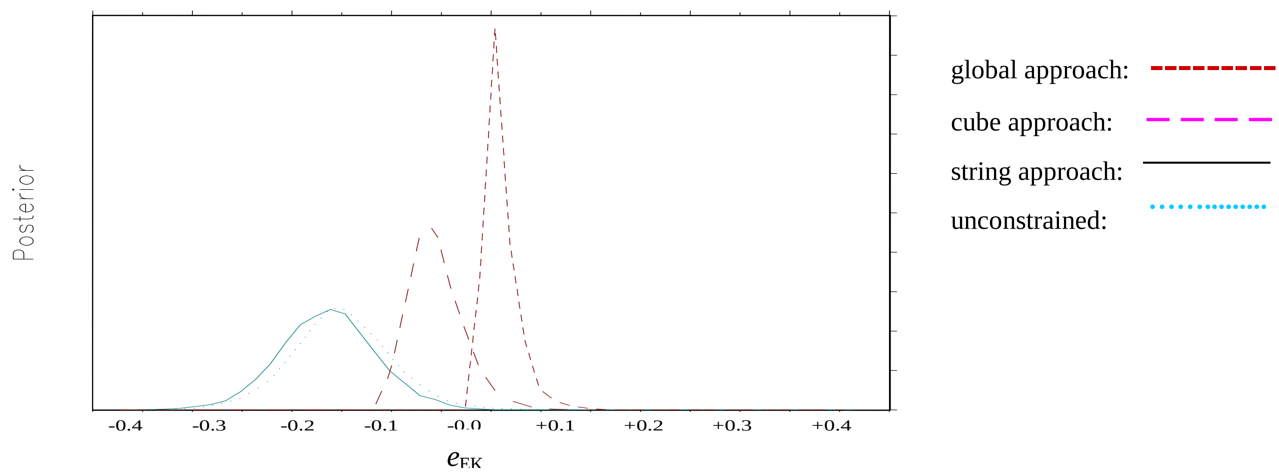


Figure 3 presents results of the cross price elasticity of energy with respect to capital e_{EK}. The global concavity approach produces the far most right distribution having a mode value at e_{EK} = 0.0145 (see Table 4) suggesting that capital and energy are weak substitutes. Instead, the entirely data driven 'unconstrained' approach suggests that capital and energy are complements with e_{EK} = -0.1315. Since the unconstrained estimate however violates duality theory, we turn towards the regional approach. Here, the regularity preserving string approach produces the distribution only slightly to the left of the unconstrained approach with the mode at e_{EK} = -0.1412. Alternatively, the cube approach produces a distribution, which is likely to be biased towards zero, as the estimated function is estimated subject to a much larger set of constraints. [18]

Table 5: Estimated changes in capital stock in millions of U.S. dollars in manufacturing sector due to 10% increase in energy price

		Global concavity approach			String approach		
	year	(lower 5% bound)	point estimate	(upper 95% bound)	(lower 5% bound)	point estimate	(upper 95% bound)
elasticity e_{EK}		(0.00102)	0.01450	(0.04007)	(-0.20928)	-0.14120	(-0.07555)
change in capital stock	1947	(1.0)	13.5	(37.3)	(-194.8)	-131.5	(-70.3)
	2001	(45.6)	646.0	(1785.2)	(-9323.9)	-6290.7	(-3365.7)

NOTE: In the first row calculations are based on the year 1947. This represents our results comparably to the tables and figures provided in previous studies (such as Berndt and Wood (1979), Diewert and Wales (1984), Barnett et al. (1991) and Terrell (1996)). Numbers in parenthesis provide the 90% coverage probability intervals of respective changes and elasticities. Intervals are computed by using 100,000 MHARA simulator outcomes.

[18] With respect to Figure 3, Kolmogorov-Smirnov tests reject the null that the distribution functions of cross price elasticity of energy with respect to capital e_{EK} are equal to each other.

Table 5 illustrates the substantial consequences of using these different estimators. The standard global concavity approach would imply that a 10% increase in energy price leads in the U.S. manufacturing sector to a 14 million U.S. dollars *increase* in capital formation. Instead using the preferred MHARA string approach predicts a significant *decrease* of the capital stock by about 132 million dollars. The second row uses more recent 2001 data provided by the Bureau of Labor Statistics. Due to the use of the different estimators, the same calculations lead to an absolute change in capital stock of about 7 billion U.S. dollars (= 6.3 + 0.6). Given that in the year 2001 the total value of the capital stock in the manufacturing sector amounts to 446 billion dollars, the change of 7 billion solely due to the use of different estimators is alarming.

Two more observations are worth noting. First, comparing the variances of the distributions, we see that a robust pattern arises: The larger the set $\boldsymbol{\psi}$, the smaller is the sample variance of the posterior distribution. Although one might be attracted by an estimator with a small variance, choosing the estimator on this basis would be very misleading! As can be seen from Table 3, the variance of the estimator is rather inversely related to the model fit statistics. Again, see Lemma 1 of Wolff et al. (2010) for a simple proof, that expanding the regularity imposing set, ceteris paribus, decreases the supremum of any statistical criterion functions measuring the model fit.

Secondly, from Figure 3, the pattern suggests that the starker the restrictions, the greater is the difference in the relative positions between the unrestricted approach and the restricted approaches. One therefore could conclude that, the unrestricted approach might be a good 'approximation', since it seems to be close to the string approach. However, such a conclusion is misleading too. For example, the own price elasticity of material of the unconstrained approach is positive (e_{MM} = 0.004, see Table 4), hence implying a perverse slightly positively sloping demand function. Instead, the preferred string approach produces a downward sloping demand function for material with an own price elasticity of -.12.

3.3 Is the KLEM data consistent with duality theory?

3.3.1 Testing Duality Theory using regionally regular estimators

For the KLEM data set, Section 3.1 demonstrated that regularity-imposing estimators are required to make the empirical model consistent with the assumed underlying economic theory. Forcing the data into such theoretical relationships should raise serious concerns, however. In order to investigate this issue, we propose to perform the following simple hypothesis test. We test the unrestricted estimate $\mathbf{b}^u$ (Table 3, column 1) against the hypothesis of 'economic theory[19]'. The null of 'economic theory' hypothesis is represented by the regionally regular estimate $\mathbf{b}^{rr}$ of the string approach (since $\mathbf{b}^{rr}$, (column 8) satisfies all shape restrictions, the model $c(z,\mathbf{b}^{rr})$ is consistent with duality theory). F, Wald and Likelihood Ratio tests (with Bartlett correction) are carried out comparing $\mathbf{b}^{\mathbf{u}}$ with $\mathbf{b}^{rr}$. The duality theory hypothesis is not rejected by any test at the 5% significance levels. Similarly, one can perform this test in the Bayesian framework[20]. Using the uninformative Bayes factor of 1, the posterior odd ratio in favor of the well-behaved model is 0.874 (Zellner, 1971).[21] This leads to the conclusion that the KLEM data in deed could have been generated by an underlying cost function that is consistent with economic theory.

An alternative is to report the percentage of data points where regularity violations occur. This percentage, however, should not be interpreted as a test statistic. For example, despite the fact that the hypothesis tests cannot reject the regular model, our unconstrained estimate violates regularity at 100% of the data points (see column (3) of Table 3).

3.3.2 Testing problems using standard estimators

In the above section we compared the unrestricted estimate $\mathbf{b}^{\mathbf{u}}$ with regionally regular estimate $\mathbf{b}^{rr}$. Assume for a moment that no regional regular estimator were available (as was the case prior to Terrell,

[19] Here 'economic theory' is defined as the cost function maintaining the shape conditions of $M_{p,y}$ and $C_{p,y}$.

[20] The Classical point estimate is exactly identical to the above defined Bayesian point estimate $\boldsymbol{\beta}^{(mode)}$ if, as we have done above, an uninformative prior distribution on $\boldsymbol{\Theta}^R|\boldsymbol{\psi}$ is employed. The Bayesian interpretation has computational advantages because the finite sample confidence intervals and standard errors of functions of $\boldsymbol{\beta}$ can be directly computed with the MCMC draws. Instead deriving the Classical distributions could be tremendously challenging and in general requires more time intensive numerical procedures (like bootstrapping). The computational burden is mainly due to the many inequality constraints $\mathbf{i} \geq \mathbf{0}$ which have to hold for all $z \in \boldsymbol{\psi}$. Also since $\boldsymbol{\beta}^{(mode)}$ could lie on the boundary of $\boldsymbol{\Theta}^R|\boldsymbol{\psi}$ further complications arise (see Geweke 1986, Andrews 1999, 2001).

[21] An alternative is to report the percentage of data points where regularity violations occur. This percentage, however, should not be interpreted as a test statistic. For example, despite the fact that the hypothesis tests cannot reject the regular model, our unconstrained estimate violates regularity at 100% of the data points.

1996). The analyst must hence use the globally regular parameter $\mathbf{b}^{gr}$ (column 2b) to represent the null of duality theory. Testing $\mathbf{b}^{u}$ against $\mathbf{b}^{gr}$ however leads to over-rejection. This is due to the fact that $\mathbf{b}^{gr}$ is a member of the much smaller subset $\mathbf{\Theta}^{R}|\Re_{+}^{2+N} \subset \mathbf{\Theta}^{R}|\boldsymbol{\psi}$. To show this fact for the KLEM data, we repeat the above procedure (from Section 3.3.1) by testing $\mathbf{b}^{u}$ against $\mathbf{b}^{gr}$. In stark contrast to the above findings, here the F, Wald and Likelihood ratio test results would lead to the erroneous conclusion that the KLEM data is not consistent with duality theory. With only 0.189 in favor of the well-behaved model, the Bayesian posterior odd gives the same result.

In conclusion, using modern shape-imposing techniques provides evidence that the KLEM data is consistent with duality theory. Instead, using standard econometric methods creates an unfortunate divide between the empirical model and the underlying economic theory. The regional regularity preserving estimators can close this gap.

4. CONCLUSION

Behavioural assumptions are the central building blocks of many economic theories, such as the assumption of profit or utility maximization. One critical implication of these assumptions is that they manifest themselves in the form of uniquely defined ‘shape conditions’. Well-known examples are curvature and monotonicity restrictions that apply to indirect utility, expenditure, production, profit, and cost functions. Unfortunately, estimated functions frequently violate these regularity conditions. Clearly, such a gap between economic theory and the empirical model is problematic for the interpretability of the results, and especially worrisome if one wishes to derive policy recommendations. In view of both, the need to produce theoretically consistent models and the empirical difficulties in implementation, Diewert and Wales (1987) observe: *One of the most vexing problems applied economists have encountered is that theoretical curvature conditions that are implied by economic theory are frequently not satisfied by the estimated cost, profit or indirect utility function.*

By reviewing and applying a series of shape imposing techniques to the estimation of the U.S. input demand system, using the “Berndt and Wood” data, this paper makes the following contributions: First, existing work either do not impose all the shape restrictions implied by economic theory or partially impose restrictions globally or locally. This paper is the first empirical paper that imposes all the shape restrictions simultaneously. This becomes feasible because of adopting the more flexible regional shape-preserving techniques developed by Gallant and Golub (1984) and Terrell (1996) as well as recently enhanced by Wolff et al. (2010). Second, this paper investigates how careful estimator selection helps eliminate the wedge between the empirical model and economic theory. In particular, this paper motivates a testing procedure that checks the plausibility of the assumed economic theory. The main finding is that the class of regional regularity estimators successfully produce an empirical demand system that is entirely consistent with the underlying economic theory, for a dataset that traditionally rejected economic theory when standard econometric approaches have been used. Third, we illustrate policy implications by assessing distributions of elasticities and find that the choice of estimator has severe implications on whether energy and capital are complements or substitutes. Finally, we show how recent advances to extend the regional approach to non-convex sets substantially improves the model fit and tractability by dramatically reducing computing time.

These estimation and inference procedures can be extended to other areas of economics. For example the method could be applied to the estimation of producer supply or consumer demand systems, which have multiple underlying shape conditions implied by economic theory. Equivalently, many functional relations in game theoretic, industrial organization and auction models exhibit curvature, quasi-convexity or monotonicity restrictions. It would also be interesting to compare these estimation results with the developing new techniques in nonparametric estimation that impose and test for shape restrictions (see Matzkin 1994, Tripathi 2000, Aït-Sahalia and Duarte 2003, and Henderson et al. 2012). This is to be explored in future research. It is hoped that the methods demonstrated in this paper promote tractability and facilitate the analysis of empirical models for which consistency with an underlying economic theory is required.

APPENDIX

Appendix I: Regional Imposition of Shape Restrictions via Numerical Integration

The distribution of interest is the regularity censored likelihood or posterior $p(\boldsymbol{\beta}|\mathbf{y},\boldsymbol{\psi})$, which depends on $\boldsymbol{\psi}$.[22] To generate outcomes from $p(\boldsymbol{\beta}|\mathbf{y},\boldsymbol{\psi})$ either a Gibbs sampler can be used, or based on Griffiths, O'Donnell and Tan-Cruz (2000) the Metropolis-Hastings Accept Reject Algorithm (MHARA) can be used to generate J (pseudo-) random outcomes, $\mathbf{b}^{(j)}$, $j = 1,\ldots,J$ from $p(\boldsymbol{\beta}|\mathbf{y},\boldsymbol{\psi})$ on the regular support $\boldsymbol{\Theta}^{R}|\boldsymbol{\psi} = \{\boldsymbol{\beta}: \mathbf{i}(\mathbf{p};\boldsymbol{\beta}) \geq \mathbf{0}\ \forall\ \mathbf{z} \in \boldsymbol{\psi}\}$. To account for the regularity condition $\mathbb{1}\{\boldsymbol{\beta} \in \boldsymbol{\Theta}^{R}|\boldsymbol{\psi}\}$, the simulator must ensure that any drawn parameter vector $\mathbf{b}^{(j)}$ implies regularity of $\mathbf{c}(\mathbf{z};\mathbf{b}^{(j)})$ for every point $\mathbf{z}$ in the predefined set $\boldsymbol{\psi}$, i.e. $\mathbf{b}^{(j)} \in \boldsymbol{\Theta}^{R}|\boldsymbol{\psi}\ \forall\ j$. Since there are an infinite number of points in $\boldsymbol{\psi}$, they cannot all be checked explicitly. The connectedness is approximated by a fine grid, denoted by the disconnected set $\boldsymbol{\psi}_{g} \subset \boldsymbol{\psi}$. Within the Gibbs or Metropolis Hastings chain an additional Accept-Reject algorithm is implemented to guarantee that $\forall\ \mathbf{b}^{(j)}$ the regularity conditions hold for any single grid point. This implies that $\mathbf{b}^{(j)} \in \boldsymbol{\Theta}^{R}|\boldsymbol{\psi}_{g}\ \forall\ j$, whereby $\boldsymbol{\Theta}^{R}|\boldsymbol{\psi}_{g}$ is the *approximated* regularity likelihood (or—in the Bayesian context—posterior) support, which will tend toward the actual set $\boldsymbol{\Theta}^{R}|\boldsymbol{\psi}$ the finer the approximation grid $\boldsymbol{\psi}_{g}$.

Appendix II: Approximating ψ

Case 1: ψ as hypercube

In the simplest setting $\boldsymbol{\psi}^{\square}$ is a hypercube (the superscript $^{\square}$ refers to the cube approach): Let $z_i(\boldsymbol{\psi}_{min})$ and $z_i(\boldsymbol{\psi}_{max})$ represent the minimum and maximum of the *i*-th right hand side variable. The grid is constructed by selecting $F = 10$ equidistant values for each variable: $z^{f}_{i} = z_i(\boldsymbol{\psi}_{min}) + (f\text{-}1)F^{-1}(z_i(\boldsymbol{\psi}_{max})\text{-}z_i(\boldsymbol{\psi}_{min}))\ \forall\ f \in \{1,2,\ldots, F\}$ and using all possible $Q = 10^6 = F^{\dim(z)}$ combinations to generate the Q-grid $\boldsymbol{\psi}^{\square}_{g} \subset \boldsymbol{\psi}^{\square}$. In order to circumvent the approximate nature of this representation, Wolff *et al*. 2010 identified conditions under which

[22] Let $\boldsymbol{\Theta}$ be the K-dimensional parameter space. If all regularity conditions hold for all values of z in $\boldsymbol{\psi}$, the regular parameter set is defined as $\boldsymbol{\Theta}^{R}|\boldsymbol{\psi} = \{\boldsymbol{\beta}\in\boldsymbol{\Theta}: \mathbf{i}(\mathbf{p};\boldsymbol{\beta}) \geq \mathbf{0}\ \forall\ z \in \boldsymbol{\psi}\}$, hence $\boldsymbol{\Theta}^{R}|\boldsymbol{\psi}$ is dependent on the choice of $\boldsymbol{\psi}$. The marginal prior on $\boldsymbol{\beta}$ is specified as an indicator function $p(\boldsymbol{\beta}|\boldsymbol{\psi}) = \mathbf{1}\{\boldsymbol{\beta} \in \boldsymbol{\Theta}^{R}|\boldsymbol{\psi}\}$ where the prior equals 1 if regularity holds at the value $\boldsymbol{\beta}\ \forall\ \mathbf{z} \in \boldsymbol{\psi}$, and equals 0 otherwise. Throughout the paper we assume the standard ignorance prior for the $N \times N$ covariance matrix $|\boldsymbol{\Sigma}|^{-(N+1)/2}$. The posterior distribution for $\boldsymbol{\beta}$ is then derived by applying Bayes rule, $p(\boldsymbol{\beta}|\mathbf{y},\boldsymbol{\psi}) \propto \int L(\boldsymbol{\beta},\boldsymbol{\Sigma}|\mathbf{y})\cdot \mathbf{1}\{\boldsymbol{\beta} \in \boldsymbol{\Theta}^{R}|\boldsymbol{\psi}\}\cdot |\boldsymbol{\Sigma}|^{-(N+1)/2}d\boldsymbol{\Sigma}$, where $L(\boldsymbol{\beta},\boldsymbol{\Sigma}|\mathbf{y})$ is the normal likelihood function.

checking a certain key point in $\boldsymbol{\psi}^{\square}$ guarantees regularity in well defined neighborhood. Following this procedure leads to a reduction in regularity checks to a total of $Q^{*} = 343900 < Q = 10^{6}$.

Case 2: $\boldsymbol{\psi}$ as a string.

The reduction of regularity checks can be further enhanced by constructing nonconvex sets: with T = 25 observations and S=10 out of sample forecasts S=10, the number of regularity checks is reduced to occur, at 316 = 1+(F-1)(T+S) grid points only in the string approach.

Point estimates and the relation to Maximum Simulated Normal Likelihood

The described techniques can be applied to the Bayesian and to the Classical frameworks. In the Classical framework one would maximize a likelihood function subject to the inequality constraints and the numerical point estimate of the maximum simulated likelihood is the mode. This Classical mode is exactly identical to the above defined Bayesian point estimate $\boldsymbol{\beta}^{(mode)}$ if, as we have done above, an uninformative prior distribution on $\boldsymbol{\Theta}^{R}|\boldsymbol{\psi}$ is employed.

References

Aït-Sahalia, Y. and J. Duarte (2003), "Nonparametric option pricing under shape restrictions," *Journal of Econometrics*, 116, pp. 9-47.

Andrews, D. (2001), "Testing When a Parameter Is on the Boundary of the Maintained Hypothesis," *Econometrica*, 69, pp. 683–734.

Andrews, D. (1999), "Estimation When a Parameter Is on a Boundary. *Econometrica*," 67, 1341–1383.

Apostolakis, B.E. (1990), "Energy-Capital Substitutability / Complementarity: The Dichotomy," *Energy Economics*, 1, pp. 48-58.

Barnett, W.A., J. Geweke, and M. Wolfe (1991), "Seminonparametric Bayesian Estimation of the Asymptotically Ideal Production Model," *Journal of Econometrics*, 49, No. 1/2, pp.5-50.

Antras, P. (2004), "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution," *Contributions to Macroeconomics*, Vol. 4, No. 1.

Barnett, W.A. (2002), "Tastes and technology: curvature is not sufficient for regularity," *Journal of Econometrics*, 108, pp. 199-202.

Barnett, W.A. and J.M. Binner (2004), "Functional Structure and Approximations in Econometrics," Elsevier.

Barnett, W.A. and M. Pasupathy (2003), "Regularity of the Generalized Quadratic Production Model: A Counterexample," *Econometric Reviews*, 22-2, pp. 135-154.

Berndt, E.R. and M.S. Khaled (1979), "Parametric Productivity Measurement and Choice among Flexible Functional Forms," *Journal of Political Economy*, 87, pp. 1220-1245.

Berndt, E.R. and D.O. Wood (1975), "Technology, Prices and the Derived Demand for Energy," *The Review of Economics and Statistics*, 57, pp. 259-268.

Blundell, R (2004), "Presidential address of the EEA/ESEM," (*European Economic Association* and *Econometric Society European Meeting*) conference in Madrid, August, 20, 2004.

Chambers (1988), "Applied Production Analysis, The Dual Analysis," Cambridge University Press.

Dasgupta, P.S. and G.M. Heal (1979), "Economic Theory and Exhaustible Resources," Cambridge University Press.

Diewert W.E. (2004), Preface. In: Barnett, W.A. and J.M. Binner (2004) "Functional Structure and Approximations in Econometrics," Elsevier.

Diewert, W. E., and Wales, T. J. (1987), "Flexible Functional Forms and Global Curvature Conditions," *Econometrica* 55, pp. 43-68.

Edwards, D. and D. Terrell (2004), "Does Theory Matter: Assessing the Impact of Monotonicity and Concavity Constraints on Forecasting Accuracy," Midwest Econometrics Group, *14th Annual Meeting, Northwestern University, Evanston, IL.*

Evans, D. S., and J. J. Heckman. (1984), "A Test for Subadditivity of the Cost Function with an Application to the Bell System," *American Economic Review* 74, pp. 615–623.

Evans, D. S., and J. J. Heckman. (1986), "A Test for Subadditivity of the Cost Function with an Application to the Bell System: Erratum," *American Economic Review* 76, pp. 856–858.

Gallant, A. R. and G.H. Golub (1984), "Imposing Curvature Restrictions on Flexible Functional Forms," *Journal of Econometrics*, 26, pp. 295-322.

Friesen, J. (1992), "Testing Dynamic Specification of Factor Demand Equations for U.S. Manufacturing," *Review of Economics and Statistics*, Vol. 74(2), pp. 240-50.

Griffiths, W.E., C.J. O'Donnell and A. Tan-Cruz (2000), "Imposing Regularity Conditions on a System of Cost and Factor Share Equations," *Australian Journal of Agricultural and Resource Economics*, 44, pp. 107-127.

Geweke, J. (1986), "Exact Inference in the Inequality Constrained Normal Linear Regression Model," *Journal of Applied Econometrics*, vol. 1, issue 2, pp. 127-41.

Henderson, Daniel J., John A. List, Daniel L. Millimet, Christopher F. Parmeter, Michael K. Price (2012), "Empirical implementation of nonparametric first-price auction models". *Journal of Econometrics,* Vol. 168(1), pp. 17-28.

Henderson, Daniel J. and Christopher F. Parmeter (2009), "Imposing Economic Constraints on Nonparametric Regression: Survey, Implementation and Extensions," *Advances in Econometrics: Nonparametric Methods*, Volume 25, pp. 433-469.

Hsieh, C. (2000), "Measuring Biased Technological Change," Princeton University, mimeo.

Koebel, B., M. Falk and F. Lasney (2003), "Imposing and Testing Curvature Conditions on a Box-Cox Cost Function," *Journal of Business and Economic Statistics*, 21(2), pp. 319-35.

Kumbhakar, S.C. and E.G. Tsionas (2005), "Measuring Technical and Allocative Inefficiency in the Translog Cost System: A Bayesian Approach," *Journal of Econometrics*, 126(2), pp. 355-384.

Lau, L.J. (1986), "Functional Forms in Econometric Model Building. Chapter 26 in Handbook of

Econometrics, "Vol. 3, Ed. Griliches, Z. and Intriligator, M.D., Elsevier Science, Amsterdam, North Holland, pp. 1515-1566.

Mankiw, G.N., D. Romer, and D. Weil (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics,* 107:2, pp. 407-437.

Matzkin, R.L. (1994), "Restrictions of Economic Theory in Nonparametric Methods," *Handbook of Econometrics*, Volume 4, chapter 42, Ed. R.F. Engle and D.L. McFadden, pp. 2524-58.

Moschini, G (1999), "Imposing Local Curvature Conditions in Flexible Demand System," *Journal of Business & Economic Statistics*, 17, pp. 487-490.

Parmeter, C., K. Sun, D. Henderson, and S. Kumbhakar (2013), "Regression and Inference Under Smoothness Restrictions", Working Paper, University of Miami.

Ryan, D.L. and T.J. Wales (1998), "A Simple Method for Imposing Local Curvature in Some Flexible Consumer Demand Systems," *Journal of Business and Economic Statistics*, 16(3), pp.331-38.

Salvanes, K.G. and S. Tjøtta (1998), "A Note on the Importance of Testing for Regularities for Estimated Flexible Functional Forms.," *Journal of Productivity Analysis*, 9, pp. 133-143.

Terrell, D. (1996), "Incorporating monotonicity and concavity conditions in flexible functional forms," *Journal of Applied Econometrics* 11, pp.179-194.

Tripathi, G. (2000), "Local Semiparametric Efficiency Bounds under Shape Restrictions," *Econometric Theory*, 16, pp. 729–739.

Wolff, H., T. Heckelei, and R.C. Mittelhammer (2010), "Imposing Curvature and Monotonicity on Flexible Functional Forms: An Efficient Regional Approach," *Computational Economics* 37(1), pp 309-339.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. John Wiley and Sons, New York.