A PROBABILISTIC NONPARAMETRIC ESTIMATOR (*)

Authors: *Juan Gabriel Rodríguez* (**) (a)
Rafael Salas (**) (b)

P. T. N.º 2/05

(*) This paper has benefited from the financial support of the European Commission under Project #ERBCHRXCT980248, the Spanish Ministry of Education under Project #PB98-0546-C0202, the Spanish Ministry of Science and Technology #SEC2003-08397, and the Fundación BBVA.

(**) We are grateful for helpful comments from M. A. Delgado; V. Dardanoni; J. M. Rodríguez Poo; E. Ferreira; P. Lambert; I. Perrote, and S. Yitzhaki. We also thank the participants in the 59th European Meeting of the Econometric Society, Madrid, 2004.

(a) Universidad Rey Juan Carlos de Madrid e Instituto de Estudios Fiscales. Madrid. Spain. jrodriguez@fcjs.urjc.es.

(b) Universidad Complutense de Madrid e Instituto de Estudios Fiscales. Madrid. Spain. r.salas@ccee.ucm.es.

N.B.: Las opiniones expresadas en este trabajo son de la exclusiva responsabilidad de los autores, pudiendo no coincidir con las del Instituto de Estudios Fiscales.

Desde el año 1998, la colección de Papeles de Trabajo del Instituto de Estudios Fiscales está disponible en versión electrónica, en la dirección: >http://www.minhac.es/ief/principal.htm.

Edita: Instituto de Estudios Fiscales
N.I.P.O.: 602-05-004-X
I.S.S.N.: 1578-0252
Depósito Legal: M-23772-2001

INDEX

ABSTRACT

This paper explores the adoption of a probabilistic nonparametric estimator in economics. First, it satisfies the *probabilistic coherence principle*, which ensures that the estimated variable can be generated by probabilistic assignment modeling from the observed variable. Second, it is proved to reduce variability in terms of noise, majorization and Lorenz dominance principles if, and only if, the estimator is probabilistic. The latter principles are universal criteria in risk and welfare economics, which expands the applicability of the estimator; for instance, to the measurement of economic discrimination. It also guarantees the symmetrical treatment of observations, a process that can produce smaller errors than positive-weight nonparametric estimators in terms of the bias-variance trade-off. This is verified by a general simulation exercise, with improvement due to the significant reduction in boundary bias. Finally, the estimator displays some other useful properties including consistency, preservation of the mean value, and multidimensional extension.

Key words: Nonparametric smoothing, majorization, Lorenz dominance, probabilistic estimator, noise.

1. INTRODUCTION

Nonparametric estimation of a regression curve has proved to be a useful tool for applied researchers in economics. For instance, Diebold and Nason (1990) have investigated the presence of nonlinearities in forecasting asset prices, Bierens and Pott-Buter (1991) and Delgado and Miles (1997) have applied nonparametric estimation of regression curves to the specification of Engel curves, and Bertschek and Entorf (1996) have used the classic Nadaraya-Watson nonparametric estimator to study the Schumpeterian link between innovation and firm size. Delgado *et al*. (2002) have also examined nonparametrically total factor productivity differences between exporting and nonexporting firms, ranking firms on the basis of stochastic dominance.

However, this paper suggests that improvements to standard nonparametric techniques are still possible. We propose a *probabilistic* nonparametric smoothing technique with the following properties. First, it satisfies the *probabilistic coherence principle*, which ensures a probabilistic link between the estimated and the observed variables. Standard nonparametric smoothers typically restrict the probabilistic link to every particular element of the estimated variable taken one by one, because of the row normalization of the weight matrix of the estimation. In the current probabilistic context, this paper extends the probabilistic link to the vector of the estimated variable as a whole, because of the double (rows and columns) normalization of the weight matrix. Therefore, it is able to generate the estimated variable from the observed variable through a probabilistic assignment modeling.[1]

The second property is that the estimator unambiguously reduces overall noise variability of the observed variable if, and only if, it is probabilistic. The estimated values are less dispersed (or more smoothed) than the observed ones in the sense that the estimated variable can be obtained from the observed one by subtracting the noise.

Moreover, the estimator reduces variability according to the robust majorization principle (Hardy *et al*., 1934; Marshall and Olkin, 1979, and Arnold, 1987), which is equivalent to the mean-preserving second-order stochastic dominance. As the mean value of the estimated variable remains unchanged, the estimation also achieves *Lorenz smoothing*, meaning that the Lorenz curve for the estimated variable dominates that of the observed variable. Accordingly, the estimated values are less dispersed than the observed ones in the sense that the observed variable can be obtained as a set of mean-preserving spreads of the estimated variable. Mean-preserving second-order stochastic dominance and Lorenz dominance are already well established in the risk (Rothschild, and Sti-

[1] We concentrate on the bivariate regression case. Analogous extensions apply to multivariate regressions. See property 8 below.

glitz, 1970) and welfare literature (Atkinson, 1970. Dasgupta; Sen, and Starret, 1973. Rothschild, and Stiglitz, 1973. Kolm, 1976. And Shorrocks, 1983).

This process allows a set of ordinal dispersion measures of the observed variable to be defined around the fitted variable, consistently with the second-order stochastic dominance criterion or with the distance between the Lorenz curves for both variables. The set of measures includes the difference of all cardinal dispersion measures for both variables, consistently with this criterion (*i.e.*, the difference of extended Gini coefficients (Donaldson, and Weymark, 1983, and Yitzhaki, 1983), Atkinson indices (Atkinson, 1970), and general entropy indices (Cowell, 1995) up to a positive monotone transformation.[2] Rodríguez *et al*. (2004), for example, propose this idea behind the probabilistic nonparametric technique to measure tax discrimination associated with a fiscal system.[3] Furthermore, its usefulness may be generalized to the measurement of economic discrimination.

Another practical property, which is a consequence of satisfying the probabilistic coherence principle, is that the estimator is observed to perform better than alternative standard nonparametric methods that use positive weights, on the basis of the bias-variance trade-off. This concentration on nonparametric methods that use positive weights is most useful because of their clear economic interpretation (see section 2) and is related to the improvement in boundary bias. The intervals are typically truncated at the boundaries so that, in general, observations in these intervals are given less importance in the construction of the estimator than the 'interior' points. As the new smoother treats all observations symmetrically, in the sense that they all receive the same aggregate weight in the process of construction of the nonparametric smoother, it gives greater weights to points near the boundary. Therefore, the estimator can alleviate the so-called *boundary bias problem*. Simulations over a set of different distributions of the explanatory variable confirm this result.

The proposed estimator has a number of additional desirable properties. First, it is consistent. Second, the mean of the estimated values is always equal to the mean of the observed values, irrespective of the number of observations (as in OLS estimation) so that the expected mean error equals zero. Third, the probabilistic estimator can be generalized to the multidimensional regression case. Moreover, the proposed probabilistic estimator is obtained from a simple low-cost modification of existing nonparametric techniques.

---

[2] Such as the coefficient of variation, which is an increasing transformation of the general entropy index for $c = 2$, or the variance when the mean of the distribution remains unchanged.

[3] In the field of inequality economics, nonparametric techniques have been used to estimate density curves. See, for instance, Hildenbrand, and Hildenbrand (1986); Cowell *et al*. (1996); Cowell, and Victoria-Fesser (1996), and Duclos, and Lambert (2000).

This paper itself is organized as follows. Section 2 provides a brief review of nonparametric smoothing. Section 3 defines the probabilistic coherence principle, and the majorization, Lorenz dominance and noise properties are examined in section 4. Section 5 deals with probabilistic nonparametric estimation and its properties while the simulation exercises are conducted in section 6. The final section includes some concluding remarks.

2. STANDARD NONPARAMETRIC SMOOTHING

Given any two-dimensional random sample, (X_1, Y_1), (X_2, Y_2), ... (X_n, Y_n),[4] the random variables, X and $Y \in \Re^n$, denote vectors of the explanatory and response variables, respectively. The *theoretical* regression curve $m(x_i)$ is defined as the expected value of Y at point $x_i \in \Re$,

$$m(x_i) = E(Y \mid X = x_i).$$

The *nonparametrically estimated* regression curve at point x_i, $M(x_i)$, can then be written as a weighted average of the observations on Y, such that:

$$M(x_i) = \sum_{j=1}^{n} W_j(x_i) Y_j$$

where the weights W_j, which downwardly weight the Y_j s if the corresponding X_j value is far from x_i, are probabilistic.[5] Reasons for concentrating on nonparametric methods that use nonnegative weights are their clear economic interpretation. For instance, when estimating an Engel curve, negative weights are either difficult to interpret or contribute to the generation of implausible (read negative) consumption values.[6] This restriction prevents working with estimators such as the local linear smoother, which optimizes the minimax risk criterion (Fan, 1993) but uses negative weights. For instance, W_j could be the Nadaraya-Watson weights (Nadaraya, 1964, and Watson, 1964):

$$W_j^{N-W}(x_i) = \frac{K\left(\dfrac{x_i - X_j}{h}\right)}{\sum_{j=1}^{n} K\left(\dfrac{x_i - X_j}{h}\right)}$$

[4] This refers to the stochastic design sample model. However, extension to the fixed design sample model is straightforward.

[5] A weight function is said to be a probability weight function if it is normal ($\Sigma_j W_j(x_i) = \Sigma_j W_{ij} = 1$) and nonnegative (see, for example Stone, 1977).

[6] In fact, Härdle (1990, p. 142) comments: "... it is highly recommended to use a positive kernel even though one has to pay a price in bias increase."

where kernel K is a continuous, bounded, and symmetric real function that integrates to unity (such as, for example, the normal density function) and h is the bandwidth smoothing value. The smoothing parameter h tends to zero as $n \to \infty$ and, for consistency, it is assumed that $nh \to \infty$ as $n \to \infty$. Consistency ensures that the estimated function converges to the theoretical one. The shape of the kernel weights is determined by K, whereas the size of the weights is parameterized by h.

However, many other nonparametric weights could be chosen, including the Priestley-Chao (1976) and Gasser-Müller (1979) smoothers or the k-th nearest-neighbor (k-NN) weights (Stone, 1977). The Priestly-Chao and Gasser-Müller estimators have more severe boundary bias problems than does the Nadaraya-Watson smoother, and for random designs has variances that are 50% higher than that of the local linear estimator (see, for example, Wand and Jones, 1995). Recall that the variance of the local linear smoother is higher than that of the Nadaraya-Watson estimator (this is confirmed by the simulation exercise below). Applying the simulation exercises used in section 6 to the Priestly-Chao and Gasser-Müller estimators obtains similar results.[7] This analysis does not, however, consider k-NN smoothing as this is equivalent to the kernel estimation when an appropriate bandwidth k parameter is used (see for example Härdle, 1990). Henceforth, nonparametric estimation is written in vector notation, $M = W \cdot Y$, where W is the weights matrix and M is the nonparametric smoother evaluated at any n points.

3. THE PROBABILISTIC COHERENCE PRINCIPLE

How should the estimated M and observed Y variables then be linked? Under the classical nonparametric estimation, the probabilistic weights criterion according to Stone's definition (see footnote 5) is satisfied. This is due to the assumption that W is stochastic; that is, the row sums are equal to unity (the row normalization feature). As a consequence, the probabilistic link is particularly restricted to any individual element of the estimated variable. Then we can write for every i:

$$M(x_i) = \sum_{j=1}^{n} W_j(x_i) Y_j = \sum_{j=1}^{n} W_{ji} Y_j, \quad \text{where} \quad \sum_{j=1}^{n} W_{ji} = 1.$$

In the current context, we extend the probabilistic link to the whole vector M with respect to Y. For every observation j we also have:

$$\sum_{i=1}^{n} W_{ji} = 1.$$

[7] The results can be obtained on request from the authors.

If (and only if) a probabilistic link over the whole vector M is assumed, we will be able to generate the estimated variable from the observed variable through a probabilistic assignment modeling. Formally we define this as follows.

DEFINITION 1. A bistochastic (or doubly stochastic) matrix is a square matrix in which all elements are nonnegative and all row and column sums are equal to unity. A particular case is a permutation matrix, which is a squared matrix with elements 0 and 1 and the row and column sums are equal to unity.

DEFINITION 2. *The probabilistic coherence principle*. Given any two variables $M, Y \in \Re^n$, M satisfies the probabilistic coherence principle, if and only if:

$$M = WY$$

where W is a bistochastic matrix.

In the nonparametric estimation context, this means that there is a probabilistic link between the observed variable Y and the estimated variable M. In other words, M can be generated from Y using probabilistic assignment modeling. Formally, we apply the following theorem.

THEOREM 1 (Birkhoff, 1946, and von Neumann, 1953). The set of bistochastic matrices is the convex hull of all permutation matrices. Formally, an n by n matrix is bistochastic if and only if it can be written in the form:

$$W = \sum_{i=1}^{n} p_i P_i$$

for some set of probabilities p_i and permutation matrices P_i; that is, the matrix is bistochastic if and only if it can be expressed as a convex combination of permutation matrices (the latter define the set of extreme points of the convex hull). A consequence of this theorem is that it is possible to decompose any bistochastic matrix additively in terms of a probability assignment model.

For example, let us assume that under a classical stochastic nonparametric estimator, the following stochastic matrix is obtained:

$$W = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0 & 0.3 & 0.7 \end{pmatrix}, \tag{1}$$

and the following bistochastic weight matrix W^B is derived from the stochastic one (according to a procedure explained below):

$$W^B = \begin{pmatrix} 0.81 & 0.11 & 0.08 \\ 0.19 & 0.68 & 0.13 \\ 0 & 0.21 & 0.79 \end{pmatrix}. \tag{2}$$

Accordingly, the bistochastic matrix W^B can be decomposed (making use of Birkhoff's algorithm; see Appendix A for details) as follows.

$$W^{B}=0.68\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}+0.13\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}+0.11\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}+0.08\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$$

A first interpretation is that the estimated variable coincides with the observed variable under a probability equal to 0.68. An adding-up property is that the overall effect of Y_1 on $M(x_1)$ of 0.81 can be now decompose on the additive 0.68 and 0.13 subeffects. Furthermore, the probabilistic coherence principle property guarantees a symmetrical treatment of observations that has a clear economic interpretation.

Symmetrical treatment of observations

The probabilistic coherence principle ensures a symmetrical (balanced) treatment of the observations, in the sense that they have the same aggregate weight, in contrast with the classical estimators that do not. This property rules out the possibility of over- or underweighted observations. For example, the weights in (1) are unbalanced (or not symmetrical) as the aggregate weight on M for observation Y_1 is 0.8, while for observation Y_2 it is 1.3 and for observation Y_3 it is 0.9. However, all observations in (2) have a symmetrical aggregate weight on M equal to unity [*i.e.*, the overall effect of Y_1 on $M(x_1)$, $M(x_2)$ and $M(x_3)$ is 0.81; 0.19 and 0.00, respectively]. As there is no *a priori* economic justification for an unequal treatment of observations, problems may arise because of difficulties in economic interpretation (see property 2 in section 5).

4. THE MAJORIZATION AND SECOND-ORDER STOCHASTIC DOMINANCE PROPERTIES

(Mean-preserving) second-order stochastic and Lorenz dominance are broadly defined criteria for smoothing or dispersion reduction in welfare and risk literature (Rothschild and Stiglitz, 1970, and Atkinson, 1970). Analogous concepts in the mathematical literature include the general majorization principle (Hardy *et al*., 1934; Marshall and Olkin, 1979, and Arnold, 1987).[8] Here there is a potential for improvement over the standard nonparametric approach, which comes from insisting that this majorization criterion be satisfied by nonparametric smoothing.

[8] Majorization is also broadly used in quantum mechanics in connection with the measurement of entropy or disorder, and it provides a natural language for expressing sharp fundamental constraints on the ability of quantum measurements to acquire information about a quantum system (see, for example, Nielsen, 2002).

We establish these principles formally. Let F and G be two n-dimensional real vectors. We use the notation $F^{\uparrow}$ to denote the vector whose elements are the entries of F reordered into increasing order, $F_1^{\uparrow} \leq F_2^{\uparrow} \leq \cdots \leq F_n^{\uparrow}$.

DEFINITION 3 (*Majorization*). Given any two vectors F and $G \in \Re^n$, *F is majorized by G*, written $F \overset{m}{\succ} G$, if and only if:

$$\sum_{i=1}^{k} F_i^{\uparrow} \geq \sum_{i=1}^{k} G_i^{\uparrow}$$

for $k = 1, ..., n$, with equality when $k = n$.[9] The last equality implies that both vectors have the same mean value. Majorization has a counterpart in economics under the title of *mean-preserving second-order stochastic dominance* (Rothschild and Stiglitz, 1970).

DEFINITION 4 (*Lorenz dominance*). Given any two vectors F and $G \in \Re^n$, and their relative transformations, say $F^R = \frac{1}{n\mu_F} F$ and $G^R = \frac{1}{n\mu_G} G$, where μ_F and μ_G denote the mean values of F and G, and sign $(\mu_F) = \text{sign}\,(\mu_G)$, *$F$ Lorenz dominates G*, written $F \overset{l}{\succ} G$, if and only if:

$$\sum_{i=1}^{k} F_i^{R\uparrow} \geq \sum_{i=1}^{k} G_i^{R\uparrow} \quad \text{for} \quad \mu_F \quad \text{and} \quad \mu_G \neq 0$$

for $k = 1, ..., n$. This Lorenz criterion is linked to majorization over relative vectors (which have the same mean value by construction).[10]

An important proposition that is used in what follows is:

THEOREM 2. Given two vectors F and $G \in \Re^n$, where $\mu_F = \mu_G \neq 0$, the following statements are equivalent.[11]

(a) $F \overset{m}{\succ} G$.

(b) $F \overset{l}{\succ} G$.

(c) F can be written as $F = V^B G$, where V^B is a bistochastic matrix.

(d) $\sum_{i=1}^{n} \Psi(F_i) \geq \sum_{i=1}^{n} \Psi(G_i)$ for every concave function $\Psi : \Re \rightarrow \Re$ (or convex function for the $\leq$ case).

---

[9] Formally, this concept is originally established in the equivalent version $\sum_{i=1}^{k} F_i^{\downarrow} \leq \sum_{i=1}^{k} G_i^{\downarrow}$, where vector F is re-ordered into decreasing order, $F_1^{\downarrow} \geq F_2^{\downarrow} \geq \cdots \geq F_n^{\downarrow}$. We adopt the definition in the main text for convenience.

[10] Note that this is a slight extension of the original Lorenz dominance criterion, which is defined for strictly positive elements. However, this extension only affects the generality of statement (b) in Theorem 2. Under the original Lorenz dominance criterion the domain should be restricted to $F, G \in \Re_{++}^n$, where $\mu_F = \mu_G \neq 0$.

[11] This theorem is, in fact, more general. The non-zero mean value is inessential to the proofs of the equivalence between (a), (c), (d), (e), (f) and (g), but it is required in the proof of the equivalence with (b).

(e) $\Phi(F) \geq \Phi(G)$ for every Schur-concave function $\Phi : \Re^n \rightarrow \Re$ (or Schur-convex for the $\leq$ case).[12]

(f) G has the same distribution as $F + R$, written $G =^d F + R$, where R is a random n-dimensional real vector that satisfies $E[R_i \mid F_i] = 0$ for all F_i.

(g) G is a mean-preserving spread of F, according to Rothschild and Stiglitz's definition (1970).

PROOF. Statement (a) is equivalent to (b) because $\mu_F = \mu_G$ by definition; statements (b), (c) and (e) are equivalent, see Dasgupta *et al*. (1973); (b) is equivalent to (f) and (g), see Rothschild, and Stiglitz (1970). Finally, propositions (a) and (d) are proved to be equivalent in Hardy *et al*. (1934), and also in Atkinson (1970).

The binary relations "$\overset{m}{\succ}$" and "$\overset{l}{\succ}$" are not complete, and they generate partial rather than total orderings of vectors. However, if F is majorized by (is more smoothed than) G, it achieves, in this mean-constant case, either a Lorenz or a mean-preserving spreads smoothing (by statements a, b and g). This is a general criterion that ensures a smoothing according to the wide class of convex, or the (even wider) class of Schur-convex, dispersion measures (statements d and e). Mean-constant variance-reducing smoothing is a particular case. Nonetheless, mean-constant variance-reducing smoothing does not imply Lorenz smoothing. In this respect, majorization and Lorenz dominance are more general or robust criteria for smoothing. Moreover, the majorization principle has a counterpart in terms of a noise-free distribution as established in statement (f). Notice that R, as defined, is usually interpreted as a noise term in the literature. Therefore, F can be viewed as a noise-free vector distribution of G.

In the next section, a nonparametric estimator is stated to satisfy majorization, Lorenz dominance and it is a noise-free estimator if and only if it is probabilistic.

5. THE PROBABILISTIC NONPARAMETRIC ESTIMATOR

The probabilistic nonparametric estimator is defined as follows.

DEFINITION 5. A nonparametric estimator, expressed in vector notation by $Z = W^B \cdot Y$, is said to be *probabilistic* if and only if W^B is a bistochastic weights matrix, which is normalized by both rows and by columns; that is $\sum_{i=1}^{n} W_{ij} = 1$ and $\sum_{j=1}^{n} W_{ij} = 1$.

[12] A function $f : \Re^n \rightarrow \Re$ is said to be Schur-concave if $x \overset{m}{\succ} y \Rightarrow f(x) \geq f(y)$. Notice that $f(\cdot)$ is Schur-convex if $-f(\cdot)$ is Schur-concave.

That is, if the curve is estimated at any n points and the weights matrix, represented by $W^B = \{W_{ij}\}_{i,j=1,\ldots,n}$, is bistochastic, then the estimator is probabilistic. The main difference between this estimator and the standard *stochastic* nonparametric estimator is that the latter is only normalized by rows.

A particular method of obtaining a probabilistic estimator is now proposed.[13] Given a nonparametric stochastic estimator, denoted by $M = W \cdot Y$, the following low-cost method of obtaining a probabilistic smoother, denoted by $Z = W^B \cdot Y$, is proposed. Of all the potential methods that could be used for double normalization of the weights matrix, the so-called iterative proportional-fitting method is adopted. In turn, this method is a special case of the algorithm proposed by Deming and Stephan (1940), which minimizes the Kullback-Liebler distance function between W and W^B (see property 5). Moreover, the algorithm converges when the observed weights are all nonnegative and estimates are consistent and asymptotically normal.

The algorithm is an iterative-fitting method applied to the initial elements W_{ij}, and it proceeds by row and column adjustments, such that at iteration t ($\forall t \in \mathbb{N}$), the new elements of the matrix of weights are:

$$W_{ij}^{(0)} = W_{ij}.$$

If t is odd,

$$W_{ij}^{(t)} = \frac{W_{ij}^{(t-1)}}{W_{i+}^{(t-1)}} = \frac{W_{ij}}{W_{i+}^{(t-1)} \times \cdots \cdots \times W_{+j}^{(1)} \times W_{i+}^{(0)}},$$

and if t is even,

$$W_{ij}^{(t)} = \frac{W_{ij}^{(t-1)}}{W_{+j}^{(t-1)}} = \frac{W_{ij}}{W_{+j}^{(t-1)} \times \cdots \cdots \times W_{+j}^{(1)} \times W_{i+}^{(0)}},$$

where $W_{+j}^{(t)} = \sum_{i=1}^{n} W_{ij}^{(t)}$, $\forall j = 1, \ldots, n$; $W_{i+}^{(t)} = \sum_{j=1}^{n} W_{ij}^{(t)}$, $\forall i = 1, \ldots, n$; and $t \in \mathbb{N}$.

The following example clarifies the proposed iterative method. Let us start with the matrix W in (1). Iteration 1 normalizes by columns (by dividing every element in the column by the column sum), but now the normalization by rows is not verified (row sums are not equal to unity). Iteration 2 renormalizes by rows (by dividing every element in the row by the row sum) and so on. Finally, after 11 iterations, the bistochastic matrix in (2) is obtained:

[13] Another nonparametric technique is the regressogram (Tukey, 1947), which guarantees that the weights matrix is bistochastic (see the proof in Appendix B). However, this estimator is rarely used because of a lack of desirable properties.

$$\begin{array}{ll} W^{(0)} = \left[\begin{array}{ccc} 0.70 & 0.20 & 0.10 \\ 0.10 & 0.80 & 0.10 \\ 0 & 0.30 & 0.70 \end{array}\right]\begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \end{array} \rightarrow & W^{(1)} = \left[\begin{array}{ccc} 0.88 & 0.15 & 0.11 \\ 0.12 & 0.62 & 0.11 \\ 0 & 0.23 & 0.78 \end{array}\right]\begin{array}{c} 1.14 \\ 0.85 \\ 1.01 \end{array} \rightarrow \\ \qquad\quad \begin{array}{ccc} 0.80 & 1.30 & 0.90 \end{array} & \qquad\quad \begin{array}{ccc} 1.00 & 1.00 & 1.00 \end{array} \end{array}$$

$$\begin{array}{ll} W^{(2)} = \left[\begin{array}{ccc} 0.77 & 0.13 & 0.10 \\ 0.15 & 0.72 & 0.13 \\ 0 & 0.23 & 0.77 \end{array}\right]\begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \end{array} \rightarrow \ldots \rightarrow & W^{B} = \left[\begin{array}{ccc} 0.81 & 0.11 & 0.08 \\ 0.19 & 0.68 & 0.13 \\ 0 & 0.21 & 0.79 \end{array}\right]\begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \end{array} \\ \qquad\quad \begin{array}{ccc} 0.92 & 1.08 & 1.00 \end{array} & \qquad\quad \begin{array}{ccc} 1.00 & 1.00 & 1.00 \end{array} \end{array}$$

There is an interesting symmetry condition relating to this particular algorithm. The convergence result for this algorithm is $W^B = \{W_{ij}^{(T)}\}_{1 \le i \le n, 1 \le j \le n}$, which applies whether one begins by normalizing by rows or columns, where *T* denotes the final iteration according to a sufficiently accurate stopping rule.[14]

What, then, is the computational cost of this algorithm? In principle, as the sample size is *n*, the curve must be evaluated at *n* points to obtain a square weight matrix. Then, $O(n^2)$ kernel evaluations are necessary before the iterative method is applied to the stochastic nonparametric estimator to obtain the bistochastic (probabilistic) one. This may make the computation of the probabilistic estimator very slow for a sufficiently large value of *n*. One way of dramatically increasing the computational speed is to compute the *binned kernel regression estimator* (see, for example, Georgiev, 1986; Fan, and Marron, 1994) before the bistochastic smoothing is applied. The binned regression smoothing technique replaces kernel estimators by approximations that can be computed quickly by using the fast Fourier transform. The intent is to replace the data by a mesh of *R* grid counts, where each grid count is a weight that represents the amount of data near the corresponding grid point. The approximation is generally very good for moderate values of *R*, and it can be made arbitrarily better by increasing the value of *R*.

Next, we analyze the consistency property of the reformulated estimator.

PROPERTY 1 (consistency of the estimator). Let $\{W_n\}$ be a consistent sequence of probability weights (as defined in footnote 5) and let $\{W_n^{(T)}\}$ be a sequence of weights of the proposed probabilistic estimator. Then, it follows that $\{W_n^{(T)}\}$ is consistent.

PROOF. The bistochastic sequence of weights can be written as:

$$W_n^{(T)} = f(W_n) \cdot W_n,$$

[14] More generally, this method can be applied to any *r* x *n*-dimensional nonsquare weights matrix *W*. Then, *W* is not properly bistochastic, but the double normalization property is retained.

where *f* is a bounded function because $\{W_n^{(T)}\}$ is a sequence of normal weights. Applying the result of Stone (1977, Corollary 2, p. 598) reveals that the sequence $\{W_n^{(T)}\}$ is consistent.

In addition, the probabilistic estimator has the following properties.

PROPERTY 2 (probabilistic coherence). The estimator satisfies the probabilistic coherence principle by construction; therefore, we can interpret the estimator in terms of a probabilistic assignment model (see section 3). A relevant implication of this property is the symmetrical treatment of observations. All observations have the same aggregate weight in the construction of the nonparametric estimator *Z*, in the sense that the sum of the (*across-interval*) weights assigned to any observation Y_i is the same (and is eventually unity). This is equivalent to imposing normalized summation across columns in the weights matrix *W*, which is what the probabilistic estimator does by definition. Note that classical stochastic estimators simply normalize weights *within intervals* because the weights matrix *W* is only normalized by rows. In the probabilistic case, both *across*- and *within-interval* weights are normalized to unity. Note that under this methodology, asymptotic unbiasedness is retained, as row normalization is verified.

Furthermore, a related issue is that the estimator can alleviate the so-called *boundary bias problem*. Aggregate weights of the extreme observations at the boundaries are typically less than unity, while those at the central distribution typically exceed unity; as in the initial *W* matrix in expression (1). The reason lies in the truncation of the intervals at the boundaries. As the new smoother gives greater weights to points near the boundary, it may improve the goodness of the fit at the boundaries. The empirical exercise used in this analysis, and applied to a set of different distributions of the variable *X*, confirms this result.

PROPERTY 3 (Majorization and Lorenz dominance consistency). The bistochastic smoothing technique is consistent with majorization, mean-preserving second-order stochastic dominance and Lorenz dominance. For instance, the Lorenz curve for *Z*:

$$L_Z(k/n) = \sum_{i=1}^{k} Z_i^{R\uparrow},$$

always lies above the Lorenz curve for *Y*, when $\mu_Z = \mu_Y \neq 0$:

$$L_Z(k/n) \geq L_Y(k/n) \quad \forall k = 1,\cdots,n.$$

PROOF. This proof applies the Theorem 2 of Dasgupta *et al.* (1973). A necessary and sufficient condition for $L_Z(k/n) \geq L_Y(k/n)\ \forall k = 1,\cdots,n$ is that $Z = W^B Y$, where W^B is bistochastic. This result is even more general as it is also second-order stochastic dominance consistent. The reasoning is that the mean of the dependent variable remains constant (see property 6 below).[15]

[15] *Second-order stochastic dominance*, *generalized Lorenz dominance* (Shorrocks, 1983) and *supermajorization* are equivalent criteria. The only difference with respect to majorization in definition 3 is the omission of the equality requirement at $k = n$.

As a consequence of this property, the estimated values can be obtained from the original values as a set of unambiguous Lorenz-variability-reducing mean-preserving spreads, if and only if the estimator is probabilistic. In particular, they can be obtained as a set of variance-reducing mean-preserving spreads. Moreover, Lorenz dominance is a general criterion that ensures the variability reduction according to the wide set of Schur-convex measures of dispersion (see section 4), such as the extended Gini coefficients, Atkinson indices or the general entropy measures.

Application

This property has potentially important applications in economics, particularly applications concerning the measurement of discrimination. In a recent paper (see Rodríguez *et al*., 2004), the probabilistic estimator has been applied to the measurement of tax discrimination (horizontal inequity associated with a tax system).

Let X and $Y \in \Re^n_{++}$ be the pre- and post-tax equivalent income distributions, respectively, and let Z be the estimated post-tax equivalent income by the nonparametric bistochastic technique. The higher the dispersion of Y around Z, the greater the discrimination associated with tax. In Rodríguez *et al*. (2004), it is shown that dispersion of Y around Z exists if and only if similar individuals with close-to-equal pre-tax income levels (associated to a particular bandwidth h) pay different taxes; that is, there is tax discrimination. In this approach, Z becomes the discrimination-free benchmark distribution.

This notion of dispersion or inequality of Y around Z can be measured by any monotone transformation of any distance function between the Lorenz curves of Y and Z. This implies, in fact, an ordinal representation of discrimination and it satisfies the important property of consistency with the standard Lorenz and mean-preserving second-order stochastic dominance criteria. The Lorenz smoothing property of the probabilistic nonparametric estimator ensures that Z Lorenz dominates Y and therefore ensures that the distance between both Lorenz curves is nonnegative. Notice that negative discrimination values have no economic meaning.

Extensions to more general discrimination frameworks, such as wage discrimination, may be done by computing the distance between the actual and the estimated wage Lorenz curves for workers with close-to-equal attributes and the same properties.[16] Both curves only differ whenever workers with close-to-equal attributes and similar productivity levels receive different wages. With im-

[16] Jenkins (1994) also uses a Lorenz curve-based methodology to measure wage discrimination.

perfect information, the unintended dispersion due to incomplete information may be discounted. This is the basis of the measurement of discrimination that the authors intend to develop in future research.

PROPERTY 4. $Z =^{d} Y - R$ where $E[R_i | Z_i] = 0$ for all Z_i. Noise is the sole element that is eliminated by the smoothing process if and only if the estimator is probabilistic. Otherwise, the estimator would probably eliminate some signal (or fail to remove all noise). See Theorem 2 in section 4 for the proof.

PROPERTY 5. The probabilistic estimator vector, to which this algorithm converges, is $Z = W^B \cdot Y$, where W^B is the closest bistochastic matrix to W, according to the Kullback-Liebler distance function $D(W^B, W)$:

$$D(W^B, W) = \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij}^B \ln\left(\frac{W_{ij}^B}{W_{ij}}\right).$$

The proof applies the result of Ireland and Kullback (1968). This distance function is widely used in the information theory approach for both generation of models and estimation of parameters.

PROPERTY 6 (zero expected estimation error). An implicit property is that the estimator Z and the variable Y have the same mean, $\mu(Z) = \mu(W^B \cdot Y) = \mu(Y)$, whatever the sample size (as in OLS estimation), because of the bistochastic matrix link between them. Hence, the expected estimated error is zero, unlike in standard stochastic estimation. This property is, in fact, implied by property 4.

In the following example, assume that $Y = (6000, 10000, 15000)$ is a vector of individual taxes. Then, suppose that the estimated stochastic W and bistochastic W^B are as in (1) and (2). In this case, the estimated taxes are $M = (7700, 11000, 13500)$ and $Z = (7160, 9890, 13950)$, respectively. The inherent adding-up property in the probabilistic estimator guarantees that the overall estimated taxes equal the actual figure of 31000, irrespective of the sample size. However, under the stochastic estimator, we obtain an overall amount of 31300. This adding-up accountant property can be also useful in the nonparametric smoothing of earnings, profits, savings, GDP, etc.

PROPERTY 7. Let (X_1, Y_1), (X_2, Y_2), …, (X_n, Y_n) be a two-dimensional random sample, let M be a nonparametric estimator of the regression curve at n different points, and let Z be the bistochastic reformulation of M. Then, each element of Z is a convex combination of the M elements. See Appendix B for the proof.

PROPERTY 8. The probabilistic estimator can be generalized to the multivariate regression case given that the Deming-Stephan algorithm can be applied to higher dimensions. Property 8 allows one to generalize the application of the probabilistic nonparametric technique to a multidimensional framework.

6. A SIMULATION EXERCISE

In this section, we test the efficiency of the proposed probabilistic estimator in terms of the bias-variance trade-off, as measured by the conditional quadratic error for different sample sizes, and compare it with standard stochastic estimators, including the Nadaraya-Watson estimator and local linear estimator.

We show that the probabilistic smoother performs better than the alternative standard nonparametric methods that employ positive weights, on the basis of the conditional quadratic error. The mean integrated square error (MISE) and other asymptotically equivalent measures are not considered in this analysis because there is no explicit expression for the probabilistic estimator described in section 5. For this reason, we undertake the following simulation exercise.

Design of the Exercise

In this simulation exercise, we evaluate the performance of the probabilistic estimator according to the conditional quadratic error (d_c),

$$d_c = E\left[n^{-1}\sum_{j=1}^{n}\left(Z(x_j) - m(x_j)\right)^2 \mid x_1,\ldots,x_n\right]$$

where $m(x)$ is the true curve and $x_1, \ldots, x_n$ is a particular sample. It is well known that d_c can be decomposed additively into bias and variance components (see, for example, Härdle, 1990, p. 148).

We compute the bias, variance and conditional quadratic error for three nonparametric smoothers. These are the Nadaraya-Watson, the bistochastic Nadaraya-Watson and the local linear estimator. Data on X have been generated from two different distributions, the standard normal, $X_i \sim N(0,1)$, and the uniform, $X_i \sim U(0,1)$. Data on Y were generated from the model, $Y_i = m(X_i)+\varepsilon_i$, where $\varepsilon_i \sim N(0,\sigma^2)$, with $m(X_i)$ being either $m(X_i) = \mathrm{Sin}(2\pi X_i)$ or $m(X_i) = \exp(X_i)$. These two different specifications imply two quite different models. We also perform simulations using two different values for σ^2, namely $\sigma^2 = 0.2$ and $\sigma^2 = 0.8$, in both models. The simulations were developed from three different sample sizes, $n = 30$, $n = 100$ and $n = 1000$. In total, we performed $2\times2\times2\times3\times3\times3 = 216$ basic computations. The results were then obtained by taking the mean of 200 independent samples or repetitions of each basic computation.[17]

Results

Results for the experiment are presented in Table 1 for the N(0,1) case. The results for the U(0,1) distribution are very similar and for the purposes of bre-

[17] Different specifications for the distribution function (lognormal) and for the $m(X)$ function (polynomials) were tested, but did not significantly alter the main results.

vity are not presented. The results are evaluated for the optimal bandwidth obtained according to the cross-validation function. However, we show below that this is not an important aspect of the analysis. The results also show that in all cases examined, the conditional quadratic error for the bistochastic smoother is lower than that for the Nadaraya-Watson estimator. In fact, the bias is substantially reduced by the bistochastic reformulation, while the increase in the variance is insufficient to offset this reduction.[18] Consequently, the probabilistic estimator provides a better fit.

Table 1

BIAS, VARIANCE AND d_c FOR THE DIFFERENT ESTIMATORS UNDER $X_i \sim N(0,1)$ (*)

n=30												
	$\sin(2\pi X_i)$						$\exp(X_i)$					
	$\sigma_\varepsilon^2 = 0.2$			$\sigma_\varepsilon^2 = 0.8$			$\sigma_\varepsilon^2 = 0.2$			$\sigma_\varepsilon^2 = 0.8$		
	Bias	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**
N-W	0.0226	0.0750	0.0976	0.1236	0.2145	0.3380	0.0929	0.0350	0.1279	0.2160	0.1085	0.3245
B N-W	0.0192	0.0762	0.0954	0.1184	0.2185	0.3370	0.0592	0.0361	0.0953	0.1153	0.1135	0.2288
L-L	0.0099	0.0911	0.1010	0.1035	0.2619	0.3654	0.0134	0.0440	0.0574	0.0214	0.1464	0.1678
n=100												
	$\sin(2\pi X_i)$						$\exp(X_i)$					
	$\sigma_\varepsilon^2 = 0.2$			$\sigma_\varepsilon^2 = 0.8$			$\sigma_\varepsilon^2 = 0.2$			$\sigma_\varepsilon^2 = 0.8$		
	Bias	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**
N-W	0.0100	0.0364	0.0464	0.0257	0.1131	0.1388	0.0469	0.0143	0.0612	0.1189	0.0453	0.1642
B N-W	0.0079	0.0371	0.0450	0.0221	0.1151	0.1372	0.0249	0.0148	0.0397	0.0567	0.0477	0.1044
L-L	0.0044	0.0421	0.0465	0.0161	0.1297	0.1459	0.0043	0.0174	0.0217	0.0116	0.0584	0.0701
n=1000												
	$\sin(2\pi X_i)$						$\exp(X_i)$					
	$\sigma_\varepsilon^2 = 0.2$			$\sigma_\varepsilon^2 = 0.8$			$\sigma_\varepsilon^2 = 0.2$			$\sigma_\varepsilon^2 = 0.8$		
	Bias	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**	**Bias**	**Variance**	**d_c**
N-W	0.0018	0.0068	0.0087	0.0037	0.0228	0.0266	0.0097	0.0028	0.0125	0.0266	0.0086	0.0353
B N-W	0.0015	0.0069	0.0084	0.0032	0.0231	0.0263	0.0039	0.0029	0.0067	0.0088	0.0090	0.0178
L-L	0.0011	0.0074	0.0085	0.0027	0.0247	0.0274	0.0006	0.0032	0.0038	0.0019	0.0103	0.0122

(*) Results are the average of 200 independent samples, and are evaluated for the optimal bandwidth according to the cross-validation function. N-W, B N-W and L-L denote the Nadaraya-Watson, bistochastic Nadaraya-Watson and local linear estimators, respectively.

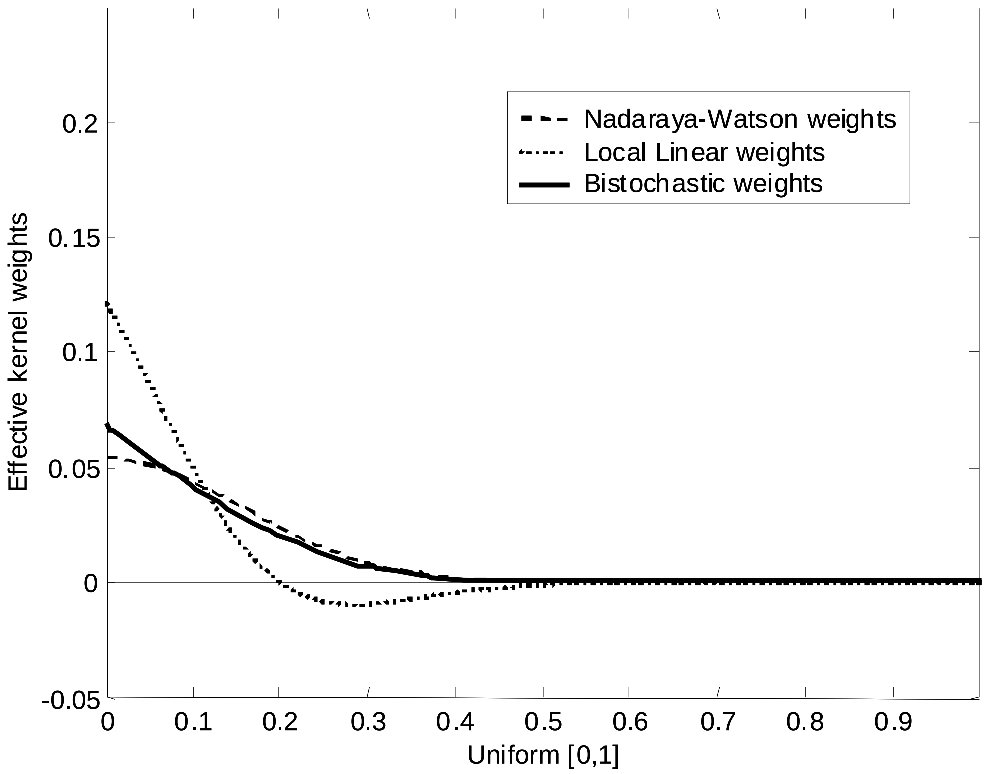
[18] Note that this increase in the variance does not contradict second-order stochastic dominance, because in this case, there is no stochastic dominance between the Nadaraya-Watson and bistochastic methods.

How are these results explained? The results are due to the way in which the bistochastic method corrects the boundary bias, which is apparent from Figure 1. Figure 1 shows the effective kernel weights associated with the lowest boundary value, for the whole range of values.[19] Probabilistic estimators give greater weights to boundary values than do standard stochastic methods. Hence, they tend to alleviate the boundary bias problem. Figure 1 shows that local linear estimators do even better (in Table 1, bias and d_c are even lower). However, we do not consider these estimators because they use negative weights elsewhere, as shown in Figure 1.

This raises the question of why the probabilistic smoother changes the boundary weights appropriately. The answer is related to the symmetrical treatment of the observations that is due to the double normalization (see property 2). The intervals are truncated at the boundaries so that, in general, observations in these intervals have less importance in the construction of the estimator than the 'interior' points. Since the new smoother gives greater weights to points near the boundary, it can improve the performance of the estimator with respect to the so-called *boundary bias problem*.

Figure 1

EFFECTIVE KERNEL WEIGHTS EVALUATED AT THE LOWEST X_i VALUE



Exercise design: $Y_i = m(X_i) + \varepsilon_i$, where $X_i \sim U(0,1)$, $m(X_i) = \exp(X_i)$ and $\varepsilon_i \sim N(0,0.2)$.

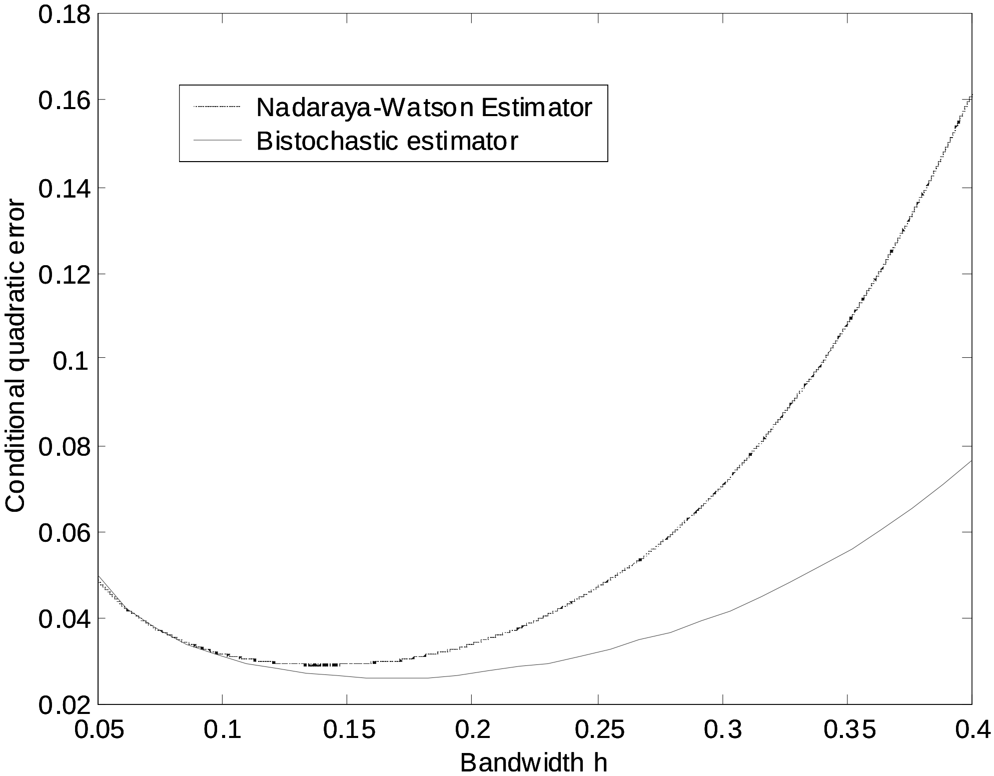
---

19 We present the uniform distribution case to avoid negative values in the *X*-variable axes. Results are similar for the normal distribution case.

The results are even more general. Figure 2 shows the conditional quadratic error for a wide range of bandwidth values. The greater efficiency of the bistochastic estimator, relative to the classical estimator, seems independent of the bandwidth value used. In particular, note especially the greater efficiency in the neighborhood of the optimal bandwidth for the stochastic estimator, whatever the optimal bandwidth criteria used. Also note that the efficiency gain may be even higher, since the optimal bandwidth for the probabilistic estimator differs from that of the stochastic smoother. This is the case in Figure 2. We further suggest computing an algorithm process for the whole range of bandwidths to achieve the most efficient probabilistic estimator.

Figure 2

CONDITIONAL QUADRATIC ERROR (d_c) FOR DIFFERENT h BANDWIDTH VALUES



Exercise design: $Y_i = m(X_i) + \varepsilon_i$, where $X_i \sim N(0,1)$, $m(X_i) = \exp(X_i)$ and $\varepsilon_i \sim N(0,0.2)$. The d_c function values correspond to the mean value for 200 independent samples.

Furthermore, the probabilistic estimator converges more quickly to the true curve than does the classical Nadaraya-Watson smoother. Note that the (negative) rate of variation of d_c for the bistochastic estimator is greater with respect to n than for the Nadaraya-Watson in all cases. For instance, the rates of variation of the bistochastic and Nadaraya-Watson estimators, between $n = 30$ and $n = 100$, are -58.3 and -52.2 percent, respectively, in the exponential and $\sigma^2 = 0.2$ case. Corresponding values between $n = 100$ and $n = 1000$ are -83.0 and -78.5, respectively, in the exponential and $\sigma^2 = 0.8$ case. By contrast, we find no evidence that the probabilistic estimator is superior to the local linear smoother in this respect.

7. CONCLUDING REMARKS

This paper analyzes the implications of adopting probabilistic nonparametric smoothing which satisfies a basic property that has been overlooked in the literature. The smoothing ensures that all noise is eliminated in the estimation if, and only if, it is probabilistic. Otherwise, the smoothing cannot avoid either eliminating some signal or leaving some noise.

The probabilistic or bistochastic smoother has also the following theoretical advantages. A probabilistic assignment model can be established between the estimated and the dependent variables. It implies an important additive decomposition property. It is also proved that the estimator reduces variability according to the robust criterion of majorization and Lorenz dominance if, and only if, it is probabilistic. In addition, it is shown that the estimator displays some other useful properties including consistency, symmetrical treatment of observations, and zero expected mean error irrespective of the number of observations (as the OLS case). Moreover, the smoothing can be generalized to the multidimensional case.

The probabilistic smoother itself imposes a double normalization of the weights matrix of the estimator, which is performed by using the low-cost iterative proportional-fitting algorithm proposed by Deming and Stephan (1940). This algorithm minimizes the Kullback-Liebler distance function with respect to the original weights matrix of the stochastic estimator.

Among the practical advantages of the probabilistic estimator, we find that the underlying Lorenz smoothing property of the estimator enhances the usefulness of the smoother in applied economics (in particular, in economic discrimination measurement). Moreover, there is an improvement over standard nonparametric methods that use positive weights in terms of the bias-variance trade-off as it is confirmed in a general simulation exercise. This raises a paradox. How can it be possible that the imposition of a new restriction (column normalization) improves the efficiency of the estimator? The probabilistic estimation takes into consideration information across intervals (see property 2 in section 5) that standard stochastic methods do not. Therefore, in the simulation exercise, the efficiency gain associated with the better use of information more than offsets the efficiency loss from the imposition of the new restriction. However, this line of inquiry requires extensive research beyond the scope of the present paper.

APÉNDIX A. THE BIRKHOFF ALGORITH (Birkhoff, 1946)

The way the Birhkoff algorithm operates can be easily seen from the example in (2):

$$W^B = \begin{bmatrix} \langle 0.81\rangle & 0.11 & 0.08 \\ 0.19 & \langle 0.68\rangle & 0.13 \\ 0 & 0.21 & \langle 0.79\rangle \end{bmatrix} = 0.68\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \langle 0.13\rangle & 0.11 & 0.08 \\ 0.19 & 0 & \langle 0.13\rangle \\ 0 & \langle 0.21\rangle & 0.11 \end{bmatrix}$$

$$= 0.68\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 0.13\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \langle 0.11\rangle & 0.08 \\ \langle 0.19\rangle & 0 & 0 \\ 0 & 0.08 & \langle 0.11\rangle \end{bmatrix}$$

$$= 0.68\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 0.13\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + 0.11\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \langle 0.08\rangle \\ \langle 0.08\rangle & 0 & 0 \\ 0 & \langle 0.08\rangle & 0 \end{bmatrix}$$

$$= 0.68\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 0.13\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + 0.11\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 0.08\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

APÉNDIX B. THE REGRESSOGRAM AS A BISTOCHASTIC ESTIMATOR

The regressogram is defined as the arithmetic mean of the Y variable across the corresponding h nonoverlapping intervals. The regressogram estimator ensures that the weights assigned to each observation of the response variable sum to unity, not only across rows but also across columns; that is, the weights matrix is bistochastic.

PROOF. Let $S = \{s_1(X), \ldots, s_h(X)\}$ be the nonoverlapping partition into h subgroups under consideration, let $U = \{n_1, \ldots, n_h\}$ be the within-groups population set, and let $V = \{\mu_1, \ldots, \mu_h\}$ be the associated response mean variable set. For the regressogram estimator,

$$z_1^i = \ldots = z_{n_i}^i = \mu_i, \quad \forall\, i = 1, \ldots, h.$$

In vector notation, $M = BY$, where B is the n-dimensional bistochastic matrix,

$$B = \begin{pmatrix} N_1 & 0 & \cdots & 0 \\ 0 & N_2 & \cdots & 0 \\ \cdots & \cdots & \ddots & 0 \\ 0 & 0 & \cdots & N_h \end{pmatrix},$$

and N_i is the n_i-dimensional square matrix,

$$N_i = \begin{pmatrix} 1/n_i & \cdots & 1/n_i \\ \cdots & \ddots & \cdots \\ 1/n_i \cdots & \cdots & 1/n_i \end{pmatrix} \quad \forall i = 1, \ldots, h.$$

APÉNDIX C. PROOF OF PROPERTY 6

The bistochastic estimator vector can be written as

$$Z = W^B \cdot Y = W^B \cdot W^{-1} \cdot W \cdot Y = W^B \cdot W^{-1} \cdot M.$$

In desegregated terms, it can be written as

$$\begin{aligned} Z(x_i) = & \left[W_{i1} \cdot W_{11}^{-1} + .. + W_{ij} \cdot W_{i1}^{-1} + .. + W_{in} \cdot W_{n1}^{-1}\right] \cdot M(x_1) + ... \\ & + \left[W_{i1} \cdot W_{1j}^{-1} + .. + W_{ij} \cdot W_{ij}^{-1} + .. + W_{in} \cdot W_{nj}^{-1}\right] \cdot M(x_i) + ... \\ & + \left[W_{i1} \cdot W_{1n}^{-1} + .. + W_{ij} \cdot W_{in}^{-1} + .. + W_{in} \cdot W_{nn}^{-1}\right] \cdot M(x_n) \end{aligned}$$

Hence, the sum of terms within the square brackets,

$$\left[W_{i1} \cdot W_{11}^{-1} + .. + W_{ij} \cdot W_{i1}^{-1} + .. + W_{in} \cdot W_{n1}^{-1}\right] + ... + \left[W_{i1} \cdot W_{1n}^{-1} + .. + W_{ij} \cdot W_{in}^{-1} + .. + W_{in} \cdot W_{nn}^{-1}\right]$$

must be unity. The above expression can be rewritten as

$$W_{i1}\left[W_{11}^{-1} + .. + W_{1j}^{-1} + .. + W_{1n}^{-1}\right] + ... + W_{in}\left[W_{n1}^{-1} + .. + W_{nj}^{-1} + .. + W_{nn}^{-1}\right]$$

We only need demonstrate that the inverse of a stochastic matrix sums to unity across rows.

Let $A = (a_{ij})_{i,j=1,..,n}$ be a stochastic matrix and let B be its inverse. Since $B{\cdot}A = I_n$ (I is the identity matrix), we know that

$$I_{ij} = b_{i+} \cdot a_{+j} = \sum_{k=1}^{n} b_{ik} \cdot a_{kj}.$$

Hence, from the stochastic property of matrix A, we obtain

$$\sum_{j=1}^{n} I_{ij} = 1 = \sum_{j=1}^{n}\sum_{k=1}^{n} b_{ik} a_{kj} = \sum_{k=1}^{n} b_{ik} \sum_{j=1}^{n} a_{kj} \Rightarrow \sum_{k=1}^{n} b_{ik} = 1.$$

REFERENCES

ATKINSON, A. (1970): "On the measurement of inequality", *Journal of Economic Theory* 2, 244-263.

ARNOLD, B. C. (1987): *Majorization and the Lorenz order: A brief introduction* (Springer-Verlag, Berlin).

BERTSCHEK, I., and ENTORF, H. (1996): "On nonparametric estimation of the Schumpeterian link between innovation and firm size: Evidence from Belgium, France and Germany", *Empirical Economics* 21(3), 401-426.

BIERENS, H. J., and POTT-BUTER, H. A. (1991): "Specification of household Engel curves by nonparametric regression", *Econometric Reviews* 9, 123-184.

BIRKHOFF, G. (1946): "Three observations on linear algebra", *Universidad Nacional de Tucumán Rev.,* Serie A 5, 147-151.

COWELL, F. A. (1995): *Measuring inequality* (Prentice Hall/Harvester Wheatsheaf, Hemel Hempstead).

COWELL, F. A.; JENKINS, S. P. and LITCHFIELD, J. A, (1996): "The changing shape of the UK income distribution: kernel density estimates", in: J. HILLS, (ed.), *New Inequalities: The Changing Distribution of Income and Wealth in the United Kingdom* (Cambridge University Press, Cambridge) 49-75.

COWELL, F. A., and VICTORIA-FESSER, M. P. (1996): "Robustness properties of inequality measures", *Econometrica* 64(1), 77-101.

DASGUPTA, P.; SEN, A., and STARRET, D. (1973): "Notes on the measurement of inequality", *Journal of Economic Theory* 6, 180-187.

DELGADO, M. A.; Fariñas, J. C., and RUANO, S. (2002): "Firm productivity and export markets: a nonparametric approach", *Journal of International Economics* 57, 397-422.

DELGADO, M. A., and MILES, D. (1997): "Household characteristics and consumption behaviour: A nonparametric approach", *Empirical Economics* 22, 409-429.

DEMING, W. E., and STEPHAN, F. F. (1940): "On a least squares adjustment of a sampled frequency table when the expected marginal tables are known", *The Annals of Mathematical Statistics* 11, 427-444.

DIEBOLD, F. X., and NASON, J. A. (1990): "Nonparametric exchange rate prediction", *Journal of International Economics* 28, 315-332.

DONALDSON, D., and WEYMARK, J. A. (1980): "A single parameter generalization of the Gini indices of inequality", *Journal of Economic Theory*, 22, 67-86.

DUCLOS, J. Y., and LAMBERT, P. (2000): "A normative and statistical approach to measuring classical horizontal inequity", *The Canadian Journal of Economics* 33, 87-113.

FAN, J. (1993): "Local linear regression smoothers and their minimax efficiency", *The Annals of Statistics* 21, 196-216.

FAN, J., and MARRON, J. S. (1994): "Fast implementations of nonparametric curve estimators", *Journal of Computational Graphical Statistics* 3, 35-56.

GASSER, T., and MÜLLER, H. G. (1979): "Kernel estimation of regression functions", *Smoothing techniques for curve estimation. Lectures notes in math* 757, 23-68 (Springer, New York).

GEORGIEV, A. A. (1986): "A fast algorithm for curve fitting", in: F. DE ANTONI; N. LAURO, and A. RIZZI, (eds.), *Proceedings in computational Statistics*, (Phisica-Verlag, Vienna) 97-101.

HÄRDLE, W. (1990): *Applied Nonparametric Regression* (Cambridge University Press, Cambridge).

HARDY, G. H.; LITTLEWOOD, J. E., and PÓLYA, G. (1934): *Inequalities* (Cambridge University Press, London).

HILDENBRAND, K., and HILDENBRAND, W. (1986): "On the mean income effect: a data analysis of the UK Family Expenditure Survey", in: W. HILDENBRAND, and A. MAS-COLELL (eds.), *Contributions to Mathematical Economics* (North-Holland, Amsterdam).

IRELAND, C. T., and KULLBACK, S. (1968): "Contingency tables with given marginals", *Biometrika* 55, 179-188.

JENKINS, S. P. (1994): "Earnings discrimination measurement: a distributional approach", *Journal of Econometrics*, 61, 81-102.

KOLM, S. C. (1976): "Unequal inequality: I and II", *Journal of Economic Theory* 12, 416-442 and 13, 88-111.

MARSHALL, A. W., and OLKIN, I. (1979): *Inequalities: Theory of majorization and its applications* (Academic Press, New York).

NADARAYA, E. A. (1964): "On estimating regression", *Theory Probability Applied* 10,186-190.

NIELSEN, M. A. (2002): *An introduction to majorization and its applications to quantum mechanics*, Department of Physics, Queensland University (Australia).

PRIESTLEY, M. B., and CHAO, M. T. (1972): "Nonparametric function fitting", *Journal of the Royal Statistical Society*, Series B 34, 385-392.

RODRÍGUEZ, J. G.; SALAS, R., and PERROTE, I. (2004): "Partial horizontal inequity orderings: a nonparametric approach", *Oxford Bulletin of Economics and Statistics*, forthcoming. Electronic version at http://www2.fundacion-centra.org/pdfs/E200401.pdf.

ROTHSCHILD, M., and STIGLITZ, J. (1970): "Increasing risk I: A definition", *Journal of Economic Theory* 2, 225-243.

– (1973): "Some further results on the measurement of inequality", *Journal of Economic Theory* 6, 188-204.

SHORROCKS, A. F. (1983): "Ranking income distributions", *Economica* 50, 3-17.

STONE, C. J. (1977): "Consistent nonparametric regression", *Annals of Statistics* 5, 595-620.

TUKEY, J. W. (1947): "Non-parametric estimation II. Statistically equivalent blocks and tolerance regions. The continuous case", *Annals of Mathematical Statistics* 18, 529-539.

VON NEUMANN, J. (1953): "A certain-zero-sum two-person game equivalent to the optimal assignment problem", in: H. KUHN, and A. TUCKER (eds.), *Contribution to the Theory of Games II* (Ann. Math. Study No. 28), (Princeton University Press) 5-12.

WAND, M. P., and JONES, M. C. (1995): *Kernel Smoothing*, Chapman and Hall.

WATSON, G. S. (1964): "Smooth regression analysis", *Sankhya Series A 26*, 101-16.

YITZHAKI, S. (1983): "On an Extension of the Gini Inequality Index", *International Economic Revie*w, 24, 617-628.

NORMAS DE PUBLICACIÓN DE PAPELES DE TRABAJO DEL INSTITUTO DE ESTUDIOS FISCALES

Esta colección de *Papeles de Trabajo* tiene como objetivo ofrecer un vehículo de expresión a todas aquellas personas interasadas en los temas de Economía Pública. Las normas para la presentación y selección de originales son las siguientes:

1. Todos los originales que se presenten estarán sometidos a evaluación y podrán ser directamente aceptados para su publicación, aceptados sujetos a revisión, o rechazados.

2. Los trabajos deberán enviarse por duplicado a la Subdirección de Estudios Tributarios. Instituto de Estudios Fiscales. Avda. Cardenal Herrera Oria, 378. 28035 Madrid.

3. La extensión máxima de texto escrito, incluidos apéndices y referencias bibliográfícas será de 7000 palabras.

4. Los originales deberán presentarse mecanografiados a doble espacio. En la primera página deberá aparecer el título del trabajo, el nombre del autor(es) y la institución a la que pertenece, así como su dirección postal y electrónica. Además, en la primera página aparecerá también un abstract de no más de 125 palabras, los códigos JEL y las palabras clave.

5. Los epígrafes irán numerados secuencialmente siguiendo la numeración arábiga. Las notas al texto irán numeradas correlativamente y aparecerán al pie de la correspondiente página. Las fórmulas matemáticas se numerarán secuencialmente ajustadas al margen derecho de las mismas. La bibliografía aparecerá al final del trabajo, bajo la inscripción “Referencias” por orden alfabético de autores y, en cada una, ajustándose al siguiente orden: autor(es), año de publicación (distinguiendo a, b, c si hay varias correspondientes al mismo autor(es) y año), título del artículo o libro, título de la revista en cursiva, número de la revista y páginas.

6. En caso de que aparezcan tablas y gráficos, éstos podrán incorporarse directamente al texto o, alternativamente, presentarse todos juntos y debidamente numerados al final del trabajo, antes de la bibliografía.

7. En cualquier caso, se deberá adjuntar un disquete con el trabajo en formato word. Siempre que el documento presente tablas y/o gráficos, éstos deberán aparecer en ficheros independientes. Asimismo, en caso de que los gráficos procedan de tablas creadas en excel, estas deberán incorporarse en el disquete debidamente identificadas.

Junto al original del Papel de Trabajo se entregará también un resumen de un máximo de dos folios que contenga las principales implicaciones de política económica que se deriven de la investigación realizada.

PUBLISHING GUIDELINES OF WORKING PAPERS AT THE **INSTITUTE FOR FISCAL STUDIES**

This serie of *Papeles de Trabajo* (working papers) aims to provide those having an interest in Public Economics with a vehicle to publicize their ideas. The rules governing submission and selection of papers are the following:

1. The manuscripts submitted will all be assessed and may be directly accepted for publication, accepted with subjections for revision or rejected.

2. The papers shall be sent in duplicate to Subdirección General de Estudios Tributarios (The Deputy Direction of Tax Studies), Instituto de Estudios Fiscales (Institute for Fiscal Studies), Avenida del Cardenal Herrera Oria, nº 378, Madrid 28035.

3. The maximum length of the text including appendices and bibliography will be no more than 7000 words.

4. The originals should be double spaced. The first page of the manuscript should contain the following information: (1) the title; (2) the name and the institutional affiliation of the author(s); (3) an abstract of no more than 125 words; (4) JEL codes and keywords; (5) the postal and e-mail address of the corresponding author.

5. Sections will be numbered in sequence with arabic numerals. Footnotes will be numbered correlatively and will appear at the foot of the corresponding page. Mathematical formulae will be numbered on the right margin of the page in sequence. Bibliographical references will appear at the end of the paper under the heading "References" in alphabetical order of authors. Each reference will have to include in this order the following terms of references: author(s), publishing date (with an a, b or c in case there are several references to the same author(s) and year), title of the article or book, name of the journal in italics, number of the issue and pages.

6. If tables and graphs are necessary, they may be included directly in the text or alternatively presented altogether and duly numbered at the end of the paper, before the bibliography.

7. In any case, a floppy disk will be enclosed in Word format. Whenever the document provides tables and/or graphs, they must be contained in separate files. Furthermore, if graphs are drawn from tables within the Excell package, these must be included in the floppy disk and duly identified.

Together with the original copy of the working paper a brief two-page summary highlighting the main policy implications derived from the research is also requested.

ÚLTIMOS PAPELES DE TRABAJO EDITADOS POR EL INSTITUTO DE ESTUDIOS FISCALES

2000

1/00 Crédito fiscal a la inversión en el impuesto de sociedades y neutralidad impositiva: Más evidencia para un viejo debate.
Autor: Desiderio Romero Jordán.
Páginas: 40.

2/00 Estudio del consumo familiar de bienes y servicios públicos a partir de la encuesta de presupuestos familiares.
Autores: Ernesto Carrilllo y Manuel Tamayo.
Páginas: 40.

3/00 Evidencia empírica de la convergencia real.
Autores: Lorenzo Escot y Miguel Ángel Galindo.
Páginas: 58.

Nueva Época

4/00 The effects of human capital depreciation on experience-earnings profiles: Evidence salaried spanish men.
Autores: M. Arrazola, J. de Hevia, M. Risueño y J. F. Sanz.
Páginas: 24.

5/00 Las ayudas fiscales a la adquisición de inmuebles residenciales en la nueva Ley del IRPF: Un análisis comparado a través del concepto de coste de uso.
Autor: José Félix Sanz Sanz.
Páginas: 44.

6/00 Las medidas fiscales de estímulo del ahorro contenidas en el Real Decreto-Ley 3/2000: análisis de sus efectos a través del tipo marginal efectivo.
Autores: José Manuel González Páramo y Nuria Badenes Plá.
Páginas: 28.

7/00 Análisis de las ganancias de bienestar asociadas a los efectos de la Reforma del IRPF sobre la oferta laboral de la familia española.
Autores: Juan Prieto Rodríguez y Santiago Álvarez García.
Páginas 32.

8/00 Un marco para la discusión de los efectos de la política impositiva sobre los precios y el *stock* de vivienda.
Autor: Miguel Ángel López García.
Páginas 36.

9/00 Descomposición de los efectos redistributivos de la Reforma del IRPF.
Autores: Jorge Onrubia Fernández y María del Carmen Rodado Ruiz.
Páginas 24.

10/00 Aspectos teóricos de la convergencia real, integración y política fiscal.
Autores: Lorenzo Escot y Miguel Ángel Galindo.
Páginas 28.

2001

1/01 Notas sobre desagregación temporal de series económicas.
Autor: Enrique M. Quilis.
Páginas 38.

2/01 Estimación y comparación de tasas de rendimiento de la educación en España.
Autores: M. Arrazola, J. de Hevia, M. Risueño y J. F. Sanz.
Páginas 28.

3/01 Doble imposición, "efecto clientela" y aversión al riesgo.
Autores: Antonio Bustos Gisbert y Francisco Pedraja Chaparro.
Páginas 34.

4/01 Non-Institutional Federalism in Spain.
Autor: Joan Rosselló Villalonga.
Páginas 32.

5/01 Estimating utilisation of Health care: A groupe data regression approach.
Autora: Mabel Amaya Amaya.
Páginas 30.

6/01 Shapley inequality descomposition by factor components.
Autores: Mercedes Sastre y Alain Trannoy.
Páginas 40.

7/01 An empirical analysis of the demand for physician services across the European Union.
Autores: Sergi Jiménez Martín, José M. Labeaga y Maite Martínez-Granado.
Páginas 40.

8/01 Demand, childbirth and the costs of babies: evidence from spanish panel data.
Autores: José M.ª Labeaga, Ian Preston y Juan A. Sanchis-Llopis.
Páginas 56.

9/01 Imposición marginal efectiva sobre el factor trabajo: Breve nota metodológica y comparación internacional.
Autores: Desiderio Romero Jordán y José Félix Sanz Sanz.
Páginas 40.

10/01 A non-parametric decomposition of redistribution into vertical and horizontal components.
Autores: Irene Perrote, Juan Gabriel Rodríguez y Rafael Salas.
Páginas 28.

11/01 Efectos sobre la renta disponible y el bienestar de la deducción por rentas ganadas en el IRPF.
Autora: Nuria Badenes Plá.
Páginas 28.

12/01 Seguros sanitarios y gasto público en España. Un modelo de microsimulación para las políticas de gastos fiscales en sanidad.
Autor: Ángel López Nicolás.
Páginas 40.

13/01 A complete parametrical class of redistribution and progressivity measures.
Autores: Isabel Rabadán y Rafael Salas.
Páginas 20.

14/01 La medición de la desigualdad económica.
Autor: Rafael Salas.
Páginas 40.

15/01 Crecimiento económico y dinámica de distribución de la renta en las regiones de la UE: un análisis no paramétrico.
Autores: Julián Ramajo Hernández y María del Mar Salinas Jiménez.
Páginas 32.

16/01 La descentralización territorial de las prestaciones asistenciales: efectos sobre la igualdad.
Autores: Luis Ayala Cañón, Rosa Martínez López y Jesus Ruiz-Huerta.
Páginas 48.

17/01 Redistribution and labour supply.
Autores: Jorge Onrubia, Rafael Salas y José Félix Sanz.
Páginas 24.

18/01 Medición de la eficiencia técnica en la economía española: El papel de las infraestructuras productivas.
Autoras: M.ª Jesús Delgado Rodríguez e Inmaculada Álvarez Ayuso.
Páginas 32.

19/01 Inversión pública eficiente e impuestos distorsionantes en un contexto de equilibrio general.
Autores: José Manuel González-Páramo y Diego Martínez López.
Páginas 28.

20/01 La incidencia distributiva del gasto público social. Análisis general y tratamiento específico de la incidencia distributiva entre grupos sociales y entre grupos de edad.
Autor: Jorge Calero Martínez.
Páginas 36.

21/01 Crisis cambiarias: Teoría y evidencia.
Autor: Óscar Bajo Rubio.
Páginas 32.

22/01 Distributive impact and evaluation of devolution proposals in Japanese local public finance.
Autores: Kazuyuki Nakamura, Minoru Kunizaki y Masanori Tahira.
Páginas 36.

23/01 El funcionamiento de los sistemas de garantía en el modelo de financiación autonómica.
Autor: Alfonso Utrilla de la Hoz.
Páginas 48.

24/01 Rendimiento de la educación en España: Nueva evidencia de las diferencias entre Hombres y Mujeres.
Autores: M. Arrazola y J. de Hevia.
Páginas 36.

25/01 Fecundidad y beneficios fiscales y sociales por descendientes.
Autora: Anabel Zárate Marco.
Páginas 52.

26/01 Estimación de precios sombra a partir del análisis Input-Output: Aplicación a la economía española.
Autora: Guadalupe Souto Nieves.
Páginas 56.

27/01 Análisis empírico de la depreciación del capital humano para el caso de las Mujeres y los Hombres en España.
Autores: M. Arrazola y J. de Hevia.
Páginas 28.

28/01 Equivalence scales in tax and transfer policies.
Autores: Luis Ayala, Rosa Martínez y Jesús Ruiz-Huerta.
Páginas 44.

29/01 Un modelo de crecimiento con restricciones de demanda: el gasto público como amortiguador del desequilibrio externo.
Autora: Belén Fernández Castro.
Páginas 44.

30/01 A bi-stochastic nonparametric estimator.
Autores: Juan G. Rodríguez y Rafael Salas.
Páginas 24.

2002

1/02 Las cestas autonómicas.
Autores: Alejandro Esteller, Jorge Navas y Pilar Sorribas.
Páginas 72.

2/02 Evolución del endeudamiento autonómico entre 1985 y 1997: la incidencia de los Escenarios de Consolidación Presupuestaria y de los límites de la LOFCA.
Autores: Julio López Laborda y Jaime Vallés Giménez.
Páginas 60.

3/02 Optimal Pricing and Grant Policies for Museums.
Autores: Juan Prieto Rodríguez y Víctor Fernández Blanco.
Páginas 28.

4/02 El mercado financiero y el racionamiento del endeudamiento autonómico.
Autores: Nuria Alcalde Fradejas y Jaime Vallés Giménez.
Páginas 36.

5/02 Experimentos secuenciales en la gestión de los recursos comunes.
Autores: Lluis Bru, Susana Cabrera, C. Mónica Capra y Rosario Gómez.
Páginas 32.

6/02 La eficiencia de la universidad medida a través de la función de distancia: Un análisis de las relaciones entre la docencia y la investigación.
Autores: Alfredo Moreno Sáez y David Trillo del Pozo.
Páginas 40.

7/02 Movilidad social y desigualdad económica.
Autores: Juan Prieto-Rodríguez, Rafael Salas y Santiago Álvarez-García.
Páginas 32.

8/02 Modelos BVAR: Especificación, estimación e inferencia.
Autor: Enrique M. Quilis.
Páginas 44.

9/02 Imposición lineal sobre la renta y equivalencia distributiva: Un ejercicio de microsimulación.
Autores: Juan Manuel Castañer Carrasco y José Félix Sanz Sanz.
Páginas 44.

10/02 The evolution of income inequality in the European Union during the period 1993-1996.
Autores: Santiago Álvarez García, Juan Prieto-Rodríguez y Rafael Salas.
Páginas 36.

11/02 Una descomposición de la redistribución en sus componentes vertical y horizontal: Una aplicación al IRPF.
Autora: Irene Perrote.
Páginas 32.

12/02 Análisis de las políticas públicas de fomento de la innovación tecnológica en las regiones españolas.
Autor: Antonio Fonfría Mesa.
Páginas 40.

13/02 Los efectos de la política fiscal sobre el consumo privado: nueva evidencia para el caso español.
Autores: Agustín García y Julián Ramajo.
Páginas 52.

14/02 Micro-modelling of retirement behavior in Spain.
Autores: Michele Boldrin, Sergi Jiménez-Martín y Franco Peracchi.
Páginas 96.

15/02 Estado de salud y participación laboral de las personas mayores.
Autores: Juan Prieto Rodríguez, Desiderio Romero Jordán y Santiago Álvarez García.
Páginas 40.

16/02 Technological change, efficiency gains and capital accumulation in labour productivity growth and convergence: an application to the Spanish regions.
Autora: M.ª del Mar Salinas Jiménez.
Páginas 40.

17/02 Déficit público, masa monetaria e inflación. Evidencia empírica en la Unión Europea.
Autor: César Pérez López.
Páginas 40.

18/02 Tax evasion and relative contribution.
Autora: Judith Panadés i Martí.
Páginas 28.

19/02 Fiscal policy and growth revisited: the case of the Spanish regions.
Autores: Óscar Bajo Rubio, Carmen Díaz Roldán y M. ª Dolores Montávez Garcés.
Páginas 28.

20/02 Optimal endowments of public investment: an empirical analysis for the Spanish regions.
Autores: Óscar Bajo Rubio, Carmen Díaz Roldán y M.ª Dolores Montávez Garcés.
Páginas 28.

21/02 Régimen fiscal de la previsión social empresarial. Incentivos existentes y equidad del sistema.
Autor: Félix Domínguez Barrero.
Páginas 52.

22/02 Poverty statics and dynamics: does the accounting period matter?
Autores: Olga Cantó, Coral del Río y Carlos Gradín.
Páginas 52.

23/02 Public employment and redistribution in Spain.
Autores: José Manuel Marqués Sevillano y Joan Rosselló Villallonga.
Páginas 36.

24/02 La evolución de la pobreza estática y dinámica en España en el periodo 1985-1995.
Autores: Olga Cantó, Coral del Río y Carlos Gradín.
Páginas: 76.

25/02 Estimación de los efectos de un "tratamiento": una aplicación a la Educación superior en España.
Autores: M. Arrazola y J. de Hevia.
Páginas 32.

26/02 Sensibilidad de las estimaciones del rendimiento de la educación a la elección de instrumentos y de forma funcional.
Autores: M. Arrazola y J. de Hevia.
Páginas 40.

27/02 Reforma fiscal verde y doble dividendo. Una revisión de la evidencia empírica.
Autor: Miguel Enrique Rodríguez Méndez.
Páginas 40.

28/02 Productividad y eficiencia en la gestión pública del transporte de ferrocarriles implicaciones de política económica.
Autor: Marcelino Martínez Cabrera.
Páginas 32.

29/02 Building stronger national movie industries: The case of Spain.
Autores: Víctor Fernández Blanco y Juan Prieto Rodríguez.
Páginas 52.

30/02 Análisis comparativo del gravamen efectivo sobre la renta empresarial entre países y activos en el contexto de la Unión Europea (2001).
Autora: Raquel Paredes Gómez.
Páginas 48.

31/02 Voting over taxes with endogenous altruism.
Autor: Joan Esteban.
Páginas 32.

32/02 Midiendo el coste marginal en bienestar de una reforma impositiva.
Autor: José Manuel González-Páramo.
Páginas 48.

33/02 Redistributive taxation with endogenous sentiments.
Autores: Joan Esteban y Laurence Kranich.
Páginas 40.

34/02 Una nota sobre la compensación de incentivos a la adquisición de vivienda habitual tras la reforma del IRPF de 1998.
Autores: Jorge Onrubia Fernández, Desiderio Romero Jordán y José Félix Sanz Sanz.
Páginas 36.

35/02 Simulación de políticas económicas: los modelos de equilibrio general aplicado.
Autor: Antonio Gómez Gómez-Plana.
Páginas 36.

2003

1/03 Análisis de la distribución de la renta a partir de funciones de cuantiles: robustez y sensibilidad de los resultados frente a escalas de equivalencia.
Autores: Marta Pascual Sáez y José María Sarabia Alegría.
Páginas 52.

2/03 Macroeconomic conditions, institutional factors and demographic structure: What causes welfare caseloads?
Autores: Luis Ayala y César Pérez.
Páginas 44.

3/03 Endeudamiento local y restricciones institucionales. De la ley reguladora de haciendas locales a la estabilidad presupuestaria.
Autores: Jaime Vallés Giménez, Pedro Pascual Arzoz y Fermín Cabasés Hita.
Páginas 56.

4/03 The dual tax as a flat tax with a surtax on labour income.
Autor: José María Durán Cabré.
Páginas 40.

5/03 La estimación de la función de producción educativa en valor añadido mediante redes neuronales: una aplicación para el caso español.
Autor: Daniel Santín González.
Páginas 52.

6/03 Privación relativa, imposición sobre la renta e índice de Gini generalizado.
Autores: Elena Bárcena Martín, Luis Imedio Olmedo y Guillermina Martín Reyes.
Páginas 36.

7/03 Fijación de precios óptimos en el sector público: una aplicación para el servicio municipal de agua.
Autora: M.ª Ángeles García Valiñas.
Páginas 44.

8/03 Tasas de descuento para la evaluación de inversiones públicas: Estimaciones para España.
Autora: Guadalupe Souto Nieves.
Páginas 40.

9/03 Una evaluación del grado de incumplimiento fiscal para las provincias españolas.
Autores: Ángel Alañón Pardo y Miguel Gómez de Antonio.
Páginas 44.

10/03 Extended bi-polarization and inequality measures.
Autores: Juan G. Rodríguez y Rafael Salas.
Páginas 32.

11/03 Fiscal decentralization, macrostability and growth.
Autores: Jorge Martínez-Vázquez y Robert M. McNab.
Páginas 44.

12/03 Valoración de bienes públicos en relación al patrimonio histórico cultural: aplicación comparada de métodos estadísticos de estimación.
Autores: Luis César Herrero Prieto, José Ángel Sanz Lara y Ana María Bedate Centeno.
Páginas 44.

13/03 Growth, convergence and public investment. A bayesian model averaging approach.
Autores: Roberto León-González y Daniel Montolio.
Páginas 44.

14/03 ¿Qué puede esperarse de una reducción de la imposición indirecta que recae sobre el consumo cultural?: Un análisis a partir de las técnicas de microsimulación.
Autores: José Félix Sanz Sanz, Desiderio Romero Jordán y Juan Prieto Rodríguez.
Páginas 40.

15/03 Estimaciones de la tasa de paro de equilibrio de la economía española a partir de la Ley de Okun.
Autores: Inés P. Murillo y Carlos Usabiaga.
Páginas 32.

16/03 La previsión social en la empresa, tras la Ley 46/2002, de reforma parcial del impuesto sobre la renta de las personas físicas.
Autor: Félix Domínguez Barrero.
Páginas 48.

17/03 The influence of previous labour market experiences on subsequent job tenure.
Autores: José María Arranz y Carlos García-Serrano.
Páginas 48.

18/03 Promoting sutdent's effort: standards *versus* torunaments.
Autores: Pedro Landeras y J. M. Pérez de Villarreal.
Páginas 44.

19/03 Non-employment and subsequent wage losses.
Autores: José María Arranz y Carlos García-Serrano.
Páginas 52.

20/03 La medida de los ingresos públicos en la Agencia Tributaria. Caja, derechos reconocidos y devengo económico.
Autores: Rafael Frutos, Francisco Melis, M.ª Jesús Pérez de la Ossa y José Luis Ramos.
Páginas 80.

21/03 Tratamiento fiscal de la vivienda y exceso de gravamen.
Autor: Miguel Ángel López García.
Páginas 44.

22/03 Medición del capital humano y análisis de su rendimiento.
Autores: María Arrazola y José de Hevia.
Páginas 36.

23/03 Vivienda, reforma impositiva y coste en bienestar.
Autor: Miguel Ángel López García.
Páginas 52.

24/03 Algunos comentarios sobre la medición del capital humano.
Autores: María Arrazola y José de Hevia.
Páginas 40.

25/03 Exploring the spanish interbank yield curve.
Autores: Leandro Navarro y Enrique M. Quilis.
Páginas 32.

26/03 Redes neuronales y medición de eficiencia: aplicación al servicio de recogida de basuras.
Autor: Francisco J. Delgado Rivero.
Páginas 60.

27/03 Equivalencia ricardiana y tipos de interés.
Autoesr: Agustín García, Julián Ramajo e Inés Piedraescrita Murillo.
Páginas 40.

28/03 Instrumentos y objetivos de las políticas de apoyo a las PYME en España.
Autor: Antonio Fonfría Mesa.
Páginas 44.

29/03 Análisis de incidencia del gasto público en educación superior: enfoque transversal.
Autora: María Gil Izquierdo.
Páginas 48.

30/03 Rentabilidad social de la inversión pública española en infraestructuras.
Autores: Jaime Alonso-Carrera, María Jesús Freire-Serén y Baltasar Manzano.
Páginas 44.

31/03 Las rentas de capital en Phogue: análisis de su fiabilidad y corrección mediante fusión estadística.
Autor: Fidel Picos Sánchez.
Páginas 44.

32/03 Efecto de los sistemas de rentas mínimas autonómicas sobre la migración interregional.
Autora: María Martínez Torres.
Páginas 44.

33/03 Rentas mínimas autonómicas en España. Su dimensión espacial.
Autora: María Martínez Torres.
Páginas 76.

34/03 Un nuevo examen de las causas del déficit autonómico.
Autor: Santiago Lago Peñas.
Páginas 52.

35/03 Uncertainty and taxpayer compliance.
Autores: Jordi Caballé y Judith Panadés.
Páginas 44.

2004

1/04 Una propuesta para la regulación de precios en el sector del agua: el caso español.
Autores: M.ª Ángeles García Valiñas y Manuel Antonio Muñiz Pérez.
Páginas 40.

2/04 Eficiencia en educación secundaria e *inputs* no controlables: sensibilidad de los resultados ante modelos alternativos.
Autores: José Manuel Cordero Ferrera, Francisco Pedraja Chaparro y Javier Salinas Jiménez.
Páginas 40.

3/04 Los efectos de la política fiscal sobre el ahorro privado: evidencia para la OCDE.
Autores: Montserrat Ferre Carracedo, Agustín García García y Julián Ramajo Hernández.
Páginas 44.

4/04 ¿Qué ha sucedido con la estabilidad del empleo en España? Un análisis desagregado con datos de la EPA: 1987-2003.
Autores: José María Arranz y Carlos García-Serrano.
Páginas 80.

5/04 La seguridad del empleo en España: evidencia con datos de la EPA (1987-2003).
Autores: José María Arranz y Carlos García-Serrano.
Páginas 72.

6/04 La ley de Wagner: un análisis sintético.
Autor: Manuel Jaén García.
Páginas 60.

7/04 La vivienda y la reforma fiscal de 1998: un ejercicio de simulación.
Autor: Miguel Ángel López García.
Páginas 44.

8/04 Modelo dual de IRPF y equidad: un nuevo enfoque teórico y su aplicación al caso español.
Autor: Fidel Picos Sánchez.
Páginas 44.

9/04 Public expenditure dynamics in Spain: a simplified model of its determinants.
Autores: Manuel Jaén García y Luis Palma Martos.
Páginas 48.

10/04 Simulación sobre los hogares españoles de la reforma del IRPF de 2003. Efectos sobre la oferta laboral, recaudación, distribución y bienestar.
Autores: Juan Manuel Castañer Carrasco, Desiderio Romero Jordán y José Félix Sanz Sanz.
Páginas 56.

11/04 Financiación de las Haciendas regionales españolas y experiencia comparada.
Autor: David Cantarero Prieto.
Páginas 52.

12/04 Multidimensional indices of housing deprivation with application to Spain.
Autores: Luis Ayala y Carolina Navarro.
Páginas 44.

13/04 Multiple ocurrence of welfare recipiency: determinants and policy implications.
Autores: Luis Ayala y Magdalena Rodríguez.
Páginas 52.

14/04 Imposición efectiva sobre las rentas laborales en la reforma del impuesto sobre la renta personal (IRPF) de 2003 en España.
Autoras: María Pazos Morán y Teresa Pérez Barrasa.
Páginas 40.

15/04 Factores determinantes de la distribución personal de la renta: un estudio empírico a partir del PHOGUE.
Autores: Marta Pascual y José María Sarabia.
Páginas 56.

16/04 Política familiar, imposición efectiva e incentivos al trabajo en la reforma de la imposición sobre la renta personal (IRPF) de 2003 en España.
Autoras: María Pazos Morán y Teresa Pérez Barrasa.
Páginas 48.

17/04 Efectos del déficit público: evidencia empírica mediante un modelo de panel dinámico para los países de la Unión Europea.
Autor: César Pérez López.
Páginas 40.

18/04 Inequality, poverty and mobility: Choosing income or consumption as welfare indicators.
Autores: Carlos Gradín, Olga Cantó y Coral del Río.
Páginas 52.

19/04 Tendencias internacionales en la financiación del gasto sanitario.
Autora: Rosa María Urbanos Garrido.
Páginas 48.

20/04 El ejercicio de la capacidad normativa de las CCAA en los tributos cedidos: una primera evaluación a través de los tipos impositivos efectivos en el IRPF.
Autores: José María Durán y Alejandro Esteller.
Páginas 68.

21/04 Explaining. budgetary indiscipline: evidence from spanish municipalities.
Autores: Ignacio Lago-Peñas y Santiago Lago-Peñas.
Páginas 44.

22/04 Local governmets' asymmetric reactions to grants: looking for the reasons.
Autor: Santiago Lago-Peñas.
Páginas 40.

23/04 Un pacto de estabilidad para el control del endeudamiento autonómico.
Autor: Roberto Fernández Llera
Páginas 48.

24/04 Una medida de la calidad del producto de la atención primaria aplicable a los análisis DEA de eficiencia.
Autora: Mariola Pinillos García.
Páginas 40.

25/04 Distribución de la renta, crecimiento y política fiscal.
Autor: Miguel Ángel Galindo Martín.
Páginas 40.

26/04 Políticas de inspección óptimas y cumplimiento fiscal.
Autores: Inés Macho Stadler y David Pérez Castrillo.
Páginas 60.

27/04 ¿Por qué ahorra la gente en planes de pensiones individuales?
Autores: Félix Domínguez Barrero y Julio López-Laborda.
Páginas 48.

28/04 La reforma del Impuesto sobre Actividades Económicas: una valoración con microdatos de la ciudad de Zaragoza.
Autores: Julio López-Laborda, M.ª Carmen Trueba Cortés y Anabel Zárate Marco.
Páginas 56.

29/04 Is an inequality-neutral flat tax reform really neutral?
Autores: Juan Prieto-Rodríguez, Juan Gabriel Rodríguez y Rafael Salas.
Páginas 40.

30/04 El equilibrio presupuestario: las restricciones sobre el déficit.
Autora: Belén Fernández Castro.
Páginas 44.

2005

1/05 Efectividad de la política de cooperación en innovación: evidencia empírica española.
Autores: Joost Heijs, Liliana Herrera, Mikel Buesa, Javier Sáiz Briones y Patrica Valadez.
Páginas 52.

2/05 A probabilistic nonparametric estimator.
Autores: Juan Gabriel Rodríguez y Rafael Salas.
Páginas 48.