TURISMO ANDALUZ

JUNTA DE ANDALUCIA
Consejería de Relaciones Institucionales

centrA: Fundación Centro de Estudios Andaluces

Documento de trabajo
E2004/22

A Bistochastic Nonparametric Estimator

Juan Gabriel Rodríguez
Rafael Salas

JUNTA DE ANDALUCIA
Consejería de Relaciones Institucionales

centrA: Fundación Centro de Estudios Andaluces

Fundación Centro de Estudios Andaluces (**centrA**)
Bailén, 50 - 41001 Sevilla

Tel: 955 055 210, Fax: 955 055 211

e-mail: centra@fundacion-centra.org
http://www.fundacion-centra.org

DEPÓSITO LEGAL: SE-108-2002

centrA:

Fundación Centro de Estudios Andaluces

Documento de Trabajo
Serie Economía E2004/22

A Bistochastic Nonparametric Estimator

Juan Gabriel Rodríguez*
Univ. Rey Juan Carlos de Madrid
e Instituto de Estudios Fiscales

Rafael Salas
Univ. Complutense de Madrid
e Instituto de Estudios Fiscales

RESUMEN
Este trabajo propone un estimador no paramétrico, que reduce la variabilidad del suavizado de manera robusta. Proponemos un estimador que satisfaga un criterio tan ampliamente asentado en la literatura del riesgo y del bienestar como el de la dominancia estocástica de segundo orden. Además, el estimador biestocástico consigue errores menores, en términos de la optimización entre el sesgo y la varianza, que el resto de estimadores no paramétricos que utilizan pesos positivos. A este resultado se llega por medio de un ejercicio de simulación. La mejora se debe a la reducción significativa del sesgo en las colas de la distribución. Por último, la consistencia, el mantenimiento de la media, y su extensión potencial a un marco multidimensional son algunas otras propiedades del estimador.

Palabras clave: estimación no paramétrica, dominancia estocástica de segundo orden, estimador biestocástico.

JEL classificación: D63, C14.

ABSTRACT
We explore the relevance of adopting a bistochastic nonparametric estimator. This estimator has two main implications. First, the estimator reduces variability according to the robust criterion of second-order stochastic (and Lorenz) dominance. This is a universally criterion in risk and welfare economics, which expands the applicability of nonparametric estimation in economics, for instance to the measurement of economic discrimination. Second, the bistochastic estimator produces smaller errors than do positive-weights nonparametric estimators, in terms of the bias-variance trade-off. This result is verified in a general simulation exercise. This improvement is due to a significant reduction in boundary bias, which makes the estimator itself useful in empirical applications. Finally, consistency, preservation of the mean value, and multidimensional extension are some other useful properties of this estimator.

Keywords: nonparametric estimation, second-order stochastic dominance, bistochastic estimator.

JEL classification: C68, D58, R13, R15.

---

* This paper has benefited from the support of the European Commission under Project #ERBCHRXCT980248, from the Spanish Ministry of Education under Project #PB98-0546-C0202, from the Spanish Ministry of Science and Technology #SEC2003-08397, and from Fundación BBVA. We are grateful for helpful comments by J.M. Rodríguez Poo, E. Ferreira, V. Dardanoni, M.A. Delgado, P. Lambert, I. Perrote and S. Yitzhaki.

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. 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 examined nonparametrically total factor productivity differences between exporting and no exporting firms. They rank distributions of firms using the concept of stochastic dominance.

However, we argue that improvements to standard nonparametric techniques are possible.[1] We propose a *bistochastic* nonparametric smoothing technique with the following two basic properties. First, the estimator unambiguously reduces overall variability in a robust sense, according to the Lorenz criterion (widely used in welfare and risk economics). We suggest that the variability reduction due to nonparametric techniques should be consistent with a wide class of dispersion measures that satisfy the mean-preserving second-order stochastic dominance criterion.[2]

The second property is that the estimator performs better than alternative standard nonparametric methods that use positive weights, on the basis of the bias–variance trade-off. We concentrate on nonparametric methods that use positive weights and have a clear economic interpretation. For instance, if we try to estimate an Engel curve, negative weights are either difficult to interpret or contribute to generating implausible (negative) consumption values.[3] This important property is related to the improvement in boundary bias that is usually associated with positive-weights-based nonparametric estimators. We show below how these two properties are linked.

Moreover, the bistochastic estimator is obtained from a simple low-cost modification of existing nonparametric techniques (such as the Nadaraya–Watson estimator). There is a relationship between the traditional estimators and the proposed class of estimators.

Standard nonparametric smoothers are typically *stochastic* because they only take into account a single normalization of the weights matrix (that is, weights sum to unity across rows). However, because we are dealing with a two-dimensional weights matrix,[4] we propose an estimator with a double normalization (in which weights sum to unity across rows and columns).[5] As a result, we obtain the bistochastic estimator. Note that under this methodology asymptotic unbiasedness is retained, as row normalization is verified.

Furthermore, in the bistochastic case, observations are treated symmetrically in the sense that they all have the same weight in the process of the construction of the nonparametric smoother. The intervals are 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. Since the new smoother gives greater weights to points near the boundary, the estimator can alleviate the so-called *boundary bias problem*. Simulations over a set of different distributions of the variable X confirm this result.

The proposed estimator has a number of other 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 bistochastic estimator can be generalized to the multidimensional regression case.

In welfare economics, the bistochastic nonparametric technique can be applied to the analysis of income distributions and economic discrimination. In fact, Rodríguez et al. (2003) propose the use of this bistochastic nonparametric technique to measure tax discrimination due to a fiscal system.[6] Furthermore, its usefulness can be generalized to other cases, such as the measurement of wage discrimination. It is indeed the underlying Lorenz smoothing property of the estimator that enhances the usefulness of the smoother in applied economics.

This paper is organized as follows. In the next section, a brief view of nonparametric smoothing is presented. In Section 3, the second-order stochastic dominance property is examined. Section 4 deals with bistochastic nonparametric estimation and its properties. In Section 5, the simulation exercises are presented. Finally, Section 6 concludes.

2. STANDARD NONPARAMETRIC SMOOTHING

Given any two-dimensional random sample, $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$,[7] the random variables, **X** and **Y**, denote vectors of the explanatory and response variables, respectively. The *theoretical* regression curve m(*x*) is defined as the expected value of Y at point *x*,

$$m(x) = E(Y \mid X = x)$$

The *nonparametrically estimated* regression curve at point *x*, M(*x*), can be written as a weighted average of the observations on Y,

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

where the weights W_j, which downweight the Y_js if the corresponding X_j value is far from x, are probabilistic.[8] For instance, W_j could be the Nadaraya–Watson weights (Nadaraya, 1964; Watson, 1964),

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

where kernel K is a continuous, bounded, and symmetric real function that integrates to unity (such as the normal density function, for example) 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) because these satisfy the probabilistic weights criterion. In particular, the Priestly–Chao and Gasser–Müller estimators have more severe boundary bias problems than does the Nadaraya–Watson smoother, and for random designs, also have 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 (which is confirmed by our simulation exercise below). However, we applied the simulation exercises used in Section 5 to the Priestly–Chao and Gasser–Müller estimators, and obtained similar results.[9] We do not consider k-

NN smoothing because this is equivalent to the kernel estimation when an appropriate bandwidth k parameter is used (see Härdle, 1990).

Henceforth, nonparametric estimation is written in vector notation, $\mathbf{M} = \mathbf{W} \cdot \mathbf{Y}$, where $\mathbf{W}$ is the weights matrix and $\mathbf{M}$ is the nonparametric smoother evaluated at any n points.

3. The Second-order Stochastic Dominance Property

Standard nonparametric methods are based on the idea of variance-reduction smoothing. Lorenz curve or (mean-preserving) second-order stochastic dominance is a broadly defined criterion for smoothing or dispersion reduction (Rothschild and Stiglitz, 1970 and Atkinson, 1970). This is a more general or robust criterion than that of variance reduction. Hence, there is a potential for improvement over the standard nonparametric approach, which comes from insisting that this robust Lorenz dominance criterion be satisfied by nonparametric smoothing.

Formally, we establish the Lorenz principle in terms of the distribution function.

DEFINITION 1: Given a distribution function, F(x), where $x \in [0,+\infty)$, we define the *second-order distribution function* as $F_2(x) = \int_0^x F(t)dt$, which is the definite integral of the distribution function, F(x).

DEFINITION 2: Given any two distribution functions with the same mean value, F(x) and G(x), where $x \in [0,+\infty)$, the distribution F(x) *mean-preserving second-order stochastically dominates* the distribution G(x) if and only if $F_2(x) \leq G_2(x) \ \forall \ x \in [0,+\infty)$.

Note that, more generally, given any two distribution functions, F(x) and G(x), where $x \in [0,+\infty)$, the distribution F(x) *second-order stochastically dominates* the distribution G(x) if and only if $F_2(x) \leq G_2(x) \ \forall \ x \in [0,+\infty)$.

Since the Lorenz curve $L_F(q)$ is the following transformation of the distribution function F(x),

$$L_F(q) = \frac{1}{\mu}\int_0^q x dF(x)$$

where $q \in [0,1]$ and μ is the mean of F(x), we can equivalently define *Lorenz dominance* of F over G as $L_F(q) \geq L_G(q)$, $\forall q$, given the same mean μ. Then, we obtain the well known result that mean-preserving second-order stochastic dominance is equivalent to Lorenz dominance; that is, $F_2(x) \leq G_2(x)$, $\forall x \in [0,+\infty)$ is equivalent to $L_F(q) \geq L_G(q)$, $\forall q \in [0,1]$.

Consider the following axiom.

AXIOM 1 (mean-preserving second-order stochastic or Lorenz dominance smoothing consistency): Suppose that the two distribution functions G(x) and F(x), which have the same mean value, correspond to the actual $\mathbf{Y}$ and estimated $\hat{\mathbf{Y}}$ distribution functions, respectively. If nonparametric smoothing unambiguously reduces the value of the second-degree distribution function, that is, if $F_2(x) \leq G_2(x)$, $\forall\ x \in [0,+\infty)$, then the smoothing technique is said to be consistent with second-order stochastic dominance.

The relevance of this axiom is related to the well known result stating equivalence between stochastic dominance and mean-preserving spreads. If the smoothing technique is consistent with second-order stochastic dominance, the estimated values can be obtained from the original values as a set of unambiguous Lorenz-variability-reducing mean-preserving spreads. In particular, they can be obtained as a set of variance-reducing mean-preserving spreads.

An important result that is used in what follows is due to Dasgupta et al. (1973).

DEFINITION 3: A bistochastic matrix is a square matrix in which all elements are nonegative and all row and column sums are equal to unity. All permutation matrices are bistochastic. In fact, the set of bistochastic matrices of order n is the convex hull of the set of permutation matrices of order n (see Berge, 1963).

PROPOSITION 1 (Dasgupta et al., 1973): Given two variables **X** and **Y** that have the same mean, **X** second-order stochastic dominates **Y** if and only if **X** can be written as $\mathbf{X} = \mathbf{V^B} \cdot \mathbf{Y}$, where $\mathbf{V^B}$ is a bistochastic matrix. Dasgupta et al. (1973) interpret variable **X** as a smoothing of variable **Y** in the sense that each income level of **X** is a convex combination of income levels in **Y**. Therefore, a nonparametric estimator is consistent with second-order stochastic dominance if and only if it is bistochastic.

4. THE BISTOCHASTIC NONPARAMETRIC ESTIMATOR

We define the bistochastic nonparametric estimator as follows.

DEFINITION 4: A nonparametric estimator, expressed in vector notation by $\mathbf{Z} = \mathbf{W^B} \cdot \mathbf{Y}$, is said to be *bistochastic* if and only if $\mathbf{W^B}$ is a bistochastic weights matrix, which is normalized by both rows and by columns; that is $\sum_{i=1}^{n} W_{ij} = 1$ and $\sum_{j=1}^{n} W_{ij} = 1$.

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

We propose a particular method of obtaining a bistochastic estimator.[10] Given a nonparametric estimator, denoted by $\mathbf{M} = \mathbf{W} \cdot \mathbf{Y}$, we propose the following low-cost method of obtaining a bistochastic smoother. Of all the potential methods that could be used for double normalization of the weights matrix, we adopt an iterative proportional-fitting method. This method is a special case of the algorithm proposed by Deming and Stephan (1940), which minimizes the Kullback–Liebler distance function.

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

$$\overline{W}_{ij}^{(0)} = W_{ij}.$$

If t is odd,

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

and if t is even,

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

where

$$\overline{W}_{+j}^{(t)} = \sum_{i=1}^{n} \overline{W}_{ij}^{(t)}$$

$\forall$ j = 1, . . ., n; t$\in$**N**

and

$$\overline{W}_{i+}^{(t)} = \sum_{j=1}^{n} \overline{W}_{ij}^{(t)}$$

$\forall$ i = 1, . . ., n; t$\in$**N**.

There is an interesting symmetry condition relating to this particular algorithm. The convergence result for this algorithm is $\mathbf{W}^{\mathbf{B}} = \{\overline{W}_{ij}\}_{1\leq i \leq n,\ 1\leq j \leq n}$, which applies whether we begin normalizing by rows or columns.[11] Hence, the bistochastic nonparametric smoother verifies Axiom 1 by construction.

How is the computational cost of this algorithm? In principle, we have to evaluate the curve at n points to obtain a square weight matrix as the sample size is n. Then, $O(n^2)$ kernel evaluations are necessary before the iterative method is applied to the stochastic nonparametric estimator to obtain the bistochastic one. This may make the computation of the bistochastic estimator very slow, for a sufficiently large value of n. One way of dramatically increasing the computational speed is to computed the *binned kernel regression estimator* (see for example Georgiev, 1986, and 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 using the fast Fourier transform. The main idea is to replace the data by a mesh of R grid counts, where each grid count is a weight which represents the amount of data near the corresponding grid point. The approximation is very good for moderate values of R and it can be made arbitrarily good 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. It follows that $\{\overline{W}_n\}$ (as defined above) is consistent.

Proof: The bistochastic sequence of weights can be written as

$$\overline{W}_n = f(W_n) \cdot W_n$$

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

In addition, the bistochastic estimator has the following properties.

PROPERTY 2 (Lorenz or mean-preserving second-order dominance consistency): The bistochastic smoothing technique is consistent with mean-preserving second-order stochastic dominance. That is, the Lorenz curve of Z always lies above the Lorenz curve of Y: $L_{F(Z)}(q) \geq L_{F(Y)}(q)$, $\forall q \in [0,1]$.

Proof. This proof applies the Proposition 1 of Dasgupta et al. (1973). A necessary and sufficient condition for $L_{F(Z)}(q) \geq L_{F(Y)}(q)$, $\forall q \in [0,1]$ is that $\mathbf{Z} = \mathbf{W^B} \cdot \mathbf{Y}$, where $\mathbf{W^B}$ is bistochastic. Therefore, the estimated values are obtained from the original values as a set of unambiguous Lorenz-variability-reducing mean-preserving spreads, if and only if the estimator is bistochastic.

Note that this result is even more general as it is also second-order stochastic dominance consistent. The reason is that the mean of the dependent variable remains constant (see property 4 below).

APPLICATION:

This property have potentially important empirical applications in the field of welfare economics. In particular, we find applications on the measurement of discrimination. In a recent paper (Rodríguez et al., 2003) the bistochastic estimator has been applied on the measurement of tax discrimination (horizontal inequity due to a tax system).

Let **X** and **Y** be the pre- and the post-tax equivalent income distributions, respectively, and **Z** be the estimated post-tax equivalent income by the nonparametric bistochastic technique. The higher the dispersion of **Y** around **Z,** the larger the tax discrimination is. In that paper, it is proved that dispersion of **Y** around **Z** exists if and only if similar individuals with close-equal pre-tax income levels (associated to a particular bandwidth h) pay different taxes and, therefore, 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 curve 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 or stochastic dominance criterion. The Lorenz smoothing property of the bistochastic nonparametric estimator ensures the dominance property and therefore, that the distance between both Lorenz curves be non-negative. Notice that negative discrimination values have no economic meaning.

Extensions to more general discrimination frameworks, such as wage discrimination, can be analogously done by computing the distance between the actual and the estimated wage Lorenz curves with the same properties. Both curves only differ whenever workers with similar productivity levels achieve different wages. In a context with imperfect information, we have to discount the dispersion due to incomplete information. This is the basics of the measurement of discrimination that we intend to develop in future research.

PROPERTY 3: The reformulated bistochastic estimator vector, to which this algorithm converges, is $\mathbf{Z} = \mathbf{W}^{\mathbf{B}} \cdot \mathbf{Y}$, where $\mathbf{W}^{\mathbf{B}}$ is the closest bistochastic matrix to **W**, according to the

Kullback–Liebler distance function. The proof applies the result of Ireland and Kullback (1968).

PROPERTY 4 (zero expected estimation error): An implicit property is that the estimator **Z** and the variable **Y** have the same mean, $\mu(\mathbf{Z}) = \mu(\mathbf{W^B} \cdot \mathbf{Y}) = \mu(\mathbf{Y})$, 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.

PROPERTY 5 (symmetrical treatment of observations): All observations have the same 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 bistochastic 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 bistochastic case, both *across*- and *within-interval* weights are normalized to unity. Our empirical exercise demonstrates the importance of this feature in the context of the boundary-bias reduction achieved by the proposed estimator.

PROPERTY 6: 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.

Proof. See Appendix B.

PROPERTY 7: The bistochastic estimator can be generalized to the multivariate regression case given that the Deming–Stephan algorithm can be applied to higher dimensions.

Property 7 allows one to generalize the application of the bistochastic nonparametric technique to a multidimensional framework.

5. A SIMULATION EXERCISE

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

We show that the bistochastic smoother performs better than alternative standard nonparametric methods that use 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 bistochastic estimator described in Section 4. For this reason, we undertake the following simulation exercise.

Design of the Exercise

In this simulation exercise, we evaluate the performance of the bistochastic 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, …, 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~N(0,1), and the uniform, X~U(0,1). Data on Y were generated from the model, Y = m(X)+ε, where ε ~ N(0,σ^2), with m(X) being either m(X) = Sin(2πX) or m(X) = exp(X). These two different specifications imply two quite different models. We also perform simulations using two different values for σ^2, namely σ^2 = 0.2 and σ^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 x 2 x 2 x 3 x 3 x 3 = 216 basic computations. The results were then obtained by taking the mean of 200 independent samples or repetitions of each basic computation.[12]

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 are not presented. 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. Results 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.[13] Consequently, the bistochastic estimator fits better.

[TABLE 1 ABOUT HERE]

How do we explain these results? 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 *x* values.[14] Bistochastic estimators give higher 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 (see also Footnote 3).

[FIGURE 1 ABOUT HERE]

This raises the question of why the bistochastic 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 5). 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*.

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 the greater efficiency in the neighborhood of the optimal bandwidth, whatever the optimal bandwidth criterion used.

[FIGURE 2 ABOUT HERE]

Furthermore, the bistochastic estimator converges more quickly to the true curve than does the classical Nadaraya–Watson smoother. Notice 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 conclusive evidence that the bistochastic estimator is superior to the local linear smoother, in this respect.

6. Concluding Remarks

In this paper, we have proposed a class of nonparametric estimators that use bistochastic smoothing, unlike standard nonparametric methods, which use stochastic smoothing. The bistochastic smoother has two main properties. First, the nonparametric bistochastic smoother reduces variability according to the robust criterion of second-order stochastic (and Lorenz) dominance. This is a universally applied dispersion-reducing criterion in risk and welfare economics that expands the applicability of nonparametric estimation. For example, Rodríguez et al. (2003) have proposed the use of this bistochastic nonparametric technique to measure tax discrimination due to a fiscal system. Furthermore, its usefulness can be generalized to other cases, such as the measurement of wage discrimination. It is indeed the underlying Lorenz smoothing property of the estimator that enhances the usefulness of the smoother in applied economics. Second, the bistochastic estimator is obtained from a given positive-weights stochastic smoother (such as the Nadaraya–Watson estimator), that reduces the estimation error. In our very general simulation exercise, the bias–variance trade-off is reduced by the bistochastic smoother. In particular, boundary bias is smaller. This is a main source of improvement in the bias–variance trade-off.

The bistochastic smoother imposes a double normalization (by rows and columns) of the weights matrix of the estimator, unlike stochastic nonparametric methods, which normalize only by rows. Double normalization 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.

Other practical properties of the bistochastic estimator are consistency, preservation of the dependent variable's mean value and multidimensional extension, which enhance its potential use in applied economics.

APPENDIX A: 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,

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

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,

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

APPENDIX B: 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

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

Hence, the sum of terms within the square brackets,

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

must be unity. The above expression can be rewritten as

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

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.

Berge, C., 1963, *Topological Spaces* (Oliver and Boyd, Edinburgh and London).

Bertschek, I. and H. Entorf, 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 H.A. Pott-Buter, 1991, Specification of household Engel curves by nonparametric regression, *Econometric Reviews* 9, 123–184.

Cowell, F.A., S.P. Jenkins and J.A. Litchfield, 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 M-P Victoria-Fesser, 1996, Robustness properties of inequality measures, *Econometrica* 64(1), 77–101.

Dasgupta, P., A. Sen and D. Starret, 1973, Notes on the measurement of inequality, *Journal of Economic Theory* 6, 180–187.

Delgado, M.A., Fariñas, J.C. and S. Ruano, 2002, Firm productivity and export markets: a nonparametric approach, *Journal of International Economics* 57, 397–422.

Delgado, M.A. and D. Miles, 1997, Household characteristics and consumption behaviour: A nonparametric approach, *Empirical Economics* 22, 409–429.

Deming, W.E. and F.F. Stephan, 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 J.A. Nason, 1990, Nonparametric exchange rate prediction, *Journal of International Economics* 28, 315–332.

Duclos, J.Y. and P. Lambert, 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 J.S. Marron, 1994, Fast implementations of nonparametric curve estimators, *Journal of Computational Graphical Statistics* 3, 35-56.

Gasser, T. and H.G. Müller, 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 COMPSTAT: *Proceedings in computational Statistics*, 97-101 (eds. F. de Antoni, N. Lauro and A. Rizzi). (Phisica-Verlag, Vienna).

Härdle, W., 1990, *Applied nonparametric regression* (Cambridge University Press, Cambridge).

Hildenbrand, K. and W. Hildenbrand, 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 S. Kullback, 1968, Contingency tables with given marginals, *Biometrika* 55, 179–188.

Nadaraya, E.A., 1964, On estimating regression, *Theory Probability Applied* 10,186–190.

Priestley, M. B. and M. T. Chao, 1972, Nonparametric function fitting, *Journal of the Royal Statistical Society*, Series B 34, 385–392.

Rodríguez, J.G., Salas, R. and I. Perrote, 2002, Partial horizontal inequity orderings: a nonparametric approach, mimeo.

Rothschild, M. and J. Stiglitz, 1970, Increasing Risk I: A definition, *Journal of Economic Theory* 2, 225–243.

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.

Wand, M.P. and M.C. Jones, 1995, *Kernel Smoothing*, Chapman and Hall.

Watson, G.S., 1964, Smooth regression analysis, *Sankhya Series A 26*, 101–16.

Yitzhaki, S., 1996, On using linear regressions in welfare economics, *Journal of Business and Economic Statistics* 14 (4), 478–486.

Table I: Bias, variance and d_c for the different estimators under X ~N(0,1) [a]

n=30

	Sin (2 X)						**exp**(X)					
	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 X)						**exp**(X)					
	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

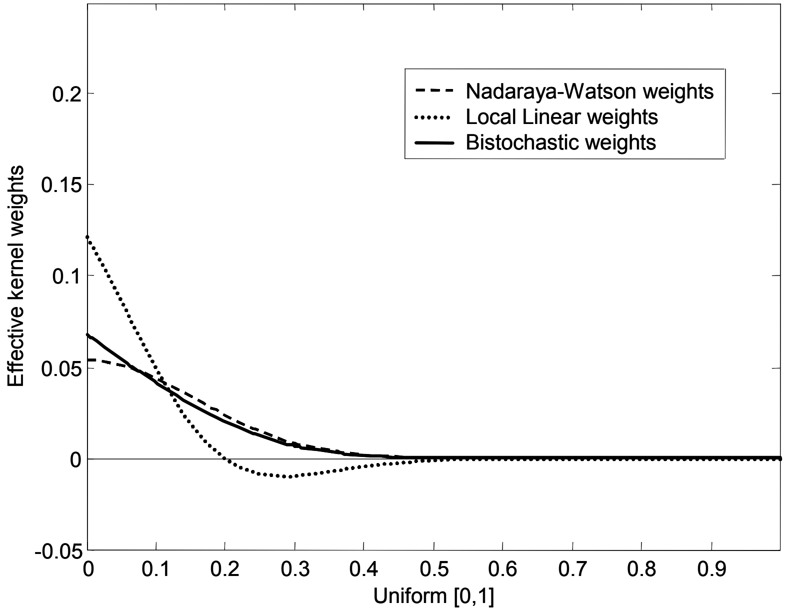
n=1000

	Sin (2 X)						**exp**(X)					
	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

[a] 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.

FIGURE 1

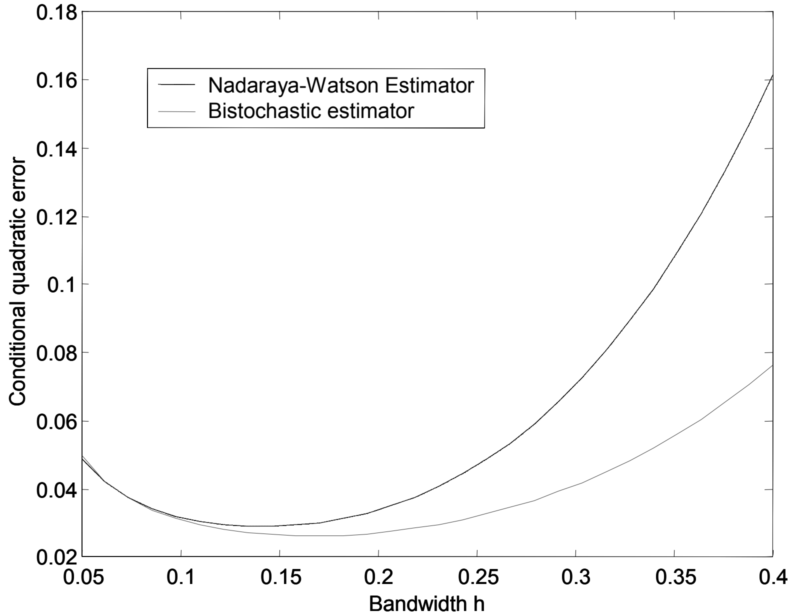
Effective kernel weights evaluated at the lowest X value.



Exercise design: $Y = m(X)+\varepsilon$, where $X \sim U(0, 1)$, $m(X) = \exp(X)$ and $\varepsilon \sim N(0, 0.2)$.

FIGURE 2

Conditional quadratic error (d_c) for different h bandwidth values.



Exercise design: Y = m(X)+ε, where X ~ N(0, 1), m(X) = exp(X) and ε ~ N(0, 0.2). The d_c function values correspond to the mean value for 200 independent samples.

FOOTNOTES

[1] A related exercise can be found in Yitzhaki (1996), which indicates the need to correct the OLS estimator, in the context of welfare economics, by using a nonparametric estimator that is consistent with the extended Gini coefficient.

[2] The second-order stochastic dominance criterion is well established in the literature on risk (Rothschild and Stiglitz, 1970) and welfare (Atkinson, 1970). The mean-preserving second-order stochastic dominance principle has counterpart equivalence with the mean-preserving spreads principle and the Lorenz dominance criterion.

[3] 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." This restriction prevents us from working with estimators such as the local linear smoother, which optimizes the minimax risk criterion (Fan, 1993), but uses negative weights.

[4] We concentrate on the bivariate regression case. Analogous extensions apply to multivariate regressions. See property 7 below.

[5] Of the methods available for undertaking double normalization of the weights matrix, we use a low-cost iterative proportional-fitting method. This is a special case of the algorithm proposed by Deming and Stephan (1940), which minimizes the Kullback–Liebler distance function (Ireland and Kullback, 1968). We find an interesting symmetry condition for this particular algorithm: the convergence result of the algorithm is independent of whether we begin normalizing by rows or columns.

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

[7] We are referring to the stochastic design sample model. However, extension to the fixed design sample model is straightforward.

[8] The weight function W_n is said to be a probability weight function if it is normalized ($\Sigma_j W_{nj}(x) = 1$) and nonnegative (see, for example Stone, 1977).

[9] These results can be obtained from the authors on request.

[10] Another nonparametric technique is the regressogram (Tukey, 1947), which guarantees that the weights matrix is bistochastic (see the proof in Appendix 1). However, this estimator is rarely used because it does not have desirable properties.

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

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

[13] 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.

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

centrA:

Fundación Centro de Estudios Andaluces

Documentos de Trabajo

Serie Economía

E2001/01 "The nineties in Spain: so much flexibility in the labor market?'', J. Ignacio García Pérez y Fernando Muñoz Bullón.

E2001/02 "A Log-linear Homotopy Approach to Initialize the Parameterized Expectations Algorithm'', Javier J. Pérez.

E2001/03 "Computing Robust Stylized Facts on Comovement'', Francisco J. André, Javier J. Pérez, y Ricardo Martín.

E2001/04 "Linking public investment to private investment. The case of the Spanish regions", Diego Martínez López.

E2001/05 "Price Wars and Collusion in the Spanish Electricity Market", Juan Toro y Natalia Fabra.

E2001/06 "Expedient and Monotone Learning Rules", Tilman Börgers, Antonio J. Morales y Rajiv Sarin.

E2001/07 "A Generalized Production Set. The Production and Recycling Function", Francisco J. André y Emilio Cerdá.

E2002/01 "Flujos Migratorios entre provincias andaluzas y entre éstas y el resto de España'', J. Ignacio García Pérez y Consuelo Gámez Amián.

E2002/02 "Flujos de trabajadores en el mercado de trabajo andaluz'', J. Ignacio García Pérez y Consuelo Gámez Amián.

E2002/03 "Absolute Expediency and Imitative Behaviour", Antonio J. Morales Siles.

E2002/04 "Implementing the 35 Hour Workweek by means of Overtime Taxation", Victoria Osuna Padilla y José-Víctor Ríos-Rull.

E2002/05 "Landfilling, Set-Up costs and Optimal Capacity", Francisco J. André y Emilio Cerdá.

E2002/06 "Identifying endogenous fiscal policy rules for macroeconomic models", Javier J. Pérez y Paul Hiebert.

E2002/07 "Análisis dinámico de la relación entre ciclo económico y ciclo del desempleo en Andalucía en comparación con el resto de España", Javier J. Pérez, Jesús Rodríguez López y Carlos Usabiaga.

E2002/08 "Provisión eficiente de inversión pública financiada con impuestos distorsionantes", José Manuel González-Páramo y Diego Martínez López.

E2002/09 "Complete or Partial Inflation convergence in the EU?", Consuelo Gámez y Amalia Morales-Zumaquero.

E2002/10 "On the Choice of an Exchange Regime: Target Zones Revisited", Jesús Rodríguez López y Hugo Rodríguez Mendizábal.

E2002/11 "Should Fiscal Policy Be Different in a Non-Competitive Framework?", Arantza Gorostiaga.

E2002/12 "Debt Reduction and Automatic Stabilisation", Paul Hiebert, Javier J. Pérez y Massimo Rostagno.

E2002/13 "An Applied General Equilibrium Model to Assess the Impact of National Tax Changes on a Regional Economy", M. Alejandro Cardenete y Ferran Sancho.

E2002/14 "Optimal Endowments of Public Investment: An Empirical Analysis for the Spanish Regions", Óscar Bajo Rubio, Carmen Díaz Roldán y M. Dolores Montávez Garcés.

E2002/15 "Is it Worth Refining Linear Approximations to Non-Linear Rational Expectations Models?" , Alfonso Novales y Javier J. Pérez.

E2002/16 "Factors affecting quits and layoffs in Spain", Antonio Caparrós Ruiz y M.ª Lucía Navarro Gómez.

E2002/17 "El problema de desempleo en la economía andaluza (1990-2001): análisis de la transición desde la educación al mercado laboral", Emilio Congregado y J. Ignacio García Pérez.

E2002/18 "Pautas cíclicas de la economía andaluza en el período 1984-2001: un análisis comparado", Teresa Leal, Javier J. Pérez y Jesús Rodríguez.

E2002/19 "The European Business Cycle", Mike Artis, Hans-Martin Krolzig y Juan Toro.

E2002/20 "Classical and Modern Business Cycle Measurement: The European Case", Hans-Martin Krolzig y Juan Toro.

E2002/21 "On the Desirability of Supply-Side Intervention in a Monetary Union", Mª Carmen Díaz Roldán.

E2003/01 "Modelo Input-Output de agua. Análisis de las relaciones intersectoriales de agua en Andalucía", Esther Velázquez Alonso.

E2003/02 "Robust Stylized Facts on Comovement for the Spanish Economy", Francisco J. André y Javier Pérez.

E2003/03 "Income Distribution in a Regional Economy: A SAM Model", Maria Llop y Antonio Manresa.

E2003/04 "Quantitative Restrictions on Clothing Imports: Impact and Determinants of the Common Trade Policy Towards Developing Countries", Juliette Milgram.

E2003/05 "Convergencia entre Andalucía y España: una aproximación a sus causas (1965-1995). ¿Afecta la inversión pública al crecimiento?", Javier Rodero Cosano, Diego Martínez López y Rafaela Pérez Sánchez.

E2003/06 "Human Capital Externalities: A Sectoral-Regional Application for Spain", Lorenzo Serrano.

E2003/07 "Dominant Strategies Implementation of the Critical Path Allocation in the Project Planning Problem", Juan Perote Peña.

E2003/08 "The Impossibility of Strategy-Proof Clustering", Javier Perote Peña y Juan Perote Peña.

E2003/09 "Plurality Rule Works in Three-Candidate Elections", Bernardo Moreno y M. Socorro Puy.

E2003/10 "A Social Choice Trade-off Between Alternative Fairness Concepts: Solidarity versus Flexibility", Juan Perote Peña.

E2003/11 "Computational Errors in Guessing Games", Pablo Brañas Garza y Antonio Morales.

E2003/12 "Dominant Strategies Implementation when Compensations are Allowed: a Characterization", Juan Perote Peña.

E2003/13 "Filter-Design and Model-Based Analysis of Economic Cycles", Diego J. Pedregal.

E2003/14 "Strategy-Proof Estimators for Simple Regression", Javier Perote Peña y Juan Perote Peña.

E2003/15 "La Teoría de Grafos aplicada al estudio del consumo sectorial de agua en Andalucía", Esther Velázquez Alonso.

E2003/16 "Solidarity in Terms of Reciprocity", Juan Perote Peña.

E2003/17 "The Effects of Common Advice on One-shot Traveler's Dilemma Games: Explaining Behavior through an Introspective Model with Errors", C. Monica Capra, Susana Cabrera y Rosario Gómez.

E2003/18 "Multi-Criteria Analysis of Factors Use Level: The Case of Water for Irrigation", José A. Gómez-Limón, Laura Riesgo y Manuel Arriaza.

E2003/19 "Gender Differences in Prisoners' Dilemma", Pablo Brañas-Garza y Antonio J. Morales-Siles.

E2003/20 "Un análisis estructural de la economía andaluza a través de matrices de contabilidad social: 1990-1999", M. Carmen Lima, M. Alejandro Cardenete y José Vallés.

E2003/21 “Análisis de multiplicadores lineales en una economía regional abierta", Maria Llop y Antonio Manresa.

E2003/22 “Testing the Fisher Effect in the Presence of Structural Change: A Case Study of the UK", Óscar Bajo-Rubio, Carmen Díaz-Roldán y Vicente Esteve.

E2003/23 "On Tests for Double Differencing: Some Extensions and the Role of Initial Values", Paulo M. M. Rodrigues y A. M. Robert Taylor.

E2003/24 "How Tight Should Central Bank’s Hands be Tied? Credibility, Volatility and the Optimal Band Width of a Target Zone", Jesús Rodríguez López y Hugo Rodríguez Mendizábal.

E2003/25 "Ethical implementation and the Creation of Moral Values", Juan Perote Peña.

E2003/26 "The Scoring Rules in an Endogenous Election", Bernardo Moreno y M. Socorro Puy.

E2003/27 "Nash Implementation and Uncertain Renegotiation", Pablo Amorós.

E2003/28 "Does Familiar Environment Affect Individual Risk Attitudes? Olive-oil Producer vs. no-producer Households", Francisca Jiménez Jiménez.

E2003/29 "Searching for Threshold Effects in the Evolution of Budget Deficits: An Application to the Spanish Case", Óscar Bajo-Rubio, Carmen Díaz-Roldán y Vicente Esteve.

E2003/30 "The Construction of input-output Coefficients Matrices in an Axiomatic Context: Some Further Considerations", Thijs ten Raa y José Manuel Rueda Cantuche.

E2003/31 "Tax Reforms in an Endogenous Growth Model with Pollution", Esther Fernández, Rafaela Pérez y Jesús Ruiz.

E2003/32 "Is the Budget Deficit Sustainable when Fiscal Policy is nonlinear? The Case of Spain, 1961-2001", Óscar Bajo-Rubio, Carmen Díaz-Roldán y Vicente Esteve.

E2003/33 "On the Credibility of a Target Zone: Evidence from the EMS", Francisco Ledesma-Rodríguez, Manuel Navarro-Ibáñez, Jorge Pérez-Rodríguez y Simón Sosvilla-Rivero.

E2003/34 "Efectos a largo plazo sobre la economía andaluza de las ayudas procedentes de los fondos estructurales: el Marco de Apoyo Comunitario 1994-1999", Encarnación Murillo García y Simón Sosvilla-Rivero.

E2003/35 “Researching with Whom? Stability and Manipulation”, José Alcalde y Pablo Revilla.

E2003/36 “Cómo deciden los matrimonios el número óptimo de hijos”, Francisca Jiménez Jiménez.

E2003/37 “Applications of Distributed Optimal Control in Economics. The Case of Forest Management”, Renan Goetz y Angels Xabadia.

E2003/38 **“**An Extra Time Duration Model with Application to Unemployment Duration under Benefits in Spain”, José María Arranz y Juan Muro Romero.

E2003/39 “Regulation and Evolution of Harvesting Rules and Compliance in Common Pool Resources”, Anastasios Xepapadeas.

E2003/40 “On the Coincidence of the Feedback Nash and Stackelberg Equilibria in Economic Applications of Differential Games”, Santiago J. Rubio.

E2003/41 “Collusion with Capacity Constraints over the Business Cycle”, Natalia Fabra.

E2003/42 “Profitable Unproductive Innovations”, María J. Álvarez-Peláez, Christian Groth.

E2003/43 “Sustainability and Substitution of Exhaustible Natural Resources. How Resource Prices Affect Long-Term R&D-Investments”, Lucas Bretschger, Sjak Smulders.

E2003/44 “Análisis de la estructura de la inflación de las regiones españolas: La metodología de Ball y Mankiw”, María Ángeles Caraballo, Carlos Usabiaga.

E2003/45 “An Empirical Analysis of the Demand for Physician Services Across the European Union”, Sergi Jiménez-Martín, José M. Labeaga, Maite Martínez-Granado.

E2003/46 “An Exploration into the Effects of Fiscal Variables on Regional Growth”, Diego Martínez López.

E2003/47 “Teaching Nash Equilibrium and Strategy Dominance: A Classroom Experiment on the Beauty Contest”. Virtudes Alba Fernández, Francisca Jiménez Jiménez, Pablo Brañas Garza, Javier Rodero Cosano.

E2003/48 "Environmental Fiscal Policies Might be Ineffective to Control Pollution", Esther Fernández, Rafaela Pérez y Jesús Ruiz.

E2003/49 "Non-stationary Job Search When Jobs Do Not Last Forever: A Structural Estimation to Evaluate Alternative Unemployment Insurance Systems", José Ignacio García Pérez.

E2003/50 “Poverty in Dictator Games: Awakening Solidarity”, Pablo Brañas-Garza.

E2003/51 “Exchange Rate Regimes, Globalisation and the Cost of Capital in Emerging Markets” Antonio Díez de los Ríos.

E2003/52 “Opting-out of Public Education in Urban Economies”. Francisco Martínez Mora.

E2004/01 "Partial Horizontal Inequity Orderings: A non-parametric Approach", Juan Gabriel Rodríguez, Rafael Salas, Irene Perrote.

E2004/02 "El enfoque microeconómico en la estimación de la demanda de transporte de mercancías. Análisis desde una perspectiva regional", Cristina Borra Marcos, Luis Palma Martos.

E2004/03 "El marco del SEC95 y las matrices de contabilidad social: España 19951", M. Alejandro Cardenete, Ferran Sancho.

E2004/04 "Performing an Environmental Tax Reform in a Regional Economy. A Computable General Equilibrium Approach", Francisco J. André, M. Alejandro Cardenete, E. Velázquez.

E2004/05 "Is the Fisher Effect Nonlinear? Some Evidence for Spain, 1963-2002", Óscar Bajo-Rubio, Carmen Díaz-Roldán, Vicente Esteve.

E2004/06 "On the Use of Differing Money Transmission Methods by Mexican Immigrants", Catalina Amuedo-Dorantes, Susan Pozo.

E2004/07 "The Motherhood Wage Gap for Women in the United States: The Importance of College and Fertility Delay", Catalina Amuedo-Dorantes, Jean Kimmel.

E2004/08 "Endogenous Financial Development and Multiple Growth Regimes", Costas Azariadis, Leo Kaas.

E2004/09 "Devaluation Beliefs and Debt Crisis: The Argentinian Case", José-María Da-Rocha, Eduardo L. Giménez, Francisco-Xavier Lores.

E2004/10 "Optimal Fiscal Policy with Rationing in the Labor Market", Arantza Gorostiaga.

E2004/11 "Switching Regimes in the Term Structure of Interest Rates During U.S. Post-War: A case for the Lucas proof equilibrium?", Jesús Vázquez.

E2004/12 "Strategic Uncertainty and Risk Attitudes: "The Experimental Connection", Pablo Brañas-Garza, Francisca Jiménez-Jiménez, Antonio J. Morales.

E2004/13 "Scope Economies and Competition Beyond the Balance Sheet: a 'broad banking' Experience", Santiago Carbó Valverde, Francisco Rodríguez Fernández.

E2004/14 "How to Estimate Unbiased and Consistent input-output Multipliers on the basis of use and Make Matrices", Thijs ten Raa, José Manuel Rueda Cantuche.

E2004/15 "Double Dividend in an Endogenous Growth Model with Pollution and Abatement", Esther Fernández, Rafaela Pérez, Jesús Ruiz.

E2004/16 "Technology Adoption in Nonrenewable Resource Management", Maria A. Cunha-e-Sá, Ana Balcão Reis, Catarina Roseta-Palma.

E2004/17 "Personal Income Tax Decentralization, Inequality and Social Welfare", Julio López Laborda, Jorge Onrubia Fernández.

E2004/18 "The Effects of Reducing Firing Costs in Spain: a Lost Opportunity?", Victoria Osuna Padilla.

E2004/19 "Un modelo input-output de precios aplicado a la economía extremeña", Francisco Javier de Miguel Vélez, Antonio Manresa Sánchez.

E2004/20 "A Model of Political Campaign Manipulation", Pablo Amorós, M. Socorro Puy.

E2004/21 "Evaluación regional del impuesto monofásico sobre las ventas minoristas de determinados hidrocarburos a través de matrices de contabilidad social", José Manuel Cansino, M. Alejandro Cardenete, Rocío Román.

E2004/22 "A Bistochastic Nonparametric Estimator", Juan Gabriel Rodríguez, Rafael Salas.

centrA:
Fundación Centro de Estudios Andaluces

Normas de publicación de Documentos de Trabajo centrA Economía

La Fundación Centro de Estudios Andaluces (**centrA**) tiene como uno de sus objetivos prioritarios proporcionar un marco idóneo para la discusión y difusión de resultados científicos en el ámbito de la Economía. Con esta intención pone a disposición de los investigadores interesados una colección de Documentos de Trabajo que facilita la transmisión de conocimientos. La Fundación Centro de Estudios Andaluces invita a la comunidad científica al envío de trabajos que, basados en los principios del análisis económico y/o utilizando técnicas cuantitativas rigurosas, ofrezcan resultados de investigaciones en curso.

Las normas de presentación y selección de originales son las siguientes:

1. El autor(es) interesado(s) en publicar un Documento de Trabajo en la serie de Economía de centrA debe enviar su artículo en formato PDF a la dirección de email: wpecono@fundacion-centra.org
2. Todos los trabajos que se envíen a la colección han de ser originales y no estar publicados en ningún medio de difusión. Los trabajos remitidos podrán estar redactados en castellano o en inglés.
3. Los originales recibidos serán sometidos a un breve proceso de evaluación en el que serán directamente aceptados para su publicación, aceptados sujetos a revisión o rechazados. Se valorará, asimismo, la presentación dei trabajo en seminarios de **centrA**.
4. En la primera página deberá aparecer el título del trabajo, nombre y filiación del autor(es), dirección postal y electrónica de referencia y agradecimientos. En esta misma página se incluirá también un resumen en castellano e inglés de no más de 100 palabras, los códigos JEL y las palabras clave de trabajo.
5. Las notas al texto deberán numerarse correlativamente al pie de página. Las ecuaciones se numerarán, cuando el autor lo considere necesario, con números arábigos entre corchetes a la derecha de las mismas.
6. La Fundación Centro de Estudios Andaluces facilitará la difusión electrónica de los documentos de trabajo. Del mismo modo, se incentivará económicamente su posterior publicación en revistas científicas de reconocido prestigio.