DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

THE PARTIALLY LINEAR REGRESSION MODEL: MONTE CARLO EVIDENCE FROM THE PROJECTION PURSUIT REGRESSION APPROACH

Dingding Li and Thanasis Stengos

Discussion Paper 2001-11

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: ++357-2-892430, Fax: ++357-2-892432
Web site: http://www.econ.ucy.ac.cy

The Partially Linear Regression Model: Monte Carlo Evidence from the Projection Pursuit Regression Approach.

Dingding Li and Thanasis Stengos*

Department of Economics

University of Guelph

Guelph, Ontario, N1G 2W1

April, 2001

Abstract

In a partially linear regression model with a high dimensional unknown component we find an estimator of the parameter of the linear part based on projection pursuit methods to be considerably more efficient than the standard density weighted kernel estimator.

Keywords: Partially Linear Model; Projection Pursuit Regression

JEL classification: C14, C15

*The second author wants to acknowledge support from SSHRC of Canada.

1 Introduction

The most extensively studied nonparameteric regression techniques are based on p-dimensional local averaging: the estimate of the regression surface at a point x_0 is the average of the responses of those observations with predictors in a neighborhood of x_0. These techniques can be shown to have desirable asymptotic properties. In high-dimensional settings, however, they do not perform well for reasonable sample sizes. The reason is the inherent sparsity of high-dimensional samples. This is the well known "curse of dimensionality" that plagues nonparametric methods.

Among the methods developed in the last few years in order to deal with the above problem is the Projection Pursuit Regression (PPR), see Friedman and Steutzle (1981) and Huber (1985) for an excellent review. The PPR is based on an approach designed to detect structure in high-dimensional data sets by viewing lower dimensional representations of the data by means of linear projections. The power of the technique stems from its ability to overcome the "curse of dimensionality" since it relies in estimation in at most tri-variate settings. The term Projection Pursuit was coined by Friedman and Tukey (1974) in the first successful implementation of the technique.

Hall (1989) proved that the common form of kernel-based PPR did estimate projections with convergence rates identical to those encountered in one-dimensional problems, although a greater degree of smoothness (in fact, an extra derivative) must be assumed to achieve this end.

In this paper we will employ PPR in a partially linear semiparametric regression (PLR) model and by means of Monte Carlo simulations evaluate the efficiency gains that are obtained using PPR as compared with the more

widely applied local constant kernel methods. In the next section we present a brief presentation of the idea behind PPR. We proceed then to present the PLR model, see Robinson (1988). Finally we present the results of the simulation.

2 The Projection Pursuit Regression Approach

The PPR approach is based on projections of the data on planes spanned by the endogenous variable y and a linear combination $\alpha^T x$ of the explanatory variables. The idea of PPR is to approximate the mean regression function by a sum of unknown ridge functions. Diaconis and Shahshahani (1984) provide an approximation theory justification as the number of ridge functions goes to infinity.

Many data sets are high dimensional. It has been a common practice to use lower dimensional linear projections of the data for visual inspection. The lower dimension is usually 1 or 2 (or maybe 3). More precisely, if $x_1, ..., x_n \in R^p$ are p-dimensional data, then a $k(< p)$-dimensional linear projection is $w_1, ..., w_n \in R^k$ where $w_i = \alpha^T x_i$ for some $p \times k$ matrix α such that $\alpha^T \alpha = I_k$, the k-dimensional identity matrix. When $k = 1$, the structure of the projected data can be viewed through a histogram; when $k = 2$, the structure can be inspected through its scatter plot; and when $k = 3$, it can be comprehensived by spinning a three-dimensional scatter plot.

Since there are infinitely many projections from a higher dimension to a lower dimension, it is important to have a technique of pursuing a finite sequence of projections that can reveal the most interesting structures of the

data. Below we will present a simple description of the mechanics involved in regression estimation for the case $k = 1$.

In a typical regression model, (X, y) is an observable pair of random variables, where $X \in R^p$ is a p-dimensional variable and $y \in R$ is the dependent variable. The goal is to estimate the unknown regression function $f(x) = E(y|X = x)$, using a random sample $(x_1, y_1), ..., (x_n, y_n)$. PPR approximates the unknown regression function $f(x)$ by a finite sum of ridge functions $g^m(x) = \sum_{j=1}^{m} g_j(x_j^T a)$. The regression function $f(x_i), (x_i \text{ is } 1 \times p)$ is written then as the sum of $m < p$ unknown nonlinear functions in scalars $w_{ij} = x_i^T a_j$

$$y_i = \sum_{j=1}^{m} g_j(x_i^T a_j) + \varepsilon_i \tag{1}$$

Suppose that we want to estimate a_1 in $w_{i1} = x_i^T a_1$. We start by conditioning (1) on w_{i1}.

$$E(y_i|w_{i1}) = \sum_{j=1}^{k} E[g_j(w_{ij})|w_{i1}] \tag{2}$$

We can write

$$y_i = E(y_i|w_{i1}) + u_i \tag{3}$$

where

$$u_i = \sum_{j=2}^{k} \{g_j(x_i^T a_j) - E[g_j(w_{ij})|w_{i1}\} + \varepsilon_i$$

Hence by construction, $E(u_i|w_{i1}) = 0$. The above can be also written as

$$y_i = f(w_{i1}) + u_i \tag{4}$$

If $f(.)$ were known in (4) one could simply apply Nonlinear Least Squares (NLS) to obtain an estimate of a_1. However, since $f(.)$ is unknown we proceed as follows.

Step 1: We start with an initial guess for a_1, say a_1^0, then form $w_{i1}^0 = x_i^T a_1^0$. We then estimate by nonparametric kernel methods $E(y_i|w_{i1}^0)$. We choose the value of a_1 than minimizes the Sum of Squared Residuals (SSR) function $\sum_{i=1}^n (y_i - E(y_i|w_{i1}))^2$. That yields an estimate of $g_1(w_{i1})$, say $\widehat{g}_1(w_{i1})$.

Step 2: We form $y_i^* = y_i - \widehat{g}_1(w_{i1})$ and we repeat the procedure of step 1 above to obtain $g_2(w_{i2})$.

Step 3: We proceed to form $y_i^{**} = y_i^* - \widehat{g}_2(w_{i2})$ and we continue as before. We stop whenever the improvement in the SSR function between successive iterations is less than some than some predetermined small tolerance value.

Note that in the implementation of the PPR method we only have to deal with univariate constant kernel regressions. They have to be computed for all n observations at each stage of the each iteration however, and for large n there is a computattional burden invlolved.

2.1 Semiparametric Estimation of Partially Linear Models

Our analysis will follow closely that of Robinson (1988). We consider a semiparametric partially linear model (e.g., Engle, et al (1986), Robinson (1988), Stock (1989)):

$$y_i = x_i'\beta + \theta(z_i) + u_i, \tag{5}$$

where x_i is of dimension $p \times 1$ and z_i is of dimension $q \times 1$. β is an unknown parameter of dimension $p \times 1$. $E(u_i|x_i, z_i) = 0$ and the conditional variance function $\sigma^2(x_i, z_i) = E(u_i^2|x_i, z_i)$ is not specified, $\theta(\cdot)$ is an unknown smooth function. We interested (mainly) in estimating β.

Following Robinson (1988), we use a two-step procedure and we will first estimate β. Taking conditional expectation of (1) (conditional on z_i) and then subtracting it from (1) yields

$$y_i - E(y_i|z_i) = \{x_i - E(x_i|z_i)\}'\beta + u_i. \tag{6}$$

Equation (2) no longer has the unknown function $\theta(\cdot)$. Therefore, we can estimate β based on equation (2). However, $E(y_i|z_i)$, $E(x_i|z_i)$ are unknown in practice. These conditional expectations can be consistently estimated using some nonparametric methods. In this paper we will use the kernel method. In order to avoid the random denominator problem associated with nonparametric kernel estimation, we choose to estimate a density weighted version of equation (2) (e.g., Powell, Stock and Stoker (1989)).

Let $\phi(\cdot)$ be the density function of z_i. Multiplying equation (2) by $\phi(z_i)$ gives

$$\phi(z_i)\{y_i - E(y_i|z_i)\} = \phi(z_i)\{x_i - E(x_i|z_i)\}'\beta + \phi(z_i)u_i. \tag{7}$$

We estimate $E(u_i|z_i)$, $E(x_i|z_i)$ and $\phi(z_i)$ by kernel estimators given by

$$\widehat{y}_i \equiv \hat{E}(y_i|z_i) = \frac{1}{Nh^q}\sum_{s\neq i} y_s K_{i,s}/\widehat{\phi}_i \tag{8}$$

$$\widehat{x}_i \equiv \hat{E}(x_i|z_i) = \frac{1}{Nh^q}\sum_{s\neq i} x_s K_{i,s}/\widehat{\phi}_i \tag{9}$$

and

$$\widehat{\phi}_i \equiv \widehat{\phi}(z_i) = \frac{1}{Nh^q}\sum_{s\neq i} K_{i,s}, \tag{10}$$

where $K_{i,s} = K(\frac{z_i - z_s}{h})$ is the kernel function and h is the smoothing parameter. We use the product kernel $K(z_i) = \prod_{l=1}^{q} k(z_{i,l})$, where k is a univariate kernel and $z_{i,l}$ is the $l-$th component of z_i.

Then replacing the unknown functions in (7) by their respective kernel estimates we get

$$\widehat{\phi}_i(y_i - \widehat{y}_i) = \widehat{\phi}_i(z_i - \hat{z}_i)'\beta + \widehat{\phi}_i u_i. \tag{11}$$

Therefore we estimate β by the least squares method of regressing $(y_i - \widehat{y}_i)\widehat{\phi}_i$ on $(x_i - \hat{x}_i)\widehat{\phi}_i$.

$$\widehat{\beta} = S^{-1}_{(x-\widehat{x})\widehat{\phi}} S_{(x-\widehat{x})\hat{f},(y-\widehat{y})\widehat{\phi}}, \tag{12}$$

where for scalar or column-vector sequences with ith elements $A_i\widehat{\phi}_i$ and $B_i\widehat{\phi}_i$, the notations $S_{A\widehat{\phi},B\widehat{\phi}} = \frac{1}{N}\sum_i A_i\widehat{\phi}_i B_i'\widehat{\phi}_i$ and $S_{A\widehat{\phi}} = S_{A\widehat{\phi},A\widehat{\phi}}$. The estimator $\widehat{\beta}$ is then shown to be $\sqrt{n} - consistent$. Below we will investigate its small sample properties when we compute it using the standard approach based on constant kernel methods as in equation (11) and applying PPR to $E(y_i|z_i)$ and $E(x_i|z_i)$ and then using least squares estimation directly on equation (6).

3 Monte Carlo Simulations

We implement the PPR method on the partially linear regression model by obtaining estimates of $E(y_i|z_i)$ and $E(x_i|z_i)$ and then using least squares estimation directly on equation (6) to obtain an estimate of β. We carry out simulations to investigate the behavior of $\widehat{\beta}$ in small to moderate sample sizes under two different cases, a low dimensional and a high dimensional $z-$vector. The standard approach is implemented using the density weigthed method outlined in the previous section. The Monte Carlo design is similar to that of Robinson (1988).

Table 1 contains the results for the model

$$y_i = x_i\beta + z_i^2\gamma + u_i, \tag{13}$$

The variables X and Z are taken to be scalars from a bivariate normal population with zero means, variances 4 and 3 respectively and covariance 2; The error term u is also chosen to be standard normal. The parameters are taken to be $\beta = \gamma = 1$ and the the sample size n is chosen as 50, 100, 200 and 300. The bandwidth selection was done by cross validation both for the standard density weighted semiparametric estimator and the PPR estimator.

For each sample size we report the mean square error (MSE) for $\overset{\Lambda}{\beta}_{semi}$, the standard density weighted semiparametric estimator of β and $\overset{\Lambda}{\beta}_{ppr}$, the PPR estimator of β. We also report a measure of relative efficiency between the two estimators as the ratio of the standard deviation $\widehat{\beta}_{ppr}$ to that of $\widehat{\beta}_{semi}$. A value greater than unity indicates that $\widehat{\beta}_{ppr}$ is less efficient than $\widehat{\beta}_{semi}$ and a value less than unity the opposite. The results suggest that for an one-dimensional z, the standard density weighted approach produces overall more efficient results, since all the values of the relative efficiency measure are greater than unity, except for the case when $n = 50$. The MSE values are also smaller for the estimates from the standard approach.

Table 2 presents the results for the case of a Z which is higher dimensional. The equation that generates the data is given by

$$y_i = \alpha + x_i\beta + \sum_{j=1}^{q} z_{ji}^2\gamma_j + u_i, \tag{14}$$

where $q = 5$, and $\beta, \gamma_j, (j = 1, ..., 5)$ are all 1; u is as before standard normal and X and the $Z_j's$ are equicorrelated identically distributed $N(1, 3)$ variables, with correlation 2/3. The results are presented in Table 2. It

becomes quite apparanet that the PPR approach produces considerable improvements in relative efficiency. As the sample size increases the measure of relative efficiency increases as well from about 0.5 for $n = 50$ to about 0.75 for $n = 300$. As it is expected as n increases the performace of the standard semiparametric estimator improves but it is apparent that for small to moderate sizes the PPR approach offers substantial efficiency gains. Hence, it can become a useful tool to improve the small sample efficiency of local averaging nonparametric estimators obtained in high dimensional settings.

References

[1] Diaconis, P. and Shahshahani, M. (1984). On Nolinear Functions of Linear Combinations. *SIAM Journal of Science and Statistical Computation*, 5, 175-191.

[2] Engle, R.F., C.W.J. Grange, J. Rice and A. Weiss, 1986, Semiparametric estimates of the relationship between weather and electricity sales, *Journal of the American Statistical Association* 81, 310-320.

[3] Friedman, J. H., & Tukey, J. (1974), A projection pursuit algorithm for explornatory data analysis. *IEEE Transactions on Computer*, C-23, 881-889.

[4] Friedman, J.H and Stuetzle, W. (1981), Projection Pursuit Regression, *Journal of the American Statistical Association*, 76, 817-823

[5] Hall, P. (1989), On Projection Pursuit Regression, *Annals of Statistics*, 17, 573-588.

[6] Huber, P. (1985), Projection pursuit, *Annals of Statistics*, 13, 435-525

[7] Powell, J.L., J.H. Stock, and T.M. Stoker, 1989, Semiparametric estimation of index coefficients, *Econometrica* 57, 1403-1430.

[8] Robinson, P., 1988, Root-N consistent semiparametric regression, *Econometrica* 56, 931-954.

[9] Stock, J.H., 1989, Nonparametric policy analysis, *Journal of American Statistical Association* 567-575.

Table 1: One-Dimensional unknown Component

n	MSE-ppr	MSE-semi	Rel. EFF
50	0.0134	0.0140	0.9720
100	0.0069	0.0062	1.0173
200	0.0032	0.0028	1.0240
300	0.0021	0.0021	1.0241

Table 2: Multi-Dimensional unknown Component

n	MSE-ppr	MSE-semi	Rel. EFF
50	0.0483	0.2000	0.5176
100	0.0245	0.0738	0.6323
200	0.0120	0.0324	0.7145
300	0.0056	0.0211	0.7498

SELECTED RECENT PUBLICATIONS

Andreou E. and E. Ghysels, Rolling Volatility Estimators: Some New Theoretical, Simulation and Empirical Results, *Journal of Business and Economic Statistics*, forthcoming 2001.

Andreou E. and A. Spanos, Testing Trend versus Difference Stationarity and Statistical Adequacy, forthcoming *Econometric Reviews*, 2001.

Andreou E., N. Pittis and A. Spanos, Modelling Stock Returns: The Empirical Literature, *Journal of Economic Surveys*, 15, 2, 187-220.

Andreou E., R. Desiano and M. Sensier, The Behaviour of Stock Returns and Interest Rates over the Business Cycle in the US and UK, *Applied Economic Letters*, 8, 233-238, 2001.

Andreou E., D. R. Osborn and M. Sensier, A Comparison of the Statistical Properties of Financial Variables in the USA, UK and Germany over the Business Cycle, *The Manchester School*, 68, 4, 396-418, 2000.

Anil K. Bera and Y. Bilias, Rao´s Score, Neyman´s C (alpha) and Silvey´s LM Tests: An Essay on Historical Developments and Some New Results, *Journal of Statistical Planning and Inference*, 97, 9-44, 2001.

Bertaut C. and M. Haliassos, Precautionary Portfolio Behavior from a Life-Cycle Perspective, *Journal of Economic Dynamics and Control*, 21, 1511-1542, 1997.

Bilias Y., Minggao Gu and Zhiliang Ying, Towards a General Asymptotic Theory for the Cox model with Staggered Entry, *The Annals of Statistics*, 25, 662-682, 1997.

Blundell R., P. Pashardes and G. Weber, What Do We Learn About Consumer Demand Patterns From Micro-Data?, *American Economic Revie*w, 83, 570-597, 1993.

Bougheas S., P. Demetriades and T. P. Mamouneas, Infrastructure, Specialization and Economic Growth, *Canadian Journal of Economics*, forthcoming.

Caporale W., C. Hassapis and N. Pittis, Unit Roots and Long Run Causality: Investigating the Relationship between Output, Money and Interest Rates, *Economic Modeling*, 15(1), 91-112, January 1998.

Caporale G. and N. Pittis, Efficient estimation of cointegrated vectors and testing for causality in vector autoregressions: A survey of the theoretical literature, *Journal of Economic Surveys*, forthcoming.

Caporale G. and N. Pittis, Unit root testing using covariates: Some theory and evidence, *Oxford Bulletin of Economics and Statistics*, forthcoming.

Caporale G. and N. Pittis, Causality and Forecasting in Incomplete Systems, *Journal of Forecasting*, 16, 6, 425-437, 1997.

Clerides K. S., S. Lach and J.R. Tybout, Is Learning-by-Exporting Important? Micro-Dynamic Evidence from Colombia, Morocco, and Mexico, *Quarterly Journal of Economics* 113(3), 903- 947, August 1998.

Cukierman A., P. Kalaitzidakis, L. Summers and S. Webb, Central Bank Independence, Growth, Investment, and Real Rates", Reprinted in Sylvester Eijffinger (ed), Independent Central Banks and Economic Performance, Edward Elgar, 416-461, 1997.

Dickens R., V. Fry and P. Pashardes, Non-Linearities and Equivalence Scales, *The Economic Journa*l, 103, 359-368, 1993.

Demetriades P. and T. P. Mamuneas, Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies, *Economic Journal*, July 2000.

Eicher Th. and P. Kalaitzidakis, The Human Capital Dimension to Foreign Direct Investment: Training, Adverse Selection and Firm Location". In Bjarne Jensen and Kar-yiu Wong (eds), Dynamics,Economic Growth, and International Trade, The University of Michigan Press, 337-364, 1997.

Fry V. and P. Pashardes, Abstention and Aggregation in Consumer Demand, *Oxford Economic Paper*s, 46, 502-518, 1994.

Gatsios K., P. Hatzipanayotou and M. S. Michael, International Migration, the Provision of Public Good and Welfare, *Journal of Development Economics*, 60/2, 561-577, 1999.

Guiso, L., M. Haliassos, and T. Jappelli, Household Portfolios: An International Comparison, forthcoming in Guiso, Haliassos, and Jappelli (Eds.), *Household Portfolios*, Cambridge, MA: MIT Press, 2002.

Haliassos M., On Perfect Foresight Models of a Stochastic World, *The Economic Journal*, 104, 477-491, 1994.

Haliassos M. and C. Bertaut, Why Do So Few Hold Stocks?, *The Economic Journal*, 105, 1110- 1129, 1995.

Haliassos M. and C. Hassapis, Non-expected Utility, Saving, and Portfolios, *The Economic Journal*, 111, 69-102, 2001.

Haliassos, M. and A. Michaelides, Portfolio Choice and Liquidity Constraints, *International Economic Review*, forthcoming.

Haliassos, M. and A. Michaelides, Calibration and Computation of Household Portfolio Models, forthcoming in Guiso, Haliassos, and Jappelli (Eds.), *Household Portfolios*, Cambridge, MA: MIT Press, 2002.

Haliassos M. and J. Tobin, The Macroeconomics of Government Finance, reprinted in J. Tobin, *Essays in Economics*, vol. 4, Cambridge: MIT Press, 1996.

Hassapis C., S. Kalyvitis and N. Pittis, Cointegration and Joint Efficiency of International Commodity Markets", *The Quarterly Review of Economics and Finance,* 39, 213-231, 1999.

Hassapis C., N. Pittis and K. Prodromidis, Unit Roots and Granger Causality in the EMS Interest Rates: The German Dominance Hypothesis Revisited, *Journal of International Money and Finance*, 18(1), 47-73, 1999.

Hassapis C., N. Pittis and K. Prodromides, EMS Interest Rates: The German Dominance Hypothesis or Else?" in European Union at the Crossroads: A Critical Analysis of Monetary Union and Enlargement, Aldershot, UK., Chapter 3, 32-54, 1998. Edward Elgar Publishing Limited.

Hatzipanayotou, P. and M.S. Michael, Public Goods, Tax Policies and Unemployment in LDC's, *Southern Economic Journal*, 68/1, 107-119, 2001.

Hatzipanayotou P., and M. S. Michael, General Equilibrium Effects of Import Constraints Under Variable Labor Supply, Public Goods and Income Taxes, *Economica*, 66, 389-401, 1999.

Hatzipanayotou, P. and M. S. Michael, Public Good Production, Nontraded Goods and Trade Restriction, *Southern Economic Journal*, 63, 4, 1100-1107, 1997.

Hatzipanayotou, P. and M. S. Michael, Real Exchange Rate Effects of Fiscal Expansion Under Trade Restrictions, *Canadian Journal of Economics*, 30-1, 42-56, 1997.

Kalaitzidakis P., T. P. Mamuneas and Th. Stengos, A Nonlinear Sensitivity Analysis of Cross-Country Growth Regressions, *Canadian Journal of Economics*, forthcoming.

Kalaitzidakis P., T. P. Mamuneas and Th. Stengos, European Economics: An Analysis Based on Publications in Core Journals, *European Economic Revie*w, 1999.

Kalaitzidakis P., On-the-job Training Under Firm-Specific Innovations and Worker Heterogeneity, *Industrial Relations*, 36, 371-390, July 1997.

Ludvigson S. and A. Michaelides, Does Buffer Stock Saving Explain the Smoothness and Excess Sensitivity of Consumption?, *American Economic Review*, 631-647, June 2001.

Lyssiotou P., P. Pashardes and Th. Stengos, Age Effects on Consumer Demand: An Additive Partially Linear Regression Model, *Canadian Journal of Economics*, forthcoming 2002.

Lyssiotou P., Dynamic Analysis of British Demand for Tourism Abroad, *Empirical Economics*, 15, 421-436, 2000.

Lyssiotou P., P. Pashardes and Th. Stengos, Testing the Rank of Engel Curves with Endogenous Expenditure, *Economics Letters*, 64, 61-65, 1999.

Lyssiotou P., P. Pashardes and Th. Stengos, Preference Heterogeneity and the Rank of Demand Systems, *Journal of Business and Economic Statistics*, 17 (2), 248-252, April 1999.

Lyssiotou P., Comparison of Alternative Tax and Transfer Treatment of Children using Adult Equivalence Scales, *Review of Income and Wealth*, 43 (1), 105-117, March 1997.

Mamuneas, Theofanis P., Spillovers from Publicly – Financed R&D Capital in High-Tech Industries, *International Journal of Industrial Organization*, 17(2), 215-239, 1999.

Mamuneas, Theofanis P. and M. I. Nadiri, R&D Tax Incentives and Manufacturing-Sector R&D Expenditures, in *Borderline Case: International Tax Policy, Corporate Research and Development, and Investmen*t, James Poterba (ed.), National Academy Press, Washington D.C., 1997. Reprinted in *Chemtec*h, 28(9), 1998.

Mamuneas, Theofanis P. and M. I. Nadiri, Public R&D Policies and Cost Behavior of the US Manufacturing Industries, *Journal of Public Economics*, 63, 57-81, 1996.

Michael S. Michael and Panos Hatzipanayotou, Welfare Effects of Migration in Societies with Indirect Taxes, Income Transfers and Public Good Provision, *Journal of Development Economics*, 64, 1-24, 2001.

Michaelides, A. and S. Ng, Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of three Simulation Estimators, *Journal of Econometrics*, 96(2), 231-266, June 1997.

Pashardes P., Equivalence Scales in a Rank-3 Demand System, *Journal of Public Economic*s, 58, 143-158, 1995.

Pashardes P., Bias in Estimating Equivalence Scales from Grouped Data, *Journal of Income Distributio*n, Special Issue: Symposium on Equivalence Scales, 4, 253-264,1995.

Pashardes P., Bias in Estimation of the Almost Ideal Demand System with the Stone Index Approximation, *Economic Journa*l, 103, 908-916, 1993.

Spanos A., Revisiting Date Mining: ´Hunting´ With or Without a License, *Journal of Methodology*, July 2000.

Spanos A., On Normality and the Linear Regression Model, *Econometric Reviews*, 14,195-203, 1995.

Spanos A., On Theory Testing in Econometrics: Modeling with nonexperimental Data, *Journal of Econometrics*, 67, 189-226, 1995.

Spanos A., On Modeling Heteroscedasticity: The Student's *t* and Elliptical Linear Regression Models, *Econometric Theor*y, 10, 286-315, 1994.

Zacharias E. and S. R. Williams, Ex Post Efficiency in the Buyer's Bid Double Auction when Demand Can be Arbitrarily Larger than Supply, *Journal of Economic Theory* 97, 175-190, 2001.