Rochester Center for Economic Research

Some Simulation Studies of Non-Parametric Estimators

Hong, Y. and A. Pagan

Working Paper No. 137
June 1988

University of
Rochester

SOME SIMULATION STUDIES OF NON-PARAMETRIC ESTIMATORS*

by

Y. Hong and A Pagan

University of Rochester

Working Paper No. 137

June 1988

*This material is based upon work supported by the National Science Foundation under Grant No. SES-8719520.

Some Simulation Studies of Non-Parametric Estimators[1]

Y.Hong and A.Pagan

University of Rochester

Address for Proofs: Dept. of Economics,

University of Rochester,

Rochester,N.Y. 14627

U.S.A.

Running Title: Non-parametric Estimator Simulation

---

[1]This material is based upon work supported by the National Science Foundation under Grant No. SES-8719520.

Abstract

This paper constructs a number of Monte Carlo studies to assess the quality of various non-parametric estimators that have been proposed recently for the estimation of non-linear econometric models. We consider both kernel and Fourier series based methods of estimation, and also examine techniques that have been suggested to improve the bias properties of the kernel estimator. The two models examined are a production function and a model emphasising the effects of risk. The Fourier estimator does very well in estimating the first of these, but not the second, while the kernel estimator shows substantial bias for the first, which is only partially alleviated by the procedures advocated for bias correction, and good results for the second

Introduction

Many problems in econometrics are characterized by an unknown functional form, and a wide variety of specifications has emerged to give empirical researchers the needed flexibility when estimating relations between observed data. For capturing production relations there are the Cobb-Douglas, CES, transcendental log, quadratic etc., functions. For demand studies using household data a variety of transformations such as Box-Cox, semi-log etc., have all emerged to estimate Engel curves. Finally, in macroeconomic models a key variable is frequently the expectation of a random variable with respect to an information set, in which case it is common to assume that the expectation is a linear function of observed variables. More recently conditional variances have assumed importance and for this quadratic formulations such as Engle's (1982) ARCH have found favor.

This parametric emphasis to modeling has been challenged in recent years with the advent of non-parametric procedures. Gallant (1982) advocated and applied a Fourier approximation approach to the determination of production

technologies while Rilstone and Ullah (1987) estimated marginal products of factors with the kernel estimator of Nadaraya (1964) and Watson (1964). Deaton (1988) applied the kernel technique to the estimation of Engel curves while Pagan and Hong (1988) employed both kernel and Fourier based methods when modeling the conditional variance of equity yields and the excess holding yield on Treasury bills. In all cases the studies seemed to show that considerable gains could be had over existing parametric formulations.

The asymptotic theory justifying these estimators is now well developed. For kernel estimators, conditions under which the estimators are consistent and asymptotically normal are summarized in Bierens (1987), while a recent paper by Andrews (1988) has provided a detailed treatment of consistency and asymptotic normality of Fourier-based estimators if data is independently and identically distributed (i.i.d), although he indicates that his results would extend to dependent observations. Frequently, it emerges from this analysis that the rate of convergence of the estimators is quite slow, and this raises an issue of whether sample sizes are sufficiently large in practice that one could take the asymptotic theory as being an accurate predictor of the sampling distributions of the estimators. Some simulation studies have appeared to examine this question e.g., Chalfont and Gallant (1985) for testing production restrictions and Rilstone and Ullah (1987) for estimating CES production functions, with generally good findings. Nevertheless, much remains to be done before these estimators could be recommended for widespread use.

This paper looks at a slightly different question, namely the relative performance of kernel and Fourier based methods in estimating a number of models that could arise in practice. Section 2 formally defines the variety

of estimators employed in the later simulation work and lists what is to be expected of the estimators. Section 3 considers the comparative performance of both estimators in estimating a CES production function, with the design of the study being derivative from White (1980). As a special case one of the variables (the capital stock) is fixed, leaving only an univariate relation between output and labor to be determined, and this produces a situation reminiscent of Engel curve studies. In section 4 we consider the estimation of the risk premium by non-parametric methods as advocated in Pagan and Ullah (1988). This experiment is of interest in its own right, but is also meant to be representative of a range of situations which are "semi-parametric," and in which the rate of convergence of the estimators of the set of parameters of interest is the same as for parametric models.

The range of models employed here is small and the purpose of the paper is to try to get some insight into how well the estimators work, whether one is superior to another in different circumstances, and what theoretical issues arise from the results. Thus our purpose is a very modest one, but we think that the results are of some interest and fill a gap in the existing literature.

2. The Estimators

Consider an unknown relation between a variable y_i and another variable x_i:

$$y_i = g(x_i) + u_i \qquad (1)$$

where u_i, $(i = 1,\ldots,N)$, is assumed i.i.d. $(0, \sigma^2)$ and is a martingale difference with respect to $\mathcal{F}_i = \{y_{i-1}, y_{i-2},\ldots, x_i, x_{i-1},\ldots\}$. For simplicity results are stated in terms of a scalar x_i, with comments being made, where necessary, about what modifications are needed when x_i is a vector of variables.

The objective is to estimate $m^*=g(x^*)$ i.e., the conditional expectation $E(y|x = x^*)$. If x_i took only discrete values $x_1^*,\ldots,x_p^*$, a simple estimator of $E(y|x = x_j^*)$ is the sample average

$$\hat{m}^* = (N_j^*)^{-1} \sum_{i\epsilon S_j^*} y_i \qquad j = 1,\ldots, p \qquad (2)$$

where S_j^* is the set of i corresponding to x_j^* and N_j^* is the number in this set. Provided $N_j^* \to \infty$, this would be a consistent and asymptotically normal estimator of $g(x_j^*)$ since $y_i(i\epsilon S_j^*)$ has $y_i \sim$ i.i.d.$(g(x_j^*), \sigma^2)$. Let us define the indicator function $I(z_i)$ as being unity if $-1/2 \leq z_i \leq 1/2$ and zero otherwise, and set $z_i = ((x_i - x_j^*)/h)$, where h is chosen to be smaller than one half of the minimum distance between any of the $x_j^*(j = 1,\ldots,p)$ so that $I(z_i)= 1$ only when $x_i = x_j^*$. With these definitions (2) could be re-written as

$$\hat{m}^{*} = \left[\sum_{i=1}^{N} I_i\right]^{-1} \sum_{i=1}^{N} I_i y_i, \qquad (3)$$

since the restriction on h ensures that only those observations corresponding to $x_i = x_j^{*}$ actually appear in the sample average (all others have $I_i = 0$).

When x_i is a continuous random variable the indicator function approach above cannot apply since x_i assumes particular values with probability zero. Nadaraya (1964) and Watson (1964) therefore suggested that I(z) be replaced by a continuous function of z, K(z), but with the same properties that $K(z) \geq 0$, $\int K(z)dz = 1$ (in fact it is also necessary that K(z) be symmetric but the heuristic argument given here to justify the kernel estimator cannot easily capture that requirement). (3) then becomes

$$\hat{m}^{*} = \left(\sum_{i=1}^{N} K_i\right)^{-1} \sum_{i=1}^{N} K_i y_i. \qquad (4)$$

Since these properties characterize density functions, there is a wide range of possible choices for K. Popular ones are the Gaussian kernel $K(z) = (2\pi)^{-1/2} \exp(-(1/2)z^2)$ and the Epanechnikov kernel $K(z) = (3/4h)(1 - z^2)I(|z| \leq 1)$ where $I(z_i)$ is an indicator function and $z_i = ((x_i - x^{*})/h)$. The former has been used in Rilstone and Ullah (1987) and Pagan and Hong (1988) while the latter was adopted by Deaton (1988). As observed by many authors the choice of a kernel is not as crucial as the choice of h, the bandwidth. Of course in the discrete case it was obvious what h should be, but in the continuous case $g(x^{*})$ will need to be estimated by "borrowing" observations on y_i corresponding to x close to x^{*}, where "closeness" is measured by the size of h. Clearly, if h is made too large one is getting a blurred image of $g(x^{*})$

(technically the bias will be large), whereas if h is made too small few observations on "x_j^*" will be available and so the variability of the estimator will be high. For this reason the selection of h can be quite important to the properties of an estimator.

The multivariate case where x_i is a q x 1 vector is easily dealt with by replacing univariate densities with multivariate ones. Hence the multivariate Gaussian kernel is $(2\pi)^{-q/2}|H|^{-1/2}\exp(-\frac{1}{2}(x_i - x^*)'H^{-1}(x_i - x^*))$, while the multivariate Epanechnikov is proportional to $(1-z'z)I(z'z \leq 1)$, where $z=(x_i-x^*)/h$ Notice from (4) that factors such as $|H|^{-1/2}$ appear on both the numerator and denominator of (4) and so disappear. In later work H is set to $\text{diag}(N^{-1/(4+q)}\hat{\sigma}_j)$, where $\hat{\sigma}_j (j=1,\ldots,q)$ is the sample standard deviation of the j'th element in x_i. The optimal choice of bandwidth has been extensively discussed in Bierens (1987), and found to be $cN^{-1/(4+q)}$, so that we are just choosing c to be the sample standard deviation. There are many ways of selecting h, either automatically (as in cross-validation) or by intervention ("eyeballing"), but the computational burdens of simulation mean that a fixed bandwidth is necessary. Further discussion is available in Singh and Ullah (1985).

The kernel is an easy way to estimate $g(x^*)$ and can be done by data transformations allied with some way of computing sums. Most standard econometric programs are therefore well designed to calculate it, and in programs such as GAUSS it is very easily implementable.

Perhaps the major problem with the kernel estimator is that asymptotically it is not centered at $g(x^*)$ unless special care is taken i.e., the limiting theory implies that $(Nh^q)^{1/2}(\hat{m}^* - E(\hat{m}^*))$ is asymptotically normal

with a zero mean, but $(Nh^q)^{1/2}(E(\hat{m}^*) - g(x^*))$ need not tend to zero. In theory it is possible to force this latter term to converge to zero by choosing h such that $\lim_{N \to \infty} h^2(Nh^q)^{1/2} \to 0$ (see Bierens (1987)) but, in practice, for observations in the 200 or so range this makes no difference to the results i.e., the bias remains.

As will be seen in the next section this bias can be a problem for the kernel estimator, and the formulation of methods to ameliorate it need to have high priority. A number of suggestions have been made. Schucany and Sommers (1977) used a jacknifing method that entails the averaging of two estimators with different bandwidths. Thus if h_1 is set to $cN^{-1/(q+4)}$ and h_2 to $cN^{-\delta/(q+4)}(\delta\epsilon(0,1))$, with corresponding estimates $\hat{m}_1^*$ and $\hat{m}_2^*$, the estimator

$$\hat{\hat{m}}^* = \{\hat{m}_1^* - N^{-2(1-\delta)/(q+4)}\hat{m}_2^*\}/(1 - N^{-2(1-\delta)/(q+4)}\} \qquad (5)$$

is asymptotically centered on $g(x^*)$.

An alternative proposal involves the construction of "higher order bias reducing" kernels that are linear combinations of a basic kernel and which can eliminate bias under certain conditions. Generally when K(z) is chosen in this way it is not possible to ensure that it is non-negative and hence a density function. Bartlett(1963) initiated this idea and Robinson(1987) has recently used it in our context. To construct such a kernel let K'(z) be a base kernel e.g., the Gaussian, and define (ℓ=2,4,6,...)

$$K(z) = \sum_{j=0}^{1/2(\ell-2)} c_j z^{2j} K'(z), \qquad (6)$$

where c_j are solutions to the system of $1/2(\ell-2) + 1$ linear equations $\sum_{j=0}^{1/2(\ell-2)} c_j \mu_{2(i+j)} = \delta_{i0}$, $0 \leq i \leq 1/2(\ell-2)$ (δ_{i0} is the Kronecker delta) and $\mu_{2j} = \int z^{2j} K'(z)dz$. Table 1 below gives values of c_j for the Gaussian kernel and a range of values for ℓ. The idea behind this suggestion is that the Taylor series expansion of $E(\hat{m}^*)-m^*$ around x^* involves a power series in h, and the coefficients of h^j will be zero if all moments of the kernel up to and including the j'th are zero. When ℓ is the order of the first non-zero moment of the kernel, $(Nh^q)^{1/2}(E(\hat{m}^*)-m^*)$ will converge to zero provided $(Nh^q)^{1/2}h^\ell \to 0$ as $N \to \infty$. Putting $h=N^{-(1/(q+4))}$ will satisfy this restriction for any $\ell > 2$.

Table 1

Values of c to Construct Kernel in (6)

	$\ell = 2$	$\ell = 4$	$\ell = 6$	$\ell = 8$
c_0	1	1.5	1.875	2.1875
c_1		-0.5	-1.25	-2.1875
c_2			0.125	0.4375
c_3				-0.02083

Another variant of non-parametric estimation is the Fourier method advocated by Gallant (1982). In this (1) is approximated by a quadratic polynomial in x_i and a linear combination of cosine and sine terms in x_i

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{j=1}^{M} (\gamma_j \cos jx_i + \delta_j \sin jx_i) + v_i . \quad (7)$$

The formulation has its origins in approximation theory where it is known that a Fourier series can approximate $g(x_i)$ arbitrarily well over the range $(0,2\pi)$ (in fact to use this method data must be scaled to be in this range). For economic analysis the Fourier technique has the extra advantage that it also approximates derivatives of $g(x_i)$ with respect to x_i arbitrarily well, and this is important as these derivatives correspond to economic quantities of interest such as the marginal product. Provided M is chosen such that $M \to \infty$ as $N \to \infty$ and $g(\cdot)$ is sufficiently smooth, the estimator of $g(x^*)$ formed as the OLS predictions from (7) is consistent. Andrews(1988) shows that, if M is set to cN^r, where c is a constant and $0<r<1$, $(N^{1-r})^{1/2}(\hat{g}(x^*)-g(x^*))$ is

asymptotically normal. The choice of r is therefore akin to that of the bandwidth in kernel regression. If the problem of bias is ignored or eliminated by bias reducing methods $N^{2/(q+4)}(\hat{m}*-m*)$ is asymptotically normal, and writing the normalizing factor as $(N.N^{(-q/(q+4))})$ suggests that r=(q/q+4) would be appropriate when comparing kernel and Fourier estimators. It is interesting to observe that the Fourier approach does not suffer the bias problems of the kernel (in theory) and this makes it attractive. Eastwood and Gallant (1987) however found that the rate of convergence to asymptotic normality could be very slow unless M was chosen as a function of the data. In our later experiments we have pre-assigned M, but the question of a "best choice" clearly deserves further investigation.

When x_i is multivariate of order q the first three terms in (7) are a quadratic form in the vector x_i i.e., they would be $\beta_0 + x_i'\beta + x_i'Bx_i$ where β is a (q x 1) vector and B is a (q x q) matrix. The Fourier terms are made functions of scalars $k_j'x_i$ where k_j are (q x 1) vectors termed "multi-indexes". The first q of these are the q elementary vectors, and an algorithm for constructing them is described in Gallant (1982) with a computer code available in Monahan (1981). An important feature of the Fourier procedure is that it is asymptotically unbiased and therefore has a potential advantage over the kernel estimator. Against this is the fact that the number of parameters to be estimated in (7) multiply rapidly as the dimension of x_i rises, and this may limit its use to small dimensional problems. In fact, past users, such as Gallant, overcome this problem of a profligate number of parameters by estimating systems of equations and imposing cross-equation constraints from economic theory.

The above discussion has concentrated upon estimating a conditional mean, but there are models estimated in econometrics which demand that a conditional variance be estimated. These are models which incorporate a risk premium and which have the generic form

$$y_i = z_i\beta + \sigma_i^2\delta + e_i \qquad (8)$$

where z_i is weakly exogenous for β and δ and σ_i^2 represents the variance of some random variable ψ_i conditional upon an information set $\mathcal{F}_i$. Pagan and Ullah (1988) survey the ways in which it has been suggested that the unobserved variable σ_i^2 be replaced by a function of data. One popular solution has been to assume that the conditional expectation of ψ_i is a linear function of some weakly exogenous variables w_i, that is $E(\psi_i|\mathcal{F}_i) = w_i\gamma$, and to then assume σ_i^2 is a function of ϕ_{i-j}^2, where $\phi_i = \psi_i - E(\psi_i|\mathcal{F}_i)$. Engle's (1982) ARCH model for example sets $\sigma_i^2 = \alpha_0 + \sum_{j=1}^{r} \alpha_j\phi_{i-j}^2$.

Pagan and Ullah (1988) advance an estimation strategy that seeks to avoid the parameterization of σ_i^2. The basic idea advanced in that paper was to replace σ_i^2 by $\phi_i^2 = (\psi_i - E(\psi_i|\mathcal{F}_i))^2$, and to then consistently estimate β and δ by applying IV with instruments constructed from $\mathcal{F}_i$ (including x_i as its own instrument). Of course instrument construction needs to be done carefully, since they must be as highly correlated with ϕ_i^2 as possible. In fact, because $E(\phi_i^2|\mathcal{F}_i) = \sigma_i^2$, σ_i^2 appeals as a good instrument, making it desirable to estimate σ_i^2 for this purpose. An important facet of the IV strategy is that it only requires $m_i = E(\psi_i|\mathcal{F}_i)$ to be estimated accurately; truncation of the instrument set to (say) $\bar{\sigma}_i^2 = E(\phi_i^2|\mathcal{F}_i')$, where $\mathcal{F}_i' \in \mathcal{F}_i$, does not affect the consistency of the estimators of β and δ, although efficiency might be

affected if $\bar{\sigma}_i^{-2}$ has only a weak correlation with ϕ_i^2. Overall, there is a trade-off between maximizing instrument correlation and restricting $\mathcal{F}_i'$ to be small enough to get estimates $\bar{\sigma}_i^{-2}$ that are not "too noisy".

Basically the IV approach requires the computation of $m_i = E(\psi_i/\mathcal{F}_i)$ and some estimate of σ_i^2 (although as noted above it is not crucial that the latter be accurate). As described earlier, m_i might be linear in the members of $\mathcal{F}_i$ i.e., $w_i\gamma$, and this is a frequent assumption in VAR modeling. If so $\hat{m}_i$ can be estimated as the predictions from the regression of ψ_i against w_i. However, in some instances it may be desirable to allow m_i to be a non-linear function of $\mathcal{F}_i$, and then it could be estimated in the same way as $g(x^*)$ was estimated earlier. Whichever way an estimate for m_i, $\hat{m}_i$, is generated, $\hat{\phi}_i^2 = (\psi_i - \hat{m}_i)^2$ replaces σ_i^2 and an instrument $\hat{\sigma}_i^2$ is used for $\hat{\phi}_i^2$. In section 4 we use as instruments the non-parametric estimate of the conditional variance of ψ_i given $\mathcal{F}_i$.

Now although neither $\hat{m}_i$ nor $\hat{\sigma}_i^2$ is likely to be estimated very precisely unless $\dim(\mathcal{F}_i)$ is small or the number of observations is large, these quantities are only being used as "regressors" in (8), and so the IV estimator of β and δ should be root-N consistent and asymptotically normal. To get this result, following Pagan and Ullah (1988) it is necessary to show that $N^{-1/2} \Sigma z_i(\hat{\phi}_i^2 - \phi_i^2) \xrightarrow{p} 0$, where z_i are the chosen instruments. When the conditional mean is parameterized as $w_i\gamma$ this is true, as discussed in Pagan(1984) and exploited in tests for heteroskedasticity, but no general proof currently exists when ϕ_i is estimated non-parametrically,particularly when ψ_i is a time series. An alternative estimator, if $\mathcal{F}_i' = \mathcal{F}_i$, might be to regress y_i against x_i and $\hat{\sigma}_i^2$, as this would give a consistent estimator of δ, β, but the standard errors would need to be adjusted because of a "generated regressor" bias

arising from the use of $\hat{\sigma}_i^2$ rather than σ_i^2. Just as in the case when $\hat{\sigma}_i^2$ is constructed parametrically from $\mathcal{F}_i$, Pagan (1984), the appropriate adjustment may be difficult to do directly, but is easily done with the IV procedure. Hence, estimation and inference is naturally performed within the IV framework.

The final issue to be addressed concerns which non-parametric method should be used to construct $\hat{m}_i$ and $\hat{\sigma}_i^2$. In section 4 we employ both the kernel based approach and the Fourier method. Both are easy to apply and have advantages and disadvantages. It should be noted that for the Fourier technique, $\hat{\sigma}_i^2$ is estimated as the predictions obtained by fitting a Fourier approximation to $(\psi_i - \hat{m}_i)^2$, since $E(\psi_i - m_i)^2 = \sigma_i^2$ is the basis of such a regression.

3. Kernel and Fourier Estimation of a CES Production Function

White (1980) set up a simulation experiment to evaluate the ability of flexible functional forms to adequately approximate unknown regression functions. He simulated data from a C.E.S. production function

$$y_i = (L_i^{\rho} + 2K^{\rho})^{1/\rho} + u_i \ . \tag{9}$$

In our first experiment K was fixed at the value 1.71, log L_i was distributed uniformly over [0, 1], u_i was i.i.d. (0, 0.01) and ρ was set to be -5. This gives an unknown functional relation between y_i and $x_i = L_i$. The variance of u_i was chosen so as to get an R^2 of around .9 when (9) is fitted with both L_i and K_i.

Our aim is to estimate the functional relation $g_i = (L_i^{-5} + 2(1.71)^{-5})^{-(1/5)}$ at selected points $L^* = 1.284, 1.568, 1.852, 2.136, 2.420$. A variety of estimators was applied. Because of the bias problem the kernel estimator was computed with jack-knifing ($\delta = .25, .5$ and $.75$) as in (5) and higher order kernels ($\ell = 4, 6, 8$) as in (6). The number of terms in the Fourier approximation (7) (M) was set to 2, making it independent of sample size. Three sample sizes N = 60, 100 and 200 were used and 500 replications performed.

Fig. 1, 2 and 3 graph g_i evaluated at the five points L^* detailed above along with the average estimates provided by the standard kernel(ℓ=2), jack-knifing (δ =.75), a higher-order kernel (ℓ = 8) and Fourier series estimators; the values of δ and ℓ represented the choices with smallest average bias. The bias problem with standard kernel estimators is immediately apparent, but neither of the bias reducing procedures is particularly successful in samples of size 60 , with noticable bias at the lower end of the grid. As the sample grows the higher order kernel approach improves but the jacknifed estimator does not. Of course for this set of parameters the relative magnitude of the bias is small: what is important is its resilience to an increase in sample size. With other parameter combinations the magnitude of the bias could probably be made much greater.

In contrast to the poor performance of the kernel estimators the Fourier method exhibits very small bias at all points in the L-space. In fact, simulations with N as low as 30 did not change this conclusion. Hence, for problems characterized by a single regressor in x_i, the Fourier estimator would seem to produce estimates of m^* that are much closer to the true values than any of the kernel based procedures. Of course one might improve upon the

kernel estimates by using variable bandwidths or selecting these by cross-validation: what the experiment emphasises is that bias is a potentially serious problem for kernel methods.

When the complete production function is estimated, results are rather similar. In this instance log K_i is distributed uniformly over [0, 1] and a grid of 25 points was taken corresponding to all bivariate combinations of the points (1.284, 1.568, 1.852, 2.136, 2.420) for L^* and K^*. Tables 2, 3 and 4 show the average estimates of m^*obtained for three of the combinations (L^*, K^*) = (1.284, 1.284), (1.852, 1.852) and (2.420, 2.420). A complete tabulation is available on request. Generally, the Fourier approximation does reasonably well, although it is noticeable that the bias does not reduce very much with the sample size and it may be that M should have been increased. The basic kernel estimator ($\ell = 2$) is not dominated by the Fourier estimator as much as it was when x_i was a scalar, while the bias reduction methods perform well for large values of K and L but remain very poor at the left end of the grid. Indeed the best strategy might well be to stay with the raw kernel, as some of the higher-order kernel estimators ($\ell > 2$) are incredibly variable, with extreme outliers in a number of replications.

These limited experiments emphasize the problem of asymptotic bias in the kernel estimator and, at least for a small number of conditioning variables, suggest that Fourier methods may well be the best way to approximate $g(\cdot)$. What is very disappointing in the results is the relatively poor performance of methods that have been proposed to reduce the bias of standard kernel estimators. Perhaps these methods would work better in larger sample sizes, but it is obvious by comparing the first columns of Tables 2 and 4 that the improvement seems very slight,even when the sample size tripled. A much more

detailed comparison between the estimators as well as investigation of other ways to reduce bias is therefore warranted.

4. Risk Estimation and ARCH Models

In these experiments the parameter δ in equation (8), with $\beta = 0$ and so z_i excluded, was estimated by non-parametric methods . The conditional variance of ψ_i, σ_i^2 , followed variants of Engle's (1982) ARCH model, while $\psi_i = \mu + \phi_i$. The density of the error term ϕ_i, conditional upon past information, was taken to be $N(0, \sigma_i^2)$, with δ , the location parameter μ , and σ_e^2 all being set to unity. Two experiments were performed ,with σ_i^2 being either an ARCH (1) or an ARCH (4) of the type used by Engle et al (1987) when modeling the excess holding yield on U.S. Treasury Bills.

$$\sigma_i^2 = \alpha_0 + \alpha_1 \phi_{i-1}^2 \qquad (10)$$

$$\sigma_i^2 = \alpha_0 + \alpha_1 \sum_{j=1}^{4} ((5 - j)/10) \phi_{i-j}^2 \qquad (11)$$

Four estimators of δ were considered.

A. Estimate σ_i^2 by a kernel estimator, with conditioning variables $\psi_{i-1}, \ldots\ldots, \psi_{i-r}$,r being the order of the ARCH process, and regress y_i against $\hat{\sigma}_i^2$.

B. Replace σ_i^2 in (8) by $\hat{\phi}_i^2 = (\psi_i - \hat{\mu})^2$, $\hat{\mu}$ being the sample mean of ψ_i, and then do instrumental variables with $\hat{\sigma}_i^2$ as instrument for $\hat{\phi}_i^2$.

C. Fit the Fourier approximation to (7) with y_i replaced by $\hat{\phi}_i^2$, take the predictions, $\tilde{\sigma}_i^2$, as an estimator of σ_i^2, and regress y_i against $\tilde{\sigma}_i^2$.

D. Fit the parameters α_0, α_1 in (10) and (11) and μ using M.L.E., compute $\bar{\sigma}_i^2$ and regress y_i against $\bar{\sigma}_i^2$.

Estimator D is meant to be a benchmark since it would be as efficient as any two-stage estimator could be, in that it exploits the fact that the conditional variance is generated by an ARCH process and the conditional density is normal. The Fourier approximation might be expected to be a good one when r = 1 since $\sigma_i^2 = \alpha_0 + \alpha_1 (\psi_{i-1} - \mu)^2 = \alpha_0 + \alpha_1 \psi_{i-1}^2 - 2 \alpha_1 \mu \psi_{i-1} + \alpha_1 \mu^2$, and therefore the quadratic terms that lead the Fourier approximation are all that is needed to estimate σ_i^2. However, it will lose something owing to the addition of unnecessary trigonometric terms. Because of this feature M was varied with the sample size, being 1 for N = 30, 2 for N = 60, 100, and 3 for N = 200. In the case of equation (11) only the diagonal terms in the quadratic form derived from $(\psi_{i-1}, \psi_{i-2}, \psi_{i-3}, \psi_{i-4})$ were retained.

Table 5 gives the mean and standard deviations (in parentheses) of the estimated δ from the four estimators for α_1 = .2 and .5, σ_i^2 as in (10), and for sample sizes N = 60, 100 and 200. Table 6 gives comparable results when σ_i^2 is from (11). The kernel based estimators do quite well in estimating δ, although it is noticeable that there is a small sample bias which is declining only slowly and that this worsens as α_1 rises. However, the move from a single (ψ_{i-1}) to four conditioning variables $(\psi_{i-1}, \ldots, \psi_{i-4})$ did not affect the performance of the estimator at all. This is to be expected from the theory, arising from the fact that the non-parametric estimators are

effectively being averaged, and therefore root-N consistency should hold. It is hard to be certain if this prediction is correct or not, but the ratio of the standard error of $\hat{\delta}$ at N = 100 to that at N = 200 is generally around 1.2 - 1.3, so that the variance would be in ratio 1.4 - 1.7, compared to the theoretical requirement of $\sqrt{2}$. It is also noticeable that the benchmark estimator D gives good results and it is certainly more efficient than any of the kernel methods, although the efficiency loss might be acceptable given the gretaer robustness of the kernel procedures to specification errors in σ_t^2.

Perhaps the most disappointing result from Tables 5 and 6 was the poor performance of the Fourier based estimator. When r=4 this might be expected since a very large number of terms appear in the expansion. But the fact that δ is poorly estimated when r=1, relative to that from the kernel and ARCH methods, when r=1 is surprising, as it has already been noted that performance might have been expected to be good in this context, in that one is merely adding on superflous trigonometric regressors in the first stage when estimating σ_i^2. In fact, if the trigonometric terms are dropped i.e. σ_i^2 is estimated by regressing $\hat{\phi}_i^2$ against a constant, ψ_{i-1} and ψ_{i-1}^2, the average values of $\hat{\delta}$ for the three sample sizes were 1.0102, .995 and .996 (α_1=.2) and 1.1079, 1.0922 and 1.0837 (α_1=.5). Hence the bias stems directly from the addition of the Fourier terms. No simple explanation of this outcome could be found and it clearly will repay further study.

5. Conclusion

This paper has tried to present some evidence upon the relative performance of kernel and Fourier based non-parametric estimators. When estimating a conditional mean theoretical considerations point to an asymptotic bias problem for kernel methods, and this was borne out in our

experiments with a production function. More disturbing however was the fact that some of the methods suggested to overcome this problem were not very effective, and in some instances the bias in the estimated mean could be worse than if no adjustment was made at all. The bias problem afflicting the kernel estimator was not present for the Fourier method and this must make it a strong candidate when the context requires the estimation of a conditional mean. If kernel methods are to be used better procedures for bias correction will need to be developed.

In contrast to this outcome, when the non-parametric estimators of a conditional variance were "averaged" by their employment in a regression to estimate the effects of risk, the kernel estimators were much superior to the Fourier ones. It is significant that, in this semi-nonparametric setting, the OLS estimator of the risk coefficient exhibits no asymptotic bias, even when the kernel method is used to generate a conditional variance for use as a proxy for the missing true variance. Hence, kernel procedures look attractive in this situation, although a wider range of specifications will need to be examined before one could confidently recommend them.

Our experiments show that non-parametric estimation may be a useful tool in econometric analysis, even in relatively small sample sizes. We have also found that there are differences in performance between the different non-parametric estimators and it will be important to determine the most effective one in any particular case. Certainly a lot more needs to be done so as to attain a good understanding of the performance of various estimators before one is likely to see them in widespread application.

Fig.1 Nonparametric Esimates of

a CES Production Function{N=60}

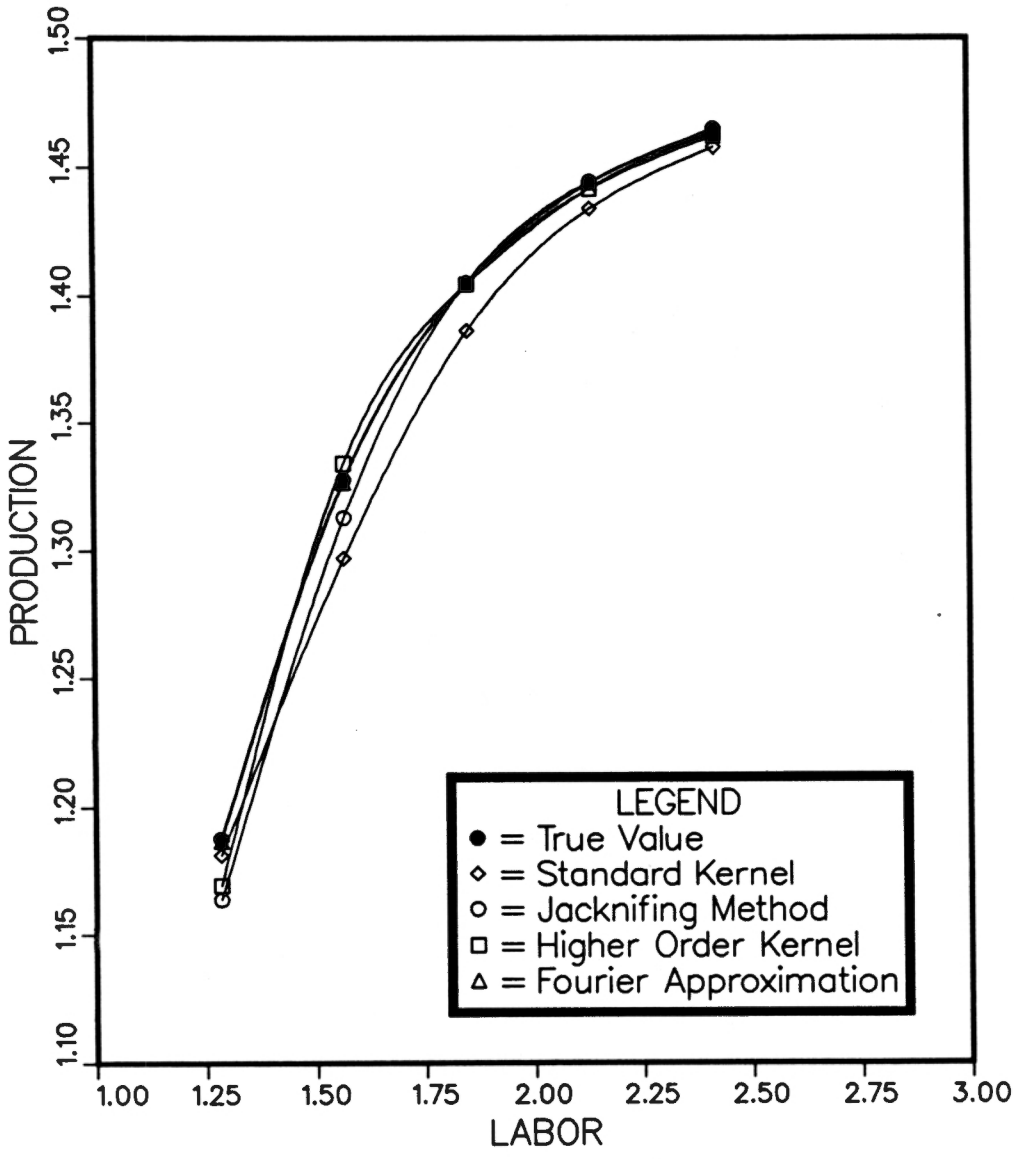

Fig.2 Nonparametric Esimates of

a CES Production Function{N=100}

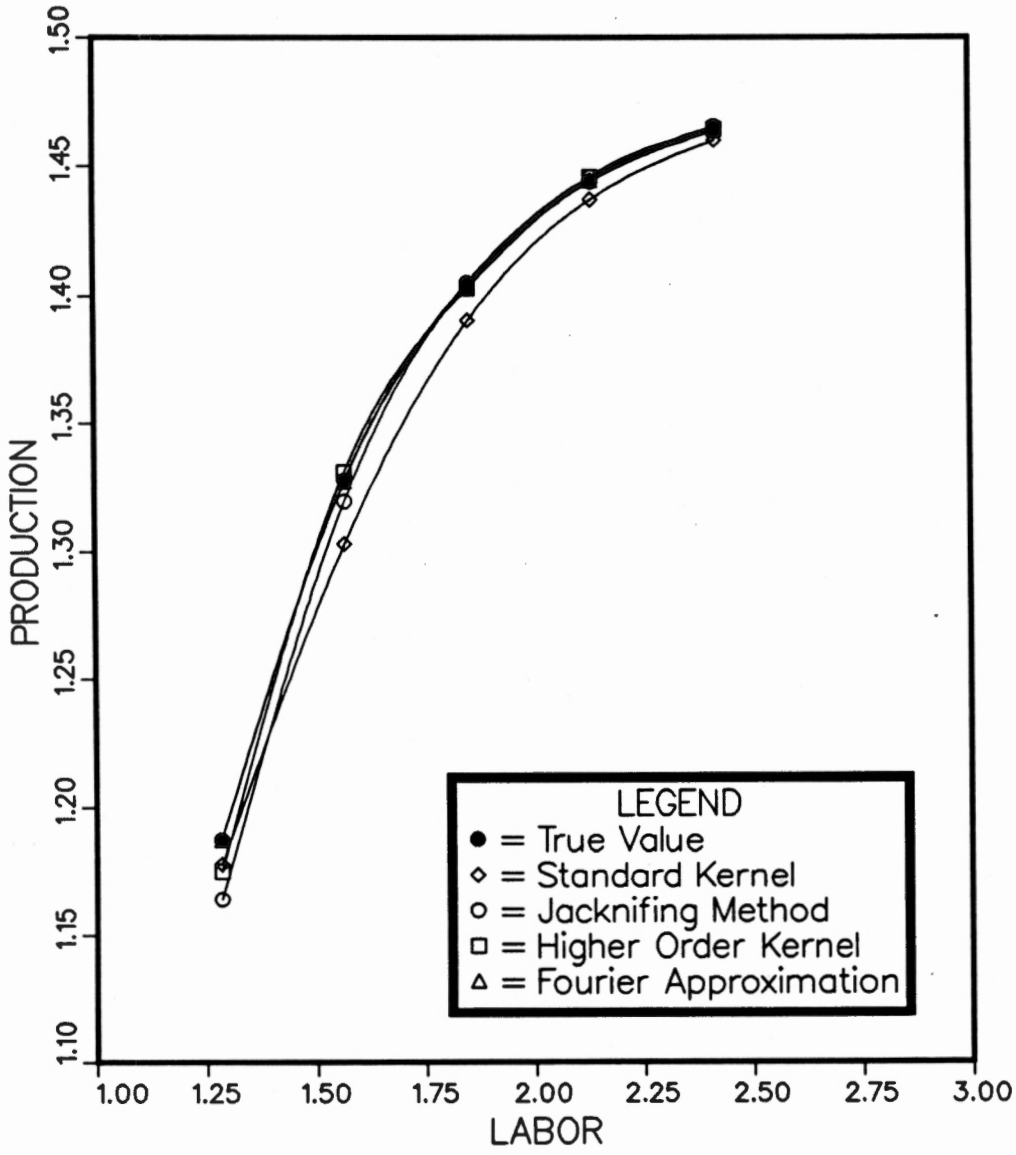

Fig.3 Nonparametric Esimates of

a CES Production Function{N=200}

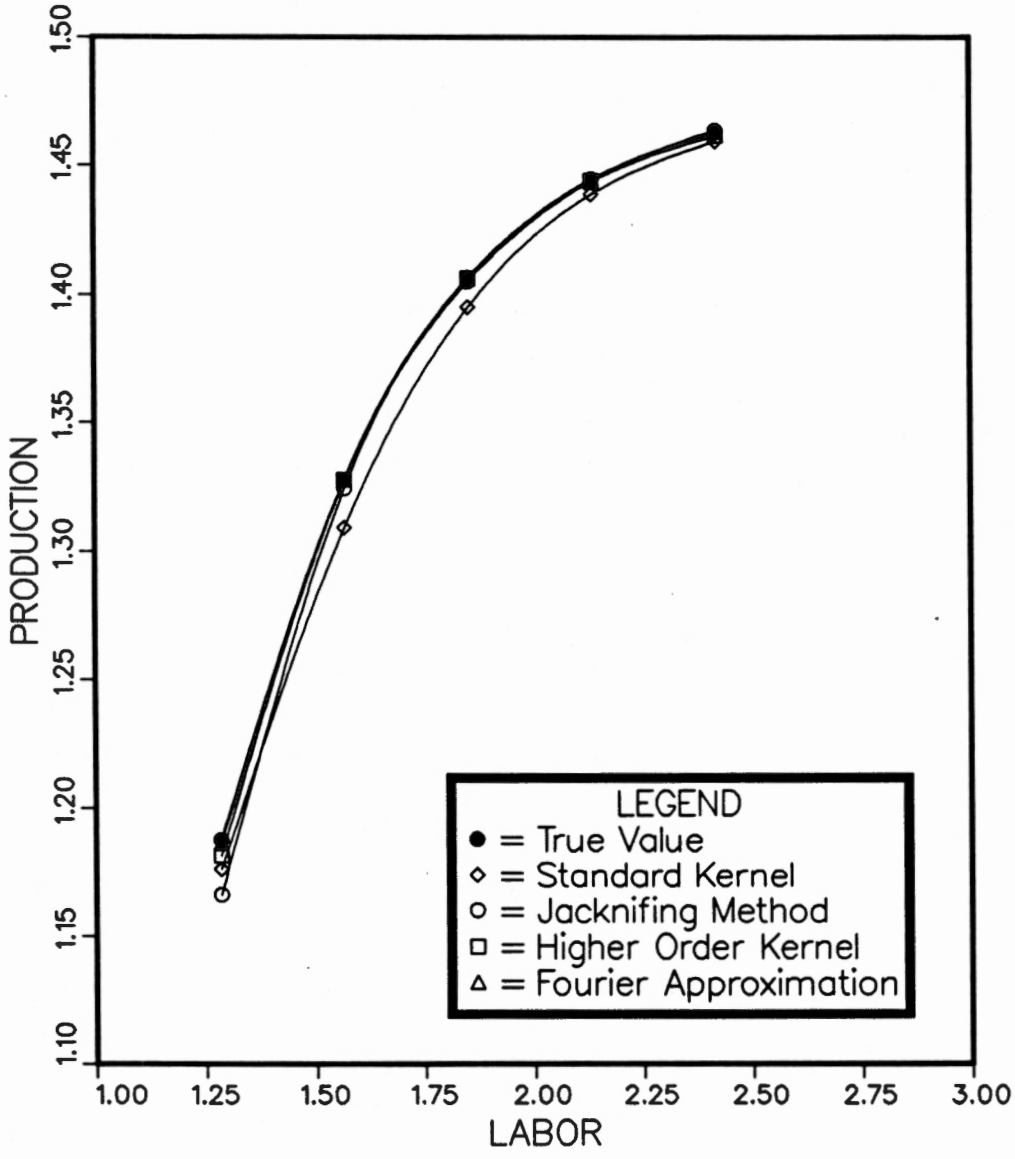

Table 2

Average Estimates of $g(\cdot)$ for (9), Various Estimators, N = 60.

	$L^* = K^* = 1.284$	$L^* = K^* = 1.852$	$L^* = K^* = 2.420$
$g(\cdot)$	1.0307	1.4867	1.9426
Fourier	1.0207	1.4758	1.9479
$\delta = .25$	1.0000	1.4634	1.9303
$\delta = .5$	.9979	1.4713	1.9339
$\delta = .75$	.9964	1.4776	1.9364
$\ell = 2$	1.0264	1.4231	1.8520
$\ell = 4$	.9871	1.6257	2.0040
$\ell = 6$	.9814	1.4980	1.9603
$\ell = 8$	3.4396	1.3363	1.5865

Table 3

Average Estimates of $g(\cdot)$ for (9), Various Estimators, N = 100

	$L^* = K^* = 1.284$	$L^* = K^* = 1.852$	$L^* = K^* = 2.420$
$g(\cdot)$	1.0307	1.4867	1.9426
Fourier	1.0197	1.4777	1.9336
$\delta = .25$	1.0005	1.4680	1.9413
$\delta = .5$	.9989	1.4753	1.9446
$\delta = .75$	.9980	1.4805	1.9470
$\ell = 2$	1.0225	1.4317	1.8735
$\ell = 4$	.9917	1.6828	1.9871
$\ell = 6$	.9950	1.4701	1.9456
$\ell = 8$	1.0056	1.5267	1.8561

Table 4

Average Estimates of $g(\cdot)$ for (9), Various Estimators, N = 200

	$L^* = K^* = 1.284$	$L^* = K^* = 1.852$	$L^* = K^* = 2.420$
$g(\cdot)$	1.0307	1.4867	1.9426
Fourier	1.0196	1.4747	1.9376
$\delta = .25$	1.000	1.4728	1.9498
$\delta = .5$	.9989	1.4789	1.9532
$\delta = .75$	.9991	1.4822	1.9556
$\ell = 2$	1.0176	1.4421	1.8939
$\ell = 4$	.9957	1.5635	1.9935
$\ell = 6$	1.0053	1.5706	1.9862
$\ell = 8$	1.0287	1.5497	1.9336

Table 5

Mean and Standard Deviation of Estimators of δ in (8)

$$\sigma_i^2 = \alpha_0 + \alpha_1 \phi_{i-1}^2.$$

	Estimator			
	A	B	C	D
$\alpha_1 = .2$				
N = 60	1.0728	1.1223	.8749	1.0235
	(.2392)	(.2578)	(.2459)	(.1827)
N = 100	1.0458	1.0718	.8902	1.0124
	(.1782)	(.1940)	(.2117)	(.1393)
N = 200	1.0269	1.0388	.9012	1.0126
	(.1246)	(.1351)	(.1626)	(.0986)
$\alpha_1 = .5$				
N = 60	1.2076	1.1960	.8963	1.0570
	(.2435)	(.2923)	(.2734)	(.2047)
N = 100	1.1589	1.1433	.8939	1.0364
	(.1984)	(.2630)	(.2599)	(.1748)
N = 200	1.1223	1.0901	.8606	1.0305
	(.1600)	(.1882)	(.2341)	(.1304)

Table 6

Mean and Standard Deviation of Estimators of δ in (8)

$$\sigma_i^2 = \alpha_0 + \alpha_1 \sum_{j=1}^{4} ((5 - j)/10)\, \phi_{i-j}^2 .$$

	Estimator			
	A	B	C	D
$\alpha_1 = .2$				
N = 60	1.0497	1.1176	.6978	.9954
	(.2818)	(.2943)	(.2027)	(.1914)
N = 100	1.0181	1.0603	.7758	1.0015
	(.2036)	(.2169)	(.1780)	(.1404)
N = 200	.9980	1.0274	.8521	1.0048
	(.1378)	(.1418)	(.1302)	(.0930)
$\alpha_1 = .5$				
N = 60	1.1686	1.1442	.7016	.9991
	(.2793)	(.2927)	(.2023)	(.1847)
N = 100	1.1344	1.0786	.7648	1.0074
	(.2009)	(.2135)	(.1892)	(.1420)
N = 200	1.1008	1.0346	.8234	1.0157
	(.1460)	(.1553)	(.1649)	(.0975)

References

Andrews, D. (1988), "Asymptotic Normality of Series Estimators for Various Nonparametric and Semiparametric Models," (mimeo, Yale University).

Bartlett, M.S. (1963), "Statistical Estimation of Density Functions", *Sankhya*, A, 25, 145-154.

Bierens, H. (1987), "Kernel Estimators of Regression Functions," Ch. 3 in T. Bewley (ed.) *Advances in Econometrics, Fifth World Congress*, Vol. I, 99-144.

Chalfont, J.A. and A.R. Gallant (1985), "Estimating Substitution Elasticities with the Fourier Cost Function," *Journal of Econometrics*, 28, 205-222.

Deaton, A. (1988), "Agricultural Pricing Policies and Demand Patterns in Thailand," (mimeo, Princeton University).

Eastwood, B.J. and A.R. Gallant (1987), "Adaptive Truncation Rules for Seminonparametric Estimators that Achieve Asymptotic Normality," (mimeo, North Carolina State University).

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometricia*, 50, 987-1008.

Engle,R.F.,D.M.Lillien and R.P.Robins, "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model", *Econometrica*,55,391-408.

Gallant, A.R. (1982), "Unbiased Determination of Production Technologies," *Journal of Econometrics*, 20, 285-323.

Monahan, J.F. (1981), "Enumeration of Elementary Multi-Indexes for Multivariate Fourier Series," Institute of Statistics Mimeograph Series No. 1338, North Carolina State University.

Nadaraya, E.A. (1964), "On Estimating Regression," *Theory of Probability and its Applications*, 9, 141-142.

Pagan, A.R. (1984), "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review*, 25, 221-247.

Pagan, A.R. and A. Ullah (1988), "The Econometric Analysis of Models with Risk Terms," *Journal of Applied Econometrics*,3,87-105.

Pagan, A.R. and Y. Hong (1988), "Non-parametric Estimation and the Risk Premium, *Rochester Center for Economic Research Working Paper* no.135, University of Rochester

Rilstone, P. and A. Ullah (1987), "Non-parametric Estimation of Response Coefficients," (mimeo, University of Western Ontario).

Robinson, P. (1987), "Root-N-Consistent Semiparametric Regression," Discussion Paper No. R. 9, London School of Economics.

Schucany, W.R. and J.P. Sommers (1977), "Improvement of Kernel Type Density Estimators," *Journal of the American Statistical Association*, 72, 420-423.

Singh, R.S. and A. Ullah (1985), "Nonparametric Time Series Estimation of Joint DGP, Conditional DGP and Vector Autoregression," *Econometric Theory*, 1, 27-52.

Watson, G.S. (1964), "Smooth Regression Analysis," *Sankhya*, Series A, 26, 359-372.

White, H. (1980), "Using Least Squares to Approximate Unknown Regression Functions," *International Economic Review*, 21, 149-170.

Rochester Center for Economic Research
University of Rochester
Department of Economics
Rochester, NY 14627

1987–88 DISCUSSION PAPERS

WP#68 RECURSIVE UTILITY AND OPTIMAL CAPITAL ACCUMULATION, I: EXISTENCE, by Robert A. Becker, John H. Boyd III, and Bom Yong Sung, January 1987

WP#69 MONEY AND MARKET INCOMPLETENESS IN OVERLAPPING–GENERATIONS MODELS, by Marianne Baxter, January 1987

WP#70 GROWTH BASED ON INCREASING RETURNS DUE TO SPECIALIZATION by Paul M. Romer, January 1987

WP#71 WHY A STUBBORN CONSERVATIVE WOULD RUN A DEFICIT: POLICY WITH TIME–INCONSISTENT PREFERENCES by Torsten Persson and Lars E.O. Svensson, January 1987

WP#72 ON THE CONTINUUM APPROACH OF SPATIAL AND SOME LOCAL PUBLIC GOODS OR PRODUCT DIFFERENTIATION MODELS by Marcus Berliant and Thijs ten Raa, January 1987

WP#73 THE QUIT–LAYOFF DISTINCTION: GROWTH EFFECTS by Kenneth J. McLaughlin, February 1987

WP#74 SOCIAL SECURITY, LIQUIDITY, AND EARLY RETIREMENT by James A. Kahn, March 1987

WP#75 THE PRODUCT CYCLE HYPOTHESIS AND THE HECKSCHER–OHLIN–SAMUELSON THEORY OF INTERNATIONAL TRADE by Sugata Marjit, April 1987

WP#76 NOTIONS OF EQUAL OPPORTUNITIES by William Thomson, April 1987

WP#77 BARGAINING PROBLEMS WITH UNCERTAIN DISAGREEMENT POINTS by Youngsub Chun and William Thomson, April 1987

WP#78 THE ECONOMICS OF RISING STARS by Glenn M. MacDonald, April 1987

WP#79 STOCHASTIC TRENDS AND ECONOMIC FLUCTUATIONS by Robert King, Charles Plosser, James Stock, and Mark Watson, April 1987

WP#80 INTEREST RATE SMOOTHING AND PRICE LEVEL TREND–STATIONARITY by Marvin Goodfriend, April 1987

WP#81 THE EQUILIBRIUM APPROACH TO EXCHANGE RATES by Alan C. Stockman, revised, April 1987

WP#82 INTEREST-RATE SMOOTHING
by Robert J. Barro, May 1987

WP#83 CYCLICAL PRICING OF DURABLE LUXURIES
by Mark Bils, May 1987

WP#84 EQUILIBRIUM IN COOPERATIVE GAMES OF POLICY FORMULATION
by Thomas F. Cooley and Bruce D. Smith, May 1987

WP#85 RENT SHARING AND TURNOVER IN A MODEL WITH EFFICIENCY UNITS OF HUMAN CAPITAL
by Kenneth J. McLaughlin, revised, May 1987

WP#86 THE CYCLICALITY OF LABOR TURNOVER: A JOINT WEALTH MAXIMIZING HYPOTHESIS
by Kenneth J. McLaughlin, revised, May 1987

WP#87 CAN EVERYONE BENEFIT FROM GROWTH? THREE DIFFICULTIES
by Herve' Moulin and William Thomson, May 1987

WP#88 TRADE IN RISKY ASSETS
by Lars E.O. Svensson, May 1987

WP#89 RATIONAL EXPECTATIONS MODELS WITH CENSORED VARIABLES
by Marianne Baxter, June 1987

WP#90 EMPIRICAL EXAMINATIONS OF THE INFORMATION SETS OF ECONOMIC AGENTS
by Nils Gottfries and Torsten Persson, June 1987

WP#91 DO WAGES VARY IN CITIES? AN EMPIRICAL STUDY OF URBAN LABOR MARKETS
by Eric A. Hanushek, June 1987

WP#92 ASPECTS OF TOURNAMENT MODELS: A SURVEY
by Kenneth J. McLaughlin, July 1987

WP#93 ON MODELLING THE NATURAL RATE OF UNEMPLOYMENT WITH INDIVISIBLE LABOR
by Jeremy Greenwood and Gregory W. Huffman

WP#94 TWENTY YEARS AFTER: ECONOMETRICS, 1966-1986
by Adrian Pagan, August 1987

WP#95 ON WELFARE THEORY AND URBAN ECONOMICS
by Marcus Berliant, Yorgos Y. Papageorgiou and Ping Wang, August 1987

WP#96 ENDOGENOUS FINANCIAL STRUCTURE IN AN ECONOMY WITH PRIVATE INFORMATION
by James Kahn, August 1987

WP#97 THE TRADE-OFF BETWEEN CHILD QUANTITY AND QUALITY: SOME EMPIRICAL EVIDENCE
by Eric Hanushek, September 1987

WP#98 SUPPLY AND EQUILIBRIUM IN AN ECONOMY WITH LAND AND PRODUCTION
by Marcus Berliant and Hou-Wen Jeng, September 1987

WP#99 AXIOMS CONCERNING UNCERTAIN DISAGREEMENT POINTS FOR 2-PERSON BARGAINING PROBLEMS
by Youngsub Chun, September 1987

WP#100 MONEY AND INFLATION IN THE AMERICAN COLONIES: FURTHER EVIDENCE ON THE FAILURE OF THE QUANTITY THEORY
by Bruce Smith, October 1987

WP#101 BANK PANICS, SUSPENSIONS, AND GEOGRAPHY: SOME NOTES ON THE "CONTAGION OF FEAR" IN BANKING
by Bruce Smith, October 1987

WP#102 LEGAL RESTRICTIONS, "SUNSPOTS", AND CYCLES
by Bruce Smith, October 1987

WP#103 THE QUIT-LAYOFF DISTINCTION IN A JOINT WEALTH MAXIMIZING APPROACH TO LABOR TURNOVER
by Kenneth McLaughlin, October 1987

WP#104 ON THE INCONSISTENCY OF THE MLE IN CERTAIN HETEROSKEDASTIC REGRESSION MODELS
by Adrian Pagan and H. Sabau, October 1987

WP#105 RECURRENT ADVERTISING
by Ignatius J. Horstmann and Glenn M. MacDonald, October 1987

WP#106 PREDICTIVE EFFICIENCY FOR SIMPLE NONLINEAR MODELS
by Thomas F. Cooley, William R. Parke and Siddhartha Chib, October 1987

WP#107 CREDIBILITY OF MACROECONOMIC POLICY: AN INTRODUCTION AND A BROAD SURVEY
by Torsten Persson, November 1987

WP#108 SOCIAL CONTRACTS AS ASSETS: A POSSIBLE SOLUTION TO THE TIME-CONSISTENCY PROBLEM
by Laurence Kotlikoff, Torsten Persson and Lars E. O. Svensson, November 1987

WP#109 EXCHANGE RATE VARIABILITY AND ASSET TRADE
by Torsten Persson and Lars E. O. Svensson, Novmeber 1987

WP#110 MICROFOUNDATIONS OF INDIVISIBLE LABOR
by Vittorio Grilli and Richard Rogerson, November 1987

WP#111 FISCAL POLICIES AND THE DOLLAR/POUND EXCHANGE RATE: 1870-1984
by Vittorio Grilli, November 1987

WP#112 INFLATION AND STOCK RETURNS WITH COMPLETE MARKETS
by Thomas Cooley and Jon Sonstelie, November 1987

WP#113 THE ECONOMETRIC ANALYSIS OF MODELS WITH RISK TERMS
by Adrian Pagan and Aman Ullah, December 1987

WP#114 PROGRAM TARGETING OPTIONS AND THE ELDERLY
by Eric Hanushek and Roberton Williams, December 1987

WP#115 BARGAINING SOLUTIONS AND STABILITY OF GROUPS
by Youngsub Chun and William Thomson, December 1987

WP#116 MONOTONIC ALLOCATION MECHANISMS
by William Thomson, December 1987

WP#117 MONOTONIC ALLOCATION MECHANISMS IN ECONOMIES WITH PUBLIC GOODS
by William Thomson, December 1987

WP#118 ADVERSE SELECTION, AGGREGATE UNCERTAINTY, AND THE ROLE FOR MUTUAL INSURANCE COMPANIES
by Bruce D. Smith and Michael J. Stutzer, February 1988

WP#119 INTEREST ON RESERVES AND SUNSPOT EQUILIBRIA: FRIEDMAN'S PROPOSAL RECONSIDERED
by Bruce D. Smith, February 1988

WP#120 INTERNATIONAL FINANCIAL INTERMEDIATION AND AGGREGATE FLUCTUATIONS UNDER ALTERNATIVE EXCHANGE RATE REGIMES
by Jeremy Greenwood and Stephen D. Williamson, February 1988

WP#121 FINANCIAL DEREGULATION, MONETARY POLICY, AND CENTRAL BANKING
by Marvin Goodfriend and Robert G. King, February 1988

WP#122 BANK RUNS IN OPEN ECONOMIES AND THE INTERNATIONAL TRANSMISSION OF PANICS
by Peter M. Garber and Vittorio U. Grilli, March 1988

WP#123 CAPITAL ACCUMULATION IN THE THEORY OF LONG RUN GROWTH
by Paul M. Romer, March 1988

WP#124 FINANCIAL INTERMEDIATION AND ENDOGENOUS GROWTH
by Valerie R. Bencivenga and Bruce D. Smith, March 1988

WP#125 UNEMPLOYMENT, THE VARIABILITY OF HOURS, AND THE PERSISTENCE OF "DISTURBANCES": A PRIVATE INFORMATION APPROACH
by Bruce D. Smith, March 1988

WP#126 WHAT CAN BE DONE WITH BAD SCHOOL PERFORMANCE DATA?
by Eric Hanushek and Lori Taylor, March 1988

WP#127 EQUILIBRIUM MARKETING STRATEGIES: IS THERE ADVERTISING, IN TRUTH?
by Ignatius Horstmann and Glenn MacDonald, revised, March 1988

WP#128 REAL EXCHANGE RATE VARIABILITY UNDER PEGGED AND FLOATING NOMINAL EXCHANGE RATE SYSTEMS: AN EQUILIBRIUM THEORY
by Alan C. Stockman, April 1988

WP#129 POST-SAMPLE PREDICTION TESTS FOR GENERALIZED METHOD OF MOMENT ESTIMATORS
by Dennis Hoffman and Adrian Pagan, April 1988

WP#130 GOVERNMENT SPENDING IN A SIMPLE MODEL OF ENDOGENOUS GROWTH
by Robert J. Barro, May 1988

WP#131 FINANCIAL DEVELOPMENT, GROWTH, AND THE DISTRIBUTION OF INCOME
by Jeremy Greenwood and Boyan Jovanovic, May 1988

WP#132 EMPLOYMENT AND HOURS OVER THE BUSINESS CYCLE
by Jang-Ok Cho and Thomas F. Cooley, May 1988

WP#133 A REFINEMENT AND EXTENSION OF THE NO-ENVY CONCEPT
by Dimitrios Diamantaras and William Thomson, May 1988

WP#134 NASH SOLUTION AND UNCERTAIN DISAGREEMENT POINTS
by Youngsub Chun and William Thomson, May 1988

WP#135 NON-PARAMETRIC ESTIMATION AND THE RISK PREMIUM
by Adrian Pagan and Y. Hong, May 1988

WP#136 CHARACTERIZING THE NASH BARGAINING SOLUTION WITHOUT PARETO-OPTIMALITY
by Terje Lensberg and William Thomson, May 1988

WP#137 SOME SIMULATION STUDIES OF NON-PARAMETRIC ESTIMATORS
by Y. Hong and A. Pagan, June 1988

To order copies of the above papers complete the attached invoice and return to Christine Massaro, W. Allen Wallis Institute of Political Economy, RCER, 109B Harkness Hall, University of Rochester, Rochester, NY 14627. Three (3) papers per year will be provided free of charge as requested below. Each additional paper will require a $5.00 service fee which must be enclosed with your order. For your convenience an invoice is provided below in order that you may request payment from your institution as necessary. Please make your check payable to the **Rochester Center for Economic Research**. **Checks must be drawn from a U.S. bank and in U.S. dollars**.

W. Allen Wallis Institute for Political Economy

Rochester Center for Economic Research, Working Paper Series

OFFICIAL INVOICE

Requestor's Name _____

Requestor's Address _____

_____

_____

_____

Please send me the following papers free of charge (**Limit: 3 free per year**).

WP# _____ WP# _____ WP# _____

I understand there is a $5.00 fee for each additional paper. Enclosed is my check or money order in the amount of $_____. Please send me the following papers.

WP# _____ WP# _____ WP# _____

WP# _____ WP# _____ WP# _____

WP# _____ WP# _____ WP# _____

WP# _____ WP# _____ WP# _____