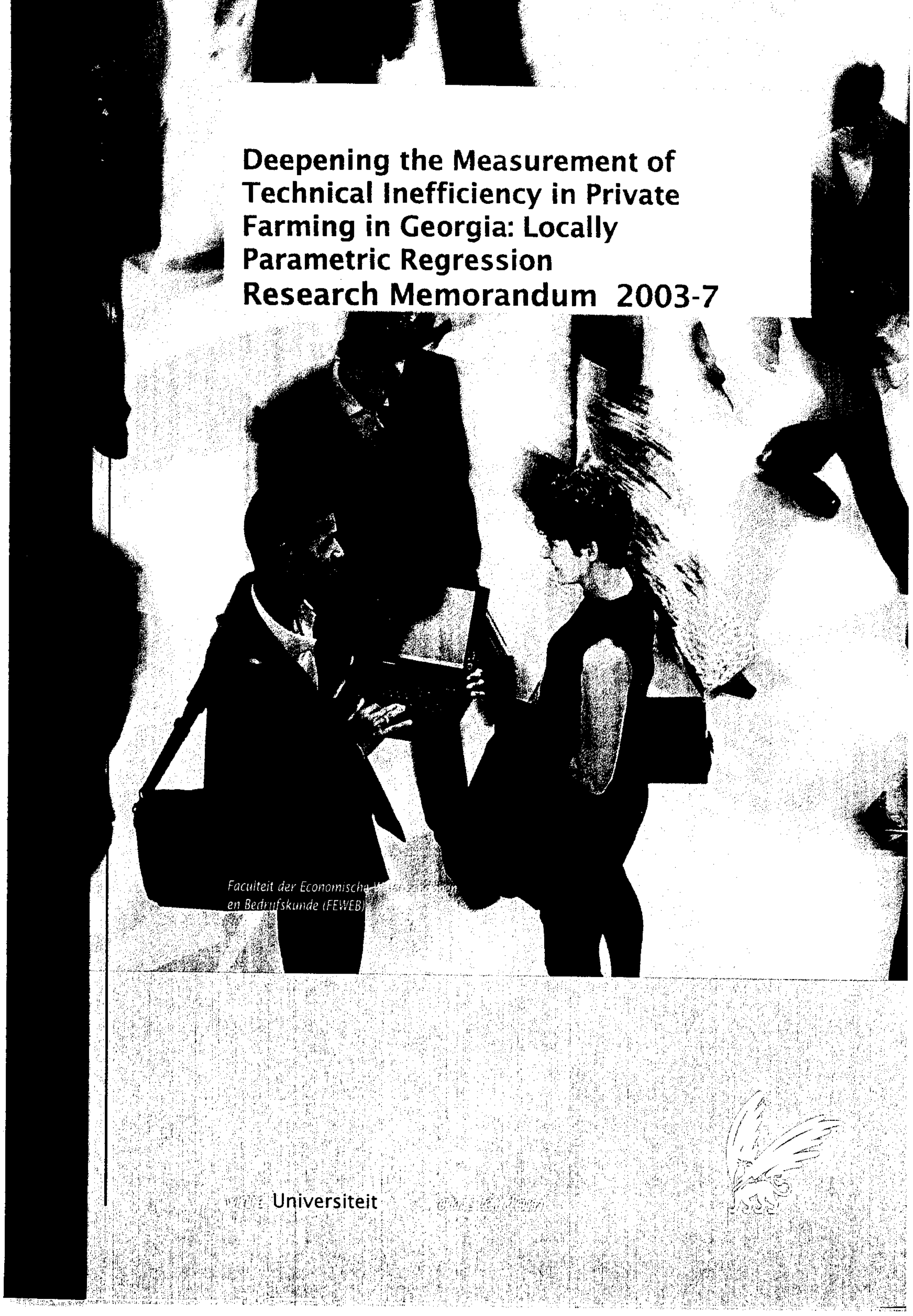
Deepening the Measurement of Technical Inefficiency in Private Farming in Georgia: Locally Parametric Regression

Research Memorandum 2003-7



Faculty of Economics and Business Administration

Deepening the Measurement of Technical Inefficiency in Private Farming in Georgia: Locally Parametric Regression

Research Memorandum 2003-7

Tugrul Temel
Andre Lucas

vrije Universiteit amsterdam

Deepening the Measurement of Technical Inefficiency in Private Farming in Georgia: Locally Parametric Regression

Tugrul Temel
International Service for National Agricultural Research
Laan van Nieuw Oost Indie 133
2593 BM The Hague
The Netherlands
t.temel@cgiar.org

Andre Lucas
Department of Finance and Tinbergen Institute
Free University of Amsterdam
De Boelelaan 1105, 1081 HV Amsterdam
The Netherlands
alucas@econ.vu.nl

February 19, 2003

Abstract

This study deepens the measurement of technical inefficiency in private maize farming in Georgia, applying locally parametric (LP) regression method, whicb builds on the stochastic frontier production function approach. Detailed survey data for 221 mixed farms for 1997 are used in the estimations. Findings suggest: (i) maize production can be further increased by breaking up large farms into smaller parcels; and (ii) increased schooling and farm experience of new private farm operators would further reduce the inefficiency. Furthermore, global and local estimations of the inefficiency suggest different policy directions as to the future of maize farming. The global estimations, revealing non-disciriminant overestimation of the inefficiency, prejudice broad-based farm reforms. On the contrary, the local estimations, pcinting out relatively large farms as the kcy source of the inefficiency, favor the design of specific policies for the effective operation of large farms.

Key words: Technical inefficiency, local and global parametric regression methods, stochastic frontier production function, after-reform maize farming in Georgia

JEL Codes: Q12, C2, Q15

1 Background and Introduction’

Farm inefficiency and its determinants have since long enjoyed prime interest in much empirical end theoretical work, beginning with Farrell (1957), continuing with Aigner, Lovell, and Schmidt (1977) and Mausen and van den Broeck (1977), and most recently including Bauer (1990), Cornwell, Schimidt, and Sickles (1990), Greene (1993), Lee and Schmidt (1993), and Battese, Malik, and Gill (1996) among others. The main goal of these studies was to estimate the amount of output foregone due to inefficient use of inputs and to determine key factors that account for this inefficiency.

Technical inefficiency is a measure of the gap between the frontier (or ideal) and the actual output levels. Consider, for example, a production relation, $y \leq f(z)$, where $f(\cdot)$ denotes production technology that translates inputs, x, into output, y. If a farm employs the optimal bundle 2, together with the most appropriate technology, the production relation would hold as $\hat{y} = f(\hat{x})$, where $\hat{y}$ is the frontier output. On the contrary, when either inputs are used sub-optimally or an inappropriate technology is adopted. the same production relation would take the form of $y = f(z) + u$, where $u = (y - \hat{y}) \leq 0$ is a measure of technical inefficiency. Such formulation of the production relation is deterministic since the inefficiency is attributed only to farmers' sub-optimal choice of input use.

In the case that the production relation is also affected by exogenous factors, those that are not under the control of farmers, the stochastic frontier production function approach, first introduced by Aigner, Lovell, and Schmidt (1977) (ALS h enceforth), becomes suitable to estimate the inefficiency. The stochastic frontier production takes on the form, $y = f(x, \beta) + \varepsilon$, where β is a vector of coefficients to be estimated; $\varepsilon = v + u$, a composite error term; v, random error; and u, the inefficiency. This formulation of the production relation assumes that v and u follow a symmetric (the normal) and an asymmetric distribution (the half-normal), respectively; and that both v and u are orthogonal to each other and to x. These assumptions would allow to disentangle v and u from the estimated composite regression errors $\hat{\varepsilon}$.

One of the disadvantages of the stochastic frontier approach is that it postulates two parametric specifications: one for the probability distribution of v and u and another for the production relation $f(\cdot)$. Adams, Berger, and Sickles (1999) relax these specifications by applying a semi-parametric efficient estimator.[2] They further use panel data to disentangle the time-varying error term, v_{it}, and the time-invariant inefficiency, u_i, from the estimated regression errors $\hat{\varepsilon}_{it}$. Unfortunately, their methodology does not allow to study the inefficiency in a cross-section context.

In recent years, nonparametric approaches began to become popular in estimating the inefficiency, as they are free of distributional assumptions and ad hoc functional specifications. Data Envelopment and kernel density regression methods are the two most commonly applied in the literature. They both attribnte all deviations from the estimated frontier to the inefficiency, thus setting the *random* error term $v \equiv 0$. A third mechod. *introduced* by Varian (1984), *incorporates* economic regnlarity conditions to the frontier analysis by finding the minimal perturbation of data that satisfies the inequality relations implied by the weak axiom of revealed preference. Although these techniques generate estimates robust to misspecification, their precision varies inversely with the number of explanatory variables and the number of observations (Härdle, 1991; Yatchew, 1998), and hence parsimony is important when such techniques are applied.

The current study introduces Tibshirani and Hastie's (1987) locally parametric (LP) regression method to estimate technical inefficiency in maize farming in Georgia, using detailed survey data for 1997. The basic idea underlying the LP method is to locally apply in input space ALS's globally parametric (GP) method. The observations sufficiently close to the postulated input vector x are used to estimate the the coefficients of the GP model and the parameters of the normal and half-normal distributions. Therefore, the LP coefficient estimations will vary in different parts of the input space, as opposed to the GP estimations fixed for the entire input space. With the application of the LP method, we introduce to the literature a way for deepening the measurement

[1]The authors like to thank Michiel Keyzer. Geert Overbosch, Maarten Nube, and seminar participants at the Center for World Food Studies - Free University Amsterdam (SOW-VU) for their comments on the earlier version of the paper. At the time a first draft of this study was completed, the first author was associated with the SOW-VU.

[2]See Greene (1990) for the implications of these assumptions.

of technical inefficiency. Finally, the study quantifies the efficiency gains attributed to the LP method by comparing it to the estimates from the GP method.

2 Model Specification and Estimation

Consider a Cobb-Douglas production function:

$$Y_i = A_i \prod_{j=1}^{Z} X_{ij}^{\alpha_j}, \text{ with } \Sigma_{j=1}^{Z}\alpha_j = 1$$

where Y and X stand for farm production and input, respectively. Subscripts $i = 1, ..., n$ and $j = 1, ..., Z$ represent the number of farms and inputs used, respectively. A_i, the coefficient of technical efficiency of farm i, is such that $A^* \geq A_i$ for all i. The production frontier is then given by

$$Y_i = e^{v_i} A_i \prod_{j=1}^{Z} X_{ij}^{\alpha_j}, \tag{1}$$

where the random error term $v_i \equiv 0$ and $A_i = A^*$ for all i. Following Aigner et al., Eq. (1) is rewritten as

$$Y_i = e^{v_i} e^{u_i} \prod_{j=1}^{Z} X_{ij}^{\alpha_j} \quad \Leftrightarrow \quad y_i = \sum_{j=1}^{Z} \alpha_j x_{ij} + \varepsilon_i, \tag{2}$$

with $y_i \equiv \ln(Y_i)$, $x_{ij} \equiv \ln(X_{ij})$, $u_i \equiv \ln(A_i/A^*)$, and $\varepsilon_i \equiv v_i + u_i$. It should be noted that v_i is an ordinary random error possibly taking on either negative or positive values; however, by definition, the measure of inefficiency, u_i, must be non-positive. Again, following Aigner et al., we model this by setting $u_i = -|\tilde{u}_i|$ and assuming that v_i and $\tilde{u}_i$ have zero-mean normal distributions with variances σ_v^2 and σ_u^2, respectively. Furthermore, v_i and $\tilde{u}_i$ are assumed to be orthogonal to each other and to the regressors x_{ij}.

Aigner et al's model is parametric in two respects. First, the specification of the maximum output function in Eq. (2) is parametric. For example, by imposing the linear structure as in Eq. (2), we are essentially imposing a Cobb-Douglas type production function. Though this might provide an adequate approximation in some cases, it may fail in others. For instance, if the true production function is of the CES-type, we are likely to end up with biased estimates of the components in Eq. (2). Second, Aigner et al's approach concerns the normal-half-normal specification imposed on (v_i, u_i). Though this specification captures the different nature of the error components, the imposed distributional shape might be overly restrictive. As discussed in the introduction, Adams et al. relax these parametric assumptions. In particular, they apply kernel estimation techniques to estimate the density of ε_i adaptively. As they employ panel rather than cross-section data, they are able to construct consistent estimates of u_i by averaging the estimation errors $\hat{\varepsilon}_{it}$ over time t, (i.e., $\hat{u}_i = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_{it}$). They also sketch an approach by which one can further relax the parametric assumptions on $f(\cdot)$ and estimate $\hat{f}(\cdot)$ rather than $f(\cdot;\hat{\beta})$ for given $f(\cdot;\cdot)$.

Following Tibshirani and Hastie (1987), the current study applies the LP approach based on Aigner et al's GP approach.[3] Let our model be given by,

$$y_i = f(x_i) + u_i + v_i, \tag{3}$$

[3] The reader is referred to Loader (1996) and Hjort and Jones (1996) for more reading on the LP approach.

where $u_i = -|\tilde{u}_i|$ and

$$\begin{pmatrix} v_i \\ \tilde{u}_i \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_v^2(x_i) & 0 \\ 0 & \sigma_u^2(x_i) \end{pmatrix}\right). \tag{4}$$

Note that the linear parametric assumption in Eq. (2) is relaxed by letting $f(\cdot)$ depend (and possibly non-linearly) on the input vector x_i. Furthermore, the parameters of the normal-half-normal distribution depend on the value of x_i. In this way, the distribution of ε_i may take different shapes in different parts of the input space.

To estimate the parameters of Eq. (3) and Eq. (4), we use locally weighted maximum likelihood based on a nearest-neighbor type weighting schema Let

$$\ell(f(x), \sigma_v^2(x), \sigma_u^2(x)) = \sum_{i=1}^{n} w(x - x_i; k_n) \cdot \tilde{\ell}(f(x), \sigma_v^2(x), \sigma_u^2(x)), \tag{5}$$

where n is the number of observations; k_n, a smoothing parameter giving the number of nearest neighbors; and $\tilde{\ell}(\cdot)$, the global liielihood of the normal-half-normal model (see Aigner et al. (1977) for the precise formulae). The weights in Eq. (5) are of the form

$$w(x \quad x_i; k_n) = \frac{0.75}{h(k_n, x)} \left(1 - \frac{d_i(x - x_i)^2}{h(k_n, x)}\right) 1_{[0,1]}(d_i(x - x_i)^2 / h(k_n, x)), \tag{6}$$

where the Mahalanobis distance is defined as

$$d_i(x - x_i) = \sqrt{(x - x_i)' Cov(x_i)^{-1} (x - x_i)}.$$

$Cov(\cdot)$ stands for the sample covariance matrix; $1_A(\cdot)$, the indicator function for the set A; and $h(k_n, x)$, a constant such that the indicator function takes on a "on-zero value for k_n observations only. The specification of Eq. (6) corresponds to a" Epanechnikov kernel estimator (see Silverman (1986)). Observations that are closer in input space to the postulated input vector x receive a larger weight in the weighted likelihood Eq. (5). To obtain consistency, we need n, $k_n \to \infty$ as well as $k_n/n \to 0$ (see Silverman (1986)).

Using the approach sketched above, we maximize the weighted local likelihood Eq. (5) for 5 = $x_1, \ldots, x_n$, thus obtaining n estimated tuples $(\hat{f}(x_i), \hat{\sigma}_v^2(x_i), \hat{\sigma}_u^2(x_i))$. We need to estimate farm i's technical inefficiency, u_i, for cross-sectional data; however, this is not possible because the regression error $\hat{\varepsilon}_i = [y_i - \hat{f}(x_i)]$ is a composite of the random error v_i and the inefficiency measure u_i. These two terms cannot be identified separately, therefore, instead of using u_i, we use either $\mathbb{E}[u_i | x = x_i, \varepsilon_i = \hat{\varepsilon}_i]$ or $\mathbb{E}[\exp(u_i) | x = x_i, \varepsilon_i = \hat{\varepsilon}_i]$. After some algebra, we derive the following two expressions:

$$\mathbb{E}[u_i | x = x_i, \varepsilon_i = \hat{\varepsilon}_i] = \lambda_i \hat{\varepsilon}_i - \frac{\sigma_i \phi(\lambda_i \hat{\varepsilon}_i / \sigma_i)}{\Phi(-\lambda_i \hat{\varepsilon}_i / \sigma_i)},$$

and

$$\mathbb{E}[\exp(u_i) | x = x_i, \varepsilon_i = \hat{\varepsilon}_i] = \exp(\lambda_i \hat{\varepsilon}_i + \sigma_i^2/2) \frac{\Phi(-(\lambda_i \hat{\varepsilon}_i + \sigma_i^2)/\sigma_i)}{\Phi(-\lambda_i \hat{\varepsilon}_i / \sigma_i)},$$

with $\phi(\cdot)$ the standard normal density function, $\Phi(\cdot)$ the normal distribution function, $\lambda_i = \sigma_u^2(x_i)/(\sigma_u^2(x_i) + \sigma_v^2(x_i))$, and $\sigma_i^2 = \lambda_i \sigma_v^2(x_i)$. It is straightforward to verify that these inefficiency measures tend to zero if $\hat{\varepsilon}_i \to \infty$, and to $-\infty$ for $\hat{\varepsilon}_i \to -\infty$.

3 An Application

Using data from a rural household survey conducted in Georgia in 1997,[4] the GP and the LP methods are applied (i) to estimate a production function for maize, which is, by far, most commonly

[4]The Rural Poverty Study of the Caucasus Countries - Georgia was conducted in 1997 by the Center for World Food Studies of the Vrije Universiteit (SOW-VU). Amsterdam, The Netherlands, in collaboration with the International Center for Reformation and Development of Georgian Economy (ICRDGE), Tbilisi, Georgia. The study was commissened by the International Fund for Agricultural Development (IFAD).

grown crop by private farmers that emerged after indepeodence in 1991, (ii) to measure farm-level technical inefficiency, and (iii) to identify the determinants of this inefficiency. Thereafter, the performance of thes two methods is compared with respect to their explanatory power.

3.1 Production function

A production function of the form, $y_i = f(x_i) + \varepsilon_i$, is estimated, where $y_i = \ln(Y_i)$ is farm i's maize production, $x_i = [\ln(K_i), \ln(N_i), \ln(L_i)]$s a vector of inputs farm i uses, snd ε_i is the disturbance term. The inputs K_i, N_i, and L_i denote total machine hours, total labor hours, and total harvested area (i.e., as a measure of farmsize), respectively.

The full-sample GP estimations of the production function are reported in Table 1.[5] Capital is found to be the only significant input; and farmsize varies inversely with the production. (This is an empirical relationship well-documented in the literature.[6]) Also reported in Table 1 are the estimates of standard errors of the stochastic error, σ_v, and of the technical inefficiency, σ_u. That the σ_u (1.05) swamps the σ_v (0.11) suggests that inefficiencies should especially be attributed to factors under the control of farmers. This clear dominance of the technical inefficiency component over the stochastic component sterns partly from the omission of other production inputs, like fertilizer, since the omitted variables can be viewed as farmers' inability in choosing the right bundle of inputs that play significant role in farm production.

In order to examine whether small farms differ from large ones with respect to technology and technical inefficiency parameters, the GP estimations of the production function are performed separately for each farmsize (see Table 2). Farms in the lowest 33 percentile of the variable L_i are classified as small; those in the middle 33 percentile, as medium; and those in the highest 33 percentile, as large farms. For small farms, capita1 and labor both have positive but statistically insignificant contribution to the production. For this group, technical inefficiency originates mostly from factors under the control of farmers, implied by $|\sigma_u| > |\sigma_v|$. For medium-size farms, only labor positively and significantly contributes to the production, and although $|\sigma_u| > |\sigma_v|$ still holds, the level is lower compared to that estimated for small farms. Lastly, for large farms, labor still positively and significantly contributes to the production, however, capital's contribution is negatie, though statistically insignificant. This negative relationship between machine hours per hectare and the production can partly be attributed to the fact that most large farms and most farm equipment, a lot of them in poor conditions, are in the hands of previous managers of large state and collective farms. A large number of respondents in the survey declared that machines often had technical problems dwing the field work, resulting in low production. For large farms, random factors seem to play more improtant role than factors under the control of farmers, implied by $|\sigma_u| < |\sigma_v|$.

To test the hypothesis that small, medium, and large farms all face the same constraints and operate under the same technology constraints, the three regression models in Table 2 are compared pair-wise by using chi-square tests. Results are reported in Table 3, indicating that none of these pair-wise comparisons is statistically significant. This suggests that similar production constraints apply to all of the private farms concerned, and hence the estimated production function in Table 1 is assumed to represent the average maize production relations in the survey at hand.

Conclusion 1 *Production relations are indifferent across farmsize. This suggests that a single production relation estimated by the GP method using the entire* sample *is to represent all farms under investigation.*

[5] Fertilizer use was excluded from the estimations as its inclusion reduces the number of observations in the sample. The new private farmers hardly apply fertilizer in their maize farming because of high price and dismanteled distribution system.

[6] For a thorough examination of the inverse relationship in the context of both farming in developing and developed countries, the reader is referred to Stanton (1978), Feder (1985), and Tavernier, Temel, and Li (1997) among others.

3.2 Technical inefficiency

Farm technical inefficiency is quantified by the conditional estimations, $\mathbb{E}[u_i|x = x_i, \varepsilon_i = \hat{\varepsilon}_i] = h(c_i)$ and $\mathbb{E}[\exp(u_i)|x = x_i, \varepsilon_i = \hat{\varepsilon}_i] = h(c_i)$, where $c_i \equiv (a_i, e_i, m_i, s_i)$ is a vector of characteristics of i^{th} farm operator, a_i is operator i's age, e_i is farming experience in years, m_i is farm management experience, and s_i is years of schooling. The hypothesis that maiz production is indifferent across male and female operators cannot be tested, as our survey has only a few female operators. Farming experience anf farm management (both measured in years on present farm), and schooling should be regarded as possible determinants of technia! change. Experience should increase maize production directly, while schooling is likely to affect the prcduction through the enhancement of farmer's allocative ability for input use decisions.

3.2.1 GP estimations

The two types of the conditional inefficiency measures given above are estimated by using $\hat{\varepsilon}$, which is obtained from the production function estimation in Table 1. The vector of characteristics c is regressed on $E(\exp(u)|\hat{\varepsilon})$ and $E(u|\hat{\varepsilon})$ separately, and the estimation results are presented in Tables 4 and 5, respectively. The signs and the levels of significance of the estimated coefficients remain the same across the two regression models. The estimations indicate that the inefficiency decreases with more schooling and more farming experience, while increasing with age and more management experience. Surprisingly, the only statistically significant variable, which is also robust across the regression models, is management experience. Such a controversial finding can, to a large extent, be attributed to the fact that the Georgian land reform entitled, without discriminating, managers of old collective farms and elderly people to receive land from the government, although these people were not able to effectively use the land for productive purposes, at least in the early years of reform. As a result, inefficiency was high among them.[7]

The two regression models of the conditional inefficiency are further estimated by using the demeaned variables (see Tables 6 and 7).[8] The intercept terms in these models measure the inefficiency of an average farm and are significant at the 0.01 leve!. Similar to the findings above, the inefficiency seems to decline with more schooling and more farm experience and to increase with age and management experience, and again the only significant variable is management experience. The difference between the intercept term (0.479) in Table 4 and that (0.496) in Table 6 amounts to 0.017, which suggests that an average farm is more efficient relative to a farm run by an operator who literally has no schooling, no farm experience, and no management experience. A similar comparison of the intercept terms in Tables 5 and 7 results in 0.118.

3.2.2 LP estimations

The LP method allows to test the hypothesis that farms close to each other on the output space operate under similar constraints and technology parameters. This method is applied using two different models, one with the dependent variable $E(\exp(u)|\hat{\varepsilon})$ and another with $E(u|\hat{\varepsilon})$, each of which is estimated across three different window sizes (50, 75, and 100 observations) (see Tables 8 and 9). Next, the same LP estimations are performed using the demeaned data (see Tables 10 and 11). The key advantage of the LP method over the GP method is that the different window sizes become instrumental in eliminating the influence of outliers on the prcduction frontier.

Estimations based on the original data are given in Tables 8 and 9. They show that the mode! with a window size of 100 observations performs the best . compared to the other two models estimated with window sizes of 50 and 75 observations. With an *F-statistic* of 2.96 (3.42) in Table 8 (Table 9), it is statistically significant at the 0.02 (0.01) leve! and explains 5 (6) percent of the variation in inefficiency. The term (0.506 + 0.007 (age)) in Table 8 and (-0.930 + 0.013 (age)) in

[7]The GP estimations in Tables 4 and 5 were also carried out by omitting "age", but the results remained the same.

[8]Let y and $\bar{y}$ stand for the original variable and its average value, respectively. The demeaned variable is defined as $y_d = (y - \bar{y})$. Therefore, the original model, $y = \alpha + \beta x + \epsilon$, can be expressed as the demeaned model, $y_d = \alpha_d + \beta_d x_d + \epsilon$, where α_d represents the level of technical inefficiency of an average farm (i.e., $(x = \bar{x}) \rightarrow x_d = 0$).

Table 9 both measure the level of inefficiency of a farmer with no schooling, no farming, and no management experience. In both models, farming experience decreases, while age and management experience increase inefficiency. Except for schooling, all other variables are significant at the 0.05 level.

Estimations based on the demeaned data are given in Tables 10 and 11. They show that the model with a window size of 100 observations performs the best. With an *F-statistic* of 2.96 (3.42) in Table 10 (Table 11), this model is statistically significant at the 0.02 (0.01) level and explains 5 (6) percent of the variation in inefficiency that is attributed to an average farm. Farming experience decreases, while age and management experience increases inefficiency. Except for schooling, all other variables are significant at the 0.05 level.

Conclusion 2 *The LP performs better thon the GP method in accounting for variation in technical inefficiency.*[9]

Conclusion 3 *Common to all the estimations is the positive relationship between increased management experience and the inefficiency. This suggests that the old Soviet style farm management is obsolete zn the current private farming environment.*

3.2.3 A comparison

The GP and LP frontier production function estimations are compared conditional on capital and land use. Figure l(a) shows the frontier estimations conditional *on low capital use and small farm size* (measured in terms of harvested land). The GP method projects a positively sloped, linear production frontier. which lies above all the projections by the LP method. The LP projections well behave in the sense that they fit into the net revenue maximizing-agents framework. The LP projection with a window size of 100 observations performs the best compared to the projections associated with window sizes of 50 and 75 observations. The optimal labor use is roughly 3.8 and the optimal output is close to 9. Figure l(b) shows the frontier estimstions conditional on *high capital use and small farm size.* The optimal labor use increases a little over 5 and the optimal output a little over 9. Figure l(c) shows the frontier estimations conditional *on lom capital use and large farm size.* Both farm output snd labor use decrease compared to the corresponding levels in Figure l(a), suggesting that small farms using low capital are more productive than large farms using low capital. This suggests that small private farms that emerged after the land reform have been more productive than large farms. Figure l(d) shows the frontier estimations conditional on *high capital use and large farm size.* The optimal labor use is 4.4, and the optimal output a little lower than 9. Comparing this to Figure l(b) indicstes that small farms with high capital employ more labor and produce more output.

Conclusion 4 *The key implication of these findings is that agricultural policies should target small farms' access to capital as these farms seem to more productive than large ones. This further suggests that breaking up large farms into smaller parcels is a viable option for increasing the aggregate maize production. Interestingly, all Of these findings are discovered only when we deepen the analysis by applying the LP method.*

The GP and LP estimations are further compared with respect to the empirical distributions of the estimated coefficients. For each observation (x, y) in the sample, an LP regression is estimated using n observations around (x, y). This procedure produces 221 sets of coefficients. Each set includes 6 elements: an intercept term, a coefficient for capital, a coefficient for labor, a coefficient for land, a coefficient for σ_v^2 and a coefficient for σ_u^2. The empirical distribution for the intercept term in Figure 2(a), for example. is nothing more than the histogram of 221 locally estimated intercepts. The vertical line at 8.52 represents the intercept term of the GP estimation in Table 1. Other three distributions around it represent the empirical distributions of the intercept terms of

[9]One should simply compare the LP estimations with window size 100 in Table 8 with the GP estimations in Table 4, considering R^2, probability of F-test, and the number of significant variables Similar comparisons should be made between Table 9 and Table 5, between Table 10 and Table 6, and between Table 11 and Table ,

the three LP models, each of which is associated with window sizes of 50, 75, and 100 observations. Similarly, the vertical liie at 0.08 in Figure Z(b) represents the coefficient of capital in the GP estimation in Table 1. Other three distributions around it represent the distributions of the LP coefficients of capital, each of which corresponds to window size 50, 75, and 100 observations. Figures 2(c) and Z(d) are for labor and land, respectively. The distributions in Figures 2(a) through 2(d) look very much like normal distributions. Therefore, from law of large numbers, we can expect that the LP *limiting* intercept term would be lower than the GP intercept term in Figure 2(a); that the LP *limiting* coefficient for capital would be very close to the GP coefficient for capital in Figure Z(b); and that the LP *limiting* coefficient for labor would be a little higher than the GP coefficient for labor in Figure 2(c). For land, however, most local estimations fall on the left of the vertical line at -0.14 (Figure 2(d)), suggesting that the LP *limiting* coefficient for land is more likely to be even lower than -0.14.

Conclusion 5 *The GP method overestimates marginal effects of production inputs due most likely to extreme values in the sample.*

Finally, the GP and LP estimations are compared with respect to the empirical distributions of the standard deviations of the stochastic error term, σ_v, and of technical inefficiency term, σ_u. Figure 2(e) shows that almost all of the LP distributions of σ_v fall on the right hand side of the vertical line representing the GP $\sigma_v = 0.11$, implying that the GP underestimates the production effects of uncontrollable factors. Figure 2(f) indicates that almost all of the LP distributions of σ_u fall on the left hand side of the vertical line at the CP $\sigma_v = 1.05$, implying that the GP overestimates the production effects of controllable factors.

Conclusion 6 *The GP and LP estimations of the inefficiency suggest different policy directions as to the future of maize farming. The GP estimations, revealing non-discirimінant overestimation of the inefficiency, prejudice broad-based farm reforms. On the contrary, the LP estimations, pointing out relatively large farms as the key source of the inefficiency, favor the design of specific policies for the effective operation of large farms.*

4 Discussion

With the demise of the forma Soviet Union (SU), Georgia launched structural adjustment policies in 1992. Large-scale state enterprises were privatized, state-owned large farms - kolhozes - broken down, new regulations and laws designed, and exchange and monetary system reformed. Land reform has been at the center of development issues, as it entails implications for rural development, agricultural production, and poverty. Land was distributed to individu&, with a radical, once-for-all, reform, and at present, small-scale farms constitute a large majority (Csaki and Lerman, 1997; Lerman, 1999; IFAD. 1998). Some of the old managers of kolhozes and politically influential people received large parcels and kept farm equipment under their control. These large farms were at the same time at an advantageous position since the existing agricultural infrastructure was still favoring them, with centralized Water sources. For private farming to develop, there was a need to establish an enabling environment in which private farmers could feel secure about the land they occupy, sell their produce, buy inputs, and involve cross-border trade.

Although private farming in Georgia is still at an early stage of development, it is important for policy makers to know how the newly created private farms have performed so far and to pinpoint the areas that need to be addressed. One thing which is obvious is that private farm operators need actual farming and management experience compatible to the newly emerging markets, as they have not had the chance to run a private entity until independence. Another thing is that these operators, who are farming on small parcels of land, are first unable, second unwilling to initiate any new farming activity, as they lack financial resources and management skills compatible to the new farming system. Therefore, the government holds most responsibility to create an enabling environment, one with an adequate infrastructure including market institutions snd regulatory bodies. To this end, an analysis of farm efficiency should consider both external (uncontrollable) and internal (controllable) factors. Unfortunately, however, databases currently available do not allow us to analyze thoroughly specific agricultural policies and their impacts on farm efficiency.

In this study we attempted to estimate the maize production function and technical inefficiency attached to is. The estimations underline the following key policy issues. First, politica1 networking played a salient role in the initial distribution of farm land, and hence a significant number of managers of the old state farms received farm land mostly large and controlled the use of farm equipment. But they lacked labor and operative farm machinery. Small farm operators, on the other hand, were mainly lacking farming experience since most of them used to hold non-farm jobs before independence. Second, as implied by the LP estimations, the inefficiency mostly stems from the factors under farmers' control, and hence any government policy and/or service, such as agricultural extension and training program, that enhances farmers' allocative ability should positively contribute to the production and increase the farm efficiency. Equally important is experience in farming for own account, which most private farmers severely lack. Third, economy-wide technology constraints apply to all farms independent of farmsize. From a policy point of view, this suggests that farms with different size should not be treated differently regarding the design and implementation of pmduction-enhancing agricultural policies. Finally, management experience is found to increase technical inefficiency. This can partly be attributed to the fact that managers of old collective farms and elderly people without actual farming experience were eligible to receive land from the government but not able to use it for productive purposes. As a result, inefficiency was high among these people.

5 Concluding Remarks

This study sought to compare the GP and LP frontier production function estimations, identify the key determinants of technical inefficiency, and estimated the gains from the application of the LP method. Empirical analysis was carried out using the data obtained from a farm household survey conducted in 1997 (IFAD, 1998).

Overall. the results imply that the maize production can be further increased by breaking up the large farms into smaller parcels, provided that small farmers' access to credit is improved and that after-reform farming infrastructure meets the needs of smal1 farms, such as the construction of a decentralized water distribution system.

Future research should grow in two directions. On the methodological account, there is the need for developing better-performing estimation techniques. A fully non-parametric estimation method, which has been receiving a wide attention in the literature, might be one alternative especially in situations where few outliers determine the production frontier. On the empirical account, lack of primary data seems to be a key constraint in analyzing the developments in private farming. Therefore, efforts should focus on the construction of databases necessary to evaluate impacts of reforms on farm technical inefficiency in Georgia.

References

[1] Adams, R. M., Berger, A. N., and Sickles, R. C. (1999). Semiparametric approaches to stochastic panel frontiers with applications in the banking industry. *Journal of Business & Economic Statistics*, 17(3), 349-358.

[2] Aigner, D. J., Lovell, K. C. A., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics 6: 21-37.*

[3] Baltagi, H. B., Hidalgo, J., and Li, Q. (1996). A nonparametric test for poolability using panel data. *Journal* of Econometrics 75: 345-367.

[4] Barrett, C. (1993). On *the price risk and the inverse farm size-productivity relationship* (Staff Paper No. 369). Wisconsin: University of Wisconsin-Madison, Department of Agricultural Economics.

[5] Bauer, W. P. (1990). Recent developments in the econometrics of estimation of frontiers. *Journal of Econometrics 46:* 39-56.

[6] Cornia, G. (1985). Farm size, land yields, and the agricultural production function: An analysis for fifteen developing countries. *World Development,* 13, 513-534.

[7] Cornwell, C., Schimidt, P., and Sickles, R. C. (1990). Production functions with cross-section and time series variation in efficiency levels. *Journal Of Econometrics* 46:185-200.

[8] Farrell, M.S. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society A* 120: 253-281.

[9] Feder, G. (1985). The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. *Journal of Development Economics, 18, 297-313.*

[10] Gozalo, P., and Linton, 0. (2000). Local nonlinear least squares: Using parametric information in nonparametric regression. *Journal of Econometrics* 99(1):63-106.

[11] Greene, H. W. (1992). Maximum likelihood estimation of stochastic frontier production models. *Journal of Econometrics* 18: 285-289.

[12] Greene, H. W. (1993). The econometric approach to efficiency analysis. In H. 0. Ried, C. A. K. Lovell, and S. S. Schimdt (Eds.) (1993). *The Measurement of Productive Efficiency: Techniques and Applications.* New York: Oxford University Press.

[13] Härdle, W. (1991). *Applied Nonoparametric Regression.* Cambridge University Press.

[14] IFAD. (1998). Rural poverty study of the Caucasus countries: Georgia. Study conducted for the International Fund for Agricultural Development by the Center for World Food Studies, The Netherlands, in collaboration with the International Center for Reformation and Development of Georgian Economy, Georgia.

[15] Lee, Y. H., and Schmidt, P. (1993). A production frontier model with Aexible variation in technical efficiency. In H. 0. Fried and C. A. K. Lovell and S. S. Schmidt (Eds.). The measurement of Productive Efficiency. Oxford University Press, New York.

[16] Schmidt. P. (1986). Frontier production functions. *Econometric Reviews* 4:289-328.

[17] Stanton, F. B (1978). Perspective on farm size. *American Journal of Agricultural Economics, 60,* 727-37.

[18] Tacis European Expertise Service (TEES). (1997). Georgian Economic Trends. Quarterly Review. Brussels: Belgium.

[19] Taslim, M. (1989j. Supervision problems and the size-productivity relation in Bangladesh agriculture. *Oxford* Bulletin *of* **Economics** *and Statistics, 51, 55-71.*

[ZO] Tavernier, M. E., Temel, T., & Li, F. (1997). The role of farm ownership in off-farm work participation. *Agricultural and Resource Economics Review,* 26(1), 67-81.

[21] Temel, T., & Albersen. P.J. (2000). *New facts for an old debate: Farm size, productivity, and geography* (Staff Working Paper WP-00-02). Amsterdam, The Netherlands: Free University, Center for World Food Studies.

[22] Varian, R. H. (1984). The nonparametric approach to production analysis. *Econometrica,* 52(3):579-598.

[23] World Bank. (1996a). Georgia, Reform in the Food and Agriculture Sector. Washington, D.C., U.S.A.

[24] World Bank. (1996b). Land reform and private farms in Georgia: 1996 status (Agricultural Policy Note #6). Washington, D.C., U.S.A.

[25] Yatchew, A. J. (1998). Nonparametric regression techniques in economics. *Journal* of *Economic Literature,* XXXVI, 669-721.

Table 1. The GP estimation of the production function (Full sample)

Parameters	Estimates	Std. err.	Est./s.e.	Prob.
Constant	8.5158	0.8644	9.852	0.0000
Capital	0.0790	0.0475	1.662	0.0483
Labor	0.0690	0.1472	0.469	0.3197
Land	-0.1441	0.2298	-0.627	0.2654
Sigma(v)	0.1051	0.2640	0.398	0.3453
Sigma(u)	1.0524	0.2473	4.256	0: 0000

Table 2. The GP estimations of the production function by farmsize

Parameters	Estimates	S.e	Est./s.e	Prob.
		(Smallest 33%)		
Constant	8.1614	1.2619	6.468	0.0000
Capital	0.0991	0.2062	0.481	0.3154
Labor	0.1549	0.3032	0.511	
Sigma(v)	0.0000	1.0000	0.000	0.3047 0.5000
Sigma(u)	-1.0481	0.2039	-5.141	0.0000
		(Medium 33%)		
Constant	7.6322	0.7739	0.468	0.0000
Capital	0.0343 0.2847	0.1773	1.605	0.3200
Sigma(v)	-0.1221	0.0312	-3.916	0.0542
		[illegible]	9.202	0.0000
Sigma(u)	0.9271	0.1007		0.0000
		(Largest 33%)		
Constant	7.6701	0.5750	-1.001	0.0000
Capital	-0.1649	0.0965	1.709	0.1584
Sigma(v)	0.5337	0.1005	5.311	0.0437
				0.0000
Sigma(u)	0.3592	0.4858	0.739	0.2298

Table 3. Chi-square tests for differences in the production relations between farmsizes

	test statistic	p-value
HO: small = medium	0.3211;	0.9560
HO: small = large	3.4070;	0.3330
HO: medium = large	4.7410;	0.1918

Table 4. The GP estimation with the dependent variable [exp(u)|e]

variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep var
Constant	0.479641	0.102415	4.683292	0.000	---	
school	-0.007399	0.005278	-1.401692	0.162	-0.101725	-0.082337
exper	-0.001950	0.001512	-1.290081	0.198	-0.133809	0.012883
age	0.002641	0.001714	1.540997	0.125	0.158044	0.082532
mngtexp	0.055550	0.026948	2.061385	0.040	0.139617	0.122895

obs.: 221; R-squared: 0.037; $F(4,216)$: 2.066; Probability of F: 0.086

Table 5. The GP estimation with the dependent variable [u|e]

variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep var
Constant	-0.965353	0.254612	-3.791470	0.000	---	
school	-0.017048	0.013122	-1.299214	0.195	-0.093711	-0.074366
exper	-0.005155	0.003758	-1.371658	0.172	-0.141400	0.019155
age	0.007640	0.004261	1.793011	0.074	0.182765	0.097266
mngtexp	0.170466	0.066994	2.544492	0.012	0.171283	0.155095

obs.: 221; R-squared: 0.049; $F(4,216)$: 2.759; Probability of F: 0.029

Table 6. The demeaned GP estimation with the dependent variable [exp(u)|e]

variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
constant	0.496767	0.015770	31.500372			
school	-0.007399	0.005278	-1.401692	0.162	-0.101725	-0.082337
exper	-0.001950	0.001512	-1.290081	0.198	-0.133809	0.012883
age	0.002641	0.001714	1.540997	0.125	0.158044	0.082532
mngtexp	0.055550	0.026948	2.061385	0.040	0.139617	0.122895

obs.: 221; R-squared: 0.037; $F(4,216)$: 2.066; Probability of F: 0.086

Table 7. The demeaned GP estimation with the dependent variable [u|e]

variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
Constant	-0.847662	0.039206	-21.620800	0.000		
school	-0.017048	0.013122	-1.299214	0.195	-0.093711	-0.074366
exper	-0.005155	0.003758	-1.371658	0.172	-0.141400	0.019155
age	0.007640	0.004261	1.793011	0.074	0.182765	0.097266
mngtexp	0.170466	0.066994	2.544492	0.012	0.171283	0.155095

obs.: 221; R-squared: 0.049; $F(4,216)$: 2.759; Probability of F: 0.029

Table 8. The LP estimations with the dependent variable [exp(u)|e]

Variable	Estimate	standard Error	t-value	Prob >|t|	standardized Estimate	Cor with Dep Var
			(window size = 50)			
constant	4.927607					
school	-0.152349	[illegible]	[illegible]	0.448	-0.093363	-0.052423
exper	-0.026093	0.034348	-0.759644	0.817	-0.079795	-0.066865
age	-0.009028	0.038945	-0.231817	[illegible]	-0.024078	-0.055710
mngtexp	0.263020	0.612301	0.429560	0.668	0.029464	0.020183
obs.: 221;	R-squared: 0.012;		F(4,216): 0.665;		Probability of F: 0.617	
			(window size = 75)			
constant	1.146575	0.520145	2.204338	0.029	---	---
school	-0.006529	0.026639	-0.245093	0.807	-0.018136	0.004333
exper	0.000178	0.007617	0.023386	0.981	0.002469	-0.022843
age	-0.002991	0.008676	-0.344756	0.731	-0.036035	[illegible]
mngtexp	0.136257	0.135779	1.003520	0.317	0.069200	0.068052
obs.: 220;	R-squared: 0.006:		F(4.215): 0.310;		Probability of F: 0.871	
			(window size = 100)			
constant	0.506984	[illegible]	[illegible]			
school	0.002810	0.01032[illegible]	0.272025	[illegible]	0.019585	0.048231
exper	-0.006567	[illegible]	-2.351991	0.027	-0.228420	-0.063963
age	0.007889	0.003354	2.269372	0.020	0.239301	0.052213
mngtexp	0.119673	0.052734		0.024	0.152482	0.157786
obs.: 221;	R-squared: 0.052;		~(4,216): 2.969:		Probability of F: 0.020	

Table 9. The LP estimations with the dependent variable [u|e]

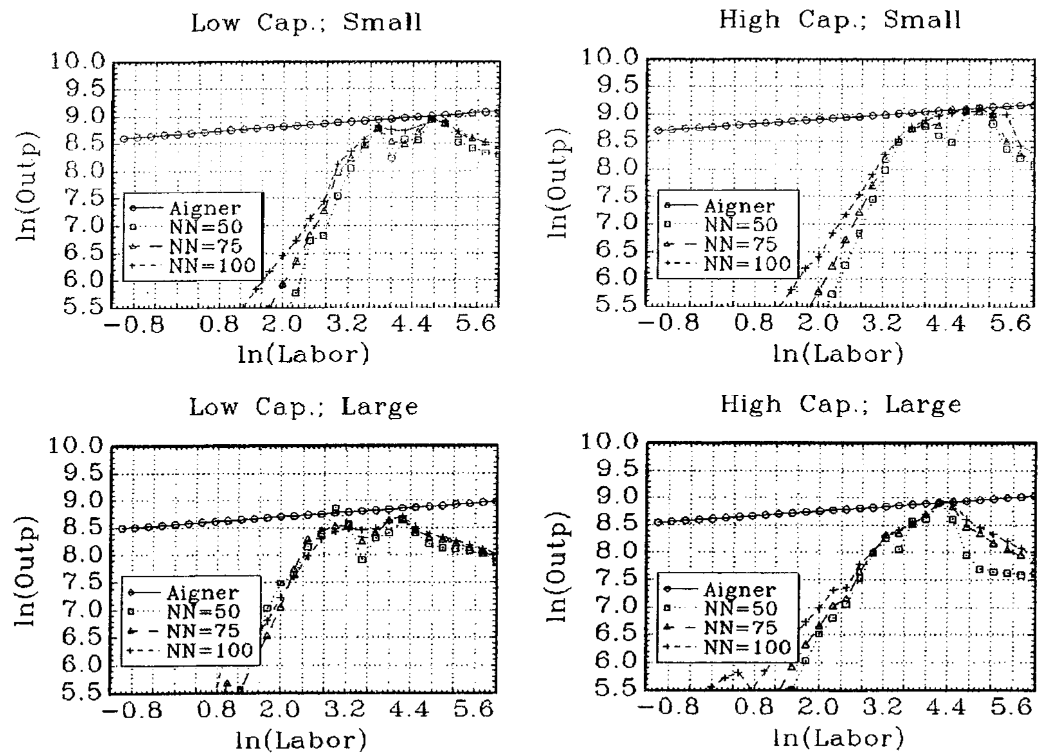
Variable	Estimate	standard Error	t-value	Prob >|t|	standardized Estimate	cor with Dep Var
			(window size = 50)	0.899		
constant	-0.079742 -0.011890	0.626831 0.032306	-0.127214 -0.368046	0.713	-0.026998 ---	0.019608 ---
school / exper	-0.007548	0.009252	-0.815791	0.416	-0.085527	-0.082751
age	-0.000089	0.010491	-0.008522	0.993	-0.000883	-0.061842
mngtexp	0.231977	0.164933	1.406488	0.161	0.096288	0.097246
obs.: 221;	R-squared: 0.016;		F(4.216): 0.876;		Probability of F: 0.479	
			(window size = 75)	0.189	---	
constant / school	[illegible]	[illegible]	[illegible]	0.643	-0.033807	-0.003846
exper	-0.00419s	0.006488	-0.646562	0.519	-0.067280	-0.012150
				0.482	0.072511	0.022651
age / mngtexp	[illegible]	[illegible]	[illegible]	0.008	0.180814	0.175542
obs.: 220;	R-squared: 0.034;		~(4,215): 1.911;		Probability of F: 0.110	
			(window size = 100)			
constant	-0.930376	0.344596	-2.699900	0.007	---	_-_
school	-0.004100	0.017760	-0.230884	0.818	-0.016557	0.006044
exper	-0.009743	0.005086	-1.915465	0.057	-0.196316	-0.020099
age	0.013707	0.005767	2.376804	0.018	0.240870	0.087977
mngtexp	0.253766	0.090671	2.798751	0.006	0.187308	0.184462
obs.: 221;	R-squared: 0.060;		~(4,216): 3.422;		Probability of F: 0.010	

Table 10. The demeaned LP estimations with the dependent variable [exp(u)|e]

Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
			(window size = 50)			
constant						
school	[illegible]	[illegible]	[illegible]	0.000	-0.093363	-0.052423
exper			5.743108	0.448	-0.079795	-0.066865
age	[illegible]	[illegible]	[illegible]	0.817	-0.024078	-0.055710
mngtex	0.263020	0.612301	0.429560	0.668	0.029464	0.020183
obs.: 421:	R-squared: 0.012;		F(4,216): 0.665;		Probability of F: 0.617	
			(window size = 75)			
Constant	0.964545	0.079599	12.117492	0.000	---	---
school	-0.006529	0.026639	-0.245093	0.807	-0.018136	0.004333
exper	0.000178	0.007617	0.023386	0.981	0.002469	-0.022843
age	-0.002991	0.008676	-0.344756	0.731	-0.036035	-0.032051
mngtexp	0.136257	0.135779	1.003520	0.317	0.069200	0.068052
obs.: 220;	R-squared: 0.006;		F(4,215): 0.310;		Probability of F: 0.871	
			(Window size = 100)			
Constant	0.002810 0.824815	0.030860 0.010329	26.727213	0.027	0.019585	0.048231
school	-0.006567	0.002958	[illegible]	0.020	-0.228420	-0.063963
age	0.007889	0.003354	2.351991	[illegible]	0.239301	0.052213
mngtexp	0.119673	0.052734	2.269372	0.024	0.152482	0.157786
obs.: 221;	R-squared: 0.052;		F(4,216): 2.969;		Probability of F: 0.020	

Table 11. The demeaned LP estimations with the dependent variable [u|e]

variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep var
			(window size = 50)			
Constant	-0.337100	0.096521	-3.492504	0.001	---	---
school	-0.011890	0.032306	-0.368046	0.713	-0.026998	0.019608
exper	-0.007548	0. 009252	-0.815791	0.416	-0.085527	-0.082751
age	-0.000089	0.010491	-0.008522	0.993	-0.000883	-0.061842
mngtex p	0.231977	0.164933	1.406488	0.161	0.096288	0.097246
obs.: 221;	R-squared: 0.016:		F(4,216): 0.876:		Probability Of F: 0.479	
			(window size = 75)			
Constant	-0.440913	0.067801	-6.503029	0.000		---
school	-0.010519	0.022691	-0.463573	0.643	-0.033807	-0.003846
exper	-0.004195	0.006488	-0.646562	0.519	-0.067280	-0.012150
age	0.005202	0.007390	0.703936	0.482	0.072511	0.022651
mngtexp	0.307715	0.115654	2.660648	0.008	0.180814	0.175542
Obs.: 220;	R-squared: 0.034;		F(4,215): 1.911;		Probability Of F: 0.110	
			(Window size = 100)			
Constant	-0.429268	0.053062		0.000		
school	-0.004100	0.017760	[illegible]	0.818	-0.016557	0.006044
exper	-0.009743	0.005086	-1.915465	0.057	-0.196316	-0.020099
age	0.013707	0.005767	2.376804	0.018	0.240870	0.087977
mngtex p	0.253766	0.090671	2.798751	0.006	0.187308	0.184462
obs.: 221;	R-squared: 0.060;		F(4,216): 3.422;		Probability of F: 0.010	


Low Cap.; Small
High Cap.; Small
Low Cap.; Large
High Cap.; Large
ln(Outp)
ln(Labor)
Aigner
NN=50
NN=75
NN=100
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
−0.8
0.8
2.0
3.2
4.4
5.6

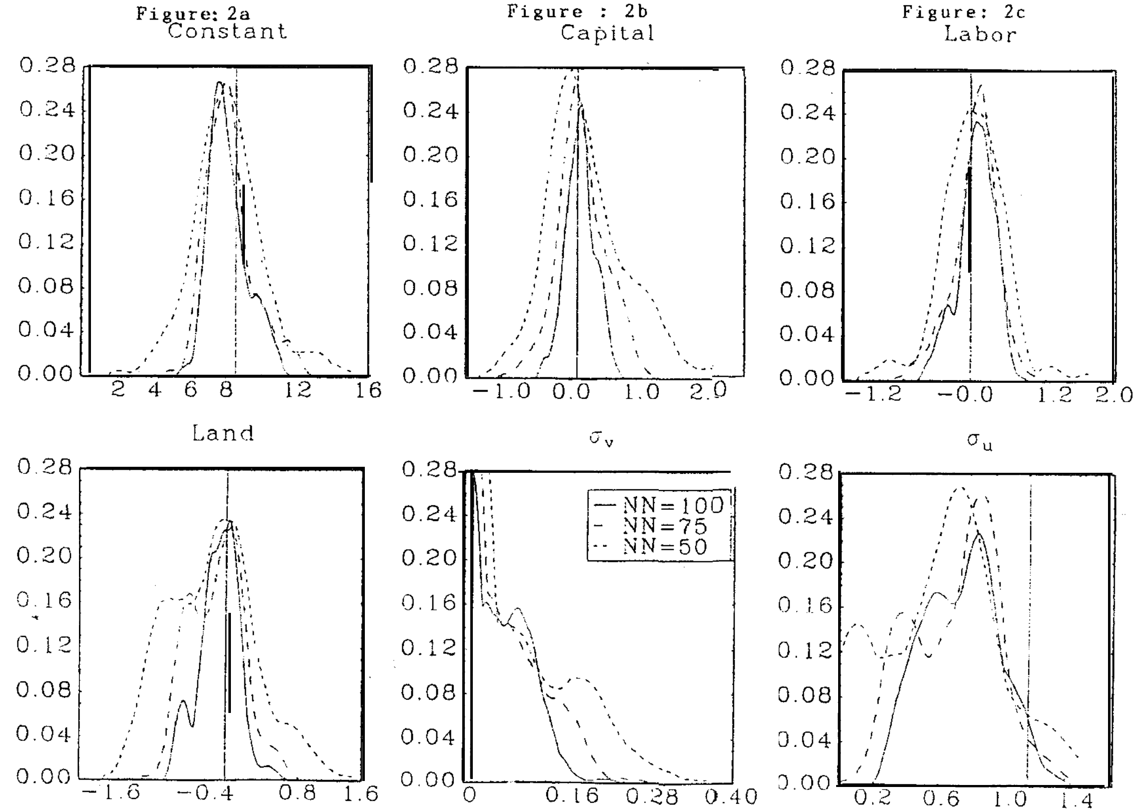


Figure: 2a

Figure : 2b

Figure: 2c

Figure:2d

Figure: 2e

Figure: 2f