A Two-stage Panel Data Envelopment Analysis Models: Wallace-Hussain, Amemiya and Swamy-Arora

Saleem Shaik (Saleem.Shaik@ndsu.edu)
North Dakota State University

Research

Keywords: Two-stage DEA, Pool and Panel DEA, Pooled, Within, Crosssection and Time-series SFA, U.S. 48 states

Posted Date: August 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-65416/v1

A Two-stage Panel Data Envelopment Analysis Models: Wallace-Hussain, Amemiya and Swamy-Arora

August 24, 2020

Abstract

This research presents two-stage panel data envelopment analysis (DEA) models for the estimation of efficiency measures. In the first stage, the variances of normal errors, v used in the transformation of the variables are recovered from pooled (Wallace-Hussain in 1969); within (Amemiya in 1971); or within, cross-sectional and time-series (Swamy-Arora in 1972) stochastic frontier analysis (SFA) models. The second stage involves estimating panel DEA efficiency measures using transformed data based on alternative two-way random effect econometric estimators in the first stage. The alternative panel DEA models might over- or under-estimate the pooled DEA model estimated efficiency measure due to the non-existence or existence of the spatial and temporal variations. An empirical application to 48 U.S. states from 1960 to 2004 suggests differences in the efficiency measure estimated by the pooled and panel DEA models. In addition, the efficiency measure estimated by the Swamy-Arora panel estimator is statistically different compared to the Wallace-Hussain and Amemiya panel estimator.

Keywords: *Two-stage DEA, Pool and Panel DEA, Pooled, Within, Cross-section and Time-series SFA, U.S. 48 states*

JEL Classification Numbers: *O3, C6, Q1*

1 Introduction

The data envelopment analysis (DEA) used in the estimation of efficiency measures is based on the output, input or graph distance function concepts developed by Malmquist (1953) and Moorsteen (1961) in the consumer context and Shephard (1953 and 1970) in the producer context. The DEA approach to the study of efficiency measures has had a history in the agriculture sector. M.J. Farrell (1957) and Farrell and Fieldhouse (1962) discussed the efficiency measures of the U.S. agriculture sector and individual farms, respectively, using multiple outputs-inputs. In 1966, at the Western Farm Management Association, four papers presented (Bressler, Boles, Seitz, and Sitorus) were related to issues associated with the different components of efficiency using a programming approach. Charnes, Cooper and Rhodes (1978) introduced the concept of efficiency to operational research. In the early 1990s, the DEA approach gained popularity, due to its ability to work without functional form, handle multiple outputs-inputs without the need for prices, and accommodate the weak and strong disposability assumptions. However, the DEA programming approach, due to its piece-wise linear approximation of the theoretical frontier, is conditioned by the number of decision making units (DMUs) and the number of constraints in the model (Shaik et al., 2012).

The DEA models are used to estimate the efficiency measure of individual DMUs in each year or for a single DMU over time. With panel data, i.e., individual DMUs over time DEA is used to estimate efficiency and productivity measures (see Briec, 1997; Bureau, Fare and Grosskopf, 1995; Ferrier and Trivitt, 2013; Fare, Grosskopf and Lovell, 1994; Lovell, 1994; Luenberger, 1992; Tulkens and Vanden, 1995). Current DEA theory and empirical models are examined with an emphasis on alternative estimators, statistical methods and bootstrapping measures (Kneipf, Simar and Wilson, 2008; Simar and Wilson, 1998, 2000 and 2011; and Bogetoft and Otto, 2011).

This research, building on the existing literature, presents two-stage panel DEA. The first stage involves recovering variance of normal errors, v given one-side errors, u from stochastic frontier analysis[1] needed for transformation of the variables. It is plausible to use the variances of one-side errors, u but the variance of normal errors, v are used in the transformation[2]. Several possibilities exist for the first-stage estimation, namely: the use of pooled errors, v as defined by

[1]The stochastic frontier analysis (SFA) model for output (y) produced by inputs (x) can be defined as $y = f(x) + v - u$, where v and u are the normal error and one-side error, respectively.

[2]We would like to thank one of the reviewer for suggesting the use of normal errors, v instead of one-side errors, u.

Wallace-Hussain (WH) in 1969; within errors, v as defined by Amemiya (AM) in 1971; or the use of within, between cross-sectional and between time-series errors, v as defined by Swamy-Arora (SA) in 1972. The second stage involves estimating panel DEA efficiency measures using transformed input and output data based on alternative two-way random effect econometric estimators in the first stage.

The development of the panel DEA models assumes there is a need to account for the spatial and temporal variations with cross-sectional time-series data. To illustrate the need to account for the spatial and temporal variations, Figures **??** and **??** presents the magnitude of spatial variation across 48 states in the U.S. and temporal variation over 45 years, respectively.

Insert Figure ?? and Figure ??

Once the spatial and temporal variations are accounted, we expect the panel DEA model estimated efficiency measure to under- or over-estimate the pooled DEA model efficiency measure. Theoretically, the pooled DEA model will over-estimate the panel DEA model estimated efficiency if and only if the common spatial and temporal variations masks the true variations in the input and output variables. The pooled DEA model estimated efficiency measure would underestimate the panel DEA model estimated efficiency measure if the true (rather than common spatial and temporal variations) variations in the input and output variables explain efficiency changes. However, this is an empirical question and needs to be evaluated as it depends on the spatial and temporal distribution of the panel data. This is the primary research question.

In addition to comparing the pooled and panel DEA efficiency measures, differences in efficiency measures estimated using alternative panel DEA models are also evaluated. These alternative panel estimators involve estimating the variance components in the first stage that are used to transform the input and output variables. Unlike the use of traditional residuals to transform the input and output variables, here the normal errors (v) given one-side errors (u) from the pooled, cross-sectional, time-series and within stochastic frontier analysis models are used to develop alternative panel DEA models.

Since the panel, DEA models are based on the use of normal errors (v) recovered from alternative stochastic frontier analysis models, we expect the efficiency measures estimated by panel WH, AM and SA data envelopment analysis models to be different. The difference in magnitude and shares of the spatial, temporal and remaining error variances are reflected in the difference in efficiency measures estimated by the alternative panel DEA models? This is the secondary research question.

2 Conceptual framework of panel DEA models

In DEA, the technology that transforms inputs, $x = (x_1, x_2, ..., x_I) \in \Re_+^I$ into outputs, $y = (y_1, y_2, ..., y_j) \in \Re_+^J$ is represented by the output $P(x)$, input $L(y)$ and graph $G(x, y)$ set. The output distance function is defined as the scalar expansion of output for given input. Input distance function is defined in terms of scalar shrinkage of observed inputs with output held fixed. The graph distance function is defined in terms of scalar expansion of output and simultaneous scalar shrinkage of observed inputs. The efficiency measure can be computed either by Shephard's concept or Farrell's concept of efficiency defined by distance function (Diewert, 1992). The output, input and graph set is presented as:

$$\begin{aligned} &P(x) = \{y : x \,\text{can produce}\, y\} \\ &L(y) = \{x : y \,\text{is produced by}\, x;\} \\ &G(x, y) = \{(x, y) : \text{collection of feasible}\, x \text{ and}\, y\} \end{aligned} \tag{1}$$

and follows the properties of strong disposability of the outputs and inputs, and is estimated under constant returns to scale (CRS) and variable returns to scale (VRS). All the DEA models are estimated under CRS and VRS technology, which allows the recovery of total, pure and scale efficiency measures. Scale efficiency is computed as the ratio of efficiency measures estimated under CRS over VRS technology.

The strong disposal graph set defined in the last line of equation (1) can be represented by the graph distance function and can be used to calculate the efficiency measure. The pooled-DEA model is used to evaluate the efficiency measures for each state and year using reference production possibilities set S represented by years $t = 1, ..., T$ and states $n = 1, ..., N$. The graph distance function is represented as:

$$\begin{aligned} &GR^S(x, y)^{-1} = \min_{\theta, z} \left\{\theta : (\theta^{-1}x, \theta y) \in GR^S(x, y)\right\} \\ &or \\ &\min_{\theta, z} \quad st. \qquad \theta^{-1}y \leq Yz \qquad \text{where}\, Y = \left(y^1, y^2, ..., y^S\right) \\ &\qquad\qquad\qquad \theta x \geq Xz \qquad\qquad X = \left(x^1, x^2, ..., x^S\right) \\ &\qquad\qquad\qquad z = 1(z \geq 0) \qquad vrs(crs) \end{aligned} \tag{2}$$

Equation (2) is used to estimate the pooled DEA model efficiency measures.

2.1 Panel DEA models

The pooled DEA model defined in equation (2) can be extended to one-way (accounts for spatial or temporal variations) and two-way (accounts for both spatial and temporal variations) random effects panel DEA models. It is pedagogical to present the cross-section (equation 4), time-series (equation 5) and within cross-section time-series (equation 6) DEA models leading to the panel DEA model (equation 3). Here, the panel DEA model is presented and followed by the three independent DEA models. To account for spatial and temporal variations, the two-way random effects panel DEA model is presented. To estimate the two-way panel random effects DEA model, equation (2) is transformed as:

$$
\begin{aligned}
GR^S(x^*_{nt}, y^*_{nt})^{-1} &= \min_{\theta,z} \ \{\theta \ : \ (\theta^{-1}x^*_{nt}, \theta\, y^*_{nt})\} \\
&\quad or \\
\min_{\theta,z} \theta \ \ s.t. \ \ \theta^{-1}\, y^*_{nt} &\leq Y^* z \qquad \text{where} \ \ Y = (y^*_{11}, y^*_{12}, ..., y^*_{NT}) \\
\theta\, x^*_{nt} &\geq X^* z \qquad\qquad X = (x^*_{11}, x^2_{12}, ..., x^*_{NT}) \\
z &= 1\,(z \geq 0) \qquad\qquad vrs\,(crs)
\end{aligned}
\tag{3}
$$

where $y^*_{nt} = \Omega^{-1/2}\, y_{nt}$; $\mathbf{x}^*_{nt} = \Omega^{-1/2}\, \mathbf{x}_{nt}$ and $\Omega \equiv \sigma^2_n\ (I_N \otimes \iota_T) + \sigma^2_t\,(I_T \otimes \iota_N) + \sigma^2_{nt}\,(I_N \otimes I_T)$ or $y^*_{nt} = y_{nt} - \theta_1 y_{n\bullet} - \theta_2 y_{\bullet\, t} + \theta_3 y_{\bullet\bullet}$ and $\theta_1 = 1 - \left(\frac{\sigma_{nt}}{\sqrt{\varphi_2}}\right)$, $\theta_2 = 1 - \left(\frac{\sigma_{nt}}{\sqrt{\varphi_3}}\right)$ and $\theta_3 = \theta_1 + \theta_2 + \left(\frac{\sigma_{nt}}{\sqrt{\varphi_4}}\right) - 1$ with $y_{i\bullet}$,$y_{\bullet\, t}$ and $y_{\bullet\bullet}$ representing the cross-section, time-series, and the overall mean of the variable and computed as $y_{n\bullet} = \frac{\sum_{t=1}^{T} y_{n\,t}}{T}$, $y_{\bullet t} = \frac{\sum_{n=1}^{N} y_{nt}}{N}$ and $y_{\bullet\bullet} = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T} y_{it}}{NT}$, respectively. The I_N and I_T (ι_N and ι_T) represent an identity matrix (vector of ones) of N and T (T and N) dimensions, respectively.

Instead of using traditional error-based variances, stochastic frontier analysis normal error (v) based variances are used to transform the data. The phi's $\hat{\varphi}_2 = T\sigma^2_n + \sigma^2_{nt}$, $\hat{\varphi}_3 = N\sigma^2_t + \sigma^2_{nt}$, and $\hat{\varphi}_4 = T\sigma^2_n + N\sigma^2_t + \sigma^2_{nt}$ are computed from the variances of between cross-section, between time-period and within cross-section time-period normal errors (v) of stochastic frontier analysis models. The phi's used in the computation of the thetas (θ_1, θ_2, and θ_3) are obtained from variances of the between cross-section, between time-period and within cross-section time-period normal errors (v) of stochastic frontier analysis models. The normal errors (v) of stochastic frontier analysis models that form the basis for phi's used in the computation of the thetas are estimated using alternative panel estimators, that is, the WH, AM and SA stochastic frontier analysis models. The WH

approach uses the normal errors (v) estimated from the pooled stochastic frontier analysis model to transform the data. The normal errors (v) estimated from within stochastic frontier analysis model is used to transform the data in the AM approach. The SA approach involves the normal errors (v) estimated from the following three stochastic frontier analysis models: between cross-section stochastic frontier analysis model (equation 4); between time-series stochastic frontier analysis model (equation 5); and within cross-section and time-series stochastic frontier analysis model (equation 6).

2.1.1 Between Cross-section stochastic frontier analysis model

The between cross-section stochastic frontier analysis model is characterized as:

$$y_n = f(x_n) + v_n - u_n \tag{4}$$

where $y_n = \frac{\sum_{t=1}^{T} y_{nt}}{T}$ and $x_n = \frac{\sum_{t=1}^{T} x_{nt}}{T}$ is the mean of each cross-section unit over time for output and input, respectively. This allows the computation of the efficiency measures of the cross-sectional units that are temporally invariant.

2.1.2 Between Time-series stochastic frontier analysis model

Similarly, the spatially invariant temporal efficiency measures are computed by the following graph distance function represented as:

$$y_t = f(x_t) + v_t - u_t \tag{5}$$

where $y_t = \frac{\sum_{n=1}^{N} y_{nt}}{N}$ and $x_t = \frac{\sum_{n=1}^{N} x_{nt}}{N}$ is the mean of each year across cross-section units for output and input, respectively.

2.1.3 Within Cross-section and Time-series stochastic frontier analysis model

Finally, the spatially and temporal adjusted, within cross-section time-series efficiency measures are computed by the following graph distance function represented as:

$$\bar{y}_{nt} = f(\bar{x}_{nt}) + v_{nt} - u_{nt} \tag{6}$$

where $\bar{y}_{nt} = y_{nt} - y_{n\bullet} - y_{\bullet t} + y_{\bullet\bullet}$ and $\bar{x}_{nt} = x_{nt} - x_{n\bullet} - x_{\bullet t} + x_{\bullet\bullet}$ is the within cross-section time-series that has taken into account the spatial and temporal variations

for output and input, respectively. The with $y_{i\bullet}$,$y_{\bullet t}$ and $y_{\bullet\bullet}$ representing the cross-section, time-series, and the overall mean of the variable and computed as $y_{n\bullet} = \frac{\sum_{t=1}^{T} y_{nt}}{T}$, $y_{\bullet t} = \frac{\sum_{n=1}^{N} y_{nt}}{N}$ and $y_{\bullet\bullet} = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T} y_{it}}{NT}$, respectively.

3 Data and Variables used in the Analysis

The U.S. Department of Agriculture's Economic Research Service constructs and publishes the state and aggregate production accounts for the farm sector. The data are available at www.ers.usda.gov/data/agproductivity/. The features of the state and national production accounts are consistent with the gross output model of production and are well documented in Ball et al. (1999).

Insert Table ??

The annual growth rate of output and input variables employed in the estimation of productivity for the period 1960 to 2004 by state is presented in Table **??**. All the variables are quantity index consistent with the production theory, which uses input and output quantities. An aggregate output and input quantity index is used in the estimation due to the sensitivity or curse of a dimensionality problem of linear programming DEA efficiency measures (Shaik et al., 2012). Annual growth rate is defined as $\left[\left(\frac{x_{t+1}}{x_t}\right)^{\frac{1}{n}} - 1\right] * 100$ where x is the input or output variable and n is the number of years in the time period. Average annual output growth rate across all 48 states is 1.31 percent with a standard deviation of 0.78 percent. Similarly, the average and standard deviation of annual input growth rate is -0.34 percent and 0.89 percent, respectively.

4 Empirical Application and Results

To evaluate the importance of spatial and temporal variations on total technical efficiency (estimated under CRS technology), pure technical efficiency (estimated under VRS technology) and scale efficiency (computed as the ratio of CRV/VRS), the panel DEA models are compared to the pooled DEA model. Second, the difference between WH, AM and SA panel DEA models are evaluated by pairwise comparison of total efficiency estimated under CRS technology. The pooled and panel DEA models are estimated by SAS and R, and the transformation of data for alternative panel models is estimated using the matrix language in SAS.

4.1 Pooled versus Alternative Panel DEA models estimated Efficiency

The efficiency measures from pooled DEA model defined in equation (2) and the three alternative panel models (WH, AM and SA) based on equation (3) are estimated. To examine the difference in magnitude, the pooled DEA model estimated efficiency measures and the ratio of the panel to the pooled model estimated efficiency measures are presented by individual states. Tables **??**, **??** and **??** present the pooled efficiency measures and the ratio of panel over pooled DEA model estimated total efficiency (estimated under CRS technology), pure efficiency (estimated under VRS technology) and scale efficiency (ratio of CRS/VRS), respectively.

Insert Tables ??, ?? and ??

Using the pooled DEA model, the average total, pure and scale efficiency measures across 48 states from 1960 to 2004 are estimated to be 80.4%, 84% and 95.7% efficient, respectively. The standard deviation is 3.2%, 3.4% and 3.8%, respectively for the total, pure and scale efficiency measures.

Tables **??**, **??** and **??** present the ratios of the panel over pooled models for total, pure and scale efficiency, respectively. The ratios of average total, pure and scale efficiency measures for the WH, AM and SA DEA models are higher compared to the pooled model. Accounting for the spatial and temporal variations by the panel WH DEA model indicates an 8.91%, 5.65% and 3.09% increase in the total, pure and scale efficiency measures over the pooled DEA model. The panel AM DEA model indicates a 9.53%, 6.11% and 3.23% increase in the total, pure and scale efficiency measures over the pooled DEA model. Similarly, the panel SA DEA model suggests a 9.73%, 6.04% and 3.49% increase in the total, pure and scale efficiency measures over the pooled DEA model. There is a consistent pattern of over-estimation of the pooled DEA model efficiency measures relative to the three panel models in the US with few exceptions.

The ratio of the total efficiency measures estimated by panel and pooled DEA model by state suggests that the pooled DEA model overestimates the true efficiency measures compared to the panel WH, AM and SA DEA models in 2, 1 and 2 states, respectively. The ratio of the scale efficiency measures suggests the pooled DEA model overestimates the true efficiency compared to panel WH, AM and SA DEA models in 9, 7 and 11 states, respectively. In contrast, the ratio of the pure efficiency suggests the pooled DEA model overestimates the true effi-

ciency compared to the panel WH, AM and SA DEA models in 1, 5 and 0 states, respectively.

Statistical tests are computed to evaluate difference in efficiency measures between pooled and alternative panel DEA models. Table **??** presents the mean, standard deviation and equal and unequal variances of efficiency measures estimated by pooled and AM, SA and WH panel models. Efficiency measures estimated using alternative panel DEA models are statistically different from pooled model efficiency estimates.

Insert Table ??

To summarize, accounting for the spatial and temporal variations in the panel DEA models provides mixed signals as it depends upon: a) if DEA models are estimated under constant returns to scale technology, variable returns to scale technology and scale; and b) the choice of alternative panel DEA model estimated efficiency measures used to transform the data.

4.2 Difference in Efficiency measures between WH, AM and SA Panel DEA models

In the context of the alternative panel DEA models, several possibilities exist to transform the data, namely the use of efficiency-based variance estimated by : a) pooled SFA model (equation 2) as in WH; b) within-SFA model (equation 6) as in AM; or c) between cross section-SFA model (equation 4), between time series-SFA model (equation 5) and within-SFA model (equation 6) as in SA. To examine the magnitude of the difference between the three alternative panel models, the ratio of the panel SA to the panel WH DEA model and the ratio of the panel SA to the panel AM DEA model estimated total efficiency measures are computed.

Statistical tests are computed to evaluate difference in efficiency measures between pooled and alternative panel DEA models. Table **??** presents the mean, standard deviation and equal and unequal variances of efficiency measures estimated by AM, SA and WH panel models.

Insert Tables ??

The ratio of the total efficiency measures estimated by the panel SA DEA model and the panel WH and AM DEA model is less than one. This suggests the use of the pooled SFA and within SFA estimated normal error-based variance

to transform the data as in the panel WH and AM DEA models respectively, overestimates the panel SA DEA model.T-test between WH or AM and SA panel DEA model estimated efficiency measures suggest statistical difference for equal and unequal variances.

The ratio of the panel WH DEA model to the panel AM DEA model estimated total efficiency is greater than one. However, the difference in efficiency measures estimated using WH and AM panel DEA models are not statistically difference based on the t-test.

Overall, there is difference in efficiency measures between alternative panel DEA models that use normal error-based variance estimated from the pooled (WH), within (AM) and between cross-section, between time-series and within cross-section time-series (SA) DEA models.

5 Conclusions

This research is unique as it integrates the statistical underpinnings of alternative two-way random effects econometric estimators with linear programming DEA models to develop the panel DEA models in two stages. The first stage involves recovering the variances of normal errors, v used in the transformation of the variables from pooled (Wallace-Hussain in 1969); within (Amemiya in 1971); or within, cross-sectional and time-series (Swamy-Arora in 1972) stochastic frontier analysis (SFA) models. The second stage involves estimating panel DEA efficiency measures using transformation based on alternative two-way random effect econometric estimators in the first stage.

The alternative panel DEA models presented help to evaluate the importance of accounting for the spatial and temporal variations over the pooled DEA model. Specifically, the total efficiency (estimated under CRS technology), pure efficiency (estimated under VRS technology) and scale efficiency (ratio of CRS/VRS) of agricultural production are estimated by the panel WH, AM and SA DEA models. The use of normal error-based variance from stochastic frontier analysis compared to the traditional residuals is proposed as an alternative for linear programming DEA models. Finally, the importance of the use of normal error-based variance estimated from the pooled SFA model (as in the panel WH DEA model), within-SFA model (as in the panel AM DEA model) or cross section-SFA model, time series-SFA model and within-SFA model (as in the panel SA DEA model) is evaluated.

An empirical application to the 48 U.S. states for the 1960 to 2004 period

suggests differences in the efficiency measures estimated by the pooled and panel WH, AM and SA DEA models. Further, the panel DEA model estimated efficiency measures are different if estimated by SA relative to the WH and AM panel estimator that uses the pooled DEA efficiency measures and within-DEA efficiency measures, respectively.

Future studies should examine the implications of large cross-sectional units compared to number of years with multiple output and input data. In addition, statistical properties need to be developed to evaluate the robustness of the pooled and panel DEA model estimated efficiency measures.

References

Amemiya, T. (1971). The Estimation of Variances in a Variance-Components Model. *International Economic Review*, 12:1-13.

Ball, V. E., F. Gollop, A. Kelly-Hawke, and G. Swinand. (1999). Patterns of Productivity Growth in the U.S. Farm Sector: Linking State and Aggregate Models. *American Journal of Agricultural Economics*, 81:164-179.7

Bogetoft, P., and Otto, L. (2011). *Benchmarking with DEA, SFA and R*, International Series in Operations Research and Management Science, Volume 157, New York.

Boles, J. N., (1966). Efficiency Squared - Efficient Computation of Efficiency Indexes. *Proceedings Western Farm Economic Association*, 137-142.

Bressler, R. G., (1966). The Measurement of Productive Efficiency. *Proceedings Western Farm Economic Association*, 129-136.

Briec, W (1997). A Graph-type Extension of Farrell Technical Efficiency Measure, *Journal of Productivity Analysis*, 8:95-110.

Bureau, J C., R. Fare, S. Grosskopf. (1995). A Comparison of Three Nonparametric Measures of Productivity Growth in European and United State Agriculture. *Journal of Agricultural Economics*, 46:309-326.

Charnes, A., W. W. Cooper, and E. Rhodes. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operations Research*, 2:429-444.

Diewert, W.E. (1992). Fisher Ideal Output, Input and Productivity Indexes Revisited. *Journal of Productivity Analysis*, 3:211-248.

Farrell, M.J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society Series A*, 120:253-281.

Farrell, M.J. and M. Fieldhouse. (1962). Estimating Efficient Production Functions under Increasing Returns to Scale. *Journal of the Royal Statistical Society*, 125:252-267.

Fare, R., S. Grosskopf and C. A. K. Lovell. (1994). *Production Frontiers*. Cambridge: Cambridge University Press, 1994.

Ferrier, G .D., and Trivitt, J. S. (2013). Incorporating Quality into the Measurement of Hospital Efficiency: A Double DEA Approach. *Journal of Productivity Analysis*, 40:337-355.

Kneip A, Simar L, Wilson PW (2008) Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models. *Economic Theory*, 24:1663-1697

Lovell, C. A. K. (1994). Linear Programming Approaches to the Measurement and Analysis of Productive Efficiency. *Top*, 2:175-224.

Luenberger, D. G (1992). Benefit function and duality, *Journal of Mathematical Economics*, 21:461-481.

Malmquist, S. (1953). Index Numbers and Indifference Surfaces. *Trabajos de Estatistica*, 4:209-242.

Moorsteen, R.H. (1961). On Measuring Productive Potential and Relative Efficiency. *Quarterly Journal of Economics*, 75:451-467.

Seitz, W.D. (1966). Efficiency Measures for Steam-Electric Generating Plants. *Proceedings Western Farm Economic Association*, 143-151.

Shaik, S., A. Mishra, and J.A. Atwood. (2012). Aggregation Issues in the Estimation of Linear Programming Productivity Measures. *Journal of Applied Economics* XV (1): 169-187.

Shephard, R.W. (1953). *Cost and Production Functions*, Princeton: Princeton University Press.

Shephard, R.W. (1970). *Theory of Cost and Production*, Princeton: Princeton University Press.

Sitorus, B.L. (1966). Productive Efficiency and Redundant Factors of Production in Traditional Agricultural of Underdeveloped Countries: A Note on Measurement. *Proceedings Western Farm Economic Association*, 153-158.

Simar, L., and Wilson, P.W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44, 49-61.

Simar, L., and Wilson, P.W. (2000). A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics*, 27:779-802.

Simar, L., and Wilson, P W. (2011). Two-stage DEA: caveat emptor. *Journal of Productivity Analysis*, 36:1-20.

Sitorus, B.L (1966). Productive Efficiency and Redundant Factors of Production in Traditional Agricultural of Underdeveloped Countries: A Note on Measurement. *Proceedings Western Farm Economic Association.*

Swamy, P.A.V.B. and S.S. Arora. (1972). The Exact Finite Sample Properties of the Estimators of Coefficients in the Error Component Regression Models. *Econometrica*, 40:261-275.

Tulkens, H., and Vanden Eeckaut, P., (1995). Non-Parametric Efficiency, Progress and Regress Measure for Panel Data: Methodological Aspects. *European Journal of Operational Research*, 80:474-499.

Wallace, T.D., A. Hussain. (1969). The Use of Error Components Models in Combining Cross-Section with Time Series Data. *Econometrica*, 37:55-72.

Table 1: Annual Growth Rate of Output and Input by State, 1960-2004

State	Output	Input	State	Output	Input
Alabama	1.61%	0.31%	Nebraska	2.24%	0.65%
Arizona	1.45%	-0.06%	Nevada	1.75%	0.53%
Arkansas	2.73%	0.81%	New Hampshire	-0.21%	-2.14%
California	2.24%	0.59%	New Jersey	-0.16%	-1.78%
Colorado	1.70%	0.61%	New Mexico	2.22%	0.79%
Connecticut	0.37%	-1.76%	New York	0.28%	-1.16%
Delaware	2.44%	0.65%	North Carolina	1.91%	0.09%
Florida	2.05%	0.62%	North Dakota	1.84%	-0.03%
Georgia	2.15%	0.26%	Ohio	1.07%	-1.04%
Idaho	2.40%	0.41%	Oklahoma	1.11%	0.54%
Illinois	1.23%	-0.69%	Oregon	2.17%	-0.37%
Indiana	1.42%	-0.82%	Pennsylvania	1.28%	-0.49%
Iowa	1.33%	-0.50%	Rhode Island	-0.39%	-2.77%
Kansas	1.70%	0.67%	South Carolina	1.07%	-0.52%
Kentucky	1.42%	-0.16%	South Dakota	1.64%	0.14%
Louisiana	1.60%	-0.29%	Tennessee	0.78%	-0.32%
Maine	0.03%	-1.81%	Texas	1.58%	0.46%
Maryland	1.47%	-0.32%	Utah	1.45%	-0.07%
Massachusetts	-0.42%	-2.63%	Vermont	0.24%	-1.34%
Michigan	1.36%	-1.00%	Virginia	1.11%	-0.39%
Minnesota	1.45%	-0.37%	Washington	2.40%	0.69%
Mississippi	1.72%	-0.23%	West Virginia	0.29%	-0.97%
Missouri	1.13%	-0.46%	Wisconsin	0.70%	-0.85%
Montana	1.27%	-0.09%	Wyoming	0.89%	0.25%

Table 2: Average pooled , panel Wallace-Hussain, Amemiya and Swamy-Arora total (CRS) efficiency

State	pooled	Ratio of Panel/pooled			State	pooled	Ratio of Panel/pooled		
	DEA	WH	AM	SA		DEA	WH	AM	SA
Alabama	0.809	1.076	1.080	1.089	Nebraska	0.830	1.044	1.035	1.048
Arizona	0.772	1.115	1.122	1.129	Nevada	0.843	1.066	1.071	1.076
Arkansas	0.792	1.000	0.957	0.982	New Hampshire	0.783	1.178	1.151	1.156
California	0.832	1.033	1.022	1.035	New Jersey	0.795	1.205	1.170	1.174
Colorado	0.821	1.053	1.048	1.058	New Mexico	0.812	1.041	1.029	1.042
Connecticut	0.755	1.193	1.223	1.222	New York	0.823	1.158	1.122	1.127
Delaware	0.785	1.032	1.005	1.025	North Carolina	0.796	1.064	1.052	1.063
Florida	0.832	1.031	1.018	1.031	North Dakota	0.763	1.060	1.043	1.058
Georgia	0.810	1.033	1.018	1.032	Ohio	0.775	1.150	1.166	1.169
Idaho	0.784	1.032	1.009	1.027	Oklahoma	0.866	1.023	1.010	1.017
Illinois	0.761	1.140	1.159	1.163	Oregon	0.743	1.068	1.048	1.065
Indiana	0.747	1.135	1.143	1.150	Pennsylvania	0.815	1.112	1.126	1.130
Iowa	0.762	1.126	1.137	1.143	Rhode Island	0.758	1.145	1.138	1.141
Kansas	0.842	1.048	1.045	1.054	South Carolina	0.816	1.081	1.070	1.075
Kentucky	0.835	1.061	1.044	1.053	South Dakota	0.810	1.077	1.082	1.090
Louisiana	0.825	1.053	1.042	1.053	Tennessee	0.817	1.126	1.134	1.137
Maine	0.776	1.187	1.173	1.178	Texas	0.850	1.078	1.083	1.087
Maryland	0.794	1.121	1.141	1.145	Utah	0.831	1.092	1.103	1.106
Massachusetts	0.756	1.252	1.219	1.223	Vermont	0.830	1.143	1.130	1.134
Michigan	0.780	1.133	1.149	1.153	Virginia	0.832	1.099	1.090	1.094
Minnesota	0.774	1.111	1.111	1.120	Washington	0.823	0.990	0.949	0.973
Mississippi	0.806	1.043	1.033	1.044	West Virginia	0.834	1.127	1.134	1.137
Missouri	0.806	1.130	1.135	1.140	Wisconsin	0.777	1.173	1.201	1.201
Montana	0.829	1.057	1.054	1.061	Wyoming	0.873	1.081	1.055	1.061

Table 3: Average pooled , panel Wallace-Hussain, Amemiya and Swamy-Arora pure (VRS) efficiency

State	pooled	Ratio of Panel/pooled			State	pooled	Ratio of Panel/pooled		
	DEA	WH	AM	SA		DEA	WH	AM	SA
Alabama	0.882	1.028	1.016	1.020	Nebraska	0.886	1.025	1.018	1.022
Arizona	0.821	1.042	1.043	1.050	Nevada	0.843	1.033	1.024	1.030
Arkansas	0.873	1.011	0.993	1.000	New Hampshire	0.869	1.042	1.035	1.037
California	0.888	1.023	1.013	1.016	New Jersey	0.814	1.121	1.106	1.106
Colorado	0.845	1.019	1.007	1.015	New Mexico	0.854	1.010	0.997	1.005
Connecticut	0.755	1.167	1.174	1.176	New York	0.865	1.060	1.051	1.053
Delaware	0.843	1.026	1.012	1.020	North Carolina	0.799	1.069	1.074	1.077
Florida	0.876	1.007	0.995	1.002	North Dakota	0.892	1.014	1.006	1.010
Georgia	0.855	1.045	1.039	1.044	Ohio	0.775	1.130	1.142	1.144
Idaho	0.872	1.005	0.991	0.998	Oklahoma	0.865	1.040	1.011	1.017
Illinois	0.804	1.103	1.112	1.114	Oregon	0.837	1.047	1.037	1.043
Indiana	0.792	1.091	1.092	1.099	Pennsylvania	0.874	1.049	1.041	1.042
Iowa	0.795	1.109	1.090	1.094	Rhode Island	0.786	1.123	1.124	1.126
Kansas	0.870	1.032	1.020	1.026	South Carolina	0.792	1.130	1.142	1.146
Kentucky	0.840	1.058	1.052	1.057	South Dakota	0.819	1.066	1.063	1.069
Louisiana	0.875	1.047	1.043	1.045	Tennessee	0.830	1.089	1.089	1.093
Maine	0.843	1.071	1.065	1.066	Texas	0.893	1.022	1.022	1.025
Maryland	0.830	1.071	1.069	1.072	Utah	0.858	1.035	1.042	1.043
Massachusetts	0.831	1.069	1.067	1.068	Vermont	0.844	1.084	1.077	1.082
Michigan	0.803	1.086	1.091	1.094	Virginia	0.816	1.087	1.103	1.106
Minnesota	0.820	1.064	1.063	1.068	Washington	0.869	1.020	1.019	1.021
Mississippi	0.843	1.053	1.058	1.062	West Virginia	0.795	1.138	1.147	1.147
Missouri	0.844	1.087	1.078	1.080	Wisconsin	0.815	1.124	1.129	1.132
Montana	0.883	1.013	0.999	1.004	Wyoming	0.857	1.047	1.027	1.032

Table 4: Average pooled , panel Wallace-Hussain, Amemiya and Swamy-Arora scale efficiency

State	pooled	Ratio of Panel/pooled			State	pooled	Ratio of Panel/pooled		
	DEA	WH	AM	SA		DEA	WH	AM	SA
Alabama	0.917	1.047	1.063	1.067	Nebraska	0.937	1.019	1.017	1.026
Arizona	0.941	1.070	1.076	1.076	Nevada	1.000	1.033	1.046	1.045
Arkansas	0.907	0.989	0.964	0.982	New Hampshire	0.901	1.130	1.111	1.115
California	0.937	1.009	1.010	1.018	New Jersey	0.976	1.075	1.058	1.062
Colorado	0.972	1.033	1.041	1.043	New Mexico	0.950	1.031	1.032	1.037
Connecticut	1.000	1.022	1.041	1.040	New York	0.952	1.092	1.067	1.070
Delaware	0.931	1.006	0.993	1.004	North Carolina	0.996	0.996	0.979	0.987
Florida	0.950	1.024	1.023	1.029	North Dakota	0.855	1.045	1.037	1.048
Georgia	0.947	0.989	0.979	0.989	Ohio	1.001	1.017	1.021	1.022
Idaho	0.899	1.027	1.018	1.029	Oklahoma	1.001	0.984	1.000	1.000
Illinois	0.947	1.034	1.042	1.044	Oregon	0.888	1.020	1.010	1.021
Indiana	0.943	1.041	1.047	1.047	Pennsylvania	0.932	1.060	1.082	1.084
Iowa	0.958	1.015	1.043	1.045	Rhode Island	0.965	1.020	1.013	1.013
Kansas	0.967	1.015	1.024	1.027	South Carolina	1.030	0.957	0.937	0.938
Kentucky	0.994	1.003	0.993	0.996	South Dakota	0.989	1.010	1.018	1.019
Louisiana	0.943	1.005	0.999	1.008	Tennessee	0.983	1.034	1.041	1.040
Maine	0.920	1.108	1.101	1.105	Texas	0.951	1.056	1.059	1.060
Maryland	0.957	1.046	1.067	1.068	Utah	0.970	1.055	1.059	1.060
Massachusetts	0.910	1.171	1.143	1.146	Vermont	0.983	1.054	1.049	1.048
Michigan	0.972	1.044	1.053	1.053	Virginia	1.019	1.011	0.988	0.989
Minnesota	0.944	1.044	1.045	1.049	Washington	0.947	0.970	0.932	0.953
Mississippi	0.956	0.990	0.976	0.983	West Virginia	1.050	0.990	0.989	0.991
Missouri	0.955	1.039	1.053	1.055	Wisconsin	0.954	1.043	1.064	1.062
Montana	0.938	1.043	1.055	1.057	Wyoming	1.018	1.032	1.027	1.028

Table 5: Statistical results of difference in Efficiency between Pool and Panel Models

Model	Method	Mean	Std Dev	Variances	t-Value	Pr¿—t—
AM panel		0.879	0.073			
Pool		0.804	0.094			
Diff (1-2)	**Pooled**	0.076	0.084	Equal	29.49	¡.0001
Diff (1-2)	**Satterthwaite**	0.076		Unequal	29.49	¡.0001
Pool		0.804	0.094			
SA panel		0.874	0.079			
Diff (1-2)	**Pooled**	-0.071	0.087	Equal	-26.71	¡.0001
Diff (1-2)	**Satterthwaite**	-0.071		Unequal	-26.71	¡.0001
Pool		0.804	0.094			
WH panel		0.881	0.074			
Diff (1-2)	**Pooled**	-0.077	0.085	Equal	-29.96	¡.0001
Diff (1-2)	**Satterthwaite**	-0.077		Unequal	-29.96	¡.0001

Table 6: Statistical results of difference in Efficiency between Panel Models

Model	Method	Mean	StdDev	Variances	t-Value	Pr¿—t—
AM panel		0.879	0.073			
SA panel		0.874	0.079			
Diff (1-2)	**Pooled**	0.005	0.076	Equal	2.16	0.031
Diff (1-2)	**Satterthwaite**	0.005		Unequal	2.16	0.031
SA panel		0.874	0.079			
WH panel		0.881	0.074			
Diff (1-2)	**Pooled**	-0.007	0.077	Equal	-2.82	0.005
Diff (1-2)	**Satterthwaite**	-0.007		Unequal	-2.82	0.005
AM panel		0.879	0.073			
WH panel		0.881	0.074			
Diff (1-2)	**Pooled**	-0.002	0.074	Equal	-0.7	0.481
Diff (1-2)	**Satterthwaite**	-0.002		Unequal	-0.7	0.481

Figure 1: Spatial Variation in Output and Input Quantity Index, 1960-2004

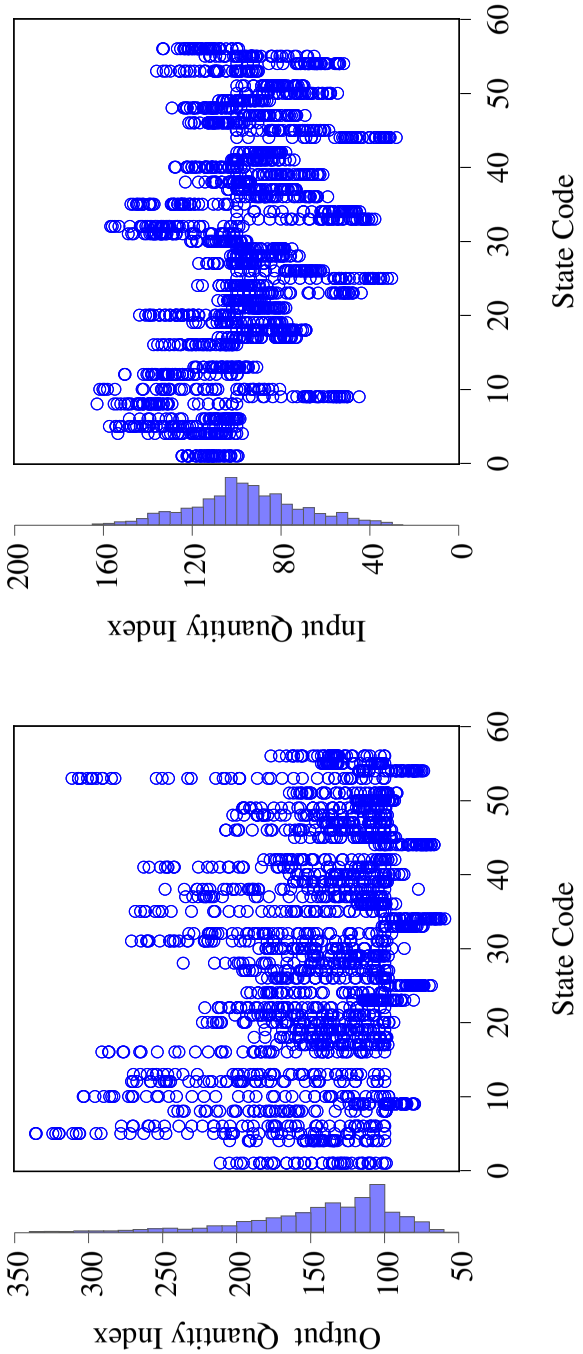
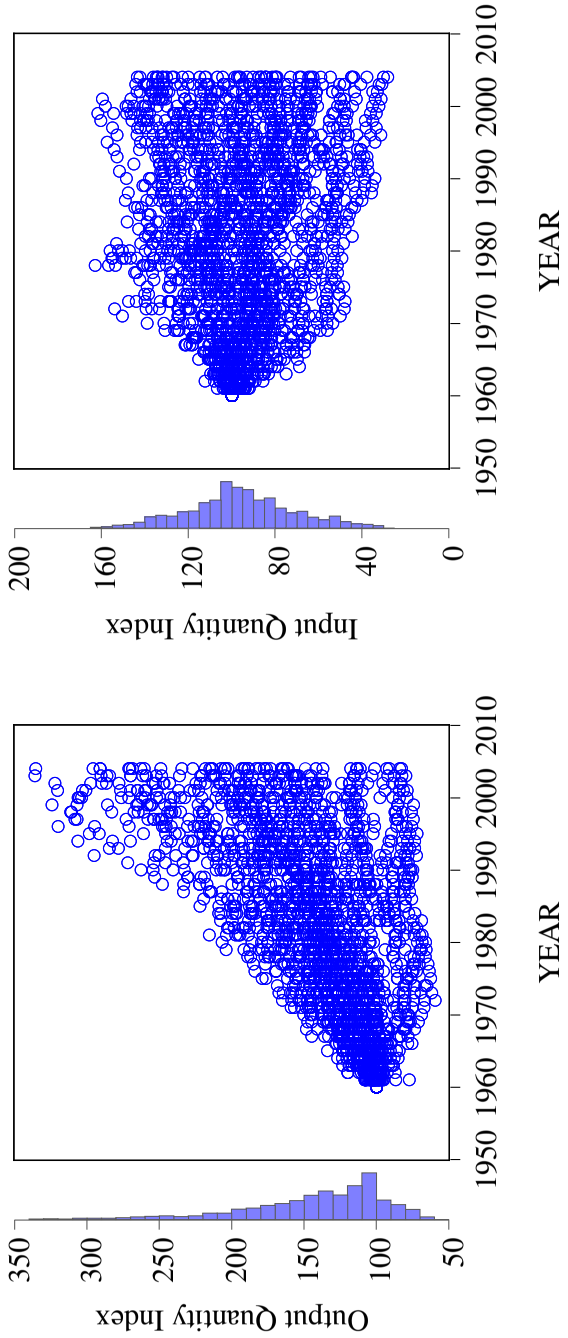


Figure 2: Temporal Variation in Output and Input Quantity Index, 1960-2004



Figures

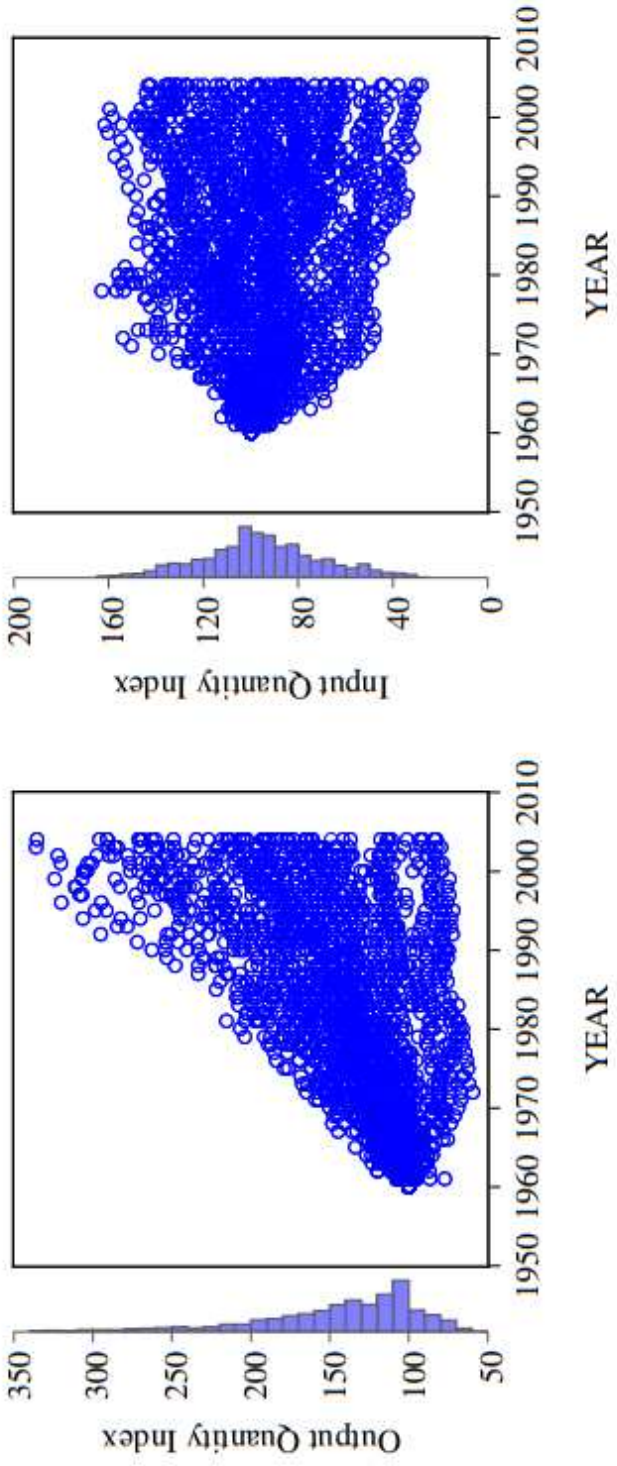


Figure 1

Spatial Variation in Output and Input Quantity Index, 1960-2004

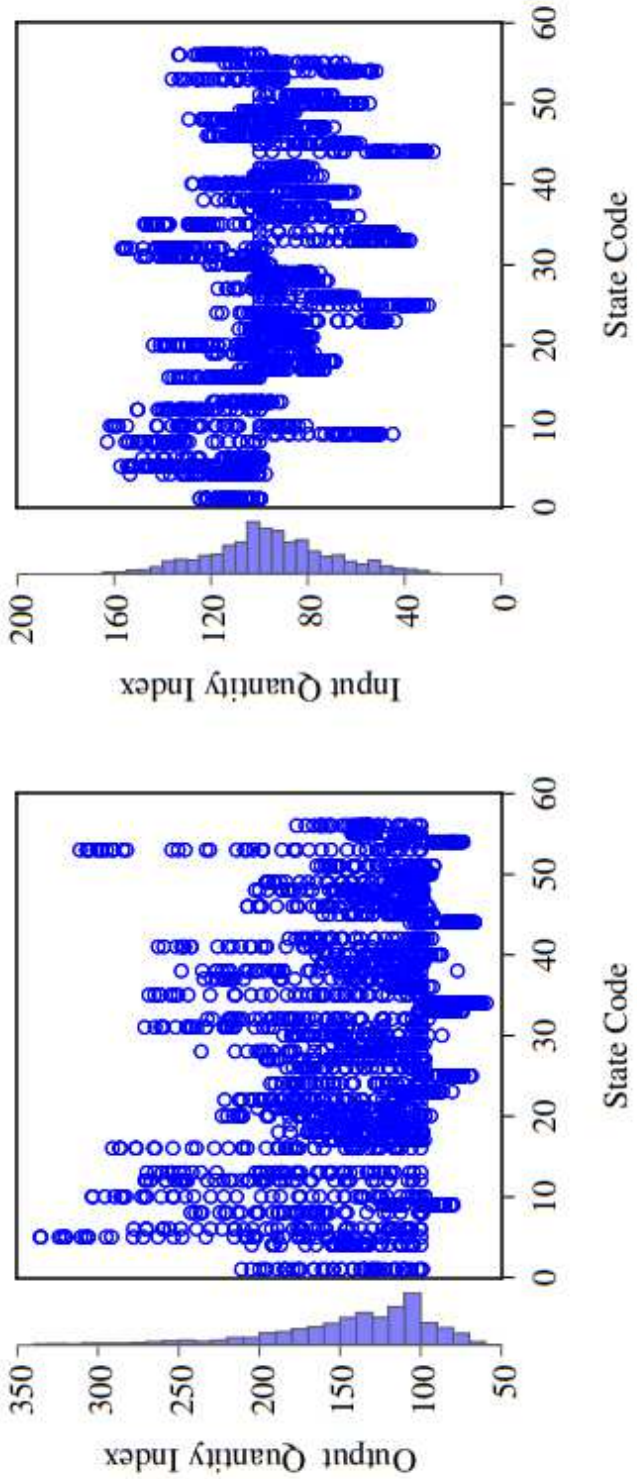


Figure 2

Temporal Variation in Output and Input Quantity Index, 1960-2004