THE INTERNATIONAL SOCIETY OF PRECISION AGRICULTURE PRESENTS THE
13th INTERNATIONAL CONFERENCE ON
PRECISION AGRICULTURE
July 31-August 4, 2016 • St. Louis, Missouri USA

Site-specific Scale Efficiency Determined by Data Envelopment Analysis of Precision Agriculture Field Data

Jacob L. Maurer[1], Terry W. Griffin[2], Ajay Sharda[3]

[1]Precision Agriculture Specialist, Highland Community College, Wamego, Kansas,

[2]Assistant Professor, Department of Agricultural Economics, Kansas State University, Manhattan, Kansas,

[3]Assistant Professor, Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, Kansas

A paper from the Proceedings of the
13th International Conference on Precision Agriculture
July 31 – August 4, 2016
St. Louis, Missouri, USA

Abstract. *Since its inception and acceptance as a benchmarking tool within the economics literature, data envelopment analysis (DEA) has been used primarily as a means of calculating and ranking whole-farm entities marked as decision making units (DMU) against one another. Within this study, instead of ranking the entire farm operation against similar peers that encompass the study, individual data points from within the field are evaluated to analyze the site-specific technical efficiencies estimated at sub-field locations. A hypothetical grid superimposed upon a field creates the DMU's so that scale efficiency can be visually assessed in a map and spatially analyzed. Input variables include as-applied inputs, geospatial data on soil characteristics, and aerial remotely-sensed imagery. Output variables were based upon yield monitor sensors from harvest equipment from one or more years and therefore one or more crops grown in rotation. Both bio-physical agronomic relationships and economic characteristics were evaluated. Analysis can be conducted on either physical units or on the dollar values of these inputs and outputs. The data here are analyzed by superimposing a grid over a production field in Kansas. Once technical efficiencies were calculated for each site-specific grid cell, the results were spatially mapped across the field to form what looks quite similar to a yield map, only instead of yield, the map now represents the site-specific technical efficiency of that particular field. From this point, tests for global and local spatial autocorrelation indicated the presence of spatial effects, further providing true economic insights into the variability generated either by nature or by the farmer.*

These results are useful for the agricultural industry as they represent the first new techniques evaluating efficiency and economics applied to precision agriculture in many years. This initial study can easily be extended to include a farmer's field with a deliberate intervention, i.e. on-farm experiment; where the technical efficiency of the experiment and in particular regression residuals can be assessed. Additional extensions to this technique can be applied to a community of farmers' fields in a big data analysis.

Keywords. *Data Envelopment Analysis, Technical Efficiency, Variable Rate*

Introduction

Using Data Envelopment Analysis (DEA) as a means of benchmarking and measuring efficiencies across firms has been used since it first gained prominence in the late 1950s. While the original framework of the concept remains intact still to this day, much work has been done to broaden the scope. This work takes into account the principles of data envelopment and adds a geographical element by presenting the result of the analysis in a more geo-spatial form rather than a traditional scatter plot. It also differs from other forms because rather than assessing decision making units (DMU) as competing firms within an industry as is typically done, the analysis is conducted within one farm tract. This one tract is then split into multiple pieces as is done in soil sampling. Then, using the exact same parameters as a traditional analysis, a DEA is completed, using the grid cells as DMUs. The results, also, are then interpolated based upon their numerical result back onto the original map so that a spatial analysis and further research may be conducted.

This new analysis layer is created by gathering all input and output data generated on a given farm for a growing year to calculate the *technical efficiency* according to Farrell's *Measurement of Productive Efficiency*. The input data includes planting, fertilizing, irrigation, and chemical application, while the output consists of all harvest yield data. The resulting TE layer can then be used for comparison against bio-physical and meta-data as a means of assisting the grower in determining whether the decisions for the year were successful in generating a profitable output, and not just if, but where, improvement needs to be made should they be necessary.

Because the analysis is being conducted within the same tract of ground instead of across multiple entities, the model increases in its effectiveness and level of realism. The more growers that are a part of a given model, the more factors that potentially hold statistical merit when making comparisons including fertility management, tillage practices, compaction, irrigation, climate, variable-rate versus flat-rate applications, and soil type just to name a few. Likewise, the more of these factors that become “statistically” ignorable due to the scope of the analysis, the less likely that growers will find the results to be relatable within the specificity of his or her operation. This new form of analysis becomes immediately pertinent to the grower because the analysis is conducted only using real data conducted by real equipment found on only one piece of ground, instead of across multiple entities.

Background

Technical Efficiency

Technical Efficiency is defined in relation to a given set of firms, in respect of a given set of factors measured in a specific way. Any change in these specifications will affect the measure (Farrell 1957). In one of the original publications describing the theory of data envelopment analysis, M.J. Farrell first described the concept by using a real-world example from agriculture, an example of which is described here.

Let us suppose that a farm is using two inputs (seed and fertilizer) to grow wheat in an environment with constant returns to scale. The quantities of each input come together as a ratio, demonstrated by the line, *OP* represented in Figure 1. The line, *AA'* refers to the slope equaling the ratio of the input prices. Because of the assumption of constant returns, the graph forms an isoquant. The isoquant *SS'* represents the combination of inputs which would be required if the grower were producing at the optimum efficiency. In mathematical terms, then, *OQ/OP* would then be defined as the technical efficiency of the firm. From there, based upon the slope of *AA'*, it could be seen that producing at the price level of *Q*', at a point still lying along the isoquant *SS'* would be seen as having the greatest price efficiency, equated by *OR/OQ*. Then, should the grower operate in a state which is perfectly efficient overall with regard to both price and cost, then *OR/OP* would represent the *overall* *efficiency* of the farm (Farrell 1957).

Technical efficiency was chosen for this analysis over price efficiency because price is subjective in nature, based upon timing, bargaining, and scale. Farrell mentioned in his research that "technical efficiency of a firm or plant indicates the undisputed gain that can be achieved by simply "gingering up" the management, while its price efficiency indicates the gain that, on certain assumptions about future price structure, can be obtained by varying the input ratios (1957)." This methodology is acceptable in this case because not all growers have the bargaining power necessary to maximize price efficiency, and improper timing in marketing a crop can have a terrific impact upon the success of a given year. The strength of using TE as a means of measuring how effective the decisions of the crop year is that there leaves more progress than simply "buying cheaper" for next year.

Data Envelopment Analysis

Much of the work done with DEA can be attributed to M.J. Farrell's study of Productive Efficiency (PE) in the 1950s, though it found new vogue in agriculture in the mid-1990s as a means of benchmarking farming operations. His research came at a time of incredible growth in industry, but also at a time when production was simply measured as the average calculated input required to produce a given output. Not satisfied with the new weighted indices being generated among similar economists and statisticians, he sought a much more theoretical approach to measure efficiency from an economic standpoint.

As Farrell noted in his Royal Statistical Society article from 1957, "more recently, attempts have been made to construct "indices of efficiency", in which a weighted average of inputs is compared with output. These attempts have naturally run into all the usual index number problems. It is the purpose of this paper to provide a satisfactory measure of productive efficiency – one which takes account of all inputs and yet avoids index number problems – and shows how it can be computed in practice. In doing so, an estimate of the relevant production function is obtained."

The modern-day agriculture industry finds itself staring this very same concept down once again. Many companies and statisticians have sought new, creative measures to calculate and represent how and where farmers are leaving opportunities out in the field. Most of these calculations, though, consist of weighted averages spread across the field reported as an index of production, possibly due to the variability in soil types or growing conditions which exist across a given acreage. The missing link, though, is the economic theory which ties these concepts together. Indeed, the input-output relationship can be easily represented quite effectively in the form of a profit index or "efficiency" conversion displayed spatially across a given geography. The difference, here, lies within the method used to calculate and translate the efficiency of the field's success in growing a crop based upon the given inputs and attributes.

While the concept of DEA may not be new, variations have been adapted of the original theory to better suit the practices of industry. Economists have used DEA to establish benchmarks for the comparison of firms, to analyze the profit-maximizing behaviors of firms, and to measure the effects of technology upon the profitability and performance of firms. In this study, it is used as a means of collectively ranking the abilities of the field to produce a crop efficiently, and then reporting the result back in the form of a geo-referenced image of the field in a manner similar to a yield map.

Methodology

A 100-by-100 foot grid was superimposed over the field. The 100 foot grid cell was chosen based on the most sparse layer of interest, fertilizer application, having a 100-foot swath width. Using the filtered yield data layer, a dissolved buffer was applied to define an area of influence. Given the 31-foot harvester swath, the buffer was set to an arbitrary 10% greater than the harvester swath width to be 17 feet (31*0.5 * 1.1). The dissolved buffer removes buffer overlaps such that the final layer is a single polygon rather than a layer of independent buffered regions. Using the buffered area around yield data layer prevents regions outside the area of interest from being considered during the analysis; areas such as end rows, waterways, and terraces are not desirable since no crop production occurred

In returning to the earlier example within this text, Farrell had stated, “it is never possible to decide precisely how far the fertility of a particular farmer’s land is due to nature and how far to good husbandry. How far the laziness and intractability of a particular firm’s labor force is ingrained and how far the product of bad management” (1957). This model helps to satisfy the burning desire to break this inference because within it there is only one labor force, one management practice, and one product grown to one standard across the entire dataset. The limitation to the accuracy of this model, of course, still lies within the variability of soil types within the field. However, the location and magnitude of the variability can be measured according to soil mapping sensors and further researched and represented using sampling and meta-data. Also, other quasi-factors would be far similar in their impact intra-field rather than intra-industry, as DEA is typically conducted.

Data

The data found within this project originates from a field in Rice County, Kansas. A single, 80-acre field was chosen for analysis due to the quantity and quality of application and harvest yield data for that particular tract. The data was all collected using a John Deere GS3 receiver with RTK over the 2015 wheat growing season. The data included within the analysis includes variable-rate liquid fertilizer application, pesticide and fungicide application, and harvest yield. The field was seeded uniformly but data was not available. The fertilizer application and yield maps are both shown in *Figure 2* and select entries from the dataset are included in *Table 1*.

Table 1 – Select observations from the farm dataset

Longitude*	Latitude*	Elevation	Yield	Finesse	Quilt	UAN	MUKEY
-98.19	38.33	1671	62.2	21.5	5.1	21.5	2733353
-98.19	38.33	1670	63.4	24.4	4.5	24.4	2733353
-98.19	38.33	1670	70.3	22.4	9.3	22.4	2733353
-98.19	38.33	1670	71.0	22.2	11.1	22.2	2733353
-98.19	38.33	1669	68.7	20.7	10.5	20.7	2733353

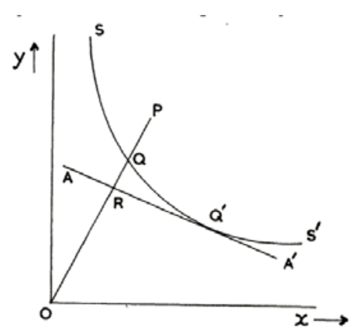

Figure 1 – Farrell’s isoquant diagram (1957)

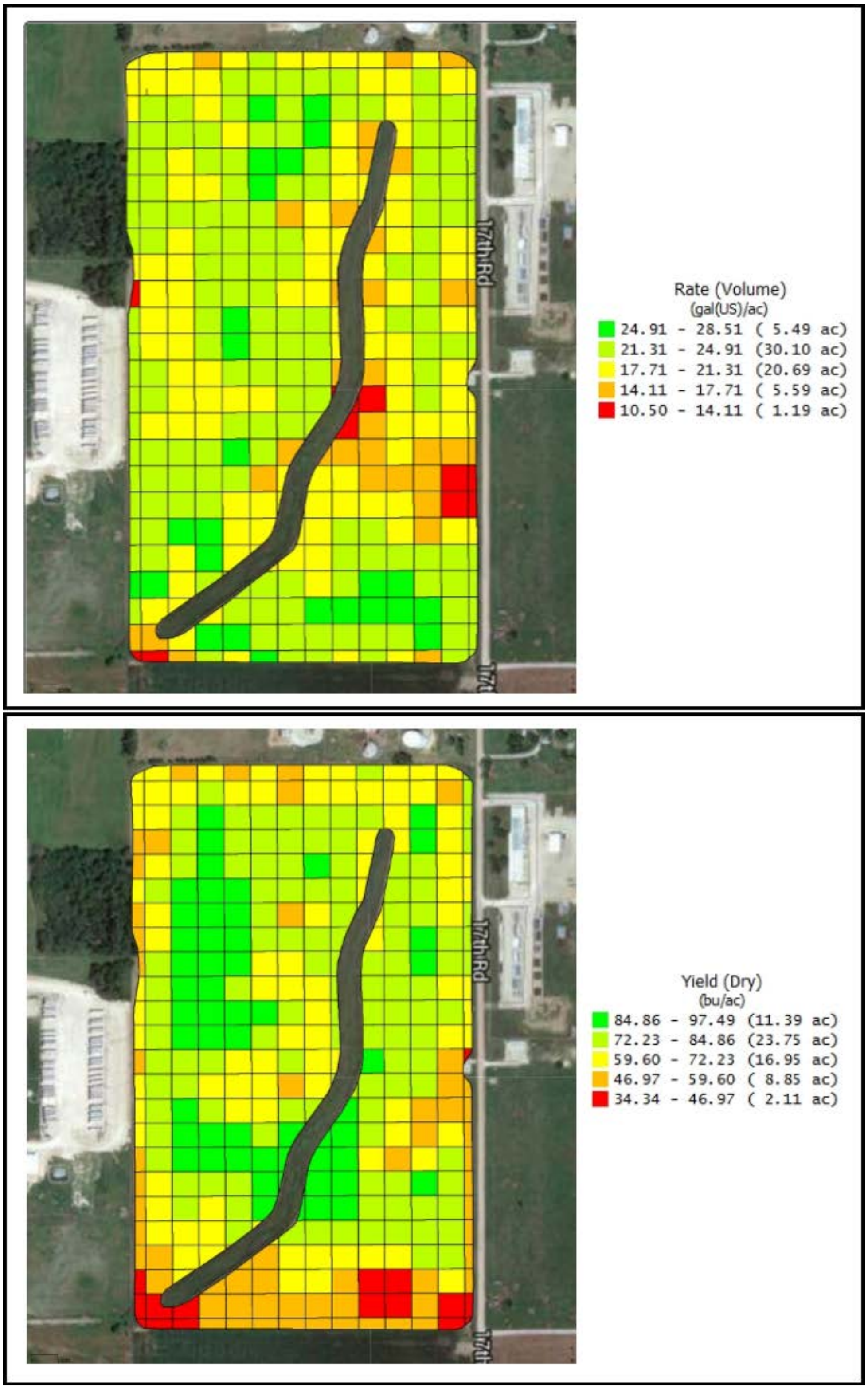

Figure 2 – Fertilizer application map and harvest yield map used in analysis

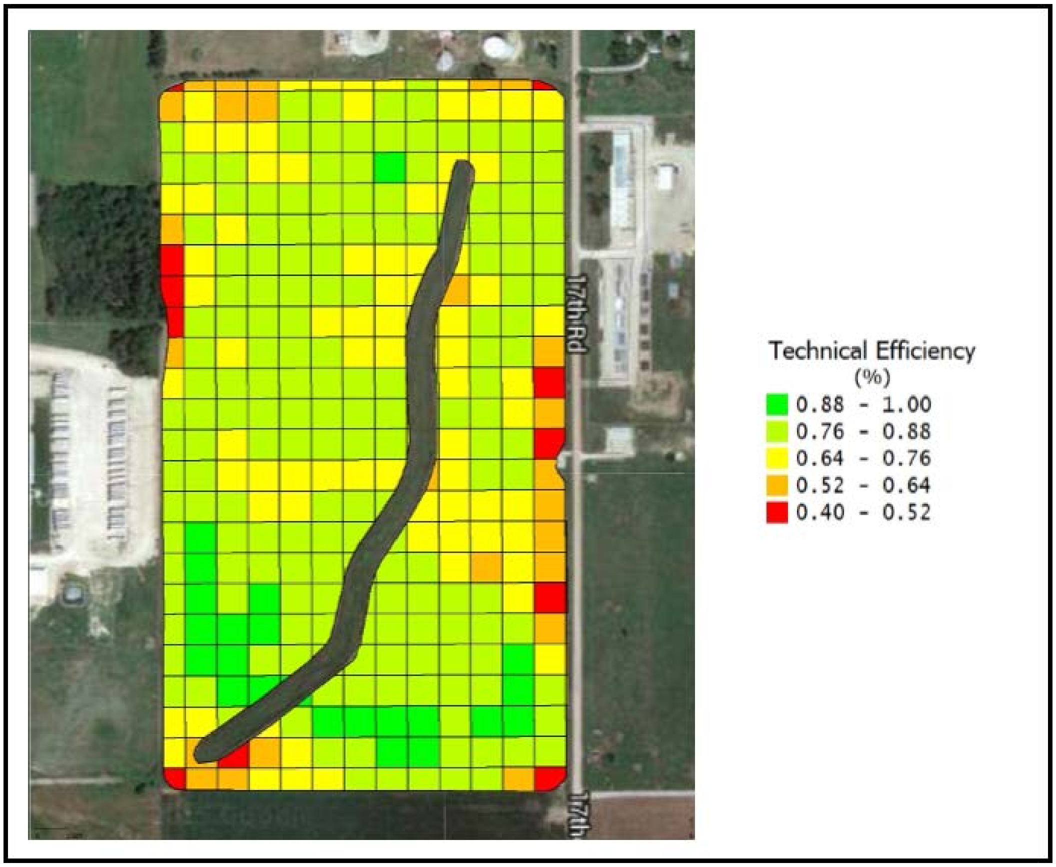

Figure 3 – 2015 wheat technical efficiency map presented in the same format as yield and application

Summary

The concept presented in this work has the potential to be successful because it eliminates as many external factors as possible in the calculation and only uses actual observations to form the model. With a format such as this, a grower would be able to complete efficiency studies using a variety of factors and parcel sizes and not have to rely upon peers to complete the research. Because the DMUs are segments within the same field rather than being parcels of ground, or differing farms altogether, farming practices, geographic factors such as climate, and other variables are not factors in the analysis because inherently they would be the same for all elements. This model also brings a much more realistic and defined result which can more effectively be acted upon than needing to change whole-farm management practices to become more efficient in comparison to one's peers. The result is tied spatially to a given locale, enabling for the grower to make true, intra-field changes in practice. There is, inevitably, variability in soil type, but those differences can be quantified using sensory devices and other meta-data in order to properly gauge and re-distribute its effect upon the model.

This project also reduces (and quite possibly eliminates) the impact of price in the measurement of efficiency when comparing DMUs. This is because output is measured by actual production data instead of accounting or market data, differentiating itself from other methods of determining efficiency.

There are a number of limitations to using this method. It relies heavily upon the quality and quantity of both production and input application data. If a grower does not have enough, or incomplete data, the model cannot function properly due to the impact that the fracture in the data has on the overall result, regardless of its inherent value. Also, one parcel of data points equally impacts the entire model, and missing a data point or an overlap can a hold an immense impact on the overall quality of the analysis.

On the other side of the argument, this method holds terrific potential for future research and offers the potential for numerous studies. Because the result of the analysis is not simply agronomic or simply yield-driven, the grower has a multitude of opportunities to make corrections or inferences based upon a variety of statistical methods. The convertibility of this platform allows for expansion to include infinite numbers of growers and farms within the study. In addition to applying these methods to sub-field areas of a single field; these sub-field or whole-field observations are evaluated across both a single farmer and multiple farmers in future studies.

References

Bogetoft, Peter, and Lars Otto. Benchmarking with DEA, SFA, and R. New York: Springer Science, 2011.

—. "Benchmarking: Benchmark and Frontier Analysis Using DEA and SFA." R package version 0.24. 2014.

Burkett, John. "Notes on choosing DEA models, writing R code, and interpreting R output." September 26, 2011.

Curtiss, Jarmila, and Ladislav Jelinek. "Cost Efficiency and Farm Self-selection in Precision Farming: The Case of Czech Wheat Production." EAAE Seminar 'Innovation for Agricultural Competitiveness and Sustainability of Rural Areas'. Prague, Czech Republic, 2012. 19.

Farrell, M J. "The Measurement of Productive Efficiency." Journal of the Royal Statistical Society Series A (General), vol. 120, no. 3 (1957): 253-290.

Featherstone, Allen M, and Md. Habibur Rahman. "Nonparametric Analysis of the Optimizing Behavior of Midwestern Cooperatives." Review of Agricultural Economics 18, no. 2 (May 1996): 265-273.

Featherstone, Allen M, and Terry L Kastens. "Non-Parametric and Semi-Parametric Techniques for Modeling and Simulating Correlated, Non-Normal Price and Yield Distributions: Applications to Risk Analysis in Kansas Agriculture." Journal of Agricultural and Applied Economics, August 2000: 267-281.

Featherstone, Allen M, Michael R Langemeier, and Mohammad Ismet. "A Nonparametric Analysis of Efficiency for a Sample of Kansas Beef Cow Farms." Journal of Agricultural and Applied Economics, July 1997: 175-184.

Fleming, Euan, Terence Farrell, Renato Villano, and Pauline Fleming. "Is Farm Benchmarking the New Acceptable Face of Comparative Analysis?" Australasian Agribusiness Review 14, no. 12 (2006): 18.

Henningsen, Arne. "censReg: Censored Regression (Tobit) Models." R package version 0.5-20. 2013.

Kutter, T, S Tiemann, R Siebert, and S Fountas. "The Role of Communication and Co-operation in the Adoption of Precision Farming." Precision Agric, 12 2011: 2-17.

Moghnieh, Ghassan A, Allen M Featherstone, and Barry K Goodwin. "A Nonparametric Analysis of Cost Minimization and Profit Maximization Behavior for a Sample of Kansas Farms." Staff Paper, Department of Agricultural Economics, Kansas State University, Manhattan, Kansas, 1991, 15.

Oh, Dong-hyun, and Dukrok Suh. "nonparaeff: Nonparametric Methods for Measuring Efficiency and Productivity." R package version 0.5-8. 2013.

Parman, Bryon James. Economies of Scale for Data Envelopment Analysis with a Kansas Farm Application. PhD Thesis, Manhattan, KS: Kansas State University, 2013.

Quintana-Ashwell, Nicolas E, and Allen M Featherstone. "Beyond Benchmarks: DEA Study of Kansas Farm Productivity." Southern Agricultural Economics Association's 2015 Annual Meeting. Atlanta, Georgia, 2015. 20.

R Core Team. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria. 2016. http://www.R-project.org/.

Schimmelpfennig, David, and Robert Ebel. "On the Doorstep of the Information Age: Recent Adoption of Precision Agriculture." EIB-80. U.S. Department of Agriculture, Economic Research Service, August 2011.

Simar, Leopold, and Paul W Wilson. "Estimation and Inference in Two-Stage, Semi-parametric Models of Production Processes." Journal of Economics 136 (2007): 31-64.

Simar, Leopold, and Paul W Wilson. "Two-Stage DEA: Caveat Emptor." Journal of Productivity Analysis, 2011.

Stacy, E W. "A Generalization of the Gamma Distribution." The Annal of Mathematical Statistics 33, no. 3 (September 1962): 1187-1192.

Trautmann, Heike, Detlef Steuer, Olaf Mersmann, and Bjorn Bornkamp. "truncnorm: Truncated Normal Distribution." R package version 1.0-7. 2014.

Watkins, K Bradley, Tatjana Hristovska, Ralph Mazzanti, and Charles E Wilson, Jr. "Measuring Technical, Allocative, and Economic Efficiency of Rice Production in Arkansas using Data Envelopment Analysis." Southern Agricultural Economics Association's 2013 Annual Meeting. Orlando, 2013.

Xue, Mei, and Patrick T Harker. "Ranking DMUs with Infeasible Super-Efficiency DEA Models." Management Science 48, no. 5 (May 2002): 705-710.