AN ABSTRACT OF THE DISSERTATION OF

Valentin Zelenyuk for the degree of Doctor of Philosophy in Economics presented on June 7, 2002.

Title: Essays in Efficiency and Productivity Analysis of Economic Systems

Redacted for Privacy

Abstract approved:_

ν Rolf Färe

Redacted for Privacy

Shawna P. Grosskopf

In this work I integrate some of my recent research developments in the theory and practice of *Productivity and Efficiency Analysis of Economic Systems*. In particular, I present some new theoretical relationships between various measures of efficiency and productivity, propose new solutions to some aggregation problems in efficiency analysis and apply the existing theory and the new findings to empirical analysis in Industrial Organization.

ESSAYS IN
EFFICIENCY AND PRODUCTIVITY ANALYSIS
OF ECONOMIC SYSTEMS

by
Valentin Zelenyuk

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 7, 2002
Commencement June 2003

Doctor of Philosophy dissertation of <u>Valentin Zelenyuk</u> presented on <u>June 7, 2002.</u>

APPROVED:

Redacted for Privacy

Co-Major Professor, representing Economics

Redacted for Privacy

Co-Major Professor, representing Economics

Redacted for Privacy

Head of the Department of Economics

Redacted for Privacy

Dean of Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Valentin Zelenyuk, Author

ACKNOWLEDGEMENT

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

ESSAYS IN EFFICIENCY AND PRODUCTIVITY ANALYSIS
OF ECONOMIC SYSTEMS

GENERAL INTRODUCTION

In this dissertation I am synthesizing some of ideas and results developed during my last three years of research. I will do it in three main chapters, or three types of essays, which in total integrate nine essays, dedicated to separate but related research questions in the theory of Productivity and Efficiency Analysis (PEA) and its applications to Industrial Organization. Let me concisely introduce you to the subject, each of the chapters and each of the essays.

PEA is a modern and fast growing area in Measurement Economics. Extensive research in this area has been around for about half of a century. Numerous measures of efficiency and productivity have been offered by researchers since then: The Farrell measure, the Russell measure, the additive, the hyperbolic measures, and the measures based on the directional distance function are among the most popular examples. It sounds logical to ask: How are all these measure related? Are they ever equal? Do they differ significantly? What if two researchers use the same data set to answer the same research question but choose different efficiency measures—will their results be equivalent? Consistent? In general or under some conditions?

Not surprisingly, these questions have been explored before (e.g., Färe, Grosskopf and Lovell (1994), Färe and Grosskopf (2000)), but some open questions remained. One of such questions sets up the first part of the dissertation; it is devoted to establishing *new relationships* between some existing measures of efficiency and productivity.

This first chapter opens up with **Essay 1** that contains a paper co-authored with Rolf Färe and Shawna Grosskopf. There, we establish precise relationships

between the Farrell and the Russell technical efficiency measures[1] as well as between the directional distance function and the additive measure of technical efficiency. This work gives a theoretical benchmark for comparison of various measures, by discovering the necessary and sufficient conditions on the technology that ensure the equivalence of the mentioned efficiency measures.

In **Essay 2**, I find a new relationship between the directional distance function and Shephard's distance functions (reciprocals of the Farrell technical efficiency measures). In particular, I find that constant returns to scale (CRS) technology is a necessary and sufficient condition for this relationship. Applying this discovery to measurement of *productivity* growth in economic systems (where CRS is a common assumption), I introduce a new Total Factor Productivity (TFP) index and show that it is a generalization of other existing TFP indexes, such as Malmquist, Fisher and Törnqvist Productivity Indexes.

Essay 3 sheds some light into another area of PEA, *scale efficiency* measurement. There, I follow the results of Färe and Grosskopf (1985) on duality between the input-scale efficiency and cost-scale efficiency measures to derive the necessary and sufficient condition on technology that ensures their finding—the equivalence of the two types of measures. Such technology turns out to exhibit a special case of the input *homotheticity*—a property that I dub as the Input Scale Homotheticity.

The second chapter of essays concentrates on a different subject in PEA— *aggregation issues*. Except for some studies (Li and Ng (1995), Russell and Blackorby (1999), and Ylvinger (2000)), these issues were rarely explored in the literature from a theoretical perspective. The questions there, however, are of fundamental importance. For example, conclusions from comparison of efficiencies of various groups (i.e., aggregation over firms) may crucially depend on the *form of aggregation* in use. Also, the use of aggregate data (i.e., aggregation over inputs or

---

[1] The Russell measure was introduced in Färe and Lovell (1978). See also Russell (1985, 1990).

outputs) may introduce a bias into the estimation of efficiency scores. Both of these matters are discussed in this chapter.

The chapter starts with **Essay 4**, a paper co-authored with Rolf Färe and Shawna Grosskopf, where we determine a relationship between the aggregate and individual efficiencies. Using the Koopmans (1957) theorem about the relationship between the aggregate and individual profit functions, we determine the relationship between the aggregate and individual efficiencies based on the directional distance functions.

From a similar perspective, the aggregation problem in PEA is addressed in **Essay 5**. Here, co-authored with Rolf Färe, we find a way to determine the relationship between the aggregate and individual *Farrell-type* efficiencies. Here, we use a particular type of technology aggregation: when the aggregate production possibility set is equal to the sum of individual production possibility sets. Such aggregation enables us to show that the revenue function defined on the aggregate technology is the sum of revenue functions defined on the individual technologies. This is an analog to the Koopmans (1957) theorem used in the previous essay, which becomes a keystone for determining the relationship between the aggregate and individual revenue, technical and allocative efficiencies.

In **Essay 6**, the paper coauthored with Rolf Färe, we use a similar technique as in **Essay 4** to show that the appropriate way of averaging Farrell-type efficiency scores (consistent with their mathematical nature) is to use the *weighted geometric* mean. Combining this with results from **Essay 5**, we are able to derive a system of weights that can be price independent and that have economic theory background and an intuitive interpretation.

So far, the focus was on the aggregation *over* decision-making *units* (firms, countries, etc). Another course of aggregation is *over goods* (inputs or outputs)— this is the subject of **Essay 7**. There, co-authored with Rolf Färe, we address the issue of bias in efficiency measurement due to input aggregation raised by Tauer (2001). Specifically, we find that the sub-vector Farrell-type efficiency measure

will yield unbiased efficiency scores *if and only if* there is no allocative inefficiency in the subvector of inputs being aggregated.

All the above-mentioned papers were theoretical in nature, but motivated by practical issues. In the third chapter of the dissertation, I make use of one of the theoretical results to address an empirical question that challenges many industrial organization economists: "What were, in practice, the primal causes of concentration in a given industry?" In particular, I adopt and slightly modify the techniques developed above (in the **Essay 5**) to measure the existence and the size of economies of scale in the U.S. brewing industry. This industry experienced a rapid and consistent rise in concentration for several decades, attracting attention of government and antitrust officials who, in turn, were seeking an explanation from economists.

Economists engaged in this issue have divided into two camps in their explanations: (*i*) those supporting the view that concentration came out primarily as a result of the existing economies of scale[2], and (*ii*) those defending the demand side cause as the primal reason for growing concentration[3]. Employing the PEA methodology and data used by Tremblay and Tremblay (1985), I find evidence that contradicts the argument of the former camp. That is, I find that although some firms were experiencing some economies of scale, they were not the firms that were causing the rise in concentration. In the conclusion of this empirical study, I attempt to connect the pieces of evidence from these as well as other researchers' results to construct a broader picture of the evolution of this industry. The resulting picture is consistent with various explanations offered by economists defending the *demand side* causes of concentration.

---

[2] See for example, Horowitz and Horowitz (1965, 1967), Greer (1971, 1981) and Tremblay and Tremblay (1985, 1987).

[3] See for example Elzinga (1973, 1977), Scherer et al. (1975) Keiththahn (1978) and Lynk (1984).

THEORETICAL ESSAYS ON RELATIONSHIP OF VARIOUS PRODUCTIVITY AND EFFICIENCY MEASURES

INTRODUCTION

An Overview

For a long time, economists had an interest in *how to evaluate the performance of an economic system* (e.g., country, region, industry, firm, public sector). Only recently, has such an interest culminated in a new and fast growing area of economic thought—*efficiency and productivity analysis (EPA)*. Despite the youth of its theoretical foundation, one can find numerous applications in almost every branch of economics: macro-, micro- and environmental economics, industrial organization, international trade, transportation economics and public policy, to mention a just few.

On the theoretical side, researchers developed various measures for evaluating efficiency and productivity, each fitting particular empirical issues and data restrictions. An important question arises: How are the various measures related to each other? This, first part of the dissertation aims to contribute to answering this question.

Before introducing each paper in this section, let me give a brief historical remark on the evolution of thought in the area of efficiency and productivity analysis.

EPA: A Historical Retrospect

While the need for efficiency and productivity analysis existed throughout the history of economic thought, most developments were made relatively recently. Origins of the fundamental ideas on the concept of efficiency and/or its

measurement are found in seminal works of Konüs and Byushgens (1926), and Koopmans (1951), Debreu (1952), Shephard (1953), Malmquist (1953), Farrell (1957) to mention just a few.

One of the earliest formal concepts of *production efficiency* is found in Koopmans (1951) who defines it as such production possibilities for which it is not feasible to increase any of the outputs without *simultaneously* increasing any of the inputs.

Perhaps the most substantial contributions to the origin of the subject are due to Farrell (1957). In particular, he suggested and justified a way to measure the *technical*, price (or *allocative*) and *overall* efficiencies of a unit in economic system (industry, country, etc.) relative to other similar units of the same system. The technical efficiency measure he suggested was later named after him, and is also often referred to as the Debreu-Farrell measure, for its conceptual relationship with the seminal Debreu (1952) work.

Nevertheless, only since 1978 has this area received broad attention. Two studies—by Färe and Lovell (1978) and by Charnes, Cooper and Rhodes (1978)—seem to be responsible for attracting such attention. Both studies emphasized the relationship of the Farrell (1957) ideas to Shephard's (1953, 1970) approach to production theory based on duality—the connection that laid out a solid economic theory foundation to the theory of EPA. Also, due to Charnes et al. (1978), the computational approach to efficiency and productivity analysis via mathematical programming obtained a new name—Data Envelopment Analysis or DEA.

An interesting evolution of thought appeared in the discussion around the measure of *technical* efficiency suggested by Farrell. This discussion starts with the aforementioned study of Färe and Lovell (1978) who challenged the Farrell (1957) approach and suggested an alternative—the Russell measure of technical efficiency. The idea was to incorporate the Koopmans (1951) notion of technical efficiency, which the Farrell measure failed to satisfy under some specific technologies (e.g., Leontief technology).

Both the Russell measure and the critique itself triggered a series of studies to rescue the Farrell approach and to criticize the new measure of efficiency (Kopp (1981), Russell (1983, 1985, 1990), Zieschang (1984), Bol (1986)). Many interesting properties of both efficiency measures were discovered due to these studies (see Russell (1990)).

One intriguing question—How are the Russell and the Farrell measures related?—was left open, however. (Exactly this question becomes the core of the first paper of this section.)

A similar intention to incorporate the Koopmans (1951) notion of technical efficiency is found in Charnes et al., (1985) when they introduced the *additive* measure of technical efficiency. This had opened another question—How this additive measure related to other measures of efficiency?—this question that is studied in the first paper of this section.

While the Farrell approach was originally input oriented (i.e., decrease inputs, for the same level of output) the methodology has been extended to the output orientation, sub-vector orientation as well to a case of simultaneous equiproportional input reduction and output expansion (e.g., see Färe et al. (1994) for details).

Further, a more general perspective to measurement of technical and other efficiencies came to EPA with the directional distance function (Chambers et al. (1996, 1998)), which was known earlier as the benefit or shortage functions in consumer welfare measurement (Luenberger, 1992). In particular, the directional distance function allowed measuring efficiency with any orientation, which could be specified with a certain directional vector. Again, an important question is how such measure of efficiency is related to other measures. This question has already been explored extensively, with a general conclusion that the directional distance function is a generalization of the Shephard distance functions (e.g., Chambers et al. (1996) Färe and Grosskopf (2000)). In the first and second papers of this chapter, a few new relationships are discovered.

Besides the technical, allocative (price) and the overall efficiency measures suggested by Farrell (1957), the theory has also been enriched with measures evaluating *productivity* changes as well as the *scale* efficiency.

The measurement of productivity change has its own history. Early studies go back to at least Fisher (1922). For a long time, two approaches have been dominating there: (*i*) measurement based on the index numbers, and (*ii*) measurement based on the parametric econometric estimation.

The two fields of EPA—productivity analysis *and* efficiency analysis—have merged perhaps with the seminal paper by Caves, Christensen and Diewert (1982) who used the Shephard's distance function to define a productivity index, naming it the Malmquist Productivity Index (MPI).

The earliest empirical implementation of the MPI is found in Färe et al. (1989). Since then, MPI has received a huge attention from the empirical researchers. On the other hand, theorists have been offering new productivity measures, many of which were in the spirit of the MPI. Chung et al., (1997) for example introduced the Malmquist Luenberger Productivity (input and output oriented) indexes (MLPI) and Chambers et al., (1996) introduced its additive analogue—the Luenberger Productivity index (LPI).

Again, a captivating question is: How are the various measures of productivity change related? This question was extensively explored in Caves et al. (1982), Diewert (1992) and Färe et al., (1997), to mention a few. The key result was that the MPI is a generalization of such popular total factor productivity indexes as Fischer, Törnquist and Hicks-Moorsteen productivity indexes.

A new relationship between productivity measures is discovered in the second paper of this section. In particular, I show that the combination of the output and input oriented MLPIs yields a measure that is a generalization of the MPI.

The history of the *scale efficiency* measurement goes back to at least Førsund and Hjalmarsson (1974, 1979) who defined the scale efficiency measure as the ratio of the Farrell technical efficiency measures estimated with respect to

constant and variable returns to scale assumptions. An alternative to this, perhaps the most popular measure of scale efficiency, was introduced by Banker et al. (1984) and recently elaborated by Sueyoshi (1999).

A dual approach to measuring the Førsund et al. type of *scale efficiency* was introduced by Färe and Grosskops (1985). They also showed the necessary and sufficient condition for equivalence of the dual and primal measures. This condition opened another intriguing question—What type of technology is capable of satisfying it?—this question is explored in the third essay of this section, by referring to the concept of homotheticity.

Finally, the following remark on the historical perspective of the EPA might be appropriate at this stage: The two seminal studies—Färe and Lovell (1978) and Charnes et al. (1978)—virtually started two competitive schools of thought in efficiency and productivity analysis. The former school tends to develop and emphasize the economic theory foundation and economic interpretation of theoretical and empirical discoveries in EPA—and thus can be called as the school of economists in EPA. The latter school tends to concentrate on the computational issues and engineering interpretation of ideas and findings—and thus can be called as the operational researchers' school in EPA. The present work follows a tradition of the economists approach.

The rest of this introduction gives a foreword to each paper of the section.

Essay 1: Finding Common Ground: Efficiency Indices

Perhaps one of the most interesting questions about relationships between various efficiency measures that remained open till now was: What is the relationship between the Farrell and the Russell measures? Or, more specifically, what type of technologies, if any, can ensure equivalence of these measures?

This question is answered in this first essay of the section—a paper coauthored with Rolf Färe and Shawna Grosskopf. Formally, we find that these two

(multiplicative) measures are equivalent *if and only if* technology is input homothetic and of the Cob-Douglas form with symmetric weights. Then, a 'parallel' result is found on the side of the *additive* efficiency measures. Specifically, we find that the directional distance function and the additive measure of technical efficiency (Charnes et al., 1985) would yield equivalent efficiency scores *if and only if* technology is translation input homothetic and linear with symmetric weights.

Another question that has not been explored was about the economic interpretation of the Russell and the additive measures of technical efficiency—which we do by showing the cost interpretation via the duality theory in economics.

In general, this essay gives a theoretical benchmark for comparison of various measures of technical efficiency, by discovering the conditions that ensure their equivalence.

Essay 2: Directional and Shephard's Distance Functions: New Link and its Implication to Productivity Measurement

In this paper, I find a new relationship between the directional distance function and the Shephard's distance functions (reciprocals of the Farrell technical efficiency measures). This discovery then helps me to find relationship of the Malmquist-Luenberger Productivity Indexes (MLPI) to the Malmquist Productivity Index (MPI), and thus to other Total Factor Productivity Indexes. In particular, I find that under constant returns to scale and for a wide range of directions, the ratio of the output oriented to input oriented MLPIs yields the output oriented Malmquist Productivity Index, thus showing that the latter is a special case of the former.

In general, this essay gives a new link and a new productivity measure that is more general than the MPI—the index that has been the most general so far.

Essay 3: Scale Efficiency: Equivalence of Primal and Dual Measures

In this essay, I investigate conditions for equivalence of the *primal* and *dual* measures of *scale* efficiency. Specifically, I follow the study of Färe and Grosskopf (1985) on duality between the input-scale efficiency and cost-scale efficiency measures. I find that the necessary and sufficient condition on technology that ensures the *equivalence* of the dual and primal scale efficiency measures is that technology must exhibit a special case of the input *homotheticity*—a property of technology that I dub as the *input scale homotheticity*.

Overall, this essay offers a precise formal interpretation of technological properties that would ensure the equivalence between the scale efficiency measure based on primal information (inputs and outputs) and the scale efficiency measure based on the dual information (cost and outputs or revenue and inputs).

Altogether, this section makes a contribution to understanding more about various measures of efficiency and productivity, and the author hope it also will be fun to read.

ESSAY 1: FINDING COMMON GROUND: EFFICIENCY INDICES

Valentin Zelenyuk[1]

Introduction

The last two decades have witnessed a revival in interest in the measurement of productive efficiency pioneered by Farrell (1957) and Debreu (1957). 1978 was a watershed year in this revival with the christening of DEA by Charnes, Cooper and Rhodes (1978) and the critique of Farrell technical efficiency in terms of axiomatic production and index number theory in Fare and Lovell (1978). These papers have inspired many others to apply these methods and to add to the debate on how best to define technical efficiency.

In this paper we try to pull together some of the variants that have arisen over these decades and show when they are equivalent. The specific cases we take up include: 1) the original Debreu-Farrell measure versus the Russell measure—the latter introduced by Färe and Lovell, and 2) the directional distance function and the additive measure. The former was introduced by Luenberger (1992) and the latter by Charnes, Cooper, Golany and Seiford (1985). We also provide a discussion of the associated cost interpretations.

Basic Production Theory Details

In this section we introduce the basic production theory that we employ in this paper. We will be focusing on the input based efficiency measures here, but the analysis could readily be extended to the output oriented case as well.

To begin, technology may be represented by its input requirement sets

$$L(y) = \{x: \ x \ can \ produce \ y\}, \ y \in \Re^{M}_{+}, \tag{1}$$

where $y \in \Re_+^M = \left\{ y \in \Re^M : y_m \geq 0, m = 1,...,M \right\}$ denotes outputs and $x \in \Re_+^N$ denotes inputs. We assume that the input requirement sets satisfy the standard axioms, including: $L(0) = \Re_+^N$, and $L(y)$ is a closed convex set with both inputs[2] and outputs[3] freely disposable (for details see Färe and Primont (1995)).

The subsets of $L(y)$ relative toward which we measure efficiency are the isoquants

$$IsoqL(y) = \left\{ x : x \in L(y), \lambda x \notin L(y), \lambda > 1 \right\}, y \in \Re_+^M, \qquad (2)$$

and the efficient subsets

$$EffL(y) = \left\{ x : x \in L(y), x' \leq x, x' \neq x \Rightarrow x' \notin L(y) \right\}, y \in \Re_+^M. \qquad (3)$$

Clearly, $EffL(y) \subseteq IsoqL(y)$ and as one can easily see with a Leontief technology, i.e., $L(y) = \left\{ (x_1, x_2) : \min\{x_1, x_2\} \geq y \right\}$, the efficient subset may be a proper subset of the isoquant.

Next we introduce two function representations of $L(y)$, namely the Shephard input distance function and the directional input distance function, and discuss some of their properties.

Shephard's (1953) input distance function is defined in terms of the input requirement sets $L(y)$ as

$$D_i(y, x) = \sup\left\{ \lambda : x / \lambda \in L(y) \right\}. \qquad (4)$$

---

[2] Inputs are freely disposable if $x' \geq x \in L(y) \Rightarrow x' \in L(y)$.

[3] Outputs are freely disposable if $y' \geq y \Rightarrow L(y') \subseteq L(y)$.

Among its important properties[4] we note the following

i) $D_i(y,x) \geq 1$ *if and only if* $x \in L(y)$, Representation

ii) $D_i(y,\lambda x) = \lambda D_i(y,x), \lambda > 0$, Homogeneity

iii) $D_i(y,x) = 1$ *if and only if* $x \in IsoqL(y)$, Indication

Our first property shows that the distance function is a complete representation of the technology. Property ii) shows that the distance function is homogeneous of degree one in inputs, i.e., the variables which are scaled in (4). The indication condition shows that the distance function identifies the isoquants.

Turning to the directional input distance function introduced by Luenberger (1992)[5], we define it as

$$\vec{D}_i(y,x;g_x) = \sup\{\beta : (x - \beta g_x) \in L(y)\}, \tag{5}$$

where $g_x \in \Re_+^N$ is the directional vector in which inefficiency is measured. Here we choose $g_x = 1^N \in \Re_+^N$. This function $\vec{D}_i(y,x;1^N)$ has properties that parallel those of $D_i(y, x)$, and are listed below. For technical reasons the indication property is split into two parts. We note that we require inputs to be strictly positive in part a) of the indication property. The proofs of these properties are found in the appendix.

i) $\vec{D}_i(y,x;1^N) \geq 0$ *if and only if* $x \in L(y)$, Representation

ii) $\vec{D}_i(y,x + \alpha 1^N;1^N) = \vec{D}_i(y,x;1^N) + \alpha, \alpha > 0$, Translation

---

[4]For additional properties and proofs, see Färe and Primont (1995).

[5]In consumer theory he calls this the benefit function and in producer theory he uses the term shortage function.

iiia) *if* $\vec{D}_i(y,x;1^N) = 0$ *and* $x_n > 0, n = 1,...N,$ *then* $x \in IsoqL(y)$, Indication

iiib) $x \in IsoqL(y)$ *implies* $\vec{D}_i(y,x;1^N) = 0$, Indication

Since we will be relating technical efficiency to costs, we also need to define the cost function, which for input prices $w \in \Re_+^N$ is

$$C(y,w) = \min\{wx : x \in L(y) \}.$$ (6)

The following dual relationships apply

$$\frac{C(y,x)}{wx} \leq 1 / D_i(y,x)$$ (7)

and

$$\frac{C(y,x) - wx}{w1^N} \leq -\vec{D}_i(y,x;1^N).$$ (8)

Expression (7) which is the Mahler inequality, states that the ratio of minimum cost to observed cost is less than or equal to the reciprocal of the input distance function. Expression (8) states that the difference between minimum and observed cost, normalized by input prices, is no larger than the negative of the directional input distance function.

These two inequalities may be transformed to strict equalities by introducing allocative inefficiency as a residual.

The Debreu-Farrell and Russell Equivalence

Our goal in this section is to find conditions on the technology $L(y), y \in \Re_+^M$, such that the Debreu-Farrell (Debreu (1957), Farrell (1957)) measure of technical efficiency coincides with the Russell (Färe and Lovell (1978))

measure. To establish these conditions we redefine the original Russell measure and introduce a multiplicative version. We do this by using the geometric mean as the objective function in its definition rather than an arithmetic mean. Thus our multiplicative Russell measure is defined as

$$R_M(y,x) = \min\left\{ (\prod_{n=1}^{N} \lambda_n)^{1/N} : (\lambda_1 x_1,...,\lambda_N x_N) \in L(y), 0 < \lambda_n \leq 1, n = 1,...,N \right\} \qquad (9)$$

The objective function here is $(\prod_{n=1}^{N} \lambda_n)^{1/N}$ in contrast to $\sum_{n=1}^{N} \lambda_n / N$ from the original specification in Färe and Lovell (1978). For technical reasons we assume here that inputs $x = (x_1, . . ., x_n)$ are strictly positive, i.e., $x_n > 0$, $n = 1,...,N$. More specifically in this section we assume that for $y \geq 0, y \neq 0, L(y)$ is a subset of the interior of $\Re_+^N$.[6]

Note that the Russell measure in (9) has the indication property

$$R_M(y,x) = 1 \text{ if and only if } x \in EffL(y) \qquad (10)$$

Recall that the Debreu-Farrell measure of technical efficiency is the reciprocal of Shephard's input distance function, i.e.,

$$DF(y,x) = 1 / D_i(y,x) \qquad (11)$$

thus it is homogeneous of degree -1 in x and it has the same indication property as $D_i(y, x)$.

---

[6] See Russell (1990) for a related assumption

Now assume that the technology is input homothetic[7], i.e.,

$$D_i(y,x) = D_i(1,x)/H(y) \tag{12}$$

and that the input aggregation function $D_i(1, x)$ is a geometric mean, so that the distance function equals

$$D_i(y,x) = (\prod_{n=1}^{N} x_n)^{1/N}/H(y). \tag{13}$$

From (4) and the Representation property it is clear that the distance function takes the form above if and only if the input requirement sets are of the following form

$$L(y) = H(y) \cdot \left\{ \hat{x} : (\prod_{n=1}^{N} \hat{x})^{1/N} \geq 1 \right\}, \hat{x} = \frac{x}{H(y)} . \tag{14}$$

The Russell characterization theorem can now be stated; the proof may be found in the appendix.

Theorem 1: Assume that $L(y)$ is interior to $\Re_+^M$ for $y \geq 0, y \neq 0$.

$R_M(y,x) = DF(y,x)$ for all $x \in L(y)$ if and only if $D_i(y,x) = (\prod_{n=1}^{N} x_n)^{1/N}/H(y)$.

Thus for these two efficiency measures to be equivalent, technology must satisfy a fairly specific form of homotheticity - technology is of a restricted Cobb-Douglas form in which the inputs have equal weights. This makes intuitive sense,

---

[7] For details see Färe and Primont (1995).

since technology must be symmetric, but clearly not of the Leontief type. That is, technology must be such that the *IsoqL(y)* =*EffL(y)*. Of course, it is exactly the Leontief type technology which motivated Färe and Lovell to introduce a measure that would use the efficient subset *EffL(y)* rather than the isoquant *IsoqL(y)* as the reference for establishing technical efficiency.

The Directional Distance Function and the Additive Measure

We now turn to some of the more recently derived versions of technical efficiency; specifically we derive conditions on the technology $L(y), y \in \Re_+^M$ that are necessary and sufficient for the directional distance function to coincide with a "stylized' additive measure of technical efficiency.

The original additive measure introduced by Charnes, Cooper, Golany and Seiford (1985)(hereafter CCGS) simultaneously expanded outputs and contracted inputs. Here we focus on a version that contracts inputs only, but in the additive form of the original measure. Although the original measure was defined relative to a variable returns to scale technology, (see p. 97, CCGS), here we leave the returns to scale issue open and impose only those conditions itemized in Section 2. Finally, we normalize their measure by the number of inputs, *N*.

We are now ready to define the stylized additive model as

$$A(y,x) = \max\left\{ \sum_{n=1}^{N} s_n / N : (x_1 - s_1, \dots, x_N - s_N) \in L(y)\right\}, \tag{15}$$

where $s_n \geq 0, n = 1, \dots, N$.

This measure reduces each input x_n so that the total reduction $\sum_{n=1}^{N} s_n / N$ is maximized. Intuitively, one can think of this problem as roughly equivalent to minimizing costs when all input prices are equal to one. We will discuss this link in the next section.

The additive measure and the modified Russell measure look quite similar, although the former uses an arithmetic mean as the objective and the modified Russell measure uses a geometric mean. The additive structure of $A(y, x)$ suggests that the directional distance function - which also has an additive structure - may be related to it.[8] To make that link we begin by characterizing the technology for which these two measures would be equivalent. We begin by assuming that technology is translation input homothetic,[9] i.e., in terms of the directional distance function we may write

$$\vec{D}_i(y, x; 1^N) = \vec{D}_i(0, x; 1^N) - F(y). \tag{16}$$

Moreover, we assume that the aggregator function $\vec{D}_i(0, x; 1^N)$ is arithmetic mean so that the directional distance function may be written as

$$\vec{D}_i(y, x; 1^N) = \frac{1}{N} \sum_{n=1}^{N} x_n - F(y). \tag{17}$$

Note that from the properties of the directional distance function, it follows that it takes the form required above if and only if the underlying input requirement sets are of the form

---

[8] Larry Seiford noted the similarity at a North American Efficiency and Productivity Workshop.

[9] For details see Chambers and Färe (1998). Chambers and Färe assumed that $F(y)$ depends on the directional vector 1^N. Here we take it as fixed and omit it.

$$L(y) = \left\{ \tilde{x} : \frac{1}{N} \sum_{n=1}^{N} \tilde{x}_n \geq 0 \right\} + F(y), \tag{18}$$

where $\tilde{x} = (x_1 - F(y), \ldots, x_N - F(y))$.

We are now ready to state our additive representation theorem (see appendix for proof),

Theorem 2:

$$\vec{D}_i(y, x; 1^N) = A(y, x) \text{ for all } x \in C(L(y)) = \left\{ \hat{x} : \hat{x} = x + \delta 1^N, x \in L(y), \delta \geq 0 \right\}$$

$$\text{if and only if } \quad \vec{D}_i(y, x; 1^N) = \frac{1}{N} \sum_{n=1}^{N} x_n - F(y).$$

Here we see that to obtain equivalence between the additive measure and the directional distance function, technology must be linear in inputs, i.e., the isoquants are straight lines with slope $= -1$.

Cost Interpretations

The Debreu-Farrell measure has a dual interpretation, namely the cost deflated cost function. Here we show that the multiplicative Russell measure and the additive measure also have dual cost interpretations.[10]

---

[10] It is straightforward to show that the original (additive) Russell measure also has a cost interpretation, despite the claim by Kopp (1981, p. 450) that the Russell measure '...cannot be given a meaningful cost interpretation which is factor price invariant.' In this section, we provide such a cost interpretation.

Recall that we define the cost function

$$C(y, w) = \min\{wx : x \in L(y)\}, \tag{19}$$

where $w \in \Re_+^N$ are input prices. From the definition it follows that

$$C(y, w) \leq wx, \forall x \in L(y). \tag{20}$$

Now since $DF(y, x) \, x \in L(y)$ it is also true that

$$C(y, w) \leq w(DF(y, x)x) = wx(DF(y, x)) \tag{21}$$

and

$$C(y, w) / wx \leq DF(y, x) \tag{22}$$

Expression (22) is the Mahler inequality expressed in terms of the cost efficiency measure *(C(y, w)/wx)* and the Debreu-Farrell measure of technical efficiency, *DF(y, x)*. This inequality may be closed by introducing a multiplicative measure of allocative efficiency, *AE(y, x, w)*, so that we have

$$C(y, w)/wx = DF(y, x)AE(y, x, w). \tag{23}$$

To introduce a cost interpretation of the multiplicative Russell measure we note that

$$(\lambda^*_1 x_1, ..., \lambda^*_N x_N) \in L(y), \tag{24}$$

where λ^*_n $(n = 1, \ldots, N)$ are the optimizers in expression (9). From the assumption that the input requirement sets are subsets of the interior of $\Re^N_+$, it follows that λ^*_n >0, $n = 1, \ldots, N$. By (20) and (24) we have

$$C(y, w) \le (\lambda^*_1 w_1 x_1, \ldots, \lambda^*_N w_N x_N) \tag{25}$$

and by multiplication

$$C(y,w)/wx \le \left(\prod_{n=1}^{N} \lambda^*_n\right)^{1/N} \left[\frac{\lambda^*_1 w_1 x_1}{\left(\prod_{n=1}^{N}\lambda^*_n\right)^{1/N} wx} + \cdots + \frac{\lambda^*_N w_N x_N}{\left(\prod_{n=1}^{N}\lambda^*_n\right)^{1/N} wx} \right] \tag{26}$$

or

$$C(y,w)/wx \le R_M(y,x) \left[\frac{\lambda^*_1 w_1 x_1}{\left(\prod_{n=1}^{N}\lambda^*_n\right)^{1/N} wx} + \cdots + \frac{\lambda^*_N w_N x_N}{\left(\prod_{n=1}^{N}\lambda^*_n\right)^{1/N} wx} \right] \tag{27}$$

Expression (27) differs from the Mahler inequality (22) in that it contains a second term on the right hand side. This term may be called the Debreu-Farrell deviation, in that if $\lambda_1 = \ldots = \lambda_N$, the deviation equals one. That is, if the scaling factors λ^*_n are equal for each n, then (27) coincides with (22). Again, the inequality (27) can be closed by introducing a multiplicative residual, which captures allocative inefficiency.

Turning to the additive measure, we note that

$$(x_1 - s_1^*, \ldots, x_N - s_N^*) \in L(y) \tag{28}$$

where $s_n^*, n = 1, \ldots, N$ are the optimizers in problem (15). Thus from cost minimization we have

$$C(y, w) \leq wx - ws^*, \tag{29}$$

where $s^* = (s_1^*, \ldots, s_N^*)$. From (29) we can derive two dual interpretations: a ratio and a difference version.

The ratio interpretation is

$$C(y, w) / wx \leq 1 - \frac{ws^*}{wx}, \tag{30}$$

which bears some similarity to the Farrell cost efficiency model in (22). Now if $w = (1, \ldots, 1)$, then it follows that the additive model is related to costs as

$$\frac{C(y, 1^N)}{\sum\limits_{n=1}^{N} x_n} \leq 1 - \frac{\sum\limits_{n=1}^{N} s_n^* / N}{\sum\limits_{n=1}^{N} x_n / N} = 1 - \frac{A(y, x)}{\sum\limits_{n=1}^{N} x_n / N} \tag{31}$$

In this case we see that Debreu-Farrell cost efficiency (the left-hand side) is not larger than one minus a normalized additive measure.

The second cost interpretation is

$$C(y,w) - wx \leq -ws^{*},$$

(32)

and when $w = (1, \ldots, 1)$ we obtain

$$\frac{C(y,1^{N}) - \sum\limits_{n=1}^{N} x_{n}}{N} \leq -A(y,x)$$

(33)

If we compare this result to (8), we see again, the close relationship between the additive measure and the directional distance function.

References

Chambers, R.G., and R. Färe (1998),"Translation Homotheticity," *Economic Theory* 11, 629-64 .

Charnes, A., W.W. Cooper, B. Golany, L. Seiford and J. Stutz (1985), "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions," *Journal of Econometrics* 30:12, 9 - 07.

Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision-making Units," *European Journal of Operational Research* 2:6, 429-444.

Debreu, G. (1951), "The Coefficient of Resource Utilization," *Econometrica* 19, 273-292.

Färe, R. and C.A.K. Lovell (1978), "Measuring Technical Efficiency of Production," *Journal of Economic Theory* 19, 1 50- 62.

Färe, R. and D. Primont (1995), *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers: Boston .

Farrell, M. (1957), "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society,* Series A, General, 1 20, Part 3, 253-28 .

Kopp, R. (1981), "Measuring the Technical Efficiency of Production: A Comment," *Journal of Economic Theory* 25, 450-452.

Luenberger, D.G. (1992), "New Optimality Principles for Economic Efficiency and Equilibrium." *Journal of Optimization Theory and Applications,* 75 ,22 -264.

Russell, R.R. (1985), "Measuring of Technical Efficiency," *Journal of Economic Theory* 35, 1109- 26.

Russell, R.R. (1987), "On the Axiomatic Approach to the Measurement of Technical Efficiency," in W. Eichhhorn, ed. *Measurement in Economics: Theory and Applications of Economic Indices,* Heidelberg: Physica Verlag, 207-2 7.

Russell, R.R. (1990), "Continuity of Measures of Technical Efficiency," *Journal of Economic Theory* 51, 255-267.

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

Appendix

Proof of (2.5):

i) See Chambers, Chung and Färe (1998, p. 354) for a similar proof.

ii)

$$\vec{D}_i(y, x + \alpha 1^N; 1^N) = \sup\{\beta : (x - \beta 1^N + \alpha 1^N) \in L(y)\}$$

$$= \sup\{\beta : (x - (\beta + \alpha) 1^N) \in L(y)\}$$

$$= +\alpha + \sup\{\hat{\beta} : (x - \beta 1^N \in L(y)\}(\hat{\beta} = \beta - \alpha)$$

$$= \vec{D}_i(y, x; 1^N) + \alpha .$$

iiia) We give a contrapositive proof. Let $x \in L(y)$ with $x_n > 0, n = 1, \ldots, N$ and

$x \notin IsoqL(y)$. Then $Di(y, x) > 1$, and by strong disposability, there is an open

neighborhood $N_\varepsilon(x)$ of x $(\varepsilon = \min\{x_1 - D_i(y,x)x_1, \ldots, x_N - D_i(y,x)x_N\})$ such that

$N_\varepsilon(x) \in L(y)$. Thus $\vec{D}_i(y,x;1^N) > 0$ proving iiia).

iiib) Again we give a contrapositive proof. Let $\vec{D}_i(y,x;1^N) > 0$ then

$x - \vec{D}_i(y,x;1^N)1^N \in L(y)$ and since the directional vector is $1^N = (1,\ldots,1)$, each

$x_n, n = 1, \ldots, N$ can be reduced while still in $L(y)$. Thus $D_i(y, x) > 1$ and by the

Indication property for $D_i(y, x)$, $x \notin IsoqL(y)$. This completes the proof.

Remark on the proof of iiia): The following figure shows that when the directional

vector has all coordinates positive, for example 1^N, then $x_n > 0, n = 1, \ldots, N$ is

required. In the Figure 1, input vector a has $x_1 = 0$, and $\vec{D}_i(y,x;1^N) = 0$, but a is

not on the isoquant.
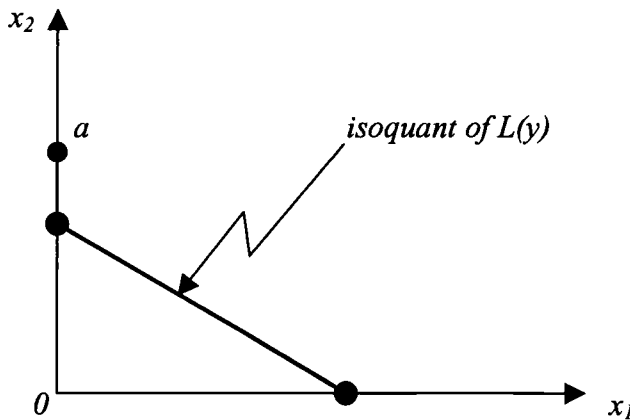
Figure 1. Remark on the proof of iiia).

This problem may be avoided by choosing the directional vector to have ones only for positive x's.

Proof of Theorem 1:

Assume first that the technology is as in (13), then

$$R_M(y,x) = \min\left\{\left(\prod_{n=1}^{N}\lambda_n\right)^{1/N} : (\lambda_1 x_1,\ldots,\lambda_N x_N) \in L(y), 0 < \lambda_n \leq 1, n = 1,\ldots,N\right\}$$

$$= \min\left\{\left(\prod_{n=1}^{N}\lambda_n\right)^{1/N} : D_i(\lambda_1 x_1,\ldots,\lambda_N x_N) \geq 1, 0 < \lambda_n \leq 1, n = 1,\ldots,N\right\}$$

$$= \min\left\{\left(\prod_{n=1}^{N}\lambda_n\right)^{1/N} : \left(\prod_{n=1}^{N}\lambda_n x_n\right)^{1/N} / H(y) \geq 1, 0 < \lambda_n \leq 1, n = 1,\ldots,N\right\}$$

$$= \min\left\{\left(\prod_{n=1}^{N}\lambda_n\right)^{1/N} : \left(\prod_{n=1}^{N}\lambda_n\right)^{1/N} \geq H(y) / \left(\prod_{n=1}^{N} x_n\right)^{1/N} 1, 0 < \lambda_n \leq 1, n = 1,\ldots,N\right\}$$

$$= H(y) / \left(\prod_{n=1}^{N} x_n\right)^{1/N} = 1/D_i(y,x).$$

Since $DF(y, x) = 1/Di(y, x)$ we have shown that (3) implies $R_M(y, x) = DF(x, y)$.

To prove the converse we first show that

$$R_M(y,\delta_1 x_1,\ldots,\delta_N x_N) = R_M(y,x) / \left(\prod_{n=1}^{N}\delta_n\right)^{1/N}, 0 < \delta_n \leq 1, n = 1,\ldots,N. \qquad (34)$$

To see this,

$$R_M(y, \delta_1 x_1, \ldots, \delta_N x_N) = \min\{ \left(\prod_{n=1}^{N} \lambda_n\right)^{1/N} : (\lambda_1 \delta_1 x_1, \ldots, \lambda_N \delta_N x_N) \in L(y),$$
$$0 < \lambda_n \leq 1, 0 < \delta_n \leq 1, n = 1, \ldots, N \}$$

$$= \left(\prod_{n=1}^{N} \delta_n\right)^{-1/N} \min\{ \left(\prod_{n=1}^{N} \lambda_n \delta_n\right)^{1/N} : (\lambda_1 \delta_1 x_1, \ldots, \lambda_N \delta_N x_N) \in L(y),$$
$$0 < \lambda_n \leq 1, 0 < \delta_n \leq 1, n = 1, \ldots, N \}$$

$$= \left(\prod_{n=1}^{N} \delta_n\right)^{-1/N} \min\{ \left(\prod_{n=1}^{N} \hat{\lambda}_n\right)^{1/N} : (\hat{\lambda}_1 \delta_1 x_1, \ldots, \hat{\lambda}_N \delta_N x_N) \in L(y),$$
$$0 < \hat{\lambda}_n \leq 1, 0 < \delta_n \leq 1, n = 1, \ldots, N \}$$

$$= R_M(y, x) \left(\prod_{n=1}^{N} \delta_n\right)^{-1/N}$$

where $\hat{\lambda}_n = \lambda_n \delta_n, n = 1, \ldots, N$. Thus (34) holds.

Next, assume that the Debreu-Farrell and the multiplicative Russell measures are equal, then

$$R_M(y, \delta_1 x_1, \ldots, \delta_N x_N) = R_M(y, x) / \left(\prod_{n=1}^{N} \delta_n\right)^{1/N} = DF(y, \delta_1 x_1, \ldots, \delta_N x_N)$$

thus

$$R_M(y, x) = DF(y, \delta_1 x_1, \ldots, \delta_N x_N) \left(\prod_{n=1}^{N} \delta_n\right)^{1/N}$$

and

$$DF(y, x) = DF(y, \delta_1 x_1, \ldots, \delta_N x_N) \left(\prod_{n=1}^{N} \delta_n\right)^{1/N}$$

Now we take $\delta_n = 1/x_n, n = 1,\ldots,N$ then

$$DF(y,x) = DF(y,1,\ldots,1)\left(\prod_{n=1}^{N} \delta_n\right)^{1/N}$$

Moreover, since the Debreu-Farrell measure is independent of units of measurement (Russell (1987), p. 215),[11] x_n can be scaled so that $x_n > 0, n = 1,\ldots,N$. Thus by taking $H(y) = DF(y,1,\ldots,1)$, and using (11) we have proved our claim.

Proof of Theorem 2:

First consider

$$A(y,x_1 - \delta_1,\ldots,x_N - \delta_N) =$$

$$= \max\left\{\frac{1}{N}\sum_{n=1}^{N} s_n : (x_1 - \delta_1 - s_1,\ldots,x_N - \delta_N - s_N) \in L(y)\right\},$$

$$= \max\left\{\frac{1}{N}\sum_{n=1}^{N}(s_n - \delta_n + \delta_n) : (x_1 - (\delta_1 + s_1),\ldots,x_N - (\delta_N + s_N)) \in L(y)\right\},$$

$$= -\frac{1}{N}\sum_{n=1}^{N}\delta_n + A(y,x),$$

where $s_n \geq 0, \delta_n \geq 0, n = 1,\ldots,N$.

---

[11] This was pointed out to us by R.R. Russell.

This is equivalent to

$$A(y,x) = \frac{1}{N}\sum_{n=1}^{N}\delta_n + A(y, x_1 - \delta_1,\ldots, x_N - \delta_N)$$

Take $\delta_n = x_n$ and define $-F(y) = A(y,0)$, then since equality between the directional distance function and the additive measure holds,

$$\vec{D}_i(y,x;1^N) = A(y,x) = \frac{1}{N}\sum_{n=1}^{N}x_n - F(y).$$

Next, let $x \in \overline{C}(L(y))$, then for some $x \in IsoqL(y)$, and $\delta \geq 0$,

$$\vec{D}_i(y,x;1^N) = \vec{D}_i(y,\hat{x} + \delta 1^N;1^N) = \vec{D}_i(y,\hat{x};1^N) + \delta.$$

Since $\hat{x} \in IsoqL(y)$, $\vec{D}_i(y,x;1^N) = \delta$.

Next,

$$A(y,x) = \max\left\{\frac{1}{N}\sum_{n=1}^{N}s_n : \sum_{n=1}^{N}(x_n - s_n)/N - F(y) \geq 0\right\}$$

$$= \max\left\{\frac{1}{N}\sum_{n=1}^{N}s_n : \sum_{n=1}^{N}(\hat{x}_n + \delta - s_n)/N - F(y) \geq 0\right\}$$

$$= \max\left\{\frac{1}{N}\sum_{n=1}^{N}s_n : \delta + \sum_{n=1}^{N}\hat{x}_n/N - F(y) \geq \frac{1}{N}s_N\right\}$$

$$= \delta,$$

since $\hat{x} \in IsoqL(y)$, thus $\vec{D}_i(y,x;1^N) = A(y,x)$.

ESSAY 2: DIRECTIONAL AND SHEPHARD'S DISTANCE FUNCTIONS: NEW LINK AND ITS IMPLICATION TO MEASURING PRODUCTIVITY CHANGES

Valentin Zelenyuk[1]

Abstract

In this study we reveal a new relationship between the directional distance function and Shephard's (1970) distance functions. We then apply this result to show a relationship between a productivity index defined in terms of the directional distance function and other popular productivity indexes.

Keywords: Malmquist and Total Factor Productivity indexes, distance functions, DEA.

JEL Code: D21, D24, D29.

Introduction

Characterization of multioutput technology in production economics is often done via Shephard's (1970) input or output distance functions, or their duals, the cost and revenue functions (Färe and Primont, 1995). Recently, a more general characterization was discovered—the *directional* distance function[2]—a dual to the profit function, and a generalization of Shephard's distance functions (Chambers et al., 1998).

In this study we unveil a new relationship between the directional and Shephard's distance functions that exist under the constant returns to scale technology (section 1). As an application (section 2), we then use this result to show a relationship between a productivity index defined in terms of the directional distance function and some popular productivity indexes.

Technology Characterizations

Let $x \in \mathfrak{R}^N_+$ and $y \in \mathfrak{R}^M_+$ be the vectors of inputs and outputs, respectively, and define the *technology* set by: $T = \{(x,y) : x \ can \ produce \ y\}$. We assume that T is a closed set with 'freely disposable' inputs and outputs. In addition, we assume that the 'no free lunch' and 'doing nothing is feasible' axioms[3] hold.

To characterize T define the Shephard's output and input distance functions, respectively, as

---

[2] For an extensive exploration of the directional distance function see Chambers et al. (1996, 1998), Briec (1999), Färe and Grosskopf (2000). It is also an analog of the shortage function introduced by Luenberger (1992).

[3] Technically, the free (or strong) disposability of inputs and outputs says: $(x,y) \in T$ and $(x',-y') \geq (x,-y) \Rightarrow (x',y') \in T$. The "doing nothing is feasible" axiom simply says: $(0, 0) \in T$, while the "no free lunch" axiom insures that: for any $(x,y) \in T$ if $x = 0$ then $y = 0$ (for details, see Chambers et al. (1998) or Färe and Primont (1995)).

$$D_o(x,y) = \inf \{\gamma : (x, y/\gamma) \in T\} \quad \text{and} \quad D_i(y,x) = \sup \{\lambda : (x/\lambda, y) \in T\}. \qquad (1.1)$$

To obtain our main result, we make use of the following properties (Färe and Primont, 1995):

$$(y,x) \in T \iff D_o(x,y) \leq 1 \iff D_i(y,x) \geq 1 \qquad \text{(complete characterizations of } T) \; (1.2)$$

$$D_o(x,ky) = kD_o(x,y) \quad \text{and} \quad D_i(y,kx) = kD_i(y,x), \; \forall \; k > 0 \qquad \text{(homogeneity)} \qquad (1.3)$$

$$CRS \iff D_o(kx,y) = (1/k)D_o(x,y), \text{ and } D_i(ky,x) = (1/k)D_i(y,x), \; \forall \; k > 0. \; (CRS) \; (1.4)$$

$$CRS \iff D_o(x,y) = 1 / D_i(y,x) \qquad \text{(reciprocal relationship under CRS)} \qquad (1.5)$$

where, *CRS* means 'technology exhibits constant returns to scale', defined as: $\lambda T = T, \; \forall \; \lambda > 0$.

Next, define the *directional* distance function

$$D_d(x,y;-g_x,g_y) = \sup\{\theta : (x-\theta g_x, y+\theta g_y) \in T\}, \qquad (1.6)$$

where (g_x,g_y) is some nonzero vector in $\mathcal{R}^N_+ \times \mathcal{R}^M_+$ that specifies the direction in which the distance between observation (x,y) and the boundary of the technology set T is measured. It was shown (Chambers et al., 1998) that $D_d(\cdot)$ is a complete characterization of T. In particular,

$$(y,x) \in T \qquad \iff \qquad D_d(x,y;-g_x,g_y) \geq 0, \qquad (g_x,g_y) \geq 0, (g_x,g_y) \neq 0. \qquad (1.7)$$

Moreover, $D_i(\cdot)$ and $D_o(\cdot)$ are special cases of $D_d(\cdot)$. Specifically (see Chambers et al., 1998), if the directional vector $(-g_x,g_y)$ is $(-x,0)$ or $(0,y)$, then (1.6) reduces, respectively, to:

$$D_d(x,y;-x,0) = 1 - 1/D_i(y,x) \qquad \text{or} \qquad D_d(x,y;0,y) = 1/D_o(x,y) - 1. \qquad (1.8)$$

Note, that this relationship requires either the input or output direction to be zero. Our goal is to find a similar straightforward relationship when both directions are accounted for (i.e., nonzero). In the following theorem we find one such relationship that exists under *CRS*. In contrast to *(1.8)*, the relationship holds for any direction *(-g$_x$,g$_y$)* that is positioned anywhere between (and including) the directions in *(1.8)*, i.e., a direction described by the linear combination *(-αx,βy) = α(-x,0) + β(0,y)*, *α* , *β* ∈ *ℜ$_+$*. We formalize and prove this claim below.

Theorem.[4]   Let *(-g$_x$, g$_y$) = (-αx, βy) ≠ 0*, *α,β* ∈*ℜ$_+$* , then *T* exhibits CRS
if and only if

$$D_d(x,y;-g_x,g_y) = (D_i(y,x)-1) / (\alpha D_i(y,x)+\beta),$$

and   *(1.9)*

$$D_d(x,y;-g_x,g_y) = (1-D_o(x,y)) / (\beta D_o(x,y)+\alpha).$$

Proof:

"⇒" part:

suppose *T* exhibits CRS, then: $D_d(x,y;-\alpha x,\beta y) \equiv sup\{\theta : (x-\theta(\alpha x),y+\theta(\beta y) \in T\}$

$= sup\{\theta : (x(1-\alpha\theta), y(1+\beta\theta)) \in T\} = sup\{\theta : D_i(y(1+\beta\theta), x(1-\alpha\theta)) \geq 1\}$ (by *(1.2)*)

$= sup\{\theta : D_i(y,x)(1-\alpha\theta) \geq (1+\beta\theta)\}$ (using *(1.3)*, and assuming *CRS* using *(1.4)*)

$= sup\{\theta : (D_i(y,x)-1)/(\alpha D_i(y, x)+\beta) \geq \theta\} = (D_i(y,x)-1)/(\alpha D_i(y,x)+\beta).$

---

[4] A special case of this theorem (when *(α,β) = (1,1)*) was independently discovered by Boussemart et al (2001).

By the same logic, it follows that:

$$D_d(x,y;-\alpha x,\beta y) \equiv sup\{\theta : (x-\theta(\alpha x), y+\theta(\beta y)) \in T\} = sup\{\theta : (x(1-\alpha\theta), y(1+\beta\theta)) \in T\}$$

$$= sup\{\theta : D_o(x(1-\alpha\theta), y(1+\beta\theta)) \leq 1\} = sup\{\theta : D_o(x,y)(1+\beta\theta) \leq (1-\alpha\theta)\}$$

$$= sup\{\theta : \theta \leq (1-D_o(x,y)) / (\beta D_o(x,y)+\alpha)\} = (1-D_o(x,y)) / (\beta D_o(x,y)+\alpha).$$

"$\Leftarrow$" part: assume *(1.9)* is true, then after simple manipulations *(1.9)* is rewritten as:

$$D_i(y,x) = (1 +\beta D_d(x,y; -\alpha x,\beta y))/(1-\alpha D_d(x,y; -\alpha x,\beta y)),$$

and   *(1.10)*

$$D_o(x,y) = (1-\alpha D_d(x,y; -\alpha x,\beta y))/(1+\beta D_d(x,y; -\alpha x,\beta y))$$

$$\Rightarrow \qquad D_i(y,x) = 1/D_o(x,y)$$

By *(1.5)*, the last statement is true if and only if T exhibits *CRS*.

Q.E.D.[5]

In words, this theorem tells us that under the *CRS* one can use Shephard's distance functions to solve explicitly for the directional distance function (and visa versa) with any direction between *(x,0)* and *(0,y)* determined by appropriate selection of positive scalars α and β. Since the assumption of *CRS* is often used in many economic studies, we expect this theorem to find many applications in analyses involving multi-output technologies--where Shephard's distance functions are currently used. In the next section we illustrate one such application.

An Implication: Measuring the Productivity Growth

Our goal here is to apply the theorem to the measurement of productivity changes (growth or a decline) of an economic system (a country, region, industry, firm, etc.). In particular we aim to show a relationship between a productivity index defined in terms of the directional distance function on one side and some popular productivity indexes on the other side. To show this relationship we use the theorem proven above. Recalling that our theorem requires only the assumption of *CRS* (in addition to standard regularity conditions), we note that this assumption is a natural one in the theory of economic growth. It is also often used in empirical productivity growth measurement if one takes the so-called 'economic approach' (Diewert, 1992b, p. 243).

In general, the idea of measuring changes in the productivity of, say, a firm is based on comparing its performance in one period relative to another. If this firm produces one output using one input then a simple but intuitive measure of productivity changes would be:

$$SFP = (y^{t+1}/y^t) / (x^{t+1}/x^t), \tag{2.1}$$

where the superscripts t, $t+1$ indicate the time periods in which x and y were observed. Here, we call this measure the Single Factor Productivity (*SFP*) index (e.g., labor productivity index). Intuitively, this index can be interpreted as the ratio of 'single output index' to the 'single input index', or as the ratio of the 'average products' in the two periods.

Although most technologies involve more than one output (input), the measure in *(2.1)* is useful as a benchmark for construction and comparison of more general productivity indexes that account for *all* the factors (inputs) and multiple outputs (Diewert, 1992a, 1992b)—the measures often called the *Total* Factor Productivity

---

[5] Also note that if $(\alpha,\beta) = (0,1)$ or $(\alpha,\beta) = (1,0)$, then *(1.8)* follows immediately from the theorem.

(*TFP*) indexes. The Fisher and Törnqvist productivity indexes are may be the most popular examples of such generalization. Another generalization of the *SFP* is the *Malmquist Productivity Index (MPI)*, a measure of productivity changes based on Shephard's distance functions. For *output* and *input* orientations it is defined, respectively, as[6]:

$$MPI_o = MPI_o(x^t, y^t, x^{t+1}, y^{t+1}) = \left[\left(\frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \right) \left(\frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^{t+1}(x^t, y^t)} \right) \right]^{1/2} \qquad (2.2)$$

and

$$MPI_i = MPI_i(x^t, y^t, x^{t+1}, y^{t+1}) = \left[\left(\frac{D_i^t(y^{t+1}, x^{t+1})}{D_i^t(y^t, x^t)} \right) \left(\frac{D_i^{t+1}(y^{t+1}, x^{t+1})}{D_i^{t+1}(y^t, x^t)} \right) \right]^{1/2}, \qquad (2.3)$$

where $D_o^s(x^l, y^l) = inf\{\theta : (x^l, y^l/\theta) \in T^s\}$ and $D_i^s(y^l, x^l) = sup\{\lambda : (x^l/\lambda, y^l) \in T^s\}$ are the distance functions *(1.1)* extended to the intertemporal framework so that they relate observations *(x,y)* in period *l* to the technology in period *s* (*l, s = t, t+1*).

Using *(1.5)*, it follows that under *CRS* it is always true that $MPI_o = 1/MPI_i$. More importantly, both are generalizations of *SFP* defined in *(2.1)* [7] and of such *TFP* indexes as the Fisher and Törnqvist productivity indexes (Färe, Grosskopf and Roos, 1997, p. 140).

Alternatively, using the *directional* distance function characterization and adopting the idea of Chung et al. (1997) we define the *output* and *input* based Malmquist-Luenberger Productivity Indexes (with the directional vector *(-g_x, g_y)* = *(-αx, βy)* ≠ *0*, *α, β* ∈ $\Re_+$) as:

---

[6] The origin of the idea of *MPI* is found in Caves, Christensen and Diewert (1982). Its computable form via the linear programming dates back to Färe et al. (1989).

[7] For details, see Berg, et al., (1992), and Färe and Grosskopf (1996).

$$MLPI_o = \left[\frac{\left(1 + \beta D_d^t(x^t, y^t; -\alpha x^t, \beta y^t)\right)}{\left(1 + \beta D_d^t(x^{t+1}, y^{t+1}; -\alpha x^{t+1}, \beta y^{t+1})\right)} \frac{\left(1 + \beta D_d^{t+1}(x^t, y^t; -\alpha x^t, \beta y^t)\right)}{\left(1 + \beta D_d^{t+1}(x^{t+1}, y^{t+1}; -\alpha x^{t+1}, \beta y^{t+1})\right)} \right]^{1/2}$$

(2.4)

and

$$MLPI_i = \left[\frac{\left(1 - \alpha D_d^t(x^t, y^t; -\alpha x^t, \beta y^t)\right)}{\left(1 - \alpha D_d^t(x^{t+1}, y^{t+1}; -\alpha x^{t+1}, \beta y^{t+1})\right)} \frac{\left(1 - \alpha D_d^{t+1}(x^t, y^t; -\alpha x^t, \beta y^t)\right)}{\left(1 - \alpha D_d^{t+1}(x^{t+1}, y^{t+1}; -\alpha x^{t+1}, \beta y^{t+1})\right)} \right]^{1/2}$$

(2.5)

respectively. Here, $D_d^s(x^l, y^l; -g_x, g_y) = sup\{ \theta : (x^l - \theta g_x, y^l + \theta g_y) \in T^s \}$ is the directional distance function *(1.6)* extended to the intertemporal framework so that it relates the observations *(x,y)* in period l to the technology in period s *(l, $s = t$, $t+1$)*. The corresponding extension of our theorem is trivial and therefore omitted. (Also, note that if *($-g_x, g_y$)* is *(0,y)* or *(-x,0)*, then due to (1.8) the *(2.4)* or *(2.5)* is reduced to (2.2) or (2.3), respectively.)

In the spirit of such *TFP* indexes as Fisher and Törnqvist, defined as ratios of an 'output quantity index' to an 'input quantity index' (Diewert, 1992a, 1992b), we use *(2.4)* and *(2.5)* to define a new measure: The *Total Malmquist-Luenberger Productivity Index*

$$TMLPI \equiv MLPI_o / MLPI_i ,$$

(2.6)

Clearly, after collecting terms of $MLPI_i / MLPI_o$, and applying the result *(1.10)* of the theorem above, it follows that under *CRS*

$$TMLPI = MPI_o \qquad \forall x \in \mathfrak{R}^N_+, y \in \mathfrak{R}^M_+ .$$

(2.7)

Thus, *TMLPI* is the generalization of the MPI_o. Recalling that MPI_o is a generalization of *SFP* and of such *TFP* indexes as the Fisher and Törnqvist

productivity indexes (Färe et al., 1997, p. 140), the result in *(2.7)* also tells us that so is the *TMLPI*.[8]

An important question now is whether *TMLPI* gives us anything new and valuable that cannot be inferred from either MPI_o or MPI_i. It seems that it does. In particular, *TMLPI* accounts for information about the *weight* of the output orientation relative to the input orientation, described by the directional vector *(-αx, βy)*. Our theorem showed that this weight does not matter under *CRS*, yielding *(2.7)* and equivalence with the *TFP* for a single-output-single-input case. What if the technology does not exhibit *CRS*, as it is assumed in some studies? Or, what if the productivity index is decomposed into parts that are based on non-CRS assumption?[9] Then the measurement based *only on the input* orientation *(*i.e., using MPI_i*)* may give quite different conclusions from those obtained using *only the output* orientation *(*i.e., using MPI_o*)*. If the subject of study suggests that *both* orientations are important, and if their relative importance can be characterized by the direction of measurement *(-αx, βy)*, then *TMLPI* might be a better choice.

References

Berg, S., F. Førsund and E. Jansen (1992), "Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking," *Scandinavian Journal of Economics* 94, S211-S228.

Boussemart, J-P., W. Briec, K. Kerstens and J-C. Poutineau (2001), "Luenberger and Malmquist Productivity Indices: Theoretical Comparisons and Empirical Illustration," working paper.

Briec, W. (1999), "Holder Distance Function and Measurement of Technical Efficiency," *Journal of Productivity Analysis* 11:2, 111-31.

---

[8] If in addition to CRS, technology is inverse homothetic, then *MPI* will be equal to the Hicks-Moorsten Productivity index (see Färe et al., 1997, p.136), thus so will *TMLPI*.

[9] For a concise review of measuring productivity changes with *MPI* under non-CRS assumption and of decomposition issues, see Grifell-Tatjé and Lovell (1999) and Färe, Grosskopf and Roos (1997). Also see Førsund (1997) and Balk (2000) for more extensive explorations.

Caves, D., L. Christensen and W. E. Diewert (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica* 50:6, 1393-1414.

Chambers, R., Y. Chung and R. Färe (1996), "Benefit and Distance Functions," *Journal of Economic Theory* 70, 407-419.

Chambers, R., Y. Chung and R. Färe (1998), "Profit, Directional Distance Functions and Nerlovian Efficiency," *Journal of Optimization Theory and Application* 98:2, 351-364.

Chung, Y, R. Färe and S. Grosskopf (1997), "Productivity and Undesirable Outputs: A Distance Function Approach," *Journal of Environmental Management* 51, 229-240.

Diewert, W.E. (1992a), "The Measurement of Productivity," *Bulletin of Economic Research* 44:3, 163-98.

Diewert, W.E. (1992b), "Fisher Ideal Output, Input, and Productivity Indexes Revisited," *Journal of Productivity Analysis* 3: 3, 211-48.

Färe, R., S. Grosskopf, B. Lindgren, and P. Roos (1989), "Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach," Discussion Paper 89-3, Southern Illinois University at Carbondale.

Färe, R., and S. Grosskopf (1996), *Intertemporal Production Frontiers: With Dynamic DEA*, Kluwer Academic Publishers, Boston.

Färe, R., S. Grosskopf and P. Roos (1997), Malmquist Productivity Indexes: a Survey of Theory and Practice, in: R. Färe, S. Grosskopf and R. Russell, eds., *Index Numbers: Essays in Honor of Sten Malmquist*, Kluwer Academic Publishers, Boston.

Färe, R., and D. Primont (1995), *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Färe, R. and S. Grosskopf (2000), "Theory and Application of Directional Distance Functions," *Journal of Productivity Analysis*, 13:2, 93-103

Førsund, F. (1997), "The Malmquist Productivity Index, TFP and scale," Memorandum NO 233, Department of Economics, School of Economics and Commercial Law, Göteborg University.

Grifell-Tatjé and Lovell (1999), "A Generalized Malmquist Productivity Index," *Sociedad de Estadística e Investigación Operativa Top* 7:1, pp. 81-101.

Luenberger, D.G. (1992), New Optimality Principles for Economic Efficiency and Equilibrium, *Journal of Optimization Theory and Application* 75, 221-264.

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

ESSAY 3: SCALE EFFICIENCY: EQUIVALENCE OF PRIMAL AND DUAL MEASUREMENTS

Valentin Zelenyuk[1]

Abstract

In a recent paper, Sueyoshi (1999) examines relations between the primal and dual measures of scale efficiency. As one of the approaches, he discusses the result of Färe and Grosskopf (1985) who provided conditions for the equivalence of such measures. Both papers opened a new question: What type of technology, if any, is consistent with such a condition? I address this question here and answer it by resorting to the concept of homotheticity.

Keywords: Scale Efficiency Measurement, Homotheticity, Duality.

Introduction

Data limitation in empirical studies is more of a rule rather than an exception. In production studies, for example, data on *all* inputs used in production are often hard or even impossible to find. Due to duality theory in economics (e.g., Shephard, 1970), vital economic information on production technology (a relation between inputs and outputs) can be retrieved from the cost-output data—information that is more often available to researchers than the input-output data.

In the context of production efficiency analysis, a methodology to extract some technology information using the cost rather than input data was first discussed by Färe and Grosskopf (1985). In their work, the authors provided conditions for equivalence of the primal (based on input-output data) and dual (based on cost-output data) *scale efficiency* measures. This condition requires that the input allocative efficiency estimated under the constant returns to scale (CRS) assumption is equal to that estimated under the variable returns to scale (VRS) assumption. This same condition is also found in Sueyoshi (1999) and, informally, in Seitz (1970). The goal of this paper is to determine what type of technology, if any, is consistent with such a condition.

Theoretical Framework

Both Färe and Grosskopf (1985) and Sueyoshi (1999) deal with technologies approximated with a convex disposal hull[2]. Here, to obtain general results, I follow Shephard (1970) and Färe and Primont (1995) characterization of technology using the input correspondence $L: \Re^M_+ \to 2^{\Re^N_+}$ that assigns to each output vector $y \in \Re^M_+$ the subset of all input vectors $x \in \Re^N_+$ that can produce this particular output level y, i.e.,

---

[2] Specifically, they use the activity analysis models or the data envelopment analysis method (DEA).

$$L(y) = \{x : x \text{ can produce } y\}, \ y \in \mathfrak{R}^M_+ . \tag{1}$$

To characterize L, I use an implicit function $F_i : \mathfrak{R}^M_+ \times \mathfrak{R}^N_+ \rightarrow \mathfrak{R}_+ \cup \{+\infty\}$, defined as

$$F_i(y,x) = \inf \{\lambda : \lambda x \in L(y)\}, \tag{2}$$

and known as the *Farrell* (1957) *input oriented measure of technical efficiency.* This function is a reciprocal of the Shephard's (1970) input distance function and, given standard regularity conditions on L, completely characterizes L due to[3],

$$L(y) = \{x : F_i(y,x) \leq 1\}, \qquad y \in \mathfrak{R}^M_+ . \tag{3}$$

The measure in *(2)* is often used to construct the *input oriented scale efficiency* measure[4]

$$S_i(y,x) = F_i(y,x|C) / F_i(y,x|V), \tag{4}$$

where, here and later, notation "$|C$" ("$|V$") is used to indicate that the function (correspondence) is estimated with the constant (variable) returns to scale assumption[5].

---

[3] For the definition of the Shephard's distance function, regularity conditions and the proof of this and other properties, see Färe and Primont, 1995.

[4] The origin of this measure goes back to at least Førsund and Hjalmarsson (1979). For alternative ways of measuring *scale* issues in DEA, see, for example, Banker et al. (1984), Førsund (1997) and Sueyoshi (1999).

[5] For the ways the constant and variable returns to scale assumptions are defined and modeled in the efficiency analysis framework, see Färe, Grosskopf and Lovell (1994). See also Sueyoshi (1999) for a recent treatment.

The approach presented above is based on the input-output data, (x,y), and is called the primal approach. To outline the dual approach, which is based on cost-output data, let w be a vector of strictly positive input prices and define the *cost function* as

$$C(y,w) = min_x \{wx : x \in L(y)\} , \quad y \in \Re^M_+ : L(y) \neq \varnothing, w \in \Re^N_{++}. \qquad (5)$$

Given convexity of $L(y)$, the dual analogue of (3) is (see Färe and Primont, 1995)

$$L(y) = \{x : \quad C(y,w) / wx \leq 1 , \quad \forall w > 0\}, \qquad (6)$$

i.e., given convexity the technology L can be completely characterized by the cost function.

The *cost efficiency* measure, a dual analogue of (2), is defined as

$$CE(y,w,x) = C(y,w) / wx. \qquad (7)$$

And the *cost-scale efficiency* (Färe and Grosskopf, 1985) is defined as

$$S_c(y,w,x) = CE(y,x|C) / CE_i(y,x|V). \qquad (8)$$

Given convexity of L, the relationship between the two approaches is obtained through the duality between $F_i(y,x)$ and $C(y,w)$, which can be stated via the Mahler inequality

$$C(y,w) / wx \leq F_i(y,x). \qquad (9)$$

This inequality can be 'closed' by defining the *allocative efficiency* measure as the 'residual',

$$C(y,w) \, / \, wx \equiv F_i(y,x) \, AE(y,x,w). \tag{10}$$

The two measures of scale efficiency, $S_i(y,x)$ and $S_c(y,w)$, are functions of different variables and, in general, may yield different efficiency scores. There is, however, a special case when they are equal. Specifically, from (4), (7), (8) and (9) it follows that

$$S_i(y,x) = S_c(y,w,x) \qquad \textit{if and only if} \quad AE(y,x,w|C) = AE(y,x,w|V). \tag{11}$$

This is exactly the result reached by Färe and Grosskopf (1985) and Sueyoshi (1999). An early reference for an intuitive explanation of this case is also found in Seitz (1970).

An imperative question now is: What type of technology can ensure the equivalence of these measures? That is, when is it true that $AE(y,x,w|C) = AE(y,x,w|V)$. An answer to this question is given in the next section.

Input Scale Homotheticity and Scale Efficiency Measurement

Consider a technology $L(y)$ that satisfies the condition that will be referred to here as *input scale homotheticity (ISH)* and defined as

$$L(y|V) = G(y)L(y|C), \qquad y \in \mathfrak{R}^M_+, \tag{12}$$

where $G(y)$ is some function $G : R^M_+ \to R_+$ consistent with standard regularity conditions on $L(y)$. Intuitively, this is a case when a characterization of a VRS technology can be decomposed into two parts: *(i)* a CRS technology

characterization and *(ii)* some real valued function reflecting the scale of the production activity.

Also note that *ISH* is related to the concept of *input homotheticity*[6] (see Färe and Mitchell, 1993, and Färe and Primont, 1995). In particular, recall that the *input homotheticity* is defined as

$$L(y) = H(y)L(1^N), \qquad y \in \mathfrak{R}^M_+ \tag{13}$$

for some function $H: R^M_+ \to R_+$ consistent with standard regularity conditions on $L(y)$. Now, note that for a single-output and multi-input case, *ISH* implies that

$$L(y|V) = G(y) \, y \, L(1|C), \qquad y \in \mathfrak{R}^I_+. \tag{14}$$

Rearranging and applying *(ISH)* to r.h.s. of (14) and then letting $H(y) = G(y) \, y \, / \, G(1)$, yields

$$L(y|V) = H(y)L(1|V), \qquad y \in \mathfrak{R}^I_+, \tag{15}$$

implying that for $y \in \mathfrak{R}^I_+$, the *ISH* technology is also input homothetic.

To see what implications such peculiar technology has towards the relationship between primal and dual scale efficiency measures, note how this technology is interpreted in terms of the cost function and the Farrell measure of technical efficiency (see appendix for proofs)

$$ISH \quad \Leftrightarrow \quad C(y,w|V) = G(y)C(y,w|C), \qquad y \in \mathfrak{R}^M_+, \; L(y) \neq \varnothing, \; w \in \mathfrak{R}^N_{++} \tag{16}$$

and

$$ISH \quad \Leftrightarrow \quad F_i(y,x|V) = G(y)F_i(y,x|C), \qquad y \in \mathfrak{R}^M_+, \; x \in \mathfrak{R}^N_+. \tag{17}$$

---

[6] I thank Rolf Färe for this insightful comment.

Using these results, it immediately follows that

$$ISH \quad \Leftrightarrow \quad AE(y,x,w|C) = AE(y,x,w|V). \qquad (18)$$

In words, the equivalence of input allocative efficiency estimated under the CRS assumption to that estimated under the VRS assumption can happen *if and only if* the technology is input scale homothetic. Combining this with the result reached by Färe and Grosskopf (1985) and Sueyoshi (1999), and restated in (11), yields the answer to the research question of this paper:

The equivalence of the dual and primal scale efficiency measures, i.e., $S_i(y,x) = S_c(y,w,x)$, can be achieved *if and only if* technology is input scale homothetic. Moreover, note that in this case both scale efficiency measures equal the reciprocal of $G(y)$—the function reflecting the scale of the production activity of such technology. Altogether,

$$S_i(y,x) = S_c(y,w,x) = 1 / G(y) \qquad \Leftrightarrow \qquad ISH. \qquad (19)$$

Concluding Remarks

In this note, I show that the *necessary and sufficient* condition for the equivalence of primal and dual scale efficiency measures provided by Färe and Grosskopf (1985) and Sueyoshi (1999) holds *if and only if* technology is of a peculiar type—*input scale homothetic.*

This study opens at least three directions for further research: *(i)* development of empirical tests for identifying *ISH* technology, *(ii)* theoretical investigation of how restrictive the assumption of *ISH* technology is, and *(iii)* identification of other areas of application of *ISH* technology.

References

Banker, R., A.Charnes, W.W. Cooper (1984), "Some Models for EstimatingTechnical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science* 30 (9), pp. 1078-92.

Charnes, A., W.W. Cooper and E.Rhodes (1978), "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research* 2 (6), pp. 429 - 444.

Färe, R. and S. Grosskopf (1985) "A Nonparametric Cost Approach to Scale Efficiency," *Scandinavian Journal of Economics 87* (4), pp. 594-604.

Färe, R. and S. Grosskopf and C.A.K. Lovell, (1994) *Production Frontier,* Cambridge, Cambridge University Press.

Färe, R. and T. Mitchell, (1993) "Multiple Outputs and Homotheticity," *Southern Economic Journal,* pp.287-296.

Färe, R. and D. Primont (1995), *Multi-Output Production and Duality: Theory and Applications,* Boston: Kluwer Academic Publishers.

Färe, R. and Sung Ko Li (2001), "A Nonparametric Test of Translation Homoetheticity *Economics Letters* 72, pp. 341-345.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, General, 120, part 3, pp.253-281.

Førsund, F. and L. Hjalmarsson (1979), "Generalized Farrell Measures of Efficiency: an Application to Milk Processing in Swedish Dairy Plants," *Economic Journal* 89, pp. 294-315.

Førsund, F. "On the Calculation of the Scale Elasticity in DEA Models" *Journal of Productivity Analysis* 7 (2-3), 283-302.

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

Seitz, W.D. (1970), "The Measurement of Efficiency Relative to a Frontier Production Function," *American Journal of Agricultural Economics* 52, pp. 505-511.

Sueyoshi (1999), "DEA Duality on Returns to Scale (RTS) in Production and Cost Analyses: An Occurance of Multiple Solutions and Differences Between Production-Based and Cost-Based RTS Estimates," *Management Science* v.45, n.11, pp. 1593-1608.

Appendix 1: Proofs

Proof of (16)

"$\Rightarrow$":

$$ISH \Rightarrow C(y,w|V) \equiv min_x \{wx : x \in L(y|V)\} = min_x \{wx : x \in G(y)L(y|C)\} \quad (by\ (ISH))$$

$$= G(y)\ min_{x'} \{wx' : x' \in L(y|C)\} = G(y)C(y,w|C), \quad where\ x'=x/G(y).$$

"$\Leftarrow$":

Now assume: $C(y,w|V) = G(y)C(y,w|C)$, then using the duality result *(6)*:

$$L(y|V) \equiv \{x : wx \geq C(y,w|V),\ \forall w > 0\} = \{x : wx \geq G(y)C(y,w|C),\ \forall w > 0\}$$

$$= G(y)\{x' : wx' \geq C(y,w|C),\ \forall w > 0\} = G(y)\ L(y|C), \text{ where } x' = x/G(y).$$

<div align="right">*Q.E.D.*</div>

Proof of (17)

"$\Rightarrow$": $ISH \Rightarrow F_i(y,x|V) = inf_\lambda\{\lambda : x\lambda \in L(y|V)\} = inf_\lambda\{\lambda : x\lambda \in G(y)L(y|C)\}$
(by (ISH))

$$= G(y)\ inf_{\lambda'}\{\lambda' : x\lambda' \in L(y|C)\} = G(y)F_i(y,x|C), \text{ where } \lambda'=\lambda/G(y).$$

"$\Leftarrow$": Now assume that $F_i(y,x|V) = G(y)F_i(y,x|C)$, then using *(3)* yields:

$L(y|V) = \{x: F_i(y,x|V) \leq 1\} = \{x: G(y)F_i(y,x|C) \leq 1\}$

$= \{x: F_i(y,x/G(y))|C) \leq 1\}$ (since $F_i(y,x)$ is homogeneous of degree -1 in x)

$= G(y) \{x': F_i(y,x')|C) \leq 1\} = G(y)L(y|C)$ (where $x' = x/G(y)$).

$Q.E.D.$

THEORETICAL ESSAYS ON AGGREGATION ISSUES IN EFFICIENCY AND PRODUCTIVITY ANALYSIS

INTRODUCTION

An Overview

Chapter 2 of this dissertation outlined various measures of efficiency and productivity, their relationship and approaches to estimation. The objective of all the measures so far was to give an efficiency estimate of a *firm*. While this is important information to know, often researchers also want to have an idea on efficiency of a *group of firms*. For example, researchers may want to know efficiency of the entire industry or its representative sample, which then may be compared to its potential, to its efficiency in a different period, to efficiency of the same industry in another region, or to another industry. In addition, researchers may be interested in comparing efficiencies of various groups in an industry. For instance, in the next chapter I attempt to measure the scale efficiency of different *strategic groups* in the same industry and the industry itself in order to understand the relationship between the scale economies and rising industry concentration.

An important question is therefore: How to measure an *efficiency of a group*? Many measures were suggested—most of them are intuitive but unfortunately ad hoc, in the sense that they were not derived from some type of economic (optimization) problem or some aggregation consistency criteria. Strictly speaking, without such economic or mathematical consistency background none of the suggested measures can be thought of as appropriate. Moreover, if some researcher's findings contradict the results of other researchers just because of using a different way of aggregation then there is no rule that can tell whose results are most accurate and reliable. In other words, a theory of *aggregation over firms* is needed to derive measures of *group efficiency* justified by economic theory or/and consistency criteria. Essays 4 through 6 make contribution into this theory.

Another aggregation question often arises when information for empirical work is available on a level of aggregation that is different from the one used in the theory. For example, instead of data on inputs in production (e.g., labor, materials) researchers may have only the corresponding cost data (e.g., labor cost, material cost)—i.e., the linearly aggregated input data with input prices being the weights. An important question is how such an aggregation impacts the estimation results. Answers to this question would constitute the theory of *aggregation over goods* (inputs, outputs). **Essays 7** and **8** attempt to contribute to building such theory for the Farrell type measures estimated with the Data Envelopment Analysis (DEA).

Figure 2 presents a taxonomy of the contributions of each essay into various aggregation theories for various efficiency measures.
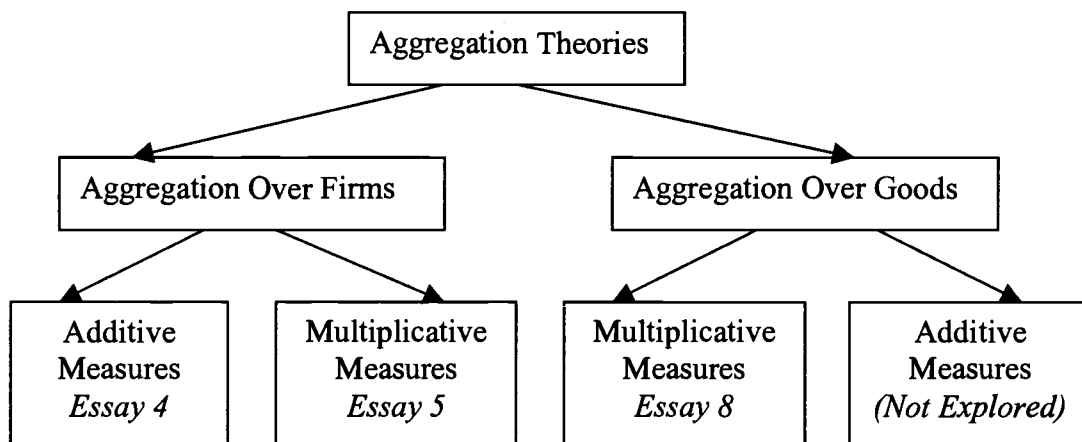
Figure 2. Taxonomy of the Aggregation Theories

The rest of this chapter is structured as following: I start with a few remarks on the aggregation issues in economics, then discuss the evolution of these issues in

the efficiency analysis, and then focus on contributions to these issues from the essays of this dissertation.

Aggregation Issues in Economics

Aggregation issues play an important role in theoretical and applied economic analysis. They frequently arise in empirical work where data often exist on a different level of aggregation than what particular theory requires. The objective of aggregation theories, in essence, is therefore to establish a theoretical link explaining at least two main questions: (1) what are the consequences of using more or less aggregated information than the theory requires, and (2) what are the conditions under which these consequences may be eliminated or reduced. The answers to these questions form the basis of aggregation theories. The two common consequences that the aggregation theorists are looking for are: (i) the bias due to aggregation (its size, direction, bounds, etc), and (ii) the properties preserved and lost due to aggregation. Both undesirable and desirable consequences are incorporated into 'consistency criteria,' which then are used to determine if a particular aggregation is consistent (with the criteria) or not.

Among seminal examples of such aggregation theories in economics are: the aggregation of capital (Klein, 1946 and Nataf, 1948), aggregation of consumer goods (e.g., that rests on the Hicksian and functional separability concepts), aggregation of consumer demands (Gorman, 1953), aggregation of production functions, etc. In general, one can distinguish two aggregation problems often studied in economic analysis: (1) the aggregation over individuals (e.g., firms, consumers) and (2) the aggregation over goods (e.g., inputs, outputs, commodities).

In the area of *Efficiency Analysis*, both problems of aggregation have been raised since the early stages of development of the area. Many fundamental questions of the aggregation analysis, however, were tackled just recently. The goal of the next subsection is to give more details of this evolution.

Aggregation over Firms in Efficiency Analysis

Perhaps, the first to talk about the *aggregation over firms* was Farrell (1957) himself, when he introduced the measures of technical and price (allocative) efficiencies that were later named after him. He was interested in a measure of efficiency of an industry. In his words,

> ... if economic planning is to concern itself with particular industries, it is important to know how far a given industry can be expected to increase its output by simply increasing its efficiency, without absorbing further resources. (p. 253.)

He particularly was motivated by a desire to create a measure that would enable economists to compare an efficiency of an industry to the same industry in another period or another country, or even to another industry. He called this measure the *Structural Efficiency of an Industry* and defined it as

> ... the technical efficiency of an industry with respect to a given efficient isoquant [that] would be simply a weighted average [weighted by output] of the [Farrell] technical efficiencies with respect to the same isoquant of its constituent firms. (p. 261.)

There were other measures defined in the spirit of Farrell (1957). For example, Carlson (1972) defines the "efficiency index for the industry" as the weighted (by the actual output) *arithmetic* average of the individual efficiency indices, where the latter are defined as the ratio of the actual to potential output for each individual decision making unit (DMU). Clearly, if the potential output is defined as the actual output multiplied by the Farrell technical efficiency score then the Farrell (1957) and the Carlson (1972) industry efficiency measures are equivalent. However, since the weights are the firms' *output* shares, both measures are applicable only for a single-output technology.

Analogously to Carlson (1972), Bjurek, Försund and Hjalmarsson (1990) introduced a measure of the "saving potential of the whole sector" defined as a weighted arithmetic average of the input saving measures, which are obtained, for each DMU, as the ratio of potentially minimal to actually used input. Since the weights for this measure were defined as the firms' input shares this measure is applicable only for a single-input technology.

A different perspective on the industry measure is given in Försund and Hjalmarsson (1974, 1979) who suggested

> "… to construct an [arithmetic] average plant for the industry and regard this average plant as an arbitrary observation on the same line as the other observations and then compute [technical and scale efficiency measures] for this average unit." (p. 300)

This latter measure attracted considerable criticism (e.g., see survey in Ylvinger, 2000) that essentially was around the fact that this 'average unit' type of industry efficiency measures may yield conclusions that the industry is not efficient even if all its units are technically efficient. On the other hand, a constructive approach to this measure is found in Li and Ng (1995), who were able to decompose (under some conditions) the 'average unit' measure of Försund and Hjalmarsson (1974, 1979) into the technical, allocative and reallocative efficiencies.

Another approach to measuring the structural efficiency of an industry is discussed in Ylvinger (2000). His goal was to determine the weights of aggregation, which he did by choosing the weights to be the shadow prices from the activity analysis models.

All of the mentioned measures were intuitive but essentially ad hoc: they were not derived from concepts of economic theory or some mathematical consistency criteria.

In fact, the controversies around the Försund and Hjalmarsson (1974, 1979) measure were essentially the challenges from the classical aggregation question: Is

there a relationship between a measure that uses aggregate data (e.g., industry average inputs and outputs) and the same measure used for each observation of the disaggregated data. Formally and on a general level, this question was first raised and answered by Blackorby and Russell (1999). Regrettably, they reached as they themselves called "discouraging" conclusion. In particular, they found that there does not exist an efficiency measure (satisfying the input or output homogeneity property) for which one can establish a relationship between the case when this measure uses the aggregate-over-firms data on inputs and outputs and when it uses the disaggregated data. Their results were a bit less discouraging for the case when only output or input is aggregated (over firms). Specifically, they found that

> "…very strong restrictions on the technology and/or the efficiency index itself [e.g., homogeneous and linear technology for aggregating the Debrue/Farrell efficiency measure] are required to enable consistent aggregation (or disaggregation)." (p. 5).

Three essays of this chapter are dedicated to finding more optimistic aggregation possibilities than the Blackorby and Russell (1999) results. Unlike the ad hoc measures, here an emphasis is given to both consistency and economic theory foundation for the aggregation.

Essay 4: Aggregation of the Nerlovian Profit Indicator

The chapter opens with Essay 4—a paper co-authored with Rolf Färe and Shawna Grosskopf, in which we consider the case of aggregation of inputs and outputs over firms. This is exactly the case where Blackorby and Russell (1999) reached their non-existence result (for any efficiency measure homogeneous of degree 1 in inputs or outputs). Our approach is different in two respects: (1) we consider a different type of efficiency measure—the additive efficiency measure based on directional distance functions, and (2) we aggregate over the *optimal* points only, and then use duality in economics to decompose the aggregate overall

efficiency measures into the technical efficiency and allocative efficiency components.

Specifically, we first define the *industry technology* as the sum of the individual technologies. Such a structure on the aggregate technology enables us to use the Koopmans (1957) theorem telling us that the industry maximal profit equals the sum of the maximal profits of all firms in the industry, which immediately yields a solution to aggregation of Nerlovian measures of profit efficiency (as defined in Chambers et al. (1998)).

Then, using the duality between the profit function and the directional distance function we decompose the aggregate profit efficiency into two components: (*i*) the sum of individual technical efficiencies, and (*ii*) the sum of individual allocative efficiencies. We then address the standard aggregation question: When does this sum of technical efficiencies equal the aggregate technical efficiency (i.e., the one based on the aggregate-over-firms data)? We show that the condition needed for this equality to hold is analogous to the Blackorby and Russell (1999) restriction on the technology. We also show that if this condition does not hold, then the difference between the two aggregate measures is always one way and bounded: the former is always smaller than the latter.

Overall, although for an 'exact' aggregation we still need quite restrictive technology, in essay 4 we were able to find a *practical way* of computing the aggregate technical efficiency—as the sum of the individual technical efficiencies—a way that is related to profit efficiency through the aggregate (sum of individual) allocative efficiencies.

Essay 5: On Aggregate Farrell Efficiencies

In Essay 5—a paper co-authored with Rolf Färe, we use a similar approach to aggregation, except that this time we aggregate the Farrell-type efficiencies. Specifically, we first define the *industry output set* as the sum of the individual

output sets. Such a structure on the aggregate technology enables us to derive a *revenue* analogue of the Koopmans (1957) theorem (used in the previous essay): Industry (group) maximal revenue equals the sum of the maximal revenues of all firms in the industry (group). This fact immediately yields a solution to the aggregation problem of revenue efficiency measures: The industry (group) efficiency is the sum of revenue efficiencies of all firms in the industry (group) weighted by the observed revenue shares of each firm.

Then, using the duality between the revenue function and the distance function we decompose the aggregate revenue efficiency into two components: (*i*) the sum of individual technical efficiencies *weighted* by actual revenue shares, and (*ii*) the sum of individual allocative efficiencies *weighted* by technically efficient revenue shares.

We then address the standard aggregation question: What is the relationship between this weighted sum of technical efficiencies and the technical efficiency based on the aggregate-over-firms data? In a single output case the answer is precise: they are equal. Moreover, in this case they both are equal to the Farrell (1957) "*Structural Efficiency of an Industry*" and to the industry revenue efficiency. In a multiple output case, however, since the former depends on prices while the latter does not, there is, in general, no relationship. Moreover, using a simple example, we show that either of them can be bigger than the other.

Overall, although there is in general no 'exact' aggregation, we were able to find a *practical* way of computing the aggregate technical efficiency—as the sum of the individual technical efficiencies weighted by the revenue shares—which is related to the revenue efficiency through the aggregate (weighted sum of individual) allocative efficiencies. In addition, we also find a way to go from the revenue shares weights to the price independent weights.

Essay 6: Averaging Farrell Scores

Essay 6—a paper co-authored with Rolf Färe—approaches aggregation problem from a purely mathematical standpoint. Specifically, we first postulate a consistency criterion on the *average* Farrell technical efficiency measure to preserve a multiplicative structure that exists on the disaggregate level and formalize this criterion as a functional equation. The solution to this equation becomes the weighted geometric average, and the objective of the paper becomes to find the appropriate weights.

We find two types of weights: (1) the revenue share weights and (2) the 'average output share' weights. The first set of weights is determined by using the result from essay 5, and then noting that the arithmetic average is the first order Taylor series approximation of the geometric average (around unity—the threshold level for efficiency measures). The second set of weights—price independent averages of firms' output shares—is derived from the first set by using the duality reflection as in Cornes (1992).

Aggregation over Goods in Efficiency Analysis

So far, the focus was on the aggregation *over* individuals (firms, countries, etc). Another course of aggregation is *over goods* (inputs or outputs). Its importance is often dictated by the data restrictions in empirical studies, where data is frequently available in a more aggregated form than theory operates with. A classical example would be the labor and material costs versus the physical amount of each type of labor and material used, respectively. If a researcher obtains efficiency results from using the aggregated data an appropriate question is: Would the results be different if the disaggregated data were used? How different (larger, smaller, etc)? How much different? In other words, will there be bias due to aggregation? What is the direction of the bias? How large would be the bias?

The problem of aggregation over commodities in efficiency analysis was tackled at least since Numamaker (1985) and Thrall (1989) who showed how the aggregation over inputs affects the computation of Farrell technical efficiency scores. Later, from a theoretical standpoint, Färe and Lovell (1988) derived a condition of unbiased aggregation over goods. Specifically, they conclude that the Farrell-type efficiency indices are invariant with respect to input (output) aggregation if and only if the cost (revenue) function is separable. Lovell, Sarkar and Sickles (1988) empirically illustrate the validity of this conclusion. The existence of the aggregation bias in empirical studies was also emphasized in Thomas and Tauer (1994). An extension to this paper is recently found in Tauer (2001), who by means of a simulated data example showed that this bias is different for different types of aggregation (exact, Divisia, linear) and, remarkably, that it *increases* as more and more inputs (outputs) are aggregated. This last paper has inspired Essay 7 of this chapter

Essay 7: Input Aggregation and Technical Efficiency

Here, in a paper co-authored with Rolf Färe, we ask a standard aggregation question: What are the conditions for an efficiency measure to have no aggregation bias? As a result, we first reformulate the Farrell technical efficiency measure as a subvector efficiency estimator and then derive *necessary and sufficient* condition under which the linear aggregator of inputs (and a similar result can be shown for aggregation of outputs) yields an unbiased outcome. This condition is interpreted as a situation with no allocative inefficiency in the subvector of inputs that is aggregated.

ESSAY 4: AGGREGATION OF NERLOVIAN PROFIT INDICATOR

Valentin Zelenyuk[1]

Abstract

In this note we show that the Nerlovian profit indicator may be aggregated over firms into an industry measure of profit efficiency. We also provide conditions under which the technical component of the indicator may also be aggregated.

---

[1] This paper is coauthored with Rolf Fare and Shawna Grosskopf. Valentin Zelenyuk is a primary author.

Introduction

The performance measure we consider here was introduced by Chambers, Chung and Färe (1998) and named after Nerlove (1965) who had introduced a related profit performance measure. What we are interested in are the conditions under which the firm Nerlovian profit indicators may be aggregated into an industry Nerlovian profit indicator. The first result follows directly from Koopmans (1957) who proved that the industry profit function is the sum of firm profit functions, which we show holds or the profit indicator as well. A more challenging task is involved when we wish to consider aggregation of the components of the Nerlovian profit indicator, namely allocative and technical efficiency. Although the results are not as straightforward we can develop bounds and conditions under which exact aggregation occurs.

We note that our success in aggregating stems in part from the fact that we adopt what we call directional distance functions as our measures of technical efficiency. The directional distance functions are closely related to what Luenberger (1992) calls benefit functions in the consumer context. The advantage of these for aggregation was noted by Luenberger (1992) 'The single normalization of the benefit function theory can be applied to all consumers, while the distance function approach requires that a given price vector be normalized differently for each consumer.' (p.480) In our context, the advantage of directional distance function is that we can choose one direction (and therefore one associated normalization) for the evaluation of each firm's efficiency; the Shephard type distance functions allow each firm to be evaluated in a different direction (namely that consistent with its input or output mix).

The Details

We define the industry technology T as the sum of the firm technologies, i.e.,

$$T = \sum_{k=1}^{K} T^k,$$
(1)

where $T^k = \{(x_k, y_k): \ input \ x^k \in \Re_+^N can \ produce \ output \ y^k\}$.

Koopmans (1957) proves that industry profit

$$\Pi(p, w) = \max_{x,y} \{py - wx : (x, y) \in T\}.$$
(2)

is the sum over firm profits, i.e.,

$$\Pi(p, w) = \sum_{k=1}^{K} \Pi^k(p, w),$$
(3)

where, given input and output prices *(w, p)*, firm *k's* profit is defined by

$$\Pi^k(p, w) = \max_{x,y} \left\{py^k - wx^k : (x^k, y^k) \in T^k \right\}.$$
(4)

By subtracting observed profit $\left(p\sum_{k=1}^{K} y^k - w\sum_{k=1}^{K} x^k\right)$ from both sides of (3) and normalizing with *(pg_y + wg_x)* where *(g_x, g_y)* is the direction in which technical efficiency is to be measured yields

$$\frac{\Pi(p,w)-\left(p\sum_{k=1}^{K}y^{k}-w\sum_{k=1}^{K}x^{k}\right)}{pg_{y}+wg_{x}}=\sum_{k=1}^{K}\frac{\Pi^{k}(p,w)-(py^{k}-wx^{k})}{pg_{y}+wg_{x}}. \quad (5)$$

The left hand side is the industry measure of profit efficiency which is equal to the right hand side which is the sum over firm profit efficiencies. These efficiency measures, introduced by Chambers, Chung and Färe (1998) are called Nerlovian measures of profit efficiency.

The Nerlovian profit efficiency index may be expressed as the sum of an allocative and a technical component. The technical component is defined for firm k as

$$\vec{D}^{k}(x^{k},y^{k};g_{x},g_{y})=\max\left\{\beta:(x^{k}-\beta g_{x},y^{k}+\beta g_{y})\in T^{k}\right\} \quad (6)$$

where (g_{x},g_{y}) is the directional vector. Luenberger (1995) calls this function (6) the shortage function. This function takes values greater than or equal to zero for feasible (x,y). As usual, the allocative efficiency index (AE^{k}) is defined as a residual, thus for firm k we have

$$\frac{\Pi^{k}(p,w)-(py^{k}-wx^{k})}{pg_{y}+wg_{x}}=AE^{k}+\vec{D}^{k}(x^{k},y^{k};g_{x},g_{y}) \quad (7)$$

The industry decomposition is similar, with its directional distance function defined on the industry technology T.

$$\frac{\Pi(p,w)-\left(p\sum_{k=1}^{K}y^{k}-w\sum_{k=1}^{K}x^{k}\right)}{pg_{y}+wg_{x}}=AE+\vec{D}\left(\sum_{k=1}^{K}x^{k},\sum_{k=1}^{K}y^{k};g_{x},g_{y}\right). \quad (8)$$

Now (5), (7) and (8) together yield

$$\vec{D}\left(\sum_{k=1}^{K}x^k, \sum_{k=1}^{K}y^k; g_x, g_y\right) = \sum_{k=1}^{K}\vec{D}^k\left(x^k, y^k; g_x, g_y\right) \qquad (9)$$

if and only if the industry allocative efficiency component is the sum of its firm components, i.e., $AE = \sum_{k=1}^{K}AE^k$.

Thus (5) and (9) show that we can derive aggregate industry efficiency from firm efficiencies. Of course, our result in (9) rests on the condition that $AE = \sum_{k=1}^{K}AE^k$, which may not be appropriate in some applications. If we relax that assumption, we can still derive the following relationship:

$$\vec{D}\left(\sum_{k=1}^{K}x^k, \sum_{k=1}^{K}y^k; g_x, g_y\right) \geq \sum_{k=1}^{K}\vec{D}^k\left(x^k, y^k; g_x, g_y\right) \qquad (10)$$

This result follows from the fact that

$$\sum_{k=1}^{K}\left(x^k - \vec{D}^k g_x, y^k + \vec{D}^k g_y\right) = \left(\sum_{k=1}^{K}x^k - \sum_{k=1}^{K}D^k g_x, \sum_{k=1}^{K}y^k + \sum_{k=1}^{K}D^k g_y\right) \in T \qquad (11)$$

and definition of the industry distance function. Thus, even if one is not willing to assume that $AE = \sum_{k=1}^{K}AE^k$, we have shown that the sum of the firm technical efficiency measures based on directional distance functions will never be greater than the corresponding industry technical efficiency measure. Note that (5), (8) and (10) yield

$$\sum_{k=1}^{K}AE^k \geq AE, \qquad (12)$$

as a general result, so if each firm is allocatively efficient, i.e., $AE^k = 0$ for all k, then since $AE \geq 0$, then the industry is allocatively efficient as well.[2]

Finally, if we assume that (9) holds, for all $x^k \in \Re_+^N$ and $y^k \in \Re_+^M$ and let $(g_x, g_y) = (1,1)$, then (9) is a Pexider functional equation in many variables. Its solution is found in Aczél (1966):

$$\vec{D}^k(x^k, y^k; 1,1) = \sum_{n=1}^{N} a_n x_{kn} + \sum_{m=1}^{M} b_m y_{km} + c_k \tag{13}$$

and

$$\vec{D}\left(\sum_{k=1}^{K} x^k, \sum_{k=1}^{K} y^k; 1,1\right) = \sum_{n=1}^{N} a_n \sum_{k=1}^{K} x_{kn} + \sum_{m=1}^{M} b_m \sum_{k=1}^{K} y_{km} + \sum_{k=1}^{K} c_k, \tag{14}$$

where a_n, b_m and c_k are arbitrary constants. This result is the directional distance function analog to the Blackorby and Russell (1999) aggregation result.

References

Aczél (1966), *Lectures on Functional Equations and Their Applications*, Academic Press, New York/London.

Blackorby, C. and R. R. Russell (1999), "Aggregation of Efficiency Indices", *Journal of Productivity Analysis* 12, 1, pp. 5-20.

Chambers, R., Y. Chung and R. Färe (1998), "Profit, Directional Distance Functions and Nerlovian Efficiency," *Journal of Optimization Theory and Application* 98:2, 351-364.

Koopmans, T.C. (1957), *Three Essays on the State of Economic Analysis*, McGraw-Hill, New York.

Luenberger, D.G. (1992), "New Optimality Principles for Economic Efficiency and Equilibrium," *Journal of Optimization Theory and Applications* 75, 221-264.

---

[2] The first proof of this fact was provided by Jesus Pastor.

ESSAY 5: ON AGGREGATE FARRELL EFFICIENCY SCORES

Valentin Zelenyuk[1]

Abstract

In this paper we establish the fact that an industry maximal revenue is the sum of its firms' maximal revenues. This fact enables us to discover conditions for aggregation of Farrell efficiencies.

Key words: Efficiency, distance function, aggregation, duality.

Introduction

In his seminal paper, Farrell (1957, p. 261) introduced a concept of the 'structural efficiency of an industry' by suggesting that "the technical efficiency of an industry with respect to a given efficient isoquant would be simply a weighted average[2] of the technical efficiencies with respect to the same isoquant of its constituent firms." His brief discussion of this concept and the final remark that "It is hoped to develop this argument further elsewhere" (Farrell, 1957, p. 262) has been given different interpretations in the literature on efficiency measurement. Ylvinger (2000) provides a survey of the topic and points out that some interpretations yield inconsistent measurements. An interesting discussion of industry efficiency is found in Førsund and Hjalmarsson (1979).

Two recent papers, Blackorby and Russell (1999) and Li and Ng (1995), have added to the understanding of the aggregate efficiency. Blackorby and Russell (1999) derive conditions on the firm technologies that are required to aggregate technical efficiency indexes. These conditions are quite stringent which is summarized by "... there does not exist a technology set such that the widely used Debreu (1951) / Farrell (1957) measure of technical efficiency can be aggregated..." Blackorby and Russell (1999, p. 7-8). Li and Ng (1995) circumvent the problem of Blackorby and Russell by introducing weights (the shadow output prices) in their aggregation.

In this paper we take a new approach and start by observing that maximized industry revenue equals the sum of maximized firms' revenues. We use this equality to derive the industry efficiency measure from the firms' measures of both technical and overall efficiencies. The resulting industry technical efficiency

---

[2] Weighted by output. (This is the original footnote of Farrell, 1957 p. 261).

measure is a multioutput generalization of the Farrell 'structural efficiency of an industry' measure.

Multiple-Output Measures of Efficiency

The Farrell (1957) efficiency framework consists of three components, a technical, an allocative and--depending on orientation--a cost or revenue component. The product of the first two makes up the last. In this section we first study the revenue component and show how the industry efficiency can be derived from the member firms' efficiency. We then turn our attention to the technical measure and again show how the industry measure is related to the firms' measures. The method we develop here can also be applied to the Farrell input oriented or cost approach, since one can prove that the industry minimum cost equals the sum of its firms' costs.

The firm technology is given by its output sets

$$(2.1) \qquad P^k(x^k) = \{y^k : x^k \text{ can produce } y^k\}, \quad x^k \in \Re^N_+,$$

where $x^k = (x_{k1}, ..., x_{kN}) \in \Re^N_+$ denote firm k's input vector and $y^k = (y_{k1}, ..., y_{kM}) \in \Re^M_+$ its output vector. We assume that there are $k = 1, ... , K$ firms in the industry, $K > 1$. (If $K=1$, then the industry consist of one firm and hence no aggregation is required.)

The industry technology is defined as

$$(2.2) \qquad \mathcal{P}(x^1, ..., x^K) = \sum_{k=1}^{K} P^k(x^k)$$

i.e., it is the sum of firm's technologies. Note that in this setting there is no reallocation of the inputs among the firms. We also note that the industry

technology $\mathcal{P}(x^1, ..., x^K)$ inherits its properties from those of the different firm technologies $P^k(x^k)$. Thus, if each $P^k(x^k)$ is a convex, compact set with inputs and outputs freely disposable, then so is $\mathcal{P}(x^1, ..., x^K)$. Note that each firm may have a different technology P^k and use different input vectors x^k.

Denote output prices by $p = (p_1, ..., p_M) \in \mathfrak{R}^M_+$, (which we assume are the same for all firms), then firm k's observed revenue is py^k and the industry revenue equals $p \sum_{k=1}^{K} y^k$. To define the industry revenue function and obtain an aggregation theorem, it is crucial that all firms face the same price vector, p. The firm's maximal revenue is defined as:

$$(2.3) \qquad R^k(x^k, p) = \max_y \{py \ : \ y \in P^k(x^k)\}$$

and its revenue efficiency is defined as the ratio of firm's k maximal revenue $R^k(x^k, p)$ to its observed revenue py^k, i.e., as

$$(2.4) \qquad R^k(x^k, p)/py^k.$$

The industry maximal revenue is

$$(2.5) \qquad \mathfrak{R}(x^1,..., x^K, p) = \max_y \{py \ : \ y \in \mathcal{P}(x^1, ..., x^K)\}$$

and industry revenue efficiency is defined as the ratio of industry maximal to industry observed revenue, i.e., as

$$(2.6) \qquad \mathfrak{R}(x^1,..., x^K, p)/p \sum_{k=1}^{K} y^k.$$

By their definitions, the firm and industry revenue efficiencies are bigger then or equal to one.

To understand how *(2.4)* and *(2.6)* are related we observe first that the industry maximal revenue *(2.5)* is the sum of the firms' maximal revenues (the proof is given in the appendix), i.e.,

(2.7) $$\Re(x^1,..., x^K, p) = \sum_{k=1}^{K} R^k(x^k,p).$$

Using *(2.7)* it follows that the industry overall output (or revenue) efficiency is the share weighted average of the firms' overall output efficiencies, i.e.,

(2.8) $$\frac{\Re(x^1,..., x^K, p)}{p\sum_{k=1}^{K} y^k} = \sum_{k=1}^{K} \frac{R^k(x^k,p)}{py^k} \cdot S^k$$

where the shares are defined by $S^k = \dfrac{py^k}{p\sum_{k=1}^{K} y^k}$. (The proof of *(2.8)* is given in the appendix.)

Blackorby and Russell (1999) introduced the concept of an aggregate indication axiom (which is a special case of the agreement property for aggregating functions in Aczél, 1990, p.24), which in our framework states:

$$\frac{\Re(x^1,..., x^K, p)}{p\sum_{k=1}^{K} y^k} = 1 \quad \text{if and only if} \quad \frac{R^k(x^k,p)}{py^k} = 1, \quad k= 1, ..., K,$$

i.e., the industry is efficient ($\frac{\Re(x^1,..., x^K, p)}{p\sum_{k=1}^{K} y^k} = 1$) if and only if each firm is efficient. We note that since our efficiency measures *(2.4)* and *(2.6)* are all larger or equal to one, our measure *(2.8)* meets the aggregate indication axiom.

To introduce the technical measures of efficiency we define the output distance functions on $P^k(x^k)$, $k = 1, ... , K$ and on $\mathcal{P}(x^1, ..., x^K)$ respectively as

$$(2.9) \qquad D_o^k(x^k, y^k) = \inf\{\theta^k : (y^k/\theta^k) \in P^k(x^k)\}$$

$$(2.10) \qquad \mathcal{D}_o(x^1, ..., x^K, y) = \inf\{\theta : (y/\theta) \in \mathcal{P}(x^1, ..., x^K)\}.$$

Following Färe, Grosskopf and Lovell (1985) we define the k's firm and the industry output oriented Farrell measures of technical efficiency as the reciprocals of *(2.9)* and *(2.10)*, respectively. In addition to these two measures, we define the share weighted output oriented industry technical efficiency as

$$(2.11) \qquad TE = \sum_{k=1}^{K} \frac{1}{D_o^k(x^k, y^k)} \cdot S^k.$$

This is a multioutput generalization of the Farrell single-output "structural efficiency of an industry", where instead of the output shares (Farrell, 1957, p. 261-262) we use the revenue shares. In some ways, *TE* is not a "good" measure of "technical" efficiency since it contains value information, and is not just a function of inputs and outputs. However, as a part of revenue measure it enters quite naturally, as we will see below.

If we formulate the Blackorby and Russell (1999) technical efficiency aggregate indication axiom as:

$$TE = 1 \quad \text{if and only if } \frac{1}{D_o^k(x^k, y^k)} = 1, \ k = 1, \ldots, K.$$

then clearly our measure *(2.11)* satisfies this condition. This follows from the fact

that $\frac{1}{D_o^k(x^k, y^k)} \geq 1$ for all feasible (x^k, y^k), i.e., $y^k \in P^k(x^k)$.

To compare the two measures of technical efficiency, (2.11) and the reciprocal of (2.10), note that in the single output (multiple input) case they are equivalent to the industry revenue efficiency measure (2.6), and precisely represent what Farrell (1957) called the Structural Efficiency of the Industry, i.e.,

$$TE \equiv \sum_{k=1}^{K} \frac{1}{D_o^k(x^k, y^k)} \cdot S^k = 1 / \mathcal{D}_o(x^1, \ldots, x^K, \sum_{k=1}^{K} y^k).$$

Such equivalence however, in general cannot be established for the multiple output

case. To see this, first note that $\sum_{k=1}^{K} \frac{y^k}{D_o^k(x^k, y^k)}$ belongs to the industry technology

$\mathcal{P}(x^1, \ldots, x^K)$, and if outputs are freely disposable, $\sum_{k=1}^{K} y^k / D_o^M$, where

$D_o^M = \max_k \{D_o^k(x^k, y^k)\}$, is also in $\mathcal{P}(x^1, \ldots, x^K)$. Thus, it follows that

$$(2.12) \qquad \mathcal{D}_o(x^1, \ldots, x^K, \sum_{k=1}^{K} y^k) \leq D_o^M.$$

i.e., the industry technical efficiency score $1/\mathcal{D}_o(x^1, \ldots, x^K, \sum_{k=1}^{K} y^k)$ is at least as large as $1 / D_o^M$. In words, the industry is at least as inefficient as the most efficient firm. Next, by an example we show that the following inequalities may hold.

(2.13) $\qquad$ $1 / D_o^M \ < \ 1 / \mathcal{D}_o(x^1, \ ..., \ x^K, \ \Sigma_{k=1}^K y^k \ )$.

(2.14) $\qquad$ $TE \ < \ 1 / \mathcal{D}_o(x^1, \ ..., \ x^K, \ \Sigma_{k=1}^K y^k \ )$.

Expression *(2.13)* tells us that firm's maximal efficiency score can be smaller than the industry efficiency score. Expression *(2.14)* shows that the share-weighted industry efficiency score *TE* may be smaller than the industry output oriented Farrell measure of technical efficiency $1 \ / \ \mathcal{D}_o(x^1, \ ..., \ x^K, \ \Sigma_{k=1}^K y^k \ )$. Figure 3 illustrates our cases.
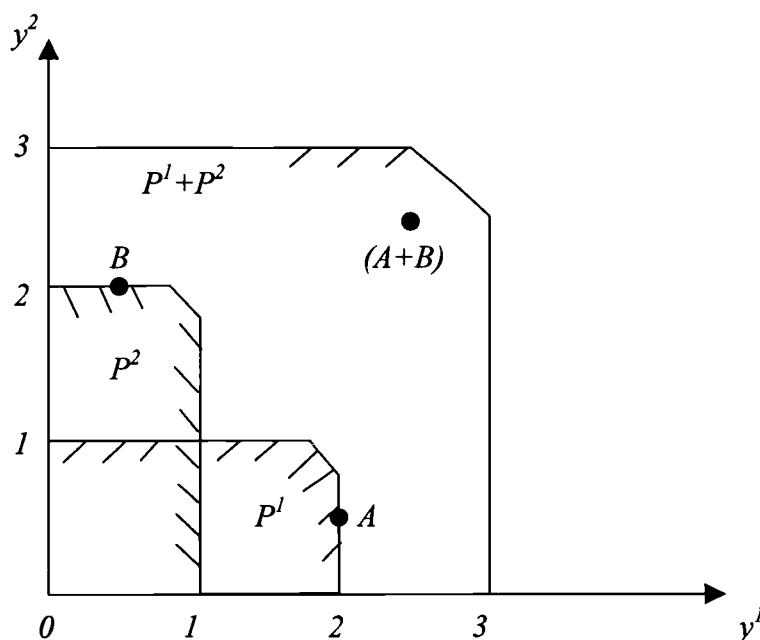
Figure 3. Measures of technical efficiency.

Output vector *A = (2, 1/2)* belongs to the output set P^1 and *B = (1/2, 2)* belongs to output set P^2. Both are efficient, so *TE = 1*, and $D_o^M = 1$. The aggregate

technology $(P^1 + P^2)$ contains $A+B = (2.5, 2.5)$ as an interior point, thus $1 / \mathcal{D}_o >$ 1, showing that *(2.13)* and *(2.14)* may hold.

By another example, we show that the inequality *(2.14)* may be reversed, i.e.,

$$(2.15) \qquad 1 / \mathcal{D}_o(x^1, ..., x^K, \textstyle\sum_{k=1}^K y^k) \; < \; TE,$$

may hold.

On the figure, if we take $A = (1,0)$ and $B = (0,2)$, then A belongs to P^1 and B belongs to P^2, with $D_o^A = 1/2$ and $D_o^B = 1$ (using (2.9)). If prices are $p = (10, 1)$ then $S^A = 10/12$, $S^B = 2/12$, and $TE = 22/12$. On the other hand, $A+B = (1, 2)$ and its efficiency score using *(2.10)* is $\mathcal{D}_o = 2/3$, showing that *(2.15)* holds.

Thus, equivalence between (2.11) and the reciprocal of (2.10) in general cannot be established for the multiple output case.

Let us now decompose the industry revenue efficiency into industry technical and industry allocative efficiency components. Following Li and Ng (1995), we define the aggregate measure of allocative efficiency as

$$(2.16) \qquad AE = \sum_{k=1}^K AE^k \cdot \hat{S}^k,$$

where

$$(2.17) \qquad AE^k = D_o^k(x^k, y^k) \frac{R^k(x^k, \; p)}{py^k},$$

and the weights are

$$(2.18) \qquad \hat{S}^k = \frac{p\left(y^k / D_o^k(x^k, y^k)\right)}{p\sum_{k=1}^{K}\left(y^k / D_o^k(x^k, y^k)\right)}.$$

i.e., the weights are based on potential outputs $\left(y^k / D_o^k(x^k, y^k)\right)$ rather than observed outputs y^k.

It now follows that the aggregate revenue efficiency can be decomposed into aggregate allocative efficiency AE and aggregate technical efficiency TE.

$$(2.19) \qquad \frac{\Re(x^1, ..., x^K, p)}{p\sum_{k=1}^{K} y^k} = AE \cdot TE.$$

To verify that *(2.19)* holds, insert *(2.11)*, *(2.16)*, *(2.17)* and *(2.18)* into *(2.19)*, use the fact *(2.7)* and the result follows.

We pointed out above that the shares S^k used in our definition of the output oriented industry technical efficiency measure is price dependent. But in the case of a single output it becomes price independent since

$$(2.20) \qquad S^k = \frac{py^k}{p\sum_{k=1}^{K} y^k} = \frac{y^k}{\sum_{k=1}^{K} y^k}, \qquad k = 1, ..., K.$$

If we want to create multioutput price independent share-weights, we may follow Cornes (1992, p.42) and choose the prices as

$$(2.21) \qquad p_m = \frac{1}{\sum_{k=1}^{K} y_{km}}, \qquad\qquad m = 1, ..., M$$

since then

$$(2.22) \qquad w^k = \frac{1}{M}\left(\frac{y_{k1}}{\sum_{k=1}^{K} y_{k1}} + \frac{y_{k2}}{\sum_{k=1}^{K} y_{k2}} + \dots + \frac{y_{kM}}{\sum_{k=1}^{K} y_{kM}} \right)$$

$$= \frac{1}{M}\left(\sum_{m=1}^{M} \frac{y_{km}}{\sum_{k=1}^{K} y_{km}} \right), \qquad k = 1, \dots, K.$$

The resulting price independent weights are the sum of each firm's share of each output normalized by the number of outputs M. They are non-negative and sum to one. Similarly, the price independent weights for aggregation of individual allocative efficiencies are

$$(2.23) \qquad \hat{w}^k = \frac{1}{M}\left(\sum_{m=1}^{M} \frac{y_{km}/D_o^k(x^k, y^k)}{\sum_{k=1}^{K} y_{km}/D_o^k(x^k, y^k)} \right), \qquad k = 1, \dots, K.$$

In our illustrative example in the next section we compare these weights with the price dependent share-weights.

A Numerical Illustration

In this section we introduce a numerical example and show how the industry efficiency may be computed from the firms' efficiencies. We assume there are twenty firms $k = 1, \dots, K$, each using two inputs (x_1, x_2) to produce two outputs (y_1, y_2). The output prices we use are $p_1 = 1$ and $p_2 = 0.1$. We compute the revenue efficiency (2.4) for each firm, and the industry efficiency (2.6).[3] We also compute the technical and allocative efficiency components for firms and the industry. We report our four sets of weights, the price dependent S^k and $\hat{S}^k$ and the price independent w^k and $\hat{w}^k$. Finally, as a comparison we included the non-weighted arithmetic average of the efficiency scores.

---

[3] The computation of efficiency scores are done on OnFront using variable returns to scale.

Table 1. Revenue, Technical and Allocative Efficiency for Firms and Industry: A Hypothetical Example

Firms	x_1^k	x_2^k	y_1^k	y_2^k	Revenue Efficiency	Technical Efficiency	Allocative Efficiency	S^k	$\hat{S}^k$	w^k	$\hat{w}^k$
1	39.00	49.00	12.00	17.53	2.200	1.772	1.242	0.008	0.013	0.008	0.014
2	37.00	45.00	19.00	22.00	1.000	1.000	1.000	0.013	0.012	0.012	0.011
3	35.00	55.00	17.29	17.00	1.929	1.891	1.020	0.012	0.020	0.010	0.018
4	34.00	63.97	25.00	12.97	1.957	1.884	1.039	0.016	0.027	0.012	0.020
5	33.00	53.00	28.00	18.72	1.000	1.000	1.000	0.018	0.017	0.014	0.013
6	70.00	50.00	35.00	43.00	1.516	1.476	1.027	0.024	0.032	0.023	0.031
7	45.00	55.56	25.00	0.00	1.915	1.712	1.119	0.015	0.024	0.009	0.013
8	60.00	62.38	45.00	37.42	1.540	1.477	1.043	0.030	0.040	0.025	0.034
9	30.00	83.33	75.00	64.03	1.000	1.000	1.000	0.049	0.045	0.042	0.038
10	40.00	90.00	34.00	59.27	2.279	1.397	1.631	0.024	0.031	0.026	0.034
11	75.00	75.00	82.00	75.00	1.107	1.072	1.033	0.054	0.053	0.047	0.046
12	45.00	125.00	78.00	101.70	1.086	1.000	1.086	0.054	0.049	0.052	0.048
13	60.00	93.75	35.00	93.54	2.417	1.256	1.924	0.027	0.031	0.035	0.041
14	87.00	53.57	75.00	120.00	1.000	1.000	1.000	0.053	0.048	0.056	0.051
15	85.00	66.18	85.00	111.00	1.009	1.000	1.009	0.058	0.053	0.057	0.052
16	91.00	99.00	100.00	171.00	1.102	1.000	1.102	0.071	0.065	0.077	0.070
17	115.20	169.00	115.00	212.00	1.198	1.000	1.198	0.083	0.075	0.092	0.085
18	86.40	240.00	80.00	151.00	1.425	1.104	1.291	0.058	0.058	0.065	0.066
19	247.00	189.00	230.00	347.00	1.069	1.035	1.033	0.161	0.151	0.165	0.157
20	240.00	180.00	247.00	359.00	1.000	1.000	1.000	0.172	0.156	0.174	0.160
Non-weighted Arithmetic Average					1.437	1.254	1.140				
Industry Efficiency					1.218	1.099	1.108				
Industry Efficiency with price independent weights					1.223	1.090	1.122				

Note: Price for y_1 is normalized to 1, and price for y_2 is set to 0.1.

References

Aczél, J. (1990), "Determining Merged Relative Scores" *Journal of MathematicalAnalysis and Applications* 150, July 15, 1990.

Blackorby, C. and R. R. Russell (1999), "Aggregation of Efficiency Indices",*Journal of Productivity Analysis* 12, 1, pp. 5-20.

Cornes, R. (1992) *Duality and Modern Economics*, Cambridge: CambridgeUniversity Press.

Debreu, G. (1951), "The coefficient of resource utilization," *Econometrica* 19, pp.273-292.

Färe, R., S. Grosskopf and J. Kirkeley (2001), "Industry Efficiency and Capacity,"*(in progress)*.

Färe, R., S. Grosskopf and C.A.K. Lovell (1985), *The Measurement of Efficiency of Production*, Kluwer Nijhoff Publishing, Boston.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, General 120, part 3, 253-281.

Førsund, F. and L. Hjalmarsson (1979), "Generalized Farrell measures of efficiency: an application to milk processing in Swedish dairy plants," *Economic Journal* 89, pp. 294-315.

Koopmans, T.C. (1957), *Three Essays on the State of Economic Analysis*, McGraw-Hill, New York.

Li, S.K., and Y.C. Ng (1995), "Measuring the Productive Efficiency of a Group of Firms," *International Advances in Economic Research* 1, 4 (November): 377-90.

Mas-Colell, A., M. Whinston and J. Green, (1995) *Microeconomic Theory*, Oxford University Press, Oxford.

Ylvinger S., (2000) "Industry performance and structural efficiency measures: Solutions to problems in firm models," *European Journal Of Operational Research* 121, 1, 164-174.

Appendix

Proof of (2.7). This proof is from Färe, Grosskopf and Kirkley (2001) and follows the outline of Mas-Colell et al (1995).

Let $y^k \in P^k(x^k)$ be arbitrary, then $\sum_{k=1}^{K} y^k \in \mathcal{P}$ $(x^1, ..., x^K)$ and since $\Re(x^1, ..., x^K, p)$ is the maximal revenue,

$$\Re(x^1, ..., x^K, p) \geq p\sum_{k=1}^{K} y^k = \sum_{k=1}^{K} py^k .$$

Now, since y^k $(k = 1, ..., K)$ is arbitrary, we have

(i) $\qquad \Re(x^1, ..., x^K, p) \geq \sum_{k=1}^{K} R^k(x^k, p).$

Conversely, let $y \in \mathcal{P}$ $(x^1, ..., x^K)$ be arbitrary, then by the definition of $\mathcal{P}$ $(x^1, ..., x^K)$ there are $y^k \in P^k(x^k)$ so that $y = \sum_{k=1}^{K} y^k$. Hence, $py = p\sum_{k=1}^{K} y^k = \sum_{k=1}^{K} py^k \leq \sum_{k=1}^{K} R^k(x^k, p)$, and by the arbitrariness of y it follows that

(ii) $\qquad \Re(x^1, ..., x^K, p) \leq \sum_{k=1}^{K} R^k(x^k, p)$

From inequalities (i) and (ii) we get

$$\Re(x^1, ..., x^K, p) \geq \sum_{k=1}^{K} R^k(x^k, p) \geq \Re(x^1, ..., x^K, p),$$

and hence

$$\Re(x^1,..., \ x^K, \ p) = \sum_{k=1}^{K} R^k(x^k,p),$$

proving our claim.

Proof of (2.8). To verify *(2.8)* requires the following steps:

(i) $\Re(x^1,..., \ x^K, \ p) = \sum_{k=1}^{K} R^k(x^k,p)$

(ii) multiply and divide the r.h.s. by py^k, then

(iii) divide both sides by $p\sum_{k=1}^{K} y^k$ and *(2.8)* follows.

ESSAY 6: INPUT AGGREGATION AND TECHNICAL EFFICIENCY

Valentin Zelenyuk[1]

Abstract

In this paper we define the notion of unbiased aggregation of inputs and provide a necessary and sufficient condition for this to apply.

---

[1] This paper is coauthored with Rolf Färe. Valentin Zelenyuk is a primary author.

In a recent article in this journal L. Tauer (2001, p. 295) wrote: "Using data simulated from a random production function it is shown that technical efficiency estimates computed by Data Envelopment Analysis are biased even if the exact aggregator function is used to aggregate inputs."

By formulating the technical efficiency estimator as a subvector estimator we can derive a necessary and sufficient condition under which the linear aggregator of inputs yields an unbiased outcome. Hereby we have an explanation to why a linear aggregation of inputs may introduce bias into the estimation of technical efficiency scores.

Suppose there are $k = 1, \ldots, K$ observations of inputs $x^k = (x_{k1}, \ldots, x_{kN}) \in \Re^N_+$ and outputs $y^k = (y_{k1}, \ldots, y_{kM}) \in \Re^M_+$ and their corresponding input prices w^k. We assume that $w^k = w \in \Re^N_+$ for all k, i.e., each firm k faces the same input prices.

Like in Tauer (2001) we assume that a subvector of inputs is aggregated using prices, i.e.,

$$(1) \qquad c_{k\hat{N}} = \sum_{n=1}^{\hat{N}} w_n x_{kn} \quad , \qquad k = 1, \ldots, K, \quad \text{and} \quad \hat{N} \leq N.$$

To define what we understand by an unbiased outcome of input aggregation, define

$$(2) \qquad SF_i(y^{k'}, x^{k'}) = \min \lambda$$

$$\text{s.t.} \qquad \sum_{k=1}^{K} z_k y_{km} \geq y_{k'm} \quad , \quad m = 1, \ldots, M,$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq \lambda x_{k'n} \quad , \quad n = 1, ..., \hat{N},$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{k'n} \quad , \quad n = \hat{N}+1, ..., N,$$

$$z_k \geq 0 \, , \quad k = 1, ..., K,$$

and

(3) $\qquad K(y^{k'}, \, c_{k'\hat{N}}, \, x_{k'\hat{N}+1}, ..., x_{k'N}) = \min \lambda$

$$\text{s.t.} \quad \sum_{k=1}^{K} z_k y_{km} \geq y_{k'm} \quad , \quad m = 1, ..., M,$$

$$\sum_{k=1}^{K} z_k c_{k\hat{N}} \leq \lambda c_{k'\hat{N}} \quad ,$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{k'n} \quad , \quad n = \hat{N}+1, ..., N,$$

$$z_k \geq 0 \, , \quad k = 1, ..., K,$$

The first problem is the subvector input oriented Farrell (1957) measure of technical efficiency (see Färe, Grosskopf and Lovell (1994)). The second problem is the measure of technical efficiency, when some inputs are aggregated as in (1).

Aggregation is unbiased if and only if

(4) $\qquad SF_i \, (y^{k'}, x^{k'}) \; = \; K(y^{k'}, \, c_{k'\hat{N}}, \, x_{k'\hat{N}+1}, ..., x_{k'N}).$

To derive conditions for (4) to hold define the subvector cost function

(5) $\qquad C(y^{k'}, w_1, ..., w_{\hat{N}}, \, x_{k'\hat{N}+1}, ..., x_{k'N}) = \min \sum_{n=1}^{\hat{N}} w_n x_n$

$$\text{s.t.} \quad \sum_{k=1}^{K} z_k y_{km} \geq y_{k'm} \quad , \quad m = 1, \dots, M,$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq x_n \quad , \quad n = 1, \dots, \hat{N},$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{k'n} \quad , \quad n = \hat{N} + 1, \dots, N,$$

$$z_k \geq 0 \quad , \quad k = 1, \dots, K,$$

From (5) we have a subvector cost index of efficiency as

$$(6) \qquad \frac{C(y^{k'}, w_1, \dots, w_{\hat{N}}, \, x_{k'\hat{N}+1}, \dots, x_{k'N})}{c_{k'\hat{N}}} = SF_i \, (y^{k'}, x^{k'}) \times SAE_i$$

where SAE_i is the subvector input allocative efficiency component.

Like in Färe and Grosskopf (1985), if $w_n > 0$, $n = 1, \dots, \hat{N}$, then

$$(7) \qquad \frac{C(y^{k'}, w_1, \dots, w_{\hat{N}}, \, x_{k'\hat{N}+1}, \dots, x_{k'N})}{c_{k'\hat{N}}} = K(y^{k'}, \, c_{k'\hat{N}}, \, x_{k'\hat{N}+1}, \dots, x_{k'N})$$

Thus (4) holds if and only if $SAE_i = 1$, i.e., if and only if there is no allocative inefficiency.

Hence, if the information on some inputs is available only in the aggregated form as in (1), then the DEA technique in (3) will yield unbiased efficiency scores if and only if there is no allocative inefficiency in the subvector of inputs that is aggregated.

References

Färe, R. and S. Grosskopf (1985), "A Nonparametric Cost Approach to Scale Efficiency," *Scandinavian Journal of Economics* 87 (4), pp. 594-604.

Färe, R., S. Grosskopf and C.A.K. Lovell (1994), *Production Frontier,* Cambridge, Cambridge University Press.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, General, 120, part 3, 253-281.

Tauer, L. (2001), "Input aggregation and computed technical efficiency," *Applied Economics Letters* 8, pp. 295-297.

ESSAY 7: AVERAGING FARRELL SCORES

Valentin Zelenyuk[1]

Abstract

In this paper we develop a method for choosing weights for aggregating Farrell scores.

Key words: Efficiency, aggregation, duality.

Introduction

Individual Farrell scores may be decomposed into sub-scores. For example, Färe, Grosskopf and Lovell (1994) multiplicatively separate the Farrell output oriented measure of technical efficiency into three components, scale, congestion and "pure" efficiencies. To preserve this decomposition in a multi-firm industry it is necessary and sufficient that a weighted geometric mean is used. In this paper we show how these weights are determined.

The Results

Let r_k and s_k ($k=1,\ 2$) be firm k's two component measures of efficiency and let their product $q_k = r_k s_k$ be the Farrell output measure of technical efficiency. Suppose we want to aggregate these measures into an industry measure while preserving the multiplicative structure. This results in the following functional equation

$$V(q_1, q_2) = V(r_1, r_2)\ V(s_1, s_2) \tag{1}$$

Let us generalize this equation by introducing a set of parameters $z = (z_1, ..., z_J) \in \Re^J$, i.e.,

$$U(q_1, q_2; z) = U(r_1, r_2; z)\ U(s_1, s_2; z) \tag{2}$$

The solution to this equation is (see Aczél, 1990, p.27 and Eichhorn 1978, p.94)

$$U(q_1, q_2; z) = q_1^{w_1(z)} q_2^{w_2(z)} \tag{3}$$

where $w_k(z)$, are arbitrary functions of z.

The purpose of this paper is to determine the weights $w_1(z)$ and $w_2(z)$.

Define the industry output set as

$$\overline{P}(x^1, x^2) = P^1(x^1) + P^2(x^2) \tag{4}$$

where x^1 and x^2 are input vectors for each firm and where $P^1(x^1)$ and $P^2(x^2)$ are the firms output sets. We assume that each firm produces (for simplicity) two outputs $y^1 = (y_{11}, y_{12})$ and $y^2 = (y_{21}, y_{22})$, respectively.

The industry and firm's revenue functions are given by

$$\overline{R}(x^1, x^2, p) = \max\{p_1 y_1 + p_2 y_2 \; : \; (y_1, y_2) \in \overline{P}(x^1, x^2)\} \tag{5}$$

and

$$R^k(x^k, p) = \max\{p_1 y_{k1} + p_2 y_{k2} \; : \; (y_{k1}, y_{k2}) \in P^k(x^k)\} \tag{6}$$

It is known, see e.g. Färe and Zelenyuk (2001), that

$$\overline{R}(x^1, x^2, p) = R^1(x^1, p) + R^2(x^2, p) \tag{7}$$

where $p = (p_1, p_2)$ is the vector of output prices.

From (7) it follows that the industry revenue efficiency is the share weighted average of the firms' efficiencies, i.e.,

$$\frac{\overline{R}(x^1,x^2,p)}{p_1(y_{11}+y_{21})+p_2(y_{12}+y_{22})} = \frac{R^1(x^1,p)}{p_1y_{11}+p_2y_{12}}S^1 + \frac{R^2(x^2,p)}{p_1y_{21}+p_2y_{22}}S^2, \qquad (8)$$

where $S^k = \dfrac{p_1y_{k1}+p_2y_{k2}}{p_1(y_{11}+y_{21})+p_2(y_{12}+y_{22})}$, $k = 1, 2.$ $\qquad (9)$

If we assume that firms are allocative efficient, then (8) becomes

$$q = q_1 S^1 + q_2 S^2, \qquad (10)$$

where q is the industry technical efficiency score.

Approximate (3) around $q_1 = q_2 = 1$, then

$$\tilde{U}(q_1,q_2;z) = w_1(z)q_1 + w_2(z)q_2 \qquad (11)$$

Thus by taking $q = \tilde{U}(q_1,q_2;z)$ we find that

$$w_1(z) = S^1, \quad w_2(z) = S^2. \qquad (12)$$

To make these weights price independent we first note that in the case of a single output

$$w_1(z) = \frac{y_{11}}{y_{11} + y_{21}}, \quad w_2(z) = \frac{y_{21}}{y_{11} + y_{21}} \tag{13}$$

where $z = (y_{11}, y_{21})$.

Hence in the single output case we have shown that the weights should be the output shares.

In the multi output case we rely on duality theory and note that the industry normalized revenue function is

$$p_1(y_{11} + y_{21}) + p_2(y_{12} + y_{22}) = 1 \tag{14}$$

To transform the prices into outputs we follow Cornes (1992, p.42) and choose

$$p_i = \frac{1}{(y_{1i} + y_{2i})}, \quad i = 1, \ 2.$$

If we insert these expressions in (9) and (12), we obtain the following *price independent* weights

$$w_1(z) = \frac{1}{2}\left(\frac{y_{11}}{y_{11} + y_{21}} + \frac{y_{12}}{y_{12} + y_{22}} \right), \quad w_2(z) = \frac{1}{2}\left(\frac{y_{21}}{y_{11} + y_{21}} + \frac{y_{22}}{y_{12} + y_{22}} \right)$$

where $z = (y_{11}, y_{12}, y_{21}, y_{22})$.

These weights sum to one and are homogeneous of degree zero, and hence are independent of the unit of measurement. They are the non-weighted average of output quantity shares.

The obvious observations are that our method of finding weights generalizes to any $k = 1, ..., K$ and $i = 1, ... , I$. Also, of course, for the input oriented case the corresponding weights are the non-weighted average of input quantity shares.

References

Aczél, J. (1990), "Determining Merged Relative Scores," *Journal of Mathematical Analysis and Applications*,150:1.

Cornes, R. (1992), *Duality and Modern Economics*, Cambridge University Press, Cambridge.

Eichhorn, W. (1978), *Functional Equations in Economics*, Addison-Wesley, Reading, MA.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, General, 120:3.

Färe, R. and V. Zelenyuk (2001), "On Aggregate Farrell Efficiencies," Department of Economics, Oregon State University.

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

EMPIRICAL APPLICATION OF EFFICIENCY AND PRODUCTIVITY ANALYSIS TO INDUSTRIAL ORGANIZATION:

ESSAY 8: CAUSES OF CONCENTRATION IN THE U.S. BREWING INDUSTRY: RECONCILING THE DEBATE WITH THE DATA ENVELOPMENT ANALYSIS

Valentin Zelenyuk

Abstract

For almost 40 years, industrial organization economists have debated whether or not cost or demand side factors are a more important cause of rising concentration in US brewing industry. In this study data envelopment analysis is applied to a panel of 22 firms form 1950 to 1985 to test cost-side forces. I find that only small firms operated in the region of economies of scale. Large (*national*) firms who grew rapidly in size and are responsible for most of the rise in concentration were larger than needed to take advantage of all economies of scale. Altogether, I find the cost-side justification for the rise in concentration to be inconsistent with the data. Upon combining the key information about the industry from this and other studies, I find these results consistent with economic theory and the demand-side argument.

1. Introduction

There has been a tremendous structural change in the U.S. Brewing industry from 1950-1990. The average size of a typical firm has increased about 14 times.[1] While the total output of the industry was growing, the number of firms in the industry was decreasing dramatically, with the greatest change in the period from 1950 to the end of 1970s. For example, in 1950, the industry consisted of 369 firms while in 1977 it declined to 49 and then to 26 firms by 1998. These striking changes were directly reflected into the 'four firm concentration ratio' of the industry, which rose form about 21% in 1950 to about 62% in 1977 and about 95% in 1996.

Previous research indicates two possible explanations for the increase in concentration. One argument is the escalation of *large economies of scale* in the late 1950s through 1970s. This argument suggests that large firms by growing larger were becoming more efficient than their smaller competitors. This eventually put smaller firms out of business.

An alternative explanation is that successful marketing (primarily product differentiation and advertising) campaigns helped some firms increase their market shares at the expense of other firms, some of which had to leave the market. Chronologically, the debate goes back at least to Horowitz and Horowitz (1965) who found no support for the economies of scale argument but did find some support for the demand-side argument. Two years later, applying a different technique Horowitz and Horowitz (1967) found that the scale economies are more profound than in their previous study, but still not as important as the demand forces. Greer (1971) pursued the same research question and concluded that the main reason for concentration in brewing was the escalating product differentiation. Ten years later, after numerous opposite views were expressed in the academic and

---

[1] Average size of a firm is computed as total output of the industry divided by the number of firms in the industry.

government literature, he compromised on allowing some "room for important contribution from economies of scale" but still placed a "greater weight to product differentiation than other analysts might think appropriate" Greer (1981, p. 89).

On the other hand, Elzinga (1973, 1977) provided evidence that supports the economies of scale argument. Based on the "survivor test" he concluded that economies of scale were the most important reason, with the product differentiation playing the secondary and complementary role. Scherer et al. (1975) opinion was a compromise between Greer and Elzinga, giving some favor to the latter author. Elzinga views were also supported by other economists, including those on the government side, for example by Keiththahn (1978) and Mueller (1978), who give even greater weight to the economies of scale as the primal reason of concentration.[2]

An interesting stimulus to the debate, was the lack of solid and consistent estimates of scale economies (e.g., see critique by Greer, 1981, p. 90). Existence of such estimates might have given some answers and perhaps stopped the debate. Exactly this motivation seemed to attract applied econometric analysts to the issue.

Lynk (1984) found empirical support for the cost side reasoning. By observing correlation between increasing concentration and increases in output along with price decrease, he concluded that the change in industry structure was due to "competitive expansion by the more efficient brewers [and not due to] anticompetitive exclusion of equally efficient but smaller brewers."[3] This reasoning

---

[2] Specifically, Mueller advocated that the scale economies played the central role up to 1970, then yielding to the influence of new marketing strategies launched by conglomerate of Philip Morris and Miller Brewing Company.

[3] Although he does not specify what kind efficiency measure was used (technical, cost, scale, etc.) I classify his arguments as those supporting the cost side reasoning for concentration. Note, that although it may be tempting to conclude that the correlation between increases in concentration with increases in output and decreases in price is likely to be caused by expansion of more efficient firms, it is not necessarily true. Same phenomenon can arise from expansion of firms that are inefficient on the cost or production side but successful on the demand or marketing side. That is, the same correlation can be observed with the demand side causes playing the main role. I elaborate on this argument in the last section of this paper.

convinced him that "concentration in brewing had been beneficial, rather than harmful, to consumers" (Lynk (1984, p. 45)).

Tremblay and Tremblay (1985, 1987) undertook econometric study of the industry demand and cost structures (respectively). By estimating the *translog* 'average' cost function, they concluded that firms' growth was mostly due to "superior marketing position", thus supporting the demand side argument. They, however, also agreed on existence of some scale economies in the industry. Unexpectedly, they also finds that large firms have had significantly greater unit costs than smaller firms had, which seem to undermine the cost side reasoning for the increase in concentration. [4]

As a logical follow up in this debate, in the present study I question the cost side reasoning of concentration. My *a priori* expectations are based on the following logic. *If economies of scale played an important role in the rise of concentration, then one must observe large firms (i.e., those who influenced concentration) to be on the decreasing portion of the average cost curve, during (at least some of) the periods of increase in concentration.* Otherwise, these firms were experiencing diseconomies rather than economies of scale, and there must be some other reasons for the increase in concentration (which may or may not be the demand side reasons).

An important question in my analysis is the choice of the *scale economies* measure. A relatively small number of observations per period encouraged me to use somewhat modern approach to measure the scale economies—via the estimation of *scale efficiencies* using the non-parametric non-stochastic efficiency measurement (Farrell, 1957, Färe and Lovell, 1978), also known as data envelopment analysis or DEA (Charnes et al., 1978). Loosely speaking, the measure of scale efficiency is a *relative* indicator of how far a firm is from the

[4] He conjectures that this is because larger ("national") producers operated on a higher (rather than lower, as expected) cost function than smaller ("regional") producers. An alternative could be that they operated far on the *increasing* portion of the *same* average cost function (a conclusion obtained in the current study). Both results may depend on the empirical specification of the cost function. A conclusion that the larger producers have had greater cost, however, seems to be unambiguous.

'best-practice' industry frontier associated with the minimum efficient scale, and after one accounts for possible technical inefficiency of this firm. This way of measuring the scale economies has some advantages and disadvantages, which I will discuss in some detail in subsequent sections.

An important characteristic of present analysis is a dissection of industry into two conceptually different groups. A review of previous studies of the US brewing industry encouraged me to take into account the existence of *strategic groups*. Following Peles (1971), Hatten and Schendel (1977), and Tremblay (1985b, 1987), in this study I distinguish between *national* and *regional* producers.

The conclusions of the study are interesting and somewhat unexpected. I found that the *national* producers constantly operated at substantial *diseconomies* rather than economies of scale. On the other hand, the *regional* producers were operating close to the *minimum efficient scale* level, although in some cases some of them were not on the best-practice cost frontier.

Overall, the application of data envelopment analysis to the industry data leads me to reject the cost-side reasoning hypothesis that the primary reason for rising concentration was economies of scale. On the contrary, the firms that significantly influenced the rise in concentration were generally oversized, but still were growing despite experiencing diseconomies of scale. These results suggest that other reasons must be tested, including the hypothesis suggested by Sutton (1991), i.e., the hybrid of the demand side and cost side arguments.

The paper is organized as follows. In the next section, I outline the methodology, its advantages and drawbacks. Then, I discuss the computational results, their empirical implications, and present the overall understanding of the studying phenomenon.

2. Methodological Background

2.1. The Goals of Measurement

The goal of this paper is to test whether or not economies of scale were the main or significant cause of rising concentration in the US brewing industry. In effect, this will be done by analyzing whether or not large (national) and smaller (regional) firms operated in the region of economies of scale in any period. If the cost-side hypothesis is true, that is if economies of scale constituted an important reason for rising concentration then one would expect the large and growing firms (whose growth has impacted the concentration) to have economies of scale at the time of growth, so that by growing more they would exploit the economies of scale and get a cost advantage over the smaller firms.

The results of such a test clearly may depend on chosen appropriateness of methodology and quality of the data. This sub-section is devoted to an intuitive explanation of what I want to measure, and the rest of the section deals with the technical details on how the measurement of scale is approached here. The second issue, the data quality, will be addressed in the next section.

Frequently, the notion of economies (diseconomies) of scale is associated with the decreasing (increasing) portion of a u-shaped long-run *average cost* (*AC*) curve. The value of output where *AC* reaches its smallest value is often called the *minimum efficient scale* (*MES*) level of output. Thus, one (and most common) way to measure existence and size of economies of scale is by identifying the *slope* of the *AC* function.

An alternative way to do this is to measure the *distance* (difference) between the observed costs on the *AC* curve relative to the least possible costs associated with the *MES* level of output. If there is no difference then the economies of scale are fully exploited (assuming convexity of *AC* curve). If the difference exists then the next step would be to determine its source. If it exists

when output is below the *MES* level, then economies of scale are present (i.e., by increasing output, *AC* may decrease) in the neighborhood of measurement. On the other hand, if it is below then the diseconomies of scale are present (i.e., by increasing output, *AC* may increase). Clearly, under convexity both derivative method and the 'distance' method are equivalent *qualitatively*. Quantitatively, however they are different, as will be seen when the latter method is defined formally in the following sections.

To illustrate, consider Figure 4. Suppose for simplicity that all firms have access to the same technology characterized by the long-run average cost function $AC(y,p|V)$, where y is (single) output, and p is a vector of input prices. (Meaning of identifier V after " | " and precise definitions is done later.)

The $AC(y,p|C)$ is a 'virtual' *AC* curve that would exist if the technology allowed producing *any* level of output at the costs associated with the *MES* of the 'true' AC curve, $AC(y,p|V)$. That is, $AC(y,p|C)$ is a *constant returns to scale* (CRS) average cost curve (note, '|C' stands for CRS) that goes through the minimum point (ac^{MES}, y^{MES}) on the $AC(y,p|V)$ curve.

Now, consider first the observation $A' = (y^A, a\hat{c}^A)$, where y^A is observed output and $a\hat{c}^A$ is observed average cost for some firm A'. Since it is on the decreasing portion of the $AC(y,p|V)$ curve, economies of scale are present for this observation, and can be measured by the slope at point A' or by the distance between A' and A^*. Similarly, for the observation C' the scale economies can be measured either by the slope at point C' or by the distance between C' and C^*. Since C' is on the increasing portion of the $AC(y,p|V)$ curve, diseconomies of scale pertain to this observation. Finally, the scale economies for the observation B' can also be measured either by the slope at point B' or by the distance between B' and B^*. Since B' is on the flat portion of the $AC(y,p|V)$ curve that coincides with $AC(y,p|C)$, no economies or diseconomies of scale are present for this observation.
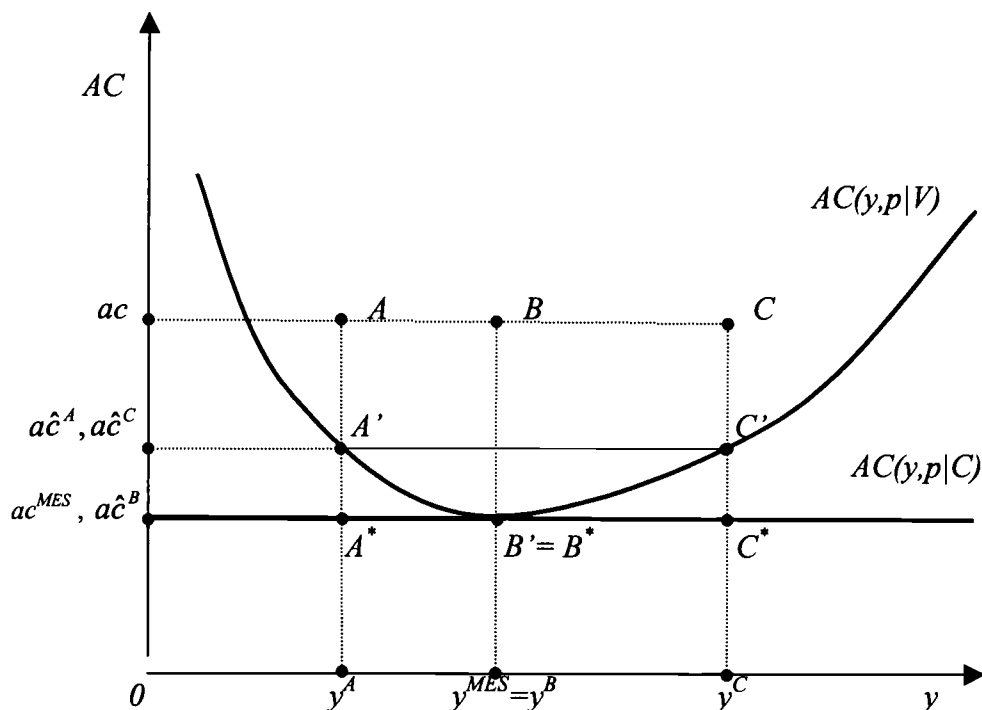
Figure 4. Measuring the Economies of Scale: Distance Approach

At this stage it is important to note two differences between the 'derivative method' and the 'distance method'. First, the derivative method indicates economies (diseconomies) of scale by giving a positive (negative) number, while it takes an additional step to get such indication for the distance method. Second, for any two observations (e.g., A' and C') having the same difference between $AC(y,p|C)$ and $AC(y,p|V)$ is not related (in general) to the fact that these observations have the same derivatives at these points, even in the absolute values. Roughly, this means that the two methods are in general *quantitatively* not related (although qualitatively they are). In this study I use the distance method.

There are many ways of measuring the distance. One simple and intuitive way of measuring the existence and size of economies of scale as the distance between the 'virtual' and 'true' average costs for a particular observation k is to

take their ratio (another natural way would be to take the difference), i.e., formally such approach gives the following measure:

$$(2.1.1) \qquad \textit{Average Cost Scale Efficiency} = \frac{AC(y^k, p \mid C)}{AC(y^k, p \mid V)}$$

which I dub as the 'average cost scale *efficiency*' measure (to distinguish it from the cost scale *elasticity* concept based on the derivative method). Note that this measure will give a number between zero and one (since by construction, $0 < AC(y^k, p \mid C) \leq AC(y^k, p \mid V)$). Unity will indicate that observation k is 'scale efficient' or has 'no scale economies and diseconomies' (e.g. as observation B'). If the number is less then unity, then the observation is dubbed as cost scale inefficient with two (mutually exclusive) possibilities or sources for this inefficiency: either due to scale economies or due to scale diseconomies. Additional step is needed to identify the source of inefficiency. In the single output case discussed here and given a U-shape AC curve, it is sufficient to compare the actual output level of the observation to the MES level of output: If the former is smaller (bigger) than the latter, then economies (diseconomies) of scale are present.

In our previous example, the average cost scale efficiency of observation A' is $a\hat{c}^A / ac^{MES} \leq 1$, for B' it is $a\hat{c}^B / ac^{MES} = 1$, and for C' it is $a\hat{c}^C / ac^{MES} \leq 1$. Also note that although, in our peculiar example, (2.1.1) gives the same efficiency scores for A' and C', the sources for inefficiency are conceptually different: for A' it is the scale economies, while for C' it is the scale diseconomies.

Now, consider the possibility of cost inefficiency, namely that observations do not necessarily 'lie' on the true cost curve, $AC(y^k, p \mid V)$, but somewhere above it. That is, some firms may be inefficient in their use of inputs. Such observations are called 'pure cost inefficient'. This type of inefficiency for a particular observation k can be measured similarly as in (2.1.1), as the ratio of the 'true' average (minimal)

costs to the actual (i.e., observed) average costs, ac^k, both associated with the observed output y^k. Technically,

$$(2.1.2) \quad \textit{Pure Average Cost Efficiency} = \frac{AC(y^k, p \mid V)}{ac^k}$$

In our hypothetical example in Figure 1, it may be that firms' actual average costs were, say, equal to ac. In this way, instead of A', B' and C', the actual observations are A, B, and C.

A standard way of measuring the cost *scale inefficiency* for observations with pure cost inefficiency is first to project the point onto the frontier (i.e., identify the where the observations should have been if they were 'pure cost efficient': e.g., points A', B' and C' for A, B and C, respectively) and then measure the cost *scale* efficiency from those points. Another way to look at this is to recognize that the *overall average cost inefficiency* for a particular observation k, can be measured as the ratio of the virtual minimal costs to the actual average costs, ac^k, where both are associated with the observed output y^k, i.e.

$$(2.1.3) \quad \textit{Overall Average Cost Efficiency} = \frac{AC(y^k, p \mid C)}{ac^k},$$

and then can be decomposed into two sources *(i)* pure cost efficiency and, and *(ii)* cost scale efficiency, i.e.,

$$(2.1.4) \quad \textit{Overall Average Cost Efficiency} = \frac{AC(y^k, p \mid V)}{ac^k} \times \frac{AC(y^k, p \mid C)}{AC(y^k, p \mid V)}$$

$$= \textit{Pure Average Cost Efficiency} \times \textit{Average Cost Scale Efficiency}$$

Let us now extend the *intuitive* analysis of scale economies into the multi-output framework. At the first glance, it may seem difficult since the average cost

are usually defined as the total cost divided by the *scalar* output (for exception see Baumol et al., 1988) and now there is a *vector* of outputs y. However, due to the ratio form of these measures, applying this definition to all above formulas yields measures defined in terms of the *total costs*, $C(\cdot)$, rather than the average costs, $AC(\cdot)$. Making appropriate changes I obtain the following decomposition of the cost efficiency measures[5]

$$(2.1.5) \qquad \textit{Overall Cost Efficiency} = \frac{C(y^k, p \mid V)}{c^k} \times \frac{C(y^k, p \mid C)}{C(y^k, p \mid V)},$$

$$= \textit{Pure Cost Efficiency} \times \textit{Cost Scale Efficiency}.$$

A general way to identify the *source* of inefficiency (i.e., whether the scale inefficiency is due to economies or diseconomies of scale) is to use the procedure outlined in Färe, Grosskopf and Lovell (1994), which will be outlined below, after more precise treatment of the efficiency concepts and its measurement is introduced.

2.2. The Means of Measurement: Firm's Level

I start with a general framework, where each firm k ($k = 1, ...,K$) in an industry consisting with K firms is allowed to have different technology characterized by its input sets

$$(2.2.1) \qquad L^k(y^k \mid r) = \{x^k : x^k \textit{ can produce } y^k \textit{ with } r \textit{ RTS}\}, \quad y^k \in \Re_+^M,$$

---

[5] Note, the type of measurement presented here is the *cost* or *input* oriented measurement, meaning that all measurement is done by looking at the potential reduction of costs, keeping the *output* constant.

where $x^k = (x_{k1}, ..., x_{kN}) \in \mathfrak{R}^N_+$ denotes firm k's input vector and $y^k = (y_{k1}, ..., y_{kM}) \in \mathfrak{R}^M_+$ its output vector ($k = 1, ... , K$). The index r will stand for description of the returns to scale (RTS) of the technology. Here, I allow for four types of returns to scale: constant ($r = C$), non-decreasing ($r = ND$), non-increasing ($r = NI$) and variable ($r = V$). The latter is the most general in the sense that it allows for existence of any other RTS locally. For definitions of these types of RTS see appendix or Färe, Grosskopf and Lovell (1994)). In particular, the following relations between the technologies with different RTS are true (see Färe, Grosskopf and Lovell (1994)):

(2.2.2) $$L^k(y^k \,|V) \subseteq L^k(y^k \,|NI) \subseteq L^k(y^k \,|C) \qquad \text{and}$$

$$L^k(y^k \,|V) \subseteq L^k(y^k \,| ND) \subseteq L^k(y^k \,|C), \;\; \forall y^k \,.$$

Given $L(y)$, the *technical efficiency* for a firm k is defined in Farrell (1957) tradition as

(2.2.3) $$F_i^k(y^k, x^k \,|r) = \min\{\lambda^k \,: \; (x^k)\lambda^k \in L^k(y^k \,|r)\}$$

Further, denote the input prices by $p = (p_1, ..., p_N) \in \mathfrak{R}^N_+$, (which I assume are the same for all firms), then firm k's observed cost is $px^k \equiv c^k$ and the firm's minimal cost is defined as

(2.2.4) $$C^k(y^k, \; p \,|r) = \min_x \{px \,: \; x \in L^k(y^k \,|r)\}.$$

This functions can be used to define the Farrell-type measure of *cost efficiency* as

(2.2.5a) $$F_c^k(y^k, c^k \,|r) = \min\{\theta^k \,: \; (c^k)\theta^k \geq C^k(y^k, p \,|r)\}$$

which is easily reduced to the following closed form

$$(2.2.5b) \qquad F_c^k(y^k, c^k \mid r) = C^k(y^k, p \mid r)/c^k$$

Assigning particular returns to scale (CRS or VRS) to (2.2.5) gives two types of measures of our interest: *'overall cost efficiency'* and *'pure cost efficiency,'* respectively:

$$(2.2.6a) \qquad F_c^k(y^k, c^k \mid C) = C^k(y^k, p \mid C)/c^k$$

and

$$(2.2.6b) \qquad F_c^k(y^k, c^k \mid V) = C^k(y^k, p \mid V)/c^k$$

(Note that these are the measures that were intuitively described in the previous section.)

Now, let $\tilde{c}^k$ be the point on $C^k(y^k, p \mid V)$ where the observations k should be if it were pure cost efficient for the output level y^k, i.e., $\tilde{c}^k = c^k F_c^k(y^k, c^k \mid V)$ then, following the intuition developed in the previous section, the cost scale efficiency is defined similar to (2.2.6a), as

$$(2.2.7) \qquad SE_c^k(y^k, c^k) \equiv F_c^k(y^k, \tilde{c}^k \mid C)$$

Using this definition and noting that for any technology, $F_c^k(y^k, c^k)$ is homogeneous of degree -1 in c^k (as can be seen from its definition), I get

$$(2.2.8) \qquad SE_c^k(y^k, c^k) = F_c^k(y^k, \tilde{c}^k \mid C) = F_c^k(y^k, c^k \mid C)/F_c^k(y^k, c^k \mid V)$$

which is the Färe and Grosskopf (1985) measure of *'cost scale efficiency'*. Substituting (2.2.6a), (2.2.6b) and (2.2.8) into (2.2.7) we get another interpretation of the scale efficiency—one that was heuristically described in the previous section, i.e.

$$(2.2.9) \qquad SE_c^k(y^k, c^k) = \frac{C^k(y^k, p \mid C)}{C^k(y^k, p \mid V)}$$

As in the previous section, the firm's 'overall cost efficiency' is *decomposed* as

$$(2.2.10) \qquad F_c^k(y^k, c^k \mid C) = F_c^k(y^k, c^k \mid V) \times SE_c^k(y^k, c^k)$$

i.e.,

Overall Cost Efficiency = Pure Cost Efficiency × Cost Scale Efficiency.

Up until now I considered only *individual* or firms efficiencies. The next section is devoted to development of the *aggregate* or industry (or group) efficiencies.

2.3. Aggregation Issues

The issues of *aggregation* in efficiency analysis are especially important when the analysis involves comparison of the efficiency of *groups* of observations. This is precisely the case of our study, where *a priori* information encourages bisecting the sample into two industry *strategic groups*: the *national* firms and the *regional* firms (e.g., as in Tremblay, 1985b, 1987).

There are two important issues regarding aggregation: *(i)* what kind of aggregation functions (additive, multiplicative, etc.) are appropriate for the cost efficiency framework, and *(ii)* what set of weights is needed in the aggregation. It is straightforward to show that both the quantitative and qualitative results of the

aggregate analysis are dependent upon answers to each of these questions. Such answers therefore must be justified with either economic and/or some technical consistency criteria. In this study I follow the aggregation theory developed by Färe and Zelenyuk (2002), adapting it to the measurement of the pure cost and cost scale efficiencies. In this section I only present the results of such adaptation, while the details are given in Appendix.

Let $k = 1, ...K$ be the index of firms in a given group (regional firms, national firms, entire Industry, etc.) and let $S^k = \dfrac{c^k}{\sum_{k=1}^{K} c^k}$, be the *observed* cost-share-weight of a firm k in this group. Let $F_c^k(y^k, c^k \mid V)$ be the firm k pure cost efficiency score, and thus define the *efficient* cost-share-weight of a firm k in the group of $k = 1, ...K$, firms as $\hat{S}^k = \dfrac{c^k \cdot F_c^k(y^k, c^k \mid V)}{\sum_{k=1}^{K} c^k \cdot F_c^k(y^k, c^k \mid V)}$, then under very general conditions (see appendix), the following aggregate efficiency measures are justified:

The *group overall cost efficiency* is obtained as

(2.3.1) $\qquad \overline{F}_c(y^1,..., \ y^K, \ \sum_{k=1}^{K} c^k \mid C) = \sum_{k=1}^{K} F_c^k(y^k, c^k \mid C) \cdot S^k.$

The *group pure cost efficiency* is obtained as

(2.3.2) $\qquad \overline{F}_c(y^1,..., \ y^K, \ \sum_{k=1}^{K} c^k \mid V) = \sum_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot S^k.$

The *group cost scale efficiency* is obtained as

(2.3.3) $\qquad \overline{SE}_c(y^1,..., \ y^K, \ \sum_{k=1}^{K} c^k) \equiv \dfrac{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid C) \cdot S^k}{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot S^k},$

or equivalently as

$$(2.3.4) \qquad \overline{SE}_c^{**}(y^1,..., \ y^K, \ \textstyle\sum_{k=1}^K c^k \mid C) = \sum_{k=1}^K SE_c^k(y^k,c^k) \cdot \hat{S}^k .$$

Thus, if one can estimate the individual scores for pure-cost, cost-scale and overall cost efficiencies, then one can use the above formulas to obtain the corresponding *aggregate* efficiencies. How to actually estimate the individual scores is the subject of the next section.

2.4. The Means of Estimation: Non-Stochastic-Non-Parametric Approach

Up until now I assumed that firms may have different technologies. Such generality still enabled us to receive consistent aggregation results relating the group efficiencies to the individual efficiencies. In empirical analysis, data availability often forces researchers to assume that all firms have the same technology, or make a weaker assumption that all firms have access to the same technology. This is based on the maintained hypothesis that firms have access to the *best practice technology*. In this study I use the non-stochastic-non-parametric activity analysis models, also known as the data envelopment analysis (DEA) models, to form such 'best practice cost frontiers.' All observations (firm's data) will then be measured relative to this frontier.

To make things more precise, the minimum cost frontier for each particular observation k', $(k' = 1, \ldots, K)$ can be estimated from the solutions of the following linear programming (LP) problem (see Färe and Grosskopf (1985) for details):

$$(2.4.1) \qquad \hat{C}(y^{k'}, p \mid C) = \min \sum_{n=1}^N p_n x_n$$

$$\text{s.t.} \qquad \sum_{k=1}^K z_k y_{km} \geq y_{k'm} \ , \quad m = 1, \ldots, M,$$

$$\sum_{k=1}^{K} z_k x_{kn} \le x_n \quad, \quad n = 1, \dots, N.$$

$$z_k \ge 0 \ , \quad k = 1, \dots, K,$$

where "hat" symbol hereafter will indicates that the function value for particular observation is *estimated*.

To incorporate the concept of VRS and compute $\hat{C}(y^{k'}, p \,|\, V)$, another constraint, $\sum_{k=1}^{K} z_k = 1$, is added to the LP problem (2.4.1). Such LP problems for each firm k, along with the fact that $c^k \equiv \sum_{k=1}^{K} p x_k$, gives all information to obtain the desired cost efficiency measures as

(2.4.2) $\qquad \hat{C}(y^{k'}, p \,|\, C) / c^{k'} \qquad\qquad$ and $\qquad\qquad \hat{C}(y^{k'}, p \,|\, V) / c^{k'}$

Such an approach requires data on all outputs, all inputs and their prices. Frequently, information on all inputs used in production for each firm may be unavailable for a researcher or very costly to find. The observed cost data (e.g., available from balance sheets), c^k, may be much easier to obtain. In fact, this is exactly the situation in our study. Färe and Grosskopf (1985) showed an alternative way for obtaining the efficiency scores in (2.4.2). In particular, under the assumption that all firms face the same prices they showed that

(2.4.3) $\qquad \hat{F}_c(y^{k'}, c^{k'} \,|\, C) = \dfrac{\hat{C}(y^{k'}, p \,|\, C)}{c^{k'}} \qquad$ and

$$\hat{F}_c(y^{k'}, c^{k'} \,|\, V) = \dfrac{\hat{C}(y^{k'}, p \,|\, V)}{c^{k'}},$$

where

$(2.4.4)$ $\qquad \hat{F}_c(y^{k'}, c^{k'} \mid C) = \min \lambda$

$$\text{s.t.} \quad \sum_{k=1}^{K} z_k y_{km} \geq y_{k'm} \quad, \quad m = 1, ..., M,$$

$$\sum_{k=1}^{K} z_k c_k \leq \lambda c_{k'} \quad,$$

$$z_k \geq 0 \quad, \quad k = 1, ..., K.$$

and $\hat{F}_c(y^{k'}, c^{k'} \mid V)$ is computed as in LP problem (2.4.4) but with additional constraint stating that $\sum_{k=1}^{K} z_k = 1$.

The estimates of the efficiency scores in (2.4.3) for each k are then used to compute the cost scale efficiency scores for each firm as described in sub-section 2.2 (see (2.2.8)) and aggregated as described in sub-section 2.4. The results from such estimation and aggregation procedures are discussed in the section 4, after I describe the sources and features of the used data set.

Further, to reveal the source of inefficiency, two approaches can be used. The simplest approach that works for the single output case is to identify the interval of full scale efficiency (MES level of output) and then compare it with the output level of a scale inefficient observation. If the output level of a scale inefficient observation is greater than the MES level of output, then clearly the source is the decreasing returns to scale. If it is smaller, then the source is increasing returns to scale.

A more general approach that works for the multiple output case is to compute the additional LP problems (2.4.4) with another constraint: $\sum_{k=1}^{K} z_k \leq 1$, thus obtaining $\hat{F}_c(y^{k'}, c^{k'} \mid NIRS)$—the cost efficiency measure obtained from the activity model with the assumption of non-increasing returns to scale (e.g., see Färe, Grosskopf and Lovell (1994) for details). If an observation k' is scale inefficient (i.e., $\hat{F}_c(y^{k'}, c^{k'} \mid V) > \hat{F}_c(y^{k'}, c^{k'} \mid C)$) and at the same time

$\hat{F}_c (y^{k'}, c^{k'} | V) > \hat{F}_c (y^{k'}, c^{k'} | NIRS)$ then the source of inefficiency is coming from increasing returns to scale. If instead, $\hat{F}_c (y^{k'}, c^{k'} | V) = \hat{F}_c (y^{k'}, c^{k'} | NIRS)$ then it is coming from the decreasing returns to scale. To make our inference on the source of scale inefficiency we use both approaches.

3. Data

The data set used in this study was received from C. Tremblay and V. Tremblay (Department of Economics, Oregon State University). See Tremblay (1985) for data sources and a description of the data. The data set is a panel consisting of 22 beer-producing companies for the period 1950-1985. [6] Table 2 gives a summary of the data. From 1950-55, the sample of firms produced about 20% of industry output. This number reached 50% by 1966 and 77% by 1985. This is due to the increase in the output shares of largest firms in the industry.

The data sets for the groups of *national* and *regional* producers have three main variables: c^k = the total cost of firm k, measured in thousands of dollars, y^k = the firm's output, measured in thousands of 31 gallon barrels of beer, and the industry output (also measured in thousands of 31 gallon barrels of beer). [7] In addition, the data set also contains information on the number of firms in the industry, the 'four-firm concentration ratio' index and the Herfindahl index of concentration available from 1950 till 1998. The sample also includes observations obtained from the survey of *small* regional producers. Two companies that agreed to provide information under conditions of confidentiality. In order to maintain their confidentiality, all firms in the sample are identified by numbers.

---

[6] Some observations that had missing values for output or cost variables were eliminated from the original data set for computational reasons.

[7] Firm's total cost was obtained as the difference between the (deflated) total revenue and profit for each particular firm, and then deflated to 1972 dollars using the 'wholesale price index' of the *U.S. Bureau of Labor Statistics*.

Table 2. Summary of the Data

Year	# of firms in Sample	# of firms in Population	Sample Mean of Output	Sample Mean of Total Costs	Industry Output	Output Share of Sample in Population
1950	8	369	1947	61733	82923	19%
1951	8	348	2136	64663	83939	20%
1952	8	300	2329	75597	84959	22%
1953	8	288	2589	84985	86209	24%
1954	8	261	2350	80230	83488	23%
1955	9	246	2213	72962	85204	23%
1956	12	236	2326	75628	85257	33%
1957	12	210	2346	75854	84668	33%
1958	12	199	2455	81261	84758	35%
1959	12	194	2748	84502	88006	37%
1960	12	181	2770	85632	88314	38%
1961	12	177	2903	88869	89473	39%
1962	12	166	3142	98090	91700	41%
1963	11	153	3637	115686	94338	42%
1964	10	142	4288	134793	99312	43%
1965	9	130	5116	153053	101059	46%
1966	10	119	5214	151602	104938	50%
1967	12	109	4776	139375	107638	53%
1968	14	94	4938	145137	112190	62%
1969	15	89	4789	138103	117066	61%
1970	15	84	5177	146727	122750	63%
1971	13	81	6057	175397	128318	61%
1972	13	77	7081	210925	132740	69%
1973	13	67	7741	206544	139600	72%
1974	11	61	9691	239482	146850	73%
1975	10	55	10992	286663	150323	73%
1976	10	51	10802	244915	152773	71%
1977	10	49	11197	276093	159460	70%
1978	11	46	12184	553213	166169	81%
1979	10	41	15425	761882	172559	89%
1980	10	38	14744	810162	177934	83%
1981	8	36	18219	1085757	181917	80%
1982	7	34	23008	1436732	182332	88%
1983	8	33	19607	1350252	183809	85%
1984	7	34	21002	1407626	182682	80%
1985	7	34	20225	1388672	183046	77%
Mean	10.472	162	7726.8	349688.8	122741.7	193%
St.Dev.	2.2	92.6	6363.1	439227.9	37610.8	286%
Min	7	49	1947.4	61733.4	82923.0	19%
Max	15	369	23008.0	1436732.1	183809.0	894%

4. Beer Industry Characteristics

One remarkable fact about the U.S. Brewing Industry that inspired this study is the dramatic decrease in the number of firms (see Figure 5). Imperatively, this decrease coincided with the *increase* in the industry output (see Figure 6).
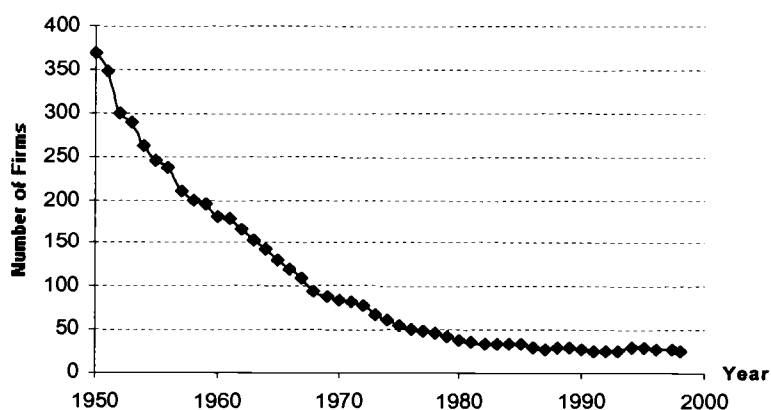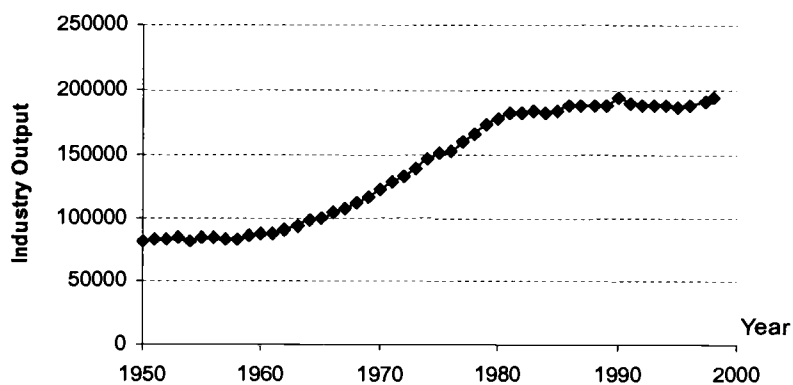
Figure 5. Number of Firms in US Brewing Industry

Figure 6. Total Output in the U.S. Brewing Industry

One implication of this concurrence is that that while some firms were leaving the industry, some of the surviving firms were growing more than needed to take the market share of the exiting firms.
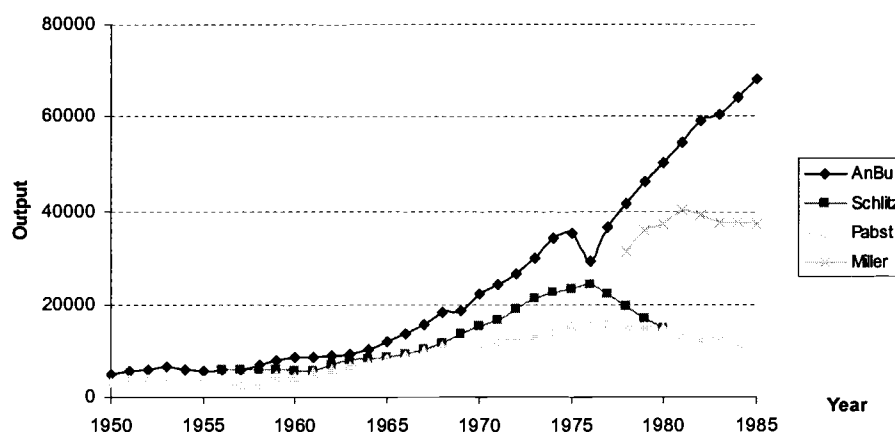
Figure 7. Total Output of 4 Largest Firms

Figure 7, gives information on the time series of the total output of the four largest (all national) firms. As one can see, the most successful was Anheuser Busch, whose output was increasing at an exponential rate during the entire period of study.[8] Two other firms, Schlitz and Pabst seemed to be trying to keep up with the industry leader, but only up to 1976, after which their output declined. The information on second largest firm, Miller, is available only from 1978 till 1985, and is telling us that it also could not keep with the expansion of the industry leader.

In fact, the average size of a firm has increased from 1950 to 1977 by about 14 times. [9] The resulting impact on industry concentration was also quite dramatic: the 'four-firm concentration ratio' has increased from about 20% in 1950 to more than 90% in 1990s (see Figure 8). The Herfindahl index presents a similar picture of industry concentration in brewing (see Figure 9).

---

[8] Exception was 1976—the year of 100-day strike at Anheuser-Busch.

[9] It is computed as the industry total output divided by the number of firms in the industry.
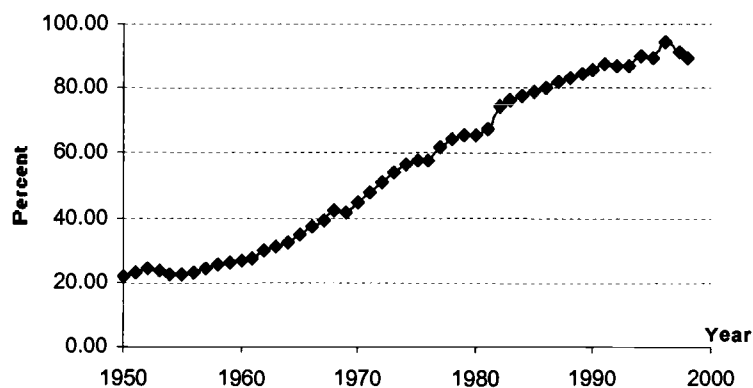
Figure 8. Four Firm Industry Concentration Ratio Index
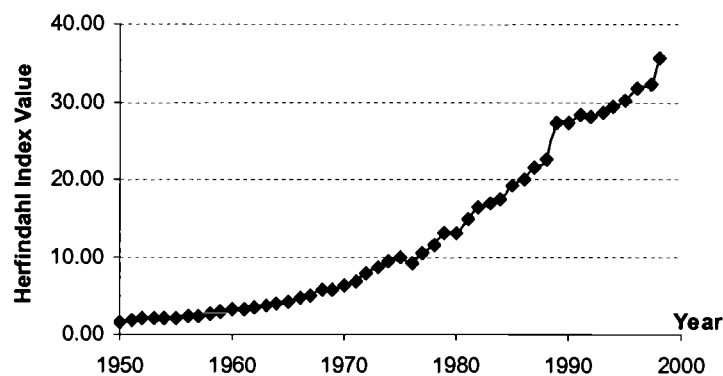
Figure 9. Herfindahl Industry Concentration Index

What could have been the reasons for such a striking structural change? As was mentioned previously, competing hypothesis that were dominating economists debates to explain such phenomenon can roughly be classified into two general types: the *demand side* and the *cost side* arguments, respectively. The goal of this study is to test the cost side argument for rising concentration.

5. Estimation Results and Implications

In this section I apply techniques presented and developed in earlier sections. In particular, I use the DEA models from section 2 to approximate the cost frontier for each year and then estimate the corresponding efficiency scores (2.4.3) for each observation in each year. Individual (i.e., for each firm) efficiency scores are then aggregated over all firms to obtain the estimate of the *industry* efficiency. Following Peles (1971), Hatten and Schendel (1977), and Tremblay (1985b and 1987), I then decompose the industry into two strategic groups: *national*[10] and *regional* producers, and obtain efficiency scores for these groups. The aggregation process was described in sub-section 2.4, where the appropriate aggregating function and unique sets of weights were derived. The aggregate efficiency scores for each group are then presented in Table 3, and depicted in the Figures 10, 11, 12.

Figure 10 and columns 2-4 of the Table 3 give an aggregate picture of the 'overall cost efficiency'. When looking at the whole sample representing the industry, one can see that its efficiency has been increasing from about 0.55 in early 1950s to about 0.65 in the late 1950s, and remained at that level (with some fluctuations) for almost the rest of the study period, with the exception of increase the late 1970's.

A more vivid picture comes with the bisection of the industry into its strategic groups. On one hand, one can see that the aggregate 'overall cost efficiency' of *regional* producers is decreasing (with some fluctuations) for most of the period (starting in the middle of 1950s from about 0.87 to about 0.62 in 1975), suddenly increased in 1978 (up to 0.91) and then went down to around 0.7. At the same time, one can observe a different picture for *national* producers:

---

[10] National firms are defined to be Anheuser Busch, Miller, Pabst, and Schlitz.

Table 3. Summary of Estimation Results

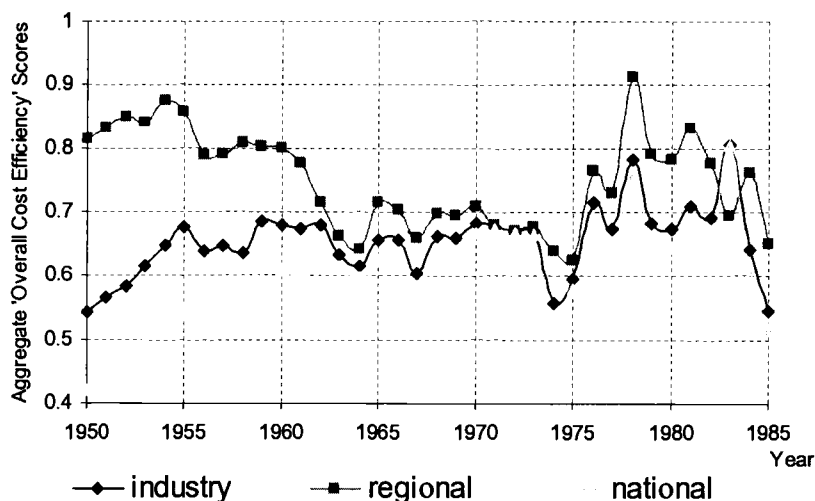| Year	Overall Cost Efficiency			Pure Cost Efficiency			Cost *Scale* Efficiency		
	industry	regional	national	industry	regional	national	industry	regional	national
1950	0.54	0.81	0.42	0.90	0.88	0.91	0.60	0.92	0.46
1951	0.57	0.83	0.45	0.96	0.89	0.99	0.59	0.94	0.45
1952	0.58	0.85	0.46	0.96	0.92	0.98	0.61	0.92	0.47
1953	0.62	0.84	0.50	0.98	0.94	1.00	0.63	0.89	0.50
1954	0.65	0.88	0.51	0.93	0.95	0.91	0.70	0.92	0.56
1955	0.68	0.86	0.54	0.92	0.92	0.92	0.73	0.93	0.59
1956	0.64	0.79	0.55	0.83	0.89	0.79	0.77	0.88	0.69
1957	0.65	0.79	0.56	0.92	0.90	0.94	0.70	0.88	0.60
1958	0.63	0.81	0.53	0.89	0.90	0.88	0.72	0.90	0.60
1959	0.69	0.80	0.61	0.96	0.88	1.00	0.72	0.91	0.61
1960	0.68	0.80	0.61	0.96	0.90	1.00	0.71	0.89	0.61
1961	0.67	0.78	0.61	0.94	0.88	0.98	0.71	0.89	0.62
1962	0.68	0.72	0.66	0.89	0.72	0.99	0.77	0.99	0.66
1963	0.63	0.66	0.61	0.87	0.67	0.99	0.73	0.98	0.62
1964	0.61	0.64	0.60	0.84	0.65	0.95	0.73	0.98	0.63
1965	0.66	0.71	0.62	0.87	0.78	0.91	0.76	0.91	0.69
1966	0.66	0.70	0.63	0.87	0.80	0.91	0.76	0.88	0.70
1967	0.60	0.66	0.57	0.83	0.71	0.90	0.72	0.93	0.64
1968	0.66	0.70	0.64	0.84	0.74	0.90	0.79	0.95	0.71
1969	0.66	0.69	0.64	0.88	0.73	0.97	0.75	0.95	0.66
1970	0.68	0.71	0.67	0.89	0.73	0.98	0.77	0.97	0.68
1971	0.68	0.68	0.67	0.90	0.71	0.99	0.75	0.95	0.68
1972	0.67	0.67	0.67	0.88	0.69	0.98	0.76	0.97	0.68
1973	0.67	0.68	0.67	0.89	0.69	1.00	0.75	0.99	0.67
1974	0.56	0.64	0.52	0.92	0.73	1.00	0.60	0.87	0.52
1975	0.59	0.62	0.58	0.90	0.68	0.99	0.66	0.92	0.58
1976	0.72	0.77	0.69	0.93	0.77	1.00	0.77	0.99	0.69
1977	0.67	0.73	0.65	0.90	0.74	0.97	0.75	0.98	0.67
1978	0.78	0.91	0.75	0.94	0.94	0.95	0.83	0.98	0.80
1979	0.68	0.79	0.65	0.94	0.95	0.94	0.73	0.84	0.69
1980	0.67	0.78	0.64	0.94	0.92	0.94	0.71	0.85	0.68
1981	0.71	0.83	0.67	0.94	0.92	0.94	0.76	0.91	0.72
1982	0.69	0.78	0.66	0.95	0.95	0.95	0.73	0.82	0.69
1983	0.81	0.69	0.80	0.95	0.95	0.94	0.85	0.73	0.85
1984	0.64	0.76	0.61	0.93	0.88	0.94	0.69	0.86	0.65
1985	0.54	0.65	0.52	0.94	0.86	0.96	0.58	0.75	0.54
Mean	**0.65**	**0.75**	**0.60**	**0.91**	**0.83**	**0.95**	**0.72**	**0.91**	**0.64**
StDev	**0.06**	**0.08**	**0.08**	**0.04**	**0.10**	**0.04**	**0.06**	**0.06**	**0.09**
Min	**0.54**	**0.62**	**0.42**	**0.83**	**0.65**	**0.79**	**0.58**	**0.73**	**0.45**
Max	**0.81**	**0.91**	**0.80**	**0.98**	**0.95**	**1.00**	**0.85**	**0.99**	**0.85**

Figure 10. Aggregate Overall Cost Efficiency

Its aggregate 'overall cost efficiency' score is fairly steadily (with some fluctuations) increasing during almost all of the period. In particular, it increased from 0.42 in 1950 to the average of about 0.65 in the rest of the period.

Interestingly, the 'overall cost efficiency' of both groups came close to each other in 1962 and moved in the same direction since then (except in 1983). In general, the average 'overall cost efficiency' over all periods was higher for regional producers than for national ones (0.75 vs. 0.60).

At least two questions arise with these results. First, what was causing such different (and then similar) pattern of efficiency distribution among groups? Second, what was causing the fluctuations in efficiency of each group, especially the sharp declines and rises? To answer these questions, I will use the decomposition of efficiency into different sources. As described in Section 2, the 'overall cost efficiency' measure can be decomposed into two sources of inefficiency: (1) due to economies (diseconomies) of scale and due to pure cost inefficiency (i.e., due to a failure to be on or 'close' to the VRS frontier).

I will start with the source of primal interest of this study—the estimate of the cost *scale inefficiency*—used to identify and measure the existence and size of the economies (diseconomies) of scale. The estimation results of the aggregated

cost scale efficiency for the industry and its two strategic groups are depicted in Figure 11 and columns 8-10 of Table 3.
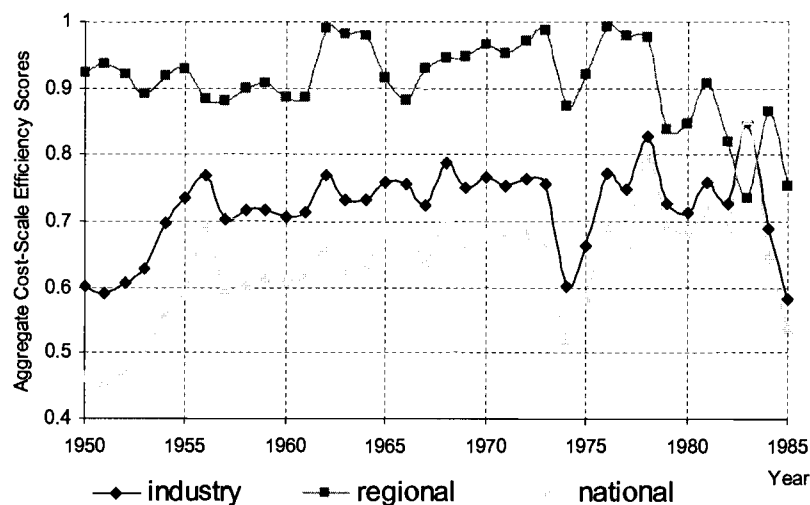

Figure 11. Aggregate Cost Scale Efficiency

One can see a distinctive picture: On average, the firms representing the group of *regional* producers were consistently more scale efficient than the *national* producers in all years, except one. Specifically, the scale efficiency of regionals was about 0.91, averaging over the whole period, while the nationals had it only about 0.64. This indicates that the regionals, on average, were operating very close to the MES output level, while the nationals were far from it. An immediate question is: Where is this source of scale inefficiency coming from? Namely, is it due to existing economies or diseconomies of scale? Additional step to identify the source of scale inefficiency reveals that none of the national firms experienced economies of scale with respect to the observed best practice frontier in any year under the study. On the contrary, all scale inefficiency was coming from diseconomies of scale, i.e., due to being *oversized*.

I will take a closer look on this issue, by looking at the efficiency of each of the national firms, after considering the other source of the overall cost

inefficiency—the *'pure cost inefficiency'*. Columns 5-7 of the Table 3 and Figure 12 give an aggregate picture regarding this type of inefficiency for entire sample (representing the industry) and with a bisection into the two strategic groups (nationals and regionals).
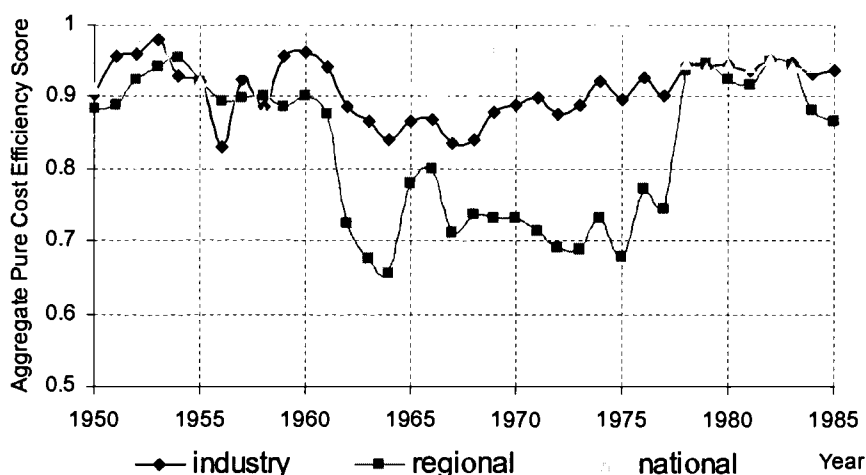

Figure 12. Aggregate Pure Cost Efficiency

Up to 1961, the 'pure cost efficiency' for the two strategic groups was similar (with slight dominance by the nationals) and quite high: about 0.9, on average. Interestingly, in earlier periods, like 1954-55 and 1957-58, the 'pure cost efficiency' of both groups was nearly the same. After that, the nationals had consistently higher efficiency (on average about 0.97) than the regional, (with the average efficiency score being about 0.74).

It might seem surprising and even contradicting to earlier conclusions that the nationals, who have just been convicted in being oversized and operating at the huge diseconomies of scale level of output, and being 'overall less efficient' than the regionals, are now having higher 'pure cost efficiency' standing. Is it really a contradiction? No, it is not. Let's recall what the 'pure cost efficiency' measure is really telling us in the context of data envelopment analysis. It gives information on how far each firm is from the *best* practice *VRS* (cost) frontier. It is quite possible

to have high "pure cost efficiency" but very low scale efficiency (and therefore low overall cost efficiency), simply because the best practice VRS frontier always passes through the largest firm (which also gets the highest weight in the aggregation). Thus, if the largest firm is very inefficient it will still be "pure cost efficient" (the inefficiency will show up in the scale and overall efficiency estimate). For the second largest firms to be efficient, it just has to beat efficiency of the largest firm, and so on. So it is possible for a group to have very high pure cost efficiency and very low cost scale and overall inefficiency.

Let's now turn to a closer look at the efficiency of the four largest firms in the sample. While Anheuser Busch was the industry leader in terms of volume of production during all the period of study, and especially since late 1960s, its overall cost efficiency standing does not look so brilliant (see Figure 13). On the contrary, most of the time it had the lowest (0.54 on average) overall cost efficiency among the national producers (and in the entire sample). Miller, second largest, firm was even less efficient when it appeared in the sample. Efficiency of the third largest firm, Schlitz, was higher than that of Anheuser Busch and Miller, but still quite low, 0.66 on average. Interestingly, the most efficient firm among the national firms, was the smallest among them—Pabst, with average of 0.81. Notably, its efficiency rapidly increased in 1959—the first wave of rapid technological change in the industry. A few years later Pabst become one of the most efficient firms in the sample.[11] Altogether, one can see that the larger the firm, the lower its overall cost efficiency.

---

[11] Also note that its efficiency was first low, than got even lower in 1958 right before the sharp increase in efficiency, which may be a result of increase in cost due to large investment into more efficient technology. Note that similar phenomenon is observed in 1974, period of the second wave of major technological change, and early 1980's, the third wave.
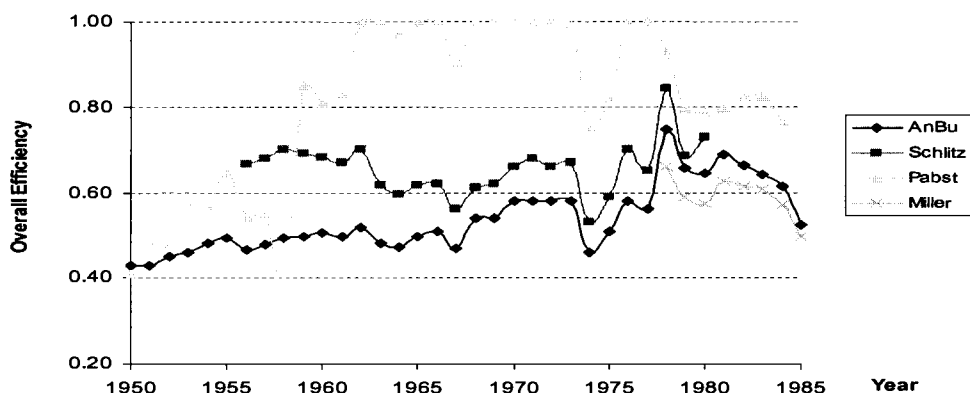
Figure 13. Overall Cost Efficiency of 4 Largest Firms

Turning attention to one of the source of the overall cost inefficiency—the pure cost inefficiency, gives us a different picture. Anheuser Busch almost always was on the best practice VRS frontier. However, as discussed above, it is just because the DEA frontier always passes through the largest observation, by construction.[12] Schlitz, the second largest firm in the sample (till 1978), was also close to the frontier for most of the years except for period between 1964 and 1968 (i.e., the period between the first and second waves of major technological change in the industry). Miller, when appeared in the sample failed to ever be on the best practice frontier.

---

[12] A sudden dip in efficiency in 1956 happened because during that particular year it was not Anheuser Busch that was the largest firm in the sample, but Schlitz.
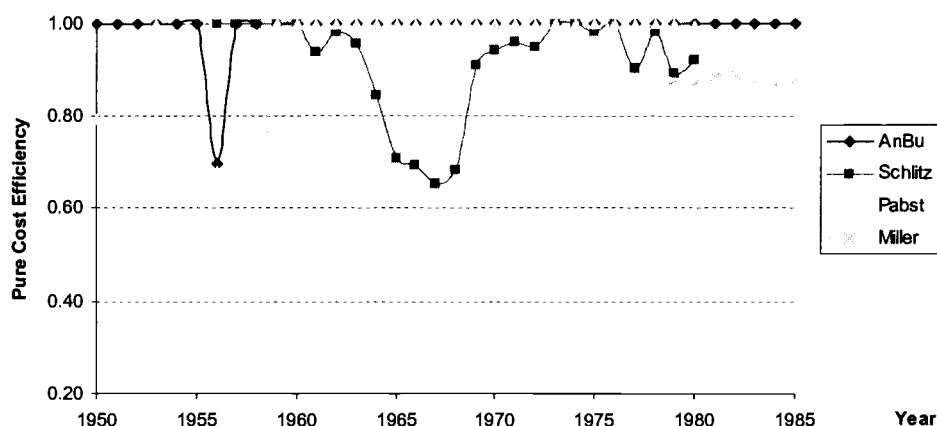
Figure 14. Pure Cost Efficiency of 4 Largest Firms

Finally, Figure 15 reveals that the largest source of the overall inefficiency of national firms is coming from the cost scale efficiency. In particular, Anheuser Busch was a true leader in terms of the cost scale *inefficiency*, with an average of 0.54 over the entire period. The second least efficient was Miller (average of 0.68), then Schlitz (average of 0.73) and finally the most efficient in the group of nationals was Pabst, with the average of 0.86

Figure15. Cost Scale Efficiency of 4 Largest Firms

All the results of the undertaken efficiency analysis can be summarized as follows. *First*, the largest (national) firms never experienced economies of scale with respect to the *best practice frontier*. On the contrary, except for the smallest of the national firms, Pabst, they were vastly oversized and experienced diseconomies of scale. *Second*, despite the lowest scale efficiency, the largest firm, Anheuser, continued to grow larger with an exponential rate, and was often presumed as the most 'successf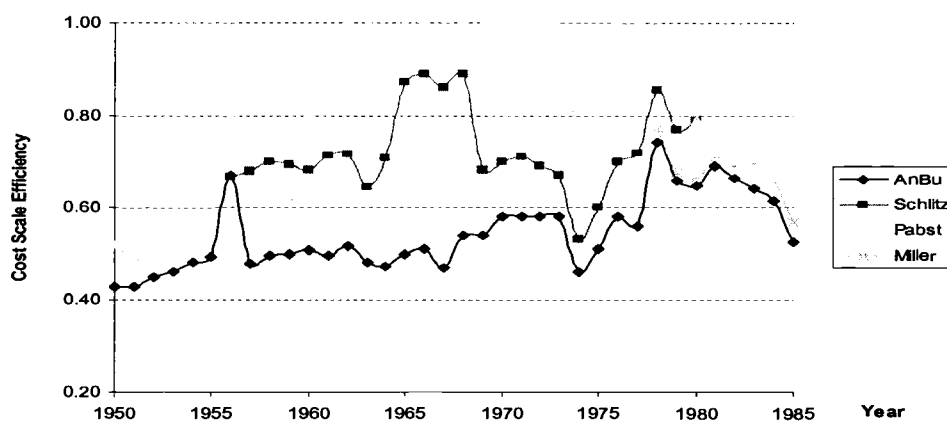ul' firm in the industry. The other large firms also tried to grow larger (but at some point failed keeping up with the leader). The only firm among the largest that ever had perfect scale efficiency was the smallest of them, which also decreased its scale efficiency (also due to diseconomies of scale) since late 1970's. *Third*, most of the regional firms experienced economies of scale, but they were very small—as indicated by high cost scale efficiency estimates (i.e., they operated close to MES level of output).

Clearly, even if the small regional producers were able to exploit all these *minor* economies of scale, they would not increase the industry concentration as dramatically as the expansion of the largest firms who had diseconomies of scale.[13] In other words, conditional on the data, my conclusion of this study is that the hypothesis that the economies of scale were predominant causes for the rise in concentration in the U.S. brewing industry during 1950-1985 must be rejected. The key to success in the industry was not growth to exploit economies of scale but something else.

---

[13] In fact, the concentration can even decrease if the expansion of the regional firms will cause contraction of the largest firms, say if measured by 'four firm concentration ratio'.

6. Conclusions and further Speculations

It is not hard to see that the results of this study are consistent with economic theory. These results provide a clearer picture of why the concentration rose so dramatically in the US brewing industry. First, the analysis indicates that national producers were operating on or close to the *increasing* portion of the best practice average cost curve. Second, findings of Tremblay (1985b, 1987) in demand and cost studies (respectively) give me another useful lead: the demand curve that a typical *national* producer faced for its produce was flatter than that for a typical *regional* producer with a slightly greater (estimated) intercept. (Intuitively, this means that on average, national producers were able to sell more output than regional producers were, for the same and every price.)

These two clues along with the standard microeconomic theory helps shed some light on what could have been happening in the US brewing Industry. I present this story in Figure 16.

The figure incorporates findings of Tremblay (1985b, 1987), by having the demand curve that a typical *national* producer faced for its produce being flatter than and above that for a typical *regional* producer (D^N vs. D^R). The figure also incorporates finding of present study that there existed (although minor) economies of scale and diseconomies of scale, i.e., the average cost curve is "U-shaped."
Consequently, under such conditions, according to economic theory, the profit-maximizing price of national producers must be greater than that of regional producers. The profit maximizing choice of output level for firms facing D^R is close to *MES* level, as it was for the regional producers in our study. On the other hand, the profit maximizing choice for firms facing D^N (the higher and flatter demand curve as national firms face) is at the level of substantial diseconomies of scale
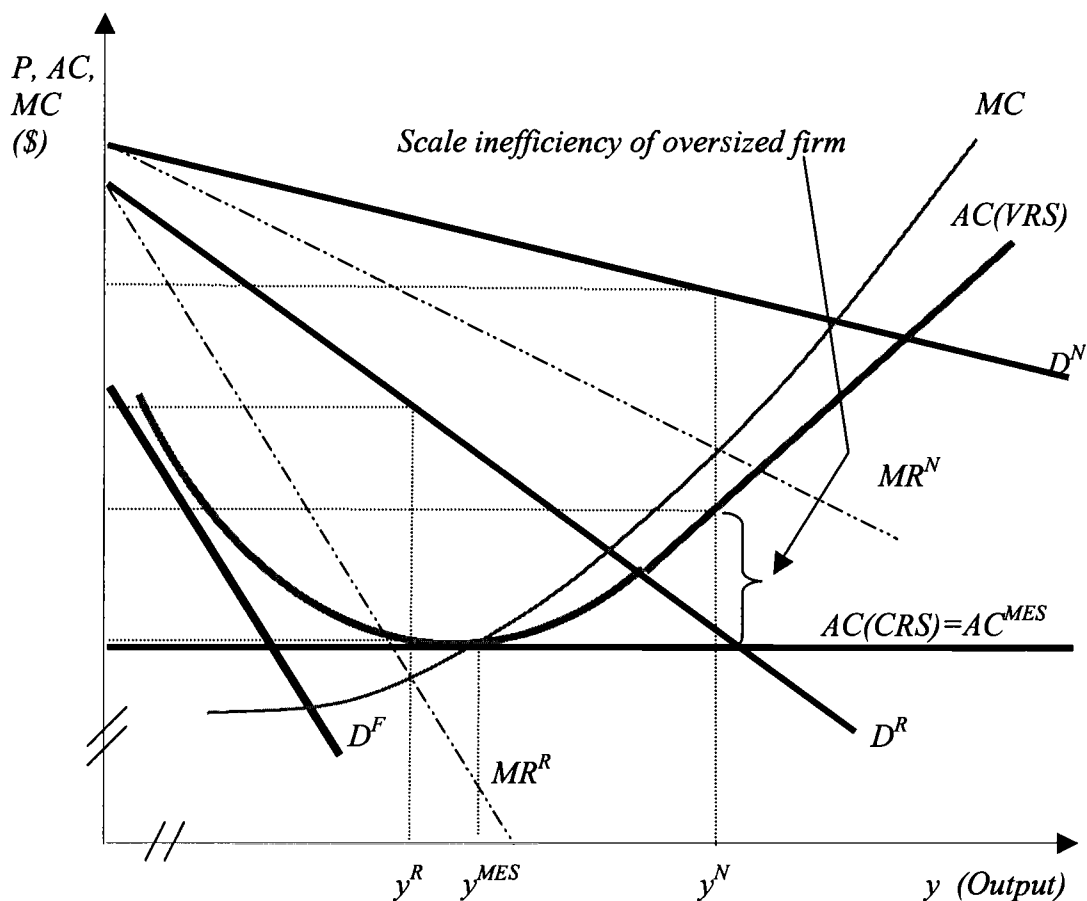
Figure 16. The Demand and Cost Structure of National and Regional Firms.

rather than at the *MES* or nearby level. This is exactly what we observed in our study with the national producers, two of which were vastly oversized, despite some expectations.

Economic theory predicts that any firm in the industry would strive to expand its demand curve (say from D^R to D^N) even for the price of incurring more and more diseconomies of scale. Aggressive marketing campaigns like product differentiation and excessive advertising could have been the tools for it. And, this is exactly what was observed in the industry by researchers (Greer 1971, 1981 and

Tremblay and Tremblay, 1996) as well as by the public observing exploding advertising, increasing variety, etc.

The shrinkage and failure of some firms can be also explained with this picture. Since some firms were growing faster than the market growth, part of these firms' expansions (shifts in the demands to the right, say due to successful advertising) was at the cost of a reduction in market shares of other firms (shifts in the demands to the left. At the extreme, some firms ended up with the demand curves below the average cost AC (D^F on the Figure 16), and therefore had to leave the market in the long run.

Taking all the arguments together, the results of the past and current studies unified with general predictions of economic theory encourage me to conclude that the dramatic structural change in the US brewing industry was mostly driven *not* from the cost reasons like scale economies but from some other sources. Possibly, all the roads lead to the demand side causes like successful marketing tactics (e.g., product differentiation and advertising) that increased the market shares for some firms and reduced or even deleted the shares of others. This demand-side argument, however, can also be broadened to include the cost side as well.

Specifically, the phenomenon of rising concentration can be viewed as a result of the strategic dynamic game between competing firms. Simply put, the competing firms may play strategies to increase the demand they face for their produce, but these endeavors might be reflected in higher cost, as we observed in our study for the largest firms. Perhaps, among the best examples of such dynamic games, where firms block the entry for and/or crowding out other firms by excessive *sunk cost* into marketing campaigns is given by Sutton (1991). A new challenge, of course, is to develop a test for verifying this hypothesis, but this is beyond the scope of the present study.

References

Aczél, J. (1990), "Determining Merged Relative Scores" *Journal of Mathematical Analysis and Applications* 150: 1, July 15, 1990.

Baumol, W., J. Panzar and R. Willig (1988), *Contestable Markets and the Theory of Industry Structure*, Harcourt Brace Jovanovich Publishers, New York.

Blackorby, C. and R. Russell (1999), "Aggregation of Efficiency Indices", *Journal of Productivity Analysis* 12:1, pp. 5-20.

Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units," *European Journal of operational research* 2: 6, pp. 429-444.

Day, D.L., A.Y. Lewin, H. Li, (1995) "Strategic leaders or strategic groups: A longitudinal Data Envelopment Analysis of the U.S. brewing industry," *European Journal of Operational Research* 80, pp. 619-638.

Elzinga, K. G. (1986), "The Beer Industry" in Adams, Walter, ed. *The structure of American industry*, pp. 203-38, Seventh edition, New York: Macmillan; London: Collier Macmillan.

Elzinga, K. G. (1973), "The Restructuring of the U.S. Brewing Industry," *The Industrial Organization Review* 1, pp.101-114.

Färe, R., Grosskopf S. and J. Kirkeley (2001), "Industry Efficiency and Capacity," *(in progress)*.

Färe, R., Grosskopf S. and C.A.K. Lovell (1985), *The Measurement of Efficiency of Production*, Kluwer Nijhoff Publishing, Boston.

Färe, R. and S. Grosskopf (1985), "A Nonparametric Cost Approach to Scale Efficiency," *Scandinavian Journal of Economics* 87:4, pp. 594-604.

Färe, R. and S. Grosskopf and C.A.K. Lovell (1994), *Production Frontier*, Cambridge University Press, Cambridge.

Färe, R. and D. Primont (1995), *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers: Boston.

Färe, R. and V. Zelenyuk (2002), 'On Aggregate Farrell Efficiencies," *European Journal of Operational Research*, forthcoming.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, General 120:3, 253-281.

Försund, F., L. Hjalmarsson (1979), "Generalized Farrell measures of efficiency: An application to milk processing in Swedish dairy plants," *Economic Journal* 89, pp. 294-315.

Greer, D. F. (1981), "The Causes of Concentration in the U.S. Brewing Industry," *Quarterly Review of Economics and Business* 21:4, pp. 87-106.

Greer, D. F. (1971), "Production Differentiation and Concentration in the Brewing Industry" *Journal of Industrial Economics* 19:3, pp. 201-19.

Hatten, K. J. and D. F. Schendel (1977), "Heterogeneity within an Industry: Firm Conduct in the U.S. Brewing Industry, 1952-71," *Journal of Industrial Economics* 26:2, pp. 97-113.

Hogarty, T. F. and K.G. Elzinga (1972), "The Demand for Beer," *Review of Economics and Statistics* 54:2, pp. 195-98.

Horowitz, I and A. R. Horowitz (1965), "Firms in a Declining Market: The Brewing Case," *Journal of Industrial Economics* 13, pp. 129-53.

Horowitz, I and A.R. Horowitz (1967), "The Beer Industry," *Business Horizons* 10.

Kerkvliet, J., W. Nebesky, C.Tremblay, and V. Tremblay (1998), "Efficiency and Technological Change in the U.S. Brewing Industry," *Journal of Productivity Analysis* 10, pp.1-18.

Koopmans, T.(1957), *Three Essays on the State of Economic Analysis*, McGraw-Hill, New York.

Li, S.K., and Y.C. Ng (1995), "Measuring the Productive Efficiency of a Group of Firms," *International Advances in Economic Research* 1:4, pp. 377-90.

Lynk, W.J. (1984), "Interpreting Rising Concentration: The Case of Beer," *Journal of Business* 57:1 (Part 1), pp. 43-55.

Lynk, W. J. (1985), "The Price and Output of Beer Revisited [Interpreting Rising Concentration: The Case of Beer]," *Journal of Business* 58:4, pp. 433-37.

Mas-Colell, A., M. Whinston and J. Green (1995), *Microeconomic Theory*, Oxford: Oxford University Press.

Peles Y. (1971), "Economies of Scale in Advertising Beer and Cigarettes," *Journal of Business* 44:1, pp. 32-37.

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

Tremblay, V.J. (1985a), "A Reappraisal of Interpreting Rising Concentration: The Case of Beer," *The Journal of Business* 58: 4, pp. 419-31.

Tremblay, V.J. (1985b) "Strategic Groups and the Demand for Beer" *Journal of Industrial Economics* 34:2,183-98.

Tremblay, V.J. (1987) "Scale Economies, Technological Change, and Firm-Cost Asymmetries in the U.S. Brewing Industry," *Quarterly Review of Economics and Business* 27:2, pp. 71-86.

Tremblay, C. H., V. J. Tremblay (1988), "The Determinants of Horizontal Acquisitions: Evidence from the U.S. Brewing Industry," *The Journal of Industrial Economics* 37:1, pp. 21-45.

Tremblay, C. H. , V. Tremblay (1996), "Firm Success, National Status, and Product Line Diversification: An Empirical Examination," *Review of Industrial Organization* 11, 771-789.

Ylvinger S. (2000), "Industry performance and structural efficiency measures: Solutions to problems in firm models," *European Journal Of Operational Research* 121:1, 164-174.

Appendix

To fulfill the task—find an economics-justifiable aggregation methods for cost efficiencies—I follow the aggregation approach suggested by Färe and Zelenyuk (2002), adapting it to incorporate the concept of returns to scale.

A critical assumption of the approach is that the *group* (entire industry, a group within an industry, etc) technology is defined as the sum of firm's technologies

(A1) $\qquad \bar{L}(y^1,..., \ y^K \mid R) = \sum_{k=1}^{K} L^k(y^k \mid r^k)$

where

(A2) $\qquad L^k(y^k \mid r) = \{x^k : \ x^k \ can \ produce \ y^k \ with \ r \ RTS\}, \ y^k \in \mathfrak{R}^M_+,$

with $x^k = (x_{k1}, \ ..., \ x_{kN}) \in \mathfrak{R}^N_+$ denoting firm k's input vector, $y^k = (y_{k1}, \ ..., \ y_{kM}) \in \mathfrak{R}^M_+$ its output vector $(k = 1, \ ... , \ K)$. The index r and R will stand for description of the returns to scale *(RTS)* of the individual and aggregate technology. To be precise, the following four types of RTS are considered (see Färe, Grosskopf and Lovell (1994) for details):

(A3) $\qquad$ CRS $\quad \Leftrightarrow \quad L^k(t \cdot y^k) = t \cdot L^k(y^k), \forall t > 0$

(A4) $\qquad$ NIRS $\quad \Leftrightarrow \quad L^k(t \cdot y^k) \subseteq t \cdot L^k(y^k), \forall t \geq 1$

(A5) $\qquad$ NDRS $\quad \Leftrightarrow \quad L^k(t \cdot y^k) \supseteq t \cdot L^k(y^k), \forall t \geq 1$

(A6) $\qquad$ VRS $\quad \Leftrightarrow \quad \neg (A1) \ \text{and} \ \neg(A2) \ \text{and} \ \neg (A3)$

where "$\neg$" is the logical operator for "not".

As a result, the following is true (see Färe, Grosskopf and Lovell (1994) for details)

$$L^k(y^k \mid V) \subseteq L^k(y^k \mid NI) \subseteq L^k(y^k \mid C) \qquad \text{and}$$

$$L^k(y^k \mid V) \subseteq L^k(y^k \mid ND) \subseteq L^k(y^k \mid C), \ \forall y^k \ .$$

The definition of RTS can also be generalized to the aggregate technology in (AL)

(A7) $\qquad$ CRS $\quad \Leftrightarrow \quad \bar{L}(t \cdot y^1,..., \ t \cdot y^K) = t \cdot \bar{L}(y^1,..., \ y^K), \forall t > 0$

(A8) $\qquad$ NIRS $\quad \Leftrightarrow \quad \bar{L}(t \cdot y^1,..., \ t \cdot y^K) \subseteq t \cdot \bar{L}(y^1,..., \ y^K), \forall t \geq 1$

(A9) NDRS $\Leftrightarrow$ $\overline{L}(t \cdot y^1,..., \ t \cdot y^K) \supseteq t \cdot \overline{L}(y^1,..., \ y^K), \forall t \geq 1$

(A10) VRS $\Leftrightarrow$ $\neg$ (A5) and $\neg$ (A6) and $\neg$ (A7)

As a result, the following will be true,

$$\overline{L}(y^1,..., \ y^K \mid V) \subseteq \overline{L}(y^1,..., \ y^K \mid NI) \subseteq \overline{L}(y^1,..., \ y^K \mid C) \quad \text{and}$$

$$\overline{L}(y^1,..., \ y^K \mid V) \subseteq \overline{L}(y^1,..., \ y^K \mid ND) \subseteq \overline{L}(y^1,..., \ y^K \mid V), \forall (y^1,..., \ y^K)$$

Note that the group technology $\overline{L}(y^1,..., \ y^K)$ inherits its properties from those of the different firm technologies $L^k(y^k)$. (For example, if each $L^k(y^k)$ is a convex, compact set with inputs and outputs freely disposable, then so is $\overline{L}(y^1,..., \ y^K)$). Thus clearly, R, an index describing the *RTS* of the *group* technology resulted from the aggregation, in general depends on the *RTS* of the individual technologies (r^k) entering the aggregation. (Recall that in our general framework, each firm may use different output vectors y^k and may have different technology L^k (including different *RTS*)). The following two general result can be established, which will be useful in further derivations.

Lemma 1.

If all individual technologies exhibit CRS then the aggregate technology in (A1) also exhibits CRS.

Proof:

Suppose every firm k $(k = 1, ..., K)$ has CRS technology, then (and only then) by (A3): $L^k(t \cdot y^k \mid C) = t \cdot L^k(y^k \mid C)$, $t > 0$, for every k. Thus,

$$\sum_{k=1}^{K} L^k(t \cdot y^k \mid C) = t \cdot \sum_{k=1}^{K} L^k(y^k \mid C), \text{ i.e. by (A7) this means that}$$

$$\overline{L}(t \cdot y^1,..., \ t \cdot y^K \mid R) = t \cdot \overline{L}(y^1,..., \ y^K \mid R), \text{ implying that the *aggregate* technology}$$

is also CRS. *q.e.d.*

Lemma 2.

If all firms have VRS technology then the aggregate technology in (A1) is also VRS.

Proof:

Suppose every firm k ($k = 1, ..., K$) has VRS technology, then (and only then) by (A6), for some points $x^1,..., x^K : x^1 \in L^1(y^1),..., x^K \in L^K(y^K)$ the following will be true:

(A11) $\quad \exists t > 1: \quad tx^1 \in tL^1(y^1) \not\subseteq L^1(ty^1),..., tx^K \in tL^K(y^K) \not\subseteq L^1(ty^1)$, and

(A12) $\quad \exists \tau > 1: \quad x^1 \in \tau L^1(y^1) \not\supseteq L^1(\tau y^1),..., \tau x^K \in \tau L^K(y^K) \not\supseteq L^1(\tau y^1)$

(A13) $\quad \exists \delta > 1: \quad x^1 \in \delta L^1(y^1) \neq L^1(\delta y^1),..., \delta x^K \in \delta L^K(y^K) \neq L^1(\delta y^1)$

but (A9) implies that

$$\exists t > 1: \quad t\sum_{k=1}^{K<\infty} x^k \in t\sum_{k=1}^{K<\infty} L^k(y^k) \not\subseteq \sum_{k=1}^{K<\infty} L^k(ty^k), \qquad \text{and therefore}$$

(A14) $\quad \exists t > 1: \quad t \cdot \overline{L}(y^1,..., y^K) \not\subseteq \overline{L}(t \cdot y^1,..., t \cdot y^K) \qquad \text{(using (A1))}.$

Similarly, (A12) implies that

$$\exists \tau > 1: \quad \tau\sum_{k=1}^{K<\infty} x^k \in \tau\sum_{k=1}^{K<\infty} L^k(y^k) \not\supseteq \sum_{k=1}^{K<\infty} L^k(\tau y^k),$$

and therefore

(A15) $\quad \exists \tau > 1: \quad \tau \cdot \overline{L}(y^1,..., y^K) \not\supseteq \overline{L}(\tau \cdot y^1,..., \tau \cdot y^K) \qquad \text{(using (A1))}.$

Analogously, (A13) implies that

$$\exists \delta > 0: \quad \delta \sum_{k=1}^{K<\infty} x^k \in \delta \sum_{k=1}^{K<\infty} L^k(y^k) \neq \sum_{k=1}^{K<\infty} L^k(\delta y^k), \quad \delta > 0,$$

and therefore

(A16) $\quad \exists \delta > 0: \quad \delta \cdot \overline{L}(y^1,..., y^K) \neq \overline{L}(\delta \cdot y^1,..., \delta \cdot y^K), \forall \delta > 0 \quad$ (using (A1)).

Finally, note that (A12) and (A13) and (A14) together imply that none of the (A7), (A8), (A9) is satisfied, implying that this is the case of (A10), i.e. aggregate technology is also VRS. $\hfill q.e.d.$

Analogously to the disaggregated level (A1), the group minimal cost is defined as

(A17) $\qquad \overline{C}(y^1,..., y^K, p \mid R) = \min_x \{px : x \in \overline{L}(y^1,..., y^K \mid R)\}$

and the *group 'overall cost efficiency'* is defined analogous to the firm's one (see 2.2.6a) as

(A18a) $\overline{F}_c(y^1,..., y^K, \sum_{k=1}^{K} c^k \mid C) = \min\{\theta : (\sum_{k=1}^{K} c^k) \cdot \theta \geq \overline{C}(y^1,..., y^K, p \mid C)\}.$

Immediately from this definition I get the closed form of the *group 'overall cost efficiency'* measure

(A18b) $\qquad \overline{F}_c(y^1,..., y^K, \sum_{k=1}^{K} c^k \mid C) = \dfrac{\overline{C}(y^1,..., y^K, p \mid C)}{\sum_{k=1}^{K} c^k}$

Similarly, the group *'pure cost efficiency'* is defined analogous to the firm's one (see 2.2.6a) as

(A19a) $\quad \overline{F}_c(y^1,..., y^K, \sum_{k=1}^{K} c^k \mid V) = \min\{\theta : \left(\sum_{k=1}^{K} c^k\right)\theta \geq \overline{C}(y^1,..., y^K, p \mid V)\}$

from which it follows that

(A19b) $\quad \overline{F}_c(y^1,..., y^K, \sum_{k=1}^{K} c^k \mid V) = \dfrac{\overline{C}(y^1,..., y^K, p \mid V)}{\sum_{k=1}^{K} c^k}$

Defining the group *scale* efficiency on the aggregate technology is a bit more tricky. Recall that while defining the individual scale efficiency, we first adjust for the "pure cost efficiency" (to bring the observation to the frontier) and then measured the scale efficiency from that "adjusted" point. On the aggregate level, there are at least two ways to do the adjustment.

The first one, starts with aggregating the observed individual costs, $\sum_{k=1}^{K} c^k$, then correcting this cost for the industry pure cost inefficiency, $\overline{F}_c(y^1,..., y^K, \sum_{k=1}^{K} c^k \mid V)$, and then measuring the scale efficiency from that point using $\overline{F}_c(\cdot)$. Hence, letting $\hat{\overline{C}} \equiv (\sum_{k=1}^{K} c^k) \cdot \overline{F}_c(y^1,..., y^K, \sum_{k=1}^{K} c^k \mid V)$ (note the resemblance with the definition on disaggregated level), one measure of *group cost scale efficiency* can be defined as

(A20) $\quad \overline{SE}_c(y^1,..., y^K, \sum_{k=1}^{K} c^k) \equiv \overline{F}_c(y^1,..., y^K, \hat{\overline{C}} \mid C)$

Using this definition and homogeneity property of the l.h.s., I get the following

$$\text{(A21)} \qquad \overline{SE}_c(y^1,...,\ y^K,\ \textstyle\sum_{k=1}^{K}c^k) \equiv \frac{\overline{F}_c(y^1,...,\ y^K,\ \sum_{k=1}^{K}c^k \mid C)}{\overline{F}_c(y^1,...,\ y^K,\ \sum_{k=1}^{K}c^k \mid V)}$$

$$= \frac{\overline{C}(y^1,...,\ y^K,\ p \mid C)}{\overline{C}(y^1,...,\ y^K,\ p \mid V)} \qquad \text{by (A18b), (A19b)}$$

i.e., it is the ratio of the group *virtual* minimal cost (associated with the *group* constant *RTS* technology and group *MES*) to the group *true* minimal cost. Hence, this measure can be used to decompose the *group 'overall cost efficiency'* as

$$\text{(A22)} \qquad \overline{F}_c(y^1,...,\ y^K,\ \textstyle\sum_{k=1}^{K}c^k \mid C)$$

$$= \overline{F}_c(y^1,...,\ y^K,\ \textstyle\sum_{k=1}^{K}c^k \mid V) \times \overline{SE}_c(y^1,...,\ y^K,\ \textstyle\sum_{k=1}^{K}c^k)$$

i.e.,

Aggregate Entire Cost Efficiency = Aggregate Pure Cost Efficiency

× Aggregate Cost Scale Efficiency

The second approach is to start not with aggregation, but with correcting the observed individual costs for the pure (individual) cost inefficiency to get $\tilde{c}^k = c^k F_c^k(y^k, c^k \mid V)$, then aggregate it into the 'corrected' industry observed costs, $\sum_{k=1}^{K}\tilde{c}^k$, and then measure the group scale efficiency from that "corrected" point using the same $\overline{F}_c(\cdot)$. The resulting measure is,

$$\text{(A23)} \qquad \overline{SE}_c^{*}(y^1,...,\ y^K,\ \textstyle\sum_{k=1}^{K}c^k) \equiv \overline{F}_c(y^1,...,\ y^K,\ \textstyle\sum_{k=1}^{K}\tilde{c}^k \mid C),$$

where $\tilde{c}^k = c^k F_c^k(y^k, c^k \mid V)$, from which it follows that

(A24) $\quad \overline{SE}_c^*(y^1,..., \ y^K, \ \sum_{k=1}^K c^k) = \dfrac{\overline{C}(y^1,..., \ y^K, \ p \mid C)}{\sum_{k=1}^K \tilde{c}^k}$

i.e., it is also the ratio of the industry minimal cost (associated with the constant *RTS* technology and industry *MES*) to the sum of individual cost corrected for pure cost efficiency.

Both measures are intuitive and are the aggregate analogs of the disaggregate measure of cost scale efficiency. They are however received by different aggregation routes, particularly, by the order when the aggregation steps in: before or after adjustment for pure scale inefficiency.

It is worthwhile to note at this point that this is not the only way the group efficiencies might be defined. In particular, what we have done is first aggregated individual technologies and then defined the group efficiency measure on it, using the same efficiency measures as we used on the disaggregated level.

Alternatively, the group efficiencies may be defined as some aggregates of the individual efficiencies. That is, perhaps there exist appropriate aggregating functions $G_i^K : \Re^K \to \Re^1$, *(i =1, 2, ...)* for which we may derive the *group* pure cost efficiencies from the corresponding individual efficiencies, i.e.,

(A25) $\quad \overline{F}_c^*(y^1,..., \ y^K, \ c^1,..., \ c^K \mid C) \equiv G_1(F_c^1(y^1,c^1 \mid C),...,F_c^K(y^K,c^K \mid C))$,

(A26) $\quad \overline{F}_c^*(y^1,..., \ y^K, \ c^1,..., \ c^K \mid V) \equiv G_2(F_c^1(y^1,c^1 \mid V),...,F_c^K(y^K,c^K \mid V))$,

and derive the *group* cost scale efficiency from the individual cost scale efficiencies, i.e.,

(A27) $\qquad \overline{SE}_c^{**}(y^1,..., y^K, c^1,..., c^K) \equiv G_3(SE_c^1(y^1,c^1),...,SE_c^K(y^K,c^K))$.

Moreover, if (A25) and (A26) can be constructed, then another alternative (forth!) definition of cost scale efficiency might be appropriate to consider:

(A28) $\qquad \overline{SE}_c^{***}(y^1,..., y^K, c^1,..., c^K) \equiv \dfrac{\overline{F}_c^*(y^1,..., y^K, c^1,..., c^K \mid C)}{\overline{F}_c^*(y^1,..., y^K, c^1,..., c^K \mid V)}$

i.e., this measure is defined analogous to (A21), but using (A25) and (A26) as the components.

A challenging question now is: Which aggregate measures to choose for our measurement? Ideally, one wants a group efficiency measure to be independent on whether it is obtained by first aggregating the technologies and then defining the efficiency measure on it or derived from the individual efficiencies. It turns out that for the pure cost and cost scale efficiencies such ideal can be achieved! Leaving all the manipulations that will follow aside, the key result that brings us to such an ideal is stated in the following proposition.

Proposition 1. Minimal cost of a group (A17) is equal to the sum of the minimal cost of all members of this group, i.e.,

(A29) $\qquad \overline{C}(y^1,..., y^K, p \mid R) = \sum_{k=1}^{K} C^k(y^k, p \mid r^k)$.

Proof:
This is a *cost* version of the proof from Färe and Zelenyuk (2002).

Let $x^k \in L^k(y^k)$, then $\sum_{k=1}^{K} x^k \in \overline{L}(y^1,..., y^K)$ and hence

$$\overline{C}(y^1,...,\ y^K,\ p) \le p\sum_{k=1}^K x^k = \sum_{k=1}^K px^k \text{ , i.e.,}$$

(A30) $$\overline{C}(y^1,...,\ y^K,\ p) \le \sum_{k=1}^K C^k(y^k,p).$$

Conversely, let $x \in \overline{L}(y^1,...,y^K)$, then by definition there are $x^k \in L^k(y^k)$ so that $x = \sum_{k=1}^K x^k$. Hence, $px = p\sum_{k=1}^K x^k = \sum_{k=1}^K px^k \ge \sum_{k=1}^K C^k(y^k,p)$, and

(A31) $$\overline{C}(y^1,...,\ y^K,\ p) \ge \sum_{k=1}^K C^k(y^k,p)$$

Inequalities *(A30)* and *(A31)* prove the claim.   *q.e.d.*

To immediately see the implication of this proposition, let S^k represent the cost-share weight of firm k ($k = 1, ... , K$) relative to the other firms in the group, i.e.

(A32) $$S^k = \frac{c^k}{\sum_{k=1}^K c^k} \ ,$$

then we get the following result.

Corollary 1. The group cost efficiency is the share weighted average of the firms' cost efficiencies, i.e.,

(A33) $$\overline{F}_c(y^1,...,\ y^K,\ \sum_{k=1}^K c^k \mid R) = \sum_{k=1}^K F_c^k(y^k,c^k \mid r^k) \cdot S^k$$

Proof:

Divide both sides of (A29) by $\sum\limits_{k=1}^{K} c^k$ and apply definition of cost efficiency)

This corollary tells us that (for the special form of aggregate technology in (A1)) the group cost efficiency defined on the aggregate technology and that derived from the individual efficiencies are equivalent. Lemma 1 and 2 above help us being more concrete on the RTS in the result (A32). Namely,

(A33a) $\qquad \overline{F}_c(y^1,\dots,\ y^K,\ \sum_{k=1}^{K} c^k \mid C) = \sum\limits_{k=1}^{K} F_c^k(y^k, c^k \mid C) \cdot S^k$

(A33b) $\qquad \overline{F}_c(y^1,\dots,\ y^K,\ \sum_{k=1}^{K} c^k \mid V) = \sum\limits_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot S^k$

thus obtaining a way to find the group overall and pure cost efficiencies from the corresponding individual efficiencies. In other words, a solution to (A25) and (A26) is found by choosing the aggregating function $G_i^K : \Re^K \to \Re^1$, *(i =1, 2)* to be weighted arithmetic average with weights defined in (A32).

Our next goal is to establish similar result for the cost scale efficiency. This task, however, is more challenging since we have four alternative definitions here.

Corollary 2. The cost scale efficiency in definition (A20) can be obtained as

(A34) $\qquad \overline{SE}_c(y^1,\dots,\ y^K,\ \sum_{k=1}^{K} c^k) \equiv \dfrac{\sum\limits_{k=1}^{K} F_c^k(y^k, c^k \mid C) \cdot S^k}{\sum\limits_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot S^k}$

Proof: Substitute (A33a-b) into (A21).

Corollary 3. The two cost scale efficiencies in definitions (A20) and (A23) are equal.

(A35) $\qquad \overline{SE}_c^*(y^1,..., \ y^K, \ \sum_{k=1}^{K}c^k) = \overline{SE}_c(y^1,..., \ y^K, \ \sum_{k=1}^{K}c^k).$

Proof:

Starting from (A24),

$$\overline{SE}_c^*(y^1,..., \ y^K, \ \sum_k c^k) = \frac{\overline{C}(y^1,..., \ y^K, \ p \mid C)}{\sum_{k=1}^{K} \tilde{c}^k} = \frac{\overline{C}(y^1,..., \ y^K, \ p \mid C)}{\sum_{k=1}^{K} c^k F_c^k(y^k, c^k \mid V)}$$

$$= \frac{\sum_{k=1}^{K} C^k(y^k, p \mid C)}{\sum_{k=1}^{K} C^k(y^k, p \mid V)} = \frac{\sum_{k=1}^{K} \dfrac{C^k(y^k, p \mid C)}{c^k} \cdot c^k / \sum_{k=1}^{K} c^k}{\sum_{k=1}^{K} \dfrac{C^k(y^k, p \mid V)}{c^k} \cdot c^k / \sum_{k=1}^{K} c^k} \qquad by \ (A29)$$

$$= \frac{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid C) \cdot S^k}{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot S^k} \qquad \text{(by definition of cost efficiency and (A32))}$$

$$q.e.d.$$

To consider the third definition of cost scale efficiency given in (A27), let $\hat{S}^k$ represent the adjusted for pure cost inefficiency cost-share-weight of firm k (k = 1, ..., K) in the group, i.e.

(A36) $\qquad \hat{S}^k = \dfrac{c^k \cdot F_c^k(y^k, c^k \mid V)}{\sum_{k=1}^{K} c^k \cdot F_c^k(y^k, c^k \mid V)},$

then we obtain the following result:

Corollary 4. If the cost scale efficiency in definition (A27) is such that $G_3^K : \Re^K \to \Re^1$ is the weighted arithmetic mean, where the weights are the cost shares adjusted for pure cost inefficiency defined in (A36), i.e.,

(A37) $\qquad \overline{SE}_c^{**}(y^1,..., y^K, \Sigma_{k=1}^K c^k \mid C) = \sum_{k=1}^{K} SE_c^k(y^k, c^k) \cdot \hat{S}^k$,

then (A27) is equal to the cost scale efficiency in definition (A20), i.e.,

(A38) $\qquad \overline{SE}_c(y^1,..., y^K, \Sigma_{k=1}^K c^k) = \overline{SE}_c(y^1,..., y^K, \Sigma_{k=1}^K c^k)$.

Proof:

Starting from result in Corollary 2, stated in (A34) we get

$$\overline{SE}_c(y^1,..., y^K, \sum_k c^k) \equiv \frac{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid C) \cdot S^k}{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot S^k} = \frac{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid C) \cdot \dfrac{c^k}{\sum_{k=1}^{K} c^k}}{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot \dfrac{c^k}{\sum_{k=1}^{K} c^k}}$$

$$= \sum_{k=1}^{K} \left(\frac{F_c^k(y^k, c^k \mid C)}{F_c^k(y^k, c^k \mid V)} \cdot \frac{c^k \cdot F_c^k(y^k, c^k \mid V)}{\sum_{k=1}^{K} F_c^k(y^k, c^k \mid V) \cdot c^k} \right)$$

$$= \sum_{k=1}^{K} SE_c^k(y^k, c^k) \cdot \hat{S}^k \qquad\qquad\qquad q.e.d.$$

Note that since (A20) can be written as (A21), its relationship to the forth definition of scale efficiency follows immediately from Corollary 4.

Finally, note that according to all the group efficiency measures outlined here satisfy the Blackorby and Russell (1999) aggregate indication axiom (which is a special case of the agreement property in Aczél, 1990, p. 24), stating that the group is efficient if and only if all firms in the group are efficient.

BIBLIOGRAPHY

Aczel, J. (1966), *Lectures on Functional Equations and Their Applications*, Academic Press, New York/London.

Aczél, J. (1990), "Determining Merged Relative Scores," *Journal of Mathematical Analysis and Applications*,150:1.

Banker, R., A.Charnes, W.W. Cooper (1984), "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science* 30 (9), pp. 1078-92.

Baumol, W., J. Panzar and R. Willig (1988), *Contestable Markets and the Theory of Industry Structure*, Harcourt Brace Jovanovich Publishers, New York.

Berg, S., F. Førsund and E. Jansen (1992), "Malmquist indices of productivity growth during the deregulation of Norwegian banking," *Scandinavian Journal of Economics* 94, S211-S228.

Blackorby, C. and R. R. Russell (1999), "Aggregation of Efficiency Indices", *Journal of Productivity Analysis*12, 1, pp. 5-20.

Boussemart, J-P., W. Briec, K. Kerstens and J-C. Poutineau (2001), "Luenberger and Malmquist Productivity Indices: Theoretical Comparisons and Empirical Illustration," working paper.

Briec, W. (1999), "Holder Distance Function and Measurement of Technical Efficiency," *Journal of Productivity Analysis* 11:2, 111-31.

Caves, D., L. Christensen and W. E. Diewert (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica* 50:6, 1393-1414.

Chambers, R., Y. Chung and R. Färe (1996), "Benefit and Distance Functions," *Journal of Economic Theory* 70, 407-419.

Chambers, R., Y. Chung and R. Färe (1998), "Profit, Directional Distance Functions and Nerlovian Efficiency," *Journal of Optimization Theory and Application* 98:2, 351-364.

Chambers, R.G., and R. Färe (1998),"Translation Homotheticity," *Economic Theory*,11, 629-64.

Charnes, A., W.W. Cooper, B. Golany, L. Seiford and J. Stutz (1985), "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions," *Journal of Econometrics*, 30:12, 91 - 07.

Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision-making Units," *European Journal of Operational Research* 2:6, 429-444.

Chung, Y, R. Färe and S. Grosskopf (1997), "Productivity and Undesirable Outputs: A Distance Function Approach," *Journal of Environmental Management* 51, 229-240.

Cornes, R. (1992), *Duality and Modern Economics*, Cambridge University Press, Cambridge.

Day, D.L., A.Y. Lewin, H. Li (1995), "Strategic leaders or strategic groups: A longitudinal Data Envelopment Analysis of the U.S. brewing industry," *European Journal of Operational Research* 80, pp. 619-638.

Debreu, G. (1951),"The Coefficient of Resource Utilization," *Econometrica* 19, 273-292.

Diewert, W.E. (1992a), "The Measurement of Productivity," *Bulletin of Economic Research* 44:3, 163-98.

Diewert, W.E. (1992b), "Fisher Ideal Output, Input, and Productivity Indexes Revisited," *Journal of Productivity Analysis* 3: 3, 211-48.

Eichhorn, W. (1978), *Functional Equations in Economics*, Addison-Wesley, Reading, MA.

Elzinga, K. G. (1973), "The Restructuring of the U.S. Brewing Industry," *The Industrial Organization Review* 1, pp.101-114

Elzinga, K. G. (1986), "The Beer Industry" in Adams, Walter, ed. *The structure of American industry*, pp. 203-38, Seventh edition, New York: Macmillan; London: Collier Macmillan.

Färe, R. and S. Grosskopf (1985) "A Nonparametric Cost Approach to Scale Efficiency," *Scandinavian Journal of Economics 87* (4), pp. 594-604.

Färe, R., and S. Grosskopf (1996), *Intertemporal Production Frontiers: With dynamic DEA*, Kluwer Academic Publishers, Boston.

Färe, R. and S. Grosskopf (2000), "Theory and Application of Directional Distance Functions," *Journal of Productivity Analysis*, 13:2, 93-103

Färe, R., S. Grosskopf and J. Kirkeley (2001), "Industry Efficiency and Capacity," *(in progress).*

Färe, R., S. Grosskopf, B. Lindgren, and P. Roos (1989), "Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach," Discussion Paper 89-3, Southern Illinois University at Carbondale.

Färe, R., S. Grosskopf and C.A.K. Lovell (1985), *The Measurement of Efficiency of Production*, Kluwer Nijhoff Publishing, Boston.

Färe, R. and S. Grosskopf and C.A.K. Lovell, (1994) *Production Frontier,* Cambridge, Cambridge University Press.

Färe, R., S. Grosskopf and P. Roos (1997), Malmquist productivity indexes: a survey of theory and practice, in: R. Färe, S. Grosskopf and R. Russell, eds., *Index Numbers: Essays in Honor of Sten Malmquist*, Kluwer Academic Publishers, Boston.

Färe, R. and Sung Ko Li (2001), "A Nonparametric Test of Translation Homoetheticity *Economics Letters* 72, pp. 341-345.

Färe, R. and C.A.K. Lovell (1978), "Measuring Technical Efficiency of Production," *Journal of Economic Theory* 19, 1 50- 62.

Färe, R. and T. Mitchell, (1993) "Multiple Outputs and Homotheticity," *Southern Economic Journal,* pp.287-296.

Färe, R. and D. Primont (1995), *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers: Boston .

Färe, R. and V. Zelenyuk (2002), 'On Aggregate Farrell Efficiencies," *European Journal of Operational Research*, forthcoming.

Farrell, M. (1957), "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society,* Series A, General, 1 20, Part 3, 253-28.

Førsund, F. (1997), "The Malmquist Productivity Index, TFP and scale," Memorandum NO 233, Department of Economics, School of Economics and Commercial Law, Göteborg University.

Førsund, F. (1998) "On the Calculation of the Scale Elasticity in DEA Models" *Journal of Productivity Analysis* 7 (2-3), 283-302.

Førsund, F. and L. Hjalmarsson (1979), "Generalized Farrell measures of efficiency: an application to milk processing in Swedish dairy plants," *Economic Journal* 89, pp. 294-315.

Greer, D. F. (1971), "Production Differentiation and Concentration in the Brewing Industry" *Journal of Industrial Economics* 19:3, pp. 201-19

Greer, D. F. (1981), "The Causes of Concentration in the U.S. Brewing Industry," *Quarterly Review of Economics and Business* 21:4, pp. 87-106.

Grifell-Tatjé and Lovell (1999), "A Generalized Malmquist productivity index," *Sociedad de Estadística e Investigación Operativa Top* 7:1, pp. 81-101.

Hatten, K. J. and D. F. Schendel (1977), "Heterogeneity within an Industry: Firm Conduct in the U.S. Brewing Industry, 1952-71," *Journal of Industrial Economics* 26:2, pp. 97-113.

Hogarty, T. F. and K.G. Elzinga (1972), "The Demand for Beer," *Review of Economics and Statistics* 54:2, pp. 195-98.

Horowitz, I and A. R. Horowitz (1965), "Firms in a Declining Market: The Brewing Case," *Journal of Industrial Economics* 13, pp. 129-53.

Horowitz, I and A.R. Horowitz (1967), "The Beer Industry," *Business Horizons* 10.

Kerkvliet, J., W. Nebesky, C.Tremblay, and V. Tremblay (1998), "Efficiency and Technological Change in the U.S. Brewing Industry," *Journal of Productivity Analysis* 10, pp.1-18.

Koopmans, T.C. (1957), *Three Essays on the State of Economic Analysis*, McGraw-Hill, New York.

Kopp, R. (1981), "Measuring the Technical Efficiency of Production: A Comment," *Journal of Economic Theory* 25, 450-452.

Li, S.K., and Y.C. Ng (1995), "Measuring the Productive Efficiency of a Group of Firms," *International Advances in Economic Research* v.1, n.4 (November): 377-90.

Luenberger, D.G. (1992), "New Optimality Principles for Economic Efficiency and Equilibrium." *Journal of Optimization Theory and Applications*, 75 ,221-264.

Lynk, W.J. (1984), "Interpreting Rising Concentration: The Case of Beer," *Journal of Business* 57:1 (Part 1), pp. 43-55.

Lynk, W. J. (1985), "The Price and Output of Beer Revisited [Interpreting Rising Concentration: The Case of Beer]," *Journal of Business* 58:4, pp. 433-37.

Mas-Colell, A., M. Whinston and J. Green, (1995) *Microeconomic Theory*, Oxford University Press, Oxford.

Peles Y. (1971), "Economies of Scale in Advertising Beer and Cigarettes," *Journal of Business* 44:1, pp. 32-37.

Russell, R.R. (1985), "Measuring of Technical Efficiency," *Journal of Economic Theory* 35, 1109- 26.

Russell, R.R. (1987), "On the Axiomatic Approach to the Measurement of Technical Efficiency," in W. Eichhhorn, ed. *Measurement in Economics: Theory and Applications of Economic Indices*, Heidelberg: Physica Verlag, 207-2 7.

Russell, R.R. (1990), "Continuity of Measures of Technical Efficiency," *Journal of Economic Theory* 51, 255-267.

Seitz, W.D. (1970), "The Measurement of Efficiency Relative to a Frontier Production Function," *American Journal of Agricultural Economics* 52, pp. 505-511.

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

Sueyoshi (1999), "DEA Duality on Returns to Scale (RTS) in Production and Cost Analyses: An Occurance of Multiple Solutions and Differences Between Production-Based and Cost-Based RTS Estimates," *Management Science* v.45, n.11, pp. 1593-1608.

Tauer, L. (2001) "Input aggregation and computed technical efficiency," *Applied Economics Letters* 8, pp. 295-297.

Tremblay, C. H., V. J. Tremblay (1988), "The Determinants of Horizontal Acquisitions: Evidence from the U.S. Brewing Industry," *The Journal of Industrial Economics* 37:1, pp. 21-45.

Tremblay, C. H. and V. Tremblay (1996), "Firm Success, National Status, and Product line Diversification: An Empirical Examination," *Review of Industrial Organization* 11, 771-789.

Tremblay, V.J. (1985a), "A Reappraisal of Interpreting Rising Concentration: The Case of Beer," *The Journal of Business* 58: 4, pp. 419-31.

Tremblay, V.J. (1985b) "Strategic Groups and the Demand for Beer" *Journal of Industrial Economics* 34:2,183-98.

Tremblay, V.J. (1987) "Scale Economies, Technological Change, and Firm-Cost Asymmetries in the U.S. Brewing Industry," *Quarterly Review of Economics and Business* 27:2, pp. 71-86.

Ylvinger S. (2000), "Industry performance and structural efficiency measures: Solutions to problems in firm models," *European Journal Of Operational Research* 121:1, 164-174.