Industrial Efficiency in Six Nations

Richard E. Caves

in association with
Sheryl D. Bailey, John Baldwin, Fabienne Fecher, Alison Green, Chris M. Harris, David G. Mayes, Sergio Perelman, Akio Torii, and Seong Min Yoo

The MIT Press
Cambridge, Massachusetts
London, England

This book was set in Times Roman by Asco Trade Typesetting Ltd., Hong Kong and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Industrial efficiency in six nations / Richard E. Caves in association with Sheryl D. Bailey ... [et al.].
p. cm.
Includes bibliographical references and index.
ISBN 0-262-03193-0
1. Industrial productivity—Case studies. 2. Efficiency, Industrial—Case studies. I. Caves, Richard E. II. Bailey, Sheryl D.
HC79.I52I53 1992
338'.06—dc20 92-3851
CIP

Industrial Efficiency in Six Nations

Contributors

Sheryl D. Bailey
City of Norfolk
Norfolk, Virginia

John Baldwin
Statistics Canada
Ottawa, Canada

Richard E. Caves
Harvard University
Cambridge, Massachusetts

Fabienne Fecher
University of Liège
Liège, Belgium

Alison Green
Polytechnic South-West
Plymouth, U.K.

Chris M. Harris
Bureau of Industry Economics
Department of Industry, Technology, and Commerce
Canberra, Australia

David G. Mayes
National Institute of Economic and Social Research
London, U.K.

Sergio Perelman
University of Liège
Liège, Belgium

Akio Torii
Yokohama National University
Yokohama, Japan

Seong Min Yoo
Korea Development Institute
Seoul, Korea

Acknowledgments

We are grateful to the Alfred P. Sloan Foundation for a grant that supported a conference and concluding editorial work on this volume. The conference provided a valuable opportunity for the contributors to compare their findings and obtain reactions from other scholars who attended the conference. In the capacity of discussants and commentators on the project Finn R. Førsund, Zvi Griliches, Lennart Hjalmarsson, Richard R. Nelson, F. M. Scherer, Peter Schmidt, and Hideki Yamawaki offered many useful suggestions and points of perspective.

Several persons made contributions to the organization and development of this project that are not fully reflected in the papers. The project had its genesis a decade ago in discussions with Derek Morris at the U.K. National Economic Development Office, and his successor David G. Mayes aided in bringing the work to completion. Masu Uekusa took part in an earlier version of the Japanese study and helped with the organization of the project, as did John Baldwin and David R. Barton, the coauthor of the preceding U.S. study. Ann Flack looked after the organization of the conference and helped coordinate the subsequent assembly of this volume.

Contents

1 Introduction and Summary

Central to the concerns of microeconomics is allocative efficiency, the study of processes and policies that distribute resources among activities and sectors so they are put to their best uses. Much less attended—indeed often dismissed as uninteresting or unresearchable—is efficiency in the popular sense of whether we accomplish a given task with the minimum effort or use of scarce resources. This issue of technical or productive efficiency of course arises in macroeconomics and the economics of development, where economists have long sought to explain why labor and other resources may be left idle and what policies may restore them to productive use. In microeconomics, however, the hypothesis of profit maximization has mutated into an axiom ever ready to deny any allegation of productive inefficiency: If it paid to do something more efficiently, someone would already have seized the opportunity.

Two developments have combined to checkmate this dismissal and make productive efficiency a subject for serious empirical inquiry. The first is theoretical research into market failures involving information costs and asymmetries, agency problems, contract and bargaining costs that cogently limit the ability of utility-maximizing economic decision makers to achieve first-best efficiency (Arrow 1977). The second is an attractive research methodology for measuring productive inefficiency and thereby assessing its extent and testing hypotheses about its determinants. That is the stochastic frontier production function (SFPF), readily estimated from the data on establishments or enterprises that are collected in every country's industrial census.

The concept of technical or productive inefficiency was defined formally by Farrell (1957). It saw little empirical use until the late 1970s, when the SFPF was proposed by Aigner, Lovell, and Schmidt (1977) and by Meeusen and van den Broeck (1977) as a tool for estimating technical efficiency on the basis of assumptions that are both parsimonious and responsive to the typical limitations of actual data. Various applications to particular industries followed, showing that the SFPF gives plausible estimates and can be used to test hypotheses about differences in the efficiency levels of an industry's members.

The SFPF was applied more broadly in the study that precedes this volume, in which Caves and Barton (1990) measured the efficiency of about 350 U.S. manufacturing industries for 1977 and set forth and tested a number of hypotheses about the factors explaining industries' efficiency levels—both factors with direct normative significance and those representing forms of heterogeneity and disequilibrium that have mainly behavioral interest.

In this book appear studies that replicate the U.S. investigation on the manufacturing sectors of five other countries, in the process developing the theory and research methodology in numerous ways. Other papers extend the approach in time and space, observing and explaining how technical efficiency changes over time and comparing its extent and determinants between countries. We feel that we have learned a great deal: procedurally about how this methodology works in practice and what are its strengths and limitations, and substantively about the factors determining efficiency in manufacturing industries and their consistency from country to country.

This chapter summarizes those conclusions. The first section deals with the research methodology, the second and third review the substantive findings, and the last reflects on lines of future research.

1.1 Research Methodology

Farrell (1957) established that any given production process can be inefficient in either or both of two ways.[1] It could be technically inefficient, employing a larger bundle of inputs than the minimum required to obtain the actual output, or it could be allocatively inefficient, selecting the wrong combination of inputs given their relative prices and marginal productivities. Farrell's work led directly to the measurement of efficiency by means of linear programming techniques that simultaneously estimate the frontier and identify the fully efficient units. That approach asks a great deal of the accuracy of the data, however, because in general the number of an industry's units that it deems fully efficient is related to the number of parameters in the production function being fitted. Should a spurious observation (due to a data error or some other sort of unsuitability) land in the efficient set, the consequent measurement of inefficiency could be substantially in error.

The SFPF escaped this difficulty by formulating the production function for statistical estimation as

$$y = f(x) \exp(v - u),$$

where y is output, x is a vector of inputs, and the error term is composed of two elements. The usual normally distributed v represents random disturbances, measurement errors, and minor omitted variables affecting the deterministic kernel. The other component $u \geq 0$ represents some one-sided distribution of technical efficiency beneath the frontier. Thus a particular data point might lie above the estimated regression plane because of a "lucky" random component; it might lie beneath the plane either due to an unlucky draw or because it is technically inefficient.[2] The simple intuition about the procedure is that if the model correctly identifies an industry's inefficiency, the residuals from its fitted production function will be negatively skewed. The second and third moments of the residuals that are used to calculate skewness are also the source of the measures of (in)efficiency that are obtained from the SFPFs. Specifically the moments yield the estimated standard deviations of the v and u components of the composed residuals (σ_v and σ_u), from which measures of technical (in)efficiency are calculated. In our analyses these measures then become dependent variables in cross-sectional (interindustry) regression models to test hypotheses about technical efficiency.

Before we identify the efficiency measures, it is important to recognize that the procedure does not always work because σ_v and σ_u cannot always be calculated. The following chapters refer to failures of two types. A type I failure occurs when the skewness of the residuals is positive, implying that $\sigma_u < 0$. A type II failure occurs when the third moment of the residuals is so large relative to the second that it implies $\sigma_v < 0$. The logic of the composed-error approach suggests the conjecture that industries subject to type I failures are likely to harbor little inefficiency, with the positive skewness reflecting an oddity of the random residuals v in the particular sample. This conjecture tempts the researcher to retain such industries in the interindustry analysis of efficiency's determinants and to score them as "fully efficient." Caution argues against this choice, however, because on another interpretation an industry with positively skewed residuals could be highly inefficient.[3] Industries with type II failures might be regarded as extremely inefficient because the one-sided component swamps the prevalent random noise, but again this interpretation is not necessary.

In the absence of these failures several measures of efficiency can be calculated and in fact are defined and used in the following chapters. The most popular has been expected technical efficiency (based on Lee and Tyler 1978), an absolute measure that depends only on σ_u and lies in the (0, 1) interval. Closely related to it is average technical inefficiency, a measure based on σ_u normalized by the mean of the dependent variable (or the estimated mean on the efficient frontier). A term used in deriving the SFPF that has been taken over as an efficiency measure is λ, which normalizes σ_u by the standard deviation of the normally distributed component of the error σ_v. Finally, since these three measures are all lost when a type I or II failure occurs in estimation, skewness has been used directly as an efficiency measure by making the assumption that the likelihood of no substantial technical inefficiency increases with the value of positive skewness.[4] That assumption is related to the assumption that industries with type I failures are fully efficient and just as much open to challenge. Nonetheless, some studies in this volume make apparently successful use of skewness as an efficiency measure.

Although the focus of this project is empirical rather than theoretical or methodological, the papers do make some noteworthy extensions of the methodology. Some of these concern the statistical distribution chosen to represent the inefficiency component of the residuals. Previous research employed the half-normal and the exponential distributions only for their simplicity and tractability. Akio Torii, however, shows (subsection 2.3.1) that such distributions of inefficiency can be derived from models of specific processes that generate inefficiency. One model turns on capital-vintage effects and fixed costs of replacement and relates the distribution of inefficiency to the distribution among production units of the slippage of capital productivity below the frontier. The other model turns on organizational inefficiency: if inefficiency tends to creep upward in the absence of specific managerial effort (a fixed cost) to combat it, the specific form of a distribution of inefficiency can again be derived.

An objection made to the use of the half-normal (or exponential) distribution to depict inefficiency is the implicit assumption that the modal level of inefficiency is zero. There is no reason, however, why the number of units that are fully efficient should exceed the number that exhibit any given positive amount of inefficiency. Chris Harris (subsection 5.2.4) makes operational the use of a truncated normal distribution rather than a half-normal, allowing the modal level of inefficiency to be strictly positive and identifying it in the estimation. (The empirical results are mentioned below.)

An econometric problem considered by Akio Torii is the possibility of bias in the estimator of σ_u obtained by the convenient (and consistent) corrected ordinary least squares (COLS) method. One problem arises from the bounded value of the third moment of the residuals from which σ_u is estimated, which can cause asymmetrical biases near the boundaries of the region in which efficiency can be estimated (that is, where the results edge toward a type I or II failure). Because the biases depend on the estimated values of σ_u and σ_v, he is able to develop (subsection 2.2.6) a way to relate the true value of σ_u to these estimates, yielding an adjusted measure of expected technical efficiency that is used throughout his empirical work. The other problem arises from the effect on the estimator of σ_u of nonlinear terms in the production function, inherent in use of the popular trans-logarithmic function to estimate technical efficiency. Although these indeed bias the COLS estimator, his Monte Carlo analysis (subsection 2.2.5) shows the magnitudes of the biases to be trivial.

Another methodological consideration is developed by John Baldwin (chapter 7). Unable to compute proper SFPFs for Canadian industries because data on plant-level capital stocks are lacking, he employs a simpler research strategy that harks back to Timmer (1971). Working with output per person employed, he assumes that some fraction (10 to 40 percent) of each industry's total output emanates from efficient plants. He can then calculate average efficiency as the ratio of the weighted average of output–employment ratios for the remaining plants that are deemed inefficient to the corresponding weighted average for plants on the assumed frontier. Should this technique perform well compared to the SFPF, it can claim two virtues. First, it does not remove (as the SFPF does) the influence of any scale inefficiency (e.g., suboptimal-scale plants) and diverse input combinations (possibly due to factor-market imperfections) from the raw variance of units' productivity levels before estimating technical inefficiency. These kinds of inefficiencies remain included in the measure of technical efficiency, and thus hypotheses about their interindustry determinants can be tested. Second, his method is much simpler than estimating SFPFs. In general, in this project we showed respect for the hazard of undue complexity; chapters 4 and 9 address the possibility that the simple dispersion of plant-productivity levels within an industry (the second moment of production-function residuals) is a more productive object of analysis than their skewness (the third moment).

1.2 Evaluating Estimated Levels of Efficiency

Now we turn to the empirical results of estimating SFPFs for manufacturing industries in six industrial countries, phase 1 in the project's jargon. A general strategy for the tactical choices in estimating frontier production functions was first worked out in the U.S. study (largely by experimentation on a small panel of industries). The results of this experience were made available to researchers in the other countries, but they did not necessarily seek to replicate the U.S. procedures in all respects. Certain features of the research design are common to all:

1. A preference for the translogarithmic production function was shared among the researchers. With each national team fitting production functions to many industries, practicality dictated selecting a specific form a priori. Because of its flexibility and the importance of enveloping the data for each industry well, the translog seemed clearly the weapon of choice.

2. Each team implemented a generally similar set of rules for editing data. Our impression is that the national census organizations vary considerably in the resources they devote to checking the correctness and consistency of data they receive from individual establishments, at least prior to their aggregation. The data editing rules sought to exclude establishments that might be reporting accurate data but be unsuitable for analysis (e.g., start-ups) as well as establishments reporting data that are internally inconsistent or simply wildly implausible. For Australia the data-editing rules were found clearly to improve the quality of the results (chapter 5). For the United States the proportion of reporting establishments excluded from the analysis proved unrelated to how well we could explain an industry's efficiency in cross section, suggesting that varying the rates of deletion did not bias the measures of efficiency (Caves and Barton 1990, 54–58, 106–107).[5]

3. The measure of plants' outputs used to estimate SFPFs could be either value of output or value added. An argument can be made for and against each. The research teams therefore carried out their computations using both of them, getting dissimilar yields of successful estimates and mean efficiency levels.[6]

4. The year chosen for the analysis was 1977 or 1978. This choice was purely fortuitous. When the U.S. project began, 1977 was the most recent year in which the full Census of Manufactures had been completed. That or an adjacent year was then chosen by the other investigators. Although macroeconomic conditions in the 1970s were undesirably disturbed in all countries, no obviously superior year was available at the time the projects were started.

Although these common decisions were expected to make the results of the studies basically comparable, many differences remain that stem from irreparable differences in the countries' methods of gathering and reporting their data. A principal difference is in the inclusion of small establishments. In the U.S. data establishments smaller than 250 employees are sampled, while the other primary data sets truncate the small establishments at some (lower) threshold. This truncation was not thought undesirable because, for small units, the proportional amount of random noise in the data was expected to decline sharply with establishment size, and SFPFs were in most cases estimated from unweighted (not size-weighted) establishment data. For Korea efficiency measures were calculated both including and excluding establishments with fewer than 20 employees. Estimated mean efficiency levels rise when they are excluded; more important and dismaying, the interindustry correlations between efficiency measures with and without the small establishments are low, at most 0.555 for expected technical efficiency (table 3.11).

Other differences among the studies arise in the measurement of the inputs of factors of production. Establishment-level capital-stock data are available for Japan and Korea as well as the United States, although their values were surely distorted by the inflation of the 1970s. Elsewhere these stocks were approximated by allocating company-level stocks (Australia) or substituting data on flows of capital expenditures (Britain). With regard to labor inputs, numerous small differences in reporting procedure clearly exist but were not investigated because we expected that they could neither be controlled nor their effects predicted.

An important comparative feature of the country studies is the incidence of type I and II estimation failures. Table 1.1 summarizes the numbers of industries available for analysis in each country and the incidence of estimation failures. The success rate varies markedly between a high of 80 percent (United States) and a low of 41 percent (Japan). The percentage of type I failures (that arguably might represent highly efficient industries) ranges more narrowly, while the prevalence of type II failures varies wildly from none to one-third. The pattern is related in no obvious way to the characteristics of the countries or their data. A simulation analysis by Akio Torii (summarized in subsection 4.1.5) suggests that the incidence of type II failures should be quite low. Despite this diversity we feel that the minority status of type I failures is important in one regard. Consider the following insidious null hypothesis: The composed-error model fundamen-

Table 1.1
Numbers of industries and rates of success in estimating stochastic frontier production functions

Country	Industries analyzed	Estimation failures		Successes
		Type I	Type II	
Australia	140	49	0	91
	100%	35%	0%	65%
Japan	351	86	121	144
	100%	25%	34%	41%
Korea[a]	242	85	29	128
	100%	35%	12%	53%
United Kingdom	151	48	31	72
	100%	32%	20%	48%
United States[b]	434	87	0	347
	100%	20%	0%	80%

Note: Some studies report alternative estimations; the one chosen here involves the deletion of outlying observations, inclusion of control variables, and use of ordinary least squares estimation.
a. Based on gross output per employee. Other countries based on value added per unit of labor input (denominator varies).
b. Taken from Caves and Barton (1990), table 4.1. Other data are from chapters 2–5.

tally fails to capture technical inefficiency, and the third moments of the SFPFs represent nothing more than sample-based skewness in residuals that are normal in the population (but not necessarily in the particular sample). If that hypothesis were correct, half the industries should incur type I failures, and the putative measures of technical efficiency would be meaningless. Ultimately it is our ability to explain cogently the inter-industry differences in measured technical efficiency that lets us proclaim this fearsome dragon to be slain. Meanwhile the fact that type I estimation failures are overall in the minority is a substantial comfort.

An international study of efficiency naturally seeks to learn which country is the most (least) efficient. We came reluctantly to the conclusion that SFPFs do not yield a reliable answer to this question. First, for any given country the mean value of efficiency was found to vary greatly with the efficiency measure chosen and the tactical choices made in estimating the SFPFs. Second, we do not know what to make of industries with estimation failures. Third, neither the data nor the estimation procedures were fully standardized between countries, and we have no way to know whether the differences substantially affected mean estimated efficiency. The question of comparative efficiency is pursued in detail (but to agnostic conclusions) for Japan and the United States in subsection 11.3.2.

What we have learned at this stage from estimating efficiency from SFPFs seems rather unpromising. The yields of estimates are somewhat puny, and the patterns differ considerably among the countries although without yielding any compensating conclusions about international differences in efficiency. Fortunately the situation changes for the better as we turn to the project's second phase, investigating the determinants of interindustry differences in efficiency.

1.2 Determinants of Interindustry Differences in Efficiency

Several families of hypotheses about the determinants of an industry's efficiency level were set forth in Caves and Barton (1990, ch. 5).[7] Although the international project sheds new light on some individual hypotheses and the variables that embody them, that taxonomy still stands. The major families of hypotheses are:

1. *Competitive conditions.* Many hypotheses connect competitive conditions to efficiency. High concentration permits inefficiency to persist, should individual firms' managers not be optimally motivated to eliminate it. Incomplete collusive bargains among firms in a concentrated industry can induce rent-seeking enlargements of those outlays that affect a competitor's position in nonprice rivalry. Finally, when the number of market participants is small, there are fewer agents to experiment and try for improved ways of doing things, and fewer peers from whom to learn.

2. *Organizational factors.* Modern theory of corporate governance provides explanations why firms may not be fully motivated to minimize costs. The slippage can arise from second-best principal-agent relations that distort the use of resources (e.g., collective-bargaining agreements) or from bargaining costs that preclude rectification. Although this project is not designed to test hypotheses about differences in efficiency between firms, it can address any firm-based differences that vary from industry to industry due to observable factors (e.g., the prevalence of trade-union organization).

3. *Structural heterogeneity.* Technical efficiency estimated from the SFPF can pick up many sorts of heterogeneity in the revenue-productivity levels of an industry's plants or firms. These include product differentiation and