THE IMPACT OF GOVERNMENT INSTITUTIONS ON DAIRY FARM EFFICIENCY: A COMPARISON OF ONTARIO AND NEW YORK DAIRY FARMS

A Thesis

Presented to

The Faculty of Graduate Studies

of

The University of Guelph

by

PETER SLADE

In partial fulfilment of requirements

for the degree of

Master of Science

August, 2011

© Peter Slade, 2011
ABSTRACT

THE IMPACT OF GOVERNMENT INSTITUTIONS ON DAIRY FARM EFFICIENCY: A COMPARISON OF ONTARIO AND NEW YORK DAIRY FARMS

Peter Slade
University of Guelph, 2011

Advisor: Getu Hailu

The impact of supply management on dairy farm efficiency in Canada is of increasing academic and public interest. This thesis decomposes the cost efficiency of Ontario and New York dairy farms into its technical and allocative components. Triangulating between linear programming and econometric techniques, we find New York farms to be more cost efficient than Ontario farms, with allocative efficiency accounting for the majority of this difference. The difference in inefficiency is in large part caused by the overcapitalization of Ontario farms, as well as their reliance on homegrown feed. We argue that the system of supply management could theoretically be at the root of this inefficiency.
Acknowledgments

While graduate students are frequently admonished to take ownership of their thesis, at the terminus of my research I cannot help but see my thesis as a product, not of individual industry, but of the conversations I have had, the instruction I have received and the intellectual environment I have lived in over the past two years. I would first like to thank my advisor, Dr. Getu Hailu, whose mentorship and friendship has been a benefit, not only to this work, but also to my own personal development.

I am also indebted to the other faculty members of FARE for their guidance throughout my research. I would especially like to thank the members of my advisory and examination committees, Drs Alan Ker, Karl Meilke, Brady Deaton, and John Cranfield. Throughout my degree I have relied on my fellow graduate students for advice, conversation and friendship; special mention is owed to Alex Cairns, James Bryan and Jessica Rosenberg.

Thank-you to my parents, Mari-Beth and Laura, who have supported me through my education, as they have in many other (less successful) ventures. Finally, I must thank my daughter Alexandra for being a source of constant joy. Alex, if have raised you at all properly, by the time you are old enough to read this you will know the marginal benefits of reading beyond this page are far outweighed by the marginal costs.
Table of Contents

List of Tables vi
List of Figures vii

1 Introduction 1

2 Policy Review 5
  2.1 Rationale for government intervention . . . . . . . . . . . . . . . . . . 5
  2.2 Institutions and regulations in Ontario . . . . . . . . . . . . . . . . . 6
    2.2.1 Pricing formula . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
    2.2.2 Production restrictions . . . . . . . . . . . . . . . . . . . . . . . 8
    2.2.3 Trade policies . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
    2.2.4 Effect of Canadian dairy policies . . . . . . . . . . . . . . . . . 11
  2.3 Institutions and regulations in New York state . . . . . . . . . . . . . 12
    2.3.1 Federal Milk Marketing Orders . . . . . . . . . . . . . . . . . . . 13
    2.3.2 Dairy Price Support Program . . . . . . . . . . . . . . . . . . . . 14
    2.3.3 Milk Income Loss Contract . . . . . . . . . . . . . . . . . . . . . 16
    2.3.4 Trade policies . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
    2.3.5 Effect of U.S. dairy policies . . . . . . . . . . . . . . . . . . . . 18
  2.4 Regional comparison . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

3 Dairy Farm Efficiency - Economic Theory and a Brief Literature Review 22
  3.1 The concept of economic efficiency . . . . . . . . . . . . . . . . . . . 22
  3.2 Regulation and efficiency . . . . . . . . . . . . . . . . . . . . . . . . 24
  3.3 Two practical considerations . . . . . . . . . . . . . . . . . . . . . . . 27
    3.3.1 The objective of firms . . . . . . . . . . . . . . . . . . . . . . . . 27
    3.3.2 Quota values . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
  3.4 Existing literature on dairy farm efficiency . . . . . . . . . . . . . . . 29
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Variables</td>
<td>35</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Milk</td>
<td>38</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Crop (output)</td>
<td>38</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Livestock (output)</td>
<td>39</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Feed</td>
<td>40</td>
</tr>
<tr>
<td>4.1.5</td>
<td>Labour</td>
<td>41</td>
</tr>
<tr>
<td>4.1.6</td>
<td>Capital</td>
<td>43</td>
</tr>
<tr>
<td>4.1.7</td>
<td>Other</td>
<td>44</td>
</tr>
<tr>
<td>5 Non-Parametric Estimation</td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>5.1</td>
<td>Data envelopment analysis</td>
<td>45</td>
</tr>
<tr>
<td>5.2</td>
<td>Statistical underpinning of DEA</td>
<td>48</td>
</tr>
<tr>
<td>5.3</td>
<td>Bootstrapping methodology</td>
<td>49</td>
</tr>
<tr>
<td>5.4</td>
<td>Bias correction</td>
<td>53</td>
</tr>
<tr>
<td>5.5</td>
<td>Confidence intervals</td>
<td>54</td>
</tr>
<tr>
<td>5.6</td>
<td>DEA and panel data</td>
<td>55</td>
</tr>
<tr>
<td>5.7</td>
<td>Outlier detection</td>
<td>61</td>
</tr>
<tr>
<td>5.8</td>
<td>Results</td>
<td>63</td>
</tr>
<tr>
<td>5.9</td>
<td>Input usage</td>
<td>68</td>
</tr>
<tr>
<td>5.10</td>
<td>Sensitivity analysis</td>
<td>68</td>
</tr>
<tr>
<td>5.11</td>
<td>Second stage regression</td>
<td>70</td>
</tr>
<tr>
<td>5.12</td>
<td>Scale economies</td>
<td>75</td>
</tr>
<tr>
<td>6 Parametric Estimation</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>6.1</td>
<td>The theory of the distance function</td>
<td>79</td>
</tr>
<tr>
<td>6.2</td>
<td>Empirical analysis</td>
<td>82</td>
</tr>
<tr>
<td>6.3</td>
<td>Second stage analysis</td>
<td>89</td>
</tr>
<tr>
<td>6.4</td>
<td>Correlation across approaches</td>
<td>89</td>
</tr>
<tr>
<td>7 Conclusion</td>
<td></td>
<td>91</td>
</tr>
<tr>
<td>7.1</td>
<td>Limitations and suggestions for future research</td>
<td>93</td>
</tr>
<tr>
<td>7.2</td>
<td>Policy recommendations</td>
<td>95</td>
</tr>
<tr>
<td>A</td>
<td>FMMO Pricing</td>
<td>109</td>
</tr>
<tr>
<td>B</td>
<td>Map of Northeast Marking Order</td>
<td>110</td>
</tr>
<tr>
<td>C</td>
<td>Density Plots of Efficiency Scores</td>
<td>111</td>
</tr>
<tr>
<td>D</td>
<td>Derivation of Cost Function Using Duality</td>
<td>116</td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Tariff rate quota for selected dairy products in Canada (2008)</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Efficiency, scale elasticity and costs of US dairy farms (Moshien and Lovell, 2009)</td>
<td>33</td>
</tr>
<tr>
<td>4.1</td>
<td>Ontario dairy farms included in analysis by year of entry</td>
<td>36</td>
</tr>
<tr>
<td>4.2</td>
<td>New York dairy farms included in analysis by year of entry</td>
<td>36</td>
</tr>
<tr>
<td>4.3</td>
<td>Descriptive statistics of variables used in this study</td>
<td>37</td>
</tr>
<tr>
<td>4.4</td>
<td>Aggregation of Crops</td>
<td>39</td>
</tr>
<tr>
<td>4.5</td>
<td>Assets of Ontario dairy farms as a proportion of assets of New York dairy farms</td>
<td>43</td>
</tr>
<tr>
<td>5.1</td>
<td>Efficiency scores from metafrontier Model</td>
<td>63</td>
</tr>
<tr>
<td>5.2</td>
<td>Efficiency scores from disaggregated Model</td>
<td>64</td>
</tr>
<tr>
<td>5.3</td>
<td>Results of Malmquist index</td>
<td>65</td>
</tr>
<tr>
<td>5.4</td>
<td>Efficiency scores from intermediate model</td>
<td>66</td>
</tr>
<tr>
<td>5.5</td>
<td>Correlation and rank correlation of the three DEA models</td>
<td>67</td>
</tr>
<tr>
<td>5.6</td>
<td>Mean proportional change in inputs necessary to achieve cost efficiency</td>
<td>68</td>
</tr>
<tr>
<td>5.7</td>
<td>Sensitivity of efficiency scores to changes in interest rates</td>
<td>69</td>
</tr>
<tr>
<td>5.8</td>
<td>Independent variables in truncated regression</td>
<td>71</td>
</tr>
<tr>
<td>5.9</td>
<td>Results of truncated regression on DEA scores</td>
<td>74</td>
</tr>
<tr>
<td>5.10</td>
<td>Average efficiencies from non-parametric estimation by farm size</td>
<td>76</td>
</tr>
<tr>
<td>6.1</td>
<td>Likelihood ratio tests comparing the translog and Cobb-Douglas functional forms across distributional assumptions</td>
<td>83</td>
</tr>
<tr>
<td>6.2</td>
<td>Likelihood ratio tests comparing the general model (truncated normal with inefficiency effects) to other distributional assumptions</td>
<td>84</td>
</tr>
<tr>
<td>6.3</td>
<td>Results of maximum likelihood estimation using inefficiency effects model</td>
<td>85</td>
</tr>
<tr>
<td>6.4</td>
<td>Efficiency scores for the Cobb-Douglas model derived using duality</td>
<td>87</td>
</tr>
<tr>
<td>6.5</td>
<td>Average efficiencies from parametric estimation by farm size</td>
<td>88</td>
</tr>
<tr>
<td>6.6</td>
<td>Results of truncated regression on allocative efficiency scores</td>
<td>89</td>
</tr>
<tr>
<td>6.7</td>
<td>Correlation and rank correlation of preferred DEA and econometric models</td>
<td>90</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Distribution of production quota in Ontario</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Milk Price Support Program purchases (1990-2008)</td>
<td>16</td>
</tr>
<tr>
<td>2.3</td>
<td>Milk income loss contract trigger price and payment (2001-2009)</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Milk income loss contract expenditure (2002-2009)</td>
<td>18</td>
</tr>
<tr>
<td>2.5</td>
<td>Mean annual milk price in Ontario and New York (2000-2009)</td>
<td>20</td>
</tr>
<tr>
<td>2.6</td>
<td>Producer subsidy equivalent for Canadian and US dairy industries (2000-2009)</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>A graphical representation of efficiency</td>
<td>25</td>
</tr>
<tr>
<td>C.1</td>
<td>Density plot of technical efficiency scores from DEA model</td>
<td>112</td>
</tr>
<tr>
<td>C.2</td>
<td>Density plot of allocative efficiency scores from DEA model</td>
<td>113</td>
</tr>
<tr>
<td>C.3</td>
<td>Density plot of technical efficiency scores from econometric model</td>
<td>114</td>
</tr>
<tr>
<td>C.4</td>
<td>Density plot of allocative efficiency scores from econometric model</td>
<td>115</td>
</tr>
</tbody>
</table>
Chapter 1  Introduction

High price levels and price increases [under the Canadian system of supply management], regardless of steps taken to reduce production costs, provide weak incentives to reduce those costs further, and this is even more true with prices based on median - rather than most efficient - producer costs. (Goldfarb, 2009, 33)

Generally the critics of supply management contend that the industries are inefficient, new product development hasn’t been forthcoming, and consumers are paying too high a price for commodities produced under supply management regulations. (Schmitz and Schmitz, 1994, 125)

Public commentators and academics have suggested that supply management in the Canadian dairy sector breeds cost inefficiency by sheltering producers from competition, and basing price changes on cost of production.\(^1\) In contrast, others argue that fixing output and reducing price risk, allow producers to focus on the achievement of cost minimization. This thesis speaks to this tension by examining the technical, allocative and cost efficiency of Ontario and New York dairy farms. The economic efficiency of Canadian dairy farms has not suffered from academic neglect (Weersink, Turvey, and Godah, 1990; Romain and Lambert, 1995; Mbaga et al., 2003; Hailu, Jeffrey, and Unterschultz, 2005). However, efficiency is a

\(^1\)Initial academic interest in the efficiency of supply management centred on "transfer inefficiency"; the measurement of deadweight loss created when transferring surplus from consumers to producers. This is a separate issue, whose literature is surveyed by Schmitz and Schmitz (1994).
benchmarking technique; studies that examine only the efficiency of dairy farms within Canada, are unable to derive any inferences regarding the effect of supply management on productive efficiency. An ideal study would compare the efficiency of producers in such a system with that of producers who exist in a comparable environment but operate without any marketing restrictions. Unfortunately, such a counterfactual does not exist. Nevertheless, New York State, whose dairies are subject to a reasonably similar climate and economic environment, though very different marketing regulations\(^2\), serves as a beneficial comparison.

The technical efficiency of Ontario and New York dairy farms was previously studied by Haghiri, Nolan, and Tran (2004), who found significant, though modest differences, between the two regions. This thesis significantly expands the analysis of Haghiri, Nolan, and Tran (2004) by employing more fulsome use of the data sets, triangulating between efficiency methods, and most critically, estimating cost efficiency (decomposed into technical and allocative components.) We find that the major differences between the two regions lie in their allocative efficiency.

The Canadian system of supply management sets an administered price for milk, based in part on estimated cost of production. This price is maintained by restricting production (through milk quota) and imports. The literature identifies three different ways in which this system could introduce inefficiency; through cost of production pricing, stifling competition and hindering expansion.

Goldfarb (2009) maintains that the pricing formula, engenders inefficiency by

\(^2\)We will argue that the policies in place in New York state are more market oriented than those in Ontario.
providing a disincentive to reduce costs. The administered price, typically provides even inefficient farmers with a stable income, thus supply management can be seen as sheltering inefficient producers from competition, and allowing otherwise efficient producers to pursue objectives other than cost minimization. Furthermore, given evidence of economies of scale in the dairy industry (Mosheim and Lovell, 2009; Tauer and Mishra, 2006), the quota system (which necessitates significant capital outlays for expansion) may prevent producers from attaining their most efficient scale.

Supporters of supply management, undergirded by neo-classical economic assumptions, have offered a repost to these criticisms. The administered price is based only in part on cost of production thus all farmers can gain by reducing their costs. Furthermore, individual farmers benefit greatly when they reduce costs relative to others. Until recently, most provincial producer organizations have allowed quota to be bought and sold freely on quota exchanges, and have allowed quota to serve as collateral for loans; permitting producers to expand without tying up capital, and providing a mechanism for removing inefficient producers. This study examines the efficiency of Ontario and New York farms using both linear programming and econometric methods. Both methods find New York farms to be significantly more technically and allocatively efficient than Ontario farms, with allocative efficiency accounting for the bulk of the overall difference. This discrepancy in allocative efficiency is explained in large part by the

3 Price changes for fluid milk in Ontario is based 40% on the cost of production
4 Recently, quota transactions have been subject to a price ceiling. This issue is further discussed in Chapter 2.
over-capitalization of Ontario farms, as well as their reliance on homegrown, as opposed to purchased feed.

The results of this study are of primary interest to dairy farmers themselves; in addition to outlining the relative standing of farms in the two regions, secondary analysis informs producers of the determinants of efficiency (such as technology and farm size), and the changes in inputs required to achieve efficiency. The findings are also relevant to those shaping dairy policy in both Canada and the United States. Finally, inasmuch as cost efficiency is reflected in the price of milk, the relative efficiency of Ontario dairy farmers is of interest to Canadian dairy processors and consumers.

The remainder of this thesis proceeds as follows: chapter two surveys dairy marketing policies in Ontario and New York. Chapter three introduces the theoretical basis for the study and reviews the literature on dairy farm efficiency. Chapter four outlines the data used in this study. Chapters five and six explain the linear programming and econometrics models used in analysis, and provide the results of this analysis. Finally, chapter seven contains a discussion of results, limitations of the study, policy recommendations and suggestions for further research.
Chapter 2  Policy Review

2.1  Rationale for government intervention

Even in the highly regulated agricultural industry, dairy is widely considered to be subject to more regulation than other commodities. As Bailey (1997) notes, milk is highly perishable and cannot be stored in fluid form much longer than one or two days without being processed. This problem is exacerbated by the “mismatch” of supply and demand for dairy products. Supply of milk is relatively constant, increasing slightly in the summer months, while demand is more variable and inversely correlated with supply. It has been argued that government intervention can curtail the market power of processors\(^1\) while reducing shortages, price fluctuations and/or spoilage. Furthermore, governments in both Canada and the United States (U.S.) have stated that providing producers a “fair return” is a primary objective of their dairy policies (Goldfarb, 2009; Federal Milk Order Study Committee, 1962).

\(^{1}\)Prior to heavy government regulation processors were often able to exercise near monopoly power, forcing farmers to accept arbitrary price reductions and quantity restrictions imposed by processors, as they had no other immediate market for their milk (McCormick, 1968).
2.2 Institutions and regulations in Ontario

The current organization of milk marketing in Ontario can be traced to local producer organizations that date to the 1870s (McCormick, 1968). As milk prices fell in the 1930s, the government encouraged the consolidation of these organizations into provincial producer associations. These associations organized producers based on the final use of their milk; in the early 1930s associations were founded representing the interests of fluid milk, cheese and condensed milk producers (producers supplying milk used in cream would later organize in 1946.) The most important role of these organizations was to negotiate price formulas with processors in their sector (McCormick, 1968, pp. 159–162).

Fluid milk, being more perishable, was subject to heavier regulation such as the Ontario Milk Control Act of 1934. It was also the first sector to experiment with quotas. In 1933 the Toronto Milk Producers Association signed the initial agreement with distributors to implement a quota system for milk sales, leading to a province-wide quota system, and pricing formula in 1954.

The differing marketing systems for fluid and industrial milk led to large price differentials, and contention between dairy farmers in the two marketing streams. This conflict and the generally scattershot organization of producers, led the provincial government to create the Ontario Milk Marketing Board (OMMB) with the view of consolidating the dairy industry under similar provincial regulations. The OMMB was established by the Milk Act of 1965, absorbing the powers of the
milk, butter and cheese producer organizations\(^2\). In its first five years of operation the OMMB became the sole buyer of raw milk in Ontario, establishing a single price for fluid milk in Southern and Northern Ontario, and taking over operation of the quota system for fluid milk (called Group I quota) (Scullion, 2004, pp. 12–14).

The federal government also sought to reshape the regulatory environment by creating the Canadian Dairy Commission (CDC) in 1965-1966. The CDC assumed responsibility for the federal government’s main policy tools: an offer-to-purchase plan (whose prices determined the industrial milk price in Ontario) and the direct subsidy program (Scullion, 2004, pp. 1–39). Under the direction of the CDC a national quota system for industrial milk emerged in the early 1970s. Quota was maintained at a level that achieved a target price determined by the federal government. In 2002 the federal government eliminated its direct subsidies to dairy farmers, following gradual reductions throughout the 1990s, leaving the quota system as the main government institution in the dairy sector (Scullion, 2004, p. 147).

The resulting regulatory system has come to be termed “supply management” and is undergirded by three pillars: a pricing formula, production restrictions and trade policies.

### 2.2.1 Pricing formula

All milk in Ontario is purchased by the Dairy Farmers of Ontario (DFO) who then sells it to processors. While all milk produced is subject to the same sanitary

---

\(^2\)The cream board maintained its independence until merging with the OMMB to form the Dairy Farmers of Ontario in 1995.
requirements, it is assigned to one of four classes\textsuperscript{3} based on its final utilization. The price that processors pay is based on the class of milk they use, while farmers receive a weighted average (blended) price of all milk sold (Dairy Farmers of Ontario, 2009). Payments are made based on the composition of milk, with the DFO specifying individual prices for butterfat, protein and solid non-fats. The target price for fluid milk is set by the DFO in accordance with a formula that bases 40\% of the price change on the cost of dairy production, 30\% on personal disposable income and 30\% on the Consumer Price Index (The Milk Producer 2009, 40.) The original base price was decided based on stakeholder negotiations (Goldfarb, 2009, p. 8).

The DFO sets the price of milk in other classes based on the CDC’s support price for butter and skim milk powder. This support price itself rests on cost of production estimates, input from stakeholders and the experience and judgment of the CDC commissioners (Scullion, 2004, p. 175).

\subsection{2.2.2 \hspace{1em} Production restrictions}

The target prices discussed above are achieved by restricting production through dairy quota. In Ontario one unit of quota entitles a producer to produce one kilogram of butterfat daily (roughly the production of one cow.) Quota, initially allocated to producers on the basis of historical production, can now be bought and sold on a quota exchange operated by the DFO (Dairy Farmers of Ontario, 2009). Since August of 2009 the DFO has imposed a ceiling on the price of quota which

\textsuperscript{3}Broadly, the four classes are: fluid milk, soft dairy products, cheese and hard dairy products.
now sits at $25,000. To maintain target prices, quota is periodically adjusted to reduce milk surpluses or shortages.

The CDC adjusts the quantity of industrial milk quota every two months based on the domestic demand for, and planned exports of, industrial milk (Canadian Dairy Commission, 2009). Changes in national quota are allocated to provinces using a formula that bases 90% of the allocation on population, and 10% on historical production (Scullion, 2004, 184). Monitoring and adjustment of fluid milk quota in Ontario is governed by an interprovincial agreement, termed the P5, whose other signatories are Quebec and the maritime provinces. Provincial adjustments to quota are passed on to provincial producers on an equal percentage basis (Dairy Farmers of Ontario, 2009, p. 2). Figure 2.1 outlines the institutional processes involved in setting the price and quantity of milk.

\footnote{The ceiling was initially set at $25,500 in August 2009 and was reduced by $100 per month, reaching its current level in January 2010. The average monthly clearing price in the twelve months before the imposition of the price ceiling was $29,980.67. The minimum monthly clearing price was $28,300, while the maximum was $33,115 (Dairy Farmers of Ontario, 2010)}
Figure 2.1: Distribution of production quota in Ontario

The CDC sets a target price for industrial milk based on:
- Production costs
- Consultation with various stakeholders
- Their own experience and judgement

National industrial milk quota is increased or decreased such that the target price will be achieved. Quota is allocated to provinces in accordance with a formula which bases 90% of the distribution to population and 10% to historical production

The P5 sets the price for fluid milk based on:
- 40% cost of production
- 30% consumers ability to pay
- 30% consumer prices index

Provincial fluid milk quota is increased or decreased such that the target price is achieved

Provincial marketing boards combine their increases/decreases in fluid and industrial quota and allocate this to farmers

To ensure the target price is achieved the CDC "smoothes" the price through purchases of butter and skim milk powder
Table 2.1: Tariff rate quota for selected dairy products in Canada (2008)

<table>
<thead>
<tr>
<th>Product</th>
<th>Quota as a share of domestic consumption</th>
<th>In-quota tariff</th>
<th>Over-quota tariff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluid milk</td>
<td>2.3%</td>
<td>7.5%</td>
<td>241%</td>
</tr>
<tr>
<td>Yogurt</td>
<td>0.1%</td>
<td>6.5%</td>
<td>238%</td>
</tr>
<tr>
<td>Butter</td>
<td>3.7%</td>
<td>6.5%</td>
<td>299%</td>
</tr>
<tr>
<td>Cheese</td>
<td>5.0%</td>
<td>1.0%</td>
<td>246%</td>
</tr>
<tr>
<td>Ice Cream</td>
<td>0.1%</td>
<td>6.5%</td>
<td>277%</td>
</tr>
<tr>
<td>Skim milk powder</td>
<td>0.0%</td>
<td>6.5%</td>
<td>277%</td>
</tr>
</tbody>
</table>

Source: Goldfarb (2009)

2.2.3 Trade policies

Trade barriers are central to the system of supply management as they preclude foreign producers from exporting at artificially high domestic dairy prices. Previous to the completion of the Uruguay Round Agreement on Agriculture these barriers took the form of import quotas. Since signing the Agreement in 1995, Canada has instituted a tariff rate quota regime. Under this structure a certain percentage of dairy products are allowed into the country paying a relatively low in-quota tariff. Additional imports are subject to a prohibitively high over-quota tariff as illustrated in table 2.1 (Barichello, 1999, p. 48).

2.2.4 Effect of Canadian dairy policies

Of interest to this study is the degree to which government institutions restrict competition; allowing uncompetitive firms to remain in the market, or allowing otherwise competitive firms to pursue goals other than cost-minimization.

According to Cairns, Meilke, and Bennett (2010), the marginal and average costs of
production for Ontario dairy farms in 2009 were $31.39/hl and $46.90/hl, respectively, while the "mail-box" price of milk received by farmers was $66.36. This would imply that an Ontario dairy farmers would need to be quite inefficient indeed (compared to their peers) for their operation to be unprofitable. Of course this cost estimate abstracts from the cost of quota itself. If one were given quota _gratis_, either from initial government distribution or through inheritance, this would not be an issue. Conversely, if quota is efficiently priced\(^5\) (assuming away risk), then a farmer who purchased quota would not benefit from quota rent.

### 2.3 Institutions and regulations in New York state

As in Ontario, milk marketing in New York at the beginning of the twentieth century was largely rationalized by producer co-ops who were able to counter the market power of distributors and retailers. During the 1930s however, the incomes of dairy farmers fell as the economic depression caused a loss of purchasing power and a precipitous drop in demand for dairy products. According to Manchester (1983, p. 21), the average price of milk fell 31% between 1929 and 1932. The collapse of milk prices and subsequent unrest among dairy farmers forced the government into direct intervention in the dairy market — passing the Agricultural Adjustment Act of 1933. While this Act was meant to be temporary, it instituted two major dairy programs that are still in existence today: Federal Milk Marketing Orders and the Dairy Price Support Program. These two programs, along with the

---

\(^5\)The efficient price of quota is the discounted future rents from quota holding.
more recent Milk Income Loss Contract and ongoing trade barriers, constitute the central tenants of U.S. dairy policy.

2.3.1 Federal Milk Marketing Orders

The Agricultural Adjustment Act of 1933 gave the Secretary of Agriculture the authority to establish voluntary marketing agreements with processors, associations of producers and other handlers of milk. According to Bailey (1997) an order is a complex set of rules that assist in the equitable pricing of milk to dairy producers and processors, and help provide a supply of pure and wholesome milk to consumers. Federal Milk Marketing Orders (FMMOs) currently set a minimum price at which milk must be purchased in a particular region. In practice, the producer price of milk is seldom much higher than this price floor (although some large cooperatives have the market power to negotiate over-order prices.)

From their inception marketing orders have been based on classified pricing and marketwide pooling. Classified pricing implies that milk is priced differently depending on its final use. Similar to Ontario, there are four classes of milk: fluid milk, soft dairy products, cheese and hard dairy products. Marketwide pooling requires that all producers be paid the same blended price regardless of the final use of their particular milk (Bailey, 1997, p. 113).

Some regions are not covered by FMMOs and instead fall under state orders. In New York State most producers fall under the Northeast FMMO, while a smaller state marketing order covers producers in the Western region. Alexander et al. (1998, p. 2) note that since the late 1980s "the state order has relied on virtually
identical pricing with surrounding federal orders.”

The system of FMMOs was most recently modified in 2000. The number of orders was reduced from 31 to 11, and the pricing scheme for milk was changed to reflect the prices of manufactured dairy products, assumed product yields, and make allowances (Jesse and Crop, 2001, p. 2).

The price floor used by FMMOs is generated by first pricing its components (butterfat, protein and other solids.) The prices of components are derived from the market price of one or two manufactured dairy products. Class III and IV prices (for cheese and hard dairy products respectively) are found by combining these component prices, while Class I and II prices (for fluid milk and soft dairy products) are based on Class III and IV prices plus a differential. The complete formulas for milk classes are found in appendix A.

One of the more integral government policies is the Class I differential (the premium paid for fluid milk), which differs both between orders and within orders. Typically the differential increases with the distance from Wisconsin, reflecting the fact that Wisconsin once had the largest fluid milk surplus in the country (Balagtas, 2007, p. 11). In New York State, the differential varies from $2.20 in Jamestown (in Western New York) to $3.15 in New York City, appendix B outlines the differentials by county (Northeast Marketing Area, 2010).

2.3.2 Dairy Price Support Program

The agricultural acts and amendments of the 1930s also authorized the first Milk Price Support Program, whose name recently changed to the Dairy Price Support
Program. The program supports milk prices through government purchases of dairy products by the Commodity Credit Corporation (CCC). Government inventories are disposed of through product sales (if prices increase), exports (although recent trade rules now preclude the large scale use of export subsidies), or foreign and domestic donation (United States Department of Agriculture, 1968).

Milk price supports were initially set to maintain milk prices between 75 and 90 percent of parity prices (with parity defined as the average milk price in the five years before World War I) (Bailey, 1997, 174-176). Parity pricing remained in effect until the 1980s when the program’s inventories and costs increased precipitously.

The Farm Bill of 1996 sought to gradually phase out the Milk Price Support Program, reducing the support price by $0.15 per year starting in 1997 (when the support price was $10.35/cwt), and provisioning the end of the program entirely in 1999 (U.S. Congress, 1996, sec 141). However, in 2000, Congress extended the program on an annual basis, and re-instituted it permanently in 2002 with a support price of $9.90/cwt (U.S. Congress, 1996, sec 1501). The Farm Bill of 2008 differed from previous bills as it specified support prices for dairy products (cheese, butter and powder) as opposed to milk. The stated component support prices are still roughly equivalent to a milk support price of $9.90/cwt (U.S. Congress, 2008, sec 1501). Figure 2.2 outlines the government purchases from 1990 to 2008. Notably, in the most recent decade government purchases have not exceeded 1% of production.
2.3.3 Milk Income Loss Contract

The Milk Income Loss Contract (MILC) is a relatively recent counter-cyclical payment introduced in the Farm Security and Rural Investment Act (Farm Bill) of 2002. Currently, when the Class I price falls below $16.94/cwt on the Boston market, farmers receive a payment equal to 45% of the difference between this trigger value and the Boston price (U.S. Congress, 2008, sec 1506). Payments are capped at 2,985,000 lbs per year per farm; approximately the output of 145 cows (U.S. Congress, 2008).
The trigger for the MILC payment is adjusted based on the National Daily Average Feed Ration calculated by the National Agricultural Statistics Survey. When feed costs rise above a specified threshold (currently set at $7.25 / cwt), the MILC trigger value will increase by 45% of the difference between the feed cost and the threshold (U.S. Congress, 2008, sec 1506). The movement of this threshold and the expenditures of the MILC program are outlined in figure 2.3.

Figure 2.3: Milk income loss contract trigger price and payment (2001-2009)

Source: [http://future.aae.wisc.edu/alliance/2012/MILC_full.pdf](http://future.aae.wisc.edu/alliance/2012/MILC_full.pdf)

### 2.3.4 Trade policies

Similar to Canada, the United States uses a system of tariff-rate quotas for dairy imports. This system has kept dairy imports to between three and five percent of
2.3.5 Effect of U.S. dairy policies

Brown (2003) estimated that the elimination of the Dairy Price Support Program would have little to no impact on dairy production and prices, finding that average production and prices between 2003-2007 would decline by less than one percent. Cox and Chavas (2001) estimated a similar effect econometrically. Given that figure 2.2 shows few purchases over the past ten years, these forecasts seem to have been borne out.
The aggregate effect of FMMOs in the national market is also estimated to be rather small; Cox and Chavas (2001) use a spatial model to simulate the effect of eliminating federal milk marketing orders on milk prices and consumption in the United States. They find that elimination of these two regulations would cause the national average price of milk to fall by 1.9% and production to fall by 0.1%. Brown (2003) finds long run declines in prices and production of less than 1%.

This national average however hides regional differences. Cox and Chavas (2001) contend that milk production in a deregulated environment would move to states with lower costs of production (such as the Midwest and California) and away from states such as New York that produce at a higher cost. For the Northeast region the price was simulated to fall by 7.1% and production by 1.8%. If the Northeast region as a whole is taken as a proxy for New York, then it can be said that the FMMO has a non-trivial impact on producer and consumer welfare in the state.

Finally, MILC program payments have averaged 4.5% of the fluid grade price of milk in New York, between 2002 and 2009. Taken as a whole these programs can therefore be seen to increase the price farmers receive by almost 12%.

### 2.4 Regional comparison

This section provided an overview of dairy policies in Canada and the United States, including evidence of the distortionary effects of such programs. The literature previously cited does not allow for direct comparison between the two policies. Two metrics that are more comparable, are the milk price received by
Figure 2.5: Mean annual milk price in Ontario and New York (2000-2009)

farmers, and producer support estimates (PSEs) calculated by the Organization for Economic Co-operation and Development (OECD). From our data sets (details of which will soon be outlined), the milk price received by farmers in Ontario averaged $0.66/litre while the price New York farms received is $0.45/litre (both figures are in Canadian dollars). Mean annual prices are shown in figure 2.5. The effect of the Milk Income Loss Contract is not felt in the milk price, however, after accounting for the mean effect of the subsidy (4.5%), the difference is still significant.

The OECD measures total transfers to farmers (either from taxpayers or consumers), by the PSE, which divides the total transfer by the value of production. Figure 2.6 shows the PSE for both countries between 2000 and 2009. The average
Figure 2.6: Producer subsidy equivalent for Canadian and US dairy industries (2000-2009)

PSE for dairy farmers in Canada measured 52.72%, while the PSE for dairy farmers in the US measured 23.97%, a difference of 28.75 percentage points. In only one year (2001) do American farmers receive more support than Canadians (OECD, 2011).\(^6\)

\[^6\text{The OECD calculation of PSE is rightly criticized by Brinkman (2010), as it assumes Canada and the US could import milk at the same price as areas that currently have “free” trade in milk. The perishability of milk, and the thinness of the world market make this assumption unreasonable. However, any adjustment to make the price more reasonable would arguably affect the two regions in the same way. Thus the statistic accurately captures relative differences in PSE.}\]
Chapter 3 Dairy Farm Efficiency - Economic Theory and a Brief Literature Review

3.1 The concept of economic efficiency

The neo-classical school, which has come to dominate much of economic thinking, is undergirded by assumptions of cost minimization and profit maximization. This conception of economic actors has, however, been challenged by empirical and anecdotal evidence that has shown that some firms produce the same output while having varying levels of inputs and cost.

Economic interest in production efficiency is generally traced to Koopmans (1951). Koopmans argues (with a nod to Pareto) that ”a producer is technically efficient if any increase in output requires a reduction in at least one other output or an increase in at least one input, and if a reduction in any input requires an increase in at least one other input or a reduction in at least one output.” (p. 60)

Economic interest quickly broadened from examining technical (or physical)
efficiency, as defined by Koopmans, to examining cost (or economic) efficiency. Cost efficiency is achieved when any increase in output requires an increase in cost, and any decrease in cost requires a decrease in output (a firm may be technically (physically) efficient, without being cost efficient\(^1\).) The link between technical and cost efficiency is explained as allocative (price) efficiency. Farrell (1957, p. 254) succinctly defined allocative efficiency as “the extent to which a firm uses the various factors of production in the best proportions, in view of their prices.”

Input-oriented technical efficiency can be described mathematically by first calculating a production possibilities set. Allowing for \(N\) inputs \((x)\) and \(M\) outputs \((y)\), the production possibilities set \((\Psi)\) can be defined as

\[
\Psi = \{(x,y) \in \mathbb{R}^{N+M} | x \text{ can produce } y\},
\]

where \(\mathbb{R}\) denotes a real number. A producer’s radial technical efficiency score\(^2\) is given by

\[
TE_I(y, x) = \min \{\theta : (\theta x, y) \in \Psi\},
\]

where \(\theta\) is a scalar. Technical efficiency is therefore the inverse of the distance function \(\delta_I\) introduced by Shephard (1953), as shown in equations 3.3 and 3.4,

\[
\delta_I(y, x) = \max \{\lambda : (x/\lambda, y) \in \Psi\}
\]

\(^1\)As opposed to cost efficiency one could also concern themselves with profit efficiency. The achievement of profit efficiency is realized when any increase or decrease in either input or output decreases profit. The measure is premised on the assumption that firms can freely move to the most efficient scale.

\(^2\)Note that this definition of technical efficiency differs from Koopmans’ definition, as it allows for “slacks”.


\[ TE_I(y, x) = 1/\delta_I(y, x), \]  

(3.4)

where \( \lambda \) is a scalar. The calculation of cost efficiency relies on the estimation of a cost frontier, or minimum cost function, given by

\[ c(y, w) = \min_x \{ wx : (x, y) \in \Psi \}, \]  

(3.5)

where \( w \) represents a vector of input prices. Cost efficiency is defined in equation 3.6 as the ratio of minimum costs (as defined by the cost frontier) to actual cost,

\[ CE(x, y, w) = c(y, w)/w^T x. \]  

(3.6)

Finally, equation 3.7 defines allocative efficiency as the ratio of cost efficiency to technical efficiency,

\[ AE_I(x, y, w) = CE(x, y, w) / TE_I(y, x). \]  

(3.7)

Figure 3.1 illustrates the concepts of technical, allocative and cost efficiency graphically.

### 3.2 Regulation and efficiency

Alchian (1950) suggests that firms should be seen as pursuing profits, as opposed to maximizing them. This pursuit, characterized by "adaptive, imitative and trial-and-error behaviour", typically results in less than maximal profits due to
In this figure a producer faces a production possibilities frontier $\delta \Psi$, and produces using the input bundle denoted $A$. His technical efficiency can be measured by the distance from the origin to the frontier, divided by the distance from the origin to his input bundle,

$$\text{Technical Efficiency} = \frac{OB}{OA}. \quad (3.8)$$

Point $D$, the tangency of the production possibilities frontier and the price ratio, represents the cost efficient point of production, and is equal (in terms of cost) to point $C$. Thus cost efficiency can be defined as,

$$\text{Cost Efficiency} = \frac{OC}{OA}. \quad (3.9)$$

Allocative efficiency is the ratio of cost efficiency to technical efficiency and is thus,

$$\text{Allocative Efficiency} = \frac{OC}{OB}. \quad (3.10)$$
bounded rationality, exogenous shocks, and learning costs. Winter (1971) adds that firms typically operate on the basis of decision rules which are changed infrequently, only when shown to be sub-optimal. Inefficiency can arise when such rules are not yet clearly shown to be imperfect. Winter goes on to equate the market with biological evolution, suggesting that firms who fail to adapt will be naturally selected from the market. Thus, in competitive industries one can imagine competition limiting the amount of inefficiency in the marketplace.\(^3\)

The evolutionary critiques of Alchian and Winter are of particular relevance for the dairy industry. Institutions that shield producers from competition, preclude inefficient producers from being naturally selected from the market. With no threat of competition, firms are free to pursue goals other than profit maximization (such goals may include maximizing utility or maximizing the utility of profits.)

Furthermore, as Hayek (1945) noted, competitive prices are a mechanism for communicating information, removing the competitive price from the marketplace eliminates a criterion that firms use to evaluate their decision rules.

We therefore expect regulatory regimes that reduce price competition, to engender more inefficiency leading to the hypothesis that Ontario dairy farmers are less efficient than their counterparts in New York.

\(^3\)This analysis calls to mind Friedman (1953)'s argument that the assumption of cost minimization should not be judged by its realism, but rather its fruitfulness in economic prediction. Thus, businessmen need not consciously be cost minimizers, so long as the competitive market forces them to act as if they were. Empirical evidence, from the dairy sector and other industries, has given reason to be less sanguine than Friedman, regarding the ability of the market to drive out inefficiency.
3.3 Two practical considerations

As mentioned in the previous chapter, the system of supply management imposes production restrictions on Ontario dairy farms, which take the form of quota. It has been argued to this author, in various forums, that this system inhibits the comparison of farms in the two regions, since farms in New York seek to maximize profit, while farms in Ontario (whose output is fixed in the short run) seek to minimize cost. A second (unrelated) argument is that the tradability of quota should drive inefficiency from the market. These two issues bear further discussion and are taken up here.

3.3.1 The objective of firms

Economic efficiency can be calculated either through the profit or cost function. In this study a cost function is used as farmers in Ontario face a constraint on their output. The obvious criticism is that New York farmers seek to maximize profit – not minimize costs. This criticism can be answered in three different ways. First, and most simply, cost minimization is a prerequisite for profit maximization. Second, empirical evidence suggests that even when output is unconstrained farmers are better characterized as cost minimizers than as profit maximizers. Tauer (1995) found that New York dairy farmers netput minimized their cost 69% of the time and maximized profits only 49.9% of the time (relative to past and future netputs.) Featherstone et al (1995) used a similar methodology in considering Kansas farms, finding that costs were minimized 72.0% of the time while profits were maximized.
only 48.2% of the time. Given that cost minimization is required for profit maximization it is not surprising that more observations violate the weak axiom of profit maximization than the weak axiom of cost minimization. The magnitude of the difference, however, does lend credence to the use of the cost function as opposed to the profit function.

Finally, the objective of the farmers is of little consequence. Farmers may seek to maximize utility, maximize the utility of profits or pursue one of a myriad of other objectives. The purpose of this study, whose relevance and audience was discussed in chapter one, is to determine their success in minimizing costs not in achieving another objective, even if that is the objective that the farmer holds.

3.3.2 Quota values

Coase (1959) theorized the initial distribution of property rights should be irrelevant in the long-run, as those producers who can derive the highest benefit from that right can buy it from those who derive smaller benefits. The preconditions for the Coase theorem to hold, including established property rights and zero transaction costs, are met fairly well by the quota exchange. Thus, the capitalization of benefits into the price of quota should actually force quota to the most efficient producers. This being said, empirical evidence has shown that quota exchanges do not necessarily lead to the efficient allocation of quota. Colman (2000) performed a simulation of the quota exchange in the UK. His model allowed firms to maintain their current size, increase their quota by up to 20%, or sell all their quota. Under these assumptions the price of quota fell by 53%, average herd size increased from
70 to 100, and 15.8% of quota was transferred to more efficient producers. Boots, Oude Lansink, and Peerlings (1997), analyzed the quota market in the Netherlands in 1992-1993, finding that 10.8% of quota would need to be traded in order to maximize industry profits. Finally, Guyomard et al. (1996) analyzed the French quota market, finding that the maximization of industry profits in France required the exchange of between 11% and 16% of quota.

These curious results (curious perhaps only to the professional economist) can be explained in part by the differing objectives that farmers hold. A farmer may derive some utility from the act of farming, thus even if the sale of quota could generate a higher profit the farmer will choose to hold quota. In the absence of a strong selection mechanism, this farmer can continue farming so long as his marginal cost is below the administered price.

### 3.4 Existing literature on dairy farm efficiency

Academic investigation of economic efficiency started in earnest thirty years ago with the acceptance of both data envelopment analysis and stochastic frontier analysis. Since this time the literature has not suffered from a lack of studies on dairy farm efficiency. This section surveys studies undertaken in Canada and New York state, and two additional papers whose findings are of interest.

average 53.20% significantly less than New York farmers who averaged 60.20%. As mentioned previously this study seeks to extend the work of this paper by examining cost and allocative efficiency, employing updated data, and triangulating between efficiency methods.

Weersink, Turvey, and Godah (1990) decomposed the technical efficiency of Ontario farmers into scale, input congestion and pure technical efficiency using data envelopment analysis. They found these efficiencies to measure 96.79%, 99.84% and 94.94% respectively. A second stage, censored non-linear least squares regression found that herd size, output per cow, proportion of purchased feed used and location all increased efficiency with statistical significance while debt to asset ratio and building expense per cow both negatively impacted efficiency with statistical significance.

Mbaga et al. (2003) employed stochastic frontier analysis to estimate the technical efficiency of Quebec dairy farms. The authors examined different functional forms and distributional assumptions, finding the generalized Leontief function to be the most appropriate specification of technology, and a truncated normal error term to be the best representation of inefficiency. Under this specification Quebec farms were on average 94.73% or 94.88% efficient (depending on if the farm was located in a maize producing region.) Notably, the paper also examined the correlation between the results of this specification and data envelopment analysis. Using data envelopment analysis the mean efficiencies for the two regions were 92.15% and 95.00%. While the mean efficiencies were not dissimilar under the two different estimators, the correlation of the two estimators was only .357 (rank correlation
Hailu, Jeffrey, and Unterschultz (2005) compared Alberta and Ontario dairy farms using stochastic frontier analysis, while testing the sensitivity of results to the imposition of curvature on the cost function. Mean efficiency scores for Ontario and Alberta measured 92% and 84% with curvature imposed, and 92% and 83% without curvature imposed. Although mean efficiency estimates were not greatly affected by curvature, input elasticities and input demands were found to be sensitive to its imposition.

Tauer and his associates, in a series of papers, have examined the efficiency of New York dairy farmers. Tauer and Belbase (1987) used correlated ordinary least squares (a deterministic econometric technique) to find an average technical efficiency of 69.3% for New York dairy farms. The authors then used ordinary least squares in a second regression to determine the factors affecting efficiency, finding that location and herd size increased efficiency, while the use of mail-in records and participation in a dairy herd improvement program reduced efficiency.\textsuperscript{4}

Tauer (1993) used data envelopment analysis to examine technical, allocative and scale efficiency of New York dairy farms in both the short and long run. The distinction between short run and long run efficiency was made by eliminating fixed inputs (capital and the operator’s labour) in the calculation of short-run efficiency. Technical efficiency was found to be 74% and 79%, allocative efficiency 87% and 70% and scale efficiency 95% and 93%, in the short and long run respectively. A logistic

\textsuperscript{4}The authors contend that participation in the dairy herd improvement program may result in the overuse of inputs, as increasing output per cow is its main goal. Alternatively, there may be a self-selection bias if it is primarily under performing farms that enrol.
regression of likely covariates on efficiency scores returned no significant variables. The increase in technical efficiency between the two horizons is not surprising as the methodology used generally returns higher efficiency when new variables are added (Banker and Morey, 1986). The decrease in allocative efficiency is explained as economic agents have less certainty about long-term prices.

Tauer (1996) used Varian’s weak axiom of cost minimization and weak axiom of profit maximization to determine if farmers act according to neo-classical cost minimizing and profit maximizing assumptions. For each year’s prices, each farmer’s netput from all sample years was compared to the current netput to determine if the current netput did indeed minimize costs or maximize profits (the implicit assumption being that the farmer’s technology allows for the attainment of any netput they have exhibited throughout the life of the sample). The weak axiom of profit maximization was violated for 50.1% of the years, while the weak axiom of cost minimization was violated 31% of the time (as cost minimization is a requisite for profit maximization the differences are unsurprising.)

Tauer and Stefanides (1998) extend Tauer’s (1995) study by using a distance function to measure the magnitude of deviation from a farmers profit maximizing point. They found that on average farmers could increase their profits by 20% if they used the most profitable netput each year. A second step tobit regression again found no significant co-variates.

Mosheim and Lovell (2009) use a shadow cost model to examine the effects of size on the efficiency of US dairy farms. The results from their study are presented in table 3.3. The results indicate that larger farms are more profitable both because
Table 3.1: Efficiency, scale elasticity and costs of US dairy farms (Moshiem and Lovell, 2009)

<table>
<thead>
<tr>
<th>Herd size</th>
<th>Technical Efficiency</th>
<th>Allocative Efficiency</th>
<th>Cost Efficiency</th>
<th>Scale Elasticity</th>
<th>AIC_{eff}</th>
<th>AIC_{ineff}</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;30</td>
<td>72%</td>
<td>53%</td>
<td>39%</td>
<td>2.42</td>
<td>$30.61</td>
<td>$82.91</td>
</tr>
<tr>
<td>30–50</td>
<td>73%</td>
<td>53%</td>
<td>39%**</td>
<td>2.48</td>
<td>$20.73</td>
<td>$55.46</td>
</tr>
<tr>
<td>50–100</td>
<td>74%</td>
<td>55%***</td>
<td>41%***</td>
<td>1.98***</td>
<td>$15.10</td>
<td>$37.88</td>
</tr>
<tr>
<td>100–200</td>
<td>74%</td>
<td>58%***</td>
<td>44%***</td>
<td>1.81</td>
<td>$10.78</td>
<td>$24.78</td>
</tr>
<tr>
<td>200–500</td>
<td>80%***</td>
<td>62%***</td>
<td>49%***</td>
<td>1.48***</td>
<td>$9.00</td>
<td>$18.33</td>
</tr>
<tr>
<td>500–1000</td>
<td>82%**</td>
<td>66%***</td>
<td>54%***</td>
<td>1.34***</td>
<td>$7.15</td>
<td>$13.25</td>
</tr>
<tr>
<td>1000–2000</td>
<td>83%</td>
<td>66%</td>
<td>55%</td>
<td>1.29</td>
<td>$6.81</td>
<td>$12.18</td>
</tr>
<tr>
<td>2000+</td>
<td>86%</td>
<td>71%*</td>
<td>60%</td>
<td>1.25</td>
<td>$7.13</td>
<td>$11.66</td>
</tr>
<tr>
<td>Average</td>
<td>75%</td>
<td>56%</td>
<td>42%</td>
<td>2.06</td>
<td>$16.76</td>
<td>$42.72</td>
</tr>
</tbody>
</table>

AIC_{eff} refers to the average incremental cost on the frontier (assuming efficiency), and AIC_{ineff} refers to the average incremental cost when technical and allocative inefficiency are included.

*, **, *** indicate significant differences between herd sizes at 10%, 5% and 1% levels

they are more efficient (technically and allocatively), and because the industry is subject to increasing returns to scale.

Kumbhakar et al. (2008) examine output growth in Norway using a distance function. The distance function was used to calculate output growth rate under three different policy regimes in Norway; a period before the quota regime was introduced (1976-1982), a period during which the most restrictive quota regime was in place (1983-1996) and a period with a more flexible quota scheme (1996-2005.) Output growth was found to be significantly affected by the particular policy regime, averaging 3.95%, 1.62% and 2.56% in the three periods respectively.

Notably they also found that all periods exhibited technological regress which averaged 0.46% over all periods. The authors posited that this regress could be due to increasing regulation, lack of external competition, increasing technical
inefficiency or a decreasing margin between output and input prices.
Chapter 4  Data

Farm level data is taken from two sources: the Ontario Dairy Farm Accounting Project\(^1\) and the Dairy Farm Business Summary\(^2\) from New York state. The Ontario Dairy Farm Accounting Project is a rotating panel consisting of a regionally stratified random sample of farms. Typically, a farm remains in the panel for three to five years. The average number of farms surveyed in any given year is around eighty (comprising 1-2% of the population.) Given this paper’s interest in dairy farms, producers who derive more than ten percent of total receipts from non-dairy operations are deleted (this threshold was previously used by Moschini (1988)).

The Dairy Farm Business Survey is a voluntary reporting program that covers approximately 5% of farms in New York State.\(^3\) Farms that enter are on average larger and more productive than the state’s average dairy farm. To redress the bias of the New York data (towards larger and more productive dairy farms), a subsample was selected from the larger sample. The subsample emulates the average farm size and productivity (output per cow) of the state, by using a

---

\(^1\)The Ontario Dairy Farm Accounting project is jointly administered by the Canadian Dairy Commission, the Dairy Farmers of Ontario, and the University of Guelph.

\(^2\)The Dairy Farm Business Summary is administered by Cornell University.

\(^3\)None of the farms in the New York state dataset have non-dairy receipts in excess of 10% of total receipts.
histogram of state farm sizes, and the state average output per cow from the Department of Agriculture and Markets, New York State (2010). The procedure for subsampling was as follows:

- Classify farms according to their size, using classes defined by Department of Agriculture and Markets, New York State (2010).
- Specify the number of farms that should be chosen from each class, each year, based on the proportion in the state population.\(^4\)
- Randomly select producers in each class for the first year.
- Verify average herd size and average output per cow are within 80% of the state average, if not resample.
- Producers included in previous subsamples are kept in subsequent years, with additional producers randomly added to satisfy the previous requirements.

The number of farms in the two (sub)samples each year, by year of entry, is given in tables 4.1 and 4.2.

### 4.1 Variables

The data is aggregated into three outputs (milk, livestock and crop) and four inputs (feed, labour, capital and other). All aggregation was done using Fisher price indices, with quantities then calculated as revenue or expenses divided by this price

---

\(^4\)In certain years there were not enough small farms (less than fifty cows) in the dataset to mimic the percent within the population, in these instances, more farms from the next size category were added.
Table 4.1: Ontario dairy farms included in analysis by year of entry

<table>
<thead>
<tr>
<th>Year of entry</th>
<th>Panel year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2000 2001 2002 2003 2004 2005 2006 2007 2008 2009</td>
</tr>
<tr>
<td>2000</td>
<td>75 49 35 26 19 0 0 0 0 0</td>
</tr>
<tr>
<td>2001</td>
<td>– 19 11 10 8 7 0 0 0 0</td>
</tr>
<tr>
<td>2002</td>
<td>– – 14 8 8 7 8 0 0 0</td>
</tr>
<tr>
<td>2003</td>
<td>– – – 30 23 20 16 12 0 0</td>
</tr>
<tr>
<td>2004</td>
<td>– – – – 16 9 7 4 2 0</td>
</tr>
<tr>
<td>2005</td>
<td>– – – – – 15 7 10 8 8</td>
</tr>
<tr>
<td>2006</td>
<td>– – – – – – 24 14 9 10</td>
</tr>
<tr>
<td>2007</td>
<td>– – – – – – – 31 20 19</td>
</tr>
<tr>
<td>2008</td>
<td>– – – – – – – – 20 15</td>
</tr>
<tr>
<td>2009</td>
<td>– – – – – – – – – 12</td>
</tr>
<tr>
<td>Total</td>
<td>75 68 60 74 74 58 62 61 59 64</td>
</tr>
</tbody>
</table>

Table 4.2: New York dairy farms included in analysis by year of entry

<table>
<thead>
<tr>
<th>Year of entry</th>
<th>Panel year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2000 2001 2002 2003 2004 2005 2006 2007 2008 2009</td>
</tr>
<tr>
<td>2000</td>
<td>75 47 40 34 31 30 29 28 24 22</td>
</tr>
<tr>
<td>2001</td>
<td>– 28 18 17 17 16 12 11 10 8</td>
</tr>
<tr>
<td>2002</td>
<td>– – 17 15 9 8 10 10 7 5</td>
</tr>
<tr>
<td>2003</td>
<td>– – – 9 8 6 5 4 2 2</td>
</tr>
<tr>
<td>2004</td>
<td>– – – – 10 8 8 6 6 4</td>
</tr>
<tr>
<td>2005</td>
<td>– – – – – 7 3 3 3 1</td>
</tr>
<tr>
<td>2006</td>
<td>– – – – – – 8 6 5 5</td>
</tr>
<tr>
<td>2007</td>
<td>– – – – – – – 7 6 5</td>
</tr>
<tr>
<td>2008</td>
<td>– – – – – – – – 12 7</td>
</tr>
<tr>
<td>2009</td>
<td>– – – – – – – – – 16</td>
</tr>
<tr>
<td>Total</td>
<td>75 75 75 75 75 75 75 75 75 75</td>
</tr>
</tbody>
</table>

index. All US dollar values are converted to Canadian dollars, using the annual average exchange rate taken from the Bank of Canada. Descriptive statistics of these variables are shown in table 4.3.
Table 4.3: Descriptive statistics of variables used in this study

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Pooled (N=1410)</th>
<th>New York (N=750)</th>
<th>Ontario (N=660)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dairy cows per farm</td>
<td>Cows</td>
<td>90 (115)</td>
<td>110 (148)</td>
<td>67 (50)</td>
</tr>
<tr>
<td>Output per cow</td>
<td>Liters</td>
<td>8,029 (1,627)</td>
<td>8,266 (1,760)</td>
<td>7,760 (1,415)</td>
</tr>
<tr>
<td>Milk price</td>
<td>Canadian Dollars</td>
<td>0.56 (0.14)</td>
<td>0.45 (0.067)</td>
<td>0.69 (0.074)</td>
</tr>
<tr>
<td>Crop revenue</td>
<td>Canadian Dollars</td>
<td>17,823 (39,255)</td>
<td>11,204 (32,966)</td>
<td>25,344 (44,193)</td>
</tr>
<tr>
<td>Livestock revenue per cow</td>
<td>Canadian Dollars</td>
<td>28,511 (42,159)</td>
<td>30,442 (52,494)</td>
<td>26,316 (25,676)</td>
</tr>
<tr>
<td>Purchased feed expense per cow</td>
<td>Canadian Dollars</td>
<td>974 (438)</td>
<td>1,184 (413)</td>
<td>735 (330)</td>
</tr>
<tr>
<td>Labour per cow</td>
<td>Implicit Fisher Index</td>
<td>104 (37)</td>
<td>101 (32)</td>
<td>107 (42)</td>
</tr>
<tr>
<td>Physical assets</td>
<td>Canadian Dollars</td>
<td>1,178,384 (1,268,622)</td>
<td>1,014,302 (1,096,149)</td>
<td>1,364,840 (1,417,725)</td>
</tr>
<tr>
<td>Physical assets per cow</td>
<td>Canadian Dollars</td>
<td>15,013 (12,657)</td>
<td>10,487 (3,889)</td>
<td>20,157 (16,598)</td>
</tr>
<tr>
<td>Other quantity per cow</td>
<td>Implicit Fisher Index</td>
<td>926 (403)</td>
<td>716 (277)</td>
<td>1,164 (390)</td>
</tr>
</tbody>
</table>

Cost shares

<table>
<thead>
<tr>
<th></th>
<th>Percentage</th>
<th>Pooled (N=1410)</th>
<th>New York (N=750)</th>
<th>Ontario (N=660)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed</td>
<td>Percentage</td>
<td>0.22 (0.11)</td>
<td>0.3 (0.09)</td>
<td>0.13 (0.062)</td>
</tr>
<tr>
<td>Labour</td>
<td>Percentage</td>
<td>0.23 (0.076)</td>
<td>0.24 (0.067)</td>
<td>0.22 (0.085)</td>
</tr>
<tr>
<td>Capital</td>
<td>Percentage</td>
<td>0.36 (0.13)</td>
<td>0.28 (0.06)</td>
<td>0.45 (0.12)</td>
</tr>
<tr>
<td>Other</td>
<td>Percentage</td>
<td>0.19 (0.056)</td>
<td>0.18 (0.055)</td>
<td>0.2 (0.055)</td>
</tr>
</tbody>
</table>

Mean values with standard deviation in parenthesis
4.1.1 Milk

Both datasets include the quantity of, and revenue from, milk sales, allowing for the calculation of an implicit price. Pounds were converted to litres at a rate of 2.275 lbs/litre (Bailey, 2002).

4.1.2 Crop (output)

The New York data set contains accrual crop revenue and production of six different crops (hay, hay silage, corn silage, grain corn, oats, wheat and other forage.) Problematically, most crops are also used as feed, with only the excess sold. To counter this problem, crops are assumed to be sold in the same proportion to which they are grown (except for pasture and forage which is assumed to be used exclusively as grown feed)\(^5\). The unindexed quantity of the \(i^{th}\) crop sold is calculated as

\[
Q_i = \frac{Q_i^g \cdot P_i}{\sum_{i=1}^{n} (Q_i^g \cdot P_i)/P_i} \cdot \frac{\text{Accrual crop sales}}{P_i}
\]  

(4.1)

where \(Q_i\) is the estimated quantity sold, \(Q_i^g\) is the quantity grown and \(P_i\) represents the price of the \(i^{th}\) crop taken from the Department of Agriculture and Markets, New York State (2010). It is assumed that the market prices of crops are relatively the same as the individual prices farmers face.

The Ontario data set offers both quantity and revenue of individual crops, allowing for the calculation of an implicit price. An examination of the implicit prices

\(^5\)This assumption is validated in the Ontario data set, where significant amounts of hay forage are grown but only 1 farm reported sales
Table 4.4: Aggregation of Crops

<table>
<thead>
<tr>
<th>New York crops</th>
<th>Ontario Crops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corn grain</td>
<td>Corn grain</td>
</tr>
<tr>
<td>Corn silage</td>
<td>Corn silage</td>
</tr>
<tr>
<td>Hay and hay silage</td>
<td>Hay, hay silage, straw, grainlage, pasture and other</td>
</tr>
<tr>
<td>Oats</td>
<td>Oats</td>
</tr>
<tr>
<td>Wheat</td>
<td>Wheat, barley and mixed grain</td>
</tr>
<tr>
<td>–</td>
<td>Other (primarily soybeans)</td>
</tr>
</tbody>
</table>

revealed some irregularities in pricing. As a result implicit prices that were 25% greater (less) than the prices reported in the annual reports of the Dairy Farmers of Ontario were reduced to the upper (lower) boundary of this range. Given that more crops were specified in this data set than in the New York State data set, individual crops were aggregated into six categories as shown in table 4.4. A producer specific price was calculated as a weighted average of the prices within the group.

4.1.3 Livestock (output)

The New York data set provided the quantity of cull cows and dairy cows sold, and revenue from cow sales (including cull and dairy cows), calf sales and other livestock. An implicit price for dairy calves was easily achieved by dividing revenue by quantity. An implicit price of cull and dairy cows was found by assuming the relative prices of these two outputs are the same as the market prices taken from

---

6The amount of variation allowed is admittedly arbitrary - in practice this decision rule changed the prices of 20% of producers
7Crops were grouped so to reduce the price variance within groups.
8While both data sets allow for the calculation of accrual livestock receipts, it was decided to use cash receipts instead. Given that livestock is also included as a capital expense, the changes in inventory are better reflected as changes in productive assets and are therefore considered in that category.
the Department of Agriculture and Markets, New York State (2010). The producer specific implicit price of dairy and cull cows was then calculated as

\[ P_i = \frac{(\text{Cash receipts from cow sales} \times \frac{Q_i \times P_{ext}^i}{\sum_{i=1}^{n} Q_i \times P_{ext}^i})}{Q_i} \]

(4.2)

where \( i \) indexes dairy and cull cows, \( Q_i \) represents the quantity sold and \( P_{ext}^i \) represents the market price.

The Ontario data set contained the quantity and revenue for several livestock categories. These categories were aggregated to match the three categories from New York, with prices calculated for each category as the weighted average of component prices. Once again, implicit prices displayed irregularities, and were therefore compared to market prices taken from the annual reports of the Dairy Farmers of Ontario. These external prices vary depending on the quality of livestock, with a typical range of $1,200 between poor and excellent quality cows. Implicit prices that were greater than 50% of the maximum or lower than 50% of the minimum of this range were reset to the nearest value within the range.\(^{10}\)

Revenue from other livestock (in both regions) were converted into producer specific cow equivalents based on the proportion of cull and dairy cows sold.\(^{11}\)

4.1.4 Feed

While the Ontario data set offered detailed information on feed, the New York data set aggregated these costs into three broad categories (grain and concentrate,

\(^{10}\)More variation was allowed in livestock prices, as quality differences tend to be larger in livestock than in crops.

\(^{11}\)Given producers were precluded from analysis if income from specialized production was greater than 10%, this procedure was deemed to be not too distorting.
roughage and non-dairy feed). With these data constraints in mind, the price of 16% dairy ration was used as a proxy for feed price.  

While previous studies of dairy farm efficiency have included home grown feed (Stefanou and Saxena, 1988; Tauer, 1993; Mosheim and Lovell, 2009) in their metric of feed costs, they are actually an intermediate input. Primary inputs in the production of home grown feed would include seed, fertilizer, pesticide and a proportion of labour and capital. Unfortunately, the proper method of accounting for these primary inputs is not clear.

Weersink, Turvey, and Godah (1990), using the Ontario Dairy Farm Accounting Project dataset, included many of these primary costs in their feed variable. Conversely, Moschini (1988) using the same dataset, included these costs as intermediate inputs, a category which also includes such variables as fuel, hydro and telephone, and veterinary and drug expenses. While the former approach is more theoretically appealing, the literature tends to favour the latter approach. This study follows the bulk of the literature in categorizing these primary feed inputs as "other" inputs. The rationale for this method is the impossibility of disaggregating the proportion of inputs (such as capital and labour) that are used in producing feed.

### 4.1.5 Labour

In measuring labour expense, the literature diverges on the costing of family labour. Most studies (such as Hailu, Jeffrey, and Unterschultz (2005) and Moschini (1984))
price family and hired labour at the same wage rate. Other studies have attempted to estimate the opportunity cost of family labour. Mosheim and Lovell (2009), for example, predict a wage for family labour by regressing off-farm income on such variables as education, age of the operator, and location (including unemployment rate and industrial composition of the region). The first method is often preferred for its ease of calculation; I argue that in efficiency studies it is more theoretically valid as well.

The prices of inputs are used to determine the optimal changes in input usage, therefore it is the marginal cost of these inputs that is of importance. If farms have significant hired labour expenses, then any reduction or increase in labour will likely first come from hired labour. Furthermore, the transaction costs involved in off-farm labour are likely quite high (especially for full-time workers), meaning the marginal benefit of off-farm labour (or the opportunity cost of on-farm labour) is likely much less than the calculated wage rate. For these reasons, this study prices family and hired labour both at the prevailing wage rate.

The average labour price from New York state is taken from Department of Agriculture and Markets, New York State (2010). The Ontario wage rate for 2010 is taken from a survey done by Ontario’s Progressive Dairy Operator’s Group (cited in Lang (2010)), this rate is deflated by Statistics Canada’s index of agricultural wage (2000-2007) and the Ontario wage rate (2008-2009). Both data sets provide the total quantity of farm labour.

13In the data the median farm had a hired labour expense of $17,090, the first and third quartile had expenses of $3,490 and $57,170, suggesting some farms use significant hired labour - others very little.

14The index of agricultural wage was discontinued in 2007
4.1.6 Capital

As has been presaged, the over-capitalization of dairy farms is of particular interest in this thesis. The descriptive statistics in table 4.3 reveal that on a per-cow basis Ontario has more than double the capitalization of New York farms. Table 4.5, breaks down this difference into three categories, real estate, machinery and livestock, revealing that Ontario farms have significantly more real estate and machinery than their counterparts in New York, but a lower value of livestock on a per cow basis.

Table 4.5: Assets of Ontario dairy farms as a proportion of assets of New York dairy farms

<table>
<thead>
<tr>
<th></th>
<th>Real estate</th>
<th>Machinery</th>
<th>Livestock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per farm</td>
<td>1.83</td>
<td>1.28</td>
<td>0.54</td>
</tr>
<tr>
<td>Per cow</td>
<td>2.98</td>
<td>2.08</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Total capital expense is the summation of the opportunity cost of holding capital stock, depreciation, rental costs and maintenance. The opportunity cost (price) of capital is calculated using the weighted average cost of capital used by Coelli et al. (2003):

\[
\text{Cost of capital} = [(1 - g) * r_e] + [g * r_d] + \delta
\] (4.3)

where \( g \) is a measure of leverage, \( \frac{\text{Debt}}{\text{Physical assets}} \), \( r_e \) and \( r_d \) represent the cost of equity and debt, and \( \delta \) is the depreciation rate. To be consistent with a long-term orientation these parameters were estimated as ten-year real averages of corporate

\(^{15}\)The New York dataset does not disaggregate land and buildings. In Ontario land and improvements account for 82% of real estate. Neither data set allows for a calculation of land prices, and thus they are not considered in this paper. This limitation is further discussed in chapter 7.
bond yields, and the prime business rate plus two percent, respectively. Both datasets offer depreciation costs for real estate and equipment. These costs are divided by the stock of the assets to generate depreciation rates.

Rental costs of land, equipment and livestock were directly available from the dataset, as were other costs such as repairs and purchases of non-depreciable equipment. These were added to the opportunity cost of capital to generate the total capital expense.

4.1.7 Other

Other inputs include fertilizer, seed, spray, insurance, utilities, fuel, bedding, veterinary services and drugs, and breeding expenses. Price indices for these inputs are given by the United States Department of Agriculture and Statistics Canada.\footnote{Data on long-term corporate bond yields, chartered bank administered interest rates - business prime, and inflation are taken from Statistics Canada. Bond yields for 2007-2009 were missing and proxied by American bond rates. Moody’s BAA bond rates and business prime rates for the United States were taken from the Federal Reserve: \url{http://www.federalreserve.gov/releases/h15/data.htm}, while inflation was proxied by the consumer price index, available from the Department of Labour: \url{http://www.bls.gov/cpi/home.htm} \# data.}

\footnote{Specific indices were available for fertilizer, seed and fuel (2000-2009), and pesticides, utilities, bedding, veterinary expenses, and breeding (2000-2007), general farm costs was used to index missing years and categories.}
Chapter 5  Non-Parametric Estimation

Empirical techniques for efficiency analysis are divided into two general schools: econometric and linear programming. Typically, econometric methods have relied on strong assumptions regarding the distribution of inefficiency, and/or the technology, being estimated. In contrast, linear programming methods make no distributional assumptions, but do not allow for stochasticity in the data generating process. Efficiency estimation using linear programming is explored in this chapter, while a more detailed discussion of econometric approaches is reserved for the subsequent chapter.

5.1 Data envelopment analysis

Data envelopment analysis (DEA) is the most widely used programming approach in efficiency measurement. Developed by Charnes, Cooper, and Rhodes (1978), DEA relies on the specification of a production possibilities set, and a corresponding frontier. It then measures a producer’s inefficiency as the distance from their observed netput (output and input) to this frontier.

As in chapter 3, the production possibilities set ($\Psi$) can be defined as
\[ \Psi = \{(x, y) \in \mathbb{R}^{N+M}_+ : x \text{ can produce } y\}, \quad (5.1) \]

where \( x \in \mathbb{R}^N_+ \) and \( y \in \mathbb{R}^M_+ \) represent the input and output vector, respectively. The feasible set is convex, monotonic, and includes all observations. Efficient firms operate on the frontier of the set (the boundary of the feasible set and its complement), while inefficient firms operate in the interior of the set. DEA estimates this set by creating conical hull around the observations such that,

\[ \hat{\Psi} = (x, y) \in \mathbb{R}^{N+M} \quad (5.2) \]

where

\[
\begin{align*}
  y &\leq \sum_{i=1}^{n} \lambda_i y_i \\
x &\geq \sum_{i=1}^{n} \lambda_i x_i \\
\lambda_i &\geq 0 \text{ for } i = 1, 2 \ldots n,
\end{align*}
\]

where \( \hat{\Psi} \) is an estimate of \( \Psi \), and \( \lambda \) is a vector of size \( n \). Additional assumptions can be made regarding returns to scale, by restricting the summation of \( \lambda \). The two most common specifications are constant returns to scale, which imposes no restriction, and variable returns to scale, which requires \( \sum_{i=1}^{n} \lambda_i = 1 \). This paper assumes variable returns to scale throughout the estimation, and later tests this assumption by computing the scale efficiency of farms (measured by dividing efficiency scores derived under the assumption constant and variable returns).
Following Farrell (1957), a producer’s technical efficiency can be measured as the radial distance from their observation, to the frontier of the input set. Allowing \( \hat{\theta} \) to be the technical efficiency score

\[
\hat{\theta}(x, y) = \min\{\theta|(\theta x, y) \in \Psi\}.
\] (5.3)

Notably, the radial distance function differs from the Koopman’s definition, as the productions possibilities set is created using inequalities, and may result in certain inputs being greater than they need be to produce the same level of output. Scale efficiency can then be calculated as the ratio of a producer’s technical efficiency when measured under constant and variable returns to scale

\[
\widehat{SE} = \frac{\hat{\theta}_{CRS}}{\hat{\theta}_{VRS}},
\] (5.4)

where \( \hat{\theta}_{CRS} \) and \( \hat{\theta}_{VRS} \) denote technical efficiency scores estimated under the assumptions of constant and variable returns to scale, respectively. Intuitively, under constant returns to scale each producer is benchmarked against another producer (or a convex combination of producers) operating at the most efficient scale. Conversely, under variable returns to scale, each producer is benchmarked against another producer (or convex combination of producers) operating at the same scale. Dividing one metric by the other captures the effect of producing at an inefficient scale.

The computation of cost efficiency is done by first specifying the minimum achievable cost \( (C_i^*) \), given the output of the firm
where $w_i$ represent the input prices of the $i$th producer. The optimal cost of producer $i$, is then divided by their observed cost ($C_i$) to generate cost efficiency scores,\

$$\text{Cost Efficiency}_i = \frac{C^*_i}{C_i}. \quad (5.6)$$

Finally, allocative efficiency is calculated as cost efficiency divided by technical efficiency.

### 5.2 Statistical underpinning of DEA

Kneip, Park, and Simar (1998) demonstrate that under the assumptions of convex technology DEA provides consistent estimation of efficiency scores. Unfortunately, the authors show that the rate of convergence is a function of the dimensionality of the model:

$$\hat{\theta}_{VRS}(x, y) - \theta(x, y) = O_p(n^{-\frac{2}{n+1}}). \quad (5.7)$$

Thus while DEA is consistent, its rate of convergence is much slower than most parametric estimators (which typically exhibit "root-n" convergence). In small samples DEA estimates of efficiency display an upward bias. Intuitively, additional data can only serve to shift the frontier further from a producer’s observation - it can never contract it.
Unfortunately, the distribution of the DEA estimator cannot be analytically determined (see Kneip, Simar, and Wilson (2003) for a discussion). In a series of papers Simar and Wilson have developed a bootstrapping technique for estimating the distribution of the DEA estimator.

To gain an understanding of the bootstrapping procedure, define the true population as $\rho$. Some sample from this population, defined $\chi_n$, is used to calculate $\hat{\Psi}$ and $\hat{\theta}$. Now suppose a subsample, $\chi^*_n$, is drawn from the sample data (which can now be termed $\hat{\rho}$) and use this to calculate $\hat{\Psi}^*$ and $\hat{\theta}^*$. Providing the bootstrap is consistent, the bootstrap estimate of the sample, will behave as the sample estimate of the population,

$$ (\hat{\theta}^*(x, y) - \hat{\theta}(x, y)) | \hat{\rho}(\chi_n) \approx (\hat{\theta}(x, y) - \theta(x, y)) | \rho(\chi_n). \quad (5.8) $$

### 5.3 Bootstrapping methodology

Simar and Wilson (2008) define two consistent bootstrapping methods for DEA scores. The first draws subsamples of size $m < n$ from a sample of size $n$. The second, termed the homogenous bootstrap, relies on smoothing techniques to generate pseudo samples. Monte Carlo simulations by Kneip, Simar, and Wilson (2003) have shown the first method to be quite sensitive to the choice of $m$, but are unable to provide guidance as to its appropriate size.\(^1\) Thus, the second methodology is employed here.

\(^1\)A recently published paper, Simar and Wilson (2011), does offer more guidance on the appropriate selection of $m$. 
The homogenous bootstrap relies on the simplifying assumption that the
distribution of technical efficiency is homogenous over the input-output space. This
allows a bootstrapping technique similar to what one might find in parametric
literature, wherein the bootstrap relies on the perturbation of the residuals. In this
instance the residuals are akin to the distance function \( \delta(x, y) = 1/\theta(x, y) \).
Essentially, after calculating distance functions for every producer, the bootstrap
contracts the inputs of a producer to the estimated frontier and then expands the
inputs according to a randomly drawn distance function. A new frontier is then
calculated using these bootstrapped values. The bootstrapped inputs for the \( i \)
producer, denoted \( x_i^* \) is calculated as
\[
x_i^* = x_i \frac{\delta_i^*}{\hat{\delta}_i},
\]
where \( x_i \) is observed input, \( \hat{\delta}_i \) is an estimated distance function and \( \delta_i^* \) is a randomly
drawn distance function. A producer’s bootstrapped efficiency score measures the
distance from \( x_i \) to this new frontier.
Naively drawing \( n \) distance functions would result in a biased estimator, due to the
mass of distance functions at one.\(^2\) To avoid this problem, Simar and Wilson (2008),
suggest reflecting distance functions around unity and generating the density
function
\[
\hat{g}_h(t) = (2nh)^{-1} \sum_{i=1}^{n} \left[ K \left( \frac{t - \hat{\delta}_i}{h} \right) + K \left( \frac{t - (2 - \hat{\delta}_i)}{h} \right) \right],
\]
\(^2\)In the true distribution, the probability of any one point, is of course zero. Given DEA is a benchmarking
technology, it typically finds multiple firms to have an efficiency score of 1.00.
where \( n \) is the sample size, \( h \) is the bandwidth and \( K \) is the standard kernel density estimator. This distribution can then be truncated, such that

\[
\hat{f}_h(t) = \begin{cases} 
2\hat{g}_h(t) & \forall \ t \geq 1 \\
0 & \text{otherwise}.
\end{cases}
\] (5.11)

Simar and Wilson (2008) note that the employment of a different kernel distribution (i.e. Epanechnikov), is of little importance in the minimization of the average mean integrated squared error. However, the result is quite sensitive to the bandwidth \((h)\).

In kernel density estimation, as \( h \) goes to infinity, the density estimate degenerates to a horizontal line. Conversely, as \( h \) goes to zero the estimate degenerates to discrete masses of \( 1/n \) at each observed point. Thus, increasing the bandwidth reduces the variance, but results in a larger bias. The data driven approach to bandwidth selection employed here, is the minimization of the mean-integrated squared error (MISE).

Following Simar and Wilson (2008), the optimal bandwidth is found by first deleting all observations which equal one, and estimating \( \hat{g}_h(t) \) with the remaining observations (letting \( m \) equal the total observations remaining after deletion). The MISE of this estimator can therefore be defined as

\[
MISE_{\hat{g}_h} = E \left[ \int_{-\infty}^{\infty} [\hat{g}_h(\delta) - g(\delta)]^2 d\delta \right].
\] (5.12)

Equation 5.12 can be estimated through least-squares cross-validation function
\[ CV(h) = \int_{-\infty}^{\infty} \hat{g}_h^2(\delta) d\delta - \frac{1}{2m} \sum_{i=1}^{2m} \hat{g}^2_{h,(i)}(\hat{\delta}_i), \] (5.13)

where \( \hat{g}_{h,(i)} \) is the leave one out estimator of \( g(\delta) \). Denoting the bandwidth that minimizes equation 5.13 as \( h_{cv} \), the optimal bandwidth \( h_{CV} \) can be defined as

\[ h_{CV} = h_{cv} \left( \frac{2m}{n} \right)^{1/5} \left( \frac{\sigma_n}{\sigma_{2m}} \right), \] (5.14)

where \( \sigma_n \) and \( \sigma_{2m} \) represent the standard deviation of the original \( n \) observations and the \( 2m \) reflected observations, respectively. Essentially, \( h_{CV} \) is simply \( h_{cv} \), adjusted to account for the size and variance differences between \( 2m \) and \( n \).

Fortunately, for the applied researcher, the kernel distributions need not be estimated (Silverman, 1986). Instead, the following algorithm proposed by Simar and Wilson (1998), yields the same outcome:

1. Compute distance functions \( (\delta_i) \) for each producer from the original sample.

2. Select a value for the bandwidth.

3. Draw \( n \) samples \( (b_1, \ldots, b_n) \), with replacement, from

\[ D_{2n} = \{ \delta_1, \ldots, \delta_n, (2 - \delta_1), \ldots, (2 - \delta_n) \}. \] (5.15)

4. Draw \( n \) samples of \( \epsilon_i \), from a standard normal distribution, distributed \((0,1)\), and compute

\[ b_i^* = b_i + h_{CV} \epsilon_i, \text{ for } i = 1, \ldots, n. \] (5.16)
5. In finite samples the mean and variance of, \( b_i^* \) will differ from \( D_{2n} \). This can be corrected by computing

\[
b_i^{**} = \bar{b}_i + \frac{b_i^* - \bar{b}_i}{(1 - h_{CV}^2 \sigma_i^2)} \delta_i.
\]  \( (5.17) \)

6. Correct for the reflection by calculating \( \delta^* \) as

\[
\delta^* = \begin{cases} 
2 - b_i^{**} & \forall b_i^{**} \leq 1 \\
b_i^{**} & \text{otherwise.}
\end{cases}
\]  \( (5.18) \)

7. Define the bootstrap sample as \( \chi_n^* = \{(x^*, y)| i = 1, \ldots, n \} \), where

\[
x_i^* = \frac{\delta_i^*}{\delta_i} x_i.
\]  \( (5.19) \)

8. Compute distance functions based on the sample \( \chi_n^* \).

9. Repeat B times.

This procedure will generate B bootstrap estimates for each producer. Given the consistency of the bootstrap, the distribution of the bootstrap mimics the distribution of the true sample.

### 5.4 Bias correction

Defining the inverse of the bootstrap estimates as \( \hat{\theta}_i^*(x, y) \), the (small sample) bias of the DEA estimator is
\[
\text{bias}_B(\hat{\theta}(x,y)) = B^{-1} \sum_{b=1}^{B} \hat{\theta}^*_b(x,y) - \hat{\theta}(x,y). \tag{5.20}
\]

Hence, bias corrected efficiency scores can be calculated by

\[
\hat{\theta}(x,y) = \hat{\theta}(x,y) - \text{bias}_B(\hat{\theta}(x,y)) = 2\hat{\theta}(x,y) - B^{-1} \sum_{b=1}^{B} \hat{\theta}^*_b(x,y). \tag{5.21}
\]

The variance of the bootstrap values is given by

\[
\hat{\sigma}^2 = B^{-1} \sum_{i=1}^{B} \left[ \hat{\theta}^*_i(x,y) - B^{-1} \sum_{i=1}^{B} \hat{\theta}^*_i(x,y) \right]^2. \tag{5.22}
\]

By equation 5.21, the bias-corrected estimates will have a variance of \(4\hat{\sigma}^2\) (assuming away correlation between the estimated efficiency scores and the summation term.)

Minimization of the mean squared error would, therefore, require bias correction only in the event that,

\[
\frac{|\text{bias}_B(\hat{\theta}(x,y))|}{\hat{\sigma}} > \frac{1}{\sqrt{3}}. \tag{5.23}
\]

### 5.5 Confidence intervals

With a known distribution one could easily compute confidence intervals by finding values of \(c\) such that

\[
\text{prob}(c_{\alpha/2} \leq \hat{\delta}(x,y) - \delta(x,y) \leq c_{1-\alpha/2}) = 1 - \alpha, \tag{5.24}
\]
where $c_\alpha$ represents the $\alpha$th quantile, of the sample distribution $(\hat{\delta}(x, y) - \delta(x, y))$.

And then defining the confidence intervals

$$CI_{\alpha/2} = \frac{1}{\hat{\delta}(x, y) - c_{\alpha/2}} \leq \theta(x, y) \leq \frac{1}{\hat{\delta}(x, y) - c_{1-\alpha/2}} = CI_{1-\alpha/2}$$

(5.25)

Of course, the distribution of $(\hat{\delta}(x, y) - \delta(x, y))$ is not known, however it is mimicked by the empirical distribution of $(\hat{\delta}^*_b(x, y) - \hat{\delta}(x, y), b = 1, \ldots, B)$. Thus $c$ can be estimated as

$$\text{prob}(\hat{c}_{\alpha/2} \leq \hat{\delta}^*(x, y) - \hat{\delta}(x, y) \leq \hat{c}_{1-\alpha/2} = 1 - \alpha.$$  

(5.26)

Empirically this merely involves finding the quantiles of the bootstrapped confidence intervals associate with $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$. Substituting the empirical confidence intervals into equation 5.25 yields,

$$CI_{\alpha/2} = \frac{1}{\hat{\delta}(x, y) - \hat{c}_{\alpha/2}} \leq \theta(x, y) \leq \frac{1}{\hat{\delta}(x, y) - \hat{c}_{1-\alpha/2}} = CI_{1-\alpha/2}.$$  

(5.27)

5.6 DEA and panel data

The inconstancy of the production possibility set over time, complicates the use of DEA. Theoretically, technological change shifts the frontier between time periods. For example, if technical change was always progressive, then aggregating across time periods would result in observations from early time periods being biased.

Fried, Lovell, and Schmidt (2008), survey the literature and offer the applied researcher several options when employing DEA with panel data. Most simply, if
one were willing to assume away technological change, the data could be pooled across time periods, and one single frontier estimated. Conversely, if one was unwilling to make any assumptions regarding technical change, separate frontiers could be estimated for each year. One intermediate option, termed "window analysis" calculates efficiency scores for each year based on multiple production possibilities sets, with each set constructed using observations from different years. Efficiency scores for any one period can then be calculated as a weighted average of the scores from each window. A second intermediate option is to assume only technological progress - allowing frontiers for any period $t^*$, to be constructed using observations from all periods $t < t^*$.

This study employs three different DEA models: a metafrontier model (in which data are pooled across time periods), a disaggregated model (which calculates separate frontiers for each year) and an intermediate model. The intermediate model introduces a new method of pooling data, based upon the computed level of technological change. For each pair of periods in the sample, the amount technical change that has occurred between these periods is measured. If significant technical progress has been found to occur, then subsequent periods will not be used in the construction of the frontier, conversely, if technical regress is found to be significant then previous periods are excluded.

Technical change is calculated through the Malmquist index. This index, originally defined by Caves, Christensen, and Diewert (1982), measures productivity changes.

---

Footnote:

3For example, in calculating scores for period $t$, one window might create a production possibilities set using observations from $t - 3$ to $t$, the next using data from $t - 2$ to $t + 1$, and so on.
between two time periods 0 and 1 by:

\[ M = \frac{\delta(x^1, y^1)}{\delta(x^0, y^0)}, \] (5.28)

where \( \delta \), is a distance function. When employing the Shepherd distance function, one faces the dilemma of whether to employ the technology of period 0 or period 1 as the benchmark. Following Färe et al. (1994), the geometric mean of the two measures is employed in this paper,

\[ M = \left( \frac{\delta^0(x^1, y^1) \delta^1(x^1, y^1)}{\delta^0(x^0, y^0) \delta^1(x^0, y^0)} \right)^{1/2}. \] (5.29)

The index can then be decomposed into efficiency change (EC) and technical change (TC);

\[ M = EC \times TC, \] (5.30)

where

\[ EC = \frac{\delta^1(x^1, y^1)}{\delta^0(x^0, y^0)}, \] (5.31)

and

\[ TC = \left( \frac{\delta^0(x^1, y^1) \delta^0(x^0, y^0)}{\delta^1(x^1, y^1) \delta^1(x^0, y^0)} \right)^{1/2}. \] (5.32)

Simar and Wilson (1999) outline a bootstrapping approach for Malmquist indices (and their components) which is similar in nature to the bootstrapping of DEA scores. The slight difference being that efficiency scores within the two periods in
question are likely highly correlated, requiring the employment of a bivariate kernel function. Once again Simar and Wilson, suggest the reflection method of Silverman (1986), this time with two boundaries. This distribution can be derived by first allowing \( A \) to be a vector of distance functions from period 1, and \( B \) a vector of distance functions from period two, and defining

\[
\Delta = \begin{bmatrix}
A & B \\
2 - A & B \\
2 - A & 2 - B \\
A & 2 - B 
\end{bmatrix}
\]

The covariance of the \( N \times 2 \) matrices \( \begin{bmatrix} A & B \end{bmatrix} \) or \( \begin{bmatrix} 2 - A & 2 - B \end{bmatrix} \) is given by

\[
\hat{\Sigma} = \begin{bmatrix}
\hat{\sigma}_1^2 & \hat{\sigma}_{12} \\
\hat{\sigma}_{21} & \hat{\sigma}_2^2 
\end{bmatrix}
\]

while the covariance matrix of the \( N \times 2 \) matrices \( \begin{bmatrix} 2 - A & B \end{bmatrix} \) or \( \begin{bmatrix} A & 2 - B \end{bmatrix} \) is given by

\[
\hat{\Sigma}_R = \begin{bmatrix}
\hat{\sigma}_1^2 & -\hat{\sigma}_{12} \\
-\hat{\sigma}_{21} & \hat{\sigma}_2^2 
\end{bmatrix}
\]

These components allow us to define the kernel distribution of the reflected values as,

\[
\hat{g}(z) = \frac{1}{4Nh^2} \sum_{j=1}^{4N} K_j \left( \frac{z - \Delta_j}{h} \right)
\]
where $\Delta_j$ is the $j^{th}$ row of $\Delta$, $h$ is the bandwidth, and $z$ has dimensions 1X2. $K_j()$ is a bivariate normal distribution with shape $\Sigma$ for $j = 1, \ldots, N, 2N + 1, \ldots, 3N$ and $\Sigma_R$ for $j = N + 1, \ldots, 2N, 3N + 1, \ldots, 4$. Finally, the consistent estimate of the density is given by:

$$
\hat{g}^*(z) = \begin{cases} 
4\hat{g}(z) & \text{for } z_1 \geq 1, z_2 \geq 1, \\
0 & \text{otherwise.}
\end{cases} 
$$

(5.37)

Fortunately, the kernel distribution need not be estimated for random sampling. Instead one need only compute the following algorithm, adapted from Silverman (1986):

1. Compute distance functions ($\delta_{ij}$) for each producer in each period, and calculate $\Delta$.

2. Select a value for the bandwidth.

3. Compute

$$
\Gamma = (1 + h^2)^{-1/2} \left( \Delta^* + h\epsilon^* - C \begin{bmatrix} \bar{\delta}_1 & 0 \\ 0 & \bar{\delta}_2 \end{bmatrix} \right) + C \begin{bmatrix} \bar{\delta}_1 & 0 \\ 0 & \bar{\delta}_2 \end{bmatrix}
$$

(5.38)

where $\Delta^*$ are a random draw of $N$ rows (with replacement) from $\Delta$. $\epsilon$ is a draw from the bivariate normal density with shape $\Sigma$ if $\Delta_i^*$ was drawn from a row $j = 1, \ldots, N, 2N + 1, \ldots, 3N$ of $\Delta$, or $\Sigma_R$ if $\Delta_i^*$ was drawn from a row $j = N + 1, \ldots, 2N, 3N + 1, \ldots, 4$ of $\Delta$. $C$ is a $N$x2 matrix of ones and $\bar{\delta}_j$ is the mean of the $j^{th}$ column of $\Delta^*$.
4. For each element $\gamma_{ij}$ of the set $\Gamma$, calculate

$$
\gamma^*_ij = \begin{cases} 
\gamma_{ij} & \text{if } \gamma_{ij} \geq 1 \\
2 - \gamma_{ij} & \text{otherwise.}
\end{cases}
$$

(5.39)

5. Modify the input set

$$
x^*_ij = \frac{\gamma_{ij}}{\delta_{ij}}x_{ij}.
$$

(5.40)

6. Calculate new distance functions and corresponding Malmquist indices.

7. Repeat $B$ times.

Bias correction and confidence intervals can be calculated as previously described.

For each pair of periods, a bootstrapped Malmquist index is calculated and for each producer, the following two hypotheses are calculated

$$
H^1_o : TC_i \leq 1 \\
H^2_o : TC_i \geq 1.
$$

(5.41)

Rejection of the first suggests that technological regress has occurred, while rejection of the second implies technological progress. Of course, the Malmquist index calculates producer specific indices, thus it occurs that the null hypothesis is rejected for some producers, while being accepted for others. The most cautious approach would be to reject technological constancy if any producer exhibits significant technological change. I invoke a more liberal, through unfortunately
arbitrary, rule under which technological constancy is rejected if at least 10% of producers experience significant (at the 10% level) technical change.

A further complication is the unbalanced nature of the data sets. Years which are far apart share few producers (i.e. only 22 producers, all from New York, are in the 2000 and 2009 samples). Given producers typically remain in the Ontario sample for only four to five years, pairs of observations more than four years apart are not considered.

5.7 Outlier detection

As noted elsewhere, efficiency scores derived from data envelopment analysis are sensitive to data outliers. As Thanassoulis, Portela, and Despic (2008) note, the detection and removal of such "outliers" is governed by the judgement of the researcher, as opposed to any hard and fast rule. Furthermore, "outliers" in and of themselves do not warrant removal, if a firm is truly much more efficient than its peers, then there is no basis for excluding it from analysis. Removal should only be undertaken if, in the opinion of the researcher, the observation has been subject to measurement error.

To make this determination I rely on the panel nature of the data and check outliers against observations on the same producer, in previous and subsequent periods. The procedure used for outlier detection is loosely based on Wilson (1995). For each year the following procedure is used:

---

4Producers in New York are more likely to span multiple years than producers in the Ontario data set. Given the maintained assumption that both regions can access the same technology this is a problem only in so much as New York producers are larger and technological change can be scale biased.
• Calculate baseline efficiency scores.

• Calculate "super-efficiency" scores by removing observation $k$ from the frontier, and calculating the efficiency score of $k$. Since $k$ is not in the frontier, its efficiency can be greater than 100%.

• Producers whose "super-efficiency" scores are above 130% are sorted in decreasing order of their "super-efficiency" scores.

• For each producer, $k$, above this threshold calculate two new DEA models; one with the $k$th observation missing, and a second with with all observations with super-efficiency scores equal or greater to $k$’s efficiency score missing.

• Compare these new models to the baseline model by examining: the mean change in efficiency and the number of observations whose efficiency scores have changed.

The procedure above identified two metrics of defining outliers - the number of producers who are projected onto the super-efficient observation, and the mean effect of removing the super-efficient producer. It was found that the first metric was a very blunt instrument, fifty-two producers were found to influence more than 30% of other observations - of these most did not appear to have any abnormalities.

The second metric was found more useful. Producers who caused the mean efficiency of the sample to change more than 1% were selected for closer examination. None of the identified producers showed any abnormalities when the super-efficient observation was compared to other observations on the producer in
past and future periods.

5.8 Results

Three DEA models are calculated: metafrontier, disaggregated and intermediate. For each model both a simple average and weighted average (based on milk output) of efficiency scores is tabulated. Across all models variable returns to scale is assumed.

The metafrontier model assumes away technical change, and constructs a production possibilities set using all observations. In contrast, the disaggregated model, constructs a unique production possibilities set for each year, using only observations from that year. The efficiency scores (both bias-corrected and uncorrected) from both of these models are reported in tables 5.1 and 5.2, where weighted average scores are reported based on milk output.

<table>
<thead>
<tr>
<th></th>
<th>Simple Averages</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Technical Eff</td>
<td>Allocative Eff</td>
<td>Cost Eff</td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>0.698 (0.757)</td>
<td>0.874 (0.885)</td>
<td>0.608 (0.668)</td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td>0.632 (0.719)***</td>
<td>0.773 (0.784)***</td>
<td>0.486 (0.56)***</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.667 (0.739)</td>
<td>0.827 (0.838)</td>
<td>0.551 (0.617)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simple Averages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Weighted Averages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>0.729 (0.823)</td>
<td>0.881 (0.892)</td>
<td>0.642 (0.735)</td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td>0.647 (0.754)***</td>
<td>0.786 (0.791)***</td>
<td>0.508 (0.597)***</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.701 (0.799)</td>
<td>0.849 (0.857)</td>
<td>0.596 (0.687)</td>
<td></td>
</tr>
</tbody>
</table>

Non-corrected scores are in parenthesis.

*, **, **** represent significant differences between regions at the .1, .05 and .01 levels, respectively.
Table 5.2: Efficiency scores from disaggregated Model

<table>
<thead>
<tr>
<th></th>
<th>Technical Eff</th>
<th>Simple Averages</th>
<th>Allocative Eff</th>
<th>Cost Eff</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>0.759 (0.844)</td>
<td>0.884 (0.88)</td>
<td></td>
<td>0.67 (0.741)</td>
</tr>
<tr>
<td>Ontario</td>
<td>0.721 (0.805)***</td>
<td>0.745 (0.743)****</td>
<td>0.536 (0.594)***</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.741 (0.825)</td>
<td>0.819 (0.815)</td>
<td></td>
<td>0.607 (0.672)</td>
</tr>
<tr>
<td><strong>Weighted Averages</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>0.769 (0.885)</td>
<td>0.917 (0.912)</td>
<td>0.704 (0.809)</td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td>0.732 (0.832)***</td>
<td>0.751 (0.736)***</td>
<td>0.549 (0.609)***</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.757 (0.868)</td>
<td>0.863 (0.855)</td>
<td>0.654 (0.744)</td>
<td></td>
</tr>
</tbody>
</table>

Non-corrected scores are in parenthesis.

*, **, **** represent significant differences between regions at the .1, .05 and .01 levels, respectively.

The intermediate model relies on the calculation of different production possibilities frontier for each of the ten years, however each set is constructed using, not only the observations for that year, but also observations from other years which have the same technology. It is assumed two years have similar technology if fewer than 10% of producers are deemed to have significant technological change (measured at the 10% level) within these years. As mentioned previously, the unbalanced nature of the data constrains us to consider periods a maximum of four years apart. Technical change is measured via the Malmquist productivity index, results from which are

\[ R_{ON/NY} = \frac{\bar{\theta}_{ON}}{\bar{\theta}_{NY}} \]  

(5.42)

Using the same methods as before construct confidence intervals can be constructed for \( R_{ON/NY} \), and if unity is not included in these intervals, then the null hypothesis, \( \bar{\theta}_{ON} = \bar{\theta}_{NY} \), can be rejected.
presented in table 5.3.
Table 5.3: Results of Malmquist index

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>0.81</td>
<td>0.805</td>
<td>0.832</td>
<td>0.811</td>
<td>0.82, 0.78</td>
<td>0.87, 0.02, 55</td>
<td>0.85, 0.02, 46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0.994</td>
<td>1.009</td>
<td>0.954</td>
<td>0.956</td>
<td>0.12, 0.17, 103</td>
<td>0.38, 0.03, 64</td>
<td>0.3, 0.07, 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>1.038</td>
<td>0.984</td>
<td>1.024</td>
<td>1.052</td>
<td>0.06, 0.02, 84</td>
<td>0.08, 0.26, 61</td>
<td>0.04, 0.5, 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>0.966</td>
<td>1.025</td>
<td>1.123</td>
<td>1.052</td>
<td>0.34, 0.112</td>
<td>0.24, 0.29, 84</td>
<td>0.04, 0.26, 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>1.165</td>
<td>1.15</td>
<td>1.138</td>
<td>1.225</td>
<td>0.79, 80</td>
<td>0.68, 59</td>
<td>0.02, 0.82, 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>1.043</td>
<td>1.026</td>
<td>1.15</td>
<td>1.061</td>
<td>0.07, 0.29, 96</td>
<td>0.2, 70</td>
<td>0.45, 49</td>
<td>0.08, 0.39, 36</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>0.955</td>
<td>1.035</td>
<td>0.979</td>
<td>0.35, 0.95</td>
<td>0.18, 0.24, 62</td>
<td>0.31, 0.148</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.997</td>
<td>0.972</td>
<td>0.31, 0.19, 95</td>
<td>0.26, 0.2, 74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>1.005</td>
<td>0.09, 97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first line in each cell represents the average technical change that occurred in the period - a value less than one implies technological regress, a value greater than one represents technological progress. The second line contains the proportion of producers who experienced statistically significant technological regress, the proportion of producers who experienced statistically significant technological progress and the number of producers who have observations in each of the two years. If the first number in the second line is less than .1, then the earlier period is included in the production possibilities set of the later. If the second number on the second line is less than .1 then the later period is included in the production possibilities set of the earlier.
Table 5.4: Efficiency scores from intermediate model

<table>
<thead>
<tr>
<th></th>
<th>Technical Eff</th>
<th>Allocative Eff</th>
<th>Cost Eff</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New York</strong></td>
<td>0.752 (0.823)</td>
<td>0.897 (0.893)</td>
<td>0.674 (0.734)</td>
</tr>
<tr>
<td><strong>Ontario</strong></td>
<td>0.722 (0.81)***</td>
<td>0.797 (0.8)***</td>
<td>0.575 (0.649)***</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>0.738 (0.817)</td>
<td>0.85 (0.849)</td>
<td>0.628 (0.694)</td>
</tr>
</tbody>
</table>

**Weighted Averages**

<table>
<thead>
<tr>
<th></th>
<th>Technical Eff</th>
<th>Allocative Eff</th>
<th>Cost Eff</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New York</strong></td>
<td>0.763 (0.87)</td>
<td>0.922 (0.916)</td>
<td>0.703 (0.798)</td>
</tr>
<tr>
<td><strong>Ontario</strong></td>
<td>0.731 (0.832)***</td>
<td>0.809 (0.802)***</td>
<td>0.591 (0.669)***</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>0.753 (0.858)</td>
<td>0.885 (0.879)</td>
<td>0.667 (0.757)</td>
</tr>
</tbody>
</table>

*, **, **** represent significant differences between regions at the .1, .05 and .01 levels

As table 5.3, indicates, the sample shows more technological regress than progress. Other studies that have found technological regress in dairies (Kumbhakar et al., 2008), have suggested increased regulations (such as environmental or marketing restrictions) may be the causal factor. In the present case, I know of no significant regulatory changes.

Using the decision rule of including years in which no more than 10% of producers experienced technical change, allowed for the inclusion of at least one additional reference year in the construction of each production possibilities set. The maximum number of years used in the construction of any set was 6 (in 2006), and the average number of additional years included was 2.6. The results from the intermediate model is shown in table 5.4, and density plots of efficiency scores from this model are shown in appendix C.

The correlation and rank correlation between the three models is found in table 5.5, demonstrating a positive, though somewhat low correlation across models. No previous studies could be found, examining the correlation of DEA scores from
Table 5.5: Correlation and rank correlation of the three DEA models

<table>
<thead>
<tr>
<th></th>
<th>Disaggregated Model</th>
<th>Intermediate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metafrontier Model</td>
<td>Correlation</td>
<td>0.461, 0.373, 0.574</td>
</tr>
<tr>
<td></td>
<td>Rank Correlation</td>
<td>0.443, 0.259, 0.553</td>
</tr>
<tr>
<td>Disaggregated Model</td>
<td>Correlation</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Rank Correlation</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td>0.552, 0.662, 0.697</td>
</tr>
<tr>
<td></td>
<td>Rank Correlation</td>
<td>0.552, 0.662, 0.697</td>
</tr>
</tbody>
</table>

Each cell contains the correlation for technical, allocative and cost efficiency, respectively.

different panel models. Studies which have compared efficiency scores from differing efficiency estimators (i.e. DEA and regression analysis), found similar magnitudes of correlation (Coelli and Perelman, 1999; Mbaga et al., 2003; Drake and Simper, 2003).

As the results indicate the three methodologies agree on a general narrative, if not actual efficiency scores. Based on the weighted averages, New York dairy farms were found to be three to eight percentage points more technically efficient than Ontario farms (significant for all weighted averages). However, the main difference between the two groups stemmed from allocative efficiency, which was over ten percentage points higher in New York across all models. Further discussion of the results is reserved for chapter 7.

Subsequent analysis is conducted using the intermediate model, which benefits from employing more data than the disaggregated model, but avoids the assumption of technological constancy, which is maintained by the metafrontier model.
5.9 Input usage

Cost efficiency is calculated by comparing the inputs of the producer in question, with the input vector that results in the lowest cost while producing the same output. Comparing the input vector of the producer in question to the cost minimizing vector, allows one to calculate the change in inputs that is necessary for the achievement of cost efficiency. These changes are outlined in table 5.6.

Table 5.6: Mean proportional change in inputs necessary to achieve cost efficiency

<table>
<thead>
<tr>
<th></th>
<th>Feed</th>
<th>Labour</th>
<th>Capital</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>-0.2690</td>
<td>-0.2158</td>
<td>-0.2851</td>
<td>-0.0763</td>
</tr>
<tr>
<td>Ontario</td>
<td>0.5589</td>
<td>-0.2196</td>
<td>-0.5539</td>
<td>-0.2819</td>
</tr>
<tr>
<td>Pooled</td>
<td>0.1185</td>
<td>-0.2176</td>
<td>-0.4109</td>
<td>-0.1725</td>
</tr>
</tbody>
</table>

New York producers are prescribed a relatively equiproportionate change in inputs. In contrast, Ontario producers would need to increase their purchased feed by 55.89%, while decreasing capital by 55.39%. Evidently, with such an increase in purchased feed, Ontario farms could grow less feed, facilitating the required reductions in labour, capital and other inputs. The dramatic reduction in capital should not be a surprise as table 4.5 showed the value of real estate and machinery on Ontario farms to be more than double the value held by New York farms, on a per-cow basis.
5.10 Sensitivity analysis

As described in chapter 4, the price of debt was assumed to equal the business prime lending rate plus two percent. In Ontario quota holdings are able to serve as collateral for business loans allowing farmers who own their quota outright, to gain access to preferential lending rates. Given that that debt is equivalent to 48.6% of total assets in Ontario (versus 34.7% in New York), these differences in interest rates may be non-trivial.

To determine the impact of lower interest rates, a sensitivity analysis is conducted wherein the interest rate is reduced by one, two and three percent. This reduction was simulated in Ontario only, and then in both Ontario and New York (resulting in six different models). Table 5.7, shows the impact of these changes on technical, allocative and cost efficiency.

<table>
<thead>
<tr>
<th></th>
<th>Ontario</th>
<th>Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>∆TE</td>
<td>∆AE</td>
</tr>
<tr>
<td>Ontario (-1%)</td>
<td>-0.0014</td>
<td>0.0015</td>
</tr>
<tr>
<td>Ontario (-2%)</td>
<td>-0.0026</td>
<td>0.0017</td>
</tr>
<tr>
<td>Ontario (-3%)</td>
<td>-0.0060</td>
<td>-0.0369</td>
</tr>
<tr>
<td>Both (-1%)</td>
<td>-0.0010</td>
<td>0.0028</td>
</tr>
<tr>
<td>Both (-2%)</td>
<td>-0.0018</td>
<td>0.0045</td>
</tr>
<tr>
<td>Both (-3%)</td>
<td>-0.0049</td>
<td>-0.0329</td>
</tr>
</tbody>
</table>

Under the assumptions made when pricing capital, reducing the price of capital has the peculiar effect of increasing the quantity of capital as expenses on rent and maintenance remain the same, but are deflated by a lower price. Thus when the
interest rate in Ontario is reduced the technical efficiency of Ontario producers falls. A change in the price of capital (or any other input for that matter), changes the first order condition for cost minimization. This change affects both producers on the cost frontier, and in the interior. Thus, overcapitalized producers may move closer to the frontier in this simulation, however, the frontier itself may shift as well, resulting in an ambiguous effect on mean efficiency.

As evidence of this ambiguity, when interest rates are reduced by one or two percentage points (in either Ontario or in both regions), the allocative efficiency of Ontario farmers increases slightly, but when interest rates are reduced by three percentage points, the mean allocative efficiency is actually reduced. Overall, this experiment shows that efficiency scores are insensitive to changes in the interest rate.

5.11 Second stage regression

In order to understand the determinants of efficiency a truncated regression is run using the efficiency scores as a dependant variable, and likely co-variates as the dependant variables. Previous studies have included such co-variates as the operator's age and education, milking and barn technology, participation in farm programs, level of debt, ratio of hired to family labour, and size. Unfortunately, the data sets in this study do not generate comparable variables for many of these co-variates. A description of the variables used in the study is given in table 5.8.

---

6Simar and Wilson (2007), prove that truncated regression provides consistent second-stage estimation, when using bootstrapped DEA scores.
Table 5.8: Independent variables in truncated regression

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Pooled (N=1410)</th>
<th>New York (N=750)</th>
<th>Ontario N=660</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>Dummy variable equal to one if producer is located in Ontario</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age</td>
<td>Age of the principal operator</td>
<td>47.99 (12.07)</td>
<td>48.72 (13.5)</td>
<td>47.17 (10.14)</td>
</tr>
<tr>
<td>Education</td>
<td>Dummy variable equal to one if producer has greater than high school education</td>
<td>0.55</td>
<td>0.51</td>
<td>0.6</td>
</tr>
<tr>
<td>bST</td>
<td>Per cow expense on bovine somatotropin, a hormone which is illegal in Canada</td>
<td>9.59 (26.05)</td>
<td>18.02 (33.53)</td>
<td>0</td>
</tr>
<tr>
<td>Herd size</td>
<td>Number of productive dairy cows</td>
<td>90.08 (115.34)</td>
<td>110 (148.27)</td>
<td>67.43 (49.89)</td>
</tr>
<tr>
<td>Tie-stall</td>
<td>Dummy variable equal to one if producer uses a tie-stall barn</td>
<td>0.67</td>
<td>0.62</td>
<td>0.74</td>
</tr>
<tr>
<td>Pipeline</td>
<td>Dummy variable equal to one if producer uses a pipeline for milking (bucket and dump system is the control technology)</td>
<td>0.67</td>
<td>0.63</td>
<td>0.72</td>
</tr>
<tr>
<td>Parlour</td>
<td>Dummy variable equal to one if producer uses a parlour for milking (bucket and dump system is the control technology)</td>
<td>0.28</td>
<td>0.29</td>
<td>0.26</td>
</tr>
<tr>
<td>Time dummies</td>
<td>Dummy variables for years 2001 to 2009 (year 2000 is the control group)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Mean values with standard deviation in parenthesis where applicable.
Two regressions are run on both technical and allocative inefficiency. The first model includes variables that are unaffected by government marketing regulations (such as location and operator characteristics) and the second model includes all the co-variates (some of which, such as size and technology, may be affected by marketing regulations). Theoretically, a farmer’s ability to choose the proper allocation of his inputs should not be influenced by his technology or use of bST, thus the regression on allocative efficiency excludes these variables.

The results of the truncated regressions is provided in table 5.9. As expected the dummy variable for Ontario, was slightly higher in the second model for both technical and allocative efficiency. Being located in Ontario is seen to increase technical inefficiency by 2.6-2.8%, while increasing allocative inefficiency by 21-23%. Age was found to increase inefficiency, with this increase being significant in both models of technical inefficiency and one model of allocative inefficiency. Education reduced technical and allocative inefficiency, though its impact was only significant on technical inefficiency.

Herd size was found to significantly reduce allocative efficiency, while having no impact on technical inefficiency; this result is further discussed in the next section. Per cow expense on bST significantly reduced technical inefficiency, this was expected as bST allows cows to produce more milk without the employment of additional inputs. The type of barn technology had no significant impact.

---

7Inefficiency is calculated as one minus (technical or allocative) efficiency. Inefficiency is used (as opposed to efficiency scores themselves), so that results are comparable with the one-step parametric regression in the next chapter.

8The table reports the coefficients of the regression - not the marginal effects. According to Greene (2008) the coefficients represent the effect of the independent variables on the population, whereas the marginal effects represent the effect of the independent variables on the (truncated) subpopulation.
Curiously, both a parlour and a pipeline milking system (thought of as more modern compared to a bucket and dump system), were found to increase technical inefficiency.\textsuperscript{9} The yearly time dummies have no economic interpretation, since DEA is a benchmarking technology, these coefficients reflect both movements in average efficiency and shifts in the frontier.\textsuperscript{10}

\textsuperscript{9}One can equally think of producers who use differing technologies as existing on separate frontiers. Following this argument, one should account for these technologies in the first stage of analysis. In keeping with the literature, these variable are included in the second stage because it lends an intuitive interpretation to the effect of technology on efficiency (something which is of interest to the audience of this study). Another practical consideration, is that DEA cannot accommodate dummy variables; their inclusion would result in separate models for each technology.

\textsuperscript{10}Technical change was previously calculated using the bootstrapped Malmquist index.
Table 5.9: Results of truncated regression on DEA scores

<table>
<thead>
<tr>
<th></th>
<th>Regression on technical inefficiency</th>
<th>Regression on allocative inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Constant</td>
<td>0.185 (0.0166)***</td>
<td>0.142 (0.0222)***</td>
</tr>
<tr>
<td>Region</td>
<td>0.0283 (0.00647)***</td>
<td>0.0259 (0.00653)***</td>
</tr>
<tr>
<td>Age</td>
<td>0.00162 (0.000255)***</td>
<td>0.00162 (0.000254)***</td>
</tr>
<tr>
<td>Education</td>
<td>-0.0137 (0.0061)***</td>
<td>-0.0102 (0.00615)*</td>
</tr>
<tr>
<td>bST</td>
<td>-0.000473 (0.000128)***</td>
<td>-0.000476 (0.000139)***</td>
</tr>
<tr>
<td>Herd size</td>
<td>–</td>
<td>2.97e-05 (3.22e-05)</td>
</tr>
<tr>
<td>Tie-stall</td>
<td>–</td>
<td>0.00876 (0.0134)</td>
</tr>
<tr>
<td>Pipeline</td>
<td>–</td>
<td>0.0407 (0.0164)**</td>
</tr>
<tr>
<td>Parlour</td>
<td>–</td>
<td>0.0297 (0.0159)*</td>
</tr>
<tr>
<td>2001</td>
<td>-0.0244 (0.0133)*</td>
<td>-0.0266 (0.0133)**</td>
</tr>
<tr>
<td>2002</td>
<td>0.042 (0.0133)***</td>
<td>0.0388 (0.0133)***</td>
</tr>
<tr>
<td>2003</td>
<td>-0.0443 (0.0136)***</td>
<td>-0.0471 (0.0136)***</td>
</tr>
<tr>
<td>2004</td>
<td>-0.00341 (0.0133)</td>
<td>-0.0067 (0.0133)</td>
</tr>
<tr>
<td>2005</td>
<td>-0.0459 (0.0137)***</td>
<td>-0.0486 (0.0136)***</td>
</tr>
<tr>
<td>2006</td>
<td>-0.0103 (0.0138)</td>
<td>-0.0117 (0.0138)</td>
</tr>
<tr>
<td>2007</td>
<td>0.0279 (0.0136)**</td>
<td>0.0258 (0.0135)*</td>
</tr>
<tr>
<td>2008</td>
<td>0.00102 (0.0135)</td>
<td>9.08e-05 (0.0135)</td>
</tr>
<tr>
<td>2009</td>
<td>0.00839 (0.0137)</td>
<td>0.00791 (0.0136)</td>
</tr>
<tr>
<td>LLF</td>
<td>1193</td>
<td>1200</td>
</tr>
<tr>
<td>Observations</td>
<td>1410</td>
<td>1410</td>
</tr>
</tbody>
</table>

*, **, *** indicate the coefficient is significantly different from zero at the .1, .05 and .01 levels, respectively.
5.12 Scale economies

The coefficient associated with herd size in the previous regression on technical efficiency, has the potential to mislead. Efficiency scores were estimated under the assumption of variable returns to scale, thus in essence large farms were benchmarked against other large farms, and small farms against other small farms, robbing the coefficient of a simple interpretation.

A better measure of scale efficiency, was previously described as the ratio of technical efficiency scores derived under constant and variable returns to scale. Using this metric, scale efficiency is found to measure .927 and .932, in New York and Ontario, respectively. This result is surprising given the economies of scale found by previous studies of the dairy industry, and the size difference between New York and Ontario farms. However, it is consistent with Haghiri, Nolan, and Tran (2004) who found that size did not significantly impact efficiency in the two regions.

Table 5.10, shows that allocative and scale efficiency is increasing in farm size. This may be due to some form of self-selection (efficient farms are the ones who are able to expand), and is consistent with the literature (for example Tauer and Mishra (2006); Mosheim and Lovell (2009)). Again, the relatively small differences in technical efficiency are expected, as variable returns to scale was imposed.
Table 5.10: Average efficiencies from non-parametric estimation by farm size

<table>
<thead>
<tr>
<th></th>
<th>Ontario</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TE</td>
<td>SE</td>
<td>AE</td>
<td>CE</td>
<td>TE</td>
<td>SE</td>
<td>AE</td>
<td>CE</td>
<td>TE</td>
<td>SE</td>
<td>AE</td>
<td>CE</td>
<td>TE</td>
<td>SE</td>
<td>AE</td>
</tr>
<tr>
<td>&lt;50</td>
<td>0.726</td>
<td>0.898</td>
<td>0.799</td>
<td>0.580</td>
<td>0.773</td>
<td>0.856</td>
<td>0.879</td>
<td>0.678</td>
<td>0.744</td>
<td>0.882</td>
<td>0.830</td>
<td>0.618</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-99</td>
<td>0.704</td>
<td>0.970</td>
<td>0.781</td>
<td>0.550</td>
<td>0.731</td>
<td>0.942</td>
<td>0.889</td>
<td>0.650</td>
<td>0.720</td>
<td>0.954</td>
<td>0.843</td>
<td>0.608</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100-199</td>
<td>0.762</td>
<td>0.978</td>
<td>0.752</td>
<td>0.572</td>
<td>0.757</td>
<td>0.981</td>
<td>0.889</td>
<td>0.672</td>
<td>0.759</td>
<td>0.980</td>
<td>0.843</td>
<td>0.639</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200-499</td>
<td>0.769</td>
<td>0.983</td>
<td>0.875</td>
<td>0.671</td>
<td>0.794</td>
<td>0.990</td>
<td>0.925</td>
<td>0.735</td>
<td>0.785</td>
<td>0.987</td>
<td>0.908</td>
<td>0.713</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.779</td>
<td>0.976</td>
<td>0.949</td>
<td>0.736</td>
<td>0.779</td>
<td>0.976</td>
<td>0.949</td>
<td>0.736</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 6  Parametric Estimation

Parametric estimation of efficiency is traced to seminal papers by Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977). These papers introduced the most common econometric method of efficiency measurement, stochastic frontier analysis (SFA). The typical SFA model differs from least squares estimation, in the addition of a second error term to capture inefficiency. The inefficiency term is assumed to follow some distribution (i.e. a half normal), allowing for the equation to be estimated through maximum likelihood.

The convoluted error term can be appended to the production function, in order to generate technical efficiency scores, or attached to the cost function, to generate cost efficiency scores. Unfortunately, the decomposition of cost efficiency, into technical and allocative components is not as straightforward.

Starting with Schmidt and Lovell (1979), many researchers have attempted to create simultaneous equation models, involving the cost function and a system of input demands. In such models, the cost function has the typical convoluted error term to capture cost inefficiency, while allocative inefficiency is captured by appending an error term to the input demand system. Greene (1980), was the first...
to note that this model relies on the assumption that the error terms of the
equations are independent. Given that allocative inefficiency is a component of cost
efficiency this independence assumption does not hold, resulting in a failure of
identification. In their survey of the literature, Kumbhakar and Lovell (2000) note
that the ”Greene problem” has yet to be remedied.

One way around the Greene problem is the shadow cost approach, recently
employed in the context of dairy farming by Mosheim and Lovell (2009). The
shadow cost approach is not a ”frontier” technique, in the sense that is does not rely
on a convoluted error term. Instead, technical efficiency is measured by the amount
by which a farmer can reduce his inputs while maintaining the same amount of
output (i.e. technical efficiency is equal to φ in \( f(\phi x) = y \)). It is further assumed
that a farmer is allocatively efficient relative to some ”shadow price”. The difference
between the shadow prices and actual prices measures the extent of allocative
inefficiency. This allows for a series of equations, estimated through a generalized
method of moments framework.

One could also decompose cost efficiency by drawing on the dual relation between
the production and cost functions. For example, Bravo-Ureta and Rieger (1991)
employ a stochastic Cobb-Douglas production function to estimate technical
efficiency, and then employ duality to calculate the cost function and resulting cost
efficiency scores analytically. This approach has been criticized by Coelli et al.
(2003), as it employs a production function while treating the input quantities as
decision variables.\(^1\) Furthermore, this specification only allows for a single output.

\(^1\)Marschak and Andrews (1944) explain that this type of formulation is subject to simultaneous equation
In the past fifteen years several studies have emerged examining technical efficiency using the distance function (as opposed to the production function). The Cobb-Douglas distance function has the benefit of allowing for multiple-output technologies while avoiding the criticism of simultaneity (Coelli, 2000). Furthermore, like the production function, the Cobb-Douglas distance function is dual to the cost function, allowing for the analytic decomposition of cost efficiency into technical and allocative components. For these reasons, and for its theoretical relation to DEA, the distance function is employed in this study.

### 6.1 The theory of the distance function

Shephard (1953) introduced the distance function as:

\[
\delta_I(y, x) = \max\{\lambda : x/\lambda \in \Psi\},
\]

where \( \Psi \) is the production possibilities set. Obviously, the reciprocal of \( \delta \) can be thought of as technical efficiency. The logarithmic Cobb-Douglas distance function (in the m-output and k-input case) can be written as:

\[
\ln \delta_{it} = \alpha + \sum_{m=1}^{M} \beta_m \ln y_{mit} + \sum_{k=1}^{K} \gamma_k \ln x_{kit}.
\]

Economic theory suggests the input-distance function is linearly homogenous inputs, requiring \( \sum \gamma_k = 1 \). This restriction can be imposed by normalizing the distance function based on an arbitrarily chosen input (the choice of the input is bias.
inconsequential to the results). A tractable econometric model can be generated by moving this arbitrarily chosen input to the left hand side, moving the distance function to the right hand side and appending an error term

\[-\ln x_{1it} = \alpha + \sum_{m=1}^{M} \beta_m \ln y_{mit} + \sum_{k=1}^{K} \gamma_k \ln \frac{x_{kit}}{x_{1it}} + v_{it} - \ln \delta_{it}. \tag{6.3}\]

Allowing \( \ln \delta_{it} \) to be distributed as a half or truncated normal, and \( v_{it} \) to be distributed as a normal with mean 0, equation 6.3 bears resemblance to the traditional convoluted error structure found in stochastic frontier analysis. During empirical analysis, three different distributional assumptions for \( \ln \delta_{it} \) are tested.

The first specification is a general model, wherein the distance function is distributed as a truncated normal, with the variance of inefficiency being a function of explanatory variables,

\[
\ln \delta_{it} \sim N[\mu_{it}, \sigma_v^2], \quad \sigma_{uit} = \exp(\sigma_u(\eta'Z_{it})). \tag{6.4}
\]

The fit of the general model can be tested against a half-normal distribution by imposing \( \mu = 0 \). Furthermore, the appropriateness of conditioning the variance on the explanatory variables can be tested by removing \( \eta'Z_{it} \).

The duality of the Cobb-Douglas distance and cost functions, and the derivation of cost efficiency scores from this approach, was first exploited by Coelli et al. (2003). The complete derivation of this dual relationship is contained in appendix D.

To remain in keeping with the empirical literature, let \( \ln \delta_{it} = u_{it} \). If \( u_{it} \) is assumed to be a truncated-normal, it takes the following distribution.
\[ f(u) = \frac{1}{\sqrt{2\pi \sigma_v}} \Phi(-\mu/\sigma_v) \exp\left\{ -\frac{(u - \mu)^2}{2\sigma_u^2} \right\}, \] (6.5)

where \( \mu \) is the mode of the distribution, \( \Phi(\cdot) \) is the standard normal cumulative distribution and \( \sigma_v \) is the standard deviation of the distribution. The half-normal distribution can be found by setting \( \mu = 0 \) in the above equation. The composed error term \( \epsilon_{it} = u_{it} + v_{it} \) can be written as

\[ f(\epsilon) = \frac{1}{\sigma} \phi \left( \frac{\epsilon + \mu}{\sigma} \right) \Phi \left( \frac{\mu}{\sigma \lambda} - \frac{\epsilon \lambda}{\sigma} \right) \left[ \Phi \left( \frac{\mu}{\sigma u} \right) \right]^{-1}, \] (6.6)

where \( \sigma = (\sigma_u^2 + \sigma_v^2)^{1/2} \) (\( \sigma_u \) is the standard deviation of the normal distribution), \( \lambda = \sigma_u/\sigma_v \), and \( \phi(\cdot) \) is the density standard normal distribution.

The resulting log-likelihood function can be written as

\[
\ln L(\alpha, \beta, \sigma, \mu) = -\frac{N}{2} \ln 2\pi + \sum_i \left[ -\ln \sigma - \ln \Phi\left( \frac{\mu}{\sigma u} \right) - \frac{1}{2} \left( \frac{\epsilon_i + \mu}{\sigma} \right)^2 + \ln \Phi\left( \frac{\mu}{\sigma u} - \frac{\epsilon_i \lambda}{\sigma} \right) \right]. \quad (6.7)
\]

Following Battese and Coelli (1988), the expected technical efficiency of a specific producer, conditional on the error term is calculated as

\[
E \left[ \exp(-u_{it}) | \epsilon_{it} \right] = \frac{1 - \Phi[\sigma_u - (\tilde{\mu}_{it} + \tilde{\sigma}_{u}^2)]}{1 - \Phi(\tilde{\mu}_{it}/\sigma_u)} \exp\left[ \tilde{\mu}_{it} + \frac{1}{2} \tilde{\sigma}_{u}^2 \right], \quad (6.8)
\]

where \( \tilde{\mu}_{it} = (-\sigma_v^2 \epsilon_{it} + \mu \sigma_{nit}^2)/\sigma^2 \) and \( \sigma_u^2 = \sigma_u^2 \sigma_v^2 / \sigma^2 \). To reiterate, the general model assumes \( \sigma_{nit} = \exp \eta(Z_{it}) \), for equations 6.5 through 6.8. Removing \( \eta'Z_{it} \) collapses the model to a simple truncated normal, while imposing \( \mu = 0 \) results in a half
normal (with inefficiency effects). Imposing both conditions results in a simple half normal model.

6.2 Empirical analysis

The empirical equation (6.3) is estimated using the method of maximum likelihood, with purchased feed as the normalizing input. Yearly dummy variables are used to control for technological change, while the covariates \( Z_{it} \) in the likelihood equation are the same as in the previous chapter.

In addition to testing the assumptions of inefficiency, empirical attention is also directed to the choice of the Cobb-Douglas function form. The Cobb-Douglas allows for an analytical decomposition of cost efficiency, however is more restrictive than other flexible representations of technology, such as the translog function. Given the Cobb-Douglas is nested within the translog system, a likelihood ratio (LR) test can be used to test the equality of the estimators.

As table 6.1 shows, the equality of the two systems is rejected at the 1% level for all distributional assumptions. This being said, the correlation between efficiency estimates from the two models is relatively high for all distributional assumptions.\(^2\)

As Maddala and Fishe (1994) note, the policy relevance of efficiency measurement is not in the level of efficiency per se, but the inter-firm ranking. Given the goal of this paper is to produce valid efficiency scores and assess differences between two regions, the insensitivity of efficiency scores to the choice of functional form, allows

\(^2\)Such a correlation has been found in a myriad of other empirical studies (Samarajeewa et al., 2011; Mbaga et al., 2003) and monte-carlo simulations Larue02
confidence in the choice of the Cobb-Douglas form. Future research could explore
decomposition of cost efficiency with a translog specification (the shadow distance
model of Atkinson and Primont (2002) is one candidate for this decomposition.)

Table 6.1: Likelihood ratio tests comparing the translog and Cobb-Douglas functional
forms across distributional assumptions

<table>
<thead>
<tr>
<th></th>
<th>Log-likelihood Translog (TL)</th>
<th>Log-likelihood Cobb-Douglas (CD)</th>
<th>LR Value</th>
<th>Correlation CD-TL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half-normal</td>
<td>467.82</td>
<td>336.21</td>
<td>263.22</td>
<td>0.91</td>
</tr>
<tr>
<td>Truncated-normal</td>
<td>470.82</td>
<td>337.00</td>
<td>267.64</td>
<td>0.89</td>
</tr>
<tr>
<td>Half-normal (with inefficiency effects)</td>
<td>514.06</td>
<td>376.57</td>
<td>274.98</td>
<td>0.85</td>
</tr>
</tbody>
</table>

The log-likelihood ratio is calculated as $2 \times (LL_{Unrestricted} - LL_{Restricted})$. In this
instance the Cobb-Douglas form is a restricted form of the translog (with 21
restrictions.) The critical chi-square value at the 5% level is 32.67.

Assumptions regarding the distribution of inefficiency can also be tested through a
likelihood ratio. These tests are performed for both the Cobb-Douglas and translog
models, with the results shown in table 6.2. I start by restricting the inefficiency
effects to zero (or removing $\eta_i'Z_{it}$ in equation 6.7), and test a simple half-normal
(found by setting $\mu = 0$ in equation 6.7) against a truncated normal. For both
functional forms we cannot reject the equality of the distributions, allowing us to
adopt the half-normal distribution. I then test the equivalence of the half-normal
without inefficiency effects to the same distribution with these effects unrestricted.
Clearly, the model with inefficiency effects is preferred to the one which omits these
covariates.
Table 6.2: Likelihood ratio tests comparing the general model (truncated normal with inefficiency effects) to other distributional assumptions

<table>
<thead>
<tr>
<th>Test</th>
<th>LR value</th>
<th>Restrictions</th>
<th>Critical value (1% level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD</td>
<td>Half v. truncated</td>
<td>1.58</td>
<td>6.63</td>
</tr>
<tr>
<td></td>
<td>Half v. half (w/ ineff)</td>
<td>80.72</td>
<td>18.48</td>
</tr>
<tr>
<td>TL</td>
<td>Half v. truncated</td>
<td>6.00</td>
<td>6.63</td>
</tr>
<tr>
<td></td>
<td>Half v. half (w/ ineff)</td>
<td>92.48</td>
<td>18.48</td>
</tr>
</tbody>
</table>

Table 6.3 contains the results of maximum likelihood estimation using both the Cobb-Douglas and translog functional forms. As expected, in the Cobb-Douglas specification, outputs have a negative sign (suggesting that they reduce the distance to the frontier), while inputs have a positive sign; all are significant at the 1% level. Consistent with the Malmquist index calculated in the previous chapter, the positive signs on the yearly dummy variables suggest technological regress, as the control group (producers in year 2000) are closer to the frontier than those in later years. In the translog specification, the interaction and square terms preclude an intuitive interpretation of the coefficients, although all the squared terms have the expected sign (with significance) for all outputs and inputs.

A positive sign on the coefficients attached to the covariates, suggests the variance of distribution of inefficiency has increased, increasing inefficiency. In the Cobb-Douglas model, the regional dummy (equal to one if the farm is located in Ontario), age and herd size, are significantly positive. The positive sign on herd size is unexpected, given the previous literature on dairy farm size. In the translog specification, the regional dummy and age are both significantly positive, while per cow expense bST is significantly negative (decreasing inefficiency).
Table 6.3: Results of maximum likelihood estimation using inefficiency effects model

<table>
<thead>
<tr>
<th></th>
<th>Coefficients (Cobb-Douglas)</th>
<th>Coefficients (Translog)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.35 (0.159)***</td>
<td>-7.33 (1.78)</td>
</tr>
<tr>
<td>Labour</td>
<td>0.291 (0.0144)***</td>
<td>0.365 (0.3)</td>
</tr>
<tr>
<td>Capital</td>
<td>0.182 (0.0122)***</td>
<td>-0.135 (0.278)</td>
</tr>
<tr>
<td>Other</td>
<td>0.23 (0.0125)***</td>
<td>0.335 (0.28)</td>
</tr>
<tr>
<td>Milk</td>
<td>-0.788 (0.0133)***</td>
<td>0.292 (0.32)</td>
</tr>
<tr>
<td>Crop</td>
<td>-0.00267 (0.000964)***</td>
<td>-0.0133 (0.0183)</td>
</tr>
<tr>
<td>Livestock</td>
<td>-0.0125 (0.00517)**</td>
<td>-0.104 (0.189)</td>
</tr>
<tr>
<td>Labour$^2$</td>
<td></td>
<td>0.0757 (0.023)***</td>
</tr>
<tr>
<td>Labour*Capital</td>
<td></td>
<td>-0.0731 (0.0294)**</td>
</tr>
<tr>
<td>Labour*Other</td>
<td></td>
<td>0.00144 (0.0282)</td>
</tr>
<tr>
<td>Labour*Milk</td>
<td></td>
<td>-0.0121 (0.0285)</td>
</tr>
<tr>
<td>Labour*Crop</td>
<td></td>
<td>0.000489 (0.00245)</td>
</tr>
<tr>
<td>Labour*Livestock</td>
<td></td>
<td>0.0246 (0.0192)</td>
</tr>
<tr>
<td>Capital$^2$</td>
<td></td>
<td>0.0576 (0.0113)***</td>
</tr>
<tr>
<td>Capital*Other</td>
<td></td>
<td>-0.0742 (0.0248)**</td>
</tr>
<tr>
<td>Capital*Milk</td>
<td></td>
<td>0.00311 (0.0228)</td>
</tr>
<tr>
<td>Capital*Crop</td>
<td></td>
<td>-0.00284 (0.0023)</td>
</tr>
<tr>
<td>Capital*Livestock</td>
<td></td>
<td>-0.00219 (0.013)</td>
</tr>
<tr>
<td>Other$^2$</td>
<td></td>
<td>0.08 (0.0166)***</td>
</tr>
<tr>
<td>Other*Milk</td>
<td></td>
<td>0.0255 (0.0245)</td>
</tr>
<tr>
<td>Other*Crop</td>
<td></td>
<td>-0.000185 (0.0021)</td>
</tr>
<tr>
<td>Other*Livestock</td>
<td></td>
<td>-0.021 (0.0166)</td>
</tr>
<tr>
<td>Milk$^2$</td>
<td></td>
<td>-0.0449 (0.0169)***</td>
</tr>
<tr>
<td>Milk*Crop</td>
<td></td>
<td>0.00165 (0.00191)</td>
</tr>
<tr>
<td>Milk*Livestock</td>
<td></td>
<td>0.00915 (0.0156)</td>
</tr>
<tr>
<td>Crop$^2$</td>
<td></td>
<td>-0.00298 (0.000407)***</td>
</tr>
<tr>
<td>Crop*Livestock</td>
<td></td>
<td>0.00119 (0.00152)</td>
</tr>
<tr>
<td>Livestock$^2$</td>
<td></td>
<td>-0.00266 (0.00102)***</td>
</tr>
<tr>
<td>2001</td>
<td>-0.191 (0.0234)***</td>
<td>-0.157 (0.0224)***</td>
</tr>
<tr>
<td>2002</td>
<td>-0.148 (0.0232)***</td>
<td>-0.123 (0.0222)***</td>
</tr>
<tr>
<td>2003</td>
<td>-0.151 (0.023)***</td>
<td>-0.124 (0.0223)***</td>
</tr>
<tr>
<td>2004</td>
<td>-0.145 (0.023)***</td>
<td>-0.117 (0.0225)***</td>
</tr>
<tr>
<td>2005</td>
<td>-0.114 (0.0234)***</td>
<td>-0.0805 (0.0229)***</td>
</tr>
<tr>
<td>2006</td>
<td>-0.0682 (0.0231)***</td>
<td>-0.0303 (0.0231)</td>
</tr>
<tr>
<td>2007</td>
<td>-0.104 (0.0229)***</td>
<td>-0.0727 (0.0229)***</td>
</tr>
<tr>
<td>2008</td>
<td>-0.141 (0.023)***</td>
<td>-0.0922 (0.0234)***</td>
</tr>
<tr>
<td>2009</td>
<td>-0.113 (0.0227)***</td>
<td>-0.0919 (0.0236)***</td>
</tr>
<tr>
<td>$\sigma_u$</td>
<td>-10.1 (3.23)***</td>
<td>-8.15 (1.24)***</td>
</tr>
</tbody>
</table>
Allocative and cost efficiency scores were calculated by inserting the coefficients from the Cobb-Douglas model into the formula presented in appendix D. The mean results are presented in table 6.4, results by farm size are presented in table 6.5, and density plots are provided in appendix C. The technical efficiency scores presented here are higher than found in the non-parametric estimation, while the allocative efficiency scores are quite similar. Once again, there are relatively small (though statistically significant differences) in technical efficiency, and large differences in allocative efficiency. Unlike the analysis in the previous chapter, there seems to be only minor differences in efficiency between farm sizes. As stated previously, the relationship between farm size and technical efficiency is negative (and significant). The subsequent section will show the relation between farm size and allocative efficiency is also negative (though insignificant). This relationship between size and

<table>
<thead>
<tr>
<th>Region</th>
<th>1.06 (0.419)**</th>
<th>0.762 (0.233)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.0542 (0.0117)**</td>
<td>0.0521 (0.00889)**</td>
</tr>
<tr>
<td>Education</td>
<td>-0.308 (0.21)</td>
<td>-0.285 (0.156)*</td>
</tr>
<tr>
<td>bST</td>
<td>-0.0078 (0.00637)</td>
<td>-0.0136 (0.00629)**</td>
</tr>
<tr>
<td>Herd size</td>
<td>0.00454 (0.0012)**</td>
<td>-0.000873 (0.00172)</td>
</tr>
<tr>
<td>Tie-stall</td>
<td>1.66 (1.35)</td>
<td>1.03 (0.519)**</td>
</tr>
<tr>
<td>Pipeline</td>
<td>1.01 (1.37)</td>
<td>0.697 (0.613)</td>
</tr>
<tr>
<td>Parlour</td>
<td>1.5 (2.03)</td>
<td>1.05 (0.764)</td>
</tr>
<tr>
<td>(\sigma_v)</td>
<td>0.172 (0.00641)</td>
<td>0.145 (0.00577)</td>
</tr>
<tr>
<td>Observations</td>
<td>1410</td>
<td>1410</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>376.57</td>
<td>514.06</td>
</tr>
</tbody>
</table>

| Average efficiency (pooled) | .918 | .900 |
| Average efficiency (New York) | .940 | .921 |
| Average efficiency (Ontario) | .900 | .877 |

Standard errors are in parenthesis

*, **, *** represent significance at the .1, .05 and .01 levels, respectively
efficiency is in contrast with other parametric studies of dairy farm efficiency (Mosheim and Lovell, 2009; Tauer and Mishra, 2006), contrary to expectations and in opposition to the findings of the previous chapter.

Table 6.4: Efficiency scores for the Cobb-Douglas model derived using duality

<table>
<thead>
<tr>
<th></th>
<th>Technical efficiency</th>
<th>Allocative efficiency</th>
<th>Cost efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>0.936 (0.912)</td>
<td>0.933 (0.937)</td>
<td>0.873 (0.854)</td>
</tr>
<tr>
<td>Ontario</td>
<td>0.898 (0.902)***</td>
<td>0.771 (0.762)***</td>
<td>0.691 (0.686)***</td>
</tr>
<tr>
<td>Pooled</td>
<td>0.918 (0.909)</td>
<td>0.857 (0.88)</td>
<td>0.788 (0.799)</td>
</tr>
</tbody>
</table>

(Weighted Average Scores in Parenthesis)

A Mann-Whitney test was used to test the equality of efficiency scores.

*, **, **** represent significant differences between regions at the .1, .05 and .01 levels, respectively
Table 6.5: Average efficiencies from parametric estimation by farm size

<table>
<thead>
<tr>
<th>Herd size</th>
<th>Ontario</th>
<th></th>
<th></th>
<th>New York</th>
<th></th>
<th></th>
<th>Pooled</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TE</td>
<td>AE</td>
<td>CE</td>
<td>TE</td>
<td>AE</td>
<td>CE</td>
<td>TE</td>
<td>AE</td>
</tr>
<tr>
<td>&lt;50</td>
<td>0.895</td>
<td>0.785</td>
<td>0.702</td>
<td>0.939</td>
<td>0.922</td>
<td>0.866</td>
<td>0.912</td>
<td>0.838</td>
</tr>
<tr>
<td>50-99</td>
<td>0.892</td>
<td>0.772</td>
<td>0.688</td>
<td>0.935</td>
<td>0.939</td>
<td>0.878</td>
<td>0.917</td>
<td>0.868</td>
</tr>
<tr>
<td>100-199</td>
<td>0.936</td>
<td>0.709</td>
<td>0.662</td>
<td>0.951</td>
<td>0.931</td>
<td>0.885</td>
<td>0.946</td>
<td>0.857</td>
</tr>
<tr>
<td>200-499</td>
<td>0.895</td>
<td>0.739</td>
<td>0.660</td>
<td>0.933</td>
<td>0.938</td>
<td>0.875</td>
<td>0.920</td>
<td>0.870</td>
</tr>
<tr>
<td>&gt;500</td>
<td>0.833</td>
<td>0.936</td>
<td>0.780</td>
<td>0.833</td>
<td>0.936</td>
<td>0.780</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.3 Second stage analysis

Similar to the analysis in the previous chapter, a truncated regression to determine the effect of likely co-variates on allocative inefficiency scores (the effect of these variables on technical inefficiency was already determined as part of the maximum likelihood regression.) The results, contained in table 6.6, find only the regional dummy to be significant. Again, herd size was found to increase inefficiency, though not significantly.

Table 6.6: Results of truncated regression on allocative efficiency scores

<table>
<thead>
<tr>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Region</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Herd size</td>
</tr>
<tr>
<td>Education</td>
</tr>
<tr>
<td>2001</td>
</tr>
<tr>
<td>2002</td>
</tr>
<tr>
<td>2003</td>
</tr>
<tr>
<td>2004</td>
</tr>
<tr>
<td>2005</td>
</tr>
<tr>
<td>2006</td>
</tr>
<tr>
<td>2007</td>
</tr>
<tr>
<td>2008</td>
</tr>
<tr>
<td>2009</td>
</tr>
</tbody>
</table>

Standard errors are in parenthesis

*, **, **** represent significance at the .1, .05 and .01 levels, respectively

6.4 Correlation across approaches

Table 6.4 contains the correlation and rank correlation between the DEA and econometric models. The correlation between technical efficiency scores is higher
than that found by Mbaga et al. (2003), who employed standard stochastic frontier analysis and data envelopment analysis in their study of Quebec dairy farms. However, it is broadly in keeping with studies from other industries that have triangulated between differing efficiency techniques (Drake and Simper, 2003; Coelli and Perelman, 1999)

Table 6.7: Correlation and rank correlation of preferred DEA and econometric models

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
<th>Rank correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical efficiency</td>
<td>0.478</td>
<td>0.434</td>
</tr>
<tr>
<td>Allocative efficiency</td>
<td>0.502</td>
<td>0.433</td>
</tr>
<tr>
<td>Cost efficiency</td>
<td>0.441</td>
<td>0.437</td>
</tr>
</tbody>
</table>

To my knowledge no previous studies have examined allocative efficiency using multiple techniques. The results show a reasonably high correlation and rank correlation between the methods. Overall the correlation found between these two methods is similar to the correlation between differing DEA models, found in table 5.5.
Chapter 7  Conclusion

The purpose of this paper was to analyze differences in cost efficiency (and its components) between Ontario and New York dairy farms. Two methodologies were used to generate efficiency scores, a linear programming approach and an econometric approach. Both methods revealed the same narrative, and the correlation between the scores of these two methodologies was moderately high. The preferred linear programming approach found New York farmers to be 3.2% more technically efficient and 11.3% more allocatively efficient while the econometric approach found these differences to measure 1% and 17.5%, respectively.\(^1\) Pooling the data, the weighted average efficiency scores from the linear programming and econometric approach measured, 75.3% and 91.8%, respectively for technical efficiency, and 88.5% and 88.0%, respectively, for allocative efficiency. The large difference in technical efficiency scores is not worrying to us; the absolute value of efficiency scores is typically quite sensitive to different estimation techniques, assumptions and methods of input aggregation. That the correlation of scores are reasonably high, and the differences between the regions robust to different methodologies, lends confidence to our general results.

\(^1\) These differences are in terms of weighted averages.
The technical efficiency scores found in this study are significantly higher than those found by Haghiri, Nolan, and Tran (2004). However, the scores from our linear programming model are in keeping with previous studies using similar empirical methods in New York state (i.e. Tauer (1993); Tauer and Stefanides (1998)).

Previous studies of Ontario farms found higher scores for technical efficiency (Weersink, Turvey, and Godah, 1990) and cost efficient (Hailu, Jeffrey, and Unterschultz, 2005). That the efficiency scores in this study are lower, was to be expected; past studies considered only Ontario farms - adding additional (and more efficient farms), will reduce scores. Overall the scores from all models, for all types of efficiency, are broadly in keeping with the previous literature surveyed in Chapter 3.

Secondary analysis in chapter 5, revealed that the cause of allocative efficiency was largely the overcapitalization of dairy farms, and their reliance on home grown feed. The over-capitalization and high debt levels of Canadian farms is not a new finding, (see for example Brinkman (2002)), however this is the first study to show its negative impact on efficiency.

These findings fit our theoretical discussion in Chapter 3. We had suggested that supply management allows farms to maintain (cost) inefficient practices, such as maintaining high capital rates, by restricting competition. Farms may continue these practices either because they derive some utility from high rates of capitalization, or because they are unaware of the inefficiency such capital stocks cause (perhaps because they are missing price information.)

The largest deviation from the past literature is our finding regarding efficiency and scale. Our linear programming approach found scale inefficiency to be rather small
compared to technical inefficiency, although scale was found to significantly increase allocative efficiency. The econometric approach however, found herd size to decrease technical efficiency (after controlling for other variables), and have no significant effect on allocative efficiency. This contrasts with previous literature (such as Tauer and Mishra (2006); Mosheim and Lovell (2009)), who found significant economies of scale in the US dairy industry.

A second surprising result was the amount of technological regress - found through both the Malmquist index and the yearly time dummies in parametric estimation. As mentioned previously, while the finding is robust across our estimators, we can offer no theoretical justification for it.

7.1 Limitations and suggestions for future research

As all theses are, this paper is constrained by data and by time. In terms of data, the most egregious omission is land values. Employing a user cost of capital abstracts from the cost of land, and has the potential to bias results by finding dairies who face high land prices less efficient. Solace can be taken in the fact that all other empirical studies listed in this paper face the same constraints.

Furthermore, the methodology generally proscribed a large reduction in capital, and (for Ontario farms) an increase in purchased feed. This likely entails a steep reduction in land. If cost efficiency requires very little real estate, then the effect of differing real estate prices will be relatively small.

As chapter four makes clear, there are further problems in the comparability of the
data sets. Often implicit prices could not be derived for New York producers, and external prices were employed. This introduces error to the extent that prices are not homogenous across the state. Furthermore, the Ontario dairy set displayed some wildly differing implied prices, suggesting some degree of measurement error. Some comfort can be taken in the fact that our outlier detection method found no anomalies in the data set.\(^2\)

Econometric estimation could be improved by employing a more flexible functional form. Our analysis showed the translog representation of the distance function was preferred to the Cobb-Douglas. The Cobb-Douglas, was however, maintained due to the ease with which the dual cost function can be estimated. Additionally, a Mann-Whitney test was used to determine if the efficiency scores of the two regions differed. A more appropriate metric would incorporate the variance of estimation, just as the procedure of Simar and Zelenyuk (2007) did for non-parametric estimation. Confidence intervals derived by Horrace and Schmidt (1996, 2000), could serve as a starting point in the derivation of such a test.

Future research could also interest itself in a longer term analysis. If data was available reaching back to the early years of supply management, the relative efficiency of Ontario and New York producers between then and now would paint a more vivid portrait of the effects of supply management. Given the concern of this paper with capitalization, the efficiency of investment in Canadian dairy farms would be an interesting subject for future research; balanced panel data could allow

\(^2\)Of course this procedure only looked for farms whose efficiency was significantly higher than average, we made no attempt to find producers whose efficiency was wildly lower than average.
for the estimation of dynamic efficiency (Silva and Stefanou, 2007; Emvalomatis, Stefanou, and Oude Lansink, 2011).

7.2 Policy recommendations

Before offering policy recommendations, it is important to note that policies are seldom based on efficiency alone. Farm policy has typically focused on protecting farm income, and maintaining domestic production. Obviously these policy goals are supported by farmers, and as Ellison, Lusk, and Briggeman (2010) note, taxpayers also tend to support farm subsidies in order to ensure domestic food security. Thus one must allow that efficiency, however much of a golden calf it is to the discipline of economics, is but one of many priorities considered by political actors. This being said, our recommendations consider only how productive efficiency can be improved and leave it for others to decide the trade-offs between efficiency and other objectives.

The boldest policy recommendation that one could make is to dismantle the system of supply management. Barichello, Cranfield, and Meilke (2009), suggest policies that would facilitate the eventual removal of supply management. Less drastic measures could also be taken to expose farmers to a price signal, such as the expansion of quota. We also recommend eliminating the ceiling on quota prices. We make this recommendation not because of economies of scale (which we did not find to be robust across our estimators), but because the quota market allows inefficient firms an incentive to exit the market,
As mentioned previously, the Dairy Farmers of Ontario allows quota to be used as collateral for loans through a process termed "letters of direction". Removing this policy and exposing farmers to higher interest rates may reduce their overcapitalization, thereby increasing their efficiency. It has been our contention that the lack of price information is one reason for productive inefficiency on Ontario farms. In lieu of a competitive price farmers should be given full information as to the extent and determinants of their inefficiency. It is our hope that this study can contribute to this objective.
Bibliography


—. 2010. “Quota Exchange Detailed Bid and Offer – Archive.”


—. 2000. “Multiple Comparisons with the Best, with Economic Applications.”


No. 73. Department of Agricultural and Applied Economics, College of
Agricultural and Life Sciences, University of Wisconsin-Madison.


Nonparametric Frontier Models.” Discussion Paper no. 0317, Institut de
Statistique, Universite Catholique de ouvain, Louvain-la-Neuve, Belguim.

Koopmans, T.C. 1951. *Activity Analysis of Production and Allocation*, New York:
Wiley, chap. An Analysis of Production as an Efficient Combination of Activities.

Milk Quotas on Output Growth: A Modified Distance Function Approach.”

UK: Cambridge University Press.

Lang, B. 2010. “Dairy Farm Wage Rates.” Available at: http:


United States Department of Agriculture. 2010. “Dairy Market News Annual Summaries.” Available at:

http://www.ams.usda.gov/AMSv1.0/ams.fetchTemplateData.do?template=


Chapter A  FMMO Pricing

### Class I prices
\[
P_{BF} = (P_{BF(ClassIII)} + \text{(Class I differential} )/100) \\
P_{SM} = (\text{Higher of } P_{SM(ClassI)} \text{ or } P_{SM(ClassIII)}) + \text{Class I differential} \\
P_{M} = (P_{SM} \times .965) + (P_{BF} \times 3.5)
\]

### Class II prices
\[
P_{BF} = P_{BF(ClassII)} + 0.007 \\
P_{SM} = P_{SM(ClassIV)} + 0.70 \\
P_{OS} = P_{SM} / 9 \\
P_{M} = (P_{SM} \times .965) + (P_{BF} \times 3.5)
\]

### Class III prices
\[
P_{BF} = (P_{BF} - 0.1715) \times 1.211 \\
P_{PR} = (PC - 0.2003) \times 1.383 + [(PC - 0.2003) \times 1.572 - (PBF \times 0.9)] \times 1.17 \\
P_{OS} = (P_{OM} - 0.1991) \times 1.03 \\
P_{SM} = (P_{PR} \times 3.1) + (P_{OS} \times 5.9) \\
P_{M} = (P_{SM} \times .965) + (P_{BF} \times 3.5)
\]

### Class IV Prices
\[
P_{BF} = P_{BF(ClassIII)} \\
P_{OS} = (P_{NFDM} - 0.1678) \times 0.99 \\
P_{SM} = P_{OS} \times 9 \\
P_{M} = (P_{SM} \times .965) + (P_{BF} \times 3.5)
\]

Where subscripts denote:
- BF: Butterfat
- B: Butter
- C: Cheese
- M: Milk
- OS: Other solids
- PR: Protein
- SM: Skim milk
- NFDM: Non-fat dry milk
Chapter B  Map of Northeast Marking Order
Chapter C  Density Plots of Efficiency Scores
Figure C.1: Density plot of technical efficiency scores from DEA model
Figure C.2: Density plot of allocative efficiency scores from DEA model
Figure C.3: Density plot of technical efficiency scores from econometric model
Figure C.4: Density plot of allocative efficiency scores from econometric model
Chapter D  Derivation of Cost Function Using Duality

We offer here the derivation of the dual relation between the input distance and cost function, in the simple two input, one output case. The extension to multiple outputs and inputs is intuitively similar. We began by setting the distance function equation equal to zero - implying technical efficiency

\[ 0 = \alpha + \sum_{m=1}^{M} \beta_m \ln y_{mit} + \gamma_1 \ln x_{1it} + \gamma_2 \ln x_{2it} + u_{it}. \]  
\[ (D.1) \]

Dropping the producer and time subscripts, and bringing \( \ln x_2 \) to the right hand side and taking the derivative with respect to \( \ln x_1 \) we find

\[ \frac{d \ln x_2}{d \ln x_1} = -\frac{\gamma_1}{\gamma_2}. \]  
\[ (D.2) \]

which can be rewritten as:

\[ \frac{dx_2}{dx_1} \frac{x_1}{x_2} = -\frac{\gamma_1}{\gamma_2}. \]  
\[ (D.3) \]

Rearranging and recalling the first order conditions for cost minimization allows,
\[ \frac{dx_2}{dx_1} = -\frac{\gamma_1 x_2}{\gamma_2 x_1} = \frac{w_1}{w_2}, \]  
(D.4)

where \( w \) are input prices. Rearranging once more

\[ w_1 x_1 \gamma_2 - w_2 x_2 \gamma_1 = 0. \]  
(D.5)

We also know that

\[ w_1 x_1 + w_2 x_2 = c. \]  
(D.6)

From equation D.6,

\[ x_1 = -\frac{w_2 p_2}{w_1} + \frac{c}{w_1}. \]  
(D.7)

Recalling that \( \sum \gamma_k = 1 \), substituting equation D.7 into equation D.5 yields

\[ x_2 = \frac{c \gamma_2}{w_2}. \]  
(D.8)

Noting that \( x_1 \) has a symmetric solution the cost function can be written as:

\[ \ln c = -\alpha - (\gamma_1 \ln \gamma_1 + \beta_2 \ln \gamma_2) - \sum_{m=1}^{M} \beta_m \ln y_m + \gamma_1 w_1 + \gamma_2 w_2 - u. \]  
(D.9)

The case of \( k \)-inputs, with producer and time subscripts, is then

\[ \ln c_{it} = -\alpha - \sum_{k=1}^{K} \gamma_k \ln \gamma_k - \sum_{m=1}^{M} \beta_m \ln y_{mit} + \sum_{k=1}^{K} \gamma_k p_{kit} - u_{it}. \]  
(D.10)
After calculating this stochastic cost function, the cost efficient input vector can be computed by Shephard’s lemma

$$x_{jit}^* = \frac{d c_{it}}{d w_{kit}} = \hat{c}_{it} \hat{\gamma_i} / w_{jit} \quad (D.11)$$

Letting $x_{it}$ represent the observed input use, cost efficiency can be calculated as

$$CE = \frac{x_{it}^* p_{it}}{x_{it} p_{it}}, \quad (D.12)$$

while allocative efficiency is calculated as the ratio of cost and technical efficiency.