Economics of Intensive Grazing In Dairy Production In the Mid-Atlantic

Erik Lichtenberg, Kota Minegishi, James Hanson, Dale Johnson

University of Maryland, College Park

Selected Paper prepared for presentation at the Agricultural & Applied Economics Associations 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011

Economics of Intensive Grazing In Dairy Production In the Mid-Atlantic

Erik Lichtenberg, Kota Minegishi, James Hanson, Dale Johnson

Department of Agricultural & Resource Economics,
University of Maryland, College Park

Preliminary. Version: May 3, 2011

Abstract: Dairy production in the US has experienced a marked increase in the size of dairy operations over time. Even as total production has grown over time, smaller operations have been disappearing. Consequently, the viability of smaller dairy farms has become an important policy concern in regions like the Mid-Atlantic where small dairy farms account for a significant share of farm enterprises. Previous studies suggest that dairy farming based on intensive (rotational) grazing, as opposed to traditional confined-feeding operations, may make it possible for smaller operations to remain economically viable. However, the short term nature of the data used in these studies limits the robustness of these findings. We utilize a unique panel of financial data collected during a 15-year-long extension program in Maryland, which allows us to investigate relative performances of intensive grazing operations with more robustness and more in depth than previous studies.

We compare technical efficiency, profitability, and risk in intensive grazing and confinement dairy operations. Using year-by-year data envelopment analysis, we assess returns to scale and compare intensive grazing and confinement systems in terms of technical efficiency. A test of returns to scale suggested by Simar and Wilson (2002) shows that constant returns to scale cannot be rejected, indicating that smaller farms are not necessarily at disadvantage at least in the observed range of operation scale in our data. A comparison of technical efficiency using non-parametric tests and regression indicates that grazers are as technically efficient as conventional confinement operators. A test of second order stochastic dominance suggested by Barrett and Donald (2003) shows that profitability is less risky in intensive grazing than in confinement operations. Overall, these results indicate that intensive (rotational) grazing is a promising approach for improving both the economic and environmental sustainability of dairy farming in regions like the Mid-Atlantic where the climate allows a relatively lengthy grazing season.

1 Introduction

Dairy production in the US has experienced a marked increase in the size of dairy operations over time. Between 2000 and 2009, for instance, the number of dairy operations in the US fell 38%. All of that decline occurred in operations with less than 500 cows; the number of operations with 1000-1999 cows increased by almost a third while the number of operations with 2000 cows or more almost doubled. This trend toward ever larger dairy operations has been generally attributed to economies of scale in the use of buildings, machinery, and other fixed factors of production (Kumbhakar, 1993; Mosheim and Lovell, 2009; Nehring et al., 2009). Some, however, have argued that greater efficiency of variable input use rather than economies of scale underlies this trend (Tauer and Mishra, 2006; Cabrera, Solís, and del Corral, 2010), suggesting that it may be possible to devise farming practices that increase the efficiency of variable input use in smaller operations sufficiently to make them competitive economically.

In the Mid-Atlantic, it has been argued that dairy farming based on intensive (rotational) grazing may make it possible for smaller operations to remain economically viable. In contrast to traditional low-intensity grazing, intensive grazing systems allow a large number of animals to graze on a small plot of land for a short period of time, e.g., 24 hours in length. The animals are then rotated to another pasture and do not return to the first pasture until the grass has fully recovered, maybe a month later, depending on the time of year and specific rotational system in use. The result is a greater reliance on pasture as feed, so that grazers typically raise less field crops than traditional confinement operators.

Studies using data from short term experiments and single-year farm records indicate that, while intensive grazing systems produce less milk per cow, they can be more profitable than confinement systems because of lower operating costs (Elbehri and Ford, 1995; Rust et al., 1995; Dartt et al., 1999; Soriano, Polan, and Miller, 2001; Tucker, Rude, and Wittayakun, 2001; Gloy, Tauer, and Knoblauch, 2002; White et al., 2002; Tozer, Bargo, and Muller, 2003; Fontaneli et al., 2005; Gillespie et al., 2009). Lower operating and capital costs may make them less risky as well as making entry into the industry easier for new farmers.

Additionally, intensive grazing has been shown to be more environmentally sustainable than confinement dairy systems; environmental benefits of intensive grazing include reduced sediment erosion (Digiacomo et al., 2001), lower phosphorus runoff (Bishop et al., 2005), and increased carbon sequestration (Guo and Gifford, 2002).

All of these previous studies have used either single-year cross-section data from actual farm enterprises or experimental data collected over a relatively short time period, limiting the inferences that can be drawn from them, especially in relation to long term economic sustainability both on average and in terms of variability of income. In contrast to these previous studies, we use a unique panel data set to compare technical efficiency, profitability, and risk in intensive grazing and confinement dairy operations in the Mid-Atlantic. The data come from IRS Form 1040 Schedule F (farm income) tax returns of dairy farmers from 1995-2009 collected as part of University of Maryland farm management extension program. Some of the 63 farmers participating in this program have used an intensive grazing system for all or part of the 15-year period covered by the study. Conclusions about the relative efficiency and economic sustainability of dairy operations obtained by analyzing data from 63 farms in Central Maryland over a 15 year period should be more robust than those currently in the literature. A previous analysis of these data shows that confinement operators produce more milk, both in total and per cow, and bring in greater revenue. At the same time, they spend more on production so that there is no statistically significant difference in profit between the two types of operations. Additionally, there are indications that cows are healthier under an intensive grazing regimen (Hanson et al., 2011).

This paper conducts a more intensive analysis of the relative efficiency of intensive grazing and confinement dairy operations. We use year-by-year data envelopment analysis to assess returns to scale and compare intensive grazing and confinement systems in terms of technical efficiency. Our data are consistent with constant returns to scale in both types of operation, indicating that smaller farms are not necessarily at disadvantage-at least in the range of operational scale observed in our data.

We compare technical efficiency non-parametrically and by regressing the technical efficiency score on an indicator for a confinement operation along with some demographic variables. We find no difference in technical efficiency between grazers and confinement operators. The results also indicate that farmers with off-farm income tend to be more inefficient while farmers with more experience and farmers with children tend to be more efficient. Surprisingly, owner-operators tend to be less efficient. Finally, a test of second order stochastic dominance (SSD) suggested by Barrett and Donald (2003) shows that profit is less risky in grazing operations than in confinement operations. Overall, these results indicate that intensive (rotational) grazing is a promising approach for improving both the economic and environmental sustainability of dairy farming in regions like the Mid-Atlantic where the climate allows a relatively lengthy grazing season.

The rest of paper proceeds as follows. We present research methodologies in section 2, describe our dataset in section 3, discuss preliminary results in section 4, and conclude the study.

2 Methods

2.1 DEA Analysis

We analyze efficiencies calculated by output-oriented data envelopment analysis (DEA). Our definition of technical inefficiency is the maximum radial expansion of observed outputs given inputs within a feasible production set. The boundary of this producible set is a piece-wise linear production possibility curve, on which maximum technical efficiency is defined. Calculated inefficiencies are used to examine returns to scale structure in dairy production and tests of relative inefficiency between grazers and confinement operators.

For a given convex set of technically feasible set $T = \{(\boldsymbol{x}, \boldsymbol{y}) : \text{inputs } \boldsymbol{x} \text{ can produce outputs } \boldsymbol{y}\}$, Farrell efficiency measurement is defined as a radial maximal radial expansion in outputs $\delta(\boldsymbol{x}, \boldsymbol{y};\, T) = \sup\{\delta : (\boldsymbol{x}, \delta\boldsymbol{y}) \in T\} \in [1, \infty)$. When an input-output combination is efficient, inefficiency δ takes a value of one.

We use year-by-year DEA to assess returns to account non-parametrically for yearly fluctuations of weather and market conditions as well as technological trends over time. That is, observations in a given year are assessed against a reference technology defined by the observations in the same year. Formally, under variable returns to scale (VRS) and strong disposability, inefficiency δ_{is} for farm i in year s is calculated by the following linear programming;

$$\delta(x_{is}, y_{is};\, T_s^{VRS}) = \text{argmax}\{\, \delta \mid\ \boldsymbol{y}_s\boldsymbol{\lambda} \geq \delta\boldsymbol{y}_{is},\ \ \boldsymbol{x}_s\boldsymbol{\lambda} \leq \boldsymbol{x}_{is},\ \ \mathbf{1}\lambda = 1,\ \ \boldsymbol{\lambda} \in \mathbb{R}_+^{N_s}\} \tag{1}$$

where $\boldsymbol{\lambda}$ is a matrix of optimal weights to construct a technological frontier in feasible set T_s^{VRS}. In the case of constant returns to scale (CRS) or non-increasing returns to scale (NIRS), the constraint $\mathbf{1}\boldsymbol{\lambda} = 1$ is removed or replaced with $\mathbf{1}\boldsymbol{\lambda} \leq 1$ respectively.

Our primary specification uses three outputs and six inputs. The three outputs are sales revenues in milk, crops, and cattle. Crops and cattle are measured only in terms of sales in our data. We measure milk in terms of sales as well in order to capture differences in quality, e.g., butterfat content; measuring milk production in quantity terms yields the same qualitative results. We use six aggregated inputs: the number of cows in the herd; paid labor; the acreage of land; the sum of depreciation, maintenance, rent, custom hire, fuel, utilities (as a measure of machinery usage); the sum of feed and veterinary expenses (as a measure of livestock-related costs); and the sum of expenditures on fertilizers, chemicals, and seed (as a measure of crop-related costs). In addressing sensitivity of the results, we also consider alternative definitions for inputs without using the acreage of land, which increases our sample size.[1]

We convert dollar-denominated inputs and outputs into quasi-quantity measurements using price indices that are normalized at year 2009 values.[2] For each aggregate input, individual-specific price indices are constructed by summing component-wise indices that are multiplied by their shares in the aggregate expenses. e.g. $p_i^{crop} = (seed_i/crop_i)\, p^{seed} + (fertilizer_i/crop_i)\, p^{fertilizer} + (chem_i/crop_i)p^{chem}$ where $crop_i = seed_i + fertilizer_i + chem_i$ is the aggregate crop-related expenses for observation i. We use the following price indices that are publicly available from National Agricultural Statistical Service at USDA: diary products (for milk sales), feed grains & hay (crop sales), meat animals (cattle sales), farm services (custom hire and veterinary expenses), farm machinery (depreciation), farm supplies & repairs (maintenance), fuels (fuels and utilities), rent (rent), seed (seed), fertilizers (fertilizers), agricultural chemicals (chemicals), feed (feed), and wage rates (hired labor).

2.1.1 Returns to Scale

We use tests developed by Simar and Wilson (2002) to investigate the returns to scale. These tests examine whether the DEA assumption on RTS makes significant difference in calculating inefficiencies. Under the null hypothesis of constant returns to scale (CRS) against the alternative of variable returns to scale (VRS), test statistics are given by (a) ratio of means $\hat{S}^{CRS} = \sum_i^N \delta_i(.;\ T^{VRS}) / \sum_i^N \delta_i(.;\ T^{CRS})$ and (b) ratio of 10% trimmed means $\hat{S}^{CRS} = \sum_i^N \delta_i(.;\ T^{VRS}) \mathbb{1}\{i \in D(5,95;VRS)\} / \sum_i^N \delta_i(.;\ T^{CRS}) \mathbb{1}\{i \in D(5,95;CRS)\}$ where $D(a,b;RTS)$ is a subset of $i \in \{1,2,...,N\}$ in which $\delta_i(.;\ T^{RST})$ is in between a-th and b-percentile values. The spirit of the test is to compare whether the test statistic significantly differs from one.

Since DEA is deterministic as opposed to being stochastic, statistical inferences are made through a specific form of bootstrapping. The basic idea is to use a bootstrapping process that imitates the data generating process so that the comparison between the test statistic and its true

[1] Some previous studies on dairy production also have examined production relationships without land size in inputs. See for example, Kumbhakar, Ghosh, and McGuckin (1991); Ahmad and Bravo-Ureta (1995); Mayen, Balagtas, and Alexander (2010).

[2] For example, we normalize price indices for dairy product for 1995-2009 by dividing them by the index value in 2009. Then, we divide observed milk revenues by the normalized indices of corresponding years.

value is approximated by the comparison between the bootstrap estimates and the test statistic. In this case, critical values are obtained from a homogeneous bootstrapping procedure described in Simar and Wilson (1998). That considers perturbations of technological frontiers under the null hypothesis (e.g. CRS) with a random rescaling factor drawn from the smoothed estimate of distribution for $\delta_i(.;\, T^{CRS})$. The process is designed to obtain smaller feasible sets (for reference technologies in DEA) under the bootstrapping process than the observed feasible set, by which the procedure mimics the fact that the observed feasible set under finite samples is inevitably a subset of underlying truly feasible set. The same procedure is applicable to the null hypothesis of non-increasing returns to scale (NIRS) against VRS, which has a nested property, $\hat{S}_i^{NIRS} \geq \hat{S}_i^{CRS}$, due to a stronger restriction on data envelopment (i.e. a smaller feasible set) under NIRS than CRS as noted above.

Theoretically, CRS should be easier to reject than NIRS as the test statistic under CRS is farther away from 1 than that of NIRS. However, critical values differ under each null hypothesis since the pseudo-frontiers from bootstrapping are obtained from random draws under each RTS assumption of null hypothesis. The researcher may find a higher p-value under CRS than NIRS as such pseudo-frontiers under CRS are more wildly perturbed away from the observed frontier than those under NIRS are, which results in less precise bootstrap counterparts for the test statistic. In practice, when the tests reject NIRS but not CRS under the alternative hypothesis of VRS, one can further test the null hypothesis of CRS against the alternative hypothesis of NIRS.

Simar and Wilson (2002) discuss test statistics using the mean ratios of efficiency scores, the mean efficiency scores, the median ratio of efficiency scores, the median of the ratios of efficiency scores, the mean ratio of efficiency scores of the sample with the top and bottom 5% trimmed, and the ratio of the 10% trimmed mean efficiency scores; the latter four serve as adjustments for skewness in the distribution of efficiency scores. We report the ratios of the mean and 10% trimmed mean efficiency scores. The remaining test statistics gave the same qualitative results.

2.1.2 Determinants of Inefficiencies

We compare relative technical efficiency in two ways. First, we examine whether the distributions of δ's systematically differ between grazers and confinement farms from the Wilcoxon-Mann-Whitney test (Wilcoxon, 1945; Mann and Whitney, 1947), a commonly used non-parametric test for differences in income distributions of two groups. Under the null hypothesis of commonly shared underlying distribution, sum of ranks for each group should be close to a theoretical value solely based on relative sample sizes.

We also compare efficiency in a regression framework that controls for a variety of demongraphic factors as well as operation type that may affect operational efficiency. Specifically, we estimate the following linear specification;

$$\delta_{it} = \boldsymbol{Z}_{it}\boldsymbol{\beta} + \varepsilon_{it} \tag{2}$$

where covariates $\boldsymbol{Z}_{it}$ contain an indicator variable that takes a value of 1 for confinement operation as well as observed operator characteristics. A correct specification of error structure ε_{it} in such a two-stage approach is still debated in the literature. While some advocate truncated regressions (Simar and Wilson, 2007), we use ordinary least squares (OLS) regressions, which consistently estimate $\boldsymbol{\beta}$ (Hoff, 2007; McDonald, 2009). Simar and Wilson (2007) also suggest correcting finite-sampling bias in DEA estimates of inefficiency. The DEA estimators are downwardly biased since an observed feasible set of input-output pairs is smaller than underlying truly feasible set. The second-stage regression serves as an integral part of data generating process that replaces homogeneous bootstrapping discussed earlier; predicted systematic inefficiencies $\boldsymbol{Z}_{it}\hat{\boldsymbol{\beta}}$ and random draw from error term ε_{it} in (2) replace a random draw of δ_{it} in constructing pseudo-frontiers and correcting for finite-sampling bias in the sense of Simar and Wilson (1998). Coefficients $\boldsymbol{\beta}$ can be estimated consistently with or without correcting such finite-sampling bias in inefficiency estimates.

2.2 Stochastic Dominance

We investigate relative riskiness of intensive grazing by testing for second-order stochastic dominance (SSD) in net profit, adjusted for inflation to 2009 dollars using the Consumer Price Index. Formally, the null hypothesis of SSD of the profit distribution for grazers $F_G(:)$ over that for confinement $F_C(:)$ is defined in the sense of (McFadden, 1989);

$$\begin{aligned} H_0: &\quad \int_{\underline{z}}^{z} F_C(t) - F_G(t)\,\mathrm{d}t \geq 0 \text{ for all } z \in [\underline{z}, \bar{z}] \\ H_1: &\quad \int_{\underline{z}}^{z} F_C(t) - F_G(t)\,\mathrm{d}t < 0 \text{ for some } z \in [\underline{z}, \bar{z}] \end{aligned} \tag{3}$$

where $H1$ is the alternative hypothesis that the two distributions violate such a SSD relationship.

We apply one of the most common SSD tests in the literature with a bootstrap-based critical values proposed by Barrett and Donald (2003). The test statistic $\hat{S}$ is the supremum of the cumulative difference in $c.d.f.$'s over the whole support of variable z;

$$\hat{S} = \sqrt{\frac{N_G N_C}{N_G + N_C}} \sup_z \left\{ \int_{\underline{z}}^{z} F_C(t) - F_G(t) \, \mathrm{d}t \right\} \tag{4}$$

where N_G, N_C are number of observations for grazers and confinement, $F_J(.;\, X) = N_J^{-1} \sum_{i=1}^{N_J} (z - X_i)\, \mathbb{1}\{X_i \geq z\}$ for $J \in \{G, C\}$ by Davidson and Duclos (2000), and the supremum is searched over sufficiently fine grids. For each bootstrap b, we calculate

$$S^b = \sqrt{N} \sup_z \left\{ \int_{\underline{z}}^{z} F_G^b(t) - F_G(t) \, \mathrm{d}t \right\} \tag{5}$$

where $F_G^b(.;\, X^b)$ is a bootstrapping counterpart for $F_G(.;\, X)$ under random sample X^b with replacement. Given enough bootstrapping estimates, we calculate the theoretical upper bound for rejection probability, or the p-value for the null hypothesis, $\hat{p} = Pr(S^b > \hat{S})$.

3 Data

We use data derived from IRS Form 1040 Schedule F (farm income) tax returns of 63 dairy farmers who have been participating in a University of Maryland Extension farm management program evaluating their financial performance on a yearly basis and discussing opportunities for improvements in each farmer's operation. Twenty of these farmers have used an intensive grazing system for all or part of the 15-year period covered by the study while the remainder use traditional confinement systems. The dataset contains the total of 580 observations on herd size, milk output, crop sales, cow sales, expenditures on numerous inputs, and profit for the period 1995-2009 as well as some demographic information (age, education, family size and composition) and risk tolerance. These are all small operations by national standards. All but three have 200 cows or less in every year of the sample; two have more than 200 cows for some, but not all years in the sample; and none have more than 500 cows in any year of the sample.

Table 1 summarizes the aggregate inputs and outputs described in the previous section and some operator characteristics. The inputs and outputs are in quasi-quantity terms. Readers interested in nominal values and the details of data descriptions are referred to Hanson et al. (2011). Information on land acreage is not available for complete series of years for some farms, which reduces sample size to 474 in our preferred specification in DEA. Operator characteristics include information on operator age, number of children, and indicators for risk-averse, owner-operator of land, and presence of off-farm source of income. The indicator of whether the operator is risk averse is derived from a Holt-Laurie experiment conducted in 2009. Almost all operators chose one of two options, indicating either risk neutrality or risk aversion at the lowest-level in our experimental design. We thus treated responses as dichotomous.

We define as intensive grazing operations those whose animals get at least 30% dry matter intake for 4 months or more. This definition is the basis for organic certification. For some farmers, intensive grazing can be a stepping stone in the transition to organic production, as it was for three of the operations in our sample.

4 Results

The DEA calculations for inefficiencies are summarized in table 2 and also graphed in figure 1 by farm type. In the upper panel, the table contains specifications with full inputs and inputs without land acreage under variable (VRS), non-increasing (NIRS), and constant (CRS) returns to scale, with and without outliers excluded.[3] Many observations are found to be efficient: the median score of 1 in all of our specifications suggest that at least half of the observations are fully technically efficient. In the preferred specifications in rows (1)-(3), the inefficiency scores indicate that the farm at 75th percentile (in inefficiency distribution) could have produced 4.8%, 9.0%, and 11.5% more outputs under the fully efficient operation at VRS, NIRS, and CRS frontiers respectively.

The graphs in figure 1 show little difference in the distributions of efficiency scores under alternative assumptions regarding returns to scale. Formal tests (reported in table 3) confirm that imposing more restrictive returns to scale assumptions makes statistically discernible difference in the distribution of efficiency scores. Specifically, we are unable to reject a null hypothesis of constant returns to scale; we therefore impose this assumption in the remainder of the analysis.

Investigations of returns to scale in US dairy farming have produced mixed results. Some studies using national data from the Agricultural Resource Management Survey (ARMS) and its predecessor, the Farm Costs and Returns Survey, have tended to find evidence of increasing returns to scale (Kumbhakar, Ghosh, and McGuckin, 1991; Mosheim and Lovell, 2009). A recent analysis using those same data, however, found increasing returns to scale for confinement operations but found also that intensive grazing operations were cost-competitive with the largest confinement operations (Nehring et al., 2009). Kumbhakar (1993) similar found evidence of returns to scale

[3]We follow the procedures proposed by Wilson (1993) to determine outliers

in confinement operations in Utah; his study did not examine intensive grazing operations. Tauer (2001) finds very small economies of scale in New York dairies. Tauer and Mishra (2006) find that differences in efficiency rather than scale economies are primarily responsible for differences in US dairy production costs, while Cabrera, Solís, and del Corral (2010) find no evidence of economies of scale in Wisconsin dairy operations with no more than 1000 cows.

The farms in our data set are all small: The largest have only a few more than 200 cows. We find no evidence of increasing returns to scale within that range; to the contrary, we find evidence of constant returns to scale, indicating no cost advantage from size. Analysis of net profits reported to the Internal Revenue Service by these farmers indicates that these dairies are consistently profitable, with intensive grazers being no less profitable overall and more profitable per cow than confinement operators (Hanson et al., 2011), a finding consistent with that of Nehring et al. (2009) at the national level. Thus, our results suggest that intensive grazing is likely to be an economically sustainable mode of operation for dairies in areas like the Mid-Atlantic, where feeding on pasture is viable for a sufficiently long period of time (up to 8 months for dairies in our sample).

We compare the relative efficiency of intensive grazing and confinement operators using a Mann-Whitney-Wilcoxon test (reported in table 2) and OLS regressions (reported in table 4). The Mann-Whitney-Wilcoxon test shows no statistically significant difference in efficiency between intensive grazing and confinement operators: the null hypothesis of equality in distributions cannot be rejected under any specification.

We also use OLS regression to analyze the determinants of inefficiencies. The regression models include farm type and other operator characteristics in addition to an indicator for confinement operations. While we maintain the assumption of CRS and use specification (3) in table 2, we obtain qualitatively similar results under alternative returns to scale specifications. Columns (3) and (4) in table 2 show the results with using standard errors corrected for finite-sampling bias in the spirit of Simar and Wilson (2007) while columns (1) and (2) are those without such corrections. The results are similar with or without such corrections. The variances of coefficients are obtained through bootstrapping from approximated normal distribution of residuals.

The coefficient of the confinement indicator is negative and significantly different from zero when that indicator for confinement operations is the only variable included in the model (along a constant term) and standard errors are corrected for finite sampling bias, a result that suggests that confinement operators are more efficient on average than grazers. Once other demographic variables are included, however, the sign of the confinement indicator coefficient becomes positive and significantly different from zero at the 10% level, suggesting that grazers are at least as efficient-and possibly more efficient-on average as confinement operators.

Most of the coefficients of the demographic variables have the expected signs. Older farmers are more efficient on average, indicating the effect of experience on technical efficiency. Farmers with off-farm income tend to be less efficient, consistent with the notion that off-farm work reduces management time and attention. Farmers with young children tend to be more efficient, perhaps because of greater demands on household income or for bequest reasons. Surprisingly, owner-operators tend to be less efficient. Finally, risk averse farmers tend to be no more or less efficient than risk neutral farmers, i.e., risk attitudes have no apparent effect on technical efficiency-as economic theory suggests should be the case.

The cumulative distributions of real net profits reported by grazers and confinement operators, adjusting for inflation to 2009 price levels using the Consumer Price Index, are shown in figure 2 along with the cumulative difference in the areas under those two cumulative distributions (which defines second order stochastic dominance). The figure indicates that the distribution of intensive grazers' profits exhibits SSD over the distribution of confinement operators' profits, a result confirmed by formal statistical testing (table 5). Thus, intensive grazing is less risky than confinement operation. This result is intuitive: Intensive grazing involves lower expenditures on virtually all categories of inputs, making it less vulnerable to variability in net returns. Additionally, intensive grazing appears to be better for animal health (as measured, e.g., by veterinary expenditures and sales of cattle) and thus less subject to variability on that score (Hanson et al., 2011).

Overall, then, we find that intensive grazing is as technically efficient as traditional confinement operations and is less risky as well. Combined with a previous finding of no difference in net profit, these results suggest that intensive grazing represents a strategy for keeping small dairy operations economically competitive, at least in areas like the Mid-Atlantic where feeding on pasture for 4 months or more is feasible.

That result is especially important in light of recent developments in environmental regulation. Failure to meet water quality standards for nitrogen and phosphorus in the Chesapeake Bay has led the Environmental Protection Agency to initiate a process of developing regulations based on total maximum daily loadings (TMDLs) that include all sources-including nonpoint agricultural sources that have heretofore been exempt. Livestock-notably dairy cattle-are responsible for a large share of nitrogen loadings into the upper Bay, so TMDL regulations will likely impose substantial effluent control requirements on dairy operations in the Chesapeake Bay watershed, which comprises much of the Mid-Atlantic region. As noted earlier, intensive grazing operations create substantially less runoff than confined feeding operations, so TMDL regulations should have much less of an effect on them than on traditional confinement operations. Thus, intensive grazing is likely to keep the dairy industry economically sustainable even under more restrictive environmental regulations.

5 Conclusions

Using financial and demographic data collected from 63 farms over 15-year period, we have investigated the relative economic viability of intensive grazing in terms of returns to scale, technical efficiency, and riskiness in net profits. First, our finding of constant returns to scale (CRS) in production indicates that smaller farms need not be less competitive in the Mid-Atlantic at least in the range of observed operational scale in our dataset. Second, from the Mann-Whitney test and two-stage DEA analysis, we find that grazers are at least as technically efficient as confinement operators. Third, formal tests of SSD demonstrate that intensive grazing is a less risky mode of production than confinement operations.

These findings contribute to the discussions on the potential for intensive grazing as a means of increasing the economic viability of smaller dairy farms in the Mid-Atlantic. The nationwide trend toward ever larger dairy operations has created concern over the long run viability of smaller dairy farms in a region like the Mid-Atlantic with high prevalence of small dairy farms that have been on the decline. This trend has been attributed to economies of scale and to shifts in the location of dairy operations to non-traditional areas with lower costs of land, feed, and labor. However, our results along with those of Cabrera, Solís, and del Corral (2010)[4] suggest that dairy farms can be competitive at smaller scales in the Mid-Atlantic, especially when environmental costs like water pollution from nutrient runoff are taken into account. Additionally, small scale farming is often believed to provide a foundation for a higher quality of life in rural communities.

Intensive grazing offers more reliable source of income, in particular for smaller farms that are sensitive to downside risk. Consistent with other studies, we find that it utilizes inputs at least as efficiently as conventional confinement dairy operations. These findings on economic viability complement other benefits of intensive grazing found in previous studies, including lower barriers to entry due to lower capital requirements, better herd health indicated by substantially smaller veterinary expenses, and lower negative environmental externalities from nutrient and chemical

[4]Cabrera, Solís, and del Corral (2010) find no evidence of scale economies in Wisconsin dairy operations with no more than 1000 cows.

runoff. Those advantages are likely to grow as the Environmental Protection Agency moves toward more comprehensive water quality regulations in areas like the Chesapeake Bay watershed. With its smaller environmental footprint, intensive grazing may provide the means for the dairy industry to survive in the face of stricter manure management regulations.

References

Ahmad, M., and B.E. Bravo-Ureta. 1995. "An Econometric Decomposition of Dairy Output Growth." *American Journal of Agricultural Economics* 77:914–921.

Barrett, G.F., and S.G. Donald. 2003. "Consistent Tests for Stochastic Dominance." *Econometrica* 71:71–104.

Bishop, P.L., W.D. Hively, J.R. Stedinger, M.R. Rafferty, J.L. Lojpersberger, and J.A. Bloomfield. 2005. "Multivariate Analysis of Paired Watershed Data to Evaluate Agricultural Best Management Practice Effects on Stream Water Phosphorus." *J Environ Qual* 34:1087–1101.

Cabrera, V., D. Solís, and J. del Corral. 2010. "Determinants of technical efficiency among dairy farms in Wisconsin." *Journal of Dairy Science* 93:387–393.

Dartt, B.A., J.W. Lloyd, B.R. Radke, J.R. Black, and J.B. Kaneene. 1999. "A Comparison of Profitability and Economic Efficiencies Between Management-Intensive Grazing and Conventionally Managed Dairies in Michigan." *J. Dairy Sci.* 82:2412–2420.

Davidson, R., and J. Duclos. 2000. "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality." *Econometrica* 68:1435–1464.

Digiacomo, G., C. Iremonger, L. Kemp, C. van Schaik, and H. Murray. 2001. "Sustainable Farming Systems: Demonstrating Environmental and Economic Performances."

Elbehri, A., and A. Ford. 1995. "Economic analysis of major dairy forage systems in Pennsylvania : the role of intensive grazing." *Journal of Production Agriculture* 4:501–507.

Fontaneli, R.S., L.E. Sollenberger, R.C. Littell, and C.R. Staples. 2005. "Performance of lactating dairy cows managed on pasture-based or in freestall barn-feeding systems." *Journal of Dairy Science* 88:1264–1276.

Gillespie, J.M., R.F. Nehring, C.B. Hallahan, C.L. Sandretto, and R.F.N.J.M. Gillespie. 2009. "Pasture-Based Dairy Systems: Who Are the Producers and Are Their Operations More Profitable than Conventional Dairies?" http://econpapers.repec.org/article/agsjlaare/57630.htm.

Gloy, B.A., L.W. Tauer, and W. Knoblauch. 2002. "Profitability of Grazing Versus Mechanical Forage Harvesting on New York Dairy Farms." *J. Dairy Sci.* 85:2215–2222.

Guo, L.B., and R.M. Gifford. 2002. "Soil carbon stocks and land use change: a meta analysis." *Global Change Biology* 8:345–360.

Hanson, J., E. Lichtenberg, K. Minegishi, and D. Johnson. 2011. "Competitiveness of Management Intensive Rotational Grazing in the Mid-Atlantic." Working paper, University of Maryland, College Park, May.

Hoff, A. 2007. "Second stage DEA: Comparison of approaches for modelling the DEA score." *European Journal of Operational Research* 181:425–435.

Kumbhakar, S.C. 1993. "Short-Run Returns to Scale, Farm-Size, and Economic Efficiency." *The Review of Economics and Statistics* 75:336–341.

Kumbhakar, S.C., S. Ghosh, and J.T. McGuckin. 1991. "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms." *Journal of Business & Economic Statistics* 9:279–286.

Mann, H.B., and D.R. Whitney. 1947. "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other." *The Annals of Mathematical Statistics* 18:50–60.

Mayen, C.D., J.V. Balagtas, and C.E. Alexander. 2010. "Technology Adoption and Technical Efficiency: Organic and Conventional Dairy Farms in the United States." *American Journal of Agricultural Economics* 92:181–195.

McDonald, J. 2009. "Using least squares and tobit in second stage DEA efficiency analyses." *European Journal of Operational Research* 197:792–798.

McFadden, D. 1989. *Testing for Stochastic Dominance*. Studies in the Economics of Uncertainty: In Honor of Josef Hadar, ed. by T B. Fomby and T K. Seo., New York: Springer.

Mosheim, R., and C.K. Lovell. 2009. "Scale Economies and Inefficiency of U.S. Dairy Farms." *American Journal of Agricultural Economics* 91:777–794.

Nehring, R., J. Gillespie, C. Sandretto, and C. Hallahan. 2009. "Small U.S. dairy farms: can they compete?" *Agricultural Economics* 40:817–825.

Rust, J., C. Sheaffer, V. Eidman, R. Moon, and R. Mathison. 1995. "Intensive Rotational Grazing for Dairy Cattle Feeding." *American Journal of Alternative Agriculture* 10:147–151.

Simar, L., and P.W. Wilson. 2007. "Estimation and inference in two-stage, semi-parametric models of production processes." *Journal of Econometrics* 136:31–64.

—. 2002. "Non-parametric tests of returns to scale." *European Journal of Operational Research* 139:115–132.

—. 1998. "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models." *MANAGEMENT SCIENCE* 44:49–61.

Soriano, F.D., C.E. Polan, and C.N. Miller. 2001. "Supplementing Pasture to Lactating Holsteins Fed a Total Mixed Ration Diet." *J. Dairy Sci.* 84:2460–2468.

Tauer, L. 2001. "Efficiency and Competitiveness of the Small New York Dairy Farm." *Journal of Dairy Science* 84:2573–2576.

Tauer, L.W., and A.K. Mishra. 2006. "Dairy Farm Cost Efficiency." *J. Dairy Sci.* 89:4937–4943.

Tozer, P.R., F. Bargo, and L.D. Muller. 2003. "Economic Analyses of Feeding Systems Combining Pasture and Total Mixed Ration." *J. Dairy Sci.* 86:808–818.

Tucker, W.B., B.J. Rude, and S. Wittayakun. 2001. "Case Study: Performance and Economics of Dairy Cows Fed a Corn Silage-Based Total Mixed Ration or Grazing Annual Ryegrass During Mid to Late Lactation." *The Professional Animal Scientist* 17:195–201.

White, S.L., G.A. Benson, S.P. Washburn, and J.T. Green. 2002. "Milk Production and Economic Measures in Confinement or Pasture Systems Using Seasonally Calved Holstein and Jersey Cows." *J. Dairy Sci.* 85:95–104.

Wilcoxon, F. 1945. "Individual Comparisons by Ranking Methods." *Biometrics Bulletin* 1:80–83.

Wilson, P.W. 1993. "Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs." *Journal of Business & Economic Statistics* 11:319–323.

Table 1: Summary Statistics of Aggregate Inputs, Outputs and Operator Characteristics

Variable	Obs	Mean	Std. Dev.	Min	Max
Whole Sample (165 Grazing & 415 Confinement Obs.)					
sales milk	580	259,218	204,215	41,279	1,602,186
sales crop	580	4,111	10,005	0	86,359
sales cattle	580	19,698	16,906	0	128,991
cow	580	108	63	22	468
labor	580	25,369	41,484	0	281,259
crop	580	38,547	36,389	0	230,954
livestock	580	142,706	113,839	4,568	746,873
machinery	580	125,629	86,606	22,063	806,265
acreage	474	321	152	90	845
$\mathbb{1}\{\text{risk averse}\}$	343	0.83	0.37	0	1
age	511	43.7	12.6	13	81
number of children	243	3.1	1.4	1	6
$\mathbb{1}\{\text{farm own}\}$	573	0.70	0.46	0	1
$\mathbb{1}\{\text{off-farm income}\}$	580	0.09	0.29	0	1
Grazers with Acreage Info.					
sales milk	161	168,943	72,916	41,279	586,887
sales crop	161	950	3,059	0	21,990
sales cattle	161	17,697	12,685	0	69,891
cow	161	87	29	37	195
labor	161	6,229	10,383	0	75,322
crop	161	18,267	17,924	0	107,151
livestock	161	94,649	47,848	7,882	255,845
machinery	161	84,517	45,558	26,717	327,004
acreage	161	285	131	115	700
$\mathbb{1}\{\text{risk averse}\}$	110	0.75	0.43	0	1
age	161	39.1	8.8	20	59
number of children	75	2.9	1.2	2	5
$\mathbb{1}\{\text{farm own}\}$	161	0.71	0.45	0	1
$\mathbb{1}\{\text{off-farm income}\}$	161	0.21	0.41	0	1
Confinement with Acreage Info.					
sales milk	313	316,528	248,363	51,110	1,602,186
sales crop	313	5,273	10,860	0	86,359
sales cattle	313	21,372	18,740	0	124,687
cow	313	122	76	22	468
labor	313	34,863	49,863	0	281,259
crop	313	47,512	40,968	0	230,954
livestock	313	167,247	134,909	13,011	746,873
machinery	313	149,177	101,677	22,063	806,265
acreage	313	339	160	90	845
$\mathbb{1}\{\text{risk averse}\}$	243	0.88	0.32	0	1
age	289	47.6	12.5	16	81
number of children	164	3.2	1.5	1	6
$\mathbb{1}\{\text{farm own}\}$	313	0.77	0.42	0	1
$\mathbb{1}\{\text{off-farm income}\}$	313	0.07	0.25	0	1

Table 2: Summary of Inefficiency Scores

Specifications			Summary Stats				Mann-Whitney Test		t-Test	
	Acreage	RTS	median	mean	75th	max	stat	p-val†	stat	p-val
(1)	Yes	VRS	1	1.050	1.048	1.612	32537	0.19	-0.25	0.80
(2)	Yes	NIRS	1	1.060	1.090	1.635	31591	0.43	-0.31	0.76
(3)	Yes	CRS	1	1.070	1.115	1.635	34121*	0.08	0.23	0.82
(4)	No	VRS	1	1.038	1.000	1.504	23313	0.41	-0.73	0.47
(5)	No	NIRS	1	1.047	1.042	1.600	22722.5	0.64	-0.53	0.60
(6)	No	CRS	1	1.052	1.073	1.600	23526	0.45	-0.48	0.63
Outliers Removed										
(7)	Yes	VRS	1	1.050	1.047	1.612	32217	0.17	-0.28	0.78
(8)	Yes	NIRS	1	1.059	1.088	1.635	31367.5	0.39	-0.36	0.72
(9)	Yes	CRS	1	1.070	1.111	1.635	33340.5*	0.07	0.24	0.81
(10)	No	VRS	1	1.037	1.000	1.504	23189	0.40	-0.81	0.42
(11)	No	NIRS	1	1.047	1.041	1.600	22610.5	0.64	-0.64	0.53
(12)	No	CRS	1	1.051	1.070	1.600	23417	0.44	-0.55	0.58

1. Summary statistics include minimum, maximum, mean, and 25th-, 50th-, and 75th-percentile values.

2. † z-value of Ranksum test (Mann-Whitney U-test) tests the hypothesis that two independent samples are from populations with the same distribution. One-sided p-value is shown for the null hypothesis confinement has higher inefficiency. Two-sided p-value for equal distributions is given by $2 * min\{p - val, 1 - p - val\}$.

3. t-stat of "H_0: two vectors x, y of *i.i.d.* normal have equal means" is calculated under unequal variances. The sign of the t-stat corresponds to $\mu(confinement) - \mu(grazers)$. (e.g. A negative t-stat indicates that confinement are more efficient than grazers, and vice versa.)

4. Statistical significance of z-value and t-stat: *** p <0.01, ** p <0.05, * p <0.1.

5. Outliers are spotted, following Wilson (1993). Number of outliers removed are 3, 5 for inputs xL and x respectively.

Table 3: Tests of Constant Returns To Scale

	Full Inputs		Without Acreage	
	(1)	(2)	(3)	(4)
H0: CRS, H1: VRS	Statistic	P-value	Statistic	P-value
(a) Ratio of means	0.988	0.260	0.983	0.553
(b) Ratio of 10% trimmed means	0.989	0.283	0.983	0.575

1. Test procedures follow Simar and Wilson (2002). Tests from other methods suggested by the authors yield qualitatively similar results.

2. Statistical significance of rejecting H_0 are represented by: *** $\alpha = 0.01$, ** $\alpha = 0.05$, * $\alpha = 0.1$ Each test is based on 400 DEA bootstrap replications under the null hypothesis.

3. Alternative hypothesis is that technology is variable returns to scale.

Table 4: Determinants of Inefficiency, OLS Estimates

			Bias Corrected	
Variables	(1)	(2)	(3)	(4)
$\mathbb{1}$\{confinement\}	-0.005	0.015	-0.017**	0.015*
$\mathbb{1}$\{risk averse\}		-0.005		-0.003
age		-0.001**		-0.001***
$\mathbb{1}$\{farm own\}		0.037***		0.059***
$\mathbb{1}$\{off-farm income\}		0.059***		0.081***
$\mathbb{1}$\{children\}		-0.074***		-0.101***
m_risk		0.003		0.021*
m_age		-0.04		-0.066***
constant	1.055***	1.093***	1.07***	1.098***
Observations	474	474	474	474
RMSE	0.103	0.097	0.085	0.078

1. Statistical significance from 400 bootstrap estimates assuming normality: *** p <0.01, ** p <0.05, * p <0.1.

2. RMSE is the standard deviation of the residuals.

3. Missing dummy for "child" is dropped due to multi-collinearity since the number of children was not recorded in the data when it is zero.

4. Finite-bias correction follows the approach in Simar and Wilson (2007).

Table 5: Second-Order Stochastic Dominance Tests

	Obs.	Min	25th	Median	Mean	75th	Max	Std. Div.
(A) Whole Sample								
Grazer	165	-81650	25490	49790	54630	83940	187900	42461
Confinement	415	-196100	19340	41310	67510	96850	640600	91538
$\hat{S}$, p-value							69729,	0.056
(B) Cows $\leq$ 200								
Grazer	165	-81650	25490	49790	54630	83940	187900	42461
Confinement	391	-196100	19280	40290	57460	89250	325000	67450
$\hat{S}$, p-value							80269,	0.034

1. Test statistic is S-hat is the maximum difference in cumulative distributions of net profits between confinement farmers and grazers. The null hypothesis is that grazer's distribution weakly second-order stochastically dominates that of confinement. Alternative hypothesis specifies violation of such a relationship.

2. P-value is calculated by a bootstrap procedure, which corresponds to bootstrap method #1 in Barrett and Donald (2003). The test is based on 1000 bootstrap replications with 1000 grid points for finding supremum.

Figure 1: Inefficiency Score Distributions For Confinement (Left) and Grazers (Right)

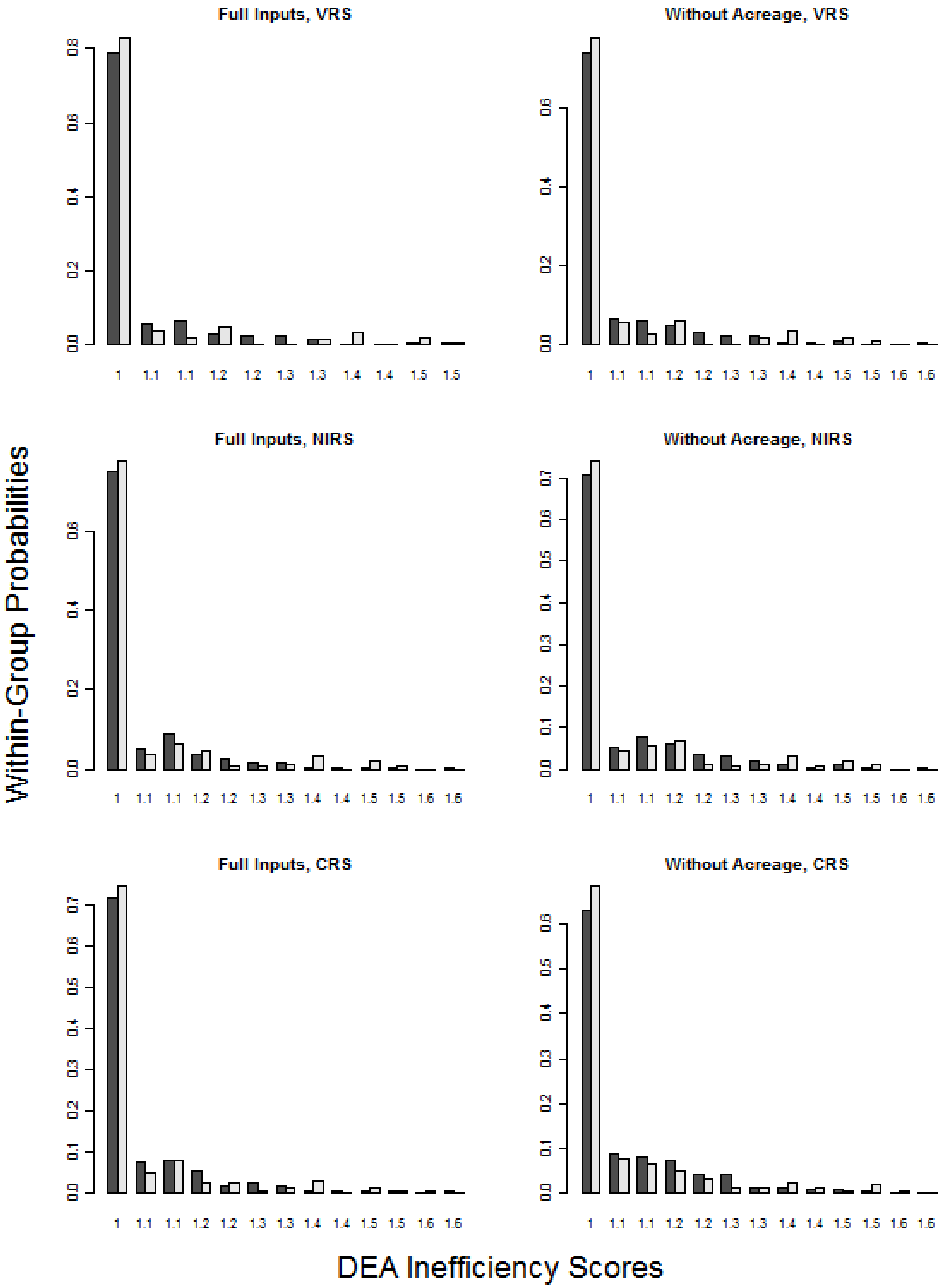

Figure 2: Distributions of Profits By Farm Type and Second-Order Stochastic Dominance