J. Dairy Sci. 99:1–7
http://dx.doi.org/10.3168/jds.2015-9942

The effect of bovine somatotropin on the cost of producing milk: Estimates using propensity scores

Loren W. Tauer[1]
Charles H. Dyson School of Applied Economics and Management, College of Agricultural and Life Sciences, Cornell University, NY 14853

ABSTRACT

Annual farm-level data from New York dairy farms from the years 1994 through 2013 were used to estimate the cost effect from bovine somatotropin (bST) using propensity score matching. Cost of production was computed using the whole-farm method, which subtracts sales of crops and animals from total costs under the assumption that the cost of producing those products is equal to their sales values. For a farm to be included in this data set, milk receipts on that farm must have comprised 85% or more of total receipts, indicating that these farms are primarily milk producers. Farm use of bST, where 25% or more of the herd was treated, ranged annually from 25 to 47% of the farms. The average cost effect from the use of bST was estimated to be a reduction of $2.67 per 100 kg of milk produced in 2013 dollars, although annual cost reduction estimates ranged from statistical zero to $3.42 in nominal dollars. Nearest neighbor matching techniques generated a similar estimate of $2.78 in 2013 dollars. These cost reductions estimated from the use of bST represented a cost savings of 5.5% per kilogram of milk produced. Herd-level production increase per cow from the use of bST over 20 yr averaged 1,160 kg.
Key words: bovine somatotropin, matching, propensity score, rbST, treatments

INTRODUCTION

The compound recombinant bST has been commercially available to US dairy producers since the year 1994. Monsanto (St. Louis, MO) marketed bST beginning in 1994 under the trademark Posilac, but sold the technology in 2008 to Eli Lilly's Animal Health Division, Elanco Animal Health (Greenfield, IN). The use of bST in lactating cows has been controversial, with opponents questioning the effect on cow health and milk composition, whereas proponents have argued it allows producing milk with fewer resources and reduces the greenhouse gas impact associated with milk production. Discussions of these arguments can be found in Dohoo et al. (2003), Collier et al. (1991), and Capper et al. (2008). St-Pierre et al. (2014) completed a recent meta-analysis of the research results addressing bST. The purpose of the current study is not to revisit those debates, which have been extensive, but to simply address one question: what has been the effect of bST on the cost of producing milk? To provide an answer, 20 yr of data on a large number of New York dairy farms was used to estimate the cost effect per 100 kg of milk produced with the use of bST compared with no use of bST using propensity score matching (**PSM**).

MATERIALS AND METHODS

The use of program evaluation methods, or treatment models, has become extensive in the literature, where researchers have been interested in the performance of treatments such as education, training, vaccination, or a new technology such as bST (Rubin, 1974; Imbens, 2000; Heckman and Navarro-Lozano, 2004). Often it is not cost effective or ethical to carry out a controlled experiment to ascertain the effect of a treatment. Thus, it is necessary to observe the effect of the treatment in the field by comparing the performance of those who received the treatment to those who did not receive the treatment. Field results can also differ significantly from experimental results under controlled conditions (Dehejia and Wahba, 2002). If participation is voluntary and not random, self-selection bias becomes an issue which may influence empirical estimates of the effect of the treatment (Imbens and Wooldridge, 2009). For example, it is generally presumed that farmers who adopt a new technology should see an improvement in their welfare or quality of life as measured by an income or leisure variable. However, because farmers themselves decide whether to adopt a new technology in a noncontrolled experiment, it may be that more highly educated and higher-income farmers are the ones who adopt, resulting in self-selection bias with an over estimate of the adoption effect. Mendola (2007) in-

Received June 10, 2015.
Accepted January 6, 2016.
[1]Corresponding author: lwt1@cornell.edu

vestigated agricultural adoption and poverty reduction in rural Bangladesh, controlling for self-selection, and determined that resource-poor farmers can improve their incomes by adopting improved seed genetics. It may also be possible that those who elect to use bST on their farm may have generated a lower cost per unit of production than those who elected not to use bST, or vice versa. Simply comparing bST users to non-bST users can lead to biased effect estimates without addressing selection bias.

A large number of techniques have been developed to estimate treatment effects. Two reviews are available from Imbens and Wooldridge (2009) and Khandker et al. (2010). In the bST effect literature, the treatment estimators that have been used include difference in differences estimation (Tauer and Knoblauch, 1997), instrumental variables (McBride et al., 2004), Heckman self-selection models (Stefanides and Tauer, 1999; Foltz and Chang, 2002; Gillespie et al., 2010), switching regression (Tauer, 2005; An, 2013), and nearest neighbor matching models (Tauer, 2009). The empirical results of these articles have been mixed with earlier studies using initial years of bST use data finding bST not profitable, whereas later studies found bST to generally be profitable. In this article, PSM, introduced by Rosenbaum and Rubin (1983), is used to measure the cost of production reduction from the use of bST on the farm. A comprehensive guide to using PSM can be found in Caliendo and Kopeinig (2008).

Treatment estimation begins with a definition of an indicator variable ($\mathbf{D_i}$), which equals 1 if individual i receives treatment and zero otherwise. The potential treatment outcomes are then measured as $Y_i(D_i)$ for each individual i, where $i = 1, \ldots, N$, with N denoting individuals, and Y the performance measure. The treatment effect (τ_i) for an individual i can be estimated as: $\tau_i = Y_i(1) - Y_i(0)$. The problem, however, is that only one of the potential outcomes is observed for each individual i. The unobserved outcome is referred to as a counterfactual. It is necessary to arrive at an estimate of a counterfactual for $Y_i(1)$ or $Y_i(0)$ depending upon whether the individual received the treatment (1) or did not receive the treatment (0). Propensity score matching techniques allow devising that counterfactual.

Propensity score matching requires estimating the probability that an individual received the treatment conditional upon characteristics of that individual. Logistic regression is used to estimate a propensity score where the dependent variable is equal to 1 if the individual received the treatment, 0 otherwise, and the independent variables are variables that are expected to determine treatment. The logistic regression estimates are used to predict the probability of each individual being treated even if they were not treated. This probability is referred to as the propensity score. Each farm that selected the treatment is then matched with a farm that did not select the treatment using the criterion that the probability of treatment selection is similar between the 2 farms; this then becomes the counterfactual. The process mimics random placement. Because the propensity score or probability is a single variable, matching is straightforward and simply entails finding the observation with the closest score. This is one-to-one matching; one-to-more than one matching was not used given the large number of one-to-one matches that were available. The difference in the performance variable provides an estimate of the treatment effect, τ_i, which is then averaged over all matches; treatment in our case was use of bST. For estimation, Stata software (version 13; StataCorp LP, College Station, TX), procedure “teffects psmatch,” was used

As stated by Khandker et al. (2010) in a World Bank Handbook on estimating treatment effects, “The validity of propensity score matching depends on two conditions: (a) conditional independence (namely, that unobserved factors do not affect participation) and (b) sizable common support or overlap in propensity scores across the participant and nonparticipant samples.” The first requirement is that the factors selected to estimate the propensity score explains treatment and no variables are omitted. The second requires overlap of the probability distribution of the propensity score so that treatment farms can be matched with nontreatment farms based on similar propensity scores.

Cornell University completes an annual dairy farm business survey in which, although voluntary, many farmers participate over multiple years. The survey collects detailed cost and revenue as well as assets and liabilities data to construct net farm income, net worth, and cash flow statements. Data on yields and characteristics of the farm are also collected. These data are primarily used for individual farm analysis, allowing individual operators to compare their farm performance to benchmark values from similar farms. This is not a random sample and generally includes farms of above performance irrespective of farming practices used on those farms. These are full-time dairy farms, because to be included milk sales must compose 85% or more of total farm receipts.

Since 1994, when bST first became commercially available, data on bST use have been collected. From 1994 through 2002, each year farmers were asked to check a category box indicating the percentage use of bST as no use, discontinued use, 0 to 25% of the herd treated, 25 to 75% of the herd treated, or 75% or more of the herd treated during the year. A measure of 100% would be treating all cows every 14 d throughout the lactation, beginning 62 d after calving and ceasing at