The Australian Journal of Agricultural and Resource Economics, 41:3, pp. 385–400

The nature of the research-induced supply shift

Michael K. Wohlgenant*

Differences among firms in a competitive industry can affect the shape of the industry supply curve. It is necessary to know how both production costs and rents are affected by research. Industry response to research will be different depending upon whether entry occurs. If the effect of entry is ignored, then the price decline from research will be overstated. Industry marginal returns can be positive with purely yield-increasing research, even when industry demand is inelastic. Standard formulas for calculating producer surplus based on linear industry supply and demand curves are strictly valid only if the analysis is restricted to short-run equilibrium behaviour.

1. Introduction

The nature of the research-induced supply shift (e.g., parallel versus pivotal) is known to have a profound and significant effect on the size and distribution of research benefits (Duncan and Tisdell 1971; Lindner and Jarrett 1978, 1980; Rose 1980; Wise and Fell 1980; Miller, Rosenblatt and Hushak 1988; Alston, Norton and Pardey 1995). While there has been much discussion in the literature about this from the perspective of relating points on a representative firm's cost curves and supply curve to the market supply curve, there has been little discussion of the effects on the shape of, and shifts in, the market supply curve when firms are different, i.e., when there are inframarginal and marginal firms in an industry. This article examines the implications of aggregation over diverse firms in an industry for measuring industry aggregate research benefits.

Since a famous exchange between Lindner and Jarrett (1980), Rose (1980), and Wise and Fell (1980), researchers have been more careful to qualify their results conditional on the type of supply shift assumed. Many researchers, however, continue to maintain the assumption of parallel supply shift – especially at the firm level – on the grounds that the simpler linear and parallel supply shift model is likely to provide a good approximation of

Michael K. Wohlgenant is William Neal Reynolds Professor of Agricultural and Resource Economics at North Carolina State University. He is also currently editor of the *American Journal of Agricultural Economics*.

the research impact (Rose 1980; Alston, Norton and Pardey 1995). The intuition of this conclusion rests on two assumptions: first, that the 'before research' equilibrium output is a large share of the new equilibrium output after adoption of the technology; and, second, that the cost impact of the research is close to the change in minimum average costs attributable to the research (Davis 1994). Given this specification of how the individual firm's cost curve, and hence supply curve, shifts, the analysis then focuses on the impact of differential adoption rates by individual firms, time lags of research adoption, spillover effects of research, etc. on aggregate research benefits to firms and society at large (Davis 1994; Alston, Norton and Pardey 1995).

While these are important considerations when moving from the individual firm or region to the aggregate level, it is important to understand that firm response to adoption of research also can depend on the nature of the firm, i.e., whether it is a marginal or inframarginal firm. In general, firms differ from one another, possessing endowments of fixed factors (e.g., soil fertility, favourable location, entrepreneurial capacity) that are in inelastic supply and may have no alternative use outside the industry in question. Thus, net returns to firms can be viewed as rents accruing to these fixed factors. Since rents are not fixed but are price-determined, adoption of research has an impact not only on production costs but also on rents received by inframarginal firms. Therefore, in contrast to the case of an industry composed of identical, marginal firms, the cost impact of the research may not be as easily approximated by the change in minimum average costs from research, and hence approximated well by a linear and parallel shift in marginal costs.

In addition to allowing for differences among firms in how research may impact costs, it is also important to take into account the effect research can have on entry of new firms into the industry. Not only may the cost impact be different for such marginal firms, but also the industry response will be different in that the supply curve from adoption of research will shift due to two components: from a change in output levels of inframarginal firms, and from a change in total output due to entry of new firms.

This article analyses the impact of different types of technical change (from research) on the industry supply curve and industry net returns (profits) when there are inframarginal and marginal firms present. Within this framework, it is found that the nature of the research-induced supply shift depends generally on how technology affects entry into the industry and the size distribution of firms in the industry, as well as outputs of both marginal and inframarginal firms. Specific forms of technical change that are popular in applied work (i.e., parallel shift and proportional shift in supply) are specified in terms of their effects on costs, and the sensitivity of the research-induced industry supply shift and industry returns to these different specifications and aggregation over firms is evaluated. A number of interesting implications are derived from the analysis, including showing that the industry supply curve is no longer linear when supply curves of individual firms are linear, and showing that producers can gain from neutral technical change when product demand is inelastic. The article concludes with discussion about the importance of aggregation over diverse firms for measurement of economic surplus.

2. The conceptual framework

Following Panzar and Willig (1978), firm diversity within the industry is modelled by assuming that the firms are distinguished by a parameter θ, indexing the firm's endowment of a fixed factor. The fixed factors are assumed to have no alternative use outside the industry in question. Strictly speaking, for this to represent long-run equilibrium, such factors could not be rented or hired by other firms and would be best interpreted as representing 'entrepreneurial capacity' (Friedman 1976, p. 90). However, in the short run, the fixed factors could represent other factors as well, so returns to the fixed factors would include returns to entrepreneurial capacity as well as returns to quasi-fixed factors (e.g., land).

A firm of type θ is characterised by the following cost function:

$$\begin{aligned} &C(q, t, \theta) \\ &\frac{\partial C}{\partial q} = C_q > 0, \frac{\partial C}{\partial t} = C_t < 0, \\ &\frac{\partial C}{\partial \theta} = C_\theta < 0, \end{aligned} \tag{1}$$

where q is output and t is an index of technical change. As indicated, costs are assumed to be increasing in output, decreasing in technical change, and decreasing with respect to θ.

Each firm of type θ maximises net returns:

$$pq - C(q, t, \theta) \tag{2}$$

where p is price of the output. The first-order conditions are represented by the Kuhn-Tucker conditions:

$$p - C_q(q, t, \theta) \leq 0, \tag{3}$$

$$q \geq 0, q(p - C_q) = 0, \tag{4}$$

$$C_{qq} > 0. \tag{5}$$

Condition (4) specifies that only firms participating in the industry are required to produce where price equals marginal cost; firms not participating (i.e., $q = 0$) face a price that is less than or equal to marginal cost. For participating firms, $q > 0$ and marginal cost will be rising according to condition (5). The firm's supply function,

$$q^*(p, t, \theta), \tag{6}$$

has the following derivatives:

$$\frac{\partial q^*}{\partial p} = q_p^* = \frac{1}{C_{qq}} > 0, \tag{7}$$

$$\frac{\partial q^*}{\partial t} = \frac{-C_{qt}}{C_{qq}}, \tag{8}$$

$$\frac{\partial q^*}{\partial \theta} = \frac{-C_{q\theta}}{C_{qq}}. \tag{9}$$

The derivative indicated by (7) is positive, implying that the firm's supply curve is positively sloped in all cases. (This will be true even when the industry supply curve is horizontal – see Friedman 1976, pp. 125–6.) We expect in all normal cases that marginal cost of output will decline as t increases; hence, the derivative in (8) is expected to be positive. The sign of the derivative in (9) is generally indeterminate because the effect of θ on marginal cost of output is in general ambiguous.

Industry equilibrium, where marginal firms earn exactly zero rents, can be characterised by the following condition:

$$pq^*(p, t, \hat{\theta}) - C[q^*(p, t, \hat{\theta}), t, \hat{\theta}] = 0 \tag{10}$$

where:

$$\hat{\theta} = \hat{\theta}(p, t) \tag{11}$$

is the value of θ that allows marginal firms to (optimally) earn zero rents (Panzar and Willig 1978, p. 475). Because the firm's net returns, equation (2) and equation (10), are an increasing function of θ, firms earning less than $\hat{\theta}$ will not participate in the industry. Using (1), (4) and (10), we find that (11) has the following properties:

$$\frac{\partial \hat{\theta}}{\partial p} = -\left[(p - C_q)\frac{\partial q^*}{\partial p} + q^*\right] \div \left[(p - C_q)\frac{\partial q^*}{\partial \theta} - C_\theta\right] = \frac{q^*}{c_\theta} < 0 \tag{12}$$

$$\frac{\partial\hat{\theta}}{\partial t} = -\left[(p - C_q)\frac{\partial q^*}{\partial t} - C_t\right] \div \left[(p - C_q)\frac{\partial q^*}{\partial \theta} - C_\theta\right] = \frac{-C_t}{C_\theta} < 0. \tag{13}$$

Conditions (12) and (13) indicate that the minimum value of θ required for firms to participate in the industry is a decreasing function of both output price and technical change.

Given this specification of marginal firms, we can now write the industry supply function as:

$$S(p, t) = \int_{\hat{\theta}}^{\bar{\theta}} q^*(p, t, \theta) f(\theta) d\theta \tag{14}$$

where $f(\theta)$ represents the (exogenous) density function, showing how many firms there are of type θ (Panzar and Willig 1978, p. 475); $\hat{\theta}$ is the maximum value of θ observed.

Finally, industry equilibrium occurs at a price, p^e, where industry quantity supplied equals quantity demanded, i.e.:

$$S(p^e, t) = D(p^e) \tag{15}$$

where $D(p)$ is the negatively sloped demand function for industry output.[1]

3. Impact of technical change on industry supply

The effect of technical change on industry supply can be ascertained through evaluating the derivative of equation (14) with respect to t, which can be written as:

$$\frac{\partial S}{\partial t} = \int_{\hat{\theta}}^{\bar{\theta}} \frac{\partial q^*}{\partial t} f(\theta) d\theta - q^*(p, t, \hat{\theta}) f(\hat{\theta}) \hat{\theta}_t \tag{16}$$

This derivative indicates that technical change has two effects on industry supply: (a) the effect on output of (existing) inframarginal firms (the first term on the right-hand side of equation (16)), and (b) the effect on output

[0] The present framework can accommodate the case where some of the inputs supplied to the industry have less than perfectly elastic supply curves. This is because the industry supply elasticity can be viewed as the supply elasticity for an industry with pecuniary diseconomies (or possibly pecuniary economies if negatively sloped). This generalisation can be taken into account by regarding the aggregate supply function in (14) as a general equilibrium supply function, showing the sum of firms' outputs at various prices with input prices for less than perfectly elastically supplied inputs adjusting to clear the market as industry output changes. None of the qualitative findings in this article are affected by ignoring this case, and the surplus formulas derived could be modified simply by changing the value of the industry supply elasticity to reflect pecuniary diseconomies.

from entry of new firms into the industry (the second term on the right-hand side of equation (16)). Effect (b) can be decomposed into the impact of technical change on θ times the proportion of output contributed by new entrants into the industry, $q^*(p, t, \theta)f(\theta)$.

To see how the industry supply function shifts at different output price levels, differentiate equation (16) with respect to p to yield

$$\frac{\partial^2 S}{\partial t \partial p} = \int_{\hat{\theta}}^{\bar{\theta}} \left[\frac{\partial^2 q^*}{\partial t \partial p} + \frac{\partial^2 q^*}{\partial t \partial \theta} \frac{\partial \hat{\theta}}{\partial p} \right] f(\theta) d(\theta) - \frac{\partial q^*(p, t, \hat{\theta})}{\partial t} f(\hat{\theta}) \hat{\theta}_p \\ - \left[\frac{\partial q^*(p, t, \hat{\theta})}{\partial p} f(\hat{\theta}) \hat{\theta}_t + q^*(p, t, \hat{\theta}) f'(\hat{\theta})(\hat{\theta}_t)(\hat{\theta}_p) + q^*(p, t, \hat{\theta}) f(\hat{\theta}) \hat{\theta}_{tp} \right] \quad (17)$$

As equation (17) shows, how the industry supply curve shifts at different price levels depends on a number of factors. In addition to the effect on the supply functions of inframarginal firms (the terms in brackets under the integral), there are also effects on the output and number of firms entering the industry that need to be considered. Even when all firms are inframarginal, i.e., when $f(\hat{\theta}) = 0$ (Panzar and Willing 1978, pp. 476–7), the effect on industry supply, which reduces to just the first term in equation (17), still is more complicated than typically perceived. In particular, note that, in addition to knowing how the slope of the supply curve of the inframarginal firm is affected by t, we would also need to know how the supply curve shifts from changes in θ, induced by changes in price. In other words, the whole term in brackets under the integral is what comprises the impact of technical change on the firm's output, not just the first term in brackets. Therefore, in view of (7), (8) and (9), it is necessary to know both the impact of technical change on marginal cost of output as well as marginal cost of fixed factors when ascertaining the effect on the shift of the supply curve.

While the special case of zero marginal firms, i.e., $f(\hat{\theta}) = 0$, may be viewed as a special case of long-run equilibrium (Panzar and Willing 1978, p. 477), it also may be thought of as applying to the short run when entry into the market is prohibited. In that case, returns to fixed factors would include quasi-rents to temporarily fixed factors as well as rents accruing to entrepreneurial capacity. Yet, as the above discussion shows, even in the short run we should expect the effect of a supply shift to be more complicated than just an assessment of the effect technical change has on production costs. Because rents, whether temporarily or permanently fixed, are price-determined, how the supply curve shifts will depend as well on how rents are affected. This clearly makes the problem of measuring the type of supply shift even more complicated than typically perceived.

Another interesting feature of equation (17) is that it implies that the industry supply curve will not necessarily be linear when supply curves of individual firms are linear. This can be seen by observing that when individual firms' supply curves are linear, only the first term under the integral disappears. In this case, the slope of the supply curve will generally only be constant if there are no marginal firms.

4. Impact of technical change on industry returns

Industry net returns from introduction of technology can be defined as follows:

$$\Pi(t) = \int_{\hat{\theta}}^{\bar{\theta}} \{p^e(t)q^*[p^e(t), t, \theta] - C(q^*, t, \theta)\} f(\theta) d\theta \tag{18}$$

The derivative of industry net returns with respect to t is:

$$\begin{aligned} \frac{\partial \Pi}{\partial t} &= \int_{\hat{\theta}}^{\bar{\theta}} \left\{ (p^e - C_q) \left(\frac{\partial q^*}{\partial p} \frac{\partial p^e}{\partial t} + \frac{\partial q^*}{\partial t} \right) + q^* \frac{\partial p^e}{\partial t} - C_t \right\} f(\theta) d\theta \\ &\quad - \left\{ p^e(t) q^* \left[p^e(t), t, \hat{\theta} \right] - C(q^*, t, \hat{\theta}) \right\} f(\hat{\theta}) \left[\frac{\partial \hat{\theta}}{\partial p} \frac{\partial p^e}{\partial t} + \frac{\partial \hat{\theta}}{\partial t} \right] \end{aligned} \tag{19}$$

which upon using (4) and (10) simplifies to:

$$\begin{aligned} \frac{\partial \Pi}{\partial t} &= - \int_{\hat{\theta}}^{\bar{\theta}} \left(q^* \frac{\partial p^e}{\partial t} - C_t \right) f(\theta) d\theta \\ &= \frac{\partial p^e}{\partial t} S(t) - \int_{\hat{\theta}}^{\bar{\theta}} C_t f(\theta) d\theta, \end{aligned} \tag{20}$$

in light of the definition of industry supply in equation (14). Recalling the equilibrium condition (15), the derivative of equilibrium price with respect to t is:

$$\begin{aligned} \frac{\partial p^e}{\partial t} &= \frac{-\partial S/\partial t}{\partial S/\partial p - D'} \\ &= - \left[\int_{\hat{\theta}}^{\bar{\theta}} \frac{\partial q^*}{\partial t} f(\theta) d\theta - q^*(p^e, t, \hat{\theta}) f(\hat{\theta}) \hat{\theta}_t \right] \\ &\quad \div \left[\int_{\hat{\theta}}^{\bar{\theta}} \frac{\partial q^*}{\partial p} f(\theta) d\theta - q^*(p^e, t, \hat{\theta}) f(\hat{\theta}) \hat{\theta}_p - D' \right] \end{aligned} \tag{21}$$

Substituting (7) and (8) into (21), and substituting the result into (20), we obtain the following expression for the impact of technical change on industry supply in terms of the underlying firms' cost functions:

$$\frac{\partial \Pi}{\partial t} = -S\left\{\int_{\hat{\theta}}^{\bar{\theta}} (-C_{qt}/C_{qq}) f(\theta) d\theta - q^*(p^e, t, \hat{\theta}) f(\hat{\theta}) \hat{\theta}_t \div \left[\int_{\hat{\theta}}^{\bar{\theta}} (1/C_{qq}) f(\theta) d\theta - q^*(p^e, t, \hat{\theta}) f(\hat{\theta}) \hat{\theta}_p - D'\right]\right\} - \int_{\hat{\theta}}^{\bar{\theta}} C_t f(\theta) d\theta \tag{22}$$

The impact of technical change, as indicated by equations (20) and (21) jointly or equation (22) by itself, depends on a number of factors. The two general effects, as shown by equation (20), consist of the impact of technical change directly on costs (the second term) and the impact of technical change on equilibrium price (the first term). As anticipated, the direct effect depends crucially on how individual firms' cost curves shift in response to t, and on how many firms there are of each type in the industry.

One of the more interesting aspects of the impact of technical change on industry net returns relates to the impact of t on industry equilibrium price. As shown in equation (21), there are not only dual effects of technical change on inframarginal and marginal firms, but there are also dual effects of price on inframarginal and marginal firms to take into account when estimating the effect of t on industry price. In particular, the first term in the denominator of equation (21), which represents the slope of the industry supply curve, shows that price has two effects on industry supply: (a) the effect on output of inframarginal firms, and (b) the effect on output from entry of new firms induced by an increase in price. The significance of the second effect of technical change on entry and output price on entry is that, *ceteris paribus*, industry equilibrium price falls less than if entry is not permitted. In order to see why this is true, note that when all firms are inframarginal (i.e., $f(\hat{\theta}) = 0$), or equivalently that entry is prohibited, then equation (21) becomes:

$$\frac{\partial p^e}{\partial t} = -\left[\int_{\hat{\theta}}^{\bar{\theta}} \frac{\partial q^*}{\partial t} f(\theta) d\theta\right] \div \left[\int_{\hat{\theta}}^{\bar{\theta}} \frac{\partial q^*}{\partial p} f(\theta) d\theta - D'\right], \tag{23}$$

which is clearly larger in absolute value than the general case where entry is permitted, equation (21). Therefore, by ignoring the differential effect entry can have on market equilibrium, the impact of the price decline from introduction of technology will be overstated.

5. Impact of technical change for specific types of cost shifts

In order to gain more insight into the impact aggregation over firms has on industry supply and returns, it is necessary to consider some special types of technological innovations. The effects of two types are considered here: (a) the same absolute decrease in marginal and average costs at all levels of output (i.e., the total cost function is additively separable in q and t such that $C = C(q) + aqt$), and (b) a proportionate reduction in costs (i.e., technical change is proportional with respect to output). The first type of cost change leads to a parallel shift in the representative firm's supply curve (ignoring effects on rents), while the second type of cost change leads to a proportionate reduction in costs and is often referred to as a strictly yield-enhancing type of technical change. The second type of cost change also may be viewed as that resulting from what Blackorby, Lovell and Thursby (1976) refer to as 'extended Hicks Neutral Technical Change' (EHN), which implies that the cost function for non-constant returns to scale technologies may be represented as follows (Lemieux and Wohlgenant 1989, p. 905):

$$C = C(q/t, \theta) \tag{24}$$

This particular cost function can be shown to have the following derivatives:

$$C_t = -C_q q \tag{25}$$

$$C_{qt} = -C_{qq} q - C_q \tag{26}$$

where the initial value of t is assumed to be unity simply for sake of notational convenience.

In order to focus the discussion further, it is assumed that research changes all inframarginal firms in the same manner, and that all firms adopt the new technology at the same time. As shown below, even under these very restrictive conditions, different effects of technical change can occur. The key to understanding the different effects is in making a distinction between inframarginal and marginal firms.

When average costs shift downward in a parallel manner, the change in average costs from research, C_t/q, equals the change in marginal cost from research, $C_t q$. Assuming this change in costs is the same across all inframarginal firms implies that equation (22) can be written as

$$\begin{aligned}\frac{\partial \Pi}{\partial t} = &-S\{[-C_{qt}(\partial S/\partial p) - (\hat{C}_{qt}\hat{\theta}_p + \hat{\theta}_t)\hat{q}^* f(\hat{\theta})] \\ &\div [\partial S/\partial p - D']\} - \int_{\hat{\theta}}^{\bar{\theta}} C_t f(\theta) d\theta\end{aligned} \tag{27}$$

where:

$$\frac{\partial S}{\partial p} = \int_{\hat{\theta}}^{\bar{\theta}} (1/C_{qq}) f(\theta) d\theta - \hat{q}^* f(\hat{\theta}) \hat{\theta}_p \tag{28}$$

Multiplying and dividing C_t by q, factoring out the common term, C_t/q, and using (14), equation (27) further simplifies to

$$\begin{aligned} \frac{\partial \Pi}{\partial t} = S\{[C_{qt}(\partial S/\partial p) + (\hat{C}_{qt}\hat{\theta}_p + \hat{\theta}_t)\hat{q}^* f(\hat{\theta})] \\ \div [\partial S/\partial p - D']\} - (C_t/q)S \end{aligned} \tag{29}$$

Using the results from (12) and (13) and the fact that $C_{tq} = C_t/q$ in the case of a parallel cost shift, we find that the second term disappears and that (29) reduces to the following expression:

$$\frac{\partial \Pi}{\partial t} = [(C_t/q)(SD')/(\partial S/\partial p - D')] > 0. \tag{30}$$

This says that for a parallel cost shift, marginal returns from research will be positive regardless if there are inframarginal firms or not.

The case of a proportional shift in costs is a little more complicated than the case of a parallel cost shift. In the parallel cost shift case both marginal cost and average cost change by the same absolute amount, so specification of change in one of the costs across firms implies the same change in the other costs. In the case of a proportional shift in costs, the change in average costs, C_t/q, must be the same across all firms (because $C_q = p$ according to equation (25)), but this does not necessarily imply the same change in marginal cost across firms (see equation (26)). For purposes of comparison, it is useful to analyse the effect on industry returns for the alternative cases where all firms' average costs change by the same absolute amount (according to equation (25), and hence marginal costs change according to equation (26)), and where all firms' marginal costs, C_{qt}, change by the same absolute amount.

Considering the case where all firms' average costs change by the same absolute amount, C_t/q, we have upon using (14), (25), (26) and (28) in (22) that:

$$\begin{aligned} \frac{\partial \Pi}{\partial t} = -S\{[S + C_q \partial S/\partial p + \hat{q}^* f(\hat{\theta})(C_q \hat{\theta}_p - \hat{\theta}_t)] \\ \div [\partial S/\partial p - D']\} - (C_t/q)S \end{aligned} \tag{31}$$

Substituting from (12) and (13) and using (4) we obtain:

$$\begin{aligned} \frac{\partial \Pi}{\partial t} &= -S(S + pD') \div (\partial S/\partial p - D') \\ &= -pS(1 + \eta)/(\varepsilon - \eta) \end{aligned} \tag{32}$$

where $\eta = (\partial D/\partial p)(p/D)$ is the price elasticity of demand and $\varepsilon = (\partial S/\partial p)(p/S)$ is the price elasticity of industry supply.

This says that, for this cost specification, industry returns will decline (increase) according as industry demand is inelastic (elastic). As in the case of a parallel cost shift, industry returns in this case do not depend upon whether the firms are marginal or inframarginal producers. This result is the same as that of Miller, Rosenblatt and Hushak (1988) when the industry returns are modelled with an industry supply curve that has a pivotal shift.

However, the above result no longer obtains when we assume research reduces all firms marginal costs by the same amount. To see this, note that equation (27) can be used to evaluate this case since the change in marginal cost is the same across all firms. In addition, from (4) and (25) we see that the change in average cost also will be the same across all firms, so that equation (29) is the relevant equation to use. In contrast to the previous cases, though, the second term in the numerator does not disappear. Rather, upon substituting from (12), (13), (25) and (26) we obtain the following expression for the marginal effect on industry returns from research:

$$\frac{\partial \Pi}{\partial t} = -pS[(1+\eta) - \mu]/(\varepsilon - \eta) \tag{33}$$

where $\mu = -(1/\varepsilon)(q^* \hat{f}/S)\hat{\theta}_p p$.

Note that this expression differs from equation (32) by the term μ, which is positive in view of (12). Therefore, we derive the new result that marginal industry returns from research can be positive when industry demand is inelastic (i.e., η less than one in absolute value) and when there is neutral technical change. This happens in this case because, as shown by equation (16) after substituting equations (8) and (13) and factoring out the common term C_{qt}, costs decline more for inframarginal firms than for marginal firms. This is different from the case of a parallel shift in costs because, in that case, costs change by the same absolute amount for inframarginal and marginal firms. In the case where marginal costs change proportionately across firms, there is also no differential effect between marginal and inframarginal firms, so the shift in the industry supply curve also will be proportionate.

6. Implications for measurement of economic surplus

Other types of technical change could also be evaluated to see how sensitive marginal returns to research are to diverse firms in the industry. The previous cases were evaluated primarily because of their popularity in applied work. The purpose of this section is to draw out some further implications of aggregation for commonly used approaches to welfare measurement and to discuss implications for applied welfare analysis.

Following Martin and Alston (1994), the approach to measurement of economic surplus (i.e., producer surplus) adopted here is based on the profit function approach, where a second-order Taylor's series expansion around the initial value of industry profits is used to approximate the change in industry returns. In contrast to Martin and Alston (1994), the specification used here takes output price as endogenous. It is also assumed that firms' cost functions can be adequately approximated by quadratic functions and that the industry demand function is linear. Given this specification, the change in industry profit associated with a parallel shift in costs can be written as:

$$\begin{aligned}\Delta\Pi &= \frac{\partial\Pi}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2\Pi}{\partial t^2}(\Delta t)^2 \\ &= \frac{\partial\Pi}{\partial t}\Delta t + \frac{1}{2}\left[\frac{\partial\Pi}{\partial t}\frac{S_t}{S} - \frac{\partial\Pi}{\partial t}\frac{S_{pt}}{(S_p - D')}\right](\Delta t)^2 \\ &= \frac{\partial\Pi}{\partial t}\Delta t\left[1 + \frac{1}{2}\frac{S_t}{S}\Delta t - \frac{1}{2}\frac{S_{pt}\Delta t}{(S_p - D')}\right]\end{aligned} \tag{34}$$

where equation (30) is differentiated with respect to t, given that C_t/q is exogenous. It is not hard to show, aside from the third term in brackets, that this has precisely the same form as the producer surplus formula for the case of a linear industry supply function with a parallel supply shift from research.[2] As can be seen from equation (17), the difference between the two formulas stems from the fact that the industry supply curve is no longer linear when supply curves of individual firms are linear.

The significance of this result can be understood upon examining equation (17) further. Note first that only the first term disappears from this equation

---

[0] For linear supply and demand functions and a parallel supply shift, change in producer surplus can be written as follows (see, e.g., Alston and Scobie 1983):

$$\Delta PS = p_0 S_0\left(\frac{\Delta p}{p_0} - k\right)\left(1 + \frac{1}{2}\frac{\Delta S}{S_0}\right)$$

where $\Delta p = p_1 - p_0$, p_0 is the initial price, p_1 is the new price, $\Delta S = S_1 - S_0$, S_0 is the initial quantity supplied, S_1 is the new quantity supplied, and k is the reduction in costs as a fraction of the initial price p_0. Let $k = \Delta t C_{qt}/p_0 = \Delta t(C_t/q)/p_0$. Since $\Delta p = (\partial p/\partial t)\Delta t$ and $\Delta S = (\partial S/\partial t)\Delta t$ for linear supply and demand functions, change in producer surplus can be written as:

$$\begin{aligned}\Delta PS &= S_0\left[\frac{\partial p}{\partial t} - (C_t/q)\right]\Delta t\left[1 + \frac{1}{2}\frac{(\partial S/\partial t)}{S_0}\Delta t\right] \\ &= \frac{\partial\Pi}{\partial t}\Delta t\left[1 + \frac{1}{2}\frac{S_t}{S_0}\right]\end{aligned}$$

in light of equation (20) when C_t/q is the same across all firms.

when supply curves of individual firms are linear. We are still left with additional terms showing the impact of technical change and output price on quantity supplied of marginal producers and the size distribution of firms in the industry (the latter reflected in the derivative of the density function). In general, the sign of $\partial^2 S/\partial p \partial t$ is indeterminate because the signs of the terms involving the derivative of the density function and the cross partial derivative of θ with respect to t and p are generally unknown. Even knowledge of the density function (e.g., uniform distribution) would not help, because the sign of the last term depends upon how marginal cost of the fixed factor θ changes as q changes (see equation 13).

Similar differences are found in the case of a proportional shift in supply. From a Taylor's series expansion around the initial level of profits, change in profit associated with neutral technical change for the two types of cost changes indicated by equations (32) and (33), respectively, are as follows:

$$\Delta\Pi = \frac{\partial\Pi}{\partial t}\Delta t\left[1 + \frac{1}{2}\frac{S_t}{S}\Delta t + \frac{1}{2}\left(\frac{S_t + D'\partial p^e/\partial t}{(S + pD')}\right)\Delta t - \frac{1}{2}\frac{S_{pt}\Delta t}{(S_p - D')}\right], \qquad (35)$$

provided $S \neq -pD'$; $\Delta\Pi = 0$ when $S = -pD'$.

$$\Delta\Pi = \frac{\partial\Pi}{\partial t}\Delta t\left\{1 + \frac{1}{2}\frac{S_t}{S}\Delta t + \frac{1}{2}\frac{[S_t(1-\mu) + D'\partial p^e/\partial t - S(\partial\mu/\partial t)]}{[S(1-\mu) + pD']}\Delta t - \frac{1}{2}\frac{S_{pt}\Delta t}{[S_p - D']}\right\} \qquad (36)$$

Provided $S(1-\mu) \neq -pD'$; $\Delta\Pi = 0$ when $S(1-\mu) = -pD'$.

Thus, we find the addition of a term related to S_{pt} and the addition of a term related to how μ in equation (33) change as t changes. As before, the signs of these terms are indeterminate.

In summary, we find that the standard formulas used for calculating producer surplus are generally incorrect when we have an industry composed of diverse firms. Additional information is required on the distribution of firms within the industry, how technical change affects inframarginal versus marginal firms, and how inframarginal rents are affected by firm output and the impact of technical change.

In general, information on these components will be lacking. In such cases, the standard formulas for calculating producer surplus may be all that are available to the researcher. The question then becomes under what circumstances would the formula be expected to yield accurate results. If the analysis is to focus on long-run response, then one would need to assess the magnitude of the error using the correct formula like those presented in equations (34) through (36), and to make assumptions about the likely impact of technical change on the entry of firms into the industry. Note that if all firms are inframarginal, or that the length of run is the short run where entry is prohibited, then equations (34) and (35) will be identical to the standard formulas used to estimate producer surplus with parallel supply shifts and proportional supply shifts. Even then, however, questions might be raised about the appropriate supply elasticities used, i.e., do the estimated elasticities reflect strictly short-run response?

Of course, if one did have the required information on individual firms, then it would be straightforward to obtain the impacts on industry returns through numerical integration of equation (21) over the relevant range for Δt. In such a case, econometric analysis could be employed to estimate cost or profit function parameters for different types of firms using an appropriate flexible functional form like the translog, normalised quadratic, or generalised Leontief. The cost or profit function then could be used directly with an exogenously estimated density function, showing the proportion of firms of each type, and an exogenously specified form of technical change, formulated in terms of how it would be expected to influence firms' costs and profits.[3]

7. Conclusions

The main conclusion from this article is that differences among firms in a competitive industry can affect the shape of, and shifts in, the industry supply curve. Presence of inframarginal firms makes the analysis of the effect of technical change on firms' costs more complicated than when the industry is composed of identical, marginal firms. In particular, with inframarginal firms, it is necessary to know how both production costs and rents (or quasi-rents) are affected by research in order to quantify the effect on firms' supply response. Among other things, this suggests that the assumption of a parallel supply shift at the individual firm level is less tenable.

Another effect from allowing for differences among firms is that the industry supply curve no longer bears a direct relationship to individual firms' supply curves. This is because industry response to a price change consists of the output responses by existing firms and output response from the number of new firms entering the industry. This means that the industry supply curve will not necessarily be linear when supply curves of individual firms are linear. Exogenous changes to the industry, like technical change from research, will cause industry response to price changes to be different depending upon whether the change induces entry or not.

Industry marginal returns also can be affected by the presence of inframarginal firms. If the effect of entry on market equilibrium is ignored,

[0] One possibility for a density function would be the log-normal distribution, which is roughly consistent with the data for farm size (Sumner and Leiby 1987).

then the price decline will be overstated and industry returns from research will be understated.

The impacts of technical change on industry marginal returns from specific types of cost shifts were evaluated. For a parallel shift in the cost function, marginal returns from research only depend on total industry response to price changes, and not the distribution of output between marginal and inframarginal firms. The same result holds for cost shifts that are proportional across firms. However, if research also decreases inframarginal firms' costs by the same absolute amount, then industry marginal returns will be affected. In particular, it is shown that industry marginal returns from research can be positive even when industry demand is inelastic and there is neutral technical change. This effect will be more likely the higher the proportion of increased output supplied by new firms and the more responsive rents are to changes in output price.

Finally, standard formulas for calculating producer surplus and industry returns were evaluated when the assumption of identical firms was relaxed. Approximating the returns by a second order Taylor's series expansion, it is shown that the standard formulas based on linear industry supply curves and parallel and proportional supply shifts only hold if there are only inframarginal firms in the industry, i.e., that the industry is in short-run equilibrium, but not necessarily long-run equilibrium. The errors result from the fact that the industry supply curve will generally not be linear when individual firms' supply curves are linear. Neither the signs nor magnitudes of the errors are known, but they depend on the distribution of firms in the industry and on how marginal costs of the fixed factors change as output changes.

Given the problems with standard formulas for calculating producer surplus from research, the preferred approach would be to obtain information on estimated cost functions by firm type, information on the distribution of firms within the industry, and information on the way in which research would be expected to shift marginal cost curves. One then could use this information, combined with appropriate information on functional forms and with the formula for industry marginal returns (equation (21)), to estimate industry returns to research via numerical integration techniques directly.

References

Alston, J.M., Norton, G.W. and Pardey, P.G. 1995, *Science Under Scarcity*, Cornell University Press, Ithaca, New York.

Alston, J.M., and Scobie, G.M. 1983, 'Distribution of research gains in multistage production systems: comment', *American Journal of Agricultural Economics*, vol. 65, no. 2, pp. 353–6, May.

Blackorby, C., Lovell, C.A.K. and Thursby, M.C. 1976, 'Extended Hicks neutral technical change', *Economic Journal*, vol. 86, no. 344, pp. 845–52, December.

Davis, J. 1994, *Disaggregation Rather than Mathematical Manipulation for Incorporating Research Impacts on Supply*, Australian Centre for International Agricultural Research, Economic Evaluation Unit Working Paper no. 3, ACT, Canberra.

Duncan, R. and Tisdell, C. 1971, 'Research and technical progress: the returns to producers', *Economic Record*, vol. 47, no. 117, pp. 124–9, March.

Friedman, M. 1976, *Price Theory*. Aldine Publishing Co., New York.

Lemieux, C.M. and Wohlgenant, M.K. 1989, 'Ex ante evaluation of the economic impact of agricultural biotechnology: the case of porcine somatotropin', *American Journal of Agricultural Economics*, vol. 71, no. 4, pp. 903–14, November.

Lindner, R.K., and Jarrett, F.G. 1978, 'Supply shifts and the size of research benefits', *American Journal of Agricultural Economics*, vol. 60, no. 1, pp. 48–58, February.

Lindner, R.K. and Jarrett, F.G. 1980, 'Supply shifts and the size of research benefits: reply', *American Journal of Agricultural Economics*, vol. 62, no. 4, pp. 841–4, November.

Martin, W. and Alston, J.M. 1997, 'Producer surplus without apology? Evaluating investments in R&D', *Economic Record*, forthcoming.

Miller, G.Y., Rosenblatt, J.M. and Hushak, L.J. 1988, 'The effects of supply shifts on producers' surplus', *American Journal of Agricultural Economics*, vol. 70, no. 4, pp. 886–91, November.

Panzar, J.C. and Willig, R.D. 1978, 'On the comparative statics of a competitive industry with inframarginal firms', *American Economic Review*, vol. 68, no. 3, pp. 474–8, June.

Rose, R.N. 1980, 'Supply shifts and research benefits: comment', *American Journal of Agricultural Economics*, vol. 62, no. 4, pp. 834–44, November.

Sumner, D.A. and Leiby, J.D. 1987, 'An econometric analysis of the effects of human capital on size and growth among dairy farms', *American Journal of Agricultural Economics*, vol. 69, no. 2, pp. 465–70, May.

Wise, W.S. and Fell, E. 1980, 'Supply shifts and the size of research benefits: comment', *American Journal of Agricultural Economics*, vol. 62, no. 4, pp. 838–40, November.