Chapter 12

THE AGRICULTURAL PRODUCER: THEORY AND STATISTICAL MEASUREMENT

RICHARD E. JUST

Agricultural and Resource Economics, University of Maryland, College Park, MD

RULON D. POPE

Department of Economics, Brigham Young University, Provo, UT

Contents

Handbook of Agricultural Economics, Volume 1, Edited by B. Gardner and G. Rausser

JEL classification: Q11

1. Introduction

Agricultural production economics grew out of the study of farm management. Farm management grew out of the study of agronomy and horticulture. Early courses in farm management particularly at Cornell were largely empirically based and sought to develop the underlying economic principles through replication of experiments [Jensen (1977)]. "As marginal analysis reached a climax with Alfred Marshall, agricultural economics was just beginning to emerge as a discipline in land-grant colleges" [Johnson (1955), p. 206]. During the 1920s and 1930s, production economics began to emerge as an integrated field that analyzed farm management and production issues from farming to and including marketing of agricultural products. As in other fields of economics, the unifying paradigms for this emerging discipline were marginal economic analysis, comparative advantage, and competition [Jensen (1977)]. That agricultural production and farm management economics embraced these central economic paradigms of the time was indisputable and as such it could properly be viewed as a subdiscipline of economics. Because the issues and problems were agricultural, most agricultural economists to this day reside in colleges of agriculture throughout the world.

The marriage of economic paradigms to farm management and production economic issues is widely viewed as successful. Agricultural economists working with other agricultural scientists have enlightened many both as to normative and positive economic choices. However, many agricultural economists particularly of older vintages likely identify more with agricultural sciences and less with economics compared to younger vintages who tend to identify more with economics as the parent discipline [Pope and Hallam (1986)].

How and why does agricultural production economics differ from the application of economic principles to other production activities in the economy? Clearly, the goods and services studied are different and that alone may justify a separate field of study. However, in a deeper sense, is the current or proper methodology for studying agricultural production different than for studying, say, manufacturing? A basic question that must be addressed in a volume such as this is, "Why is the study of agricultural economics different than the study of the economics of any other sector?" and in particular, "What are the distinguishing features of agricultural production economics?"

In this chapter, we emphasize the production issues that differentiate agriculture from manufacturing. We begin in the following section by identifying a number of unique features of agricultural production – features not necessarily unique in their existence but unique by their combination and predominance in agriculture. While some mathematical characterizations in this section facilitate understanding, they are merely illustrative with formal analysis delayed to later sections. The purpose of Section 2 is to raise issues and questions related to the unique features of agriculture that are addressed in subsequent sections. The general conclusion is that agricultural production is heavily structured because of spatial, temporal, and stochastic issues. Section 3 develops a set of economic principles that are needed to address a sector dominated by such features. Some examples are used to illustrate the points with no attempt to achieve generality.

The general conclusion is that serious errors can be made if structural issues are ignored in analysis. Section 4 then develops some fundamental theoretical considerations needed to address the principles identified in Section 3 with generality at least in a short-run context. This backdrop is used to discuss the extent to which agricultural production economics, as depicted by the previous chapters in this Handbook, has addressed these needs. The implications of these results are that (i) reduced-form approaches that initiate empirical work from an arbitrary specification of the production possibilities frontier cannot determine many important characteristics of technology, (ii) approaches that under-represent structure are not useful for policy analysis because they embed policy assumptions, (iii) both early primal applications and standard current applications of duality have tended to focus on reduced-form representations, (iv) both dual and primal approaches should be expanded to consider a qualifying degree of structure, and (v) examination of structure is limited by data availability. In Section 5 we consider other needed generalizations that come into play in moving beyond the short run and the extent of related empirical progress thus far. This leads to a critical evaluation of the state of data for agricultural production analysis, a call for action to improve the scope of data, and a conclusion that the current state of agricultural production analysis is heavily limited by data availability.

2. Uniqueness of agricultural technology

Perhaps the most important reason for studying agricultural production separately is the uniqueness of agricultural technology associated with its biological nature and exposure to widely varying and unpredictable elements of nature. This section discusses some of the main features that differentiate agricultural production: (i) lags and intertemporal complexity with limited observability caused by biological processes, (ii) uncertainty in biological processes related to weather and pests, (iii) multiple outputs with cyclical flexibility in the output mix related to growing seasons, (iv) technological change with fragmented and mixed adoption associated with both physical and biological capital adjustment, and (v) atomistic heterogeneity in major characteristics such as soil productivity, climate, infrastructure, environmental sensitivity, farmer abilities, etc. While some limited parallels can be found with some of these features in other sectors, the combination and extent found in agriculture have critical implications for the ability to represent them empirically. They dramatically affect all other aspects of the agricultural sector including domestic markets, international trade, finance, environmental concerns, and policy issues. For example, unanticipated national crop failures cause dramatic swings in world markets and trade as in the commodity boom of the 1970s [Chambers and Just (1981)], and the spatial correlations of production practices with environmental characteristics dramatically influence environmental quality and response to policies [Just and Antle (1990)].

During the first half century of agricultural economics study, many agricultural production economists cooperated with the biological and soil science disciplines to integrate representations of biological and chemical processes and better represent the intricacies of relevant biological and physical relationships. As in engineering economics, there was a substantive interest in understanding and describing technology in cooperation with other disciplines. This interdisciplinary communication described technology in primal form. Some of the earliest production studies used agronomic data to estimate fertilizer response functions and optimal fertilization rates [Day (1965)]. Over time, a greater understanding of the science of input interactions has been accumulated to allow further economic insights into basic production problems [Berck and Helfand (1990); Paris (1992)]. As agricultural economics has evolved, dual methods have become prominent because of their simplicity, convenience, and power [Binswanger (1974)]. These methods have been widely applied but the applications typically lack the biological and dynamic detail that often accompanies other optimization or econometric models [Bryant et al. (1993); Woodward (1996); Burt (1993); Foster and Burt (1992)]. As a result, questions arise about whether agricultural production economists are now in a poorer position than earlier to assess plausibility of estimates and add cumulatively to a store of stylized facts regarded by the profession to describe agricultural technology. For example, Mundlak's review (2001) of the early production function literature emphasizes elasticity estimates and portrays the cumulative characterization of both production and supply-demand elasticities from that literature. Though no such similar review is available for recent literature, estimates of simple concepts such as elasticities are *remarkably* disparate even when similar methods (e.g., duality) and data are used [Shumway and Lim (1993), Table 3]. In this state of affairs, one must question whether agricultural production economists are approaching or losing track of the goal of better understanding and measuring behavior.

2.1. Sequential biological stages, temporal allocation, and limited observability

Agriculture in much of the world thrives with little division or specialization of labor [Allen and Lueck (1998)] because of (i) the sequential nature of production stages, (ii) non-overlapping annual growing seasons imposed by weather conditions, (iii) long time lags from application of variable inputs to harvest of finished outputs, (iv) relative unobservability of the state of production during this lag, and (v) moral hazard associated with using hired labor in certain stages of production where monitoring the effect on output is difficult.[1] Typically, a single person or family decides what to produce given the current capital stock and available services, and then applies variable inputs stage-by-stage through sequential production stages to produce the final product. A stage-wise delineation of the production process is possible in many cases because a relatively small number of sequential rather than concurrent operations are required. Such a production structure is typically imposed by the biological nature of agricultural

[1] For example, harvest labor for fruits and vegetables may be easy to monitor when wages are paid at a piece rate for the amount harvested. However, labor required to seed a crop may be harder to monitor because errors in application rates are largely unobserved until much later when crop stands are apparent or final production is realized.

production. By comparison, manufacturing with a small number of sequential rather than concurrent operations is hard to imagine and likely inefficient because assembly lines are precluded.

For some annual non-irrigated crops, few inputs are applied during the five to nine months between the time of planting and harvesting. For other annual crops, inputs such as pesticides may be applied for preventative reasons before or at planting as well as for prescriptive purposes after planting. A simplifying characteristic of crop production is that application of most inputs involves a costly trip over a field. Thus, most inputs cannot be economically applied continuously (irrigation and inputs applied through irrigation water are exceptions), but rather the timing of input applications is a crucial production decision because of weather.

Because input responses are weather-dependent and harvests are seasonal, production and revenue depend on the timing of input applications. Thus, an m-stage technically efficient input-output relationship might be described by the smooth function,

$$y = f\big(f_1(\boldsymbol{x}^1, t_1), \ldots, f_m(\boldsymbol{x}^m, t_m)\big), \tag{1}$$

where $\boldsymbol{x}^i$ is the variable input vector at time t_i and $\boldsymbol{x}^m$ is harvest inputs applied at harvest time t_m. Note that both the quantity of each $\boldsymbol{x}^i$ and the associated time of application t_i are decision variables. In other words, timing as well as quantities are input choices. The chosen harvest date may not correspond to maximum possible production not only due to time preferences and interest rate incentives but because of labor and machinery scheduling problems, weather, and uncertainty of crop maturity. Because of lags, each $\boldsymbol{x}^i$ is relevant to final output, $\partial f/\partial \boldsymbol{x}^i = (\partial f/\partial f_i)(\partial f_i/\partial x^i) \neq 0$.

In one of only a few studies that have treated timing of operations as decision variables, Just and Candler (1985) demonstrate that agricultural production functions tend to be concave in the timing of both planting and harvesting operations so a unique timing exists that is technically efficient. Antle, Capalbo, and Crissman (1994) similarly investigate optimal timing and suggest an efficiency dimension of input timing. Interestingly, optimal timing in the context of the whole farm operation may not be technically efficient when the availability of resources such as labor or machinery services is constrained. That is, available labor and machinery may not be sufficient to harvest all plots at the same time if they should all mature at the same time.[2]

Also unlike manufacturing where the quality of a continuous or intermediate-stage output is observable, the implications of the current state of a crop for final production are highly subjective at each stage of the growing cycle. In most manufacturing processes, the time it takes to create a finished product, $t_m - t_1$, is relatively short. Additionally, intermediate productivity is more observable compared to agriculture, e.g., how far an item has moved on an assembly line or how well an intermediate step of assembly has been accomplished. Thus, continuous monitoring of input productivity and

[2] One could define technical efficiency to include any non-price constraints but this seems at variance with typical technologically based definitions.

making related adjustments at each stage of the production process is more effective. In other words, technical efficiency is best achieved by examining carefully each stage's output as it occurs or by testing to reach conclusions about the technical efficiency of individual production stages.

In contrast, the long delay from input application to observed productivity tends to confound the observed effects of inputs applied in multiple stages of agricultural production processes. As a result, one cannot easily infer from output which stage is inefficient. Moreover, the effects of inputs observed on other farms may not apply because of differing soil and climatic features. The focus of management is thus more on following recommended guidelines, experimentation to adapt recommended guidelines to specific farm or plot circumstances, and monitoring exogenous and uncontrollable inputs such as weather and pests in order to formulate counter measures.

To better represent intraseasonal unobservability, suppose the representation of the production process assuming technical efficiency in the intermediate states of production follows[3]

$$y = f^*\big(f_1(\boldsymbol{x}^1, y_0), \ldots, f_m(\boldsymbol{x}^m, y_{m-1})\big), \tag{2}$$

where the timings of input applications are implicit decision variables suppressed for simplicity. That is, efficient management at stage i involves maximizing the intermediate output, y_i, where the technology set at stage i is represented by $y_i \leqslant f_i(\boldsymbol{x}^i, y_{i-1})$ and y_0 represents initial conditions [Antle and Hatchett (1986)].[4] One way of conceptualizing the difference between agricultural and manufacturing production in this framework is that the intermediate outputs in agriculture, the y_i's, are largely unobservable. In many manufacturing contexts, the separate stage production functions are readily observed, estimated, and applied separately for management purposes. Thus, efficient farming is directed toward learning well the stage technology through acquiring information available from beyond the farm (such as guidelines from technology developers and universities), experimentation, monitoring uncontrollable inputs, and estimating optimal adjustments accordingly.

This recursively separable structure of production whereby inputs $\boldsymbol{x}_i$ in stage t_i are separable from inputs $\boldsymbol{x}^j$ in stage t_j $(j > i)$ has important implications for agricultural production analysis. For example, labor and capital services applied during pre-planting cultivation will be separable from labor and capital services applied to post-planting herbicide application. This property allows experiment station or extension scientists or scientists from input supply firms to make recommendations on specific input choices that are clear and relevant to farmers assuming that the state variable from the previous

[3] For convenience, we use the expression in (2) to represent a production process of the form

$$y = f_n(x_n, f_{n-1}(x_{n-1}, f_{n-2}(x_{n-2}, \ldots, f_1(x_1, y_0)\ldots))).$$

[4] This yields a variant of recursive separability [Blackorby et al. (1978)].

stage is typical (the case of experiment station guidelines) or monitored (the case of professional pesticide applicators).

Timing of operations has been largely ignored in agricultural production economics. Rather, public agricultural production data are recorded on an annual basis. Accordingly, the timing of input applications as well as the intermediate outputs are unobserved. To utilize such data, the firm is typically presumed to solve:

$$y = f^0(\boldsymbol{x}, y_0) = \max_{\{x^i\}} \left\{ f^*\left(f_1\left(\boldsymbol{x}^1, y_0\right), \ldots, f_m\left(\boldsymbol{x}^m, y_{m-1}\right)\right) \,\middle|\, \sum_i \boldsymbol{x}^i = \boldsymbol{x} \right\}. \tag{3}$$

Initial conditions are typically ignored because data are unavailable in which case the estimated technology corresponds to $y = f^0(\boldsymbol{x})$. In this approach, the aggregate input vector $\boldsymbol{x}$ is treated as the decision variable in the related profit maximization problem (possibly some elements of $\boldsymbol{x}$ are treated as fixed or quasi-fixed inputs).

Interestingly, the assumptions implicit in (3) for input aggregation tend to be inadequate as a representation of family farming, the predominant form of agricultural production. The reason is that some inputs such as family labor and fixed-capital service flows present recurring input constraints through the growing season rather than across the entire production season. As a result, the shadow price (or opportunity cost) of resources can vary considerably through the growing season. For example, farm machinery is typically idle or underutilized through much of the year but is used heavily during several weeks. A grain farmer's most expensive piece of equipment may be a combine that is used only 3 or 4 weeks of the year. Tractors may be used to capacity only at planting or cultivation time of the few dominant crops grown on a farm. In spite of low average use rates, farmers find ownership advantageous because all farmers in an area tend to need the same machinery services at the same time due to local climate and soil conditions that tend to dictate crop timing. Capital services may be hired to relax such constraints in some cases, but custom machinery service markets do not operate in many cases because demands are too seasonal. The implication is that available service flows from such equipment are constrained by fixed investments but the shadow prices caused by such constraints may vary widely through a crop season. For example, the shadow price of the service of a combine may be almost comparable to or even higher than custom hiring rates in the peak use season, but yet much lower in a secondary harvesting season where excess capacity is available. These possibilities explain why farmers choose to hold stocks of expensive machinery even though average use is light.

Likewise, family labor may have distinct advantages over hired labor for specific functions because of moral hazard. That is, additional labor may be hired for such needs as harvesting where productivity is easily monitored and rewarded by piece rates, but moral hazard problems may make hired labor a poor alternative for other types of labor needs such as seeding. Indeed, the superiority of using family labor for carrying out certain functions is an important explanation for survival and predominance of the family farm [Allen and Lueck (1998)]. As a result, family labor within the conventional production model (which typically does not consider moral hazard) can be reasonably treated as a recurring constraint through the growing season that is far more limiting at some times than at others. Thus, the shadow price of family labor may vary widely through the growing season. The widely varying nature of implicit prices of farmer-controlled resources across labor periods (stages of production) has been well-recognized in programming models used to represent agricultural technology [McCarl et al. (1977); Kutcher and Scandizzo (1981); Keplinger et al. (1998)].[5]

If the implicit shadow prices of recurring farmer-controlled inputs vary widely from stage to stage, then the implicit formulation in Equation (3) may be inadequate. Mundlak (2001) emphasizes the need for this generalization in his discussion regarding the representation of capital inputs as stocks versus flows. To emphasize this difference, let $\boldsymbol{x}^i$ represent a vector of purchased variable inputs in stage t_i, and let z^i represent a vector of uses of farmer-controlled inputs such as family labor and capital services in stage t_i. Also, let $\boldsymbol{k}$ be a vector of maximum uses or availability of services made possible by the fixed stock of farmer-controlled resources in each stage.[6] Then technology can be represented by

$$\begin{aligned} y &= f^0(\boldsymbol{x}, y_0 \mid \boldsymbol{k}) \\ &= \max_{\{x^i, k^i\}} \left\{ f^*\left(f_1\left(\boldsymbol{x}^1, y_0, z^1\right), \ldots, f_m\left(\boldsymbol{x}^m, y_{m-1}, z^m\right)\right) \,\middle|\, \sum_i \boldsymbol{x}^i = \boldsymbol{x}; z^i \leqslant \boldsymbol{k} \right\}. \end{aligned} \tag{4}$$

This formulation makes clear that varying implicit prices of fixed farmer-controlled inputs is likely. In some periods, the optimal choice may be $z^i = \boldsymbol{k}$ with a high implicit price while in others it is some $z^i < \boldsymbol{k}$ with a zero implicit price.

[5] Mathematical programming models of agricultural decisions have largely given way to econometric models of decisions as indicated by a review of the literature. Several reasons are as follows. First, there is a great desire for statistical inference whereas inference with inequality constraints is a daunting task [Amemiya (1985); Diewert and Wales (1987)]. Second, in traditional practice, programming approaches have typically used subjective and ad hoc approaches to calibrate models, which some regard as falling short of scientific standards. Third, a primary purpose of production economics has become development of aggregate models of behavior with which to undertake policy analysis. Aggregate programming models tend to generate supplies and demands with large and irregular steps that are regarded as implausible. To the extent firm-level heterogeneity can be handled by smooth econometric models, programming models are less useful. However, recent developments in data envelopment analysis and Bayesian applications have spawned greater interest in merging programming and econometric methods [Fried et al. (1993); Chavas and Cox (1988); Paris and Howitt (1998)]. We note also that modern computer technology is rapidly making possible the boot-strapping of statistical properties of programming models with realistic components such as intermittently binding inequality constraints [Vanker (1996)]. For the purposes of this chapter, we consider primarily the econometric approach to empirical work. However, the principles apply to programming models as well and may be ultimately implemented by some merger of programming and econometric methods.

[6] For simplicity, we assume that farmer-controlled resources and thus maximum uses are constant across production stages in the same growing season. If this is not the case, then time subscripts must be added to the limits of use.

These considerations raise questions about how explicitly models must depict the stage-wise production problem and what types of data are needed to do so. For example, if capital service input data are not available by stages, then Equation (4) suggests that capital input data must measure the state of the capital stock (which determines the maximum possible flow of capital services in each stage) rather than the aggregate flow of capital services over the entire growing season. Modeling the stage-wise choices of capital service flows given these stocks may greatly improve understanding of production decisions if data are available for analysis. But if data are unavailable, how can models represent these implicit production choices sufficiently?

2.2. *Flexibility in the output mix and spatial allocation*

In principle, all firms conceptually choose among producing and marketing multiple final outputs because, at least in principle at the capital investment stage, they decide what to produce. However, much of agriculture throughout the world involves actually producing multiple products simultaneously. While measures of diversification are beginning to decline in many areas, particularly in the post-war period in the United States and most notably for livestock firms [White and Irwin (1972)], crop farming remains highly diversified. An important factor in choosing an agricultural output mix is spatial allocation of inputs among plots. This aspect of agricultural production makes agriculture an interesting case for study of scope economies and the effect of scale on scope economies [Chavas (2001)].

Many multiple-product manufacturing settings involve products that are produced in fixed or limited proportions determined by fixed plant and equipment or physical properties of production processes such as chemical reactions. In others, multiple products result from abruptly switching an entire plant from the manufacturing of one product to another (where simultaneous production of several outputs is not feasible or economical). In agriculture, a few production processes lead to related joint products with limited flexibility such as meat in combination with hides or cotton in combination with cottonseed. However, farmers often have great flexibility in switching among annual crops from season to season and in allocating land, machinery services, and family labor among crops in the same season. Flexible capital leads to large elasticities of product transformation (and, hence, large supply elasticities) because farmers can readily change their relative output mix from one crop season to the next. Much of this flexibility occurs because allocated inputs have similar marginal revenue product schedules in the production of several crops.[7] For example, land and land preparation machinery have similar marginal values in production of corn and soybeans in the corn belt or in production of wheat and sorghum in the southern Great Plains. This is why other considerations are sufficient to cause farms to rotate plots of land and diversify production among such crops.

[7] Flexibility also implies that capital has relatively large marginal products in various states of nature as well.

Marshallian joint production is generally presumed to be a reasonable explanation for many economies of scope and the implied optimality of multi-product farms. Baumol, Panzar and Willig (1988) define inputs for such processes as public inputs because they can be costlessly redirected from one industry to another. Clearly some purchased capital such as buildings or tractors may have some of these characteristics when congestion effects are not present. Some aspects of management skill and information have these properties. Clearly weather is a classic public input [Pope (1976)]. However, the timing and nature of demands on private inputs (or public inputs with congestion effects) can also promote diversification.

For example, when several crops compete for the same farmer-controlled resources, constraints on allocation of these resources can play an important role in determining diversification of the product mix. Farmers must generally allocate farmer-controlled resources consisting of land, management ability, machinery services, and family labor among plots of land. Because these inputs must be allocated spatially among plots, and plots are generally planted to distinct crops (or distinct crop mixes in some developing agriculture), these allocations usually amount to allocations among crops as well. Producing multiple outputs, which have different peak input-use seasons according to their varied stages of production, thus provides a way of more fully utilizing farmer-controlled resources and allowing more off-farm labor possibilities. For example, by producing several crops with different growing seasons, or by producing both crops and livestock which have different seasonality requirements, a farmer may be able to use smaller-scale, less expensive machinery and more fully utilize available family labor and management ability than if the entire farm had to be covered with the same operation at one time. Such considerations can be so important that, when coupled with price incentives, they lead to diversification when specialization otherwise occurs [Pope (1976); Pope and Prescott (1980); Baker and McCarl (1982)].

Interestingly, most agricultural production scientists focus on the rate of application of inputs or input services to a particular plot on which a particular crop is grown. For example, extension specialists recommend different rates of fertilizer application for different crops and soil conditions. Pesticides are often regulated with specified application rates per acre under legal licensing requirements. In this context, the representations in (1)–(4) may apply where y is a vector of outputs and $\boldsymbol{x}^i$ is a vector which distinguishes not only type of input but location (plot and thus crop) of application. With non-jointness of all inputs, Equation (4) becomes

$$\begin{aligned} & y^j = f_j^*\big(f_{1j}\big(\boldsymbol{x}^{1j}, y_0^j, z^{1j}\big), \ldots, f_{mj}\big(\boldsymbol{x}^{mj}, y_{m-1}^j, z^{mj}\big)\big), \quad j = 1, \ldots, \\ & \quad \text{subject to} \sum_j z^{ij} \leqslant \boldsymbol{k}, \quad i = 1, \ldots, m, \end{aligned} \tag{5}$$

where $y^j = y_m^j$ is the quantity produced of output j, $\boldsymbol{x}^{ij}$ is a vector of purchased variable inputs allocated to output j in stage i, and z^{ij} is a vector of uses of farmer-controlled resources allocated to output j in stage i. That is, uses of farmer-controlled resources across all production activities must satisfy availability constraints jointly. Note for convenience and to represent availability constraints appropriately in (5), the first subscript of $\boldsymbol{x}$ is assumed to represent a common timing choice across all production activities so that $t_{ij} = t_i$ is the timing choice for operations in stage i of production for all outputs.[8] Also, note more generally that each y^j could represent a vector of outputs with j indexing additively separable production activities in which case (5) does not imply nonjointness of inputs.

The framework of (5) reflects the notion of allocated fixed inputs introduced by Shumway, Pope, and Nash (1984) and investigated by Just, Zilberman, and Hochman (1983), Leathers (1991), and Just et al. (1990). This literature recognizes that inputs such as land are typically measured and must generally be allocated to the production of specific crops (or crop mixes). While the nonjointness assumptions of these papers may be debatable, the need for farmers to allocate at least some purchased variable inputs and at least some farmer-controlled resources among plots is clear.

Public data typically report inputs and outputs for a region but generally do not give allocations of inputs to crops, plots, or production activities as does farm-level accounting and management data. As a result, problems of estimation of multi-output production relationships in agriculture typically have been simplified to eliminate the allocation problem. As in the case where temporal allocation of inputs is unobserved, elimination of spatial allocation variables presumably assumes implicitly that inputs are allocated among plots to achieve efficiency given that inputs have identical prices across plots. Thus, the firm is treated as solving an allocation problem of the form

$$\begin{aligned} y^1 &= f\left(\boldsymbol{x}, k, y^2, y^3, \ldots\right) \\ &= \max\left\{ y^1 \mid y^j = f_j^*\left(f_{1j}\left(\boldsymbol{x}^{1j}, z^{1j}\right), \ldots, f_{mj}\left(\boldsymbol{x}^{mj}, z^{mj}\right)\right),\ j = 1, \ldots; \right. \\ &\qquad \left. \sum_j z^{ij} \leqslant \boldsymbol{k},\ i = 1, \ldots, m;\ \sum_i \sum_j \boldsymbol{x}^{ij} \leqslant \boldsymbol{x} \right\} \end{aligned}$$

in the typical case where initial conditions are ignored. These practices raise questions regarding how explicitly allocation decisions must be represented in production models and how much understanding of the production problem can be gained by representing allocations explicitly. Can greater econometric efficiency be gained thereby? What data are required?

2.3. Fragmented technology adoption and embodied technology

Much has been written about technical change and technology adoption in agriculture [see the reviews by Sunding and Zilberman (2001) and by Evenson (2001)]. Such phenomena explain both the successes and failures of the "green revolution" and explain

[8] This represents no loss in generality because the input vector may be zero for some production activities in some production stages.

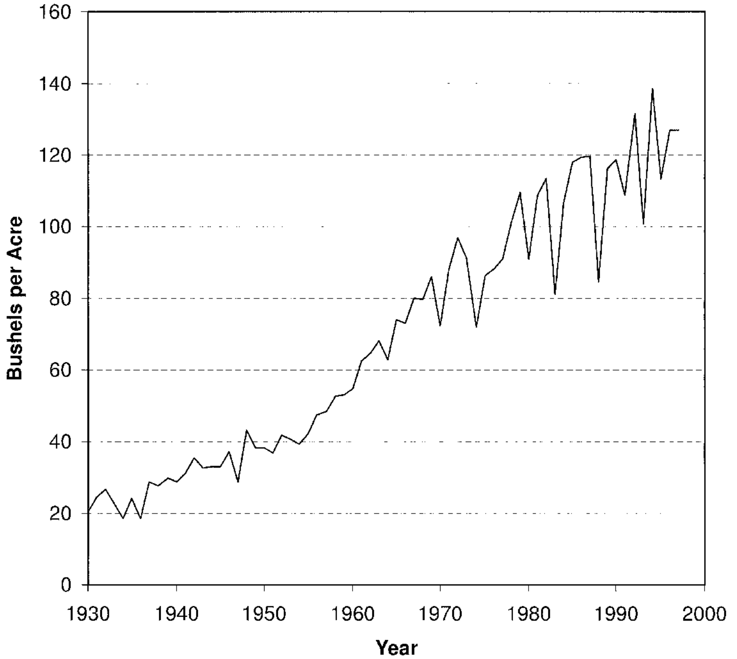

Figure 1. U.S. corn yield growth.

the dramatic growth in agricultural productivity in the twentieth century. More currently, they have much to say about potential agricultural responses to environmental problems, food safety, genetically modified organisms, and the induced innovation that is likely to occur as a result [Sunding and Zilberman (2001); Chavas (2001)]. The dramatic growth in productivity due to technology is illustrated in Figure 1 by the sixfold increase in average U.S. corn yields since 1930. Figure 2 illustrates how much higher the rate of growth in productivity per worker has been in agriculture compared to manufacturing and business.[9] Much of the growth in productivity in developing agriculture has come in the form of higher-yielding seeds, fertilizer use, tube wells for irrigation, and replacement of traditional crops by modern crops. A major explanation in developed agriculture lies in the development of larger-scale machinery, improved crop varieties and livestock breeds, and new inputs such as pesticides and growth hormones. In each case, the technology is embodied in either variable production inputs or in the capital stock.

[9] To construct Figure 2, U.S. Bureau of Labor Statistics indexes for productivity per hour in manufacturing and business are used. The productivity per worker index for agriculture is constructed from U.S. Economic Research Service data by dividing the index of total farm output by the index of farm labor input (see the 1999 Economic Report of the President). All indexes are then adjusted to 100 in 1949. Because the number of hours in the work week in manufacturing and business has been falling, a fair comparison would imply an even greater divergence in growth of output per worker.

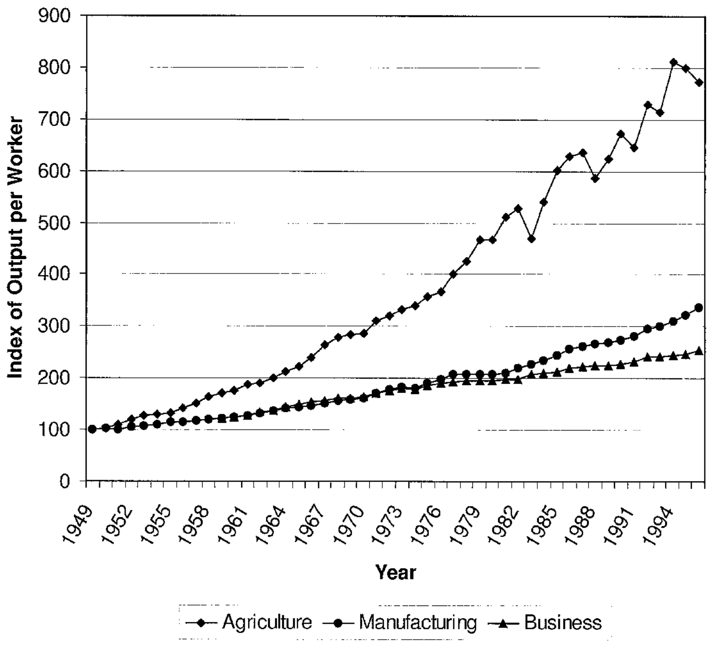

Figure 2. U.S. output per worker.

One of the core features of modern production processes is that production decisions often lag years behind capital decisions. For example, in automobile production, the cycle time from product design to production often takes at least two years. However, once the plant and equipment are in place, the application of inputs typically yields a finished output with very little lag. For example, the typical auto assembly plant produces a car every few minutes and the complete cycle time including pre-assembly of important components is measured in days. For mature industrial processes, this process evolves largely into "quality control".

Some aspects of agricultural production resemble the manufacturing paradigm. For example, producing tree crops and vineyards requires considerable time to put the capital stock in place (e.g., mature trees and vines). Similarly, livestock production involves considerable time to grow mature breeding animals (for gestation, birth, weaning, etc.). These biologically induced lags introduce some interesting and lengthy nonlinear dynamics into the production process [Chavas and Holt (1991, 1993)]. However, a unique aspect of agriculture (compared to manufacturing) is that once the physical capital (perennial crop stands and breeding herds in addition to machinery and buildings) are in place, the lag from the application of variable inputs to the finished output is relatively long.

Another largely unique feature of agricultural production – particularly annual crop production – is that the technology choice is described by a lengthy list of piecemeal decisions that must be made with each new growing season on each plot of land [Sunding and Zilberman (2001); Feder et al. (1985)]. For example, each time a crop is planted a producer can choose to grow a different crop, use a different seed variety, apply fertilizer, use herbicides, apply insecticides, or employ plant growth regulators. A typical grower may choose among 3 to 5 economical crops for the area, each crop may have from 3 to 5 prominent crop varieties with different levels of resistance to unforeseen weather and crop disease conditions, and the farmer may face from 3 to 5 attractive alternative choices each for fertilizers, herbicides, insecticides (if needed), etc. A farmer can choose to use low tillage methods or a variety of tilling operations to control weeds and conserve moisture. Some of these choices are influenced by the stock of equipment (variety and size). The stock of equipment is typically adjusted in piecemeal fashion because most tractors can accept a wide variety of equipment (although the size of equipment is constrained by the size of tractor). The variety of equipment on hand can constrain either the feasible or economical crop set. The size of equipment as well as family labor availability can constrain the amount of land that can be economically allocated to a particular crop/technology combination.

To complicate farmers' choices further, new technology is constantly being developed. New seed varieties and new pesticides are being developed every year and in some cases have dramatic effects on yields.[10] These effects explain the dramatic increase in crop yields as illustrated for corn in Figure 1. A typical problem, however, is that new technologies are unproven and are thus viewed as more risky. Farmers may delay adoption and observe responses obtained by neighbors or allocate small test plots to new technologies. For characterizations of technology to be consistent with such behavior, these uncertainties and options must be represented. Furthermore, technology embodied in machinery or perennial crops is largely fixed by vintage of the capital stock. Some farmers may adopt technologies on a small scale and then increasingly with learning by doing [Foster and Rosenzweig (1995)]. As a result of the complex nature of the technology choice and lags in adoption, a large number of technologies are employed concurrently by different farmers and on different plots by the same farmer [Feder et al. (1985)]. These phenomena complicate drawing inferences from agricultural production data that has been aggregated across heterogeneous farms as discussed further in the section on heterogeneity below. How explicitly does the distribution of technology need to be represented in production models? How much does the distribution of technology depend on the capital distribution? What data on technologies can improve production modeling and how?

2.4. Uncertainty: The role of weather and pests in biological processes

One way agricultural production differs from most manufacturing production is in the magnitude of the impact of uncontrollable factors – many of which are highly stochastic and unpredictable. The dominance of uncertainty in agricultural production is one

[10] The term 'pesticide' is a generic term that includes insecticides, herbicides, fungicides, rodenticides, and crop growth regulators.

reason the study of production under risk has flourished in agricultural economics [Moschini and Hennessy (2001)]. The highly unpredictable nature of agricultural production is illustrated by the yearly national-level corn yields depicted in Figure 1. Furthermore, the data in Figure 1 understate variability because averaging at the national level washes out variation among individual farms. Empirical evidence suggests that variability at the farm level is from 2 to 10 times greater than indicated by aggregate time series data [Just and Weninger (1999)]. The most important uncontrollable factors are weather, pests, and unpredictable biological processes, all of which vary from farm to farm. Weather and pests can cause either localized or widespread crop failures or shortfalls through hail storms, high winds, drought, crop disease, insects, and weed infestations.

Production variability translates also into relatively larger price variability in agriculture as well. The difference in price variability among sectors is highlighted by U.S. producer price indexes at the finished goods and consumer foods level. The variance of annual percentage changes in prices over 1989–1998 was 37.7 for crude consumer foods compared to 4.7 for finished consumer goods, 2.1 for finished capital equipment, 3.8 for processed consumer foods (which represents primarily non-food inputs of processing and packaging), and 5.7 for finished consumer goods excluding foods.

To illustrate the extent of uncontrollable random variation at the state level, the coefficients of variation (CVs) for corn and wheat yields in the United States in Table 1 suggest that farmers on average can have only about a 68 percent probability that production will be in an interval equal to 30 percent of expectations (as implied by normality when CVs average about .15). Furthermore, these coefficients of variation are considerably higher in some states (ranging from .04 to about .25 for both crops). The lower coefficients of variation occur mostly in states where expensive irrigation technology is used to compensate for low and irregular rainfall. Furthermore, it should be noted that much of the variation experienced by individual farmers is washed out by the statewide aggregate data summarized in Table 1 so that the statistics in Table 1 represent a significant underestimate of the effect of uncontrollable factors at the farm level.

Weather and pests are continuous inputs that affect crop growth throughout the entire growing season. Characterizing technically efficient decisions on the basis of ex post random draws of output is difficult because the impact of any vector of inputs $\boldsymbol{x}^i$ on output $(\partial y/\partial \boldsymbol{x}^i)$ is obscured by previous weather occurrences embodied in a largely unobserved state of the crop at time t_i and future weather occurrences embodied in the ultimate observed production, y. Drawing on the well-known literature under uncertainty, $(\boldsymbol{x}_i', G')$ is technically inefficient in an ex ante sense in stage t_i if $G(y_i \mid \boldsymbol{x}^i, y_{i-1}) < G'(y_i \mid \boldsymbol{x}^{i\prime}, y_{i-1})$ for all y_i where G and G' are cumulative distribution functions associated with y_i and $\boldsymbol{x}_i' \geqslant \boldsymbol{x}_i$. This relationship, however, merely represents first-degree stochastic dominance. First-degree stochastic dominance holds for a particular distribution (for a particular input vector) if it yields the largest output for every state of nature. A similar notion can be developed using conventional input distance measures [Färe (1996)]. If first-degree stochastic dominance fails, higher orders of dominance may provide potentially useful comparisons [Eeckhoudt and Gollier (1992)].

Table 1
Coefficients of variation and average yields for U.S. corn and wheat, 1988–97

State	Coefficient of variation		Mean yield	
	Corn	Wheat	Corn	Wheat
Alabama	0.23	0.19	75.2	38.1
Arizona	0.06	0.04	164.5	91.1
Arkansas	0.13	0.22	112.3	43.5
California	0.04	0.06	161.5	77.4
Colorado	0.11	0.10	143.0	33.0
Delaware	0.19	0.15	107.8	55.7
Florida	0.14	0.17	75.4	34.7
Georgia	0.19	0.16	89.6	40.8
Idaho	0.07	0.07	133.5	72.9
Illinois	0.19	0.19	124.3	49.5
Indiana	0.17	0.15	121.9	52.0
Iowa	0.20	0.18	122.3	38.4
Kansas	0.09	0.19	133.7	33.9
Kentucky	0.17	0.19	107.7	49.5
Louisiana	0.13	0.21	106.8	33.9
Maryland	0.22	0.14	104.2	54.6
Michigan	0.14	0.17	106.6	50.3
Minnesota	0.21	0.25	114.6	35.5
Mississippi	0.18	0.24	85.7	37.3
Missouri	0.18	0.16	107.0	43.0
Montana	0.16	0.21	114.7	30.4
Nebraska	0.10	0.12	126.4	34.5
Nevada		0.11		81.7
New Jersey	0.16	0.14	106.2	47.1
New Mexico	0.06	0.22	161.0	27.1
New York	0.10	0.10	101.5	50.7
North Carolina	0.14	0.13	87.7	44.6
North Dakota	0.23	0.25	78.0	29.1
Ohio	0.16	0.13	117.7	53.6
Oklahoma	0.13	0.18	120.1	28.0
Oregon	0.09	0.11	160.9	62.7
Pennsylvania	0.19	0.11	100.6	48.5
South Carolina	0.26	0.17	76.2	41.8
South Dakota	0.21	0.21	81.1	29.3
Tennessee	0.18	0.18	101.2	42.7
Texas	0.12	0.14	112.3	28.3
Utah	0.09	0.14	129.6	41.7
Virginia	0.20	0.13	97.7	54.5
Washington	0.04	0.13	182.0	57.2
West Virginia	0.17	0.09	93.1	47.8
Wisconsin	0.19	0.16	110.9	48.1
Wyoming	0.15	0.15	111.8	27.8
Average	0.15	0.16	111.2	45.8

One of the pressing issues in the measurement of efficiency across firms is that firms may have access to the same technology but may not have access to identical distributions of weather, i.e., identical distributions of outputs given inputs. To denote dependence on local random weather, the production response can be represented by

$$y = f\left(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^m, \boldsymbol{k}, \boldsymbol{\varepsilon}\right), \tag{6}$$

where $\boldsymbol{k}$ represents all relevant capital inputs, $\boldsymbol{\varepsilon}$ is a vector of weather occurrences on a particular plot or farm, and the choice of timing of inputs is suppressed for convenience.

Adding intermediate temporal detail, a more informative representation is

$$y = \left\{ f^*\left(f_1\left(\boldsymbol{x}^1, y_0, z^1, \boldsymbol{\varepsilon}^1\right), \ldots, f_m\left(\boldsymbol{x}^m, y_{m-1}, z^m, \boldsymbol{\varepsilon}^m\right)\right) \mid z^i \leqslant \boldsymbol{k} \right\}, \tag{7}$$

where $\boldsymbol{\varepsilon}^i$ represents local weather events occurring during stage t_i of the production process. The possibility for weather events to cause significant variation in final or stage output is large. Weather can cause certain operations (stages) to be largely ineffective or consume excessive resources unless choices of timing are altered. For example, trying to cultivate a field that is too wet can cause tillage to be ineffective or consume excessive labor. Or trying to plant a crop before adequate rain can result in an inadequate stand of seedlings. The associated consequences for output can be dramatic. For example, delaying planting of corn in Iowa beyond the average optimum of May 1 to May 20 implies more than a 10 percent decline in yields [Burger (1998)]. Weather can also reduce plant growth with excessive heat or inadequate rain or destroy crops at any stage through hail storms.

An important result following from the lags in (7) is that realized output may not be monotonically increasing in input variables. For example, bad weather (pest infestations) can reduce yields while motivating managers to use more labor (pesticides). Thus, a regression of output y_{t_n} on some total input vector $\boldsymbol{x} = \sum_i \boldsymbol{x}^i$ may suggest a negative association for some variables even though $\partial \mathrm{E}_i(y_{t_n})/\partial \boldsymbol{x}^i$ is positive, where E_i is the expectation of y_{t_n} taken at time t_i (using information available at time t_i). This has led some economists to model particular inputs as controlling the damage to normal growth [Feder (1979)].

Considerable early efforts were made to determine the relationship between yields and weather [Doll (1967); McQuigg and Doll (1961); de Janvry (1972)]. These studies try to use weather and ex post measurements of yields to model the conditional distributional of yields given controlled inputs. Voluminous data compiled by the U.S. National Weather Service include hourly temperature, wind, and precipitation data at a large number of weather stations in the United States. Because the data is so voluminous and detailed, suitable aggregator functions are needed but have not been developed. Alternatively, recent work has been content to consider a Taylor's series approximation of (6) at, say, $\mathrm{E}_n(\boldsymbol{\varepsilon}) = 0$ and estimate functions such as $y = \tilde{f}(\boldsymbol{x}, y_0, \boldsymbol{k}) + \tilde{g}(\boldsymbol{x}, y_0, \boldsymbol{k})\boldsymbol{\varepsilon}$ where y_0 is typically not measured. This leads to a function in terms of controllable inputs plus a heteroscedastic error [Just and Pope (1978, 1979); Antle (1983)]. More formally, a first-order Taylor series approximation of (7) is

$$y = \big\{f^*\big(f_1(\boldsymbol{x}^1, y_0, z^1, 0), \ldots, f_m(\boldsymbol{x}^m, y_{m-1}, z^m, 0)\big) + f^*_{\boldsymbol{\varepsilon}}\big(f_1(\boldsymbol{x}^1, y_0, z^1, 0), \ldots, f_m(\boldsymbol{x}^m, y_{m-1}, z^m, 0)\big)\boldsymbol{\varepsilon} \mid z^i \leqslant \boldsymbol{k}\big\}, \tag{8}$$

where subscripts of f^* represent differentiation, transposition is ignored for simplicity, and $\boldsymbol{\varepsilon}$ is a vector composed of $\boldsymbol{\varepsilon}^1, \ldots, \boldsymbol{\varepsilon}^m$. The key marginal effect of $\boldsymbol{x}^i$ on the variance (mean-preserving spread) of y is thus

$$\partial \operatorname{var}(y)/\partial \boldsymbol{x}^i = 2 f^*_{\boldsymbol{\varepsilon}}(\cdot) f^*_{\boldsymbol{\varepsilon}\boldsymbol{x}^i}(\cdot) \mathrm{E}_m\big(\boldsymbol{\varepsilon}^2\big) \tag{9}$$

and has signs determined by elements of $f^*_{\boldsymbol{\varepsilon}}(\cdot) f^*_{\boldsymbol{\varepsilon}\boldsymbol{x}^i}(\cdot)$.

While the structure of (8) appears quite complex for empirical purposes, considerable common structure between the first and second right-hand terms can be exploited for efficiency purposes. For example, the same separable structure is preserved in both the mean and the shock portion of production because it is generated by the same recursive structure of the production stages. Thus, if seeds are separable from labor and machinery in mean wheat production, the same should also be true for the variance. As suggested by Antle's (1983) work, the framework in (8) and (9) can also be expanded in a straightforward way to consider higher moments of the output distribution. The more recent work of Chambers and Quiggin (1998) can also be considered as a generalization of this characterization of production because it characterizes stochastic production by the production set in every state of nature [see Moschini and Hennessy (2001) in this Handbook for a brief explanation].[11]

More importantly, Equations (8) and (9) highlight a central issue in decision making when mean-variance decision models are appropriate; namely, that an input may contribute to the mean differently than it contributes to variance. Indeed, the contributions may be opposite in sign. An input in which (9) is negative (positive) is typically called risk reducing (increasing), following Just and Pope (1978, 1979). Another related possibility is to classify inputs based upon the marginal effect of risk aversion on use [Loehman and Nelson (92)]. A large body of research has developed on risk-reducing marketing, production, and financial strategies. Further examples of empirical research measuring the stochastic characteristics of inputs are found in Love and Buccola (1991); Regev et al. (1997); Horowitz and Lichtenberg (1993); and Nelson and Preckel (1989). However, for the most part, these studies have explored possibilities on a piecemeal basis and have not produced a coherent and widely used framework for agricultural production analysis. Many questions remain. How explicitly do stochastic elements of production have to be represented? Does the source of stochasticity make a difference? How can the micro-level stochastic production problem be represented adequately with available data?

[11] Assuming technical efficiency, characterizing all the moments of output is equivalent to characterizing efficient production in every state of nature because of the uniqueness of moment-generating functions.

2.5. Interseasonal complexity of biological and physical processes

A host of longer-run (inter-year) issues also complicate matters [Nerlove and Bessler (2001)]. Like manufacturing, these involve evolutions of the capital stock represented in k from one production period to the next and how these affect technology. The state of the capital stock is affected by how heavily it has been used in previous periods (which determines the likelihood of time-consuming breakdowns and costly repairs) as well as by net investment. However, an important consideration in agriculture is that initial crop-year soil conditions and pest infestations/resistances and perennial crop states in y_0 are dependent on previous cropping choices, fertilizer and pesticide applications, and soil tillage. The state of the machinery capital stock may be largely observable through inventory records and by inspecting wear, while the state of soil and pest conditions is largely unobservable except through extensive (and in some cases impractical) testing.[12]

In this context, both y_0 and $\boldsymbol{k}$ are affected by decisions in earlier growing seasons. Regarding t now as spanning growing seasons, output is Markovian through both y_0 and $\boldsymbol{k}$. This phenomenon is manifest by crop rotation practices where weed or insect cycles are broken by switching a given plot among crops on a regular basis. The need for such rotation is typically realized on the basis of previously observed infestations that occur otherwise, rather than direct indications of carry-over soil or pest conditions. Rotation actions are often undertaken on a preventative basis because once a serious problem occurs it affects an entire growing season before corrections can be made. Thus, a careful delineation of intertemporal production possibilities implies consideration of inter- and intra-year effects. Implied models contain non-linear dynamics with accompanying instability [Chavas and Holt (1991, 1993)].

The static or short-run generic description of technology in Sections 2.1–2.4 is consistent with this depiction of inter-cycle production because the choices made in a previous growing season are fixed in the current growing season. However, this simplification in theoretical modeling does not imply that initial conditions in y_0 can be ignored in empirical work as most production studies have done. Empirical work documenting the importance of inter-cycle production phenomena through carry-over conditions has been limited. See, e.g., Chambers and Lichtenberg (1995) for a rare study of interseasonal investment in soil capital.

The forward impacts of input choice are essential to many agricultural economic problems. For some inputs a positive future effect is clear, $\partial y_{i+j}/\partial \boldsymbol{x}^i > 0,\ j > 0$, while for others it is negative. For example, fertilizer has both initial and future positive effects (ignoring externalities) due to the carryover of nutrients in the soil [Woodward (1996)]. However, many pesticides have negative dynamic effects by inducing pest resistance [Hueth and Regev (1974); Clark and Carlson (1990)]. In addition, interpreting

[12] Often limited spot testing of soil is used as a basis for prescribing fertilizer needs but the results give only a crude estimate of the inventory of soil conditions. The extent of weed seed carry-over and gestating insect infestations are impractical to assess.

the y_i's as outputs in a given year, it is clear that nitrogen fixation of legumes and other crop rotational issues have positive marginal dynamic effects on future outputs. Seemingly, micro-studies of crop choice must consider these effects if they influence observed farmer choices. Finally, capital decisions have important interyear effects [Vasavada and Chambers (1986); Vasavada and Ball (1988); Morrison and Siegel (1998)].

While many advances have been made in conceptual representation of these interseasonal issues of production, many questions are not well understood. Data for investigating these issues empirically has been lacking, particularly for crop production, and accordingly few empirical studies have been undertaken. While livestock production has been examined with more dynamic detail, models have been conceptually less elegant and, thus, of less interest. Accordingly, little is known about interseasonal behavioral preferences, particularly where risk issues are important.

2.6. *Atomistic heterogeneity and aggregation*

While the concepts of production theory generally are developed at the individual firm level, much of the empirical work in agricultural production is done at the aggregate level of a state or nation. Use of aggregate data has occurred because few firm-level data sets have been developed and access to them is limited or conditional.[13] Thus, the discussion of agricultural production analysis cannot be complete without considering the problem of aggregation.

Agriculture is atomistic with respect to most products. That is, the number of firms is large and each is individually unimportant at aggregate levels. Nevertheless, farms differ in many ways. The wide distribution of technology employed simultaneously across farms suggests one dimension of this problem. Another dimension is the wide variation in climate and soil quality across farms. Differences in soil quality have been highlighted historically by U.S. Natural Resources Conservation Service (formerly U.S. Soil Conservation Service) classifications of soil and land characteristics but are increasingly highlighted by precision farming techniques, localized incentives for environmental preservation and conservation practices, and location-specific environmental policy. The implications of variation in climate and soil for crop production and variability are depicted by the variation of both average yields and coefficients of variation of major crops among states in the United States as illustrated in Table 1. These variations explain much of the dramatic difference in crop mixes chosen by farmers from one location to another.

[13] There are exceptions such as the Agricultural Resource Management Study (formerly the Farm Cost and Returns Survey) data compiled by the U.S. Economic Research Service and the Kansas State University Farm Management Survey data in developed agriculture, and the ICRISAT Household Survey data and various other World Bank surveys in developing agriculture. However, access to such farm-level surveys tends to be limited to those with in-house affiliations, willingness to analyze the data in-house, or willingness to collaborate, and thus such data has been explored only to a limited extent.

Another dimension of heterogeneity imposed by geographic differences in climate and soil quality is the heterogeneity in prices induced thereby. As a result, land rents, the opportunity cost of labor and the price of services do not follow the law of one price. A considerable amount of output price variation also occurs due to differences in climate-induced product quality and timing of production [Pope and Chambers (1989); Chambers and Pope (1991)].

Though many inputs have similar marginal products across farms as well as space and enterprises, others such as chemicals, purchased services, and some machinery are highly and increasingly specialized. Some pesticides have primarily pre-emergent uses and others have primarily post-emergent uses. Most pesticides are used primarily on only a few crops and thus differ across farms. Aerial spraying equipment is primarily used for post-emergent applications while ground operations are primarily used for pre-emergent application. Aggregating such pesticide uses or machinery services over farms as well as time and space can be problematic. As technology turns more toward genetically engineered seeds, such as Roundup-ready soybeans which introduce dependence on specific pesticides, the allocation of a given total input quantity to enterprises to achieve technical efficiency may be trivial on individual farms but underrepresented by aggregates.

To represent heterogeneity, let $G(\boldsymbol{\varepsilon}, \boldsymbol{k} \mid \theta)$ represent the distribution across all farms of characteristics such as weather and pests, capital and technology, management ability, and other policy and input constraints imposed on farms by external conditions. Then following the representation in (7), aggregate production response is described by

$$y = \int \left\{ f^*\left(f_1\left(\boldsymbol{x}^1, y_0, z^1, \boldsymbol{\varepsilon}^1\right), \ldots, f_m\left(\boldsymbol{x}^m, y_{m-1}, z^m, \boldsymbol{\varepsilon}^m\right)\right) \mid z^i \leqslant \boldsymbol{k} \right\} \mathrm{d}G(\boldsymbol{\varepsilon}, \boldsymbol{k} \mid \theta). \tag{10}$$

In this framework, standard regularity conditions fail at the aggregate level even under profit maximization but distribution-sensitive aggregation such as in (10) can preserve practical versions of regularity conditions [Just and Pope (1999)]. These results raise questions about how specifically and explicitly stochastic sources of variation must be represented in production models. Of course, related considerations of heterogeneity in prices, expectations, and risk preferences are necessary to derive supplies and demands from representations of the production technology.

2.7. *Implications of the unique features of agricultural production*

The discussion in Sections 2.1–2.6 emphasizes a number of unique features of agricultural production that require specific attention. Time lags and stages imposed on the production process by biological characteristics of production suggest that one should appropriately represent the allocation of inputs over time within a crop season. Flexibility of crop mixes and specificity of inputs by crop (or location) highlight the importance of representing allocations over these dimensions. The role of farmer-owned resources such as land and labor, and their allocation, may imply significant economic constraints that, in turn, complicate the aggregation of capital service flows and family labor over time. How appropriate is production modeling when based only on data aggregated across these dimensions? How limited are the sets of issues that can be investigated?

Though empirical economic practices must of necessity work with aggregates at some level, we believe that agricultural economists have often been cavalier about temporal and spatial (biological) structure and heterogeneity in agriculture. This has led to inappropriate grouping of inputs and outputs over space and time. Spatial dimensions of input groupings may be particularly important in agriculture because inputs must be tailored to the heterogeneity of farm resources, which differ substantially by climate and land quality (location). For example, ignoring these circumstances may lead economists to conclude that too much land is used by a large farm with heterogeneous land quality in comparison to a "best practice" when in reality economists are not using "best practice" methods of aggregation. Such practices have implications for measuring technical or other inefficiencies as well as for measuring behavior. Similarly, time dimensions of input groupings may be more important in agriculture because production lags tend to be longer and thus encounter more price heterogeneity.

At a minimum, the conclusions that are being drawn must be carefully and fully qualified given these possibilities. One purpose of this chapter is to identify the extent of needed qualifications. In some cases, existing data allows more careful consideration of aggregation issues. For example, inventories of land qualities and of land allocations can be used to enhance economic understanding. However, data are rarely collected on intraseasonal input choices nor is it generally reported on spatial allocations. Thus, for issues that require data on intraseasonal or spatial choice, limitations of current data imply that there is a clear tension between the description of agriculture in Sections 2.1–2.6 and available data. Perhaps more importantly, these issues raise concern about whether approaches that require the specification of technology, either explicitly or implicitly, can correctly reflect technology when temporal and spatial aggregates are used.

The next few sections of this paper investigate conceptual and theoretical implications of the issues raised thus far regarding temporal (Section 2.1) and spatial (Section 2.2) allocation of inputs, the potential differences among generic inputs represented by embodied technology (Section 2.3), and the stochastic nature of production (Section 2.4). The nature of constraints imposed on service flows by farmer-owned resources is considered explicitly. A set of principles is developed in Section 3 that address these issues in agricultural production analysis. Then fundamental theoretical results are developed to apply these principles in Section 4. In Sections 3 and 4, we provide a critique of current approaches to agricultural production analysis, identify the limitations imposed by data availability, and suggest appropriate qualifications that must be attached to agricultural production studies given data limitations. In some cases, these qualifications invalidate many empirical findings to date. We suggest this is one reason for the poor performance of duality models noted by Mundlak (2001) and that some of these problems arise from trying to apply the concepts of production economics without sufficient attention to the unique features of agriculture discussed in this section.

Following Sections 3 and 4, we address in Section 5 the extent of generalizations that have been achieved in agricultural production analysis regarding the other unique features of agriculture identified above (Sections 2.5–2.6) and related issues. Relatively less emphasis is placed on these issues in this chapter because they are emphasized heavily by Sunding and Zilberman (2001) and Nerlove and Bessler (2001). However, we suggest implications of the principles of this paper that are applicable and which call for more structural and detailed analysis, empirical investigation in the context of a broader maintained model, and more adequate representation of heterogeneity as data allows.

3. Principles of agricultural technology representation

Before proceeding to consider appropriate principles for agricultural production analysis, introduction of some conventional concepts and definitions is useful to facilitate discussion. Following the seminal work of Nobel laureate Gerard Debreu, we define an economic good not only physically but also temporally and spatially [Debreu (1959, pp. 28–32)]. In other words, date and location in addition to physical identification of a commodity are essential. Debreu emphasizes that this distinction "should always be kept in mind" in his comprehensive mathematical representation of economic phenomena (p. 32). Thus, fertilizer applied to a particular wheat field at planting time is considered distinct from post-emergent fertilizer applied to the same wheat field at a later date and from fertilizer applied in planting a barley field or another wheat field even if on the same date and farm. Debreu also emphasizes with a long list of examples that the physical identification of goods must be complete. As an example, he emphasizes that land must be described completely by the nature of the soil and subsoil characteristics, crop residues, etc. (p. 30). These considerations have important implications for the analysis of agricultural production because of the unique features of agricultural production involving long time lags in the production process, wide variation in prices and local weather conditions, and great heterogeneity both among and within farms.

In contrast to Debreu's clear conceptual definitions, we note that carrying the distinction of space and time and even many of the attributes of physical identification has been largely dropped from empirical agricultural production studies. For example, it is not unusual to represent output as a single aggregate commodity, a two-dimensional measure with crop and livestock aggregates, or a short vector consisting of the aggregate production of several crops. Inputs often consist of four to six aggregate annual input quantities such as land, fertilizer, pesticides, seeds, labor, and machinery services. The role of weather and pests is usually swept under the guise of an ad hoc homoscedastic error term. Examples of such studies include many widely referenced studies of U.S. agriculture over the past two decades and virtually all of the production studies referenced in the survey by Shumway (1995). Thus, the specifications in (1)–(10) are rarely employed empirically in agriculture.

For better or worse, agricultural production studies have increasingly relied upon readily available data to illustrate advances in technical methods of analysis. Readily available data, however, tends to consist of highly aggregated public data in which temporal detail (within a growing season) and spatial detail (among plots, farms, or land used to produce specific crops) is lost. As economists have become more heavily focused on policy-related applications, the aggregate level of analysis has taken on elevated importance. By comparison, as economists have focused relatively less on supporting the management of firms, emphasis on analyses of individual firm data has decreased. In consequence, as production economists have become more focused on aggregate responses and aggregate productivity, technology has tended to be described by a production possibilities frontier (PPF) where both inputs and outputs have been stripped of their temporal and spatial dimensions.

Our purpose is to consider the validity and implications of such practices and illustrate the need for development of data sets that allow the underlying hypotheses to be tested. Results demonstrate that the failure of modern agricultural production analysis, and of typical implementations of duality theory in particular, may be due to such practices. We note also, however, that potentially necessary generalizations can be made in a variety of frameworks including dual models, but such generalizations are likely to require better firm-level data than has hitherto been available.

3.1. A general framework for production analysis

To facilitate formal analysis, we define a general notation that applies through the remainder of this chapter. Following the distinctions emphasized by Debreu, available technology for production is represented by $\boldsymbol{W} \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})$ where

$\boldsymbol{W} \in R^n$ is a netput vector of inputs (if negative) and outputs (if positive);

$\Im$ is the feasible short-run production set;

$\boldsymbol{k} \in R_+^{n_k}$ is a vector of firm-controlled resources;

$\boldsymbol{\varepsilon} \in R_+^{n_\varepsilon}$ is a vector of uncontrolled inputs that describes the state of nature;

both $\boldsymbol{W}$ and $\boldsymbol{k}$ are defined with temporal, spatial, and physical detail; firm-controlled resources in $\boldsymbol{k}$ include family attributes and capital (land, buildings, and machinery) that determine recurring availability of family labor and capital service flows with temporal, spatial, and physical detail; and uncontrolled inputs include weather and pest infestations, also with temporal, spatial, and physical detail. Revisions in the capital stock through investment may change the amount of recurring capital service flows, for example, following a putty-clay framework, but need not be considered in the short-run case of Sections 3 and 4. Because firm-controlled resources are available in fixed amounts depending on capital stocks and family composition, the service flows from them must be allocated either temporally or spatially across competing production activities. These are called allocated fixed factors following the terminology of Shumway, Pope, and Nash (1984).

Typically, the netput vector is partitioned into inputs and outputs. Following Section 2.1, we further distinguish purchased variable inputs from allocated fixed factors so that $\boldsymbol{W} = (\boldsymbol{Y}, -\boldsymbol{X}, -\boldsymbol{Z})$ where

$\boldsymbol{Y} \in R_+^{n_y}$ is a vector of output quantities,
$\boldsymbol{X} \in R_+^{n_x}$ is a vector of purchased variable input quantities, and
$\boldsymbol{Z} \in R_+^{n_z}$ is a vector of allocations of allocated fixed factors,

where $n = n_y + n_x + n_z$. Thus, allocations of recurring capital service flows and family labor appear in $\boldsymbol{Z}$. Netputs if negative represent net inputs which impose costs or deplete firm-controlled resources, and if positive represent net outputs which generate revenue. Thus, the partitioning of netputs into inputs and outputs is regarded as a local convenience and may not apply globally. For example, some goods may be purchased as inputs in some circumstances and produced as outputs in others, and some capital services may be provided fully by firm-owned machinery in some circumstances and purchased by means of commercial contracting in others.

We argue that temporal, spatial, and physical distinctions should not be dropped until doing so can be demonstrated to be appropriate. The purpose of Section 3 is to demonstrate some of the problems encountered by doing so while Section 4 presents a more general disaggregated analysis. Interestingly, temporal and spatial distinctions are commonplace in the agricultural marketing literature although largely ignored in the agricultural production literature. To investigate the importance of temporal and spatial detail, the input and output vectors require further partitioning following

$$\boldsymbol{Y} \equiv \{\boldsymbol{y}^1, \dots, \boldsymbol{y}^m\}, \quad \boldsymbol{y} \equiv \sum_{i=1}^{m} \boldsymbol{y}^i,$$
$$\boldsymbol{X} \equiv \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^m\}, \quad \boldsymbol{x} \equiv \sum_{i=1}^{m} \boldsymbol{x}^i,$$
$$\boldsymbol{Z} \equiv \{\boldsymbol{z}^1, \dots, \boldsymbol{z}^m\}, \quad \boldsymbol{z} \equiv \sum_{i=1}^{m} \boldsymbol{z}^i,$$

where $i = 1, \dots, m$ indexes time and/or location and $\boldsymbol{y}$, $\boldsymbol{x}$, and $\boldsymbol{z}$ represent aggregate vectors of outputs, purchased variable inputs, and allocated fixed factors, respectively, which include only physical detail. For example, each $\boldsymbol{x}^i$ may specify quantities of specific seeds, fertilizers, and pesticides to apply as inputs at a time and/or location indexed by i and $\boldsymbol{y}^i$ may give quantities of specific types of outputs that occur at a time and/or location indexed by i. Interesting questions arise in considering aggregation over time and/or space. For example, one may consider when the feasible production set can be adequately represented by $\boldsymbol{w} \equiv (\boldsymbol{y}, -\boldsymbol{x}, -\boldsymbol{z}) \in \Im_{-m}$ where $\Im_{-m}$ represents a feasibility set devoid of temporal and/or spatial detail.

For purposes of facilitating discussion of practical implications of technical efficiency, corresponding price vectors are also defined. Let $\boldsymbol{P}$ be a price vector corresponding to output vector $\boldsymbol{Y}$ and let $\boldsymbol{R}$ be a price vector corresponding to input vector $\boldsymbol{X}$. Then short-run profits can be represented by $\pi = \boldsymbol{PY} - \boldsymbol{RX}$. Suppose also that price vectors are partitioned temporally and/or spatially as in the case of $\boldsymbol{Y}$ and $\boldsymbol{X}$ so that short-run profits are equivalently expressed as $\pi = \sum_{i=1}^{m}(\boldsymbol{p}^i \boldsymbol{y}^i - \boldsymbol{r}^i \boldsymbol{x}^i)$. Finally, if and only if $\boldsymbol{p} = \boldsymbol{p}^i$ and $\boldsymbol{r} = \boldsymbol{r}^i$ for $i = 1, \ldots, m$ can profits be generally expressed with temporal and/or spatial aggregation as $\pi = \boldsymbol{py} - \boldsymbol{rx}$.

Agricultural economics has long-standing traditions of pursuing production analysis using both set theoretic models (often represented by normative mathematical programming models) and smooth econometric representations of either average or frontier technologies. For example, the production set is commonly represented by the transformation function F where $\Im = \{\boldsymbol{W} \mid F(\boldsymbol{W}) \leqslant 0\}$ in the general netput case or $\Im = \{(\boldsymbol{Y}, -\boldsymbol{X}, -\boldsymbol{Z}) \mid \boldsymbol{Y} \leqslant f(\boldsymbol{X}, \boldsymbol{Z})\}$ in the partitioned case with explicit inputs and outputs.[14] For smooth econometric representations, equality in the transformation relationship defines the boundary or frontier of the production set, i.e., $F(\boldsymbol{W}) = 0$ or $\boldsymbol{Y} = f(\boldsymbol{X}, \boldsymbol{Z})$ are the boundaries of production sets defined by $F(\boldsymbol{W}) \leqslant 0$ or $\boldsymbol{Y} \leqslant f(\boldsymbol{X}, \boldsymbol{Z})$, respectively. Representations of average technologies follow $\boldsymbol{Y} = f(\boldsymbol{X}, \boldsymbol{Z}) + \boldsymbol{\varepsilon}$ where $\boldsymbol{\varepsilon}$ represents random or uncontrolled inputs and $\mathrm{E}(\boldsymbol{\varepsilon}) = 0$ is used to represent average efficiency conditions. Such models are popular in general agricultural production problems. Alternatively, one-sided error term models such as $\boldsymbol{Y} = f(\boldsymbol{X}, \boldsymbol{Z}) + \boldsymbol{\varepsilon}$ where $\boldsymbol{\varepsilon} \leqslant 0$ have been used where efficiency is of primary interest, in which case all deviations denote random deviations from the case of efficient production, $\boldsymbol{Y} = f(\boldsymbol{X}, \boldsymbol{Z})$ [see Fried et al. (1993)].[15] In these various models, the technology is described by measures such as production elasticities, scale and scope economies, factor substitution, productivity and

[14] In cases with only one output where F is a scalar function, the existence of the function f follows from continuity of F by the implicit function theorem based on the classification of $\boldsymbol{W}$ into $\boldsymbol{Y}$, $\boldsymbol{X}$, and $\boldsymbol{Z}$. As argued below, however, in some cases consideration of a vector-valued f function is appropriate for representing multiple outputs.

[15] Of the smaller programming-based literature in agricultural production economics, data envelopment analysis (DEA) is becoming prominent. The strength of DEA is that it imposes less structure on the form of production and follows the language of leading graduate micro-theory texts [e.g., Varian (1992)]. Its current weakness seems to be that it is not easily used to address the breadth of questions usually considered by agricultural production economists and it generally assumes that all variation in observed production relationships is due to technical inefficiency rather than random, unmeasured, or uncontrollable factors (the same assumption typically applies also to econometric models with one-sided error terms).

Until recently, another weakness was that procedures for statistical inference were not available [Vanker (1996)]. While statistical inference in data envelopment analysis (DEA) models is not yet fully developed, DEA is likely to become fully integrated with econometric methods eventually. In this chapter, our discussion tends to focus on smooth functional representations of production, possibly including cases such as Leontief fixed-proportions production in practice. Thus, we build upon the simple generic representation above to draw implications directly for typical econometric practices. However, we note that the principles and results implied by the unique features of agriculture have applicability for other approaches to modeling agricultural production.

technical change bias, distance functions, separability, and (non)jointness, which we hereafter call "standard characteristics".[16]

3.2. Traditional concepts of efficiency

Traditionally, the frontier of the production set has been used as a representation of technical efficiency. Samuelson (1967), for example, defined aggregate technology by the PPF denoted by $F(\boldsymbol{y}, \boldsymbol{x}) = 0$ where $\boldsymbol{y}$ and $\boldsymbol{x}$ are aggregate vectors of outputs and inputs, respectively (the service flow vector, $\boldsymbol{z}$, is temporarily dropped for convenience and congruence). Samuelson's PPF was determined by various separate technologies for each individual output denoted by $y_i = f_i(\boldsymbol{x}^i)$ where y_i is an element of $\boldsymbol{y}$, and $\boldsymbol{x}^i$ is a nonnegative vector of factor allocations to production activity i. The PPF is thus a smooth function determined by aggregation of individual production functions [Samuelson (1967, pp. 230–231)], e.g.,

$$
\begin{aligned}
&F(\boldsymbol{y}, \boldsymbol{x}) \equiv y_1 - f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}), \\
&f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}) \equiv \max\left\{ y_1 \;\middle|\; y_i = f_i(\boldsymbol{x}^i),\ i = 1, \ldots, n_y;\ x = \sum_{i=1}^{n_y} \boldsymbol{x}^i \right\},
\end{aligned}
\tag{11}
$$

where $\boldsymbol{x}$ is the aggregate input vector that determines the PPF, and $\boldsymbol{y}_{-1}$ is a vector containing all elements of $\boldsymbol{y}$ other than y_1, i.e., $\boldsymbol{y} = (y_1, \boldsymbol{y}_{-1})$. Thus, Samuelson did not maintain the industry (or production activity) distinction when referring to F in (11). For example, where i indexes production activities by time or location, the technology representation in (11) does not retain temporal or spatial detail, respectively.

Following Samuelson, the early literature on production efficiency was based largely on a comparison of actual production to the PPF [e.g., Farrell (1957)]. These boundary points may or may not be technically efficient. For example, the efficient set is at best a subset of the boundary consisting of all $\boldsymbol{W} \in \Im$ such that there is no distinct $\boldsymbol{W}' \in \Im$

[16] Typical assumptions employed to make these technology representations meaningful include (suppressing the arguments of $\Im$ for convenience): (1) $\Im$ is nonempty, i.e., there exists at least one $\boldsymbol{W} \in \Im$; (2) $\Im$ is convex in $\boldsymbol{W}$, i.e., if $\boldsymbol{W}, \boldsymbol{W}' \in \Im$ then $\lambda \boldsymbol{W} + (1-\lambda)\boldsymbol{W}' \in \Im$ where $\lambda \in [0, 1]$; (3) $\Im$ is closed, i.e., if $\boldsymbol{W}^k \in \Im$ such that $\boldsymbol{W}^k \to \boldsymbol{W}$ then $\boldsymbol{W} \in \Im$; (4) inaction, i.e., $\boldsymbol{0} \in \Im$; (5) free disposal or monotonicity, i.e., $\boldsymbol{W} - R_+^n \subset \Im$ if $\boldsymbol{W} \in \Im$, which implies that if $\boldsymbol{X}$ can produce $\boldsymbol{Y}$ then $\boldsymbol{X}' \geqslant \boldsymbol{X}$ can produce at least $\boldsymbol{Y}$; (6) additivity, i.e, if $\boldsymbol{W}, \boldsymbol{W}' \in \Im$ then $\boldsymbol{W} + \boldsymbol{W}' \in \Im$; and some subset of (7a) nonincreasing returns to scale, i.e., if $\boldsymbol{W} \in \Im$ then $\lambda \boldsymbol{W} \in \Im$ where $\lambda \in [0, 1]$; (7b) nondecreasing returns to scale, i.e., if $\boldsymbol{W} \in \Im$ then $\lambda \boldsymbol{W} \in \Im$ where $\lambda \geqslant 1$; and (7c) constant returns to scale, i.e., if $\boldsymbol{W} \in \Im$ then $\lambda \boldsymbol{W} \in \Im$ where $\lambda \geqslant 0$. As an example of a subset of the latter three, decreasing returns to scale occurs when (7a) but not (7c) holds. It is important to note that various combinations of these assumptions have distinct implications. For example, additivity plus constant returns to scale implies that $\Im$ is a convex cone. Also, adding setup costs to any one of the other properties may yield a very different property. For example, setup costs with inaction introduces nonconvexities into $\boldsymbol{W}$. Note also that items (6) and (7) can be rewritten using a production transformation function as: (6′) if $F(\boldsymbol{W}) \leqslant 0$ and $F(\boldsymbol{W}') \leqslant 0$ then $F(\boldsymbol{W} + \boldsymbol{W}') \leqslant 0$; (7a′) if $F(\boldsymbol{W}) \leqslant 0$ and $\lambda \in [0, 1]$ then $F(\lambda \boldsymbol{W}) \leqslant 0$; (7b′) if $F(\boldsymbol{W}) \leqslant 0$ and $\lambda \geqslant 1$ then $F(\lambda \boldsymbol{W}) \leqslant 0$; and (7c′) if $F(\boldsymbol{W}) \leqslant 0$ and $\lambda \geqslant 0$ then $F(\lambda \boldsymbol{W}) \leqslant 0$.

with $\boldsymbol{W}' \geqslant \boldsymbol{W}$. This conceptualization of production has carried through to the modern production literature with slight generalization and now seems to permeate most empirical production analyses of both the data envelopment analysis and conventional econometric approaches. That is, with these implicit Samuelsonian foundations, the modern production literature has evolved toward representation of production in terms of aggregates of allocations of inputs over time and location. For example, a typical empirical model in agricultural production has aggregate output (either for a commodity group or for total agricultural output) depending on aggregate annual inputs such as fertilizer and pesticides rather than allocations of those inputs to specific locations (e.g., to land in specific crops) and stages of production [e.g., Shumway (1983)].

For example, Diewert (1974) gives a commonly cited definition of the PPF, sometimes called the transformation function, as $y_1 \equiv f^*(y_{-1}, x)$ where (when it exists)

$$f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}) \equiv \max\{y_1 \mid (y_1, \boldsymbol{y}_{-1}, \boldsymbol{x}) \in \Im\}. \tag{12}$$

The dual input distance function is

$$D_I(\boldsymbol{y}, \boldsymbol{x}) \equiv \max\{\alpha \mid \boldsymbol{x}/\alpha \in \upsilon(\boldsymbol{y})\},$$

where $\upsilon(\boldsymbol{y})$ is the set of all inputs $\boldsymbol{x}$ that will produce at least $\boldsymbol{y}$, e.g.,

$$\upsilon(\boldsymbol{y}) = \{\boldsymbol{x} \mid f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}) \geqslant y_1\} = \{\boldsymbol{x} \mid (\boldsymbol{y}, \boldsymbol{x}) \in \Im\}.$$

A corresponding dual representation of the PPF is thus given by $D_I(\boldsymbol{y}, \boldsymbol{x}) = 1$. Because Diewert explicitly used the language that f^* and D_I characterize the production possibilities set and because he immediately applied f^* and D_I to international trade, it seems clear that he considered the components of $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\Im$ as neither spatially nor industry specific following the tradition of Samuelson.[17]

The elimination of spatial and temporal distinction is especially apparent in the celebrated paper by Lau (1978). Lau's functional representation of efficient production is based on the input requirement function. Suppose that a primary input such as labor is denoted separately by λ and that the associated input requirement function is $\lambda = \omega^*(\boldsymbol{y}, \boldsymbol{x}_{-1})$ where $\boldsymbol{x}_{-1}$ represents all other inputs such that the full input vector is $\boldsymbol{x} = (\lambda, \boldsymbol{x}_{-1})$. Clearly, λ and $\boldsymbol{x}_{-1}$ represent total use of the respective inputs because Lau's definition of nonjointness states that production is nonjoint in inputs if there exist individual production functions, $\lambda_i = \omega_i^*(y_i, \boldsymbol{x}_{-1}^i)$, such that

$$\omega^*(\boldsymbol{y}, \boldsymbol{x}_{-1}) = \min\left\{\sum_i \omega_i^*(y_i, \boldsymbol{x}_{-1}^i) \;\Big|\; \boldsymbol{x}_{-1} = \sum_i \boldsymbol{x}_{-1}^i\right\}, \tag{13}$$

[17] Of course, Diewert's (1982) notation could be used to define a PPF in higher dimensions that include both spatial and temporal distinctions. However, this is not what he did nor has this been the practice in modern production applications to date. Our purpose is to explore the implications of following one practice or the other.

where $x_i = (\lambda_i, x^i_{-1})$. Thus, each element of x is clearly a sum over industry (or individual production activity) uses.

Lau's definition suggests another way to obtain the PPF from an underlying optimization model under nonjointness – that is, by minimizing the use of one input subject to technology constraints for all outputs and endowment constraints for all other inputs. Clearly, when ω^* and ω_i^* are continuous and monotonic in y and y_i, the relationships $\lambda = \nu^*(y, x_{-1})$ and $\lambda_i = \nu_i^*(y_i, x^i_{-1})$ are equivalent to $F(y, x) = 0$ and $y_1 = f^*(y_{-1}, x)$ in (11), respectively, by the implicit function theorem. Hence, societal (or firm) level efficiency is characterized equivalently by an input requirement function ω^* defined over total uses of other inputs. Again, spatial, temporal, and physical detail in x is omitted in the definition ω^*.

The concept of efficiency to this point has been discussed without regard to prices of either inputs or outputs. A central tenet of this section, however, is that prices are critically important if the leading concepts of production efficiency are to have practical meaning. Otherwise, the efficiency concepts that correspond to aggregation of inputs or outputs may not be consistent with economic optimization of either costs or profits.

Consider first the possibility of functional aggregation in (1) such that inputs can be simply aggregated additively following the approach used for most agricultural data. In such a case,

$$y = f\big(f_1(x^1, t_1), \ldots, f_m(x^m, t_m)\big) = \tilde{f}\big(x^1 + \cdots + x^m\big). \tag{14}$$

Such additivity (a special case of strong separability) implies that the marginal product of an input in stage i is equal to the marginal product of an input in stage j. Hence, a mean-preserving spread in the distribution of an input across stages will have no impact on output. This assumption implies that generic inputs applied across stages are perfect substitutes, which seems unreasonable in virtually all agricultural production. Thus, such a rationalization for adding generic inputs is summarily rejected.[18] Perhaps more general forms of functional aggregation following Blackorby, Primont, and Russell (1978) are appropriate, but additivity seems unreasonable.

By far, the most common reason to presume that technology can be written with the sum of inputs is based upon efficient allocation across outputs. This explanation commonly proceeds by assuming positive use of each x_i in each stage in (3). Efficiency implies that annual aggregate inputs are allocated among production stages to equate marginal products and rates of technical substitution across stages. Thus, optimization results in the efficient description of technology in terms of the aggregate or added inputs. However, this conclusion crucially hinges on a notion similar to Hick's composite commodity theorem [Hicks (1956)].

[18] However, some inputs are apparently perfect substitutes or near perfect substitutes within a stage. For example, two brands of fertilizers with the same chemical content may be perfect substitutes.

Samuelson (1967, p. 231) made clear the assumption that all inputs had to have the same factor prices across industries in his development of the PPF. Under such conditions, efficiency concepts can serve as the first stage of a two-stage optimization process. To this end, if $\boldsymbol{x}$ enters as a sum in the PPF definition, it must enter as the same sum in the calculation of costs. For example, in the Lau problem, if and only if the wage rate is the same for all industries is (13) equivalent to minimizing the cost of an input subject to technology and the endowments of other inputs (e.g., by multiplying both ω^* and ω_i^* by the same wage rate) which then, in turn, is consistent with profit maximization if all input allocations aggregated in $\boldsymbol{x}$ have the same prices.

3.3. *Instructive examples of within-firm aggregation*

Several examples can illustrate the implicit problem of within-firm aggregation across commodities and allocations in agricultural data. Consider the case with two outputs, y_1 and y_2, distinguished over time or space with corresponding prices p_1 and p_2; and two inputs, x_1 and x_2, distinguished by time or space with corresponding prices r_1 and r_2.

3.3.1. *Case 1: Price homogeneity allows additive aggregation independent of prices*

Suppose technology follows $y_i = f_i(x_i)$, $i = 1, 2$, so the profit maximization problem is

$$\pi = \max_{x_1,x_2}\left\{p_1 f_1(x_1) + p_2 f_2(x_2) - r_1 x_1 - r_2 x_2\right\}.$$

If $p_1 = p_2 = p$ and $r_1 = r_2 = r$, then inputs and outputs can be aggregated additively with $x = x_1 + x_2$ and $y = y_1 + y_2$ so the problem can be represented as

$$\begin{aligned}\pi &= \max_{x_1,x_2}\left\{p\left[f_1(x_1) + f_2(x_2)\right] - r(x_1 + x_2)\right\}\\ &= \max_{x}\left\{pf(x) - rx\right\}\\ &= \pi^*(p, r),\end{aligned}$$

where the aggregate technology $f(x)$ is defined independent of prices by an implicit maximization,

$$f(x) = \max_{x_1,x_2}\left\{f_1(x_1) + f_2(x_2) \mid x = x_1 + x_2\right\}. \tag{15}$$

The maximization in (15) requires equating the marginal products across input uses. In many instances, economists (and statisticians who produce the data they use) aggregate inputs or outputs simply by adding them as in Equation (14). As noted earlier, such practices are typical in conceptual descriptions of aggregate technical efficiency. In empirical work, perhaps the most common examples of simple adding across space are generic inputs like fertilizer, water or land. On the other hand, capital service or labor categories are often summed across time [Shumway (1983, 1988)]. As noted by this case, the ability to do so properly hinges on the equality of prices.

3.3.2. Case 2: Price heterogeneity requires index aggregation

Reality requires consideration of the case where prices are not equal. For example, land typically has heterogeneous quality. Even hard red No. 2 winter wheat prices differ by location and time. Suppose prices are not identical, $p_1 \neq p_2$ and $r_1 \neq r_2$, but other assumptions follow Case 1. In this case, quantity aggregation requires use of price weights. Assuming the existence of index numbers consistent with Fisher's weak factor reversal property [Fisher (1922)], the profit maximization problem can be represented as

$$\begin{aligned} \pi &= \max_{x_1,x_2}\{p_1 f_1(x_1) + p_2 f_2(x_2) - r_1 x_1 - r_2 x_2\} \\ &= \max_{x}\{pf(x) - rx\} \qquad (16) \\ &= \pi^*(p,r), \qquad (17) \end{aligned}$$

where p is an index of output prices, r is an index of input prices, aggregate quantities are defined with index weights such that $x = x_1^* + x_2^*$ and $y = y_1^* + y_2^*$ where

$$\begin{aligned} y_1^* &= (p_1/p)y_1, \qquad y_2^* = (p_2/p)y_2, \\ x_1^* &= (r_1/r)x_1, \qquad x_2^* = (r_2/r)x_2, \end{aligned} \qquad (18)$$

and the aggregate technology is represented by $f(x)$ as defined by an implicit maximization,

$$f(x) = \max_{x_1^*,x_2^*}\{(p_1/p)f_1(rx_1^*/r_1) + (p_2/p)f_2(rx_2^*/r_2) \mid x = x_1^* + x_2^*\}. \qquad (19)$$

In this case, price-weighted marginal products are equated across input uses. Thus, the problem is represented accurately with aggregates but the definitions of the aggregates depend crucially on the price weights. Dependence on the price weights means that an estimate of the technology in (16) or (19) or the profit function in (17) under one price regime may not serve well to forecast the response to a different (possibly unobserved) price regime, i.e., where the corresponding $y = f(x)$ is not known or observed. For example, a very different distribution of prices among the outputs (inputs) could lead to the same price index p (r) as in (17) but a very different $f(x)$.

Examples abound where aggregates are not simply summed but are formed as price-weighted aggregates. Examples are index numbers in Divisia, Paasche or Laspeyres form. Often, Fisher's weak factor reversal property is used so that a quantity (price) index can be implicitly calculated from a price (quantity) index. The important point is explicit recognition that aggregation involves heterogeneous prices and products. Index numbers that are exact for particular technologies have been explored by Diewert (1976), who coined the phrase "superlative indexes" for those that correspond to homogeneous second-order flexible technical forms.

The development of these indexes is typically based upon cost minimization assuming all prices and quantities are positive. When inputs and outputs are positive and separable from one another, index procedures may exist that are exact for both aggregators of inputs (in the production function) and outputs (in the input requirements function). Clearly, the more aggregate the data, the less likely is observation of a zero production. These indexes are clearly useful to represent aggregate output or inputs or even to aggregate a portion of each. However, they do not generally specify technology in a form that can be used to illuminate allocative technical efficiency. Also, exact indexes are not robust with respect to behavioral preferences. Profit maximization is a crucial assumption. For example, risk aversion where inputs affect risk is sufficient to cause failure.

Currently, publicly reported agricultural data at county, state, regional, and national levels of aggregation contain many Laspeyres aggregations [Shumway (1988)] that are exact only for Leontief or linear technology [Diewert (1976)]. Some data particularly at state or lower levels of aggregation are constructed using simple summation aggregators such as a simple average [Pope and Chambers (1989); Chambers and Pope (1991)]. Neither is exact under flexible functional form technology. In recent years, some public data aggregated with Tornqvist–Theil indexes across groups of inputs or outputs has appeared. This approach is exact for homogeneous translog technology [Ball (1985); Ball et al. (1997); Ball et al. (1999)]. However, this and other index approaches are limited by the fact that data on many of the groups that go into these calculations are constructed with simple summations.

In lieu of using public aggregate data, some studies investigate agricultural production using one of the few farm-level data sets that have been collected (e.g., the Kansas State University farm accounting data). Farm-level data is scarce and, in most cases, access is limited. Moreover, from the standpoint of the discussion in this section, farm-level data is typically derived from expenditure and receipt information in accounting records. Expenditures and receipts are typically aggregated additively over input categories, time, and/or spatially separated production activities. Because no indexing of prices is used, the implicit assumptions necessarily correspond to Case 1.

3.3.3. *Case 3: Price homogeneity with short-term fixities or corner solutions*

Unfortunately, convenient rationalizations that accompany exact index numbers or simple sums in the production possibilities frontier fail when fixed or corner solutions arise. Suppose prices are identical across time or space as in Case 1 but that at least one of the production activities is constrained by short-term fixities. Where the inputs and outputs represent temporal heterogeneity, the fixities could represent family labor or capital service flow constraints that vary by time period. Where the different inputs and outputs represent spatial heterogeneity, the fixities could represent land allocation constraints imposed by government policy (acreage set asides, diversion requirements, environmental restrictions such as pesticide use near surface water, etc.). Suppose technology follows $y_i = f_i(x_i, z_i)$, $i = 1, 2$, where z_i represents, say, the amount of land allocated to production activity i. Suppose further that allocation of a fixed input quantity, z, between the two production activities is limited by a restriction, $z_1 \leqslant z_1^*$, that turns out to be binding. If $p_1 = p_2 = p$ and $r_1 = r_2 = r$, then the profit maximization problem is

$$\pi = \max_{x_1, x_2} \{ p[f_1(x_1, z_1) + f_2(x_2, z_2)] - r(x_1 + x_2) \mid z_1 = z_1^*, z = z_1 + z_2 \}.$$

In this case, the inputs and outputs can be aggregated additively with $x = x_1 + x_2$ and $y = y_1 + y_2$ as in Case 1 but the problem requires a more complicated representation:

$$\begin{aligned} \pi &= \max_{x_1, x_2} \{ p[f_1(x_1, z_1^*) + f_2(x_2, z - z_1^*)] - r(x_1 + x_2) \} \\ &= \max_{x} \{ pf(x, z_1^*, z) - rx \} && (20) \\ &= \pi^*(p, r, z_1^*, z). && (21) \end{aligned}$$

Here the aggregate technology, $f(x, z_1^*, z)$, can be defined independent of prices but not independent of the short-term fixities,

$$f(x, z_1^*, z) = \max_{x_1, x_2} \{ f_1(x_1, z_1^*) + f_2(x_2, z - z_1^*) \mid x = x_1 + x_2 \}.$$

The latter implicit maximization requires equating the marginal products of x across input uses but the marginal products depend on how fixities affect land allocation. If factors affecting these fixities (z_1^* or z) vary over observations (time or space) used to estimate the production problem, then the specification and estimation of (20) or (21) is not as simple and elegant as standard methodologies imply. Specifically, estimation of (20) and (21) is not generally valid unless the disaggregated allocation of land is considered explicitly. This implies that the state-level practice of simply adding acreage for the estimation of crop technologies is problematic unless land is homogeneous. The constrained problem becomes particularly complicated if such constraints are intermittently binding across observations that represent different land qualities or are intermittently imposed across time or space by government policy.

As noted in Section 2.1, agricultural production economics has compiled substantial conceptual and empirical support for treating capital and family labor service flows as constrained at crucial times during the growing season. Thus, in certain stages of production in (2), labor or capital service constraints may be binding. These will likely have different shadow values because constraints will be binding in some periods and not others. Any attempt to represent efficiency in terms of total availability or total use of a service is inappropriate.

3.3.4. Case 4: The case of corner solutions with ex post adjustment

Now suppose that a random state of nature is introduced to which the producer can respond, e.g., by applying pesticides if a pest infestation is observed. Suppose production follows $y_i = f_i(x_i, z_i, \varepsilon_i) = z_i^{\alpha_i}[\beta_i + (1-\beta_i)(1-\mathrm{e}^{-x_i})]^{\varepsilon_i}$ where ε_i is a random state of nature equal to zero or one depending on whether a pest infestation occurs, $\alpha_i > 0$, $0 < \beta_i < 1$. Thus, if $\varepsilon_i = 0$, then production is $y_i = z_i^{\alpha_i}$. If $\varepsilon_i = 1$, then a portion of the crop is lost resulting in (i) production $y_i = \beta_i z_i^{\alpha_i}$ if no pesticide is applied or (ii) production asymptotically approaching the case of no pest infestation as large amounts of pesticides are applied. Suppose land allocation must be determined prior to realization of the state of nature and must satisfy the binding land constraint, $z_1 + z_2 = z$. Then the profit maximization problem is represented by

$$\begin{aligned}\pi &= \max_{z_1,z_2}\left\{\mathrm{E}\left[\max_{x_1,x_2}\left\{\sum_i p_i z_i^{\alpha_i}\left[\beta_i + (1-\beta_i)\left(1-\mathrm{e}^{-x_i}\right)\right]^{\varepsilon_i}\right.\right.\right.\\ &\qquad\qquad \left.\left.\left. - \sum_i r_i x_i\right\}\right]\,\middle|\, z_1 + z_2 = z\right\}\\ &= \max_x\left\{pf(z, x, \varepsilon_1, \varepsilon_2) - rx\right\} &(22)\\ &= \pi^*(z, p, r, \varepsilon_1, \varepsilon_2), &(23)\end{aligned}$$

where E is an ex ante expectation, p and r are price indexes for outputs and inputs, respectively, and aggregate quantities are again defined with index weights so that $x = x_1^* + x_2^*$ and $y = y_1^* + y_2^*$ following (18). For this problem, the aggregate technology must be defined by the implicit maximization,

$$f(x, \varepsilon_1, \varepsilon_2) = \max_{x_1^*, x_2^*}\left\{\sum_i (p_i/p)(z_i^*)^{\alpha_i}\left[\beta_i + (1-\beta_i)\left(1-\mathrm{e}^{-rx_i^*/r_i}\right)\right]^{\varepsilon_i}\,\middle|\, x = x_1^* + x_2^*\right\}, \tag{24}$$

where

$$\begin{aligned}(z_1^*, z_2^* \mid x) &= \operatorname*{argmax}_{z_1,z_2} \mathrm{E}\left[\max_{x_1^*,x_2^*}\left\{\sum_i (p_i/p) z_i^{\alpha_i}\left[\beta_i + (1-\beta_i)\right.\right.\right.\\ &\qquad \left.\left.\left.\times\left(1-\mathrm{e}^{-rx_i^*/r_i}\right)\right]^{\varepsilon_i}\,\middle|\, x = x_1^* + x_2^*\right\}\right].\end{aligned}$$

For this problem, the price-weighted marginal products that are equated across input uses are dependent on the states of nature. For example, for (22), (23) or (24) to correctly reflect technology, they must be conditioned on disaggregated states of nature.

These cases make clear that simple index procedures may not be empirically appropriate when corner solutions, fixities, price heterogeneity or ex post adjustments are present.[19] The essential point relevant to representation of production technologies in terms of aggregate inputs and outputs is that either (i) prices (or shadow prices) of those goods that are aggregated additively must be homogeneous or (ii) aggregation must follow index forms appropriate to the (unknown) technology. In the latter case, production choices must not be constrained by fixities and cannot involve ex post adjustment to states of nature. Otherwise, disaggregated data is required to represent efficiency, i.e., production possibilities frontiers expressed solely in terms of simple aggregates are not well defined.

Some important principles implied by the above cases are as follows:

PRINCIPLE 1. *Each unconstrained input aggregation in the efficiency concept should be composed of allocations that have identical prices if fixities or ex post adjustments affect those allocations.*

PRINCIPLE 2. *Each output aggregation in the efficiency concept should be composed of output quantities that have identical prices if fixities or ex post adjustments affect their production.*

It is tempting to state each of these principles in a form that requires identical prices generally. Indeed, the basic concept of technology is typically stated in terms of sets or functions defined over inputs and outputs alone (not depending on prices). Clearly, Case 1 illustrates what is required to represent such technologies. However, Case 2 and the exact aggregation literature clearly show that prices can be appropriately used to aggregate inputs or outputs when, for example, the fixities and ex post adjustments of Cases 3 and 4 are not present. In these cases, theorems are required to rationalize procedures for aggregation using prices as illustrated by the influential work of Diewert (1976). However, these theorems for aggregation via price indexes require knowledge of the functional form of technology, the class of behavior, and all prices that provide behavioral incentives. A number of circumstances limit the practical usefulness of these results. For example, markets for risk are generally believed to be incomplete in agriculture so that necessary prices may not exist. An example can illustrate these index problems when production is random.

3.3.5. *Case 5: Dependence of exact indexes on behavior and technology*

Suppose technology is quadratic and random of the form

$$y = f(\boldsymbol{x})\varepsilon = (\boldsymbol{x}A\boldsymbol{x})\varepsilon, \quad \mathrm{E}(\varepsilon) = 1,$$

[19] Differences in marginal returns due to failure to adjust to identical prices was recognized as a significant problem in published productivity indexes by Griliches (1963) who attempted to estimate the difference in marginal returns among input allocations and to correct index number measurements accordingly.

(where transposition is suppressed for convenience) and that the firm is an expected utility maximizer solving

$$\max_{x} \mathrm{E}\big[U(w_0 + pf(\boldsymbol{x})\varepsilon - \boldsymbol{rx})\big],$$

where all prices are certain, p is output price, $\boldsymbol{r}$ is a vector of input prices, U is utility, E denotes expectation, and w_0 is certain initial wealth. The first-order conditions (assumed sufficient) are

$$pf_x(\boldsymbol{x}) - \boldsymbol{r} - R_x(p, \boldsymbol{r}, w_0) = 0,$$

where R_x is the marginal risk premium. Following the quadratic lemma of Diewert (1976), the difference in output from some base period is given by

$$\mathrm{E}(y) - \mathrm{E}(y_0) = .5[f_x(\boldsymbol{x}) + f_x(\boldsymbol{x}_0)](\boldsymbol{x} - \boldsymbol{x}_0), \tag{25}$$

where $\boldsymbol{x}_0$ is the base period vector of inputs and y_0 is the corresponding output. Inserting the first-order conditions from expected utility maximization into (25) gives

$$\mathrm{E}(y) - \mathrm{E}(y_0) = .5\big[(\boldsymbol{r} + R_x)/p + (\boldsymbol{r}_0 + R_{x_0})/p_0\big](\boldsymbol{x} - \boldsymbol{x}_0).$$

Assuming that inputs are normalized so that $\mathrm{E}(y_0)$ is a known constant, expected output is known only if the two marginal risk premiums are known. Under risk neutrality, R_x and R_{x_0} are zero and (changes in) output is (are) simply represented by

$$\mathrm{E}(y) - \mathrm{E}(y_0) = .5(\boldsymbol{r}/p + \boldsymbol{r}_0/p_0)(\boldsymbol{x} - \boldsymbol{x}_0),$$

which is a simple index of observed relative input prices and inputs. This result illustrates that knowledge of the proper behavioral model is required for use of index numbers and all dual methods that infer the form of technology using them. Of further interest is that restricting the form of technology can lead to a standard index number. For example, consider homogeneous quadratic production, $y = f(\boldsymbol{x}) = (\boldsymbol{xAx})^{.5} + (\boldsymbol{xBx})^{.25}\varepsilon$, $\mathrm{E}(\varepsilon) = 0$ where $\boldsymbol{A}$ and $\boldsymbol{B}$ are parameter matrices. Given constant absolute risk aversion and normality, $p(\boldsymbol{xAx})^{.5} - .5\lambda p^2\sigma^2(\boldsymbol{xBx})^{.5}$ is obtained as Fisher's (1922) ideal price index, which is the geometric mean of the Paasche and Laspeyres price indexes where σ^2 is the variance of ε and λ is the absolute risk aversion coefficient. Here the index is standard but its meaning is not. The index recovers expected revenue reduced by the risk premium.

Even when exact index forms are known, it must be recognized that many prices may not be known to firms when inputs are applied or when outputs are planned, implying again that behavior may well not follow optimization of cost, revenue, or profit functions based on index numbers. That is, explicit technological parameters may be eliminated using first-order conditions for optimization. Risk preferences, moments of the distribution of prices and even technological parameters in the marginal risk premium may remain. These also will vitiate the convenience of index number aggregation. The simple fact is that prescription or prediction for unobserved price scenarios (such as are necessary in ex ante policy analyses) cannot be usefully addressed with the exact index number approach. Although one might initially think that these problems refer only to outputs, Case 4 makes clear that input prices in a dynamic world can be subject to many of the same concerns as output price uncertainty.[20]

Principles 1 and 2 must be enlarged when they apply to aggregation of service flows from farmer-owned resources. Suppose the farm production problem is described by

$$\begin{aligned} &F(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}) \equiv y_1 - f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}), \\ &f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}) \equiv \max\{y_1 \mid \boldsymbol{y} = f(\boldsymbol{x}, z^1, \ldots, z^m),\ z^i \leqslant \boldsymbol{k},\ i = 1, \ldots, m\}. \end{aligned} \tag{26}$$

Here i indexes time. Use of recurring allocated fixed factors must meet availability constraints. These may be limiting in some time periods and slack in others. An obvious question is when can this problem be represented by time-aggregated service flows in the form

$$y_1 = f^{**}(\boldsymbol{y}_{-1}, \boldsymbol{x}, z), \tag{27}$$

[20] To highlight the severity of this problem, we note that almost all agricultural production studies combine pesticides into one variable. This is problematic because (i) at least some pesticides are applied after commencement of the growing season when some random conditions are already observed (as in Case 4), (ii) many pesticides have highly specific uses, and (iii) individual pesticide prices may not be highly correlated because of the role of patents and market concentration. To illustrate the specificity of uses, grasses on corn lands are typically controlled by Eradicane or Sutan at the pre-planting stage, by Lasso, Dual, or Prowl in other pre-emergent applications, and by Beacon or Accent in post-emergent applications. In contrast, grasses on soybean lands are typically controlled by Treflan or Prowl at pre-planting, by Lasso or Dual in pre-emergent applications, and by a number of additional herbicides in post-emergent applications. Broadleaf weeds on corn lands, on the other hand, are typically controlled by Atrazine at pre-planting, by Atrazine or Bladex in pre-emergent applications, and by Banvel, 2,4-D, Buctril, Permit, or Exceed in post-emergent applications. To illustrate the magnitude of unrelated price movements among leading pesticides, note that newly patented pesticides come onto the market almost every year while patents on others expire, leading to generic competition. Either can cause the price of a specific pesticide competing in a specific use to decline by as much as 20–50 percent while other pesticide prices are rising. Examples include a 43 percent decline in glyphosate price due to patent expiration while the price of atrazine increased 25 percent in response to a dramatic reduction in the number of selling firms during 1989–1992. Such dramatic differences in price variation are not the exception. For example, generic entry following patent expiration caused a price decline of 20 percent for atrazine, 26 percent for diuron, 40 percent for simazine, and 25 percent for trifluralin (not simultaneously) while the pesticide price index and prices for leading pesticides such as Lasso and Lorsban were rising [United States Senate (1987)]. Such examples are likely to increase in importance because of the increasing predominance of pesticide-dependent no-till technologies and because of genetic engineering which is creating niche products such as Roundup-ready soybeans that introduce dependence on specific products. To illustrate this trend, note that pesticides now account for 17–21 percent of the variable costs of corn production and 30–35 percent of the variable costs of soybean production in most areas of the United States [Economic Research Service (1999a and 1999b). Thus, serious concern may be warranted when crop- or location-specific variation in pesticide prices is aggregated or swept under the guise of an error term.

where $z = \sum_{i=1}^{m} z^i$? The answer is that the behavior of a profit-maximizing farmer can be represented generally with such a production function only when the shadow prices of service flows are constant over time. The reasons are that (i) profit-maximizing service flows in each time period will have different marginal productivities corresponding to the different shadow prices and (ii) the constraints are typically binding only intermittently. Once aggregated, the differences in shadow prices are lost.[21] A simple example can illustrate.

3.3.6. Case 6: Aggregation of service flows

Let output be additive in the stage outputs with output price equal to 1 where the first stage output is given by $z_1 - .5z_1^2$ and the second stage output is $3z_2 - .5z_2^2$. The first stage has marginal product $1 - z_1$ and the second stage has marginal product $3 - z_2$. Where $k = 2$ the optimal solution is

$$z_1^* = 1, \quad \lambda_1^* = 0, \quad z_2^* = 2, \quad \text{and} \quad \lambda_2^* = 1$$

with output 4.5 where λ_i^* represents the shadow value of the service flow constraint in time period i. One cannot maximize profit or output subject to an aggregate service flow availability constraint, $z = 4$, or an aggregate service flow use constraint, $z = 3$, because the different magnitudes of multipliers in different time periods cause different marginal products. In both cases in this example, the use constraint in the second time period would be violated. However, if the shadow prices and total shadow value of available service flows are observed, one could correctly maximize profit or output subject to $v = \lambda_1^* z_1 + \lambda_2^* z_2$ where v is the total shadow value of available service flows ($v = 2$ in the example above). The problem here is that both the optimal shadow prices and the total shadow value of available service flows depend on parameters of the problem that are likely to vary among observations used for estimation both across stages of production and across farms.

The principle implied by this discussion of shadow prices is as follows.

PRINCIPLE 3. *For inputs or outputs that are constrained, each output and input aggregation in the efficiency concept should be composed of those allocations that have identical shadow prices. For service flows that are constrained, if shadow prices of service flows are not observed intraseasonally and used for weighted aggregation, each service flow aggregation in the efficiency concept should be composed of service flows*

[21] The assumption implicit in the allocated fixed input constraint is a "use it or lose it" concept, e.g., if operator labor is not used this time period it does not add to operator labor available for next period. While this assumption applies quite well to labor and land, some types of machinery if used less may have more machinery life available for future time periods. In such cases, the recurring fixity constraint would apply because of fixed machinery capacity but some further user costs would need to be considered among variable costs to reflect how much of a machine's life is exhausted with use.

that have identical shadow prices. Alternatively, the efficiency concept should be based on intraseasonal service flow constraint levels as in (4) *rather than actual service flows.*

As Case 6 implies, it is not the total seasonal flow or stock that is relevant. It is the maximal capacity service flow in each period (which for notational convenience we have assumed is identical in each period). This capacity is what enters (4). The representation of technical efficiency will not be smooth in maximal service flow capacities represented in $\boldsymbol{k}$ because constraints bind in some seasons and do not bind in others.

The implications of not following this principle are difficult to determine because shadow prices are not readily observable. Furthermore, shadow prices typically depend on market prices and behavioral objectives. Thus, aggregating service flows makes the description of technology dependent on prices, policy and behavior. In other words, the associated efficiency concept is not generally a technical efficiency concept when aggregate service flow data are used.

We note that agricultural production analysis has been increasingly turning toward representing aggregate production relationships in terms of aggregate service flows. That is, agricultural production is increasingly being modeled using capital service flows as variable inputs as in (27) rather than with capital investments as fixed inputs as in (26). This movement has both motivated and been motivated by the development of public agricultural production data as measurements of capital service flows rather than capital stocks [Ball (1985); Ball et al. (1999)]. Thus, the prevailing direction of emphasis both in agricultural production analysis and data generation appears to be leading away from a valid representation as in (26) and toward a representation as in (27), the underlying assumptions of which are inapplicable according to Principle 3 except in a narrow and unlikely set of circumstances.

The principles of this section highlight the critical nature of heterogeneity due to spatial or temporal price variation in agricultural production. As discussed in Section 2, some major price variations over typical spatial and temporal aggregations of both inputs and outputs appear to be relatively large and thus render those aggregates of questionable value for testing technical efficiency of production or for simply representing the standard characteristics of technology.

These examples and principles lead to a set of conditions that are sufficient for simple aggregate representations of technology dependent on observed data.

3.3.7. The aggregation qualification condition

A. Simple input or output aggregation devoid of prices requires:
   1. Functional separability [Blackorby et al. (1978)], or
   2. Equality of prices and marginal conditions across all aggregated quantities. If ex post adjustments under uncertainty apply, then all aggregated quantities must also be adjusted according to the same ex post information.

B. Conventional input or output aggregation using prices and observable production data requires the cases for which index forms yield exact cost or revenue optimization in terms of aggregates, e.g, as in the quadratic approximation lemma of Diewert (1976). These aggregations must not be over variables affected other than as aggregates by intertemporal or activity-specific policy constraints or ex post adjustments under uncertainty and must not depend on preferences. Such aggregations are not useful when disaggregated prices are unobserved.

3.4. The production possibilities frontier as a representation of technology

Consider next the typical practice of representing multi-output technologies by their PPFs. With the development and application of tractable flexible forms using dual methods, a number of studies based upon the PPF of multiproduct firms have ensued [e.g., Antle (1984); Ball (1988); Shumway (1983); Weaver (1983)]. That is, in virtually all multi-output empirical applications of duality, allocations of inputs are ignored [Chambers and Just (1989) is an exception]. For empirical purposes (when smoothness is imposed), efficient technology is characterized implicitly by a single-equation representation of the product transformation function involving only aggregate inputs and aggregate outputs. These studies examine issues for which the measurement of $F(\boldsymbol{y}, \boldsymbol{x})$ is beneficial, including measurement of total factor demands and product supplies, various forms of separability, productivity and technical change, and the standard characteristics of the PPF.[22]

However, examination of an economic sector or firm as a whole by means of the PPF using $F(\boldsymbol{y}, \boldsymbol{x})$ or its dual profit function cannot answer a number of interesting questions that have policy or management relevance. To illustrate, suppose production is truly nonjoint so that the existence and notion of an underlying technology is clear – the f_i's in Samuelson's case in (11) or the ω_i^*'s in Lau's definition in (13). Total profit, π, is the sum of industry (or production activity) profits where each industry (or production activity) profit, π_i, depends only on the corresponding output price, p_i, and input price vector $\boldsymbol{r}$,

$$\pi(\boldsymbol{p}, \boldsymbol{r}) = \sum_i \pi_i(p_i, \boldsymbol{r}), \tag{28}$$

where $\boldsymbol{p}$ is the output price vector. The dual to the left side of (28) is the PPF or transformation function, $F(\boldsymbol{y}, \boldsymbol{x}) = y_1 - f^*(\boldsymbol{y}_{-1}, \boldsymbol{x})$, while the dual to an element of the sum on the right-hand side of (28) is the industry production function, $y_i = f_i(\boldsymbol{x}^i)$.

Note that any structure found to be present in π or F says almost nothing about the structure of any π_i or f_i. For example, separability of π in some partition of $\boldsymbol{r}$ does not

[22] For these purposes, however, one must attach a *ceteris paribus* qualification as demonstrated in the following section. That is, changes in policy or behavior can alter the apparent PPF.

imply separability of π_i in that partition nor vice versa. In other words, learning about the structure of $F(\boldsymbol{y}, \boldsymbol{x})$ either directly or implicitly through $\pi(\boldsymbol{p}, \boldsymbol{r})$ reveals little about the structure of any $f_i(\boldsymbol{x}^i)$.

PRINCIPLE 4. *The structure of a production possibilities frontier, which is the level at which production technology is represented in most modern production studies, does not reveal the structure of any distinct underlying (industry- or production-activity-specific) technologies.*

3.5. An illustration of the technical content of a production frontier

The point of Principle 4 can be illustrated with an example including two competitive industries (or production activities). Using an underlying technology that is nonjoint and symmetrically separable in inputs from outputs, the PPF exhibits separability in inputs from outputs when the partial production elasticities in both sectors are equal. Thus, a test for separability of the PPF may be only a test about the relationship of production elasticities rather than separability of the underlying technology.

Consider a production technology with two outputs and two allocated inputs, one variable and one fixed, following Cobb–Douglas technology,

$$y_1 = a x_1^{\alpha_1} z_1^{\alpha_2}, \qquad y_2 = b x_2^{\beta_1} z_2^{\beta_2}, \tag{29}$$

where x_1 and x_2 are amounts of the variable input allocated to the respective production activities and z_1 and z_2 are amounts of the fixed input allocated to the respective production activities. The aggregate amounts of the two inputs are thus $x = x_1 + x_2$ and $z = z_1 + z_2$, respectively. Suppose the latter must satisfy the allocated fixed input constraint, $z = k$. These relationships can be considered as constraints on the technology in any behavioral optimization or substituted into (29) to represent technology by

$$y_1 = a x_1^{\alpha_1} z_1^{\alpha_2}, \tag{30}$$

$$y_2 = b(x - x_1)^{\beta_1}(k - z_1)^{\beta_2}, \tag{31}$$

where $\boldsymbol{X} = (x_1, x_2)$ and $\boldsymbol{Z} = (z_1, z_2)$. To maximize profits, $\pi = p_1 y_1 + p_2 y_2 - rx$, subject to the constraints, the first-order conditions corresponding to (29) are

$$\alpha_1 p_1 a x_1^{\alpha_1 - 1} z_1^{\alpha_2} - \beta_1 p_2 b(x - x_1)^{\beta_1 - 1}(k - z_1)^{\beta_2} = 0, \tag{32}$$

$$\alpha_2 p_1 a x_1^{\alpha_1} z_1^{\alpha_2 - 1} \beta_2 p_2 b(x - x_1)^{\beta_1}(k - z_1)^{\beta_2 - 1} = 0, \tag{33}$$

$$\beta_1 p_2 b(x - x_1)^{\beta_1 - 1}(k - z_1)^{\beta_2} - r = 0. \tag{34}$$

Combining (30)–(34) and solving out prices obtains the relationship corresponding to (11) or (12).

Another representation of technology is to solve two of these relationships for prices, e.g., p_1 and p_2, after normalizing the other, e.g., setting $w = 1$, to obtain three remaining relationships devoid of prices. One such representation includes (30), (31) and

$$\frac{\beta_1(k - z_1)}{\alpha_1 z_1} = \frac{\beta_2(x - x_1)}{\alpha_2 x_1}, \tag{35}$$

which follows from combining (32) and (33). To obtain (11) or (12) from (30), (31), and (35), one can solve (35) for z_1. This result can be substituted into (30) which can then be solved for x_1. Then both of these results can be substituted into (31) to obtain an equation in y_1, y_2, x, and k.

Even though this step is possible in principle, an explicit solution cannot be found in practice without constraining the parameter space. Since an example suffices, let $\alpha_1 = \alpha_2$. Then solving (35) for z_1 obtains

$$z_1 = \frac{\beta_1 \alpha_2 x_1 k}{\alpha_1 \beta_2(x - x_1) + \beta_1 \alpha_2 x_1} \tag{36}$$

and solving (30) for x_1 obtains

$$x_1 = (y_1/a)^{1/\alpha_1}/z_1. \tag{37}$$

Solving (36) and (37) simultaneously yields

$$z_1 = \frac{2y_1^\gamma}{cx}\left[1 + \sqrt{1 + ckxy_1^{-\gamma}}\right], \qquad x_1 = \frac{cxa^{-\gamma}}{2}\left[1 + \sqrt{1 + ckxy_1^{-\gamma}}\right]^{-1},$$

where $\gamma = 1/\alpha_1, c = 4\beta_2 a^{1/\alpha_1}/(\beta_2 - \beta_1)$. The negative root is ruled out by positivity constraints. Substituting these results into (31) obtains

$$\begin{aligned} y_2 = b\left\{x - \frac{cxa^{-\gamma}}{2}\left[1 + \sqrt{1 + ckxy_1^{-\gamma}}\right]^{-1}\right\}^{\beta_1} \\ \times \left\{k - \frac{2y_1^\gamma}{cx}\left[1 + \sqrt{1 + ckxy_1^{-\gamma}}\right]\right\}^{\beta_2}. \end{aligned} \tag{38}$$

The relationship in (38) illustrates the problem with implicit representation of technology. While the underlying technology in (29) is separable in both inputs and outputs, and nonjoint in inputs, the implicit representation of technology by (38) satisfies none of these properties except in special circumstances.[23] For example, if one further assumes $\beta_1 = \beta_2$, then (38) reduces to

$$(y_1/a)^{1/(2\alpha)} + (y_2/b)^{1/(2\beta)} = (xk)^{1/2}$$

[23] Shumway, Pope, and Nash (1984) have previously shown that the presence of allocated fixed inputs can induce an apparent jointness even though the underlying technology is nonjoint. This result is somewhat more

which satisfies separability in both inputs and outputs. Thus, a test for separability in (38) may simply test whether parameters have particular relationships even though the underlying technology satisfies separability regardless.

This misleading conclusion occurs because additional information must be imposed together with technology to obtain (38) from (29). This additional information may be viewed as relatively harmless. For example, the assumption of profit maximization is not needed to obtain (38) from (29). Simple Pareto efficiency is enough or, equivalently, following Chambers (1988, p. 261) one can simply assume inputs and outputs are chosen to maximize, say, y_1 given y_2 and x subject to (30) and (31). Nevertheless, the implied relationship in (38) obscures the underlying technology and makes detection of its standard characteristics misleading and impossible. This example thus verifies Principle 4.

3.6. *Prescription versus description*

Ignoring or subsuming allocations has led to an ever larger division of interests and methods between farm management economists and production economists. Farm management economists have concentrated on strategies and prescriptions for input allocation across production stages and production activities (which they call enterprises) such that both technical and price efficiency is maintained. Production economists, on the other hand, have tended to assume efficient allocation implicitly in order to concentrate on properties of the multi-output efficiency frontier. An excellent example of this approach is the creation and subsequent analysis of aggregate agricultural productivity by Ball (1985) and Ball et al. (1999). Production economists, while often allowing for technical inefficiencies, typically have had little to say about the allocations of a given input over the growing season or across production activities. While this practice by production economists is due in part to data limitations, the data limitations are at least partially endogenous. Those designing data set construction and reporting have chosen to ignore allocations.

The most fundamental definition of economics involves the allocation of productive resources to the satisfaction of competing wants. In the study of production, application of this practice involves determining the optimal allocation of aggregate inputs to various industries or production activities in addition to simply determining the optimal aggregate. Historically, one of the important motivations for studying agricultural production economics was indeed prescriptive – to improve farm management and help farmers make better decisions. More recently, efforts have been devoted to helping regional and national policymakers formulate better policies. We note, however, that the

general because all that is changed in this example if both inputs are variable inputs allocated to separate production activities is that k is replaced by z in (38). The additional first-order condition is used in solving for the additional price of the second variable input. Thus, presence of allocated fixed inputs is not crucial in these results. Rather, ignoring allocations, whether of fixed or variable inputs, is the cause of incorrectly reflecting the properties of technology.

PPF is often inadequate in a prescriptive sense when inputs and outputs are aggregated. For example, it does little good for a water economist to determine the optimal capacity of a water system under rationing if no guidance is available for allocation of rations among jurisdictions or farms. Similarly, it does little good for a farmer to know the profit-maximizing aggregate use of fertilizer if no information is available on how to allocate it among crops of different productivities or plots of different soil capacities. Optimal benefits are generally unattainable without proper allocation.

The same principle applies to allocation of aggregate inputs over time. In the framework of Equation (3), knowledge of f^0 does not reveal the nature of the stage production functions nor do deviations from the frontier in f^0 reveal where inefficiencies occur in allocations.[24] Agricultural economists typically measure or estimate f^0 rather than f^*. Knowledge of f^0 is sufficient to address many interesting questions if it is well defined, but the existence, meaning, and appropriate measurement of f^0 hinges crucially on an implicit assumption of constant input prices within aggregates or lack of corner solutions throughout the stages. During growing seasons with high interest rates and varying input prices, f^0 may not be well defined.

This discussion implies an additional principle broadly derived from Principles 1–4 and put in context as follows.

PRINCIPLE 5. *Descriptions of technology expressed only with aggregates over time and location are not conducive to prescription for farm management and they limit meaningful analysis of policy controls.*

In summary, the conventional PPF that subsumes allocative efficiency is not the object of interest in many economic analyses. In practice, knowing what is good may be of little help without knowing how to get there. Conventional analysis of the PPF leaves out direct information on most allocation decisions. It cannot be used to uncover the structure of any underlying sub-technologies. Furthermore, the conventional PPF is not robust in the presence of various policies, behavioral preferences, and environments. When complexities of behavior or environment are introduced, one must proceed from a more basic notion of production efficiency to determine if the usual concept and calculation of allocative efficiency is appropriate or should be amended. Knowledge of any underlying sub-technologies is essential in this process. Hence, knowledge of the sub-technologies is always relevant but knowledge of the PPF may not be relevant. Moreover, the PPF may not be well defined because dependence on policies and behaviors may not be specified but yet affect empirical observations. The above discussion motivates the need to explore alternative concepts of technical and production efficiency which may be useful in distinguishing underlying technologies from the conventional PPF.

[24] Of course being on a production function is not sufficient for allocative technical efficiency.

3.7. *Eliminating behavior and policy from representations of technology*

Principles 4 and 5 are important because some inquiries are required at the level of a single production activity or of a single input to a single production activity that are not sensible at the aggregate PPF level. For example, inquiries regarding technical efficiency need to be sensitive to the extent of price variation across time, space, and production activities in order to have practical implications for overall firm efficiency or social efficiency. When an environmental agency considers prohibiting use of a single pesticide on a group of crops (perhaps the most common type of economic benefit analysis used by an environmental protection agency), a PPF that aggregates use of all pesticides across all crops will be of little use for analyzing the implications. On a more technical level with respect to the properties of production, homotheticity is essentially about the scaling of inputs and/or outputs leaving ratios unchanged. Examination of homotheticity of agricultural technology, for example, using a regional PPF seems to have little policy (or "what if") relevance due to the fixity of land. Nonjointness as implied by $\partial y_i/\partial p_j = 0$ $(j \neq i)$ is likely not present in the PPF due to land fixities even when technologies for individual production activities are nonjoint [Shumway et al. (1984)].

From a practical standpoint, the primary intent of many policies is to alter specific input allocations. For example, acreage controls in agriculture (allotments, set asides, and base acreages) apply to the use of a specific input (land) in a specific production activity (crop). Also, pesticide use standards apply at the crop- and sometimes location-specific levels. For example, EPA registrations allow a pesticide to be used only on crops that appear on its registration label. Other EPA requirements limit how close to surface water certain pesticides can be applied. With respect to outputs, government policy instruments such as target and support prices cause the same crop to be sold at more than one price in the same season (not all of a farmer's crop may qualify for the higher subsidized price). Turning to more recent crop and revenue insurance policies, the alteration of effective prices by indemnity payments is crop-specific in some cases and farm-wide in others. In each of these cases, the focus on a PPF following the modern practices of production economics effectively eliminates the relevant policy consideration by aggregating over decisions that are treated distinctly by policies.

Similar considerations apply to behavioral preferences that treat different production activities differently. While much of the modern production literature is based on profit maximization following standard dual approaches, one of the unique features of agriculture is risky production. If some production activities involve more risk than others, then risk averse farmers will tend to allocate fewer inputs to more risky activities, i.e., expected marginal productivities of inputs will be higher among more risky activities. With either a change in behavior (e.g., an increase in risk aversion with operator age) or an enhancement in policies such as crop insurance or government disaster assistance that mitigate risk, the differences in marginal productivities among production activities of different risks could change. Descriptions of technology that do not reflect individual production activities but only aggregate production possibilities cannot be used to analyze such policies or phenomena. Furthermore, analyses of technical efficiency based on revealed preference data affected by such policies is of questionable import when the effects of such policies are ignored.

Probably the most important reason to explore the underlying technology rather than the PPF has to do with robustness. Unless coupled with estimation of disaggregated production technologies, the observable PPF is policy- and behavior-dependent. For example, data envelopment analysis would tend to identify the production efficiency of the least risk averse farmer or the farmer least affected by policy parameters as "the" PPF. Alternatively, if the basic underlying technologies and preferences are estimated conditional on the specific policies affecting them, then a host of alternative policies can be evaluated, including those that address a specific type of behavior (e.g., like crop insurance addresses risk aversion). Pope and Just (1996) demonstrate that even production uncertainty with risk neutrality has fundamental implications for conceptualization and estimation of the cost function. Risk aversion is critical in evaluating, for example, changes in crop insurance. A conventional PPF (not conditioned on policy or behavior) may be clearly irrelevant for such analyses. It may serve only to indicate a potential that can never be reached or that is irrelevant in practice and, if so, will hold no useful information of social benefit.

Although there might be broad conceptual agreement that the PPF represents technology parameters and technical efficiency, distinguishing between a PPF conditioned on policy and one that is purely technological may be very difficult. They may appear observationally equivalent. For example, the constraint in Case 3 above could represent heterogeneity of land quality or an acreage policy control. In the former case it would be a part of technology while in the latter it would not. If policy controls are mingled with technology then a shift in the PPF has an uncertain source and estimates of the PPF are not useful for policy analysis. Productivity could increase due to either a technical change or a policy change such as elimination of the control. Principle 6 summarizes the advantages of a representation of efficiency that depends solely on technological relationships.

PRINCIPLE 6. *A useful concept of production efficiency for policy and management purposes corresponds to the first stage of a two-stage characterization of the producers optimization problem where the first stage fully reflects technical possibilities and the second stage includes all impacts of policies and behavior on decisions.*

3.8. An example with production errors

The point of Principle 6 can be illustrated by a simple one-input, two-output example using multiplicatively random nonjoint production functions. Let $y_1 = f_1(x_1)\varepsilon_1, \mathrm{E}(\varepsilon_1) = 1$, and $y_2 = f_2(x_2)\varepsilon_2$, $\mathrm{E}(\varepsilon_2) = 1$, where each f_i is strictly increasing. The PPF can be written as

$$y_2 = f_2\big(x - f_1^{-1}(y_1/\varepsilon_1)\big)\varepsilon_2,$$

where $x = x_1 + x_2$. Assuming prices are certain, if uncertainty is ignored in the second stage then the firm is assumed to produce on the PPF described by

$$-\partial y_2/\partial y_1 = f_{x_2}/f_{x_1} = p_1/p_2, \tag{39}$$

$$y_2 = f_2\big(x - f_1^{-1}(y_1)\big). \tag{40}$$

However, because y_1 and y_2 are random, (39) and (40) are not consistent with expected utility maximization. A risk neutral firm will produce where

$$-\partial \mathrm{E}(y_2)/\partial \mathrm{E}(y_1) = f_{x_2}/f_{x_1} = p_1/p_2,$$

$$y_2 = f_2\big(x - f_1^{-1}(y_1/\varepsilon_1)\big), \tag{41}$$

because y_1/ε_1 is $\mathrm{E}(y_1)$ given x_1. Given the nonlinear transformation of y_1 in (40), $\mathrm{E}(y_2)$ is not equal on average to the right-hand side of (41). Thus, ignoring uncertainty is inconsistent with two-stage expected utility maximization.

In general, to build a PPF in the Samuelsonian fashion consistent with expected utility maximization under price and production risk, one must identify all of the relevant moments of wealth that enter expected utility and develop a two-stage maximization approach consistent with the overall expected utility maximization problem. For example, if input prices are certain and equal as in the typical generic input case, and ε_1 and ε_2 are independent and have two parameter distributions, then

$$\mathrm{E}[U(w)] = U^*(m_{11}, m_{12}, m_{21}, m_{22}, w_0, r),$$

for some function U^* where m_{ij} is the ith moment of revenue for good j ($i, j = 1, 2$), w_0 is additive initial certain wealth, r is the generic input price, U is utility, and w is wealth (assuming cross-moments are zero). An appropriate two-stage procedure is defined by

$$\max_{m_{12}, m_{22}, m_{21}, x} \quad \max_{m_{11} | m_{12}, m_{22}, m_{21}, x} \quad \mathrm{E}[U(w)], \tag{42}$$

where x is total input use. If production is nonjoint and described by

$$y_1 = f_1(x_1) + h_1(x_1)\varepsilon_1, \quad \mathrm{E}(\varepsilon_1) = 0,$$
$$y_2 = f_2(x_2) + h_2(x_2)\varepsilon_2, \quad \mathrm{E}(\varepsilon_2) = 0,$$

and output prices are certain, then $m_{1j} = p_j f_j(x_j)$ and $m_{2j} = p_j^2 h_j(x_j)^2 \mathrm{E}(\varepsilon_j^2)$. Thus, h_1, f_2, h_2, and x can be effectively constrained in the first stage of (42).

In summary, the appropriate PPF concept for risk neutrality must be based on expected production but, more generally, the PPF must be tailored to the way risk enters production and the extent of risk aversion. This implies that an empirically useful PPF is necessarily dependent upon behavior and the environment.

4. Fundamentals of modeling agricultural production

This section builds upon the principles of Section 3 to suggest needed advances in models of agricultural production. Some of these advances may be feasible with present limitations while application of others is constrained by data availability. Finally, the meaning of existing empirical work when more general specifications apply is discussed. Traditionally, multi-output technologies were represented either by single-equation forms such as $F(\boldsymbol{y}, \boldsymbol{x}) = 0$, e.g., Klein (1947), or by individual production functions for each output where all inputs are allocated among individual outputs such as in (11), e.g., Pfouts (1961). Regarding these two cases as extremes, we suggest an intermediate premise based on the assertion that multi-output production problems typically exhibit at least one of the following properties: (i) some input(s) must be allocated among production processes or points of application in the production process either temporally or spatially,[25] (ii) some output(s) are produced by more than one production process or at more than one location or time in the production process, and/or (iii) some output(s) are produced as by-products so that their production is related in some way to the production of one or more other outputs.

As an example of (i), land in farms must be allocated among crops or (in developing agriculture) among crop mixes; automobile factory assembly lines are allocated among makes or models of cars; and chemical production plants are allocated among primary chemical processes. As an example of (ii), corn production on a farm is diffused among locations while the output of most manufacturing processes is diffused over time. As an example of (iii), many chemical production processes produce both a primary and one or more secondary chemicals; cotton ginning produces both cotton and cottonseed; and soybean crushing produces both soy oil and soy meal. In some activities, the producer may be able to influence the mix of outputs by adjusting the application of inputs (e.g., the choice of seed variety affects the oil content of soybeans) but, in others, the outputs may be constrained to fixed physical relationships (e.g., chemical reactions). While these examples are sufficient to verify validity of the premise, the discussion in Section 2 suggests that these features of agriculture are widespread and dominant.

This section explores the implications of this premise for technology representation. Results show that typical indirect or single-equation representations in such circumstances can, at best, provide reduced-form "as if" representations of technology that facilitate characterization of supply and demand in perfectly competitive markets but cannot identify the technology itself. At worst, such representations of technology are not

[25] Typically, some inputs are allocated to distinct production processes while others apply jointly. Knudsen (1973) argues that full nonjointness in inputs is unlikely because it assumes away technological reasons for the observed existence of multi-output firms. For example, training for management or automated control equipment in a multi-output plant or multi-production-process firm may simultaneously enhance production of all outputs. However, Leathers (1991) shows that a sufficient reason for existence of multi-output firms is short-run fixity of allocated factors.

well defined and are useless for investigation of a host of policy, management, and market structure issues in an imperfect world where credit constraints apply, contingency markets are missing, etc. The true underlying technology may provide more flexibility (the typical effect of unrepresented input allocations) or less flexibility (the typical effect of unrepresented by-product relationships). To develop these results requires a substantial development of conceptual groundwork to permit sorting out behavior from technology and to identify the meaning of various functions of aggregate variables. For this purpose, we place the technical detail in an Appendix but describe results in the following sections.

4.1. Structural concepts and efficiency of production

To facilitate clarity of discussion for the case where a firm's technology is possibly composed of several sub-technologies, several alternative concepts of efficiency must be defined. Sub-technologies are defined as production activities where, more generally than in specifications such as (11), each production activity can have more than one output thus allowing input jointness within sub-technologies. When the technology of a firm is composed of sub-technologies, we will say that the technology has *structure*. Typically, this structure can be exploited to understand the implications of alternative policies and preferences.

Suppose the production set $\Im$ can be described by sub-technologies $(\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon})$ where $\boldsymbol{y}^i$ and $\boldsymbol{x}^i$ are subvectors of $\boldsymbol{Y} \equiv \{\boldsymbol{y}^1, \dots, \boldsymbol{y}^m\}$ and $\boldsymbol{X} \equiv \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^m\}$, $\Im_i(z^i, \boldsymbol{\varepsilon})$ represents all possible combinations of $\boldsymbol{y}^i$ and $\boldsymbol{x}^i$ regardless of values taken by other elements of $\boldsymbol{Y}$ and $\boldsymbol{X}$, and aggregate outputs and inputs satisfy $\boldsymbol{y} \equiv \sum_i \boldsymbol{y}^i$ and $\boldsymbol{x} \equiv \sum_i \boldsymbol{x}^i$, respectively. This structure is sufficient to explore some possible implications of technologies where an important step in choosing the output mix is spatial allocation of inputs among plots as in Section 2.2. The different sub-technologies may represent various crop production activities on different plots. For example, one sub-technology may produce wheat and another soybeans by single cropping techniques and another might produce both wheat and soybeans at different times by double cropping. Clearly, the same principles apply to temporal allocation among time periods as in Section 2.1.[26]

[26] For added generality, this framework can easily add dependence of each sub-technology on outputs of lower sub-technologies. For example, the dependence of each successive stage of production on the intermediate outputs of the previous stage can be represented by

$$\{(\boldsymbol{Y}, \boldsymbol{X}) \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})\} \equiv \Big\{(\boldsymbol{y}^1, \dots, \boldsymbol{y}^m, \boldsymbol{x}^1, \dots, \boldsymbol{x}^m) \mid (\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon}, \boldsymbol{y}^{i-1}),\ i = 1, \dots, m,$$
$$\sum_s z_t^s \leqslant \boldsymbol{k},\ t = 1, \dots, T\Big\}.$$

Note that a suitable definition of $\boldsymbol{y}^{i-1}$ can permit dependence of each stage on all $\boldsymbol{y}^{i-j}$ for $j > 1$ as typical

For simplicity of presentation, this notation does not represent explicitly the possible presence of public inputs that cause jointness across sub-technologies, i.e., inputs that jointly affect multiple sub-technologies simultaneously. For example, the production set of each sub-technology might be described more completely by $(\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(\boldsymbol{x}^0, z^i, \boldsymbol{\varepsilon})$ where $\boldsymbol{x}^0$ is a vector of public inputs and the detailed use of variable inputs is described by

$$\boldsymbol{X} = \{\boldsymbol{x}^0, \boldsymbol{x}^1, \ldots, \boldsymbol{x}^m\}.$$

For purposes of this chapter, such public inputs may be present but are suppressed from notation to focus on the implications of allocations that are required for production decision implementation.

For notational simplicity, aggregations of the spatial and temporal allocation detail in $\boldsymbol{Y}$ and $\boldsymbol{X}$ vectors are represented by $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{Y}$ and $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{X}$, respectively, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are full row rank matrices of ones and zeros. The vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ maintain only physical distinction of outputs and inputs. Because each sub-technology may potentially produce only one or a few physical outputs using some subset of physical inputs, this notation can be suitably collapsed to eliminate identically zero elements of $\boldsymbol{Y}$ and $\boldsymbol{X}$ and related columns of $\boldsymbol{A}$ and $\boldsymbol{B}$.

In addition to descriptions of sub-technologies, the available technology set is assumed also to be constrained by availability of allocated fixed factors such as machinery services and operator labor. For example, if sub-technologies are indexed strictly by location, then the constraints on allocated fixed factors follow $\sum_i z^i \leqslant \boldsymbol{k}$ as in the case where a farmer's tractor services or labor must be allocated across plots so as not to exceed availability. If sub-technologies are indexed strictly by time and represent the stages of production, then these constraints follow $z^i \leqslant \boldsymbol{k}$, $i = 1, \ldots, m$, as in the case where tractor services or operator labor are available with recurrence in each successive time period. Where $\boldsymbol{Z} = \{z^1, \ldots, z^m\}$ represents the allocation of fixed factors to m sub-technologies with both spatial and/or temporal detail, the constraints on allocated fixed factors may be represented generally and compactly by $\boldsymbol{C}\boldsymbol{Z} \leqslant \boldsymbol{K}$, where $\boldsymbol{C}$ is a matrix of ones and zeros with full row rank, and $\boldsymbol{K}$ is a vector that duplicates $\boldsymbol{k}$ for each time period (or modifies it as appropriate if capacities differ by time period) and is thus a function of $\boldsymbol{k}$. For example, the first several rows of $\boldsymbol{C}\boldsymbol{Z} \leqslant \boldsymbol{K}$ may constrain the total allocation of labor and machinery services in time period 1 by $\boldsymbol{k}$, the next several rows may do the same for time period 2, and so on.[27] With this framework, one way of

in Markovian frameworks. We forgo the generality of this representation for simplicity of exposition but note the importance of this generality for empirical applications following Section 2.1.

[27] As for purchased variable inputs, the presence of any public fixed factor inputs is suppressed from the explicit notation for simplicity of presentation. For example, each sub-technology production set might be described more completely by $(\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(\boldsymbol{x}^0, z^0, z^i, \boldsymbol{\varepsilon})$ where both $\boldsymbol{x}^0$ and z^0 are vectors of public inputs and the detailed use of fixed factors is represented by $\boldsymbol{Z} = \{z^0, z^1, \ldots, z^m\}$.

describing the technology is[28]

$$
\begin{aligned}
&\{(y, x) \in \Im_{-i}(k, \varepsilon)\} \\
&\quad \equiv \{(y, x) \mid (y^i, x^i) \in \Im_i(z^i, \varepsilon),\ y = AY,\ x = BX,\ CZ \leqslant K\},
\end{aligned} \tag{43}
$$

where $\Im_{-i}(k, \varepsilon)$ represents the set of potential choices of aggregate output and input vectors, i.e., total amounts of physical outputs and inputs after aggregation over time and space.

We refer to descriptions of technology as on the left-hand side of (43) as reduced-form representations because the underlying structure on the right-hand side is solved out of the problem. Structured technologies can be represented by reduced-form production sets devoid of temporal and/or spatial detail as on the left-hand side of (43), but without the right-hand side structural detail the implications of policy instruments that impose limitations on specific y^i's, x^i's, or z^i's cannot be considered nor can preferences that value specific y^i's, x^i's, or z^i's such as peak operator labor. Furthermore, the specific production plan that attains any distinct $(y, x) \in \Im_{-i}(k, \varepsilon)$ is not apparent without the right-hand side detail in Y, X, and Z.

Alternatively, the technology can be represented completely by

$$
\begin{aligned}
&\{(Y, X, Z) \in \Im(k, \varepsilon)\} \\
&\quad \equiv \{(y^1, \ldots, y^m, x^1, \ldots, x^m, z^1, \ldots, z^m) \mid (y^i, x^i) \in \Im_i(z^i, \varepsilon), CZ \leqslant K\},
\end{aligned} \tag{44}
$$

where the selection of an element of the technology set prescribes the production plan completely. Also, the elements of the technology set excluded by any particular policy that limits inputs or outputs at specific times or locations can be clearly imposed on (44) but not on the left-hand side of (43).

We submit that the differences in (43) and (44) are fundamentally important. Clearly, if (43) is appropriate, then it substantially reduces the dimension of the efficient choice set. This is a welcome convenience for the study of some issues. However, serious errors can occur from use of (43) when the Aggregation Qualification Condition fails. We note that virtually all empirical applications of duality to agricultural production use the reduced-form representation on the left-hand side of (43) rather than the structural representation of (44). If the efficiency standards of (43) are inappropriate, the state of the empirical agricultural production literature must be seriously questioned. These

[28] While more general descriptions of technology structure with nonlinear relationships in place of A, B, and C are easily possible, such generalizations needlessly complicate the points made below without adding insight. We leave extension to these obvious cases to the reader. It should also be noted that the left-hand side of (43) is defined by k rather than $z = \sum_i z^i$ because the right-hand side of (43) embeds the determination of the z^i's and because the use of allocated fixed factors cannot be freely reallocated among the alternative time periods represented in $CZ \leqslant K$.

differences are best illuminated by defining several concepts of technical efficiency. We begin with the strongest technical efficiency concept imposed by (43). Corresponding formal definitions are given in the Appendix.

Reduced-form technical efficiency corresponds to operating on the efficient frontier of $\Im_{-i}$ defined by (43), which under continuity and monotonicity can be represented as a production possibilities frontier, $F^*(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) = 0$.[29]

Note that reduced-form efficiency is the typical concept of production efficiency and is defined in terms of aggregate inputs and outputs. An example is given by (11) for the case where allocated fixed inputs are not present (or are ignored) and production is conjoint. Several weaker concepts of output-oriented technical allocative efficiency can be defined depending on which allocations are considered: (i) allocation of purchased variable inputs, (ii) allocation of allocated fixed inputs, and/or (iii) allocation of production among sub-technologies. Each holds one vector of allocations fixed while optimizing another:

Fixed factor technical allocative efficiency holds for a production plan $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ if no other production plan $(\boldsymbol{Y}', \boldsymbol{X}, \boldsymbol{Z}')$ achieves more of at least one output with no less of others $(\boldsymbol{Y}' \nleqslant \boldsymbol{Y})$ without using more allocated fixed factors $(\boldsymbol{C}\boldsymbol{Z}' \leqslant \boldsymbol{C}\boldsymbol{Z})$.[30]

Variable input technical allocative efficiency holds for a production plan $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ if no other production plan $(\boldsymbol{Y}', \boldsymbol{X}', \boldsymbol{Z})$ achieves more of one output and no less of others $(\boldsymbol{Y}' \nleqslant \boldsymbol{Y})$ without using more purchased variable inputs $(\boldsymbol{B}\boldsymbol{X}' \leqslant \boldsymbol{B}\boldsymbol{X})$.

Output technical allocative efficiency holds for a production plan $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ if no other production plan $(\boldsymbol{Y}', \boldsymbol{X}', \boldsymbol{Z}')$ achieves more of one aggregate output and no less of other aggregate outputs $(\boldsymbol{A}\boldsymbol{Y}' \nleqslant \boldsymbol{A}\boldsymbol{Y})$ without using more purchased variable inputs $(\boldsymbol{B}\boldsymbol{X}' \leqslant \boldsymbol{B}\boldsymbol{X})$ or more allocated fixed factors $(\boldsymbol{C}\boldsymbol{Z}' \leqslant \boldsymbol{C}\boldsymbol{Z})$.

In these definitions, technical allocative efficiency is differentiated from standard concepts of allocative efficiency that depend on prices and correspond to operating at tangencies of price lines with physical trade-off possibilities, e.g., the tangency of the output price line with the PPF. These concepts of technical allocative efficiency are weaker

[29] Following Chambers (1988, p. 261), the PPF is defined by $F^*(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) = y_1 - f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) = 0$ where $\boldsymbol{y} = (y_1, \boldsymbol{y}_{-1})$ and $y_1 = f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) = \max\{y_1 \mid (\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\}$. The term "efficient frontier" in this context refers to the upper right-hand frontier of the set of possible aggregate outputs and purchased variable inputs, $(\boldsymbol{y}, -\boldsymbol{x})$.

[30] Consistent with Principle 3, it should be noted in the definition of fixed factor technical allocative efficiency that $z = \sum_i z^{i'} \leqslant z = \sum_i z^i$, which does not include temporal detail, is not an appropriate condition in place of $\boldsymbol{C}\boldsymbol{Z}' \leqslant \boldsymbol{C}\boldsymbol{Z}$, which imposes allocated fixed factor constraints by time period. The reason is that allocated fixed inputs cannot be freely reallocated among time periods as might be the case for purchased variable inputs. The implication is that data on aggregate flows of machinery services are not appropriate for modeling production if decisions are made in a reality of intraseasonal constraints on machinery service flows.

and merely correspond to operating on the physical trade-off frontier. The reason for using these weaker definitions is to identify a measure of technical efficiency that is sufficiently independent of prices, policy, and behavior for various circumstances.

To verify that these concepts of technical allocative efficiency are weaker than reduced-form efficiency, consider a somewhat stronger concept of fixed factor technical allocative efficiency:

Feasible fixed factor technical allocative efficiency holds for a production plan $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ if no other production plan $(\boldsymbol{Y}', \boldsymbol{X}, \boldsymbol{Z}')$ achieves more of at least one output with no less of others $(\boldsymbol{Y}' \nleqslant \boldsymbol{Y})$ given feasibility of allocated fixed factors $(\boldsymbol{C}\boldsymbol{Z}', \boldsymbol{C}\boldsymbol{Z} \leqslant \boldsymbol{K})$.

Reduced-form efficiency is obtained by combining this stronger concept of feasible fixed factor technical allocative efficiency with variable input and output technical allocative efficiency. Thus, all of the above technical allocative efficiency concepts are implied by reduced-form efficiency.

The potential inappropriateness of reduced-form technical efficiency can thus be studied by considering potential inappropriateness of the technical allocative efficiency concepts. Each of the various forms of technical allocative efficiency (which are implied by corresponding standard concepts of price-based allocative efficiency) may be inconsistent (i) with plausible preferences, (ii) with restrictions imposed by government policies, and (iii) even with profit maximization in absence of policy restrictions. In particular, if the allocated fixed inputs, variable inputs, or outputs that are aggregated over time and space by physical characteristics do not satisfy the Aggregation Qualification Condition, then the respective technical allocative efficiency concept is inappropriate. This condition implies that aggregation is not appropriate in cases where (i) generic input prices are heterogeneous over space and time and disaggregated prices are unobserved, (ii) allocation-specific government policy controls apply, (iii) allocation-specific ex post adjustments respond to unanticipated conditions, or (iv) behavioral criteria more general than profit maximization have allocation-specific considerations (such as risk aversion with allocation-specific risk effects of inputs).

These failures occur because technical allocative efficiency employs a standard of minimizing physical aggregates of fixed allocated resources and/or variable inputs, and/or maximizing physical aggregates of outputs under the assumption of equal marginal productivities and possibly also equal marginal risk effects. If these standards are inappropriate due to, say, spatial or temporal price variation, then the assumption of equal marginal productivities is typically not satisfied. If profit maximization fails due to risk aversion, then the assumption of equal marginal risk effects may not be satisfied. When the Aggregation Qualification Condition is not satisfied, a weaker concept of technical efficiency can be satisfactory.

Feasible disaggregated input-output efficiency corresponds to operating on the efficient frontier of $\Im$ where $\Im$ is defined by (44).

Feasible disaggregated input-output efficiency implies operating on the upper right-hand frontier of the set of possible disaggregated outputs and purchased variable inputs, $(\boldsymbol{Y}, -\boldsymbol{X})$, given feasible allocations of fixed factors. It is likely the strongest concept of technical efficiency devoid of policy or behavioral content among those above. Similarly, feasible disaggregated input-output efficiency is also likely the strongest concept of efficiency clearly independent of (typically unobserved) spatial and/or temporal price distributions among those above.

If the producer has preferences over leisure as well as profit (and thus, implicitly, over operator labor), then feasible disaggregated input-output efficiency is also inappropriate because the producer may choose a level of operator labor inside the associated fixed allocated input constraint. In this case, fixed factor technical allocative efficiency, which does not require exhausting constraints, may hold while feasible fixed factor technical allocative efficiency fails. For this case, the following weaker concepts of sub-technology and structural technical efficiency are appropriate. If the Aggregation Qualification Condition fails for allocated fixed factors, then these may be the strongest appropriate concepts of technical efficiency.

Sub-technology efficiency corresponds to operating on the efficient frontier of $\Im_i$, which under continuity and monotonicity can be represented as $F_i(\boldsymbol{y}^i, \boldsymbol{x}^i, z^i, \boldsymbol{\varepsilon}) = 0$.[31]

Structural technical efficiency corresponds to operating on the efficient boundary of all sub-technologies simultaneously which under continuity and monotonicity can be represented as

$$F(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon}) \underset{m}{\equiv} \begin{bmatrix} F_1(\boldsymbol{y}^1, \boldsymbol{x}^1, z^1, \boldsymbol{\varepsilon}) \\ \vdots \\ F_m(\boldsymbol{y}^m, \boldsymbol{x}^m, z^m, \boldsymbol{\varepsilon}) \end{bmatrix} \underset{m}{=} 0. \tag{45}$$

Note that, to avoid confusion, a subscript is added to these equalities to denote vector dimensionality of the equalities.

Intuitively, sub-technology efficiency is appropriate for any objective function that is monotonically increasing in the elements of $\boldsymbol{y}^i$, a highly plausible condition. The same can be said for the more expansive concept of structural technical efficiency. Note that feasible input-output technical efficiency is obtained by adding feasible fixed factor technical allocative efficiency to structural technical efficiency. While this stronger concept of technical efficiency appears highly plausible because production plans that violate fixed production resource constraints are not feasible, the example above where the producer has preferences with respect to use of particular fixed resource service flows such as operator labor gives an example where it is not.

[31] Specifically, define $F_i(\boldsymbol{y}^i, \boldsymbol{x}^i, z^i, \boldsymbol{\varepsilon}) = y_i^i - f_i(\boldsymbol{y}_{-i}^i, \boldsymbol{x}^i, z^i, \boldsymbol{\varepsilon})$ where $\boldsymbol{y}_{-i}^i$ is the $\boldsymbol{y}^i$ vector with y_i^i deleted and $y_i^i = f_i(\boldsymbol{y}_{-i}^i, \boldsymbol{x}^i, z^i, \boldsymbol{\varepsilon}) = \max\{y_i^i \mid (\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon})\}$. The term "efficient frontier" in this context refers to the upper right-hand frontier of the set of possible $(\boldsymbol{y}_i, -\boldsymbol{x}_i)$.

Some important points evident from this discussion are as follows (see Appendix Section A.1 for sketches of proofs).

PROPOSITION 1. *Preferences, policies, and spatial and/or temporal price variation can affect allocation under technologies with structure, which renders typical concepts of technical allocative efficiency (and thus standard concepts of price-based allocative efficiency) inapplicable.*

PROPOSITION 2. *For technologies with structure (technologies composed of sub-technologies), reduced-form technical efficiency, i.e., operating on the aggregate production possibilities frontier, is not necessarily consistent with profit maximization.*

PROPOSITION 3. *If there is at least one allocated fixed (variable) input and the output(s) of at least two sub-technologies are strictly monotonic in that input, then structural technical efficiency is not equivalent to output technical allocative efficiency nor fixed factor (variable input) technical allocative efficiency.*

4.2. The purpose of characterizing production efficiency

Presumably, the major objective of characterizing production efficiency is to decompose the producer's problem usefully into technical, behavioral, and policy components. Without this decomposition, microeconomic models of supply and demand cannot predict or analyze the effects of changes in technology and/or policy. According to the Aggregation Qualification Condition, decomposition whereby the first stage is strictly technical may be correctly accomplished only under particular circumstances. Suppose that Aggregation Qualification Condition A.2 holds prior to imposing any constraints and that all functionals of the decision variables subject to distinct policy controls or behavioral preferences are retained as decision variables in the second stage (Principle 6). For example, if (i) the producer is a profit maximizer, (ii) government policy controls are fully expressed by $(\boldsymbol{y}, \boldsymbol{x}) \in G$, and (iii) prices are identical among sub-technologies ($\boldsymbol{p} = \boldsymbol{p}_i, \boldsymbol{r} = \boldsymbol{r}_i$ for $i = 1, \ldots, m$), then the first stage defined by (43) is devoid of policy and behavioral content and is sufficient to reflect the full generality of the remaining decisions in a second-stage problem of the form $\pi = \max_{y,x}\{\boldsymbol{py} - \boldsymbol{rx} \mid (\boldsymbol{y}, \boldsymbol{x}) \in G \cap \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\}$.

By comparison, if either a full expression of government policy controls requires $(\boldsymbol{Y}, \boldsymbol{X}) \in G$ or prices (market or shadow) are not identical among sub-technologies, then the first stage must retain the detail of (44). In the case of profit maximization, the corresponding second stage is then of the form $\pi = \max_{Y,X}\{\boldsymbol{PY} - \boldsymbol{RX} \mid (\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}) \in G \cap \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})\}$. For example, policy might constrain the use of a particular input such as a fertilizer, pesticide, or tillage practice differently depending on the proximity of an individual plot to surface water resources. Similar conclusions apply to aggregations over time and space as well as over sub-technologies.

If behavioral alternatives to profit maximization are admitted, then additional generalities must be preserved by the first stage. For example, under risk aversion some functional must be included describing how risk is fully determined by second-stage decisions. If this functional is affected differently according to which sub-technology an input is applied (e.g., if fertilizer affects risk on a corn field differently than it affects risk on a wheat field), then distinction in the input vector must be carried to the second stage if behavioral content is to be avoided in the first stage.

Alternatively, suppose the production problem is decomposed so that the first stage is not purely technical but also admits policy constraints or behavioral preferences. For example, where the first-stage decision set G_{-i} is defined by

$$\begin{aligned}&\{(\boldsymbol{y},\boldsymbol{x}) \in G_{-i}(\boldsymbol{k},\boldsymbol{\varepsilon})\}\\ &\quad \equiv \{(\boldsymbol{y},\boldsymbol{x}) \mid (\boldsymbol{y}^i,\boldsymbol{x}^i) \in G \cap \Im_i(z^i,\boldsymbol{\varepsilon}),\ \boldsymbol{y}=\boldsymbol{A}\boldsymbol{Y},\ \boldsymbol{x}=\boldsymbol{B}\boldsymbol{X},\ \boldsymbol{C}\boldsymbol{Z}\leqslant \boldsymbol{K}\},\end{aligned}$$

the description of technology carried to the second stage by $(\boldsymbol{y},\boldsymbol{x}) \in G_{-i}(\boldsymbol{k},\boldsymbol{\varepsilon})$ clearly carries policy content. If so, then determination of whether $(\boldsymbol{y},\boldsymbol{x})$ choices are on the frontier of $G_{-i}(\boldsymbol{k},\boldsymbol{\varepsilon})$ has little to say about technical efficiency. Policy-constrained behavior may be technically inefficient. Further, statistical tests about the structure of $G_{-i}(\boldsymbol{k},\boldsymbol{\varepsilon})$ have little to say about the structural properties of technology. Decisions may be on the frontier of $G_{-i}(\boldsymbol{k},\boldsymbol{\varepsilon})$ but yet be technically inefficient. Finally, accurate estimation of the second stage decision equations for this problem will be of little value for analyzing the effects of changes in policies that affect G_{-i}. The parameters of such equations will be dependent on the policies that determine $G_{-i}(\boldsymbol{k},\boldsymbol{\varepsilon})$ and thus inappropriate for analyzing alternative policies following the Lucas critique. The important point of this discussion is summarized by the following proposition (see Appendix Section A.2 for a sketch of the proof).

PROPOSITION 4. *If the (first-stage) description of technology depends on policy or behavior, then statistical tests regarding efficiency and structural characteristics do not necessarily apply to technology nor are estimated (second-stage) models useful for policy analysis.*

Proposition 4 points out a problem that applies to many agricultural production studies in the literature to date because of crop- and/or spatial- and/or time-specific policy instruments associated with commercial agricultural and environmental policy. Of course, for estimation, sufficient variation in policy instruments and variables affecting preferences must be observed to facilitate identification and distinction of technical relationships from policy- or preference-induced relationships. In other words, the problem is not whether inputs or outputs are aggregated but that the dimensions and configuration of A, B, and C are likely wrong in most empirical studies. "Wrong" in this context means that either the Aggregation Qualification Condition is violated or that observed data are inadequate for identification because of excessive detail. With limited data, distinction may not be possible.

These considerations motivate the definition of criteria for technical allocative efficiency that satisfy policy- and behavior-independence where feasible aggregation is undertaken to conserve degrees of freedom for estimation. That is, aggregation of $\boldsymbol{Y}$, $\boldsymbol{X}$, and $\boldsymbol{Z}$ is appropriate within groups that have common prices and that enter the policy and preference calculus as aggregates. Suppose the technology choice is summarized by $(\boldsymbol{y}^*, \boldsymbol{x}^*, z^*) = \boldsymbol{H} \cdot (\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}) \in R_+^{n*}$ where $n^* < n$ and $\boldsymbol{H}$ is a full row rank aggregator matrix of ones and zeros similar to $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$. If $\boldsymbol{H}$ preserves distinction for all input and output quantities that have distinct prices, distinct policy controls, distinct ex post adjustment possibilities, or distinct behavioral preferences and implications, then the full flexibility of the technology for responding to price, policy, and behavioral concerns is preserved by the first stage of a production problem that satisfies

$$\begin{aligned}&\{(\boldsymbol{y}^*, \boldsymbol{x}^*, z^*) \in \Im^*(\boldsymbol{k}, \boldsymbol{\varepsilon})\}\\&\quad\equiv \{\boldsymbol{H} \cdot (\boldsymbol{y}^1, \ldots, \boldsymbol{y}^m, \boldsymbol{x}^1, \ldots, \boldsymbol{x}^m, z^1, \ldots, z^m) \mid (\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon}), \boldsymbol{C}\boldsymbol{Z} \leqslant \boldsymbol{K}\}.\end{aligned} \tag{46}$$

Two additional definitions facilitate this distinction.

An aggregation $(\boldsymbol{y}^*, \boldsymbol{x}^*, z^*)$ *is policy- and behavior-relevant* if it satisfies the Aggregation Qualification Condition. The efficient frontier of $\Im^*$ defined in (46) thus provides a standard of technical allocative efficiency independent of policy and behavior.

An aggregation $(\boldsymbol{y}^*, \boldsymbol{x}^*, z^*)$ *is policy- or behavior-dependent* if it does not satisfy the Aggregation Qualification Condition. The efficient frontier of $\Im^*$ defined in (46) thus does not provide a standard of technical allocative efficiency independent of policy and behavior.[32]

Aside from the extreme assumptions of functional separability, the Aggregation Qualification Condition implies distinction must be preserved for all input and output quantities that have distinct prices, distinct ex post adjustments, distinct policy controls, distinct ex post adjustment possibilities, and distinct behavioral preferences and implications. According to Proposition 4, this concept of policy- and behavior-relevance must be satisfied in order to investigate technical efficiency or properties of the technology in a meaningful and relevant way.

For the remainder of this paper, we emphasize that imposing efficiency concepts in the definition of the technology is inappropriate whenever it is incongruent with policy- and behavior-relevance. For example equating marginal rates of technical substitution or marginal value products across allocated fixed factors such as land is inappropriate if (i) agricultural policy restrictions impose crop-specific acreage limitations, (ii) environmental policy imposes land-use restrictions or acreage-specific conservation measures,

[32] Note that policy- and behavior-relevance is the opposite of policy- and behavior-dependence. A representation is policy- and behavior-relevant if it applies regardless of the particular policy or behavior in effect.

(iii) the farmer has crop-specific preferences, or (iv) the farmer values leisure and different crops have different returns to operator labor. In each of these cases, policy- or behavior-related considerations cause implicit prices to vary across allocation variables. Similarly, if allocation-specific prices of variable inputs are unobserved then similar marginal conditions may be inappropriate for variable input allocations. Because some aggregation is required for practical and tractable representation of most production problems, we assume from this point forward that the disaggregated description of the production problem includes all aggregation that is policy- and behavior-relevant. That is, the notation of (44) will be assumed to represent a policy- and behavior-relevant description of the production problem as in (46) where asterisks are dropped for convenience.[33]

4.3. Functional representation of technology

A common practice in production economics has been to switch readily from set theoretic notation to smooth functional representation of technology for econometric purposes upon assuming continuity and monotonicity. Technologies with structure can be analyzed somewhat more generally using the dual set theoretical framework developed by Chambers, Chung, and Färe (1996). Related empirical applications are possible along the lines of Chambers and Just (1989). However, the bulk of our presentation uses functional notation to facilitate accessibility for the broader agricultural economics profession and to relate better to common empirical practices. In practice, the switch to a functional representation is typically made arbitrarily with little regard for the structure of production.

From its earliest consideration in economics, multi-output efficiency has been characterized by single-equation multi-output production functions of the form

$$F(\boldsymbol{Y}, \boldsymbol{X}) = 0. \tag{47}$$

Samuelson (1967) argued that such forms are very general and can be derived from a host of underlying production functions and optimal conditions. Some have taken these arguments to mean that (47) can contain a host of distinct functions and conditions of the form, $F_i(\boldsymbol{Y}, \boldsymbol{X}) = 0$, $i = 1, \ldots, m$, which are imposed simultaneously by, say,

$$F(\boldsymbol{Y}, \boldsymbol{X}) \equiv \sum_{i=1}^{m} \left[F_i(\boldsymbol{Y}, \boldsymbol{X})\right]^2 = 0 \tag{48}$$

[33] Accordingly, the policy- and behavior-relevant description of a sub-technology corresponds to

$$\left\{(\boldsymbol{y}^{i*}, \boldsymbol{x}^{i*}) \in \Im_i^*(\boldsymbol{z}^{i*}, \boldsymbol{\varepsilon})\right\} \equiv \left\{\boldsymbol{H} \cdot (0, \ldots, 0, \boldsymbol{y}^i, 0, \ldots, 0, \boldsymbol{x}^i, 0, \ldots, 0, \boldsymbol{z}^i, 0, \ldots, 0) \mid (\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(\boldsymbol{z}^i, \boldsymbol{\varepsilon})\right\}$$

and the physical sums of allocations are represented by $\boldsymbol{c}^* \boldsymbol{z}^* = \boldsymbol{C}\boldsymbol{Z}$ where $\boldsymbol{z}^* = (\boldsymbol{z}^{1*}, \ldots, \boldsymbol{z}^{n_z *})$.

[Mittelhammer et al. (1981)].[34] If so, then a simple direct specification that completely determines n_y outputs from n_x inputs following n_y distinct scalar relationships,

$$y_i = f_i(\boldsymbol{X}), \quad i = 1, \ldots, n_y, \tag{49}$$

can be represented by (48) where $F_i(\boldsymbol{Y}, \boldsymbol{X}) = y_i - f_i(\boldsymbol{X})$, $i = 1, \ldots, n_y$. For practical purposes, however, the representation in (48) is not useful because it yields $\partial F/\partial \boldsymbol{Y} = 0$ and $\partial F/\partial \boldsymbol{X} = 0$ whenever $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$. Such single-equation forms are not consistent with many standard manipulations of production problems and, in particular, violate the standard convexity assumption of duality (see Appendix, Section A.3, for details).

PROPOSITION 5. *Single-equation implicit production functions cannot represent technologies with structure in ways that lend themselves to standard assumptions of duality or other standard manipulations of production problems using Lagrangians, Kuhn–Tucker conditions, or the implicit function theorem.*

To represent technologies with structure, ambiguity about how many functional conditions are imposed by the technology must be resolved. In spite of the potential generality of (47), common single-equation specifications of technology render representations such as (48) and (49) inapplicable. For example, Klein's (1947) multi-product generalization of the Cobb–Douglas production function,

$$F(\boldsymbol{Y}, \boldsymbol{X}) = y_1 y_2^{\delta} - A x_1^{\alpha_1} x_2^{\alpha_2},$$

or, indeed, any single-equation form that is separable in inputs and outputs,

$$F(\boldsymbol{Y}, \boldsymbol{X}) = h(\boldsymbol{Y}) - g(\boldsymbol{X}) = 0,$$

cannot represent structures such as (48) and (49). Alternatively, the structure in either (48) or (49) is represented unambiguously by (45) without requiring $\partial F/\partial \boldsymbol{Y} = 0$ and $\partial F/\partial \boldsymbol{X} = 0$ when $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$. As a result, the standard assumptions and manipulations identified in Proposition 5 are not excluded. Thus, the form in (45) is used below.

To see that similar implications apply in the dual approach, consider the special input nonjointness case of (49) where $y_i = f_i(\boldsymbol{x}^i)$, $i = 1, \ldots, n_y$. With the dual approach of

[34] Samuelson (1967) is somewhat ambiguous on this point. Clearly, Samuelson interprets (11) as giving the maximum amount of any one output given amounts of all inputs and all other outputs. In other words, (11) characterizes the production possibilities frontier associated with a given input vector $\boldsymbol{X}$. This interpretation alone, however, does not identify whether more than one condition may be required to reflect, say, a technology where 2 outputs follow a particular by-product relationship in addition to a typical implicit production possibilities frontier relationship. We note also that Samuelson also uses standard Lagrangian techniques which, as shown below, are not applicable for representations such as (11) that combine multiple conditions.

Chambers, Chung, and Färe (1996), which is sufficiently general to handle technologies with structure, the input distance function becomes[35]

$$D_I(\boldsymbol{y},\boldsymbol{x}) = \max\left\{\min_i\{D_I^i(y_i,\boldsymbol{x}^i)\}\;\middle|\;\sum_i \boldsymbol{x}^i = \boldsymbol{x},\, \boldsymbol{y} = (y_1,\ldots,y_{n_y})\right\},$$

where $D_I^i(y_i,\boldsymbol{x}_i)$ is the input distance function associated with sub-technology i. Each of the sub-technology distance functions corresponds to one of the production relationships in Lau's (1978) development illustrated in Equation (13). The left-hand side distance function cannot reflect the structural characteristics of the multiple sub-technology distance functions on the right-hand side. Thus, multiple functions are required to fully represent multiple sub-technology structure in dual as well as primal approaches.

4.4. Structural versus reduced-form representation of technology

A typical view that has followed from the duality emphasis on PPFs has been that the input vector determines the output possibilities set rather than a specific output vector [e.g., Chambers (1988)]. Indeed, this view is appropriate as a reduced-form representation where allocations of aggregate inputs both spatially and temporally as well as among production activities represent a corresponding structural determination of the output mix. The contrast between these reduced-form and structural concepts of technology are analogues of reduced form and structure in econometric models. Each has its appeal. However, unlike econometric models, the just-identified case occurs here only when technologies have trivial structure (i.e., either there are no sub-technologies or the sub-technologies have mutually exclusive sets of inputs and outputs). Otherwise, if technical inefficiency is measured in the reduced form in (43), one cannot identify whether the sub-technologies are inefficient, or whether the inefficiency comes from allocative inefficiency, or whether the Aggregation Qualification Condition is violated.

At the basic level of management decision making, the manager must control decisions that determine which mix of outputs is produced (for given magnitudes of uncontrollable factors). Otherwise, the typical tangency conditions of price lines with production possibilities frontiers cannot be attained by deliberate choice, in which case the

[35] To see this result, let the input requirement set for each sub-technology be represented by $\upsilon_i(y_i) = \{\boldsymbol{x}^i \mid (y_i,\boldsymbol{x}^i) \in \Im_i\}$ in which case the overall input requirement set is $\upsilon(\boldsymbol{y}) = \sum_i \upsilon_i(y_i)$. The input distance function is

$$\begin{aligned} D_I(\boldsymbol{y},\boldsymbol{x}) &= \max\left\{\alpha > 0 \;\middle|\; \boldsymbol{x}/\alpha \in \sum_i \upsilon_i(y_i)\right\} \\ &= \max\left\{\alpha > 0 \;\middle|\; \boldsymbol{x}^i/\alpha \in \upsilon_i(y_i),\ i = 1,\ldots,n_y;\ \sum_i \boldsymbol{x}_i = \boldsymbol{x}\right\}. \end{aligned}$$

Since α can only be feasible if $\boldsymbol{x}^i/\alpha \in \upsilon_i(y_i)$, $i = 1,\ldots,n_y$, it must satisfy $\alpha \leqslant D_I^i(y_i,\boldsymbol{x}^i)$, $i = 1,\ldots,n_y$, where $D_I(\boldsymbol{y},\boldsymbol{x}) = \max\{\min_i\{D_I^i(y_i,\boldsymbol{x}^i)\} \mid \sum_i \boldsymbol{x}^i = \boldsymbol{x},\, \boldsymbol{y} = (y_1,\ldots,y_{n_y})\}$.

bulk of multi-output production theory is inapplicable. Thus, given convexity of sub-technology production sets (quasi-concavity of the associated production functions), no generality is lost by assuming that input decisions for the underlying sub-technologies determine the output vector with structural technical efficiency uniquely for given magnitudes of uncontrollable factors. This determination is made by the allocation of inputs both spatially and temporally among sub-technologies as in the following axiom.

The Fundamental Axiom of Multi-output Production: For given magnitudes of uncontrollable factors, the complete vector of input decisions uniquely determines the technically efficient vector of outputs.

The complete vector of decisions includes all spatial and temporal allocations of both purchased variable inputs and allocated fixed factors including allocations of recurring service flows from firm-owned resources (as well as any non-allocated fixed factors not represented explicitly here). From a practical standpoint, this axiom simply implies that a farmer can determine the production mix of, say, corn and soybeans (subject to uncontrolled exogenous and random forces such as weather, pest infestations, illness, variations in work quality, and errors in applying decisions) by making all available input decisions including allocations of land, fertilizer, pesticides, labor, machinery services, etc., both spatially and temporally. This axiom, in effect, simply divides all forces affecting production into two groups – controlled and uncontrolled – and assumes that these two groups of forces determine production uniquely. In other words, once all production decisions are made and uncontrolled forces act, the producer is not left with an *ex post* ability to adjust the output mix.

As realistic and innocuous as this axiom seems, it has been the focus of an implicit ongoing debate [compare the PPF-based duality literature to Mittelhammer et al. (1981); Just et al. (1983); Shumway et al. (1984)]. Adopting this Fundamental Axiom, however, immediately leads away from the PPF and toward sub-technology characterizations of technology. The basic points are that (i) PPFs represent a reduced-form summary of the implications of a more basic representation of technology involving spatial and temporal allocations of inputs among sub-technologies, whereas (ii) the structures of sub-technologies have potential implications for production analysis and related policy analysis.

The Fundamental Axiom permits the use of sub-technology representations to examine implications. An immediate implication of the Fundamental Axiom is that sub-technology efficiency under continuity and monotonicity can be represented by

$$y^i \underset{k_i}{=} f_i\left(x^i, z^i, \varepsilon\right), \tag{50}$$

where f_i is a multivariate function determining the complete vector of output quantities and k_i is the number of outputs that are not identically zero in sub-technology i. Thus, if $F_i(y^i, x^i, z^i, \varepsilon)$ in (45) includes multiple implicit relationships along the lines of (48) then the representation in (50) is assumed to make these relationships explicit yielding non-trivial derivatives for purposes of standard manipulations of production problems. Thus, whereas $F_i(\boldsymbol{y}^i, \boldsymbol{x}^i, \boldsymbol{z}^i, \boldsymbol{\varepsilon})$ in (45) represents a sub-technology by its reduced-form PPF, Equation (50) represents the structure of a sub-technology explicitly with the multiple equations of a vector valued function where appropriate. In turn, structural technical efficiency is represented by[36]

$$\boldsymbol{Y} \underset{\sum_i k_i}{\equiv} \begin{bmatrix} \boldsymbol{y}^1 \\ \vdots \\ \boldsymbol{y}^m \end{bmatrix} \underset{\sum_i k_i}{=} \begin{bmatrix} f_1(\boldsymbol{x}^1, \boldsymbol{z}^1, \boldsymbol{\varepsilon}) \\ \vdots \\ f_m(\boldsymbol{x}^m, \boldsymbol{z}^m, \boldsymbol{\varepsilon}) \end{bmatrix} \underset{\sum_i k_i}{\equiv} f(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon}), \tag{51}$$

which specifies the complete vector of output quantities of the firm. The number of non-identically-zero outputs in $\boldsymbol{Y}$ is thus $\sum_i k_i$. For example, the simple Samuelsonian case of (11) where each sub-technology produces a unique single output yields $k_i = 1$ for all i, which is the case of full input nonjointness.[37] In addition, fixed allocated inputs must obey $\boldsymbol{CZ} \leqslant \boldsymbol{K}$ so that feasible disaggregated input-output efficiency is given by

$$\boldsymbol{Y} \in \left\{ f(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon}) \mid \boldsymbol{CZ} \leqslant \boldsymbol{K};\ \nexists \boldsymbol{Z}' \ni f(\boldsymbol{X}, \boldsymbol{Z}', \boldsymbol{\varepsilon}) \nleq f(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon})\ \&\ \boldsymbol{CZ}' \leqslant \boldsymbol{K} \right\}.$$

A brief example can illustrate the richness of (51). Suppose a farm has two sub-technologies: one for production of wheat and one for cow-calf production. The wheat sub-technology may produce both grain and straw (both are outputs of harvesting). The cow-calf operation produces both bull calves and heifer calves. The farmer faces decisions of how much labor to allocate to each of the two sub-technologies but each sub-technology has two outputs. Thus, each $f_i(\boldsymbol{x}^i, \boldsymbol{z}^i, \boldsymbol{\varepsilon})$ is two-dimensional whereas $\boldsymbol{Y}$ is four-dimensional.

The characterization of technology by (50) and (51) employs possibly numerous equations to describe a firm's technology compared to the more traditional single equation reduced-form description of a PPF as in (47). The purpose of the next several sections is to show that the multitude of equations in (50) and (51) have much to say about the structure and properties of technology that can only be uncovered by examining sub-technologies. Furthermore, typical econometric efficiency considerations suggest advantages to estimation of as much of this structure as data availability permits. By comparison, single-equation reduced-form PPF estimation of (47) or duality based supply-demand estimation based on a PPF characterization of the production set suffers from econometric inefficiency [Mundlak (1996)].[38]

[36] Again, the reader should bear in mind the possible presence of public inputs, which are suppressed for notational convenience. For example, each $f_i(\boldsymbol{x}^i, \boldsymbol{z}^i, \boldsymbol{\varepsilon})$ in (50) and (51) might be described more completely by $f_i(\boldsymbol{x}^0, \boldsymbol{z}^0, \boldsymbol{x}^i, \boldsymbol{z}^i, \boldsymbol{\varepsilon})$ where $\boldsymbol{x}^0$ and $\boldsymbol{z}^0$ are public variable and fixed inputs, respectively.

[37] Note that from this point forward $\boldsymbol{Y}$ is assumed to include only outputs that are not identically zero. Thus, $n_y = \sum_i k_i$ where the dimensions of the $\boldsymbol{y}^i$'s are not all the same. Nevertheless, $\boldsymbol{y} = \boldsymbol{AY}$ is assumed to represent aggregation by physical attributes across time, space, and sub-technologies.

[38] The underlying econometric principle has been developed by, for example, Dhrymes (1973), who showed that more efficient estimates of reduced forms are obtained by estimating the underlying structure.

To examine the structural representation of technology for empirical purposes, however, requires careful specification of allocation decisions, by-product relationships, and related concepts of controllability and rank of technologies. These issues are discussed in detail in Appendix Section A.4 but are outlined intuitively here to facilitate remaining discussion. While relatively little may be known regarding specific functional forms, the dimensions of allocations, by-product relationships, controllability and rank can typically be determined quite well on the basis of actual farming practices and information available from production scientists. For example, purchased inputs for crop production must typically be allocated among plots and time (i.e., among specific trips over specific plots with specific farm equipment). When one crop is grown at a time as in developed agriculture, a major decision is how much land to allocate to each crop in each growing season as well as how much seed, fertilizer, and pesticides to apply per unit of land on each crop or plot and when to apply it. These simple observations determine much about the structure of production and the dimension of the producer's decision vector. Additionally, some products like cotton and cottonseed or bull calves and heifer calves are produced in tandem. Cottonseed is not produced as a by-product of wheat production nor is cottonseed produced independent of cotton. These relationships, in effect, reduce the producer's flexibility in choosing the decision vector in substantive ways.

With this background, we say a sub-vector of decisions is locally controllable if the producer is free to vary any part of the vector by a small amount in any direction. The existence of by-product relationships reduces the producer's controllability in choosing output mixes. Assuming continuity and monotonicity and partitioning the Y vector as $Y = (\widetilde{Y}, \widehat{Y})$ where $\widehat{Y} \in R_+^{n_b}$, the outputs in $\widehat{Y}$ are called by-products of $\widetilde{Y}$ under technology $\Im$ if there exists a non-trivial relationship $\widehat{Y} \underset{n_b}{=} g(\widetilde{Y}, \varepsilon)$ that uniquely determines $\widehat{Y}$. For example, in the case of a wheat sub-technology, if grain and straw are produced in fixed proportions aside from uncontrolled forces, then either output may be considered a by-product of the other. If, on the other hand, the choice of inputs determines the mix of grain and straw output, then the two equations describing their production differ substantively (in rank) and neither output is a by-product of the other. In the fixed proportions case, straw is typically considered the by-product because of its lower value, but this designation is price-dependent and not appropriate in a pure description of technology.

Although this partitioning of the output vector is not unique, results in the Appendix show that the dimension of controllability is determined uniquely by the rank of the technology, i.e., $n_a = \operatorname{rank}(f_X, f_Z)$ where subscripts of f denote differentiation, $n_y = n_a + n_b$, and $\widetilde{Y} \in R_+^{n_a}$. The useful purpose of defining the rank of a technology is to determine how many equations are required to represent it empirically and, in turn, what the dimension is of the investigation required to determine production efficiency. This information is also necessary to determine econometric efficiency (how many equations are required to represent structure fully). This framework gives a constructive way to test for the existence of by-products. For example, nonparametric estimates of f can be used to test for the rank of (f_X, f_Z) following Cragg and Donald (1997).

A similar issue of controllability applies to inputs. For example, constraints imposed on allocated fixed inputs by fixed resources of the firm limit the producer's flexibility in making input choices because they must satisfy $\boldsymbol{CZ} \leqslant \boldsymbol{K}$. Because the individual constraints contained therein may or may not be locally binding, suppose that $\widetilde{\boldsymbol{C}}$ and $\widetilde{\boldsymbol{K}}$ represent subsets of the rows of $\boldsymbol{C}$ and $\boldsymbol{K}$, respectively, corresponding to locally binding constraints. Then $\widetilde{\boldsymbol{C}}\boldsymbol{Z} \underset{n_c}{=} \widetilde{\boldsymbol{K}}$ summarizes all locally binding constraints on allocated fixed inputs. Assuming without loss of generality that $\boldsymbol{C}$ includes no redundant restrictions, $\widetilde{\boldsymbol{C}}$ has rank n_c so that $\widetilde{\boldsymbol{C}}\boldsymbol{Z} \underset{n_c}{=} \widetilde{\boldsymbol{K}}$ can be solved uniquely for $\widehat{\boldsymbol{Z}} \underset{n_c}{=} h(\widetilde{\boldsymbol{Z}}, \boldsymbol{K})$ where $\boldsymbol{Z}$ is partitioned as $\boldsymbol{Z} = (\widetilde{\boldsymbol{Z}}, \widehat{\boldsymbol{Z}})$, $\widetilde{\boldsymbol{Z}} \in R_+^{n_z - n_c}$, and $\widehat{\boldsymbol{Z}} \in R_+^{n_c}$. Obviously, an arbitrary fixed input allocation vector is not fully controllable unless no fixed inputs are limiting. That is, even though there are n_z allocated fixed input decisions, only $n_f = n_z - n_c$ of them are freely controllable.

The presence of allocated fixed input constraints explains why the responses of seemingly independent production activities may appear dependent.[39] That is, input constraints across production activities can induce jointness between them even when the production activities are fully nonjoint [Shumway et al. (1984)]. This is true whether the constraints result from allocated fixed inputs or other sources such as policy parameters imposed on a firm (such as water use restraints, acreage controls or pesticide standards). For this reason, statistical testing for restrictions on the input space appears advisable. For example, nonparametric tests can be used to determine controllability of inputs. Alternatively, such tests can be based upon the existence of a non-trivial h function by regressing $\widehat{\boldsymbol{Z}}$ on $\widetilde{\boldsymbol{Z}}$ for hypothesized partitions of $\boldsymbol{Z}$.

With this background, a standard form for the structure of technology is useful.

A canonical form for the local structure of multi-output technologies consists of[40]
(i) the controllable production technology,

$$\widetilde{\boldsymbol{Y}} \underset{n_a}{=} \tilde{f}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon}), \tag{52}$$

(ii) the byproduct relationships,

$$\widehat{\boldsymbol{Y}} \underset{n_b}{=} g(\widetilde{\boldsymbol{Y}}, \boldsymbol{\varepsilon}), \tag{53}$$

[39] In other words, a producer's decisions result in a specific vector of output quantities – a production possibilities surface. Even if aggregate inputs produce a production possibilities surface (as in some dual developments), an additional decision must be made to determine a particular point on the production possibilities surface if the standard multi-output profit maximization theory is appropriate.

[40] Note that Equations (52) and (53) are jointly equivalent to Equation (51) in the sense that either can be solved for the other. Equations (52) and (53) correspond to a representation that solves the $\boldsymbol{x}^i$'s and z^i's out of as many individual equations of (51) as possible. Note, however, that $n_z \geqslant m$ because each sub-technology must have at least one substantive production relationship.

and (iii) the binding input restrictions,

$$\widehat{\boldsymbol{Z}} \underset{n_c}{=} h(\widetilde{\boldsymbol{Z}}, \boldsymbol{K}), \tag{54}$$

where $\widetilde{\boldsymbol{Y}}$ and $\widetilde{\boldsymbol{Z}}$ are locally controllable, $\boldsymbol{Z} = (\widetilde{\boldsymbol{Z}}, \widehat{\boldsymbol{Z}})$, and the Jacobians of $\tilde{f}$, g, and h have full row rank.

Appendix Section A.4 derives the following proposition.

PROPOSITION 6. *Every technology that satisfies*
(i) *the Fundamental Axiom of Multi-output Production,*
(ii) *continuity, and*
(iii) *differentiability*
can be characterized locally in canonical form.

This characterization of technology is convenient for applying the various measures of efficiency defined above. For example, each of the individual equations in the controllable technology corresponds to sub-technology efficiency. The equations in the controllable technology plus the by-product relationships correspond to structural technical efficiency. Combining the controllable technology with by-product relationships and input restrictions corresponds to feasible disaggregated input-output efficiency. This description of technology is thus policy- and behavior-relevant.

Before turning to applications of this framework, we consider one remaining generality of the controllable technology. Some inputs may be allocated so that a distinct portion is applied to each production activity, i.e., is allocated to a specific time and location within that activity. However, other distinct input applications may have positive marginal products in more than one output equation as in the case of a public input. This gives rise to joint output relationships that may connect some output equations in the controllable technology. Appropriate modeling of such relationships is essential for proper investigation of issues such as diversification. Indeed, such modeling is essential for understanding issues of scope and scale (to the extent that economies of scope depend on scale). Chavas (2001) mentions several examples of processes that determine economies of scope such as nitrogen fixation, pest control, and crop-livestock interactions. As explained above, single-equation representations cannot convey useful understanding of such multi-dimensional interaction. Economies of scope and diversification may depend on many factors including public inputs as well as binding input restrictions and by-product relationships [even under profit maximization as in Pope (1976)] in addition to typical risk aversion explanations. We submit that the approach of description versus technical detail in Chavas' review is indicative of (i) the poor state of understanding of these issues and (ii) the lack of true explanation provided by PPF approaches.

4.5. The problem with unobservable decision variables

A typical problem for empirical analysis of production is that some variables are not observed. For example, a typical case in aggregate agricultural production is where temporal and spatial allocations of purchased variable inputs within growing seasons are not observed. This lack of data availability seems to motivate the focus on PPFs in typical production studies. As noted above, however, such approaches as typically practiced have not led to policy- and behavior-relevant representations of technology.

Here we investigate the feasible approach to estimation and identification of policy- and behavior-relevant technology when some decisions are not observed. An appropriate approach in this case is to solve unobserved variables out of the system in (52)–(54). We argue that this is the only feasible approach if the resulting representation of technology is to be policy- and behavior-relevant, i.e., truly a representation of nothing more than technology.[41]

Consider, for example, the case where an individual output is not observed. Then the corresponding individual equation in (52) or (53), which explains that output, is not observable and must be dropped from any estimable system. Alternatively, suppose an individual input variable is not observed. If the input appears among the binding constraints, then one of the binding constraints can be solved for that input variable; that result can then be used to substitute for the unobserved variable in (52). If a constraint with known coefficients is solved, such as a simple aggregation constraint for an allocated fixed input, then values of the variable can be calculated from the others to substitute into (52). When a constraint that has unknown coefficients is solved for an unobserved input variable, then an estimable form must be substituted for the unobserved variable in (52). This process may complicate estimation of (52) because numerous parameters may appear in individual equations after substitution thus requiring more observations for identification. But estimation is possible in principle and no justifiable alternative is apparent.

If additional input variables are unobserved, remaining input constraints can be used one-by-one if the unobserved variables appear among the remaining input constraints. Otherwise, one of the remaining equations in (52) must be solved for the unobserved input variable. Such a relationship may include an unobservable error term and add to the stochastic complications of estimation of remaining relationships. But again, estimation is possible in principle with sufficient numbers of observations.

Continuing inductively with this approach which is applicable under the assumptions of the Implicit Function Theorem obtains the following proposition.

[41] We remind the reader that all policy- and behavior-relevant aggregations are assumed to be included in the problem representation at this point as in (46). If further policy- and behavior-relevant aggregation is possible, then simple Pareto efficiency among those allocations may yield additional policy- and behavior-relevant structural equations.

PROPOSITION 7. *Aside from by-product relationships, which do not characterize the effects of inputs, the maximum number of non-redundant observable equations that can characterize purely technological relationships is equal to the number of observable controllable outputs* (n_a) *plus the number of purely technological binding input constraints* (n_c) *minus the number of unobservable decision variables (if non-negative).*

An immediate implication is as follows.

COROLLARY 1. *Aside from by-product relationships, if the number of observable controllable outputs plus the number of purely technological binding input constraints minus the number of unobservable decision variables is non-positive, then no purely technological relationship is estimable. Any estimable relationship between outputs and decision variables must embody non-technical relationships imposed on the observed data, for example, by behavioral and policy criteria.*

4.6. The typical agricultural production problem

In typical agricultural production problems involving multiple outputs, farmers choose not only a production possibilities set, but choose a production point in that set. In typical dual representations, the choice of a production possibilities set is made by choosing an aggregate input vector. The concept of efficiency based on profit maximization is used to restrict choices to the frontier of the production possibilities set. Then the choice of a point on the frontier is represented implicitly by the choice of an output vector. Outputs, however, are *ex post* observations of the production problem and thus do not characterize the actual process of production or decision making. Choice of an output vector implicitly involves determining other choices relating to how the aggregate input vector is used. These implicit choices typically involve allocation of aggregate inputs over production activities, i.e., over space and time. For example, in examples of basic economic principles with two production processes using the same input, the production possibilities frontier depends on the aggregate amount of input available, and the choice of a point on the frontier is determined by how much of the aggregate input is allocated to one production process versus the other.

With the allocation of fixed inputs discussed by Shumway, Pope, and Nash (1984), farmers must determine how much land to allocate to each crop or how many tractor hours and hours of labor to allocate to each plot, etc. Variable inputs must also be allocated among production activities (plots) and times of application to choose a point on the production possibilities frontier [Just et al. (1983)]. For example, farmers must generally determine how much fertilizer, pesticides, and labor to apply to each crop and plot as well as how much of each variable input to use in the aggregate. We do not contend that all inputs must be allocated but argue that at least some allocation decisions are required to determine the mix of outputs in most agricultural production problems. Such decisions must be considered part of the detailed $\boldsymbol{X}$ and $\boldsymbol{Z}$ vectors that determine outputs.

Allocations of both fixed and variable inputs are typically treated as unobserved in common applications of duality. For example, a common specification of the profit function in multi-output production problems is

$$\max_{y,x}\{\boldsymbol{py} - \boldsymbol{rx} \mid (\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\}$$

where $\boldsymbol{p}$ is an output price vector taken to apply to the entire growing season, $\boldsymbol{r}$ is an input price vector taken to apply uniformly to the entire aggregate of input quantities used, $\boldsymbol{x}$ represents choice of aggregate input quantities without regard to temporal or spatial allocations, and $\Im_{-i}$ represents possible choices of aggregate inputs and aggregate outputs with available technology and fixed inputs [e.g., Shumway (1983); Ball and Chambers (1982); Weaver (1983)]. This specification yields maximum profit as a function of $\boldsymbol{p}$, $\boldsymbol{r}$, $\boldsymbol{k}$, and $\boldsymbol{\varepsilon}$ if Aggregation Qualification Condition A.2 holds.

Alternatively, the production framework in (52)–(54) reveals all detailed allocations and decisions that demonstrate how outputs are determined *ex ante* (aside from uncontrollable factors). Clearly, if variable or fixed allocation decisions must be made and allocations are ignored as in a typical dual framework, then the allocations must either be considered unobservable or the econometric efficiency that can be attained with full structural estimation is lost. We note, however, that there is nothing about the modern dual approach that prevents this more detailed empirical investigation. For example, Chambers and Just (1989) use a dual approach to investigate allocations of an observed allocated fixed factor. Similar techniques can also be used to investigate price differences among allocated quantities where they apply following the general theoretical framework of Chambers, Chung, and Färe (1996).

Undoubtedly, some of the elegance and simplicity of the typical reduced form (dual or primal) approach is lost by considering a full structure for production technology as in (52)–(54). However, unobserved or ignored allocations of inputs have dramatic implications for estimation of technology as the following proposition demonstrates (see the Appendix for the proof).

PROPOSITION 8. *If* (i) *two or more inputs (whether variable or fixed) must be allocated among sub-technologies,* (ii) *the allocations are unobserved or ignored in estimation, and* (iii) *the number of controllable outputs is less than the number of allocated inputs times the number of sub-technologies, then no purely technological relationship other than by-product relationships is estimable. In particular, no purely technological relationship is estimable in the input nonjointness case of* (11).

The conditions of Proposition 8 appear to be broadly applicable and cast doubt on the ability to estimate purely technological relationships from aggregate data. Furthermore, a much stronger result applies if physical inputs must be allocated over space and time within sub-technologies. Proposition 8 focuses only on allocations of inputs over sub-technologies. The problem is that many allocations of inputs over crops as well as space and time are generally not recorded in aggregate data. For example, aggregate public data are generally not available on the allocation among crops (or plots) of variable inputs such as labor or of allocated fixed inputs such as tractor hours. We note, however, that allocation of land among crops is usually available and has not been exploited by typical dual production studies. Thus, the failure to utilize allocation data cannot be blamed entirely on data unavailability.

While we note that allocation data for land among crops is generally available and unutilized (and was a prominent subject of study prior to the duality revolution), the principle of Proposition 8 also suggests that specific assumptions may be required for its use. For example, suppose that land allocations are observed but that other allocations are unobserved as in the following corollary to Proposition 8 (see the Appendix for a proof).

COROLLARY 2. *If* (i) *three or more inputs (whether variable or fixed) must be allocated among sub-technologies,* (ii) *only the allocations of one input are observed and used in estimation, and* (iii) *the number of controllable outputs is less than the number of observed allocated inputs times the number of sub-technologies, then no purely technological relationship other than by-product relationships is estimable. In particular, no purely technological relationship is estimable in the input nonjointness case of* (11).

The implication of Corollary 2 is that if only land allocations are observed and at least two other input allocations are unobserved, then purely technological relationships are generally unobservable unless specific restrictions are imposed on the technology. For example, one could assume that other allocated inputs are applied in fixed proportions with land. Since such assumptions must be imposed to observe technology, it follows that hypotheses such as fixed input proportions among allocated inputs cannot be rejected with observable data under the conditions of Corollary 2.

In Mundlak's (2001) review (this Handbook), he characterizes the modern dual approach as having not delivered its promised benefits in the empirical analysis of production. We agree but argue that the criticism should not be of the potential of the dual approach but of the failure to pursue understanding of the structure of production. It is simply the *typical practice* of duality (the focus on the PPF alone) that has been limiting. We argue that this practice is, at least in part, a self-imposed limitation of the profession. But it is also, in part, a result of public data limitations. With proper consideration, some hypotheses that have been entertained in the literature may not be testable with available data.

4.7. Estimable relationships among inputs and outputs

To examine additional implications of the criticism in Section 4.6, the production problem can be further characterized by determination of the decision vectors $\boldsymbol{X}$ and $\boldsymbol{Z}$ according to some behavioral criterion given available technology.

A behavioral criterion is a rule sufficient to determine production decisions in $\boldsymbol{X}$ and $\boldsymbol{Z}$ uniquely given the full description of technology in (52)–(54).

For example, in the case of profit maximization (ignoring uncontrollable factors for purposes of illustration), the problem is

$$\max_{Y,X,Z}\left\{PY - RX \mid \widetilde{Y} \underset{n_a}{=} \tilde{f}(X, Z, \varepsilon),\ \widehat{Y} \underset{n_b}{=} g(\widetilde{Y}, \varepsilon),\ \widehat{Z} \underset{n_c}{=} h(\widetilde{Z}, K)\right\}. \tag{55}$$

After substitution of the constraints and assuming well-behaved technology, this problem generates a set of $n_o = n_x + n_f$ first-order conditions for optimization of the form

$$\zeta(X, \widetilde{Z}, P, R, K, \varepsilon) \underset{n_o}{=} 0. \tag{56}$$

These n_o relationships together with the n_c binding constraints in (54) uniquely determine X and Z. The remaining $n_a + n_b$ relationships in (52) and (53), in turn, determine Y following the Fundamental Axiom of Multi-output Production. This framework clearly differentiates the relationships defining technology in (52)–(54), which appear as constraints in (55), and the behavioral relationships in (56).

Alternatively, if all allocations of input and output quantities have identical prices over space and time in the production cycle (the typical assumption), the decision problem can be represented by[42]

$$\max_{y,x}\left\{py - rx \mid y \underset{n_y^*}{=} Af\big(X, \widetilde{Z}, h(\widetilde{Z}, K), \varepsilon\big),\ x \underset{n_x^*}{=} BK\right\}, \tag{57}$$

where $y \in R_+^{n_y^*}$, $x \in R_+^{n_x^*}$. Several points are important in comparing this problem to typical analysis of aggregate production problems. First, the decision variables in this problem, after substituting constraints, are not simply y and x but rather X and $\widetilde{Z}$. Thus, the number of first-order conditions is $n_x + n_f$. In this set of first-order conditions, the price of one input or output can be arbitrarily normalized (set to 1) because of homogeneity of supplies and demands in prices, which follows from (57). Then, in principle, $n_y^* + n_x^* - 1$ of the $n_x + n_f$ first-order conditions can be solved for the non-normalized prices and substituted into the remaining first-order conditions obtaining $n_o^* = n_x + n_f - n_y^* - n_x^* + 1$ relationships expressed solely in terms of y, X, and $\widetilde{Z}$, say,

$$\zeta(y, X, \widetilde{Z}, K, \varepsilon) \underset{n_o^*}{=} 0. \tag{58}$$

Typically, the number of relationships in (58) is large when there are allocations because n_x^* represents the number of aggregate variable inputs whereas n_x represents the number of variable factor allocation variables summed over all variable inputs and, similarly, n_y^*

[42] Note that $y \underset{n_y^*}{=} Af(X, Z, \varepsilon) = A \cdot (\widetilde{Y}, \widehat{Y})$ where $\widetilde{Y} \underset{n_a}{=} \tilde{f}(X, Z, \varepsilon)$ and $\widehat{Y} \underset{n_b}{=} g(\widetilde{Y}, \varepsilon)$. Also note that $Z = (\widetilde{Z}, \widehat{Z}) = (\widetilde{Z}, h(\widetilde{Z}, K))$.

represents the number of aggregate outputs whereas n_f represents the number of fixed factor allocation variables summed over all production activities.

Note, however, that the relationships in (58) cannot be purely technological relationships even though they include only input and output quantities because they are derived from first-order conditions based on the behavioral criterion. Clearly, these relationships include more information than reflected in the pure statement of technology, $\boldsymbol{y} \underset{n_y^*}{=} \boldsymbol{A} f(\boldsymbol{X}, \widetilde{\boldsymbol{Z}}, h(\widetilde{\boldsymbol{Z}}, \boldsymbol{K}), \boldsymbol{\varepsilon})$ because the rank of first-order conditions leading to (58) is $n_x + n_f - n_y^* - n_x^* + 1$ (if they can be solved uniquely for all decisions) whereas the reduced-form statement of technology in (57) has at most rank n_y^*.

From these results, estimation and hypothesis testing based on first-order conditions, including all dual methodology, does not necessarily reveal information about technology. For example, apparent nonjointness or apparent nonseparability suggested by estimates of any subset of these relationships may simply reflect an interaction among variables induced by the maintained behavioral hypothesis (see the example of Section 3.4). These results are summarized by Proposition 9 (see the Appendix for a proof).

PROPOSITION 9. *Under the conditions of Proposition* 8, *no hypotheses about the structure of technology are testable. All observable relationships of inputs and outputs are policy- or behavior-dependent.*

All hypothesis tests on the structure of agricultural technology relating to jointness and separability of which we are aware are made in problems where the presence of two or more inputs with unobserved allocations cannot be ruled out. On the basis of Proposition 9, the associated conclusions are invalid.

4.8. The "technology" estimated with standard dual applications

In standard dual applications assuming differentiability, technology is implicitly represented by a scalar PPF relationship of the form derived in (44), $F(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) \underset{1}{=} 0$, i.e., one involving only aggregate inputs and outputs [e.g., Shumway (1983); Ball and Chambers (1982); Weaver (1983)]. From the $n_a + n_b + n_c$ relationships describing technology and firm-controlled resource constraints in (52)–(54) and the $n_o = n_x + n_f$ first-order conditions in (56), exactly one such relationship involving only aggregate inputs and outputs is observable in general. To find this relationship starting from (55), one must aggregate not only the outputs as in (57) but also eliminate all the allocations. In other words, after obtaining (57), the $n_o^* = n_x + n_f - n_y^* - n_x^* + 1$ relationships among all input decisions and outputs that are devoid of prices can be used together with $\boldsymbol{y} \underset{n_y^*}{=} \boldsymbol{A} f(\boldsymbol{X}, \widetilde{\boldsymbol{Z}}, h(\widetilde{\boldsymbol{Z}}, \boldsymbol{K}), \boldsymbol{\varepsilon})$ and $\boldsymbol{x} \underset{n_x^*}{=} \boldsymbol{B}\boldsymbol{X}$ to solve for all $n_x + n_f$ allocations, which are then substituted into a remaining single condition.

The resulting single condition is of the form $F^*(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) \underset{1}{=} 0$ and is regarded as characterizing the production possibilities frontier. However, under the conditions of

Propositions 7, 8, and 9, this relationship may be determined at least partially by the behavioral criterion. For example, if one of the inputs is an allocated fixed input or involves allocation of a variable input that does not have the same price across all locations or times, then any single-equation representation of technology using only aggregate variables will be policy- or behavior-dependent.

An interesting question is whether knowledge of this frontier can reveal information about the structure of technology for which it is commonly used to test. In general, the answer is no. At best (when the Aggregation Qualification Condition holds), it can only answer very limited questions about the reduced-form structure. The illustrative example of Section 3.4 demonstrates clearly the difference between structure and reduced form.

4.9. Congruent modeling of econometric errors and inefficiencies

Thus far, we have largely ignored uncertainty issues related to agricultural production. In reality, agricultural production is highly subject to random forces such as weather and pests. The presence of such forces in worldwide agricultural production causes prices also to be random and unpredictable – particularly because of long lags between commencement of production and realization of output. Adding unanticipated stochastic variation in production reveals further problems with typical practices.

To illustrate, note that the profits in (55) after substituting constraints can be represented by $\boldsymbol{P}f(\boldsymbol{X},\widetilde{\boldsymbol{Z}},\boldsymbol{K},\boldsymbol{\varepsilon})-\boldsymbol{R}\boldsymbol{X}$ where

$$\begin{aligned} f(\boldsymbol{X},\widetilde{\boldsymbol{Z}},\boldsymbol{K},\boldsymbol{\varepsilon}) &= \big[\tilde{f}(\boldsymbol{X},\widetilde{\boldsymbol{Z}},h(\widetilde{\boldsymbol{Z}},\boldsymbol{K}),\boldsymbol{\varepsilon}),g(\tilde{f}(\boldsymbol{X},\widetilde{\boldsymbol{Z}},h(\widetilde{\boldsymbol{Z}},\boldsymbol{K}),\boldsymbol{\varepsilon}),\boldsymbol{\varepsilon})\big] \\ &= \big\{(\widetilde{\boldsymbol{Y}},\widehat{\boldsymbol{Y}}) \mid \widetilde{\boldsymbol{Y}}=\tilde{f}(\boldsymbol{X},\boldsymbol{Z},\boldsymbol{\varepsilon}),\widehat{\boldsymbol{Y}}=g(\widetilde{\boldsymbol{Y}},\boldsymbol{\varepsilon}),\widehat{\boldsymbol{Z}}=h(\widetilde{\boldsymbol{Z}},\boldsymbol{K})\big\}. \end{aligned} \tag{59}$$

Prices and production disturbances are assumed to be random at the time decisions are made and, for simplicity in this section, the behavioral criterion is assumed to be expected profit maximization. The expected profit function and resulting demands and allocations are

$$\pi(\boldsymbol{I},\boldsymbol{K})=\max_{\boldsymbol{X},\widetilde{\boldsymbol{Z}}}\big\{\mathrm{E}_I[\boldsymbol{P}f(\boldsymbol{X},\widetilde{\boldsymbol{Z}},\boldsymbol{K},\boldsymbol{\varepsilon})-\boldsymbol{R}\boldsymbol{X}]\big\},$$

$$\big[\boldsymbol{X}^*(\boldsymbol{I},\boldsymbol{K}),\widetilde{\boldsymbol{Z}}^*(\boldsymbol{I},\boldsymbol{K})\big]=\operatorname*{argmax}_{\boldsymbol{X},\widetilde{\boldsymbol{Z}}}\big\{\mathrm{E}_I[\boldsymbol{P}f(\boldsymbol{X},\widetilde{\boldsymbol{Z}},\boldsymbol{K},\boldsymbol{\varepsilon})-\boldsymbol{R}\boldsymbol{X}]\big\},$$

respectively, where $\boldsymbol{I}$ represents information (e.g., a subjective distribution) upon which the producer's expectations of $\boldsymbol{P}$, $\boldsymbol{R}$, and $\boldsymbol{\varepsilon}$ are based.

Assuming mean expected prices are included among $\boldsymbol{I}$, differentiation of the profit function with respect to them obtains demands, $\boldsymbol{X}=\boldsymbol{X}^*(\boldsymbol{I},\boldsymbol{K})$, by the envelope theorem consistent with Hotelling's lemma.[43] Chambers and Just (1989) demonstrate how

[43] Depending on the stage of production in a dynamic representation, some of the random disturbances may have already been realized in which case $\boldsymbol{I}$ can include some actual values of some elements of $\boldsymbol{\varepsilon}$.

the allocation equation specifications, $\widetilde{\boldsymbol{Z}} = \widetilde{\boldsymbol{Z}}^*(\boldsymbol{I}, \boldsymbol{K})$, can also be derived consistent with sub-technology profit function specifications. Because prices and output are random, however, simple differentiation of the profit function does not generally obtain consistent output supply specifications following Hotelling's lemma. Rather, substituting input demands and allocations into (59) yields the actual output supplies,

$$\boldsymbol{Y} = \boldsymbol{Y}^*(\boldsymbol{I}, \boldsymbol{K}, \boldsymbol{\varepsilon}) = f(\boldsymbol{X}^*, \widetilde{\boldsymbol{Z}}^*, \boldsymbol{K}, \boldsymbol{\varepsilon}),$$

which have expectation $\mathrm{E}_I(\boldsymbol{Y}) = \overline{\boldsymbol{Y}}^*(\boldsymbol{I}, \boldsymbol{K}) = \mathrm{E}_I[f(\boldsymbol{X}^*, \widetilde{\boldsymbol{Z}}^*, \boldsymbol{K}, \boldsymbol{\varepsilon})]$. This specification generally differs from the derivative of the profit function with respect to mean output prices because of correlation among output prices and quantities, and nonlinearities of output in the production disturbance.

Two often-overlooked problems arise in subjecting this framework to estimation. First, the need to treat allocations differently than variable input demands is typically ignored assuming their prices can be represented implicitly as constants across space and time. Since this practice was criticized above, we abstract from the case with allocations for the remainder of this section because most readers are more familiar with notation that ignores allocations.

The second typically-overlooked problem relates to stochastic specification for estimation. Because input demands are derived by maximizing expected profits rather than actual profits, random variation is removed, leaving the resulting specification devoid of the random disturbances necessary for econometric purposes. The typical practice has been to append arbitrarily an econometric disturbance vector, say $\boldsymbol{\delta}$, to the vector of demand equations so the estimated specifications follow $\boldsymbol{X} = \boldsymbol{X}^*(\boldsymbol{I}, \boldsymbol{K}) + \boldsymbol{\delta}$. Alternatively, the profit function has been treated as a deterministic problem in mean prices, say $\overline{\boldsymbol{P}} = \mathrm{E}(\boldsymbol{P})$ and $\overline{\boldsymbol{R}} = \mathrm{E}(\boldsymbol{R})$, so that application of Hotelling's lemma obtains

$$\begin{aligned}
&\overline{\pi}(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K}) = \max_{\boldsymbol{X}}\{\overline{\boldsymbol{P}} f(\boldsymbol{X}, \boldsymbol{K}, \boldsymbol{\varepsilon}) - \overline{\boldsymbol{R}}\boldsymbol{X} \mid \boldsymbol{\varepsilon} = \boldsymbol{0}\},\\
&\boldsymbol{Y} = \overline{\boldsymbol{Y}}^*(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K}) = \partial\overline{\pi}(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K})/\partial\overline{\boldsymbol{P}},\\
&\boldsymbol{X} = \overline{\boldsymbol{X}}^*(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K}) = -\partial\overline{\pi}(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K})/\partial\overline{\boldsymbol{R}}.
\end{aligned}$$

This approach leaves each specification lacking an econometric disturbance for purposes of estimation. Typical practice has been to simply append disturbances to each relationship obtaining an estimation system of the form

$$\boldsymbol{Y} = \overline{\boldsymbol{Y}}^*(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K}) + \boldsymbol{\nu}, \qquad \boldsymbol{X} = \overline{\boldsymbol{X}}^*(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K}) + \boldsymbol{\delta}, \tag{60}$$

where $\boldsymbol{\nu}$ and $\boldsymbol{\delta}$ are vector-valued disturbances with zero expectations, e.g., $\boldsymbol{\varepsilon} = (\boldsymbol{\nu}, \boldsymbol{\delta})$.

A major problem with arbitrarily appending disturbances to a profit-function-based system is suggested by McElroy (1987) who initiated work on congruent specifications of input and output disturbances in the context of cost function estimation. The problem is that after arbitrarily appending disturbances to supply and demand specifications as in

(60), they no longer integrate back to the same underlying profit function.[44] In the spirit of McElroy, the profit function that yields via Hotelling's lemma both random supplies and factor demands as in (60) is of the form

$$\overline{\pi}(\overline{\boldsymbol{P}}, \overline{\boldsymbol{R}}, \boldsymbol{K}) = \max_{\boldsymbol{X}}\left\{\overline{\boldsymbol{P}}[f(\boldsymbol{X}, \boldsymbol{K}) + \boldsymbol{\nu}] - \overline{\boldsymbol{R}}(\boldsymbol{X} + \boldsymbol{\delta})\right\}. \tag{61}$$

The remaining problem with McElroy's approach is that disturbances are arbitrarily inserted to satisfy a particular theoretical convenience rather than to correspond to how random forces actually affect decision makers. In particular, the specification in (61) imposes additive errors in the production relationship and thus cannot admit risk-reducing or risk-increasing effects of inputs [Just and Pope (1978)]. Also, if the demand disturbances in (60) represent errors in optimization, then the specification in (61) is inappropriate because it has profits monotonically decreasing in errors, i.e., the decision maker is better off making large negative errors thus contradicting the concept of optimization.

To explore this problem further, an assessment of potential sources of error is instructive. Typically errors in agricultural production systems can be expected to arise from errors in decision making by farmers, random variation in uncontrolled forces such as weather that affect the production process, and errors in variables (measurement errors in data).[45] In each of these cases, the role of disturbances may be different. Yet typical a priori information hardly allows exclusion of one or the other.

The errors-in-optimization (EIO) case. To illustrate, if disturbances represent errors in decision making, then optimization errors can be simply appended to the profit-maximizing input levels as in the demand system in (60). In this case, however, the supply specification in (60) is no longer appropriate because the errors in input levels affect output following $\boldsymbol{Y} = f(\boldsymbol{X}^* + \boldsymbol{\delta}, \boldsymbol{K}, \boldsymbol{\nu})$, which surely differs from the $\boldsymbol{Y} = f(\boldsymbol{X}^*, \boldsymbol{K}, \boldsymbol{\nu})$ that generates the supply system in (60).[46]

[44] While all the discussion here is in terms of profit functions for simplicity, as illustrated by McElroy's work the same principles apply to cost and revenue function estimation as well.

[45] Another source of error in modeling is econometrician error. Perhaps these errors dominate all others but we refrain from a substantive discussion because (i) a major goal of this entire chapter is to improve econometric modeling, and (ii) the effects of modeling errors are dependent on the particular type of econometrician error and thus present too many alternatives to discuss here. For example, one possible econometrician error is made by assuming disturbances follow EIO (EIV) when EIV (EIO) applies. Another typical example is when, following the practice of modern duality theory, the econometrician specifies a profit function with little thought about the underlying technology because the profit function is not estimated but only used instrumentally to specify demands and supplies. Thus, the factor demands are obtained up to a random error but the profit function depends on this error because the supply or production depends on actual inputs. This is also an econometrician error.

[46] This result showing failure of Hotelling's lemma when input errors are transmitted to production functions is developed formally by Pope and Just (2000a).

The errors-in-variables (EIV) case. Suppose the disturbances represent errors in variables. For example, let $\boldsymbol{v}$ represent additive errors in measurement for $\boldsymbol{Y}$ and let $\boldsymbol{\delta}$ represent additive errors in measurement for $\boldsymbol{X}$, which are thus not part of the disturbances in $\boldsymbol{\varepsilon}$ that affect the true production problem. Then the specification in (60) is appropriate for the case where prices are nonstochastic. However, the profit function does not then follow (61) because the errors in $\boldsymbol{v}$ and $\boldsymbol{\delta}$ are not errors that actually affect decision makers and actual outcomes.

The errors-in-uncontrolled-conditions (EIU) case. If disturbances represent errors in uncontrolled conditions affecting the production process that are not observed until after decision making, then the representation of $\boldsymbol{\varepsilon}$ as an argument of f above is appropriate. In this case, the errors possibly interact with other input choices to alter production responses and marginal risk effects of inputs. For this problem, practical wisdom implies that the researcher is not free to choose an arbitrary representation (or point of insertion) of an ad hoc disturbance because the role of the disturbance is a substantive part of the economic problem. For this problem, a first-order Taylor series approximation of $\boldsymbol{Y}$ about $\boldsymbol{\varepsilon} = 0$ yields a Just–Pope production function, $\boldsymbol{Y} = f(\boldsymbol{X}, \boldsymbol{K}, \boldsymbol{0}) + f_{\boldsymbol{\varepsilon}}(\boldsymbol{X}, \boldsymbol{K}, \boldsymbol{0})\boldsymbol{\varepsilon}$, which provides a minimal yet tractable level of flexibility in the production specification.

Considering these three sources of error begs a discussion of which are most likely to be manifest in agricultural data. Based on the discussion in Section 2.4, it seems that the highly unpredictable effects of weather and pests inherent in the EIU case are most important to admit unless variables that reflect weather and pest conditions are included as measured variables rather than disturbances. While the other two sources of error seem less essential, they cannot be ruled out. Thus, the most conservative approach is to consider all three simultaneously. For example, one might start with a specification for $\pi(\boldsymbol{I}, \boldsymbol{K}) = \max_X\{\mathrm{E}_I[\boldsymbol{P} f(\boldsymbol{X}, \boldsymbol{K}, \boldsymbol{\varepsilon}) - \boldsymbol{RX}]\}$ and derive a specification for $\boldsymbol{X}^*(\boldsymbol{I}, \boldsymbol{K}) = \operatorname{argmax}_X\{\mathrm{E}_I[\boldsymbol{P} f(\boldsymbol{X}, \boldsymbol{K}, \boldsymbol{\varepsilon}) - \boldsymbol{RX}]\}$, which explicitly recognizes the potential randomness of prices. Then the supplies and demands might be estimated following

$$\boldsymbol{X} = \boldsymbol{X}^*(\boldsymbol{I}, \boldsymbol{K}) + \boldsymbol{\delta} + \boldsymbol{\zeta}, \qquad \boldsymbol{Y} = f(\boldsymbol{X}^* + \boldsymbol{\delta}, \boldsymbol{K}, \boldsymbol{\varepsilon}) + \boldsymbol{\nu}, \tag{62}$$

where $\boldsymbol{\delta}$ represents errors in optimization (which enter through the decisions and thus affect outputs through the technology that describes output responses to inputs), $\boldsymbol{\zeta}$ represents errors in measurement of inputs that do not affect observed outputs, $\boldsymbol{\varepsilon}$ represents uncontrolled inputs such as weather, and $\boldsymbol{v}$ represents errors in measurement of outputs.

Misspecification of the role of disturbances in production problems can cause considerable misinterpretation of data and empirical results. For example, Pope and Just (1996) developed what appears to be the first approach for consistent estimation of ex ante cost functions in the EIO case of stochastic production. Moschini (forthcoming) later showed that a different estimator was required for consistent estimation in the EIV case.[47] The contrast of these two papers and the bias and inconsistency resulting from using the wrong estimator demonstrates the importance of focusing carefully on the source of errors in production problems.

Moreover, these results underscore the need to develop robust estimation methods that can address a more general model such as in (62) for the case where the correct disturbance specification is not known a priori. Then statistical inference can be used to determine the correct error specification. In such an effort, Pope and Just (2000b) employ a specification similar to (62) by combining $\delta + \zeta$ into a single disturbance, say ξ, and then including $\lambda\xi$ as the embedded disturbance in place of δ in the production function. Their estimate of λ using aggregate U.S. agricultural data is .919 with a standard error of .322 implying that the pure EIV case is soundly rejected. The EIO case is not rejected even at the .001 level.

The results in this section are derived for the case where all decisions are made ex ante and all uncertainty is then resolved to determine final production and profit. More realistically, decision makers make some decisions, then observe some resolution of uncertainty. Then further decisions are made and further uncertainty is resolved, and so on until the end of a production cycle. Many of the principles in this section can be developed for this more complex and realistic case but space does not allow development here.

Based on the points in this section, we suggest that agricultural production economists have been far too cavalier about inserting disturbances in econometric specifications to facilitate estimation. The form in which disturbances enter has dramatic effects on estimated technology and on the statistical properties of estimators. The form in which disturbances enter can ultimately be answered by statistical inference. Until such answers are forthcoming and accepted, agricultural production estimation should seek for robust specifications or at least specifications consistent with accepted wisdom regarding the nature of agricultural production.

5. Other generalizations and empirical progress

Thus far, we have focused on the static production problem to demonstrate some fundamental principles and show how the structural implications and usefulness of agricultural production analyses depend on specification. In reality, the agricultural production problem is more complex. This section considers briefly several important additional frontiers of generalization: (i) dynamic interseasonal considerations related to physical and biological processes and investment, (ii) market uncertainties and characterization of information regarding them, (iii) implications of imposing behavioral criteria in agricultural production analyses, and (iv) changing technology with atomistic heterogeneity of adoption.

[47] Moschini (forthcoming) shows that the Pope and Just (1996) estimator is inconsistent in his EIV case. But Moschini's estimator is inconsistent for EIO cases covered by Pope and Just's estimator under risk aversion [Pope and Just (1998)]. The properties of Moschini's estimator clearly depend on risk neutrality.

5.1. Investment, asset fixity, biological growth, and fertility carryover

In general, agriculture presents a complicated problem of modeling production over time because of partial fixity and limited flexibility of physical production capital, the dynamic nature of biological capital (e.g., perennial plants and animals), accumulations of pest populations and resistance, and evolution of soil fertility and erosion. For example, machinery and buildings may be highly subject to asset fixity considerations [Chavas (2001)] but yet some assets may be highly flexible in application to production of a variety of crops. For example, for the most part the same machinery is used to cultivate wheat, sorghum, barley, oats, rice and most other small grains. Other types of equipment such as hay balers, milking equipment, and tomato harvesters may be highly output specific. Because of the dramatic role that physical capital plays in agricultural production, understanding investment in machinery, buildings, and land is likely the most important step to understanding agricultural production in a time series context. Specifically, lags and dynamic processes appear to be at the heart of understanding large-animal livestock and perennial crop production problems. Similarly, as Evenson (2001) states, lags and dynamic processes are also at the heart of understanding such broad policy questions as the economic aspects of R&D. Indeed, they are at the heart of understanding agricultural productivity.

Because new machinery can often be purchased with little delay, is highly lumpy (many farms have a single combine or high-horsepower tractor), and embodies unique technologies (as in the case of the tomato harvester and related color-sorting equipment), machinery investment may fit the putty-clay model well [Johansen (1972)] and require sophisticated discrete-continuous modeling of physical capital investment [see Just and Zilberman (1983) for a primitive such model]. For example, the problem of machinery replacement appears to be one of comparing the cost of new equipment less salvage value of old equipment (along with the higher productivity of new embodied technology) to the cost of continuing operations with old equipment given its higher repair costs and down time. Similar principles apply to constructing new buildings. The obstacle to analyzing these problems is that available data typically do not report machinery or building vintages (ages). Thus, for example, neither the relative technological improvements embodied in new machinery nor the salvage value of old equipment can be considered adequately in explaining machinery investment. Nor can repair costs be explained adequately by the machinery age distribution because it is unobservable.

Alternatively, development of biological capital (e.g., breeding stock, perennial stands of trees, or fertility content of soil) is constrained by biological and physical laws of nature and may require long lags for biological growth and adjustment. This is why such problems are typically modeled with difference equations that describe the number of animals or (acres of) plants that survive from one time period or age cohort to the next [see, e.g., Nerlove and Bessler (2001)]. With respect to these investments, costs and supply response may follow the traditional model of short- versus long-run cost curves [Viner (1931)]. Thus, knowledge of biological growth functions from the agricultural sciences may greatly improve empirical modeling of agricultural production and allow economists to focus estimation on features of the problem about which economic knowledge is weak (behavior and expectations).

To date, however, relatively little research has been devoted to understanding many of these longer-term problems of agricultural production. Several studies have examined asset fixity in agriculture both theoretically and empirically [e.g., Johnson (1956); Johnson and Quance (1972); Chambers and Vasavada (1983)]. Competing conceptual models with putty-putty, putty-clay, and clay-clay properties have been proposed [e.g., Johansen (1972); Fuss (1977)]. But little recent work has focused on fundamental empirical representations for some important classes of outputs. For example, the work of French, King, and Minami (1985) is essentially the last substantive work on perennial crops. Again, perhaps the major obstacle is lack of data regarding the age distribution of perennial crops. We note also that perennial crops and large-animal capital stocks have hardly been addressed with the modern tools of duality, in part because some of the elegance of duality is lost in doing so. For example, embedding a known biological process in a more complex production problem essentially requires a primal representation of part of the process.

Perhaps agricultural production economists occupied with simple dual approaches have been reluctant to tackle such problems. We suggest that more work is needed to enhance models for perennial crop and large-animal livestock production by combining known aspects of the age distribution evolution of biological capital with the advances in representations of the short-run production problem, dual or otherwise. For example, the canonical form of the short-run production problem remains as in (52)–(54) after adding τ subscripts to each variable to denote crop season (e.g., year). What must be added is the state equation,

$$\boldsymbol{K}_\tau = \kappa(\boldsymbol{K}_{\tau-1}, \boldsymbol{X}_{\tau-1}, \boldsymbol{Z}_{\tau-1}, \boldsymbol{\varepsilon}_{\tau-1}), \tag{63}$$

which describes buildings and machinery (by age and wear attributes), livestock and perennials (by age, size and health attributes), pest populations (by accumulated resistance attributes), soil quality (by accumulated fertility attributes, which depend on previous crop use and inputs), and accumulated debt and credit limitations (which depend on previous decisions to defer or accelerate repayment).[48]

A dual quasi-profit function may represent the short-run production problem if the state equation adequately represents interseasonal aspects of the problem. Such a representation of the production problem would not be complete, however, without adding a representation of how behavioral criteria determine implicit and explicit investment decisions, conservation behavior, crop rotation decisions, etc. (see Section 5.3). That is, behavioral criteria must be supplemented with long-term objective criteria that depend on $\boldsymbol{K}_\tau$; and the behavioral relationships in (56) must be supplemented accordingly

[48] While we consider only one lag in defining the state equation, as in any Markov process individual elements of the $\boldsymbol{K}$ vector can represent individual vintages of arbitrary age for any capital stock variable. Thus, the complete age distribution of various capital assets can be included.

with preferences that relate to choices in the stock equation. While the state equation in (63) may be complex, in some cases substantial knowledge of biological growth functions from the agricultural sciences can greatly improve empirical modeling and allow economists to focus estimation on features of the problem about which economic knowledge is weak (behavior and expectations).

5.2. *Expectations formation and information acquisition*

Representing production problems with price and output risk requires modeling both producer information (expectations) and producer behavior. A variety of approaches to modeling expectations have been used to model short-run (annual) production under uncertainty with some success [see the review by Nerlove and Bessler (2001)]. However, the problem of modeling expectations is more difficult in longer-term dynamic problems because (i) expectations are, in general, not directly observable, (ii) different producers may follow different approaches to forming expectations, and (iii) individual producers may switch among different information bases (or expectations mechanisms) depending on circumstances.

Modeling aggregate behavior is particularly difficult when producers' expectations are neither directly observable nor identical. The problem is that no data are typically available to explain even indirectly how expectations may be distributed among producers. However, Nerlove (1983) presents evidence of considerable heterogeneity in individual expectations. So, in many cases, the present state of knowledge simply does not reveal how vulnerable agricultural production analysis is to this problem.

Just and Rausser (1983) further suggest that rationality with costly information implies endogeneity of the operative expectation mechanism at the individual level. For example, some decision makers may find rational expectations require too much costly information in periods of stability compared to, say, naive expectations, but yet are worth the cost in periods of instability. Nerlove and Bessler (2001) also suggest that separation of expectations and optimizing behavior is not theoretically correct. Rather, the formation of expectations depends on the use to which expectations are put.

These considerations imply that agricultural economists are far from unraveling the role of expectations and the process of expectations formation particularly in heavily dynamic problems. The hope of doing so with aggregate data and current limitations on availability of firm-specific data appears dim [Nerlove (1983)]. Nevertheless, the role of information is becoming of increasing interest in this "age of information". More efforts are focusing on understanding individual information demand and vendor choice [Salin et al. (1998); Wolf et al. (forthcoming)]. We predict an increasing importance of these efforts in both aggregate and broad farm-specific models of agricultural production. For example, suppose the profits in (55) are represented using (59). Then the information choice problem might be represented as

$$\max_{I, X, \widetilde{Z}} \mathrm{E}_I\big[\boldsymbol{P} f(\boldsymbol{X}, \widetilde{\boldsymbol{Z}}, \boldsymbol{K}, \boldsymbol{\varepsilon}) - \boldsymbol{R}\boldsymbol{X} - c_I(\boldsymbol{I}) \mid \boldsymbol{I} \in \Phi\big],$$

where $\boldsymbol{I}$ represents a choice among various available sets of information in Φ, information is acquired with cost $c_I(\boldsymbol{I})$, and E_I represents a subjective assessment of expectations over $\boldsymbol{P}$, $\boldsymbol{R}$, and $\boldsymbol{\varepsilon}$ given information vector $\boldsymbol{I}$. The concept here is one of forming an expectation for the benefits of each information set when the actual information set is unknown and perhaps untried. In forming subjective assessments of the benefits of various information choices, a variety of experimentation and learning-by-doing possibilities arise akin to the problem of learning about new technologies in the technology adoption problem [see Sunding and Zilberman (2001)]. Clearly, much remains to be done to address these issues.

5.3. *Imposed versus revealed behavioral criteria*

Much of the traditional body of economic theory and empirical modeling, whether by input share equations, duality, or non-parametric estimators, implicitly imposes competitive profit maximization [see Mundlak (2001)]. This behavioral assumption apparently has been quite robust in the general economics literature for problems where certainty approximates reality in short-run production problems. Because of the importance of uncertainty in agriculture, however, this robustness may not apply. Most studies in agricultural economics that recognize this possibility have modeled agricultural production assuming either expected profit maximization or expected maximization of von Neuman–Morgenstern utility under risk aversion. Very little statistical testing against more general maintained behavioral hypotheses has been done, although a few studies have attempted to measure properties of risk aversion (absolute risk aversion, relative risk aversion, and partial risk aversion) and determine whether such measures are constant, increasing or decreasing. For example, Pope and Just (1991), Chavas and Holt (1996), and Bar-Shira, Just and Zilberman (1997) have attempted to determine the structure of risk preferences from actual production data, and Binswanger (1980, 1981) has attempted to determine the structure of risk preferences from revealed preferences for manipulated lotteries.

Outside of the expected utility hypothesis (which has expected profit maximization as a special case), however, few alternative behavioral hypotheses have been considered empirically. However, numerous studies have criticized the expected utility hypothesis on positive grounds because it fails to describe observed behavior [Kahneman and Tversky (1979); Moschini and Hennessy (2001); Chambers and Quiggin (1998)]. One approach is to introduce a different weighting of outcomes in different states following the generalized expected utility approach [Quiggin (1982); Machina (1987)]. While alternatives have been proposed, little comprehensive empirical evidence has been generated in direct comparative support of alternatives. Most recently Buschena and Zilberman (2000) have shown that generalized expected utility models lose much of their predictive dominance over expected utility when a heteroscedastic error structure is used. While the expected utility model has been criticized because it is informationally demanding [Moschini and Hennessy (2001)], generalized approaches tend to be even more informationally demanding at least when many states of nature are considered. An approach that reduces information demands on both decision makers and researchers is to rely on rules of thumb and recommendations of agricultural extension specialists. Just et al. (1990) show for Israeli agricultural data that such behavioral hypotheses tend to better fit observed behavior than the expected utility hypothesis.

Still other generalizations of behavior are appealing. Some of these are suggested by the multiple-goal programming models of farm management [e.g., Candler and Boehlje (1971); de Koning et al. (1995)]. For example, in a business where family labor appears to be a qualitatively different input for some tasks because of moral hazard considerations, farmers may prefer to trade off profit for labor depending on the amount of family labor needed to maximize (expected) profit. Thus, the utility function may have more arguments than profit that must be considered to explain behavior. Similarly, because of complex dynamics caused by biological production relationships, some farmers may prefer to trade off present profits for future wealth or long-term financial security. The large number of alternative objective criteria considered by Barry and Robison (2001) are evidence of such considerations. In recent decades, hobby farming has also become more important in which case farmers may have preferences for specific outputs (e.g., horses) or inputs (e.g., picturesque white fences).

With the possibility of such concerns in farmer preferences, we suggest that agricultural economists have been cavalier regarding behavioral criteria in most standard production studies. Forging ahead with the convenience and intuitive appeal of the profit maximization hypothesis in agricultural production analysis may be subject, at least for some problems, to the McCloskey (1998) criticism of searching under a lamppost for a lost wallet merely because the light is brighter there.

Evenson (2001) states that models of diffusion based on revealed preferences depend on properly sorting out technology, behavior, and expectations from one another. Barry and Robison (2001) emphasize the need for the study of agricultural production to support policy analysis by correctly sorting out (i) the role of constraints such as collateral limits or other credit rationing, (ii) the role of policy in altering behavior, (iii) the role of risk and risk preferences, and (iv) the role of intertemporal behavior. The central points of this paper further demonstrate that sorting out the properties and structure of production depends on sorting out technology from behavior. When behavioral criteria are imposed rather than determined empirically, models may be far from robust and results may fall far short of sorting out this crucial distinction. Moreover, imposing a false behavioral criterion may cause results to suggest a false representation of technology [Alston and Chalfant (1991); Smale et al. (1994)].

To suggest a framework in which observed data rather than assumptions are used to uncover behavioral criteria, recall the canonical representation of the production problem in (52)–(54). From a representation of technology that is complete and yet devoid of behavioral content, the description of the production problem (possibly an econometric system representing it) is properly closed by adding the behavioral relationships (and policy constraints) that determine choices given the technology. However, rather than assuming fixed and known relationships for this purpose, the relationships representing behavior can be made a matter of inference. Models that estimate a risk aversion coefficient (or risk preference structure) take a step in this direction but allow only one (or a few) estimated parameter(s).

Specifically, under the Fundamental Axiom of Multi-output Production, the full production system is closed by supplementing the purely technical equations in (52)–(54) with behavioral relationships such as (56). Under (expected) profit maximization, the researcher assumes that no additional unknown parameters appear in these behavioral equations, e.g., in the case of an interior solution,

$$\frac{\partial}{\partial(\boldsymbol{X}, \widetilde{\boldsymbol{Z}})} \mathrm{E}\big[\boldsymbol{P} f\big(\boldsymbol{X}, \widetilde{\boldsymbol{Z}}, \boldsymbol{K}, \boldsymbol{\varepsilon}\big) - \boldsymbol{R}\boldsymbol{X}\big] = 0, \tag{64}$$

where E is the producer's expectation with respect to $\boldsymbol{P}$, $\boldsymbol{R}$, and $\boldsymbol{\varepsilon}$ and f is defined as in (59). With von Neuman–Morgenstern expected utility maximization, the researcher assumes only one or a few unknown parameters are introduced in a utility function U so that the behavioral relationships in (56) follow

$$\frac{\partial}{\partial(\boldsymbol{X}, \widetilde{\boldsymbol{Z}})} \mathrm{E}\big[U\big(\boldsymbol{P} f(\boldsymbol{X}, \widetilde{\boldsymbol{Z}}, \boldsymbol{K}, \boldsymbol{\varepsilon}) - \boldsymbol{R}\boldsymbol{X}\big)\big] = 0.$$

Strangely, the production literature (as represented by the typical duality approach) has tended over the past few decades toward introducing greater parametric flexibility into (52)–(54), e.g., second-order flexible forms, while imposing total inflexibility in (56). In principle, the behavioral equations can be made a matter of inference by estimating a general and perhaps flexible form for them and then testing for expected profit or expected utility maximization in the context of a broader maintained behavioral hypothesis. For example, suppose U is specified as a second-order flexible form in profit, family labor, creditworthiness, and ending wealth. In this context, wealth differs realistically from initial wealth plus profit by including the productive value of physical capital and soil fertility that have distinctly lower salvage or liquidation values. Then the behavioral relationships in (56) may follow

$$\frac{\partial}{\partial(\boldsymbol{X}, \widetilde{\boldsymbol{Z}})} \mathrm{E}\big[U\big(\boldsymbol{P} f(\boldsymbol{X}, \widetilde{\boldsymbol{Z}}, K, \boldsymbol{\varepsilon}) - \boldsymbol{R}\boldsymbol{X}, z_{i'}, \omega(\boldsymbol{K}), \eta(\boldsymbol{K})\big)\big] = 0, \tag{65}$$

where $z_{i'}$ represents total family labor, $\omega(\boldsymbol{K})$ represents ending wealth as a function of stocks and assets (asset prices are suppressed for convenience), and $\eta(\boldsymbol{K})$ represents creditworthiness as a function of stocks in $\boldsymbol{K}$ (i.e., asset quantities and accumulated debts). In this case, the complete representation of the problem, which closes the system, includes (52)–(54) and (63) in addition to (65). This approach allows inferences about preferences regarding the difficult practical question of how much profit family farms choose to use for consumption versus reinvestment in the operation (as opposed to simply imposing, say, either maximization of the discounted value of profits or maximization of terminal wealth).

Although space in this Handbook is inadequate for presenting a detailed example of this approach, we suggest that balance in the flexibility of technical and behavioral modeling is needed. In the longer-term planning horizons considered by Barry and Robison (2001) for agricultural finance problems, simulation approaches are often found preferable to optimization. One reason is that little has been determined empirically about (i) the importance of current income and consumption versus net worth, (ii) how farmers trade off short-run returns and riskiness with long-run security, and (iii) how asset fixity versus flexibility are used as tools for accomplishing these trade-offs. By estimating the complete production system with flexible behavioral approaches such as in (65), data can begin to sort out empirical applicability of the variety of simulation criteria identified by Barry and Robison. Also, in this context, the need to consider simultaneity in the combined production system as discussed by Mundlak (2001) becomes clear as does the need to use estimation methods that correct for it.

As an additional consideration, dynamic optimization under uncertainty typically assumes additive temporal separability of utility in order to treat dynamic problems of uncertainty [Nerlove and Bessler (2001)]. Better formal modeling depends on understanding the dynamic aspects of risk preferences and how short-term risk trades off with long-term risk given agricultural producers' preferences. Additive temporal separability of utility and risk preferences may not apply. In reality, a farmer may prefer an income stream with low or negative serial correlation rather than high positive serial correlation given the same overall risk because some types of capital investment, debt payment or consumption can be postponed without great difficulty if they can be made up in the near future. On the other hand, postponing such items for many years can cause reduced production, business failure or severe loss in welfare. No satisfactory approach for addressing such problems has yet been proposed.

To date, only the simplest of models have been developed that permit mid-course corrections as specific risks are resolved. As the review by Moschini and Hennessy (2001) shows, even a two-period model that permits one ex post choice has outcomes that depend on third derivatives of production technology. While statistical significance might be obtained in estimating a third derivative of the production technology in a single production study, the variety of results typically obtained by fitting even second-order flexible specifications leaves a great deal open to question. As suggested by Mundlak (2001), the profession has barely, if at all, come to agreement on many elasticities of production, which are determined by first derivatives. Duality has permitted flexibility in estimation of second derivatives but little agreement has been reached on characterizing second derivatives. The profession has hardly crossed the threshold of trying to identify third derivatives. Admitting needed interaction in estimation of technology and preferences and pursuing it with more balance may make clear why models estimated to date do not forecast as well as statistics of fit suggest they should.

5.4. *Technology adoption and technical progress*

In addition to dynamic intertemporal relationships, expectations, and behavioral criteria, modeling technology adoption is also a complicated and complex problem [Sunding and Zilberman (2001); Feder et al. (1985)]. Some technology is embodied in physical capital such as machinery and irrigation so adoption depends on long-term financing opportunities. Some is embodied in variable inputs such as seeds, fertilizer, and pesticides so short-run financing is critical. Depending on how well known and locally applicable is the performance of a technology, adoption can depend heavily on subjective risk, experience, and the extent of rents on technology included in input prices. Some technology is adopted through improved breeding methods and is thus relatively costless but requires years of implementation through succeeding production cycles to realize benefits. Other technology can be implemented only after acquiring costly information or acquiring skills of learning by doing, in which case limited experimentation is a prudent way to proceed [see Foster and Rosenzweig (1995)].

In each of these cases, adoption depends on different factors and constraints that affect an individual farm's production. The role of these factors and the extent to which they apply at the individual level is crucial to understanding the aggregate rate of adoption and agricultural productivity growth. Similarly, each of these cases enters differently through behavioral criteria, production constraints, and modifications of production functions. Again, sorting out technology from behavior from external constraints on the firm is crucial. Because of the complexity of factors potentially affecting technology adoption, space in this overview is not adequate for a critical evaluation of the technology adoption literature beyond the principles already developed throughout this chapter. However, we underscore that technology adoption is a highly heterogeneous problem because of heterogeneous physical capital and differing abilities to take advantage of individual technologies among farms, heterogeneous abilities to learn and thus make new technologies work quickly when information is limited, heterogeneous access to information based on education and other factors, heterogeneous credit constraints that limit financial ability to adopt, etc. The role of experimentation and heterogeneity in technology adoption underscores the importance of considering allocation variables, risk preferences, appropriate long-term as well as short-term preferences, etc. All of these issues fall squarely among the topics addressed in this chapter. For example, because much new technology is embodied in inputs that are subject to financial constraints, the associated principles in Section 4 are relevant. Because much new technology is embodied in capital investment with long-term implications and uncertainties, the principles of Sections 5.1–5.3 are relevant. Accordingly, we suggest many remaining avenues to improving understanding of technology adoption.

6. Heterogeneity and data limitations

As much of this essay has concluded, perhaps the most significant obstacle to further progress in agricultural production analysis is lack of better and more detailed data. Mundlak, Moschini and Hennessy (2001), Nerlove and Bessler (2001), Sunding and Zilberman (2001), and Barry and Robison (2001) (all in this Handbook) each emphasize the problem of trying to learn about micro-level behavior from aggregate data and/or modeling aggregate behavior when individual firms are heterogeneous. As pointed out by Moschini and Hennessy, these problems are difficult under certainty but are more difficult under uncertainty. Considering the other surveys in this part of the Handbook, Deininger and Feder (2001) emphasize heterogeneity of farms associated with soil fertility, soil degradation, liquidity, and transactions costs. Huffman (2001) underscores heterogeneity in human capital and education. Schultz (2001) highlights differences in sex, age, and quality of labor among households and household members, and the associated off-farm labor opportunities. Evenson (2001) also emphasizes soil factors, farmer skills, climatic factors, and infrastructure. Given this heavy recognition of heterogeneity, we finally turn to considerations of heterogeneity and a related call for action.

6.1. Heterogeneity and aggregation across firms

In this section, we examine some remaining issues of heterogeneity and suggest that failure to consider heterogeneity across firms causes errors in aggregation so that estimated forms not only misrepresent technology but fail to support the assumptions used to recover technology from estimated structures.[49] Typically, statistical tests have rejected the standard regularity conditions of homogeneity, monotonicity, symmetry, and convexity of profit functions. Since these regularity conditions are typically used to integrate estimated supplies and demands back to the profit function for purposes of inferring properties of technology, such statistical results call into question the associated inferences regarding technology.[50] In this section, we show that exact aggregation across firms fails when heterogeneity among firms is not represented adequately, which explains one source of failure of the standard regularity conditions. The problem is due to over-summarizing micro-level behavior in publicly reported aggregate data.

Consider the disaggregated static profit maximization problem $\pi = \max_{Y,X}\{\boldsymbol{PY} - \boldsymbol{RX} \mid (\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}) \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})\}$ with resulting vector-valued firm-level supplies $\boldsymbol{y}_i = \boldsymbol{y}(\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{k}_i)$ and demands $\boldsymbol{x}_i = \boldsymbol{x}(\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{k}_i)$ where an i subscript is now added to index firms. For simplicity of notation, let supplies and demands be combined into a netput vector, $\boldsymbol{w}_i = \boldsymbol{w}(\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{k}_i) = (\boldsymbol{y}_i, -\boldsymbol{x}_i)$, let elements of $\boldsymbol{w}_i$ be denoted by w_{ij} where j indexes netputs, and let the netput price vector corresponding to $\boldsymbol{w}_i$ be denoted by $\boldsymbol{q} = (\boldsymbol{p}, \boldsymbol{r})$. Thus, netput functions are denoted compactly by $\boldsymbol{w}_i = \boldsymbol{w}(\boldsymbol{q}, \boldsymbol{k}_i) = \boldsymbol{w}_i(\boldsymbol{q})$. With standard assumptions on technology, profit maximization, and differentiability, individual firm netputs satisfy the four standard regularity conditions of homogeneity, $w_{ij}(\lambda \boldsymbol{q}) = w_{ij}(\boldsymbol{q})$, $\lambda > 0$; monotonicity, $\partial w_{ij}/\partial q_j \geqslant 0$; symmetry, $\partial w_{ij}/\partial q_{j'} =$

[49] This section draws on Just and Pope (1999) where further results and detail are found.

[50] As shown elsewhere in this chapter, standard approaches for aggregation within the firm fail if behavioral preferences follow various alternatives to profit maximization (as discussed in Section 5.3) or firms face various types of constraints such as policy constraints and imperfect capital market constraints (Section 3.3). Just and Pope (1999) show further that the standard regularity conditions generally fail at the firm level when these conditions are present.

$\partial w_{ij'}/\partial q_j$; and convexity, $\{\partial w_{ij}/\partial q_{j'}\} \geqslant 0$, i.e., positive semidefiniteness of the matrix of cross partials.

Defining aggregate netputs across firms as $\overline{\boldsymbol{w}} = \sum_i \boldsymbol{w}_i$, it follows immediately that the four standard regularity conditions must hold at the aggregate level if they hold at the firm level:

$$\overline{w}_j(\lambda \boldsymbol{q}) = \sum_i w_{ij}(\lambda \boldsymbol{q}) = \sum_i w_{ij}(\boldsymbol{q}) = \overline{w}_j(\boldsymbol{q});$$

$$\partial \overline{w}_j/\partial q_j = \sum_i \partial w_{ij}/\partial q_j \geqslant 0;$$

$$\partial \overline{w}_j/\partial q_{j'} = \sum_i \partial w_{ij}/\partial q_{j'} = \sum_i \partial w_{ij'}/\partial q_j = \partial \overline{w}_{j'}/\partial q_j;$$

$$\{\partial \overline{w}_j/\partial q_{j'}\} = \left\{\sum_i \partial w_{ij}/\partial q_{j'}\right\} = \sum_i \{\partial w_{ij}/\partial q_{j'}\} \geqslant 0.$$

Thus, exact aggregation preserves the four standard properties but requires knowledge of all micro variables and functions. The implication is that statistical failure of the regularity conditions must be due to either bias in aggregation of factors and characteristics or failure of the regularity conditions at the firm level. Indeed, the regularity conditions can fail at the firm level because of inapplicability of profit maximization, inappropriate (within-season) temporal aggregation, discrete start-up/shut-down decisions, imperfect capital markets (resource constraints), or errors in measurement [Just and Pope (1999)]. These reasons for failure of standard theory at the firm level have been largely explored in earlier sections. Here we focus on reasons for theoretical failure at the aggregate level assuming regularity conditions hold at the firm level. Results show how aggregation bias and failure of aggregate regularity conditions occur because of the typical approach to representing both price and non-price heterogeneity.

Non-price heterogeneity occurs because of differences among firms in physical capital, technology (including farmer ability and soil productivity), information, and constraints (possibly due to government policy). If such factors are constant across firms, then their effects can be captured in constant parameters. However, investment and technology tend to change over time and differ among firms. Government restrictions change from one policy regime to another and depend on individual farm characteristics such as planting and yield histories or proximity to water resources. These differences cause firms to respond differently to changes in prices.

Suppose $\boldsymbol{k}_i$ represents all short-run fixed factors such as physical capital stock and embodied technologies, family labor constraints, debt constraints, and other attributes of the farm and farmer that explain differences in productivity and profits among individual producers after accounting for variable input choices and allocations of fixed factors. If each firm faces the same price vector, an accurate aggregate netput specification is $\overline{w}_j(\boldsymbol{q}, \boldsymbol{k}_1, \ldots, \boldsymbol{k}_\eta) = \sum_i w_j(\boldsymbol{q}, \boldsymbol{k}_i)$ where η is the total number of firms. However, estimation of an aggregate equation of the form $\overline{w}_j(\boldsymbol{q}, \boldsymbol{k}_1, \ldots, \boldsymbol{k}_\eta)$ is likely impractical both because complete firm-specific data is typically not available and because too many parameters require estimation (without considerable simplification).

A feasible approach is to model the distribution of non-price factors. Where $G(\boldsymbol{k})$ represents the joint distribution of such factors among firms, an accurate specification of aggregate netputs is $\overline{w}_j(\boldsymbol{q}, G) = \int \eta w_j(\boldsymbol{q}, \boldsymbol{k})\, \mathrm{d}G(\boldsymbol{k})$. If this distribution has a parameter vector, say $\boldsymbol{\theta}$, then aggregate netputs follow

$$\overline{w}_j(\boldsymbol{q}, \theta) = \int \eta w_j(\boldsymbol{q}, \boldsymbol{k})\, \mathrm{d}G(\boldsymbol{k} \mid \boldsymbol{\theta}). \tag{66}$$

From this result, exact aggregation and the standard regularity conditions are preserved if aggregation considers the full distribution of characteristics among firms. While a full distribution would require complete sampling of all firms, if $\boldsymbol{\theta}$ is a sufficiently short parameter vector it can be estimated from a random sample of $\boldsymbol{k}$. Thus, (66) facilitates tractable empirical representation under heterogeneity. Aggregation is then exact aside from errors in estimating $\boldsymbol{\theta}$ so that regularity conditions are preserved. For example, if G can be represented by, say, a two-parameter distribution such as a log-normal, then the two parameters can be usefully estimated from survey data over a limited random sample of firms.

Alternatively, aggregate demand is typically estimated in the form $\overline{w}_j(\boldsymbol{q}, \overline{\boldsymbol{k}})$ where $\overline{\boldsymbol{k}}$ is a vector of non-price indexes. A relevant question is whether some choice of $\overline{\boldsymbol{k}}$ can achieve exact aggregation, $\overline{\boldsymbol{w}}_j(\boldsymbol{q}, \overline{\boldsymbol{k}}) = \sum_i \boldsymbol{w}(\boldsymbol{q}, \boldsymbol{k}_i)$, where $\overline{\boldsymbol{k}}(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_\eta)$ is an aggregate index vector of firm characteristics. Such macro indexes typically consist only of sums or means (e.g., total or per capita physical capital). Unfortunately, neither exact aggregation nor the standard regularity conditions are preserved when all moments in $\boldsymbol{\theta}$ other than the first are ignored (assuming $\boldsymbol{\theta}$ contains two or more parameters). Following (66), other moments corresponding to each of the moments in $\boldsymbol{\theta}$ are generally needed for exact aggregation.

This result implies that aggregate netput specifications based on distribution-insensitive indexes cannot, in general, represent the aggregate marginal effects of either price or non-price factors. Aggregate netput specifications based only on total, per capita, or average characteristics cannot represent aggregate marginal effects because aggregate marginal effects depend on how increments in aggregate characteristics are allocated among firms. Similarly, incomplete models depending only on single-moment indexes cannot represent the aggregate marginal effects of prices because marginal price effects depend on the distribution of non-price factors among firms. For example, consider the case where shut-down conditions vary among firms because of differences in characteristics. In such a case, both aggregation and standard regularity conditions fail [see Just and Pope (1999)].

In reality, some of the factors that differentiate farms and farmers such as management ability or soil fertility may be hard to observe. However, other public data on farm characteristics is routinely collected. For example, data on physical capital are compiled by sampling individual farms. Typically, public data report only means or totals for such data collection efforts. Additionally reporting, say, the standard deviation and skewness would be relatively costless. The full data set would be useful but is usually not made available because of right-to-privacy restrictions. However, the major cost is in conducting the survey – a cost that must be incurred whether one or many moments of the distribution are reported – so a more complete reporting of the distribution appears feasible with minor costs of reporting. The results here suggest that models of production and estimates of supplies and demands could possibly be improved substantially as a result.

While the above discussion considers one-dimensional differences among firms, in reality firms differ in multiple ways. Note, however, that $G(\boldsymbol{k})$ represents the joint distribution of all characteristics among farms including capital structure and technology, information, constraints, farmer abilities, and farm fertility. Thus, the right-hand side of (66) considers cross-characteristic relationships among firms, e.g., between factors such as capital and family labor availability. Therefore, the result in (66) further implies that aggregate netput specifications may depend on correlations among characteristics. By implication, correlation-insensitive indexes of non-price factors cannot, in general, represent the aggregate marginal effects of either price or non-price factors [see Just and Pope (1999) for details].

These results imply that expanded data reporting efforts should focus not only on own-moments of characteristic distributions among firms, but also on cross-moments. For example, if $G(\boldsymbol{k})$ follows a multivariate log normal distribution, then the mean and covariance matrix of characteristics across firms would be sufficient to facilitate exact aggregation following (66). Unfortunately, much agricultural data is reported in a way that does not reflect correlations of characteristics. This is particularly true of the relationship of productivity characteristics to environmental characteristics because these two sets of characteristics tend to be collected by independent surveys and even by independent government agencies [Just and Antle (1990); Antle and Just (1992)]. For roughly the same data collection costs, correlations could be estimated if data were indexed by farms, and efforts were made to include the same farms in samples. Apparently, more exact aggregation is possible with little additional data collection cost if data reporting efforts are sensitive to these possibilities. If so, more congruence of theory and empirical results seems likely.

A similar additional generalization permits consideration of price heterogeneity. Regardless of competition, firms may face different prices because of transportation costs, volume discounts, and seasonality.[51] Where individual netputs follow $w_j(\boldsymbol{q}_i, \boldsymbol{k}_i)$, an

[51] The potential magnitude of this problem is illustrated by spatial variations of output prices due to geographic variation in seasonality of crop production. For example, because of typical weather patterns, the wheat harvest in the U.S. typically starts in Texas in May and continues gradually northward to North Dakota in September. If wheat prices vary throughout the year, then southern farmers are not responding to the same price signals as northern farmers. A dramatic example of wide price variation in a single crop season was caused by the Soviet grain deals in the 1970s. As the Soviet Union bought more and more grain in 1972, wheat prices increased from \$1.56 per bushel in Texas to \$1.70, \$1.68, \$1.74, \$1.81, and \$1.90 in Oklahoma,

accurate aggregate netput specification is $\overline{w}_j(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_\eta, \boldsymbol{k}_1, \ldots, \boldsymbol{k}_\eta) = \sum_i w_j(\boldsymbol{q}_i, \boldsymbol{k}_i)$. While complete data on heterogeneity of both prices and characteristics among farms is typically not available, a tractable approach is again available if a joint distribution of prices and characteristics among firms can be estimated. Where $G(\boldsymbol{q}, \boldsymbol{k})$ represents this joint distribution, an accurate specification of aggregate netputs is $\overline{w}_j(G) = \int \eta w_j(\boldsymbol{q}, \boldsymbol{k})\, \mathrm{d}G(\boldsymbol{q}, \boldsymbol{k})$. If this distribution is parameterized by a vector $\boldsymbol{\theta}$ that can be estimated for each aggregate observation, then aggregate netputs can be represented as $\overline{w}_j(\boldsymbol{\theta}) = \int \eta w_j(\boldsymbol{q}, \boldsymbol{k})\, \mathrm{d}G(\boldsymbol{q}, \boldsymbol{k} \mid \boldsymbol{\theta})$, which facilitates accurate aggregation to the extent that $\boldsymbol{\theta}$ is accurately estimated. With this approach, aggregate netputs preserve homogeneity in mean and spread parameters of the price distribution; and monotonicity, symmetry and convexity are preserved in mean prices [see Just and Pope (1999)]. For other results on aggregation with price heterogeneity, see Pope and Chambers (1989).

In lieu of this approach, most aggregate specifications attempt to represent netputs as functions of aggregate price indexes, $\overline{\boldsymbol{q}}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_\eta)$, as well as indexes of characteristics, $\overline{\boldsymbol{k}}(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_\eta)$. The related problem is whether the standard linear aggregation condition, $\boldsymbol{w}(\overline{\boldsymbol{q}}, \overline{\boldsymbol{k}}) = \sum_i \boldsymbol{w}_i(\boldsymbol{q}_i, \boldsymbol{k}_i)$, holds. Such aggregate indexes typically include only average prices or characteristics and include only one index for each price and each characteristic that differentiates individual firms. Again, more accurate aggregation is possible and standard properties are more likely to hold if the indexes used to represent prices as well as characteristics reflect all of the moments in $\boldsymbol{\theta}$ needed to differentiate the distribution of prices and characteristics among aggregate observations used for estimation. Again, because price data are collected at a disaggregated level, at least some measures of dispersion could easily be reported in addition to the simple or weighted averages now reported with no additional data collection costs and small additional reporting costs.

Finally, we suggest the potential for heterogeneity of information. While a non-trivial role of information can be posed under certainty, many interesting information problems in agriculture arise under uncertainty. Agricultural producers must make decisions affecting output before uncertain output prices are known. Producers likely have different expectations for both prices and technology performance. Such heterogeneity can have important implications even under risk neutrality as demonstrated by Pope and Just (1996, 1998).

Suppose the firm maximizes expected profit as in (64). Then the resulting expected netput vector of the firm can be represented by $\boldsymbol{w}(\boldsymbol{q}, \boldsymbol{k}_i, \boldsymbol{I}_i)$ where $\boldsymbol{I}_i$ denotes the information by which farmer i formulates expectations regarding production responses and uncontrolled production effects (disturbances). Assuming farmers' expectations are unbiased, an accurate specification for expected aggregate netput j is

Kansas, Nebraska, South Dakota, and North Dakota, respectively, as the harvest moved north. In 1973, prices increased from \$3.04 in Texas to \$3.56, \$3.75, \$3.80, \$4.24, and \$4.82 in Oklahoma, Kansas, Nebraska, South Dakota, and North Dakota, respectively [Economic Research Service (various years)]. Aggregating inputs and outputs across these farmers based only on the national average price, one would thus expect such volatile price years to appear technically inefficient falsely even if all individual farmers are fully efficient [Chambers and Pope (1991)].

$\overline{w}_j(\boldsymbol{\theta}) = \int \eta \mathrm{E}_I[w_j(\boldsymbol{q}, \boldsymbol{k}, \boldsymbol{I})] \, \mathrm{d}G(\boldsymbol{q}, \boldsymbol{k}, \boldsymbol{I} \mid \boldsymbol{\theta})$ where $\overline{w}_j$ now represents an expected aggregate netput and G represents a joint distribution of prices $\boldsymbol{q}$, characteristics $\boldsymbol{k}$, and information $\boldsymbol{I}$ over all farmers. Thus, similar conclusions follow as for other cases of heterogeneity.

Characterizing the distribution of information among producers, however, is a daunting task. Only recently has work such as Wolf, Just, and Zilberman (forthcoming) attempted to characterize sources and choices of information by individual firms. However, no systematic and recurring efforts have been developed to compile such data for use in comprehensive production studies. Other studies [e.g., Just (1974)] have attempted to describe producer information by including regression functions explaining moments of subjective price or yield distributions. To date, however, these approaches have been implemented only at the aggregate level and thus introduce potential aggregation problems in information. Perhaps if other firm-level information were sufficiently complete, differences in information among firms could be inferred with these approaches. In either case, it seems that information heterogeneity is a source of aggregation bias that will be difficult to overcome empirically without more complete firm-level data.

This section demonstrates several generalizations whereby congruence of theory and empirical work can be (better) achieved by better data and aggregation. In each case, empirical implementation is constrained by current data availability. The most promising step to improving aggregation appears to be generalizing data reporting to include at least second own- and cross-moments of producer characteristics. Then aggregate supply/demand specifications can be based on at least two-parameter distributions of characteristics among firms. Seemingly, reporting independent distributional data for capital, prices, government controls, and many determinants of technology (e.g., land quality) is possible with little additional public expense. On the other hand, characterization of some factors such as farmer ability and information at the firm level will likely be more difficult.

6.2. *Data limitations: a call for action*

That existing data seriously limits agricultural production research may be surprising given that Leontief (1971, p. 5), while president of the American Economics Association, pronounced agricultural economic data to be a model which other economic subdisciplines could/should emulate: "Official agricultural statistics are more complete, reliable, and systematic than those pertaining to any other major sector of our economy". The part of this statement that now seems implausible is related to the word "complete". Though agricultural economists' appetite for data is probably insatiable, a brief evaluation of the sources of agricultural production data is worthwhile in assessing whether Leontief's 1971 evaluation is accurate today.

Secondary aggregate data for both crops and livestock are abundant. For example, data on crops include acres planted and harvested, inventories, trade, storage, disappearance, and price. Though there are differences in quality and availability, such data are generally available throughout the world. They are summarized annually in the U.S. Department of Agriculture's publication, Agricultural Statistics, and are available for most countries from the FAO. A second source of aggregate U.S. data is the *Census of Agriculture*, which is published at roughly five-year intervals. Additionally, county and state data on individual commodities and crop and livestock aggregates are widely available therein. However, individual farm data are not released by public sources because of right-to-privacy concerns.

As previous sections have shown repeatedly, aggregate data is a poor substitute for disaggregated data for understanding agricultural production. This is particularly true for problems where allocations within firms over time (production stages) and space (plots) are important but unrecorded, and for problems where variation among firms is crucial (e.g., where risk and heterogeneity of characteristics are important). For example, Just and Weninger (1999) show that farm-level yield variances are from two to ten times greater than reflected by aggregate data so that most of the risk faced by individual farmers is averaged out of aggregate data, and the structure of risk facing farmers is often significantly mischaracterized. Theoretical models suggest that response to risk is unlikely to be measured effectively with secondary aggregate data because it (i) tends to obfuscate individual responses and risk and (ii) offers very poor measurement of wealth on which risk aversion likely depends. In addition, conceptual studies are finding that representation of heterogeneity is of crucial structural importance for policy analysis, particularly when environmental concerns are important, because both actual and contemplated controls depend on localized land characteristics [e.g., Hochman and Zilberman (1978); Just and Antle (1990); Antle and Just (1992)]. Yet the vast majority of agricultural production studies are done using aggregate data without apology. The primary reason is lack of adequate firm-level data.

At the firm level, the "Agricultural Resource Management Study" (formerly "The Farm Cost and Returns Survey") conducted by the U.S. Department of Agriculture's Economic Research Service contains extensive data on individual farms, but these are not available for use outside of the agency as a public use sample. Furthermore, these surveys are limited in scope because of governmental sampling exposure concerns. Lacking are microeconomic data that will allow a more thorough understanding of farm behavior. As discussed throughout this chapter, needed data must be capable of representing considerable heterogeneity. Yet the very identifying data that could permit merging of these observations with the extensive data base on land quality compiled by the U.S. Natural Resources Conservation Service (formerly the U.S. Soil Conservation Service) is typically restricted. As well, for many issues, the data needs to include intertemporal continuity. To avoid excessive survey exposure, observations are typically drawn on different farms from year to year so no information is available to track investment and productive asset replacement over time. In absence of obtaining such data, a reliable analysis of productive asset acquisition and replacement is difficult and doubtful. Panel data is necessary to do a careful and comprehensive analysis of agricultural investment behavior.

In the U.S., some state land grant institutions have developed farm-level data sets across both time and farms by offering farm-level accounting and management assistance (e.g., Kansas State University). However, since participation is farmer-selected these samples are not random. Furthermore, these data are typically not publicly available and the data are not organized around a broad set of recurring economic issues. For example, such data typically record only external transactions of the farms whereas some additional recording of internal decisions (e.g., allocations of variable inputs) and characteristics (e.g., soil quality) could greatly enhance the value of the data. Yet in spite of these limitations, judging by publications in the leading agricultural economic journals, these samples are heavily used by those who have access to them. Such studies try to understand a variety of behaviors ranging from consumption and wealth accumulation to risk response [e.g., Jensen et al. (1993); Saha et al. (1994)].

Perhaps the best approximation of a comprehensive panel data base for agricultural production is the ICRISAT household data base, which represents primitive developing agriculture. With these data, many aspects of developing agriculture have been investigated and considerable additive debate has emerged accordingly. Developed agriculture, however, is considerably more complex because of scale heterogeneity, policy variability, complex finance and investment, greater scope of inputs and outputs, etc. Furthermore, understanding policy, markets and prices in all countries depends heavily on understanding agricultural production in the major developed countries because of their domination of world trade.

We propose that a significant and complete data base for developed agricultural production needs to be developed as an investment by/for the agricultural economics profession, and that access to such data should be made freely available to all in order to facilitate debate. Debate could be additive because researchers would be forced to compare their maintained hypotheses when working with the same data. Such a data set would allow students to hone their research skills more comprehensively and allow the leading contributions of the profession to add cumulatively to a set of commonly held stylized facts. From these, additional knowledge would spring. Such a data base could facilitate investigation of many issues identified by this study as blocked by data unavailability. By comparison, the current proliferation of studies with uncommon data bases and incongruent maintained hypotheses has led to endless speculative explanations of differences in results with little comprehensive comparison [Alston and Chalfant (1991); Smale et al. (1994).

Such a data base could serve much like public labor data have served the labor economics discipline to facilitate debate and development of a set of stylized facts for the discipline and its policy analysis efforts.[52] Labor economics is a field of economics that

[52] An example of the usefulness of stylized facts is given for the marketing arm of the agricultural economics discipline by the focus and debate about elasticities of supply and demand during the 1950s and 1960s. Prior to the flexibility fad in supply and demand estimation, empirical production and marketing studies were heavily judged and criticized on the basis of accepted wisdom regarding supply and demand elasticities and whether they added to the profession's knowledge of them.

has aggressively developed a useful set of microeconomic data. These data include the public use samples of the Census and Current Population Survey (CPS). The CPS is a monthly survey of approximately 60,000 households in all 50 states. Approximately one in 1600 households are surveyed. These data are extensive regarding wages, labor force participation, and socio-demographic data. Other panel-type data are found in the "Panel Study of Income Dynamics" (PSID), the National Longitudinal Survey (NLS and NLS Y2), the High School and Beyond Survey, and many special purpose instruments as well. In comparison, the dearth of microeconomic agricultural data makes understanding agricultural production, a seemingly more complex problem, very difficult.

Any effort to create a broad and complete public panel of agricultural production data will likely require more resources than state land-grant efforts could/should devote. Furthermore, state-level development is likely not to lead to the public access that is needed to facilitate a broad professional and cumulative debate. Because the benefits of such data would be broadly applicable, such an effort seems to be merited at the national or even international level. However, because of excessive survey exposure and right-to-privacy restrictions applied to government surveys, a non-governmental organization may be a more effective means of developing such a data set.

If these possibilities are pursued, the agricultural economics profession can once again lead the general economics discipline as an example of empirical excellence. Many of the issues raised throughout this chapter regarding the structure of technology and preferences can be addressed under assumptions much more consistent with practical agricultural knowledge. And many of the thorny generalizations (representation of investment, information acquisition, and the role of disturbances) yet needed to represent the agricultural production problem meaningfully and comprehensively can then be addressed sensibly.

7. Conclusions

Economists have a primary responsibility to discover behavioral relationships. In practice, this has led to use of methodologies that require minimal or no resources for understanding the underlying structure of technology. Ironically, the effort to represent technologies with maximal flexibility has resulted in empirical approaches that exhaust the identifying potential of available data in capturing that flexibility. Little or no identifying potential remains for discovering behavior.

Presumably, all production economists agree that understanding the essential elements of technology is important to economic thought and measurement. Indeed, the concepts and measurement of productive and technical efficiency and the creation and adoption of technology all seem to be undergoing a considerable rebirth of interest in recent years. A fundamental question in these pursuits is, "What elements of technology should economists consider essential?" That no consensus exists is evident by perusing the *American Journal of Agricultural Economics*, the *Journal of Productivity Analysis*, and the *International Journal of Production Economics*. We have argued that agricultural technology is fundamentally different than for most industrial production and that potentially large gains may come from understanding more of the structure that underlies aggregate reduced-form concepts of production technology. Questions regarding economies of scale and scope, prescriptions for farm management, adoption of technology, productivity and technical change, input demand, output supply, outsourcing [Coase (1937)] and the structure of the firm are only properly understood in the context of technology descriptions that include dynamics, risk, technical structure, input allocation, and constraints associated with policy controls and firm-owned resources. If technology, behavior, and policy instruments are confounded in specification and estimation, then models are not useful for investigating the effects of changes in policy, technologies or industry structure.

As an example, one of the most important issues for future policy is the rapid evolution in the nature of the farm firm. Many farms, particularly those in the livestock sector, increasingly resemble the large-scale specialized manufacturing model. Many farms (e.g., those involved in contract farming) resemble component suppliers to manufacturers. Some (e.g., in the poultry industry) specialize as proprietors of technology. These developments likely have explanation in the framework proposed by Coase (1937). Careful representation and analysis of structured technology in the presence of information asymmetries appear to be crucial to understanding why some services are purchased, why others are produced within the farm, and yet others are produced by the operator or owner of the farm [Allen and Lueck (1998)].

If the agricultural economics profession lacks either relevant theory or evidence, it is a profession without science. Improved congruence of theory and evidence is needed to (i) enable researchers to better understand behavior, (ii) provide better support for policymakers, and (iii) facilitate greater appreciation of classroom theory by students. Some of the most basic theoretical properties of production theory – for example, monotonicity, homogeneity, convexity, and symmetry – are rejected by a predominance of empirical work [Shumway (1995)]. Rejection could be due to flawed theory, flawed empirical analysis, or flawed data. We have suggested several possibilities of theoretical failure beginning in Section 3, several possible failures of empirical practices beginning in Section 4.4, and some major shortcomings of available data in Section 6. Likely some combination of these explanations accounts for the poor performance of agricultural production models noted by Mundlak (2001). Without further research – some of which may not be possible with present data – the extent of failure caused by each is almost impossible to determine. Thus, enhancement of data seems to be a first priority.

We noted in our introduction that there is an increasing gulf between farm management economists on one hand and (agricultural) production economists on the other. Economists are accustomed to arguing for the benefits of division of labor. However, we have argued that much of this gulf is due to cavalier empirical treatment by agricultural production economists of the structure of technology, behavioral preferences of producers, and the constraints and policies they face. Our point of departure is the Fundamental Axiom of Multi-output Production. If this axiom is taken seriously, then methods used by production economists and the data required for analysis are fundamentally different.

The results of this paper underscore the need to develop farm-level data and data on input allocations. One of the greatest problems is inappropriate aggregation and inappropriate representation of heterogeneity imposed by present data availability. Public data are mostly aggregate data describing only the first moment of the underlying distribution among farms. Furthermore, data rarely record allocations of inputs except for land. Even most farm-level survey data such as the Agricultural Resource Management Study (formerly the Farm Costs and Returns Survey) carried out by the Economic Research Service do not record input use by crop or application rates. A few private services (e.g., Doane Marketing Research, Inc.) provide data on pesticide use or application rates by crop but this information is rarely if ever used in the journals of the agricultural economics profession in part because of the expense and in part because the data cannot be provided to others as required by some journal policies.

Lack of data on allocations has tended to cause agricultural production analysis to use aggregate implicit representations of technology. Conceptually, we have demonstrated that implicit representation of technology can lead to deceiving conclusions when some producer decisions are unobserved (most particularly, allocations). Hypothesis tests of technology structure using standard dual and implicit representations of technology are shown to be invalid for typical cases. Under-representing the dimensions of the producer's decision problem can cause inappropriate conclusions. If a producer does not simply decide how much fertilizer to use, but must decide how much fertilizer to use on corn and how much to use on wheat, or how much to use at planting and how much to use during the growing period, then these considerations must be taken into account in specifying the technology before solving out the unobserved variables to reach estimable forms. Allocations as well as aggregate use must be considered in testing for technology structure.

While policy- and behavior-relevant aggregations are appropriate in representing technology, the typical practice has been to ignore allocations and characterize technology with purely aggregate variables. While the set notation of duality lends itself to a high level of generality in theory, the typical step to empirical representation has ignored that potential by assuming technology is neatly described by a single equation devoid of allocations. Standard implicit or explicit specifications of scalar product transformation functions of the form $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$ do not permit generality with respect to the rank of the relationship between $\boldsymbol{X}$ and $\boldsymbol{Y}$.[53] Implicit representation is particularly distorting if some producer decisions are unobserved. That is, when some unobserved variables are solved out of the structural representation before computing the reduced form, the apparent structure of the observable production possibilities frontier may not reflect characteristics of the underlying technology. However, implicit representation is

[53] That is, all common scalar specifications of $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$ imply a Jacobian for the transformation from $\boldsymbol{X}$ to $\boldsymbol{Y}$ of rank 1.

an important problem even if all producer decisions are observed and included in the scalar implicit representation of technology. If such generality is not admitted, then implicit forms arbitrarily exclude the potential nonjointness of, say, (45) for which they are used to test. Alternatively, explicit representations such as (11) can be estimated and used to determine the rank of the relationship between X and Y.

More importantly, single-equation and indirect representations of multi-output production can under-represent the dimensionality of the decision problem. Most often, inputs are represented only by aggregate variables that under-represent the dimensionality of the production technology (and the associated decision problem) when inputs must be allocated in some way over space, time, or production activities. As a result, estimates are policy- or behavior-dependent implying that "technology" models are unstable across observations where policy differs (as is typical in time series data) or behavior differs (as is likely in cross section data). In fact, if there are two or more unobserved allocated inputs, then no purely technological relationship is likely observable. With the present state of data, this may be a major constraint to any meaningful analysis of technical efficiency. Also, if decisions are changing frequently because of changes in policy instruments, then time series data and typical dual (PPF) approaches may offer little hope for estimating a stable "technology."

Dual methods, while not inherently tied to this problem, have led to flexible but indirect representations of technology in practice because flexible forms are not self dual [McFadden (1978); Blackorby et al. (1978)]. Because these approaches start from a PPF representation of technology, most estimates of production technologies in the literature likely include behavioral criteria, are contaminated by policy heterogeneity either across firms or time, and are not pure estimates of technology. Associated hypothesis tests about technology are therefore invalid and actually represent joint tests about technology, policy, and behavioral criteria. For example, rejection of a hypothesis of, say, technical change could, in fact, imply rejection of the profit maximization assumption on which standard duality is based.

Because a large part of the empirical agricultural production literature is based on a PPF approach (e.g., the typical PPF dual approach), the limits of usefulness of PPFs need to be recognized. A PPF permits (i) estimation of total factor demands and supplies and (ii) measurement of industry rents, but even these are valid only if the Aggregation Qualification Condition is met. By comparison, estimates of the PPF alone do not permit (i) examination of nonjointness, homotheticity, or separability of the technology, (ii) prescription of decisions, (iii) analysis of effects of changes in policy instruments, or (iv) explanation of how technical change affects decisions. The reason is that PPFs, because they do not represent allocations, may be policy- or behavior-dependent. In any case, tests of nonjointness, homotheticity, and separability on the frontier do not determine similar properties of the underlying technology.

More seriously, under-representing technological dimensionality may induce structural characteristics such as jointness and non-separability on the aggregate variables when similar characteristics do not apply to underlying technology. These possibilities invalidate some tests and limit the usefulness of almost all tests of technology structure to date for multi-output production problems. These results also offer a likely explanation for why empirical methodologies have not delivered according to their conceptual promises [see Mundlak's (2001) criticism]. That is, if the technology description implicitly includes policy and behavioral criteria, then it is not surprising that empirical estimates are not stable and are inappropriately interpreting observed empirical relationships as implausible relationships in the data. Seemingly the practice noted by Moschini and Hennessy (2001) of sophisticated theoretical modeling with simplistic empirical modeling has led to few recognized empirical regularities. Given the potential invalidating implications of ignored realities, we fear that the current state of empirical knowledge of agricultural production sums up to little more than an empty box.

Appendix. Describing technology independent of policy and behavior

This appendix gives a brief formal treatment of some of the points in Sections 4.1–4.4 using the notation introduced in Sections 3.1 and 4.1. The overall technology is assumed to have a structure composed of sub-technologies $(\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(\boldsymbol{z}^i, \boldsymbol{\varepsilon})$ that yield aggregate output $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{Y}$ using aggregate purchased inputs $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{X}$ given fixed allocated resource constraints $\boldsymbol{C}\boldsymbol{Z} \leqslant \boldsymbol{K}$, i.e.,

$$\big\{(\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\big\} \equiv \big\{(\boldsymbol{y}, \boldsymbol{x}) \mid (\boldsymbol{Y}, \boldsymbol{X}) \in \cup_i \Im_i\big(\boldsymbol{z}^i, \boldsymbol{\varepsilon}\big), \boldsymbol{y} = \boldsymbol{A}\boldsymbol{Y}, \boldsymbol{x} = \boldsymbol{B}\boldsymbol{X}, \boldsymbol{C}\boldsymbol{Z} \leqslant \boldsymbol{K}\big\}, \tag{A.1}$$

which is equivalent to (43). Under continuity and monotonicity, the upper right-hand (efficient) boundary of feasible $(\boldsymbol{y}, -\boldsymbol{x})$ associated with $\Im_{-i}$ is described by

$$\begin{aligned} &F(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) \equiv y_1 - f(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) = 0, \\ &y_1 = f(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) \equiv \max\big\{y_1 \mid (\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\big\}, \end{aligned} \tag{A.2}$$

where $\boldsymbol{y} = (y_1, \boldsymbol{y}_{-1})$. To identify the specific production plan necessary to attain any distinct $(\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}$, the spatially and temporally detailed vectors $\boldsymbol{X} = (\boldsymbol{x}^1, \dots, \boldsymbol{x}^m)$, $\boldsymbol{Z} = (\boldsymbol{z}^1, \dots, \boldsymbol{z}^m)$, and $\boldsymbol{Y} = (\boldsymbol{y}^1, \dots, \boldsymbol{y}^m)$ not included in (A.2) must be determined. Also, to facilitate determination of the implications of policy instruments that impose limitations on specific $\boldsymbol{x}^i$'s or $\boldsymbol{y}^i$'s, such as $(\boldsymbol{Y}, \boldsymbol{X}) \in G$, the representation in (A.2) does not suffice. Alternatively, this technology can be represented with spatial and temporal detail,

$$\{(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}) \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})\} \equiv \big\{(\boldsymbol{Y}, \boldsymbol{X}) \in \cup_i \Im_i\big(\boldsymbol{z}^i, \boldsymbol{\varepsilon}\big)\big\} \cap \{\boldsymbol{Z} \mid \boldsymbol{C}\boldsymbol{Z} \leqslant \boldsymbol{K}\}, \tag{A.3}$$

which is equivalent to (44). To represent spatially and temporally detailed technology in functional form under continuity and monotonicity, the upper right-hand (efficient) boundaries of feasible $(\boldsymbol{y}^i, -\boldsymbol{x}^i)$ associated with $\Im_i$ are described by

$$\begin{aligned} F_i(\boldsymbol{y}^i, \boldsymbol{x}^i, z^i, \boldsymbol{\varepsilon}) &\equiv y_1^i - f_i(\boldsymbol{y}_{-1}^i, \boldsymbol{x}^i, z^i, \boldsymbol{\varepsilon}) = 0, \\ y_1^i = f_i(\boldsymbol{y}_{-1}^i, \boldsymbol{x}^i, z^i, \boldsymbol{\varepsilon}) &\equiv \max\{y_1^i \mid (\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon})\}, \end{aligned} \tag{A.4}$$

where $\boldsymbol{y}^i = (y_1^i, \boldsymbol{y}_{-1}^i)$. The combination of these conditions across all sub-technologies is represented by (45).

A.1. Alternative concepts of efficiency

The alternative concepts of technical efficiency in Section 4.1 are defined formally by

DEFINITION A.1.
- (E1) $(\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon})$ is sub-technology efficient if there does not exist $(\boldsymbol{y}^{i\prime}, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon})$ such that $\boldsymbol{y}^{i\prime} \nleq \boldsymbol{y}^i$. With continuity and monotonicity, this condition reduces equivalently to (A.4).
- (E2) $(\boldsymbol{Y}, \boldsymbol{X})$ satisfies structural technical efficiency for a given $(\boldsymbol{Z}, \boldsymbol{\varepsilon})$ if there does not exist $(\boldsymbol{y}^{i\prime}, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon})$ such that $\boldsymbol{y}^{i\prime} \nleq \boldsymbol{y}^i$ for any sub-technology. With continuity and monotonicity, this condition reduces to (45).
- (E3) $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ satisfies fixed factor technical allocative efficiency if there does not exist an alternative allocation $\boldsymbol{Z}'$ such that $(\boldsymbol{Y}', \boldsymbol{X}, \boldsymbol{Z}') \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})$, $\boldsymbol{CZ}' \leqslant \boldsymbol{CZ}$, and $\boldsymbol{Y}' \nleq \boldsymbol{Y}$.[54]
- (E4) $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ satisfies variable input technical allocative efficiency if there does not exist an alternative allocation $\boldsymbol{X}'$ such that $(\boldsymbol{Y}', \boldsymbol{X}', \boldsymbol{Z}) \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})$, $\boldsymbol{BX}' \leqslant \boldsymbol{BX}$, and $\boldsymbol{Y}' \nleq \boldsymbol{Y}$.
- (E5) $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ satisfies output technical allocative efficiency if there does not exist an alternative allocation among sub-technologies such that $(\boldsymbol{Y}', \boldsymbol{X}', \boldsymbol{Z}') \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})$, $\boldsymbol{BX}' \leqslant \boldsymbol{BX}$, $\boldsymbol{CZ}' \leqslant \boldsymbol{CZ}$, and $\boldsymbol{AY}' \nleq \boldsymbol{AY}$.
- (E6) $(\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})$ is technically efficient in a reduced-form sense if there does not exist an alternative allocation of the aggregate input vector $\boldsymbol{x}$ that will produce an aggregate output vector $\boldsymbol{y}' \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})$ such that $\boldsymbol{y}' \nleq \boldsymbol{y}$. This condition corresponds to the efficient boundary of (A.1) which reduces equivalently to (A.2).
- (E7) $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z})$ satisfies feasible disaggregated input-output efficiency if in addition to (E2) there does not exist an alternative allocation $\boldsymbol{Z}'$ such that $(\boldsymbol{Y}', \boldsymbol{X}, \boldsymbol{Z}') \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})$, $\boldsymbol{CZ}' \leqslant \boldsymbol{K}$, and $\boldsymbol{Y}' \nleq \boldsymbol{Y}$. This condition corresponds to the efficient boundary of (A.3) and includes (E3).

Because Propositions 1–3 state negative results, they can be proved by examples. For purposes of brevity, proofs of propositions are only outlined.

[54] The relationship $\boldsymbol{Y}' \nleq \boldsymbol{Y}$ means $\boldsymbol{Y}' \geqslant \boldsymbol{Y}$ with strict inequality in at least one element.

PROOF OF PROPOSITION 1. Consider the full profit-maximization problem with temporal and spatial price detail in $\boldsymbol{P} = (\boldsymbol{p}^1, \ldots, \boldsymbol{p}^m)$ and $\boldsymbol{R} = (\boldsymbol{r}^1, \ldots, \boldsymbol{r}^m)$ ignoring for the moment inability to forecast $\boldsymbol{\varepsilon}$,

$$\pi = \max_{\boldsymbol{Y},\boldsymbol{X}} \{\boldsymbol{PY} - \boldsymbol{RX} \mid (\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}) \in \Im(\boldsymbol{k}, \boldsymbol{\varepsilon})\}. \tag{A.5}$$

Considering aggregation over sub-technologies, the corresponding profit maximization problem,

$$\pi = \max_{\boldsymbol{y},\boldsymbol{x}} \{\boldsymbol{py} - \boldsymbol{rx} \mid (\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\},$$

is clearly not an aggregate of the solution to problem (A.5) when $\boldsymbol{p} \neq \boldsymbol{p}_i$ and $\boldsymbol{r} \neq \boldsymbol{r}_i$ for $i = 1, \ldots, m$. Accordingly, (E4) and (E5) may be inconsistent with standard profit maximization behavior. Similarly, imposing a policy or behavioral constraint on a specific element of the $\boldsymbol{X}$ or $\boldsymbol{Z}$ vector as in $(\boldsymbol{Y}, \boldsymbol{X}) \in G$ renders (E4) or (E3) inapplicable, respectively. □

PROOF OF PROPOSITION 2. See the proof of Proposition 1 and note that (E4) and (E5) are required by profit maximization of (A.1). □

PROOF OF PROPOSITION 3. This proof follows by noting that (45) has no condition equating marginal productivities of allocated fixed factors and, conversely, equating marginal productivities of allocated fixed factors does not necessarily satisfy (45). □

As noted in Section 3.4, the problems in Propositions 1–3 may be encountered in aggregating inputs or outputs spatially and/or temporally over allocations that are not subject to the same prices, shadow prices, policy constraints, or preference relationships (see the Aggregation Qualification Condition).

A.2. *Two-stage representation of the producer's problem*

This section illustrates the issue concerning policy and/or behavioral content in the first-stage of a two-stage decomposition of a production problem. Suppose, for example, that policy and/or behavioral considerations impose only one constraint on the first input in the first sub-technology represented by $h(x_1^1) \leqslant \alpha$. Adding this constraint to (A.1) obtains

$$\begin{aligned}&\{(\boldsymbol{y}, \boldsymbol{x}, x_1^1) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\}\\&\quad = \{(\boldsymbol{y}, \boldsymbol{x}, x_1^1) \mid h(x_1^1) \leqslant \alpha, (\boldsymbol{y}^i, \boldsymbol{x}^i) \in \Im_i(z^i, \boldsymbol{\varepsilon}), \boldsymbol{y} = \boldsymbol{AY}, \boldsymbol{x} = \boldsymbol{BX}, \boldsymbol{CZ} \leqslant \boldsymbol{K}\}.\end{aligned}$$

Under continuity and monotonicity, (A.2) thus becomes

$$F(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}, \alpha) = y_1 - f(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}, \alpha) = 0, \tag{A.2'}$$

$$y_1 = f(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}, \alpha) = \max\{y_1 \mid h(x_1^1) \leqslant \alpha, (\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\}.$$

This form is not dependent on the specific policy or behavioral constraint. That is, the constrained level α can be imposed, adjusted, or dropped in the second-stage problem. However, the form in (A.2′) is substantially different than that in (A.2). Indeed, the domain is a different space.

An alternative way to proceed is to define F with the constraint in place,

$$y_1 = f(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) = \max\{y_1 \mid h(x_1^1) \leqslant \alpha, (\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\}, \tag{A.2''}$$

with F defined as in (A.2). However, if this is done, the frontier is clearly policy- or behavior-dependent. The resulting PPF corresponding to (A.2′) will depend not only on the existence of a policy affecting x_1^1 as in (A.2″) but also on the policy level, α. Likewise, the profit function dual to F will also depend on α. Thus, the true PPF will be policy-dependent and what may appear as technical change or inefficiency in the PPF associated with (A.2″) may be due to changes in policy. Similarly, if the PPF from the first stage is based on (A.2″), then second-stage analyses cannot consider changes in policy or behavior associated with α.

PROOF OF PROPOSITION 4. Because Proposition 4 is stated in negative form, the above example suffices as a proof. □

Using the partial aggregation definitions associated with (46), the definitions of technical allocative efficiency can be generalized and divided into policy- and behavior-relevant and policy- or behavior-dependent categories under Aggregation Qualification Condition A.2.

DEFINITION A.1′.

(E3′) $(\boldsymbol{y}^*, \boldsymbol{x}^*, \boldsymbol{z}^*)$ satisfies fixed factor technical allocative efficiency for a given $\boldsymbol{\varepsilon}$ if there does not exist an alternative allocation $\boldsymbol{z}^{*\prime}$ such that $(\boldsymbol{y}^{*\prime}, \boldsymbol{x}^*, \boldsymbol{z}^{*\prime}) \in \Im^*(\boldsymbol{k}, \boldsymbol{\varepsilon})$ in (46) and $\boldsymbol{y}^{*\prime} \not\leqslant \boldsymbol{y}$.

(E4′) $(\boldsymbol{y}^*, \boldsymbol{x}^*, \boldsymbol{z}^*)$ satisfies variable input technical allocative efficiency for a given $\boldsymbol{\varepsilon}$ if there does not exist an alternative allocation $\boldsymbol{x}^{*\prime}$ such that $(\boldsymbol{y}^{*\prime}, \boldsymbol{x}^{*\prime}, \boldsymbol{z}^*) \in \Im^*(\boldsymbol{k}, \boldsymbol{\varepsilon})$, $\boldsymbol{x}^{*\prime} \in C(\boldsymbol{x}^*)$, and $\boldsymbol{y}^{*\prime} \not\leqslant \boldsymbol{y}$ where $C(\boldsymbol{x}^*)$ defines the set of possible purchased variable input allocations with a given vector of aggregate purchases, i.e., $C(\boldsymbol{x}^*) = \{\boldsymbol{x}^{*\prime} \mid (\boldsymbol{y}^{*\prime}, \boldsymbol{x}^{*\prime}, \boldsymbol{z}^*) = H(\boldsymbol{Y}', \boldsymbol{X}', \boldsymbol{Z}), \boldsymbol{B}\boldsymbol{X}' \leqslant \boldsymbol{B}\boldsymbol{X}\}$.

(E5′) $(\boldsymbol{y}^*, \boldsymbol{x}^*, \boldsymbol{z}^*)$ satisfies output technical allocative efficiency for a given $\boldsymbol{\varepsilon}$ if there does not exist an alternative allocation among sub-technologies $\boldsymbol{y}^{*\prime}$ such that $(\boldsymbol{y}^{*\prime}, \boldsymbol{x}^{*\prime}, \boldsymbol{z}^{*\prime}) \in \Im^*(\boldsymbol{k}, \boldsymbol{\varepsilon})$, $\boldsymbol{x}^{*\prime} \in C(\boldsymbol{x}^*)$, and $\boldsymbol{y}^{*\prime} \in Y(\boldsymbol{y}^*)$ where $Y(\boldsymbol{y}^*)$ defines the set of outputs that correspond to a dominant aggregate output vector, i.e., $Y = \{\boldsymbol{y}^{*\prime} \mid (\boldsymbol{y}^{*\prime}, \boldsymbol{x}^{*\prime}, \boldsymbol{z}^*) = H(\boldsymbol{Y}', \boldsymbol{X}', \boldsymbol{Z}), \boldsymbol{A}\boldsymbol{Y}' \not\leqslant \boldsymbol{A}\boldsymbol{Y}\}$.

(E8) $(\boldsymbol{y}^*, \boldsymbol{x}^*, \boldsymbol{z}^*)$ is policy- and behavior-relevant if it satisfies the Aggregation Qualification Conditions, i.e., preserves distinction for all input and output quantities that have distinct prices, distinct policy controls, distinct ex post adjustment possibilities or distinct behavioral preferences and implications so that (46) preserves full generality of policy and behavioral issues to the second stage.

(E9) (y^*, x^*, z^*) is policy- or behavior-dependent if it does not satisfy the Aggregation Qualification Condition, i.e., does not preserve distinction for all input and output quantities that have distinct prices, distinct policy controls, distinct ex post adjustment possibilities or distinct behavioral preferences and implications, in which case (46) does not preserve the full generality of policy and behavioral issues to the second stage.

Using these definitions jointly one can define, for example, technologies that satisfy various concepts of technical allocative efficiency and are also policy- and behavior-relevant. We submit that policy- and behavior-relevance as defined by (E8) must be a prerequisite requirement for investigating technical efficiency following (E3′), (E4′) or (E5′). Otherwise, tests of technical efficiency are actually confounded tests of policy and behavior that are not meaningful for investigating properties of technology.

A.3. Implicit representation of technical efficiency by scalar functions

All specific single-equation multi-output production functions in the literature of which we are aware satisfy $\partial F/\partial \boldsymbol{Y} \neq 0$ and $\partial F/\partial \boldsymbol{X} \neq 0$ whenever $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$. By comparison, a form such as

$$F(\boldsymbol{Y}, \boldsymbol{X}) \equiv \sum_{i=1}^{m} \left[F_i(\boldsymbol{Y}, \boldsymbol{X})\right]^{2\nu}, \tag{A.6}$$

where ν is a positive integer can impose multiple constraints of the form $F_i(\boldsymbol{Y}, \boldsymbol{X}) = 0$ but yields $\partial F/\partial \boldsymbol{Y} = 0$ and $\partial F/\partial \boldsymbol{X} = 0$ whenever $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$. Mittelhammer, Matulich, and Bushaw (1981) have shown that such forms do not lend themselves to Lagrangians, Kuhn–Tucker conditions, nor the implicit function theorem. To illustrate, the profit maximization problem $\max_{\boldsymbol{Y},\boldsymbol{X}}\{\boldsymbol{PY} - \boldsymbol{RX} \mid F(\boldsymbol{Y}, \boldsymbol{X}) = 0\}$ can be expressed as the Lagrangian $\mathcal{L} = \boldsymbol{PY} - \boldsymbol{RX} + \lambda F(\boldsymbol{Y}, \boldsymbol{X})$ for which first-order conditions cannot be solved when $\partial F/\partial \boldsymbol{Y} = 0$ and $\partial F/\partial \boldsymbol{X} = 0$ at $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$.[55]

Two approaches to implicit representation are possible: (i) require non-zero derivatives of F with respect to all relevant inputs and outputs when $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$, or (ii) develop methods to deal with zero derivatives. The first approach facilitates standard mathematical manipulations, but obtains a model that can neither reflect input allocations nor

[55] More technically, the Jacobian of the constraint does not have full rank when $F(\boldsymbol{Y}, \boldsymbol{X})$ has all zero derivatives so the first-order conditions of the Lagrangian are not necessary. For Kuhn–Tucker problems, constraint qualification fails.

impose by-product relationships. By the implicit function theorem, $F(\boldsymbol{Y}, \boldsymbol{X}) = 0$ yields a function such as $y_i = f_i(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_{n_y}, \boldsymbol{X})$ where all partial derivatives are non-zero. Thus, any input change (even though it may represent an allocation of a distinct input to production of a distinct output) can be transformed into a change in any other output (rather than the output to which it is allocated). Similarly, any transformation between distinct outputs (possibly between a primary output and a by-product that, in reality, can be produced only in fixed proportions) is allowed.

Alternatively, one can argue that assuming non-zero derivatives for F is very restrictive. By comparison, the representation in (A.6) clearly has all zero derivatives and yet implies[56]

$$F(\boldsymbol{Y}, \boldsymbol{X}) \underset{m}{\equiv} \begin{bmatrix} F_1(\boldsymbol{Y}, \boldsymbol{X}) \\ \vdots \\ F_m(\boldsymbol{Y}, \boldsymbol{X}) \end{bmatrix} \underset{m}{=} 0. \tag{A.7}$$

Conversely, one can always transform (A.7) into (A.6) but, interestingly, the Implicit Function Theorem (which is only sufficient, not necessary) can apply in (A.7) even if not in (A.6).

An interesting question is, can the product transformation functions implied by standard profit and cost function specifications admit forms such as (A.7)? The answer is no. A standard axiom of duality is convexity of the feasible technology set $\Im$ which is defined as the set of all feasible combinations of $(\boldsymbol{Y}, \boldsymbol{X})$. Diewert has shown for the single-output case that if $\Im$ is convex then the corresponding product transformation function, $F(\boldsymbol{Y}, \boldsymbol{X})$, is a convex function. However, $F(\boldsymbol{Y}, \boldsymbol{X})$ in (A.6) is not necessarily convex even in the single-output case.[57] Furthermore, a form such as (A.6) is not monotonic in $\boldsymbol{Y}$ and $\boldsymbol{X}$. Thus, disposability does not correspond to an inequality in (A.6) and the producible output set cannot be defined as $\{\boldsymbol{Y} \mid F(\boldsymbol{Y}, \boldsymbol{X}) \leqslant 0\}$ as in the multi-output development of Chambers (1988). Moreover, all dual developments that derive production functions and product transformation functions do so by implicitly imposing a technical efficiency criterion leading to the convex hull of the technology

[56] Mittelhammer, Matulich, and Bushaw (1981) argue that vector-valued implicit functions are required to represent multi-output technology because single-valued functions with non-zero derivatives are restrictive. The result in (26) shows that the vector-valued representation they propose can be derived directly from a single-valued representation if zero derivatives are allowed. In either case, what really matters is the rank of the technological relationship (defined below) rather than the number of equalities used to describe it. Even if expressed in a vector-valued implicit form, one must still verify full rank of the system and make sure no equation discarded in getting to full rank is of a form such as (11) that embodies further restrictions implicitly. (Note that Mittelhammer, Matulich, and Bushaw apparently allow zero derivatives of individual scalar-valued functions contained in vector-valued representations.)

[57] To illustrate where x and y are scalar, let $F(y, x) = [y - f(x)]^2$ where $f(x)$ is concave and let $y_1 = f(x_1)$, $y_2 = f(x_2)$, $y_0 = \theta y_1 + (1-\theta) y_2$, and $x_0 = \theta x_1 + (1-\theta) x_2$. Then $F(y_0, x_0) = \{\theta y_1 + (1-\theta) y_2 - f[\theta x_1 + (1-\theta) x_2]\}^2 > 0 = \theta F(y_1, x_1) + (1-\theta) F(y_2, x_2)$. Thus, $F(y, x)$ is concave. Convexity is obtained, for example, if $F(y, x) = y - f(x)$.

set [e.g., Diewert (1974, 1982)]. Additionally, while not a serious problem for single-output problems, imposing both structural technical efficiency and technical allocative efficiency for multi-output production problems limits the rank of the resulting technology representation to unity so that structure (related to sub-technologies) cannot be represented (see Sections 4.6 and 4.7 for further details).[58]

PROOF OF PROPOSITION 5. Proposition 5 follows from the discussion of this section and results in Mittelhammer, Matulich, and Bushaw (1981). □

A.4. Controllability and rank of structural technology representations

To represent structure of technologies meaningfully, possible redundancies in (50) and (51) must be considered. For example, under continuity, the Jacobian of f_i may not be of full rank, e.g., one of the outputs of the sub-technology may be some function of another as in the case where $y_1^i = g(y_2^i)$, $g' > 0$. In this case, $\operatorname{rank}\{\partial(y_1^i, y_2^i)/\partial X; \partial(y_1^i, y_2^i)/\partial Z\} = 1$. Thus, the decision maker is not able to control the second output independently of the first. Such relationships represent the case of by-products. Similarly, constraints on allocated fixed inputs reduce the dimension of the input space when constraints are binding. Thus, it is helpful to consider the following definitions:

DEFINITION A.2. Let $\widetilde{Y} \in R_+^{n_a}$ be a subset of outputs in $Y \in R_+^{n_y}$ and let $N(\widetilde{Y}) \subset R_+^{n_a}$ denote a neighborhood of $\widetilde{Y}$. The mix of outputs $\widetilde{Y}$ is locally controllable in $R_+^{n_a}$ if $N(\widetilde{Y}) \subseteq \Im$.

If $n_a = 1$, then this neighborhood would correspond to an open set on the real line. If $n_a \geqslant 2$, then N corresponds to an open ball in multi-dimensional space. When controllability is not met, the producer does not have the flexibility to attain all output mixes in $N(\widetilde{Y})$.

DEFINITION A.3. Let $\widehat{Y} \in R_+^{n_b}$ be a subset of outputs in $Y \in R_+^{n_y}$. The outputs in $\widehat{Y}$ are by-products of $\widetilde{Y}$ under technology $\Im$ if there exists a non-trivial relationship in $\Im$ such that only one $\widehat{Y}$ exists for each $\widetilde{Y}$ given uncontrollable factors, i.e., $\widehat{Y} \underset{n_b}{=} g(\widetilde{Y}, \boldsymbol{\varepsilon})$.

The existence of by-product relationships reduces the producer's flexibility in choosing output mixes. The remaining flexibility after taking these relationships into account is described by the rank of a technology.

[58] If the PPF is defined conventionally by $F^*(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) \equiv y_1 - f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon})$ where $\boldsymbol{y} = (y_1, \boldsymbol{y}_{-1})$ and $y_1 = f^*(\boldsymbol{y}_{-1}, \boldsymbol{x}, \boldsymbol{k}, \boldsymbol{\varepsilon}) \equiv \max\{y_1 \mid (\boldsymbol{y}, \boldsymbol{x}) \in \Im_{-i}(\boldsymbol{k}, \boldsymbol{\varepsilon})\}$ and if no fixed factors are allocated, then the same function is obtained as in (A.2) upon imposing technical allocative efficiency with respect to inputs and outputs.

DEFINITION A.4. The rank of a technology is the dimension of the largest locally controllable mix of outputs.

To proceed, suppose that the relationships in (50) and (51) are continuous and differentiable.

LEMMA A.1. *Under continuity and differentiability, the rank of a technology*

$$\boldsymbol{Y} \underset{m \cdot n_y}{=} f(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon})$$

is given by $\operatorname{rank}(f_X, f_Z)$.

PROOF OF LEMMA A.1. Let $\boldsymbol{W} = (\boldsymbol{X}, \boldsymbol{Z})$ and $\operatorname{rank}(f_X, f_Z) = \operatorname{rank}(f_W) = \rho$. Then there exists a nonsingular $\rho \times \rho$ Jacobian f_W^* as a submatrix of f_W after appropriate reordering of $\boldsymbol{Y}$ and $\boldsymbol{W}$. Corresponding to f_W^* are equations $\boldsymbol{Y}^* \underset{\rho}{=} f^*(\boldsymbol{W}^*, \boldsymbol{W}^{**})$ which together with $\boldsymbol{Y}^{**} = f^{**}(\boldsymbol{W}^*, \boldsymbol{W}^{**})$ represent a partitioning of $\boldsymbol{Y} = f(\boldsymbol{W})$ where $\boldsymbol{W} = (\boldsymbol{W}^*, \boldsymbol{W}^{**})$ is a corresponding partitioning of $\boldsymbol{W}$. By the Implicit Function Theorem, the equation $\boldsymbol{Y}^* \underset{\rho}{=} f^*(\boldsymbol{W}^*, \boldsymbol{W}^{**})$ has a solution where $\boldsymbol{W}^*$ are the active or endogenous variables. Let $\boldsymbol{W}_0^*, \boldsymbol{W}_0^{**}, \boldsymbol{Y}_0^*$ be such a solution. By the Implicit Function Theorem, there is an open ball $\mathcal{B}(\boldsymbol{Y}_0^*)$ such that $\boldsymbol{Y}^* = \{f^*(\boldsymbol{W}^*, \boldsymbol{W}^{**}) \mid \boldsymbol{W}^{**} = \boldsymbol{W}_0^{**}\}$ is a one-to-one transformation for all $\boldsymbol{Y}_0^* \in \mathcal{B}(\boldsymbol{Y}_0^*)$ and $\boldsymbol{W}^*$ near $\boldsymbol{W}_0^*$. Hence, $\boldsymbol{Y}^*$ of dimension ρ is locally controllable. □

LEMMA A.2. *Where the rank of a technology is n_a and $\widetilde{\boldsymbol{Y}}$ is a locally controllable vector of outputs in $R_+^{n_a}$, the complete output vector can be partitioned into $\boldsymbol{Y} = (\widetilde{\boldsymbol{Y}}, \widehat{\boldsymbol{Y}})$ where the choice of $\widetilde{\boldsymbol{Y}}$ determines the other outputs in $\widehat{\boldsymbol{Y}} \in R_+^{n_b}$ and $n_b = n_y - n_a$, i.e., the number of by-products in a multi-output technology is equal to the number of outputs minus the rank of the technology.*

PROOF OF LEMMA A.2. Consider the production relations in (51) and assume $\operatorname{rank}(f_X, f_Z) = n_a$. By the Inverse Function Theorem and Definition A.2, $\widetilde{\boldsymbol{Y}}$ can be found as a function of $(\boldsymbol{X}, \boldsymbol{Z})$ for a given $\boldsymbol{\varepsilon}$, say $\widehat{\boldsymbol{Y}} \underset{n_a}{=} \tilde{f}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon})$. This relationship is simply a subset consisting of n_a of the individual equations contained in (51). Using Definition A.3, $\widehat{\boldsymbol{Y}} \underset{n_b}{=} g(\tilde{f}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\varepsilon}))$, where $n_b = \sum_i k_i - n_a$. □

From the proof of Lemma A.2, the gradient of f only spans an n_a-dimensional space. In particular, the Jacobian of $\widehat{\boldsymbol{Y}}$ is $g_{\widetilde{Y}} \cdot (\tilde{f}_X, \tilde{f}_Z)$ which is a linear transformation of the Jacobian of $\widetilde{\boldsymbol{Y}}$ given by $(\tilde{f}_X, \tilde{f}_Z)$, which itself has only rank n_a. Next consider input controllability.

DEFINITION A.5. Let $\widetilde{Z} \in R_+^{n_f}$ be a subset of inputs included in $Z \in R_+^{n_z}$ and let $N(\widetilde{Z}) \subset R_+^{n_f}$ denote a neighborhood of $\widetilde{Z}$. The mix of inputs $\widetilde{Z}$ is locally controllable in $R_+^{n_f}$ if $N(\widetilde{Z}) \subseteq \Im$. A subset of inputs is locally restricted if it is not locally controllable.

Even though there are n_z allocated fixed input decisions, only $n_f = n_z - n_c$ of them are freely controllable. Generalizing to the possibility of nonlinear constraints, let $\widehat{Z} \underset{n_c}{=} h(\widetilde{Z}, K)$ represent the set of allocated fixed inputs determined by the choice of $\widetilde{Z}$ given K. Under continuity and differentiability, a parsimonious nonlinear representation of the binding (non-redundant) constraints will have a Jacobian of full rank.

LEMMA A.3. *Let the vector of all constrained inputs be denoted by $Z \in R_+^{n_z}$ and let all locally binding input constraints in $\Im$ be summarized by $\widehat{Z} \underset{n_c}{=} h(\widetilde{Z}, K)$ with full rank Jacobian, $h_{\widetilde{Z}}$. Then the input vector can be partitioned into $Z = (\widetilde{Z}, \widehat{Z})$ where $\widetilde{Z} \in R_+^{n_z - n_c}$ is locally controllable and the choice of $\widetilde{Z}$ determines the other inputs in $\widehat{Z} \in R_+^{n_c}$.*

PROOF OF LEMMA A.3. The proof is omitted because it is similar to Lemma A.1. □

Note that this lemma is worded generally so as to apply to all forms of constraints whether associated with firm-owned resources, policy instruments, behavior, or market rationing and whether applicable to allocated fixed inputs or purchased variable inputs.

PROOF OF PROPOSITION 6. The proof of Proposition 6 follows the Fundamental Axiom, which permits technology to be represented as in (51), and from Lemmas A.1–A.3 under continuity and differentiability. □

PROOF OF PROPOSITION 7. This proof is omitted for brevity since it is sketched clearly in the text. □

PROOF OF PROPOSITION 8. If n_c fixed inputs must be allocated among m subtechnologies, then at least $n_c(m-1)$ allocation variables are unobserved. From Proposition 7, the maximum number of non-redundant observable controllable equations is thus $n_a + n_c - n_c(m-1)$. This number is greater than zero only if $n_a \geqslant n_c m$. In the case of nonjointness, $m = n_y = n_a$ in which case this condition reduces directly to $n_c \leqslant 1$. In the case where some variable input allocations are unobserved but their aggregates are observed, a similar proof applies where (i) $x = BX$ is used to substitute into (52)–(54), (ii) the number of such variable inputs is added to n_c, and (iii) the associated number of unobserved allocations are considered in the calculation. For allocated fixed inputs without binding restrictions, note that m rather than $m-1$ allocation variables are unobserved so even more variables are unobserved. □

PROOF OF COROLLARY 2. Since n_c fixed inputs must be allocated among m subtechnologies, then at least $(n_c - 1)(m-1)$ allocation variables are unobserved. From

Proposition 7, the maximum number of non-redundant observable controllable equations is thus $n_a + n_c - (n_c - 1)(m - 1)$. This number is greater than zero only if $n_a + m - 1 \geqslant n_c m$. If $n_a < (n_c - 1)m$, then $n_a + m - 1 < n_c m - 1$. Other assertions follow as in the proof of Proposition 8. □

PROOF OF PROPOSITION 9. Under the conditions of Proposition 8, no purely technological relationship among inputs and outputs is estimable. All estimable relationships of $\boldsymbol{y}$ and $\boldsymbol{x}$ obtained from solving the production problem with conditions (51) and (58) must embody policy or behavioral criteria. Thus, hypothesis tests on the relationship of input and output variables cannot test the structure of technology alone, but rather test the relationship of variables induced by a combination of behavioral criteria and technology. □

References

Allen, D.W., and D.L. Lueck (1998), "The nature of the farm", Journal of Law and Economics 41: 343–386.

Alston, J.M., and J.A. Chalfant (1991), "Unstable models from incorrect forms", American Journal of Agricultural Economics 73:1171–1181.

Amemiya, T. (1985), Advanced Econometrics (Harvard University Press, Cambridge).

Antle, J.M. (1983), "Testing the stochastic structure of production: A flexible moment-based approach", Journal of Business and Economic Statistics 1:192–201.

Antle, J. (1984), "The structure of U.S. agricultural technology", American Journal of Agricultural Economics 66:414–421.

Antle, J.M., S.M. Capalbo and C.C. Crissman (1994), "Econometric production models with endogenous input timing: An application to Ecuadorian potato production", Journal of Agricultural And Resource Economics 19:1–18.

Antle, J., and S. Hatchett (1986), "Dynamic input decisions in econometric production models", American Journal of Agricultural Economics 68:939–949.

Antle, J.M., and R.E. Just (1992), "Conceptual and empirical foundations for agricultural-environmental policy analysis", Journal of Environmental Quality 21:307–316.

Baker, T.G., and B.A. McCarl (1982), "Representing farm resource availability over time in linear programs: A case study", North Central Journal of Agricultural Economics 4:59–68.

Ball, V.E. (1985), "Output, input, and productivity measurement in U.S. agriculture, 1948–79", American Journal of Agricultural Economics 67:425–436.

Ball, V.E. (1988), "Modeling supply response in a multiproduct framework", American Journal of Agricultural Economics 70:813–825.

Ball, V.E., J.-C. Bureau, R. Nehring and A. Somwaru (1997), "Agricultural productivity revisited", American Journal of Agricultural Economics 79:1045–1063.

Ball, V.E., and R.G. Chambers (1982), "An economic analysis of technology in the meat products industry", American Journal of Agricultural Economics 64:699–709.

Ball, V.E., F. Gollop, A. Kelly and G. Swinand (1999), "Patterns of productivity growth in the U.S. farm sector: Linking state and aggregate models", American Journal of Agricultural Economics 81:164–179.

Barry, P.J., and L.J. Robison (2001), "Agricultural finance: Credit, credit constraints, and consequences", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 513–571.

Bar-Shira, Z., R.E. Just and D. Zilberman (1997), "Estimation of farmers' risk attitude: An econometric approach", Agricultural Economics 17:211–222.

Baumol, W.J., J.C. Panzar and R.D. Willig (1988), Contestable Markets and the Theory of Industry Structure, Revised Edition (Harcourt Brace Jovanovich, San Diego).

Berck, P., and G. Helfand (1990), "Reconciling the Von Liebig and differentiable crop production functions", American Journal of Agricultural Economics 72:985–996.

Binswanger, H.P. (1974), "The measurement of technological change biases with many factors of production", American Economic Review 64:964–976.

Binswanger, H.P. (1980), "Attitudes toward risk: Experimental measurement in rural India", American Journal of Agricultural Economics 62:395–407.

Binswanger, H.P. (1981), "Attitudes toward risk: Theoretical implications of an experiment in rural India", Economic Journal 91:867–890.

Blackorby, C., D. Primont and R.R. Russell (1978), Duality, Separability and Functional Structure (North-Holland, New York).

Bryant, K.J., J.W. Mjelde and R.D. Lacewell (1993), "An intraseasonal dynamic optimization model to allocate irrigation water between crops", American Journal of Agricultural Economics 75:1021–1029.

Burger, B. (1998), "Delayed corn planting – what are your options?", Novartis Seeds, http://www.nk.com/nkwired/1998/05_04_delay.html.

Burt, O. (1993), "Decision rules for the dynamic animal feeding problem", American Journal of Agricultural Economics 75:190–202.

Buschena, D., and D. Zilberman (2000), "Generalized expected utility, heteroscedastic error, and path dependence in risky choice", Journal of Risk and Uncertainty 20:67–88.

Candler, W., and M. Boehlje (1971), "The use of linear programming in capital budgeting with multiple goals", American Journal of Agricultural Economics 53:325–330.

Chambers, R.G. (1988), Applied Production Analysis (Cambridge, New York).

Chambers, R.G., Y. Chung and R. Färe (1996), "Benefit and distance functions", Journal of Economic Theory 70:407–419.

Chambers, R.G., and R.E. Just (1981), "Effects of exchange rate changes on U.S. agriculture: A dynamic analysis", American Journal of Agricultural Economics 63:32–46.

Chambers, R.G., and R.E. Just (1989), "Estimating multioutput technologies", American Journal of Agricultural Economics 71:980–995.

Chambers, R.G., and E. Lichtenberg (1995), "The Economics of sustainable agriculture in the Mid-Atlantic", Department of Agricultural and Resource Economics (University of Maryland, College Park).

Chambers, R.G., and R. Pope (1991), "Testing for consistent aggregation", American Journal of Agricultural Economics 73:808–818.

Chambers, R.G., and J. Quiggin (1998), "Cost functions for stochastic technologies", American Journal of Agricultural Economics 80:288–295.

Chambers, R.G., and U. Vasavada (1983), "Testing asset fixity for U.S. agriculture", American Journal of Agricultural Economics 65:761–769.

Chavas, J.-P. (2001), "Structural change in agricultural production: Economics, technology and policy", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 263–285.

Chavas, J.-P., and T.L. Cox (1988), "A nonparametric analysis of agricultural technology", American Journal of Agricultural Economics 70:303–310.

Chavas, J.-P. and M.T. Holt (1991), "On non-linear dynamics: The case of the pork cycle", American Journal of Agricultural Economics 73:819–828.

Chavas, J.-P., and M.T. Holt (1993), "Market instability and nonlinear dynamics", American Journal of Agricultural Economics 75:113–120.

Chavas, J.-P. and M.T. Holt (1996), "Economic behavior under uncertainty: A joint analysis of risk preferences and technology", Review of Economics and Statistics 78:329–335.

Clark, J.S., and G.A. Carlson (1990), "Testing for common versus private property: The case of pesticide resistance", Journal of Environmental Economics and Management 19:45–60.

Coase, R. (1937), "On the nature of the firm", Economica 4:386–405.

Cragg, J.G., and S.G. Donald (1997), "Inferring the rank of a matrix", Journal of Econometrics 76:223–250.
Day, R.H. (1965), "Probability distributions of field crop yields", Journal of Farm Economics 47:713–741.
de Janvry, A. (1972), "Optimal Levels of fertilization under risk: The potential for corn and wheat fertilization under alternative price policies in Argentina", American Journal of Agricultural Economics 54:1–10.
de Koning G.H.J., H. Van Keulen, R. Rabbinge and H. Hanssen (1995), "Determination of input and output coefficients of cropping systems in the European Community", Agricultural Systems 48:485–502.
Debreu, G. (1959), Theory of Value (Wiley, New York).
Deininger, K., and G. Feder (2001), "Land institutions and land markets", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 287–331.
Diewert, W.E. (1974), "Applications of duality theory", in: M.D. Intrilligator and D.A. Kendrick, eds., Frontiers of Quantitative Economics, Vol. II (North-Holland, Amsterdam) 106–199.
Diewert, W.E. (1976), "Exact and superlative index numbers", Journal of Econometrics 4:115–145.
Diewert, W.E. (1982), "Duality approaches to microeconomic theory", in: K. Arrow and M. Intrilligator, eds., Handbook of Mathematical Economics, Vol. 2 (North-Holland, Amsterdam).
Diewert, W.E., and T.J. Wales (1987), "Flexible functional forms and global curvature conditions", Econometrica 55:43–68.
Dhrymes, P.J. (1973), "Restricted and unrestricted reduced forms: Asymptotic distribution and relative efficiency", Econometrica 41:119–134.
Doll, J. (1967), "An analytical technique for estimating weather indexes from meteorological measurement", American Journal of Agricultural Economics 49:79–88.
Eeckhoudt, L., and C. Gollier (1992), Risk: Evaluation, Management and Sharing (Harvester Wheatsheaf, New York).
Economic Research Service (1999a), Corn Costs and Returns: 1997–98 Costs of Production from the Agricultural Resource Management Study (U.S. Government Printing Office, Washington D.C.).
Economic Research Service (1999b), Soybean Costs and Returns: 1997–98 Costs of Production from the Agricultural Resource Management Study (U.S. Government Printing Office, Washington D.C.).
Economic Research Service (various years), Wheat Yearbook (U.S. Government Printing Office, Washington D.C.).
Evenson, R.E. (2001), "Economic impacts of agricultural research and extension", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 573–628.
Färe, R. (1996), "The opportunity cost of duality", Journal of Productivity Analysis 7:213–224.
Farrell, M.J. (1957), "The measurement of productive efficiency", Journal of the Royal Statistical Society, Series A, General 120:253–281.
Feder, G. (1979), "Pesticides, information, and pest management under uncertainty", American Journal of Agricultural Economics 61:97–103.
Feder, G., R.E. Just and D. Zilberman (1985), "Adoption of agricultural innovations in developing countries: A survey", Economic Development and Cultural Change 33:255–298.
Fisher, I. (1922), The Making of Index Numbers (Houghton Mifflin, Boston).
Foster, A.D., and M.R. Rosenzweig (1995), "Learning by doing and learning from others: Human capital and technical change in agriculture", Journal of Political Economy 103:1176–1209.
Foster, K.A., and O.R. Burt (1992), "A dynamic model of investment in the U.S. beef-cattle industry", Journal of Business and Economic Statistics 10:419–426.
French, B., G. King and D. Minami (1985), "Planting removal relationships for perennial crops: An application to cling peaches", American Journal of Agricultural Economics 67:215–223.
Fried, H.O., C.A.K. Lovell and S.S. Schmidt (1993), The Measurement of Productive Efficiency: Techniques and Applications (Oxford University Press, Oxford).
Fuss, M.A. (1977), "The structure of technology over time: A model for testing the putty-clay hypothesis", Econometrica 45:1797–1821.
Griliches, Z. (1963), "The sources of measured productivity growth: U.S. agriculture, 1940–1960", Journal of Political Economy 71:309–322.
Hicks, J.R. (1956), A Revision of Demand Theory (Oxford University Press, London).

Hochman, E., and D. Zilberman (1978), "Examination of environmental policies using production and pollution microparameter distributions", Econometrica 46:739–760.

Horowitz, J.K., and E. Lichtenberg (1993), "Insurance, moral hazard, and chemical use in agriculture", American Journal of Agricultural Economics 75:926–935.

Hueth, D., and U. Regev (1974), "Optimal agricultural pest management with increasing pest resistance", American Journal Agricultural Economics 56:543–552.

Huffman, W.E. (2001), "Human capital: Education and agriculture", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 333–381.

Jensen, H. (1977), "Farm management and production economics, 1946–70", in: Lee Martin, ed., A Survey of Agricultural Economics Literature (University of Minneapolis Press, Minneapolis).

Jensen, F.E., J.S. Lawson and L.N. Langemeier (1993), "Agricultural investment and internal cash flow variables", Review of Agricultural Economics 15:295–306.

Johansen, L. (1972), "Production functions: An integration of micro and macro, short-run and long-run aspects" (North-Holland, Amsterdam).

Johnson, G.L. (1955), "Results from production economics analysis", Journal of Farm Economics 37:206–222.

Johnson, G.L. (1956), "Supply functions – some facts and notions", in: E. Heady et al., eds, Agricultural Adjustment Problems in a Growing Economy (Iowa State University, Ames).

Johnson, G.L., and L.C. Quance (1972), The Overproduction Trap in U.S. Agriculture (Johns Hopkins University Press, Baltimore).

Just, R.E. (1974), "An Investigation of the importance of risk in farmers' decisions", American Journal of Agricultural Economics 56:14–25.

Just, R.E., and J.M. Antle (1990), "Interactions between agricultural and environmental policies: A conceptual framework", American Economic Review 80:197–202.

Just, R.E., and W. Candler (1985), "Production functions and rationality of mixed cropping: The Nigerian case", European Review of Agricultural Economics 12:207–231.

Just, R.E., and R.D. Pope (1978), "Stochastic specification of production functions and economic implications", Journal of Econometrics 7:67–86.

Just, R.E., and R.D. Pope (1979), "Production function estimation and related risk considerations", American Journal Agricultural Economics 61:276–283.

Just, R.E., and R.D. Pope (1999), "Implications of heterogeneity for theory and practice in production economics", American Journal of Agricultural Economics 81:711–718.

Just, R.E., and G.C. Rausser (1983), "Applied commodity price analysis, forecasting, and market risk management", NCR-134 Conference (Des Moines, IA).

Just, R.E., and Q. Weninger (1999), "Are crop yields normally distributed?", American Journal of Agricultural Economics 81:287–304.

Just, R.E., and D. Zilberman (1983), "Stochastic structure, farm size, and technology adoption in developing agriculture", Oxford Economic Papers 35:307–328.

Just, R.E., D. Zilberman and E. Hochman (1983), "Estimation of multicrop production functions", American Journal Agricultural Economics 65:770–780.

Just, R.E., D. Zilberman, E. Hochman and Z. Bar-Shira (1990), "Input allocation in multicrop systems", American Journal Agricultural Economics 72:200–209.

Kahneman, D., and A. Tversky (1979), "Prospect theory: An analysis of decision under risk", Econometrica 47:263–291.

Keplinger, K.O., B.A. McCarl, M.E. Chowdhury and R.D. Lacewell (1998), "Economic and hydrologic implications of implementing a dry year option for the Edwards aquifer", Journal of Agricultural and Resource Economics 23:191–205.

Klein, L.R. (1947), "The use of cross-section data in econometrics with application to a study of production of railroad services in the united states" (National Bureau of Economic Research, Washington D.C.).

Knudsen, N. (1973), Production and Cost Models of a Multi Product Firm (Odense University Press, Odense, Denmark).

Kutcher, G.P., and P. Scandizzo (1981), The Agricultural Economy of Northeast Brazil (Johns Hopkins Press, Baltimore).

Lau, L.J. (1978), "Applications of profit functions", in: M. Fuss and D. McFadden, eds., Production Economics: A Dual Approach to Theory and Applications (North-Holland, Amsterdam) 133–216.

Leathers, H. (1991), "Allocatable fixed inputs as a cause of joint production: A cost function approach", American Journal of Agricultural Economics 73:1083–1090.

Leontief, W. (1971), "Theoretical assumptions and nonobserved facts", American Economic Review 41:1–7.

Loehman, E., and C. Nelson (1992), "Optimal risk management, risk aversion, and production function properties", Journal of Agricultural and Resource Economics 17:219–231.

Love, H.A., and S. Buccola (1991), "Joint risk preference-technology estimation with a primal system", American Journal of Agricultural Economics 73:763–773.

Machina, M.J. (1987), "Choice under uncertainty: Problems solved and unsolved", Journal of Economic Perspectives 1:121–154.

McCarl, B.A., W.V. Candler, D.H. Doster and P.R. Robbins (1977), "Experiences with farmer oriented linear programming for crop planning", Canadian Journal of Agricultural Economics 25:17–30.

McCloskey, D. (1998), The Rhetoric of Economics, 2nd edn. (University of Wisconsin Press, Madison).

McElroy, M. (1987), "Additive general error models for production, cost, and derived demand or share systems", Journal of Political Economy 95:737–757.

McFadden, D. (1978), "Cost, revenue and profit functions", in: M. Fuss and D. McFadden, eds., Production Economics: A Dual Approach to Theory and Applications (North-Holland, Amsterdam).

McQuigg, J.D., and J.P. Doll (1961), "Economic analysis involving random weather inputs,", Journal of Farm Economics 43:909–20.

Mittelhammer, R.C., S.C. Matulich and D. Bushaw (1981), "On implicit forms of multiproduct-multifactor production functions", American Journal of Agricultural Economics 63:164–168.

Morrison, C.J., and D. Siegel (1998), "Knowledge capital and cost structure in U.S. food and fiber industries", American Journal of Agricultural Economics 80:30–45.

Moschini, G. (forthcoming), "Production risk and the estimation of ex-ante cost functions", Journal of Econometrics.

Moschini, G., and D.A. Hennessy (2001), "Uncertainty, risk aversion, and risk management for agricultural producers", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 87–153.

Mundlak, Y. (1996), "Production function estimation: Reviving the primal", Econometrica 64:431–438.

Mundlak, Y. (2001), "Production and supply", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 3–85.

Nelson, C.H., and P. Preckel (1989), "The conditional beta distribution as a stochastic production function", American Journal of Agricultural Economics 71:370–378.

Nerlove, M. (1983), "Expectations, plans and realizations in theory and practice", Econometrica 51:1251–1279.

Nerlove, M., and D.A. Bessler (2001), "Expectations, information and dynamics", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 155–206.

Pfouts, R.W. (1961), "The theory of cost and production in the multiproduct firm", Econometrica 39:650–668.

Paris, Q. (1992), "The von Liebig hypothesis", American Journal of Agricultural Economics 74:1019–1028.

Paris, Q., and R. Howitt (1998), "An analysis of ill-posed production problems using maximum entropy", American Journal of Agricultural Economics 80:124–138.

Pope, R. (1976), "An analysis of factors affecting farm diversification: With special reference to San Joaquin Valley crop farms", Unpublished Ph.D. Dissertation (University of California, Berkeley).

Pope, R., and R. Chambers (1989), "Price aggregation when price-taking firms' prices vary", Review of Economic Studies 56:297–310.

Pope, R., and A. Hallam (1986), "A confusion of agricultural economists? A profession interest survey and essay", American Journal of Agricultural Economics 68:572–594.

Pope, R.D., and R.E. Just (1991), "On testing the structure of risk preferences in agricultural supply analysis", American Journal of Agricultural Economics 73:743–748.

Pope, R.D., and R.E. Just (1996), "Empirical implementation of ex ante cost functions", Journal of Econometrics 25:231–249.

Pope, R.D., and R.E. Just (1998), "Cost function estimation under risk aversion", American Journal of Agricultural Economics 80:296–302.

Pope, R.D., and R.E. Just (2000a), "Random profits with duality", Working paper (Brigham Young University and University of Maryland).

Pope, R.D., and R.E. Just (2000b), "Distinguishing errors in measurement from errors in optimization", Working paper (Brigham Young University and University of Maryland).

Pope, R., and R. Prescott (1980), "Diversification in relation to farm size and other socioeconomic characteristics", American Journal of Agricultural Economics 62:554–559.

Quiggin, J. (1982), "A theory of anticipated utility", Journal of Economic Behavior and Organization 3:323–343.

Regev, U., N. Gotsch and P. Rieder (1997), "Are fungicides, nitrogen, and plant growth regulators risk-reducing? Empirical evidence from Swiss wheat production", Journal of Agricultural Economics 48:167–178.

Samuelson, P.A. (1967), Foundations of Economic Analysis (Atheneum, New York).

Saha, A., C.R. Shumway and H. Talpaz (1994), "Joint estimation of risk preference structure and technology using expo-power utility", American Journal of Agricultural Economics 76:173–184.

Salin, V., A. Thurow, K. Smith and N. Elmer (1998), "Exploring the market for agricultural economics information: Views of private sector analysis", Review of Agricultural Economics 20:114–124.

Schultz, T.P. (2001), "Women's roles in the agricultural household: Bargaining and human capital investments", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 383–456.

Shumway, C.R. (1983), "Supply, demand, and technology in a multi-product industry: Texas field crops", American Journal of Agricultural Economics 65:748–760.

Shumway, C.R. (1988), "The statistical base for agricultural productivity research: A review and critique", in: S.M. Capalbo and J.M. Antle, eds., Agricultural Productivity Measurement and Explanation (Resources for the Future, Washington, D.C.) 138–156.

Shumway, C.R. (1995), "Recent duality contributions in production economics", Journal of Agricultural and Resource Economics 20:178–194.

Shumway, C.R., and H. Lim (1993), "Functional form and U.S. agricultural production elasticities", Journal of Agricultural and Resource Economics 18:330–337.

Shumway, C.R., R.D. Pope and E.K. Nash (1984), "Allocatable fixed inputs and jointness in agricultural production: Implications for economic modeling", American Journal of Agricultural Economics 66:72–78.

Smale, M., R.E. Just and H.D. Leathers (1994), "Land allocation in hyv adoption models: An investigation of alternative explanations", American Journal of Agricultural Economics 76:535–546.

Sunding, D., and D. Zilberman (2001), "The agricultural innovation process: Research and technology adoption in a changing agricultural sector", in: B. Gardner and G. Rausser, eds., Handbook of Agricultural Economics, Vol. 1 (Elsevier, Amsterdam) 207–261.

United States Senate (1987), Hearings before the Committee on Agriculture, Nutrition, and Forestry, Federal Insecticide, Fungicide, and Rodenticide Act, 100th Congress, First Session, S. 1516, Part II, April 29, May 20, and July 30, p. 188–189.

Vanker, R.D. (1996), "Hypothesis tests using data envelopment analysis", Journal of Productivity Analysis 7:139–160.

Varian, H. (1992), Microeconomic Analysis, 3rd edn. (W.W. Norton and Co., New York).

Vasavada, U., and V.E. Ball (1988), "A dynamic investment model within a multi-commodity framework", Agricultural Economics 2:123–137.

Vasavada, U., and R.G. Chambers (1986), "Investment in agriculture", American Journal of Agricultural Economics 68:939–949.

Viner, J. (1931), "Cost curves and supply curves", Zeitschrift für Nationalökonomie and Statistik 3:23–46.

Weaver, R. (1983), "Multiple input, multiple output production choices and technology in the U.S. wheat region", American Journal of Agricultural Economics 65:45–56.

White, T., and G. Irwin (1972), "Farm size and specialization", in: G. Ball and E. Heady, eds., Size, Structure and Future of Farms (Iowa State University Press, Ames).

Wolf, S., D.R. Just and D. Zilberman (forthcoming), "Between data and decisions: The organization of agricultural economic information systems", Research Policy.

Woodward, S. (1996), "A dynamic nutrient carryover model for pastoral soils and its application to optimising fertiliser allocation to several blocks with a cost constraint", Review of Marketing and Agricultural Economics 64:75–85.