APPENDIX II

# The InSTePP Production Accounts

Version 4 of the InSTePP production accounts is the latest in a series of accounts all developed under the leadership of Philip Pardey at the University of Minnesota's International Science and Technology Practice and Policy (InSTePP) center as a joint effort with colleagues now at Oberlin College (Barbara Craig), the University of Wyoming (Matthew Andersen), and the University of California, Davis (Julian Alston). Jason Beddow, Connie Chan-Kang and Michelle Hallaway, all from InSTePP, and Louise Letnes, from the Department of Applied Economics, University of Minnesota, also gave generously of their time and creativity in the development of this data set, as did numerous individuals at the USDA's Economic Research Service (notably Eldon Ball) and the National Agricultural Statistical Service.[1]

Version 4 consists of state-specific measures of the prices and quantities of 74 categories of outputs (including two non-commodity outputs) and 58 categories of inputs for the 48 contiguous U.S. states. The input dataset covers the period 1949–2002 while the output dataset runs from 1949 to 2006. This version of these data (see Andersen 2005) represents a revised, expanded and updated version of the dataset published by Acquaye, Alston and Pardey (2002 and 2003), which ran from 1949 to 1991. The original version of the dataset, which ran from 1949 to 1985, was first described and used by Craig and Pardey (1990) and then used (with updates and revisions) more formally by Craig and Pardey (1996a and 1996b) and Alston, Craig and Pardey (1998). Here we provide a brief overview of the InSTePP production accounts, emphasizing some of the more important data construction choices used to assemble the data. More complete descriptions of the data sources and development details can be found in Pardey et al. (2009a).

The primary motivation for constructing these production accounts was to enrich our understanding of the national and state-specific economic history of inputs, outputs, and

1 These production accounts were first developed during the period 1986–1989 and subsequently reworked and expanded through three significant rounds of revisions with the generous financial and institutional support of the Minnesota agricultural experiment station (MAES), the Dean's office of the University of Minnesota's College of Food, Agriculture and Natural Sciences (CFANS), InSTePP, the USDA Economic Research Service, the USDA National Research Initiative, the Farm Foundation, the University of California (UC) Pacific-Rim Research Program, the UC Division of Agriculture and Natural Resources, the UC Davis College of Agriculture and Environmental Sciences, and the Giannini Foundation of Agricultural Economics.

productivity in American agriculture, and to gain a better understanding of the relationship between investments in R&D and agricultural output and productivity. To that end, special attention was given to accounting for variation in the composition of input and output aggregates, with particular reference to the quality of inputs and outputs and the spatial dimension. The work of Star (1974) motivated a decision to disaggregate the data as finely as possible to serve the analytical objectives of this research. Star showed that one is safe in using pre-aggregated data only if all of the inputs (or outputs) in the class are growing at the same rate or are perfect substitutes for one another. If, for example, the rate of growth of the higher-priced inputs (outputs) exceeds the rate of growth of the lower-priced inputs (outputs), the estimated rate of growth of the group will be biased downwards when pre-aggregated data are used. Hence, growth rates of agricultural productivity will tend to be overstated if the quantities of higher-priced (i.e., higher-quality) inputs are growing relatively quickly. Using earlier versions of these data, Craig and Pardey (1996b) and Acquaye, Alston and Pardey (2002 and 2003) investigated the implications of carefully accounting for price variation across qualities or U.S. states for measures of aggregate national and regional output and input growth, and that assessment is pushed further in this book.

## Variable Description

For much of the analysis in this volume, the 58 categories of inputs are grouped into four broad categories: land, labor, capital, and materials inputs. The land input is subdivided into service flows from three basic types of land, namely pasture and rangeland, non-irrigated cropland, and irrigated cropland. This measure of the quantity of land in agriculture differs from the more-traditional measures of land in farms in that it excludes non-grazed forest and woodlands (areas which, although in farms, are not in agriculture), and includes federally owned land rented or leased for rangeland grazing purposes. Also included in the cropland measures are land acres idled for whatever reason.[2] The price weights used for aggregation of the land input are annual, state- or region-specific, cash rents for each of the three land types.

Annual data for a total of thirty-two categories of labor—comprising thirty categories of farm operators, as well as hired labor and family labor—were used to form the labor aggregate. Data on days worked off-farm by farm operators were used to adjust the measure of operator labor, and to account for the substantial but uneven shift towards part-time farming. State-specific wages were obtained for the hired and family labor, whereas implicit wages for operators were developed using national income earned by farm, ranch, and other agricultural managers categorized by age and educational attainment.

The twelve different capital inputs in this data set include seven classes of physical capital and five classes of biological capital. A physical inventory method, based on counts of assets purchased or in place was used to compile the data as described in some detail

2 At times under U.S. farm commodity programs (e.g., as they operated under the 1985 Farm Bill), farmers who opted to participate in crop programs were obliged to set aside some land as a condition for receiving farm program loans or subsidy payments. Under these policies, the decision to participate in the program involved the choice to take a support price instead of a market price and, jointly, to commit some additional land as set aside. In this way a program crop can be seen as higher-priced but effectively lower-yielding in terms of the acreage that must be committed to produce a bushel that is eligible for the high price. Having used distorted prices to value program crops, on the grounds that they are the decision prices and represent the "opportunity cost for the producer," it is logically consistent also to treat the set-aside acres as an input, because they are also part of the producer's decision calculus. Footnote 3 in this section provides additional justification for this decision.

in Andersen, Alston and Pardey (2009) and Pardey et al. (2009b). In addition, we adjusted inventories of the physical capital classes to reflect quality change over time, depending on the nature of the data available and the service-flow profile of each capital type. For all farm machinery, automobiles, and trucks, inventories on farm were converted to equivalent head counts of a new machine, using information on the average age of machines. For tractors and combines, additional information on the productive characteristics of machines on farms was used to adjust for quality as well, which led to inventories measured in equivalent numbers of new, numeraire tractors (two-wheel drive, 55 horsepower) or combines (medium capacity). Biological capital, here consisting of breeding, milking, and egg-laying livestock assets, were also included. For only one capital class, buildings, the quantities were implicit quantities derived by dividing the total rental value of buildings on farms by a price index developed for the purpose. Rents for capital items were taken to be specific fractions of the purchase price, fractions that varied among capital types. Purchase prices were assumed to reflect the expected present value of real capital services over the lifetime of the specific type of capital.

Eleven types of materials inputs are included in this data set. Apart from fertilizers, measured as quantities of elemental nitrogen, phosphorous, and potash, the purchased input quantities were implicit quantities derived from state-specific expenditure totals. The miscellaneous category was preaggregated, and included a list of disparate inputs, such as fencing, irrigation fees, hand tools, veterinary services, and insurance costs, among others. In the materials input category, state-specific prices were available only for electricity; all other input prices were national prices or price indices based on national prices paid by farmers.

In the disaggregated form, the output data cover 74 output categories, including 16 field crops, 22 fruits and nuts, 22 vegetables, implicit quantities of nursery and greenhouse products, nine livestock commodities, and four miscellaneous items including implicit quantities of machines rented out by farmers, and CRP acreage.[3] Because data were not available, horses and mules are not included in this data set either as biological capital inputs or as outputs even though they have been important, particularly in some states.

The prices used as weights to form various output aggregates are state-specific prices received by farmers for all commodities, except machines for hire and nursery and greenhouse products. For these composite commodities a price index based on national prices for their subcomponents was constructed and used to deflate reported current dollar values. A standard (but somewhat questionable) choice was to use the (policy-distorted) prices paid and received by farmers as the relevant prices for aggregating inputs and outputs (see footnotes 2 and 3 in this appendix for some related discussion). Table II-1 summarizes the input and output variables and their groupings into various categories. Table II-2 summarizes the groupings of states into the regions used throughout this book.

## Aggregating Inputs and Outputs

Index number theory provides insight into how to minimize bias from the procedure used to aggregate inputs and outputs by choosing an appropriate index procedure, carefully

3 As a departure from Craig and Pardey (1996a and 1996b) and other studies, in this version of the production accounts CRP acreage is treated as an output for which producers receive revenue equal to the CRP payments from the government. This is consistent with treating idled acres under the Acreage Reduction Program as an input.

**Table II-1 InSTePP Input and Output Categories**

| Input and Output Categories | Details |
|---|---|
| *Inputs* | |
| Land (3) | |
| Cropland | |
| Irrigated cropland | |
| Pasture and grassland | |
| Labor (32) | |
| Family labor | |
| Hired labor | |
| Operator labor (30) | Thirty classes characterized by:<br>Education: 0–7 years, 8 years, 1–3 years of high school, 4 years of high school, 1–3 years of college, 4 years or more of college<br>Age: 25–34, 35–44, 45–54, 55–64, or 65 or more years of age |
| Capital (12) | |
| Machinery (6) | Automobiles, combines, mowers and conditioners, pickers and balers, tractors, trucks |
| Biological capital (5) | Breeding cows, chickens, ewes, milking cows, sows |
| Buildings | |
| Materials (11) | Electricity, purchased feed, fuel, hired machines, pesticides, nitrogen, phosphorous, potash, repairs, seeds, miscellaneous purchases |
| *Outputs* | |
| Crops (61) | |
| Field crops (16) | Barley, corn, cotton, flax, field beans, oats, peanuts, rice, rye, sugar beets, sugarcane, sorghum, soybeans, sunflowers, tobacco, wheat |
| Fruits and Nuts (22) | Almonds, apples, apricots, avocados, blueberries, cherries, cranberries, grapefruit, grapes, lemons, nectarines, oranges, pears, peaches, pecans, pistachios, plums, prunes, raspberries, strawberries, tangerines, walnuts |
| Vegetables (22) | Asparagus, bell peppers, broccoli, carrots, cantaloupes, cauliflower, celery, cucumbers, garlic, honeydews, lettuce, onions, peas, potatoes, snap beans for processing, spinach (processed), sweet corn (fresh and for processing), sweet potatoes, tomatoes (fresh and for processing), watermelons |
| Nursery and greenhouse products | Aggregate of nursery and greenhouse products |
| Livestock (9) | Broilers, cattle, eggs, hogs, honey, milk, sheep, turkeys, and wool |
| Miscellaneous (4) | Hops, mushrooms, machines rented out, Conservation Reserve Program acreage |

*Notes:* Numbers in parentheses indicate the number of items in each category.

selecting value weights for all inputs and outputs, and disaggregating inputs and outputs as finely as possible. Throughout most of this volume, we use indexes of the quantities and prices of aggregate output and input that were formed using a Fisher discrete approximation to a Divisia index.[4] To illustrate the nature and magnitude of the biases involved in using a

4 The Fisher indexes of multi-factor productivity reported and used throughout this volume are not adjusted for procyclical changes in capital asset utilization in U.S. agriculture, although such changes have been studied in some detail by Andersen (2005) and Andersen, Alston and Pardey (2007).

**Table II-2 Regional Groupings of States**

| Region | States in region |
|---|---|
| Pacific | California, Oregon, Washington |
| Mountain | Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming |
| Northern Plains | Kansas, Nebraska, North Dakota, South Dakota |
| Southern Plains | Arkansas, Louisiana, Mississippi, Oklahoma, Texas |
| Central | Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, Wisconsin |
| Southeast | Alabama, Florida, Georgia, Kentucky, North Carolina, South Carolina, Tennessee, Virginia, West Virginia |
| Northeast | Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont |

fixed-weight index, in some instances Laspeyres indexes are also presented for comparison. Using the formation of an output quantity aggregate to illustrate the development of these indexes, let $q_{i,t}$ = quantity of output item $i$ = 1,2,…,$N$, in year $t$, $p_{i,t}$ = the price of item $i$ in year $t$, and $p_{i,0}$ = price of item $i$ in the base year = 0. A Laspeyres quantity index, using base-period prices as weights, is defined as:

$$\text{(II.1)} \qquad Q_L = \frac{\sum_{i=1}^{N} q_{i,t}\, p_{i,0}}{\sum_{i=1}^{N} q_{i,0}\, p_{i,0}}$$

and a chained Fisher quantity index is defined as:

$$\text{(II.2)} \qquad Q_F = \left( \frac{\sum_{i=1}^{N} q_{i,t}\, p_{i,t}}{\sum_{i=1}^{N} q_{i,t-1} p_{i,t}} \right)^{1/2} \left( \frac{\sum_{i=1}^{N} q_{i,t}\, p_{i,t-1}}{\sum_{i=1}^{N} q_{i,t-1} p_{i,t-1}} \right)^{1/2} .$$

Price indexes are constructed similarly by switching the roles of prices and quantities.

The detailed InSTePP database of state- and commodity-specific prices and quantities of goods also enabled an examination of the sensitivity of the aggregation procedure to two types of errors (a) using an identical average price for each distinct item included in a particular sub-aggregate (such as using an average price for all 32 types of labor included in the aggregate of labor versus a price specific to each of the 32 classes of labor), and (b) using a national average price for each item (such as using the national average price of wheat versus a state-specific price of wheat). For example, the use of national prices ignores variation in the quality and varietal composition of the wheat crop among states that is captured by using state-specific prices.

## Major Sources of Data

The price and quantity data for agricultural outputs are annual estimates from the U.S. Department of Agriculture (USDA) Economic Research Service (ERS) and National

Agricultural Statistics Service (NASS). The estimates come principally from two publications, *Agricultural Statistics* and *Statistical Bulletins*, supplemented with NASS and USDA occasional commodity reports.[5] The output price and quantity data are all state- and commodity-specific except for the machines hired out category, for which we used a national average price.

The agricultural input data come from a host of sources, most importantly various issues of the *U.S. Census of Agriculture* (since 1997 conducted under the auspices of the USDA, and the U.S. Census Bureau for earlier years). For almost all of the input data, quinquennial *Census of Agriculture* estimates were also used to benchmark each of the time series.

To compile the data on land, state-specific estimates of acres devoted to agricultural production were obtained primarily from an ERS publication titled, *Major Land Uses*.[6] Rental rates for agricultural land include estimates of average cash rent per acre for irrigated and non-irrigated cropland, and pastureland, for USDA regions and selected states. For more recent years, estimates of cropland and pastureland rental rates were drawn from the *Agricultural Land Values and Cash Rents* series included online as part of the USDA's Economics, Statistics, and Market Information System.[7] Hard copies of this and related reports were used to compile data for earlier years.

Estimates of operator labor differentiated into 30 classes consisting of five age and six education cohorts were developed using NASS data from the Census of Agriculture and the Agricultural Resource Management Survey (ARMS), as well as data on earnings by occupation and education data from the U.S. Department of Commerce, *Census of Population*. Estimates of days worked off farm taken from various issues of the *Census of Agriculture* were used to calculate the extent of part-time farming by farm operators. Data on hired and family labor were taken from various *Farm Labor* reports published for many years by the USDA Crop Reporting Board and now available online.[8]

*U.S. Census of Agriculture* reports were the primary source for the data on stocks of machinery on farms as used to estimate the series of capital service flows. For the period after 1963, the *Census of Agriculture* data were supplemented with unpublished information purchased from the Association of Equipment Manufactures (AEM) on annual, state-specific sales of tractors and combines (differentiated by machine type and size). *U.S. Census of Agriculture* reports, supplemented with information from a range of other sources (including *Agricultural Statistics*) were used to compile the data used to construct estimates of the capital service flows derived from service structures (including buildings on farm net of farm dwellings) and biological capital.

---

5 Data for the initial years were compiled by hand from the respective reports. For the more-recent years (1992 and thereafter) these publications, and associated data files, are available online at http://www.nass.usda.gov/Publications/Ag_Statistics/index.asp and http://www.nass.usda.gov/Publications/Statistical_Bulletins/index.asp.

6 The ERS land-use data are online at http://www.ers.usda.gov/Data/MajorLandUses/.

7 For the years 1997 and after, the USDA estimates of land values and cash rents for irrigated and non-irrigated cropland and pastureland are available online at: http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1446.

8 See http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1063.

Implicit quantities of most of the materials inputs were developed using state-specific expenditure data from the Economic Research Service (ERS) and a national average price.[9] Price indexes for most of the materials inputs were obtained from the Bureau of Labor Statistics and the Bureau of Economic Analysis. Notable exceptions are the fertilizer and electricity inputs. Fertilizer data consist of state-specific quantities (tons of plant nutrient content) of nitrogen, potash, and phosphate and national average prices (dollars per ton) specific to each of these components. The electricity data include state-specific prices (dollars per kilowatt) and implicit quantities developed using electricity expenditures in each state.

## Satellite InSTePP Value-of-Production Accounts

To support an assessment of the long-run change in the composition of U.S. agricultural output (Chapter 4) and the intensity of state-specific investments in agricultural R&D (Chapter 6) we developed a satellite set of accounts representing state- and commodity-specific value-of-production (VOP) estimates for the period 1924–2005. The VOP accounts span 89 agricultural commodities in 1924 growing to 128 agricultural commodities in 2005—compared with 72 commodities (excluding CRP acreage and machines hired out) in the InSTePP production accounts—as the availability of commodity data improved.[10] The VOP series ostensibly spans all agricultural outputs (including horticultural and ornamental and nursery production, but excluding aquaculture and forestry outputs); commodities not explicitly listed were included in respective "other" output sub-totals.

The VOP accounts constitute a blending of derived value-of-production estimates using the specific price and corresponding quantity data for 72 agricultural commodities from the InSTePP production accounts for the period 1949–2005 supplemented with farm cash receipt data from http://www.ers.usda.gov/data/FarmIncome/finfidmu.htm and various occasional publications for nursery and greenhouse products, including the NASS *Census of Horticulture*. For some commodities (for example, cotton, field and sweet corn, mushrooms, onions, peaches, rice, sugarbeets, tomatoes and wool) we detected a significant discrepancy between the value of production and the farm cash receipts data (representing the value of commodities produced that were sold off farm) in the period 1949–2005. In these cases, state- and commodity-specific prices-received and quantities-produced data were compiled by hand for the period extending back to 1924, mainly from USDA *Agricultural Statistics* publications. In addition, some data were missing for some commodities in some states in some years, and were estimated by linear interpolation, but these observations never exceeded 0.1 percent of the total value of production in any year.

---

9 The ERS data on purchased inputs are now available on-line at: http://www.ers.usda.gov/Data/FarmIncome/finfidmu.htm.

10 The 128 commodities in the VOP accounts for 2005 consist of 74 specialty crops (including fruits, vegetables, and a nursery and greenhouse marketing category), 40 other crops (including various field crops), and 14 livestock products. The InSTePP series treats fresh and processed tomatoes and fresh and processed sweet corn as separate items, while data limitations allowed for the inclusion of only one tomato (fresh and processed combined) category and one (fresh and processed sweet corn, combined) category in the VOP series. The additional commodities in the VOP versus the InSTePP series were mainly niche specialty crops and various turf and pasture grasses that individually constituted a small share of the national total and were typically produced in just a handful of states. The value of production totals in the InSTePP production accounts ranged between 91 and 97 percent of the 48-state value of production total in the VOP satellite accounts.