Economic forecasting in agriculture

P. Geoffrey Allen*

Department of Resource Economics, University of Massachusetts, Amherst, MA 01003, USA

Abstract

Forecasts of agricultural production and prices are intended to be useful for farmers, governments, and agribusiness industries. Because of the special position of food production in a nation's security, governments have become both principal suppliers and main users of agricultural forecasts. They need internal forecasts to execute policies that provide technical and market support for the agricultural sector. Government publications routinely provide private decision makers with commodity price and output forecasts at regional and national levels and at various horizons. Routine forecasts are not found in the agricultural economics journals that are the sources for most of this review. The review emphasizes methodological contributions and changes.

Short-term output or 'outlook' forecasting uses a unique form of leading indicator. Because the production process has long been well understood, production forecasts are based on the quantifiable features of livestock or a growing crop. Price forecasts are largely made by conventional econometric methods, with time series approaches occupying minor roles. Because of the dominance of agricultural economists, there has been an overemphasis on explanation, and little interest in the predictive power of models. In recent years, some agricultural economists have begun to compare forecasts from different methods. Findings generally conform to widely held beliefs. For short-term forecasting, combining leads to more accurate forecasts, better than those produced by vector autoregression, which surprisingly is the best single method. Also surprising is that econometric models and univariate methods both do badly compared with naive models.

Key words: Agricultural prices; Agricultural production; Forecast comparisons; Econometric forecasting; Judgmental forecasting; Meta-analysis; Sector modeling

1. Introduction

Economic forecasting in agriculture has some features in common with business forecasting and with macroeconomic forecasting. But over time, it has developed a focus of its own. Just (1993) and Just and Rausser (1993) characterize the first quarter century of agricultural economics research (from about 1925–1950) as prescriptive; recommendations were made to farmers and managers in order to increase profits. During the second quarter century, the profession shifted toward prediction, broadly defined, including use of econometric techniques for estimating elasticities and forecasting prices. The third quarter century, from 1975 onwards, has been characterized by research on policy, trade and the global economy and expansion to environmental and resource problems. Throughout

* Tel: (413) 545-5715; fax: (413) 545-5853.
the entire period, and more markedly of late, explanation of past behavior has been the dominant focus of agricultural supply modeling, which is the area to which most agricultural forecasting belongs.

Because an assured food supply is important to national security, governments have attempted to quantify agricultural production and to exert some control over it. In the beginning, simply collecting and tabulating data on the current agricultural situation was a major challenge, and agricultural statisticians played a major role in the development of statistical methods [USDA (1969)]. Data revision was frequent. Estimates of production, for example, were subject to revision after a new census had been tabulated. The large number of ‘Situation reports’ or similarly titled publications indicates the fascination of agricultural statisticians with estimating the current status of a data series.

Most agricultural forecasters were trained as either statisticians or agricultural economists. The two professions have formed what has been, at times, an uneasy alliance. Statisticians have been largely responsible for developing the approach to outlook forecasting that relies on indicator analysis. Agricultural economists have tended to emphasize ever more complicated econometric models. They have worried a great deal about providing convincing explanations of economic phenomena, with the assumption (generally untested) that this would be useful not only for decision making but also for forecasting.

1.1. A brief history

The aim of the review is to provide a summary of the main approaches used by agricultural forecasters, with an assessment of the strengths and weaknesses of each approach. The review tries to answer two opposing questions: which results from research into agricultural forecasting can be generalized to all kinds of forecasting? Which conclusions from general forecasting research apply to agriculture? The major sections discuss the methods of forecasting as they appeared chronologically. Correspondingly, the methods become increasingly complex.

Judgmental forecasts appeared first and are still significant components of short-term outlook forecasts. There is a long history of econometric analysis, starting with single equation studies. Greater computing power saw larger multiequation analyses appear. First came studies that performed partial analysis on a single commodity sector. The interrelation among certain sectors, particularly livestock and feed, was recognized early on. Early studies on the agricultural sector in aggregate contained few equations and were of simple form. Later, multiequation, multisectoral econometric studies appeared.

Although trend extrapolation methods were widely used in commodity outlook studies, agricultural applications of modern time series methods did not appear until the early 1970s. Gradually, more sophisticated efforts were made by a handful of agricultural economists, their work ranging from various forms of composite forecasts to vector autoregression (VAR) and state space models.

Because of the historical interest in decision-making by micro-economists, there has been a scattering of articles relating forecasting to the making of choices, including those concerning probabilistic forecasts, value of information and comparison of forecasting methods when used to aid a specific purchase or sales decision. Present work in agricultural forecasting reflects the culmination of two strands of research. From earliest times, statisticians have analyzed agricultural data, in part because it was available, but also because the results were of value to farmers and other business people. Second, predicting the outcomes of different policies is a major activity of many agricultural economists.

1.2. Scope of the review

Articles on the methods or results of forecasting were extracted from an exhaustive search of the main agricultural economics journals. Searches of DIALOG databases from 1969 (Agricola) or 1970 (Journal of Economic Literature) to 1992 and of Government indexes added to the list of studies. Coverage of journals and other sources is shown in the appendix. Expanded
tables that list the source studies for each table entry can be obtained from the author.

The review covers those agricultural commodities and inputs that are the subject of forecasts regularly made by government departments of agriculture and reported in their publications. These include, at national, regional and local aggregations, the production and value (equivalently area, yield and price) of crops, and livestock numbers, production and value. Also included in the review are industrial products such as vegetable oils and meals, grain by-products, and agricultural inputs (e.g. fertilizer, pesticides, but not general petroleum products). I attempted to include every study that compared forecasts of agricultural commodities or inputs done by different methods, as well as all articles that evaluated the performance of agricultural forecasters and their methods.

Studies that focus strictly on market efficiency are excluded, as are studies that use the commodity futures markets as a test of efficiency rather than in a comparison of forecasting ability. Readers interested in such issues should refer to two recent studies on livestock futures price movements around the date of release of USDA inventory reports [Colling and Irwin (1990), Schroeder et al. (1990)]. These studies also review earlier work in that area. Also excluded from this review are commodity forecasts where the focus is the industrial use of food and fibre products. Forestry, fishing and aquaculture sectors are omitted as well.

The review relies on published work. It would be a mammoth task to collect and assess a representative sample of published government forecasts and even more difficult to acquire private company forecasts and unpublished government forecasts. [Two studies that have performed this task on USDA short-term outlook forecasts are included in the comparative review: Gunnelson et al. (1972), Surls and Gajewski (1990).]

1.3. A note on evaluation

Evaluation of the different forecasting approaches is a key feature of the review. Making fair comparisons can be difficult for two reasons. First, the only true test of a forecasting model is its forecasting performance in the post-estimation period. Partly, no doubt, because of short data series, testing, if done at all, has often been by within-sample simulation. With post-sample testing, particularly of econometric models, several problems arise. Parameter updating may be done as data become ‘known’ in the post-sample period, or updating may be totally omitted. A form of forecast contamination frequently occurs when actual values of exogenous variables are used, though these would be unknown at the time of the forecast. We do not at present know how much such contamination misleads model selection and model accuracy. In real-world forecasting, unknown exogenous variables would themselves need to be forecast. And forecasts may be modified by the analyst’s judgment before release—a kind of informal combined forecast, even if not acknowledged as such.

Second, studies use various criteria for measuring how well a method makes point forecasts and turning point predictions. The criterion used by each study to make comparative rankings is noted in the detailed tables available from the author. Root mean square error (RMSE) is the most widely reported accuracy measure and was used to construct the table entries wherever possible. Aggregating rankings from studies using different criteria is cavalier, to say the least. Even worse, the careful study by Armstrong and Collopy (1992) shows that neither of the commonest criteria (RMSE and mean absolute percentage error, MAPE) is the most reliable for choosing the best method. Similar sensitivity to choice of criterion occurs when attempts are made to rank methods for their ability to forecast turning-points (discussed further in section 7.2).

2. Agriculture’s special features

The nature of agricultural production and the historical relations among the different groups of participants in agriculture make agriculture different from most economic activity. Most prod-
uct is unbranded and sold in markets where individual suppliers have no say in price determination. Both nature and government policy can have a major impact on a farmer’s production and profits. Farmers and others connected with agriculture are used to receiving technical and economic information from publicly supported institutions.

2.1. Characteristics of agricultural production

Agricultural production is unusual compared with most business activity in its strong dependence on biological processes. Farmers have minimal ability to alter the rate of development of a crop or animal. Second, for most commodities, the production cycle is measured in months or years. Other features impose dynamic structure, especially on prices: seasonal impacts on production, high cost of adjustment once production is underway and the need to carry inventory. Estimation of leading indicators therefore became a major part of short-term agricultural production forecasting, dominating any work on price forecasting. The estimation of leading indicators was a natural extension of the data gathering activity concerning current production or inventories. For example, estimation of acres planted to spring wheat is a good indication of harvested acreage. In no other sector has leading indicator analysis found such long-term and widespread use.

Agricultural production appears to meet the four conditions laid down by Armstrong (1985, p. 196) for good forecasts by econometric methods: there should be strong causal relationships, relations should be capable of being measured accurately, causal variables should change substantially and it should be possible to forecast changes in causal variables. Unfortunately, econometric methods do poorly at forecasting agricultural production and prices. The most likely reason is the great influence on production of random shocks. Relative to most manufacturing activity, agriculture is greatly influenced by unpredictable random events such as droughts, floods and attacks by pests. The consequence of these shocks on production can be assessed reasonably well after they have occurred, which is useful in making post-harvest production estimates, but not pre-harvest forecasts.

2.2. Producers of agricultural forecasts

The predominant forecaster of production, prices and trade of agricultural commodities and inputs in most countries is central government. The Economic Research Service of the United States Department of Agriculture (USDA-ERS) contains the largest agglomeration of agricultural economists and produces the greatest number of agricultural forecasts. Government commodity specialists are the main providers of outlook information in Australia, Canada and the US [Johnson et al. (1982)]. Reports on the situation and outlook for commodity and input markets at local, national and world levels are issued from one to twelve times a year depending on commodity and country. Some agencies issue regular medium-term forecasts (2–5 years ahead). For example, Agriculture Canada has issued medium-term outlook reports twice a year since 1987 [Cluff (1990)]. Long-term projections are generally issued only irregularly, and usually for groups of commodities. Although governments publish many forecasts, often as regular series, they also make many forecasts solely for internal use, for example, the USDA forecasts of the budgetary cost of the farm program.

Other public agencies, from the Food and Agricultural Organization of the United Nations to regional or provincial governments, also produce forecasts. University faculty and (in the US) extension economists prepare forecasts for general release as part of short-term outlook programs for local farmers and agribusinesses. They may also present forecasts in scholarly publications; these usually have a methodological focus.

Private companies that process or trade commodities or supply inputs produce forecasts for in-house use, typically with relatively simple models combined with judgment. They are probably closest to business forecasters in both approach and objectives. Private consultants also produce forecasts for sale, most frequently as
adjuncts to large-scale macroeconomic models. Farmers practically never produce formal forecasts, though most of them doubtlessly form a judgment about future outcomes of their business choices.

2.3. Users of agricultural forecasts

Farmers may rarely make forecasts, but they form the largest group of users. They need to make production and marketing decisions that may have financial repercussions many months in the future. Short-run commodity outlook forecasts, at least in the US, have tended to emphasize production and inventory information. Farmers have more use for price forecasts. Once committed to a product, farmers are price takers. They produce goods that are homogeneous or highly substitutable with the goods of their competitors, who may either be their neighbors or live halfway round the world. They have no concern with problems common in manufacturing, such as the amount of sales of a branded product or what quantity of a specific model to keep in inventory. But farmers, especially those in developed countries, must also be concerned with the ways in which changes in government policy will alter their business conditions.

Agricultural journalists represent a second kind of audience for commodity forecasts. They are not users in the sense of being makers of decisions based on forecast information. They provide an indirect way for readers and listeners (mainly farmers) to receive outlook forecasts.

Processors of food and fiber, and others in the marketing chain, need forecasts to aid in their purchasing and storing decisions. They too would probably like price forecasts, but would be able to make greater use of production forecasts in their decisions than would farmers. Larger businesses also supplement public forecasts with their own in-house ones.

Governments in many countries intervene in agricultural production to protect domestic agriculture and provide food security. For this they need two kinds of information. First, for legislation and, to a much lesser extent, for program implementation, governments need to know the consequences of different policy choices on different groups in society. Agricultural economists have been especially willing, over the last 30 years, to build ever larger models to provide answers to policy questions. Emphasis has been placed on comparing proposed policies via simulations, which has measurably assisted legislators. Forecasts of output and prices are conditional on the policy actually selected. To date, efforts to forecast which policy will be selected have been minimal. [See Rausser (1982), Chapter 18 for a review of the theory and empirical applications of endogenous government behavior.] Neither have government or academic economists done much to evaluate a model's ability to forecast the actual consequences of an adopted policy. Second, in monitoring the progress of farm programs designed to control supplies or support prices, governments would like to know about the effectiveness of the program and anticipated budget outlays.

3. Short-term production forecasting

Government agencies have issued short-term forecasts of prices and production for many years. In the early years, the reports contained much about current situation and little about outlook [Hudson and Furniss (1966)]. The development of methods of estimating and forecasting agricultural production in the United States forms the basis for the organization of this section. [For a politically oriented statistical history, see US Department of Agriculture (1969). For detailed technical descriptions of data gathering and analysis, see USDA (1983). For a summary of statistical methods and detailed information on timing and content, e.g. estimates, forecasts and intentions of crop and livestock reports, see USDA (1989).]

The first agricultural forecasts were estimates of crop appearance, referred to as 'condition'. Initially these were purely judgmental assessments made by crop reporters, who compared the crop's current appearance and vitality with that of a normal year. These assessments were soon used to calibrate formal production fore-
casts. The ultimate development of short-term crop forecasts was based on ‘objective yield’ estimates. Agronomic studies related observable intermediate characteristics, such as number of flowers or number of ears, to ultimate yield. Mid-season sampling of observable characteristics enabled forecasters to improve their predictions on ultimate yield. For longer horizons, a second type of judgmental forecast resulted from surveys of farmers’ intentions to plant specific crops or to breed specific animals. Formal correction for sampling bias and the relation of past intentions to past actual performance followed.

The national annual outlook conference became a feature in most developed countries. It was run by the appropriate government agency and attended by government, academic and private agricultural economists. The first such conference occurred in the US on 20–21 April 1923, and by 1929 had evolved into a standard procedure [Kunze (1990)]. It was later moved to February and is now held (more usefully for production planning) in December. In Canada, the first federal and provincial conference was held in Ottawa in February 1934. Today’s typical conference features a number of commodity-specific sessions in which government analysts review the present situation and forces of change. The analysts then present forecasts and receive feedback from the audience.

3.1. Judgmental reporting of condition

In 1862, the editor of the American Agriculturalist sought and published monthly crop ‘condition’ summaries from May to September using information submitted by farmer subscribers [Ebling (1939)]. The following year, the USDA took over the task. Its first monthly crop report stated the condition, as of May 1863, of 19 crops in 21 Northern states and the Nebraska Territory [Newell and Warrington (1962)].

In 1910, the crop reporting agency (at that time the USDA Bureau of Statistics) was issuing quantitative estimates of acreage, production and value for 13 crops, condition reports for 23 crops and pasture, and inventory estimates for five livestock species. By 1920, the number of estimates had roughly doubled [Becker and Harlan (1939)]. Judgmental estimates of the state of a growing crop, as assessed by farmer or government crop reporters, provided the USDA with its main means of crop production forecasting for nearly 100 years. Condition summaries are still important in early-season estimation of field crop yields [USDA (1983)].

3.2. Quantitative analysis of condition and use of par

The first USDA forecast, as opposed to condition report, was made in May 1912 for winter wheat and from the following month for most field crops, excluding cotton, a politically sensitive crop. Between 1912 and 1929, the reported condition of various crops was interpreted as a forecast of yield based on the ‘par’ method [Becker and Harlan (1939)]. Essentially, the monthly condition of each crop in each state was converted to ‘full’ or 100% equivalent yield, with adjustment for trend. For example, a condition value of 80% for winter wheat on 1 July, when the final yield in that state for that year was 28 bushels, gave a 100% equivalent yield of 35 bushels (28 + 80 \times 100). By taking, for example, a 5 year moving average of 1 July ‘full’ yields, the statistician obtained the 1 July ‘par’ yield. The following year’s yield forecast was simply par yield multiplied by condition. Because some crop reports covered more acreage of a given crop than others, USDA statisticians first calculated a yield forecast within a crop reporting district, then aggregated using estimated acreage in the same area as weights [USDA (1983)]. Field crop production forecasts were calculated as the estimated acres available for harvest multiplied by the yield forecast.

Although acreage weighting removed some of the biases caused by the non-probability sampling procedure, it could not accommodate biases from problems in forming judgments or from technology changes. After 1929, correlation of condition and final yield replaced the par method as the means of yield forecasting [USDA (1969)]. Later, statisticians used charts of monthly condition and final yield plotted against time.
to judgmentally adjust yield forecasts. Even later, especially where yield trends were noted, multiple regression of condition and time variables on final yield was used to make yield forecasts.

For fruit production, a modified par method was used [Becker and Harlan (1939)]. Each past year, the condition at time of harvest, summarized from crop reporters' judgments, was divided into estimates of final production. The result was the historical par or percentage of full production for that year. The regression of reported par on historical par showed that reporters tended to slightly underestimate par (leading to a consequent underestimate of yield) coupled with over-optimism in good years. Historical par values also trended upwards over time. These two patterns were charted and used first to predict the next historical par from the latest reported condition and then to predict total production in the current year. This approach was found to give forecasts as good as those relating reported condition to yield or to actual production [Palmer and Schlotzhauer (1950)].

3.3. Objective yield forecasts

In 1925, Frank Parker proposed a plan to improve cotton yield forecasts by counting the number of plants and number of bolls of cotton. Such data have been regularly collected since 1928 [Becker and Harlan (1939)]. In 1951, the USDA Crop Reporting Service made estimates of cotton production based on these data that turned out to overstate actual production by about 15%. Since the commodity market relies heavily on crop forecasts, dealers were paying farmers relatively low prices for the supposed bumper crop. Farmers were estimated to have lost about $125 million in revenues [U.S. Congress (1952)]. Presumably the purchasers of cotton gained the windfall $125 million once the forecast error was revealed by the end of the year. However, the cost of the error prompted a Congressional inquiry from which several recommendations were made, principally the establishment of a special unit within the Bureau of Agricultural Economics to examine problems with the present methods and devise improvements. Although the USDA makes cotton price forecasts for internal use, Congress still prohibits their publication. Commentators at the time [Wallace (1953)] suggested that the USDA already had the means to make improvements, since it had developed the area probability sampling method and had access to studies on the effect of weather on yield of cotton, corn and wheat.

By 1956, 'objective forecasting' of crops was advancing on several fronts. The method essentially required a detailed quantitative understanding of plant development so that observable characteristics measured earlier in the season could be related to final harvest weight by regression analysis. One problem was that the forecast was required for the entire United States on the same date each year. On 1 August, for example, plant development might be delayed compared with a normal year. Also, plants at the northern limit of a crop's production area would be less developed than plants in states to the south. Finally, to overcome the biases and uncertainties which accompany judgmental assessments, it had to be possible to count, weigh or measure the characteristics in a standard way.

As the original culprit, cotton was the first crop to be investigated. By 1 September, the final number of bolls has appeared and the boll count is a good predictor of total yield. However, for the 1 August forecast, fruits are in different stages of development and the numbers visible exceed the final number of bolls. Surveys conducted in 1954 and 1955 established the relation between counts of different kinds of fruits on 1 August and final numbers of bolls, and between fruit count per plant and average weight per fruit. In 1956, the relations were used for a state by state forecast of cotton yield in a ten state region [Hendricks and Huddleston (1957)].

Similar problems in relating observable characteristics of young corn and young soybeans to yield were reported by, respectively, Huddleston (1958) and Kelly (1957). Objective yield surveys became operational for cotton and corn
yield forecasts in 1961, for wheat in 1962 and for soybeans in 1967, and have since expanded to include potatoes, several tree nuts and citrus fruits [USDA (1983)].

3.4. Producer intentions

In 1918, the USDA sent out a questionnaire in order to find out how great an acreage of spring wheat farmers intended to plant. USDA administrators must have had second thoughts about the effort because they kept the results secret [Ebling (1939)]. In 1923, the USDA published the first report on intended acreage for nine spring-sown crops, including cotton, based on a non-probability survey of individual farmers. Farmers reported that they intended to increase cotton acreage by 12%, an underestimate of the actual increase. In a response that was to be echoed almost 30 years later, the forecast caused activity on the cotton exchanges and some reduction in price. The following year, Congress passed legislation prohibiting future intentions reports for cotton, on the grounds that such reports were more harmful than beneficial [Becker and Harlan (1939)]. The legislation was not repealed until 1958 [USDA (1969)].

One danger became apparent in focusing on a series of intentions to plant. The series was constructed by summarizing farmers’ responses to the question: ‘Compared with the acreage of (name of crop) you harvested last year, how much percentage increase or decrease in acreage do you intend to plant this year?’ While the harvested acreage can never exceed planted acreage, it can sometimes be much less, when drought or disease results in a yield too small to harvest profitably. A large percentage increase in intentions to plant in the following year might only represent an attempt to return to the normal pattern. The obvious solution of comparing planting intentions with actual planted acres had to await data on planted acreage. By 1938, the USDA had sufficient statistics on acres planted to be able to convert planting intentions reported by farmers into ‘prospective plantings’ [Becker and Harlan (1939)]. At present, acreage intentions or prospective plantings are reported for all major field crops except cotton, processing vegetables and mushrooms, with planted acreage estimates for fresh vegetables and melons [USDA (1983)].

In the US, the first pig crop report was issued in 1922, based on a survey delivered to pig farmers by rural mail carriers. Breeding intentions have since been surveyed quarterly in the major producing states, and semi-annually elsewhere. Estimates of intended breeding are supplemented by inventory surveys for all classes of stock. Semi-annual inventory surveys are the main method of forecasting cattle production.

3.5. Probability sampling

Non-probability sampling by mail is cheap, particularly when dealing with specialized types of production. Its disadvantages are the difficulty of expanding sample findings to the population and the inability to estimate sampling errors. Probability sampling requires definition of a proper random sample. Because the sample unit may be a collection of fields and not necessarily an entire farm, enumeration is sometimes by interview rather than by mail. More accuracy can be achieved with a smaller sample and standard errors can also be computed. Probability sampling for acreage estimates started in June 1961 in 15 states [Trelogan (1963)], reaching the rest of the US by 1965. The USDA has continued to refine its sampling techniques to maintain sampling accuracy while reducing cost. Multiple-frame sampling supplemented the livestock mail surveys, and probability surveys entirely replaced the non-probability mailings from about 1979. In multiple-frame sampling, all producers in a randomly selected area (the area frame) are identified as belonging or not belonging to a master list of names (the list frame). By knowing the inventories or intentions of each farmer in the area frame, the list frame can be expanded to represent the entire population [USDA (1983)].

3.6. Evaluation of short term forecasting for outlook work

Farmers and economists have criticized the timing, usefulness and accuracy of USDA outlook reports. Criticisms on timing made some 25
years ago [Bottum (1966), Daly (1966)] may no longer be valid. The annual outlook conference has been moved forward to December. Prospective plantings reports appear at the beginning of March, before many farmers have begun to plant. Estimates useful for livestock producers appear frequently, from quarterly (for hogs and pigs) or monthly (cattle on feed) to weekly (broiler hatchery). Movements in futures prices when hogs and cattle reports appear [reviewed by Schroeder et al. (1990)] suggest that market participants use the outlook information (perhaps because it provides a more accurate estimate of current situation). Once crop or animal production is underway, farmers’ responses to price signals are limited, as is the impact of their actions on forecasted prices. Some actions are possible and perhaps profitable. For example (as discussed later in section 8.2), forecasts can be used for crop storage and livestock rearing decisions.

A continuing problem is ensuring the usefulness of forecasts. Farmers want price forecasts when the planting or breeding decision is being made. Planting or breeding intentions are reported instead. And forecasts of acres or animal numbers need to be translated into total production and then into price. Historically, both the US and Australian outlook programs seemed deliberately to leave the more difficult step of price forecasting to individual farmers. For example, the agricultural outlook for 1930 stated [quoted in Kunze (1990) p. 257]:

These reports are not designed to tell individual farmers what to do, but to give them the basic facts upon which to make intelligent decisions in view of their local conditions.

Matters have improved only slightly. Today, the discretion of the commodity analyst appears to determine whether or not outlook reports contain quantitative price forecasts in the narrative. As a compromise, government agencies could issue a price forecast and then explain the logic behind it [Freebairn (1978)].

Improved accuracy through the use of better data and techniques was an early concern [Bottum (1966), Daly (1966), Freebairn (1978)]. A number of beliefs exist [Bullock et al. (1982)]: (1) production forecasts must be perfectly accurate to be of value; (2) if outlook reports were not released, prices to farmers would be higher; (3) inaccurate reports are a major cause of short-run resource misallocation. These prove to be myths. A widely accepted psychological explanation is that people explain their successes as a result of their own efforts and their failures as a result of things outside their control. In a simple two-period model, Bullock et al. (1982) show diagrammatically how perfect information (a perfect forecast) can be used to determine the inventory carryover (say, for grain) at which marginal social value (benefits to consumers) equals marginal social cost (storage charges). With no information, the carryover decision cannot be avoided, but the carryover quantity will be sub-optimal. As long as the information in a forecast causes decision makers to store a quantity closer to the optimal carryover, such information has value. Using historical information only, decision makers would on average choose to carryover the correct amount, but from year to year price would be either too high or too low, not consistently higher. The third myth is more difficult to demolish, since optimal resource allocation in a risky environment depends on the decision-maker’s risk–reward tradeoff as well as knowledge of the production response. Studies of production on individual farms have concluded that actual production is both inefficient in use of inputs and too conservative compared with decision-makers’ stated risk preferences. The studies described in section 8.2 show limited gains in profitability from using forecast information.

Using phrases such as ‘ample supplies will likely keep prices below last year’s levels’, outlook reports have provided a form of probabilistic information for a long time, especially for prices. Although there was early recognition of the need to provide probabilistic information [Bottum (1966), Timm (1966)], Nelson (1980) was the first to suggest how such an outlook program might be set up. The need is as great today. Point estimates of production and yield are almost never accompanied by confidence intervals or similar indications of reliability. Business forecasts suffer from the same problem. Dalrymple’s (1987) survey of 134 US businesses
found that only 22% of sales forecasts were ‘frequently’ or ‘usually’ accompanied by interval estimates.

The top panel of Table 1 summarizes studies that compared forecasts from mechanical methods with outlook [Baker and Paarlberg (1952), Elam and Holder (1985), Freebairn (1975), Jolly and Wong (1987)]. Sometimes the comparison is with a naive no change forecast, compared with which outlook is generally, though not universally, better. In the most comprehensive study of this kind, covering seven crops over 42 years, Gunnelson, Dobson and Pamperin (1972) found that the first USDA crop production forecast of the year was better than naive 70% of the time. Freebairn (1975), who studied annual forecasts of price and output for ten crop and livestock products over 8 years, found slightly better results for (Australian) BAE forecasts. Later in this review, a group including outlook forecasts along with other expert forecasts will be shown to be more accurate than naive and exponential smoothing methods, but generally worse than a range of methods. As Armstrong (1985, pp. 92–96) maintains, experts and commodity analysts do seem better at determining current status than at making forecasts.

Evidence ranging from capital investment, consumer durable purchases and political voting shows that intentions can act as good forecasts. Certain conditions must be met: the event is important, responses can be obtained, respondents have a plan that they can fulfil, that they report correctly and that they are unlikely to change [Armstrong (1985)]. Acreage intentions data meet these conditions all too well. When prospective plantings data are reported to the public, most farmers are unable to change their production plans, even though this was one of the purposes of reporting intentions data [USDA (1969) pp. 67–68].

Use of planting intentions data led to noticeable improvements in forecast accuracy. Foote and Weingarten (1958) made forecasts of 19 crops using production data only and compared these with forecasts based on planting intentions.

<table>
<thead>
<tr>
<th>Type</th>
<th>Number studies</th>
<th>Number series</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison with naive</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quant</td>
<td>6</td>
<td>36 cr</td>
<td>Accuracy better than naive about 70% of time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 lv</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>4</td>
<td>15 cr</td>
<td>Accuracy better than naive or trend about 70% of time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 lv</td>
<td></td>
</tr>
<tr>
<td>Quant</td>
<td>3</td>
<td>15 cr</td>
<td>Turning point better than naive about 85% of time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 lv</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>3</td>
<td>9 cr</td>
<td>Turning point better than naive, trend or random about 70% of time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 lv</td>
<td></td>
</tr>
<tr>
<td><strong>Econometric studies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quant</td>
<td>3</td>
<td>25 cr</td>
<td>Including producer intentions as explanatory variable always improved accuracy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 lv</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>1</td>
<td>1 cr</td>
<td></td>
</tr>
<tr>
<td>Quant</td>
<td>1</td>
<td>6 cr</td>
<td>Including producer intentions as explanatory variable improved turning point accuracy 85% of time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 cr</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1**

Outlook forecasts compared with others

cr = crops, lv = livestock products.

Accuracy; (root) mean squared error (R)MSE, mean absolute percentage error MAPE, or the average of Theil’s $R = \frac{(F(t) - F(t - 1))/(A(t) - F(t - 1))}{\text{average of Theil's R}}$, where $F(t)$ is current forecast, $F(t - 1)$ is previous forecast (naive forecast in this case) and $A(t)$ is actual production. A value of $R$ between 0 and 2 indicates that the current forecast is an improvement.

Turning point ratio; number of correct one step ahead increases or decreases divided by total number of changes forecast.
Generally, use of intentions data explained about 60–80% of the actual variation in production; methods that did not use intentions data explained only 20–50% of the variation. Ladd and Kongtong (1979) reached similar conclusions for within-sample predictions on six crops. Livestock intentions might appear to be more useful, since the decision to sell or retain potential breeding stock can be made at several points in the year. However, Trapp (1981) showed that intentions data forecast beef marketings more accurately than do objective indexes of growth rate, weights or inventory numbers. These studies are summarized in the lower panel of Table 1.

Errors in USDA crop forecasts decreased with each monthly revision of production and exports of wheat, coarse grains and soybeans [Surls and Gajewski (1990)] and of production of these crops plus late potatoes [Gunnelson et al. (1972)]. Lowenstein (1954) reported similar results for cotton. These results support the advice given in the forecasting literature that it is best to use the most recent data.

4. Single equation econometric forecasts

Regression analysis and deterministic trend analysis share the common origin of correlation analysis. This section describes the dominant causal modeling approach. Trend analysis was a rarely reported but probably widely used forecasting method.

4.1. Early history

Henry Moore, widely recognized as the founder of statistical economics, presented the first econometric forecast for an agricultural commodity [Moore (1917)]. His regressions of cotton yield on rainfall and temperature in selected months made better forecasts than USDA forecasts based on condition reports. Later, agricultural statisticians estimated several true single equation forecasting models [Sarle (1925), hog prices; Smith (1925), cotton acreage; Ezekiel (1927), hog prices; Hopkins (1927), cattle prices]. These models specified lagged explanatory variables whose values were known at the time of making the forecast, a feature that is missing from many later studies. Ezekiel (1927) compared short-run price forecasts (1–6 months ahead) from this ‘empirical formula’ approach with forecasts based on the survey indicators described in section 3.4. The ‘empirical formula’ approach appeared to be more accurate, based on six forecasts, though a footnote to the paper hinted that a bigger analysis might reverse the findings. Ezekiel also recognized the value of combining, noting (p. 29) “... eventually the most satisfactory results may be obtained by some combination of the... methods.” After this pioneering effort, dynamic supply response became the main line of time-related single equation work. Its history is almost entirely one of explanation and policy analysis [Nerlove (1958), chapter 3]. One of the few efforts of pure forecasting was by Cox and Luby (1956), whose specifications for 6 and 12 month ahead price forecasts for hogs also relied on explanatory variables known at the time of the forecast. They reported average errors (probably corresponding to MAPE) of 8.1% to 9.3% for 16 annual and 32 semi-annual within-sample forecasts. All but nine of the 48 forecasts indicated the correct direction of price movement.

4.2. Crop and livestock production and price

Because most crops have an annual growing season, crop response models are typically annual. The generic supply response model is

\[ Q_i^* = f(P_i^*) \]

where \( Q_i^* \) is anticipated output, \( P_i^* \) is expected price and \( i \) is the time period. Other crop and input prices (or indexes) often appear as explanatory variables. Production response is frequently disaggregated into a two-equation recursive system, first of acreage response, then yield response. In the simplest model, farmers’ price expectation is assumed to correspond to the naive no change model and \( P_i^* \) is replaced by lagged price. Use of a futures price, if one exists
for the commodity, has occasionally been tried. Slightly more sophisticated is the adaptive expectations model for price developed by Cagan and Nerlove [Nerlove (1958)]:

\[ P_t^* - P_{t-1}^* = \alpha (P_{t-1} - P_{t-1}^*) \quad 0 < \alpha < 1. \]

Equivalently, expected price is the previous expected price plus a fraction (\( \alpha \)) of the previous error in expectation. Some algebra shows that expected price is also an exponentially decaying function of past prices. The adaptive expectations model corresponds to simple exponential smoothing of observed prices.

By a suitable transformation, output is a function of lagged price and lagged output. Similarly, output can be expressed as previous output plus a partial adjustment of the difference between anticipated and previous output, the technical rigidities model. The final result is that output is a function of price lagged one period and output lagged one and two periods. [See Askari and Cummings (1977) for an extensive review.] While such equations could readily be used for one step ahead output forecasting, this was rarely done and even more rarely published. Cape1 (1968) actually produced a forecast of Canadian wheat acreage, although, since it was published before the forecast date, he could report no comparison with actual acreage.

For price forecasting, a demand equation is added. Price and quantity are usually assumed to be recursive in agriculture, though simultaneous specifications exist. L'Esperance (1964) found forecasts from a reduced form to be slightly, though not consistently, more accurate than single equation forecasts. A common alternative is to estimate a single reduced form equation for price, based on a simultaneous system. This will typically contain contemporaneous variables for income and other commodity prices. Since explanation or policy analysis was the usual purpose of any study, econometricians ignored the need to first forecast contemporaneous variables before the price equation could be used in prediction.

Livestock production has been modeled by the same partial adjustment as described for crops. Naive price expectations were used initially to explain the existence of hog and beef cycles. Since livestock production is year round, in contrast to crop production, studies soon came to use quarterly or monthly data series and different methods of describing seasonal and cyclical patterns were employed. Livestock production and more especially prices have been popular subjects for single equation econometric studies.

4.3. Evaluation

Institutional forecasters produce many forecasts, but far fewer reports detailing the methods used. The single equation econometric approach has been popular, though the published record contains insufficient information to state what proportion of forecasts are produced from this approach. Most official government forecasts are, in any case, the consensus of a committee [Newell and Warrington (1962)]. Tables 2 and 3 summarize all single equation forecasting studies located, aside from the earliest studies cited in section 4.1. Since 1964, 13 studies used single equation methods to forecast 17 quantities (acres, production, yield and export quantities) and eight prices of crops. On the livestock side, there have been 39 studies since 1952. Fifteen studies have examined production, 20 investigated price and four, both. The production studies were almost equally divided between hog and beef production, with a few on lamb, milk and wool production. If the studies summarized can be taken to be representative of the state of the art, a number of weaknesses are evident that limit their usefulness. Over half (46 of 85 specifications) require ancillary forecasts of contemporaneous independent variables to make forecasts of the dependent variable. This statistic is slightly misleading, since a proportion of these studies only require standard macroeconomic forecasts such as disposable income or a consumer price index. Almost all of the studies that test forecasting performance do so within-sample. Typically, actual values of explanatory variables are used (ex post forecasting), even though these would be unknown in a real forecasting.
context. In an article that demonstrated a largely ignored point, Fox (1953) observed that error of forecast was greater in ex ante forecasting, when explanatory variables needed to be forecast. In his example, when corn production and personal income both needed to be forecast, the standard error of forecast for 12 months ahead corn price increased by two and a half times. Though logically defensible, the finding contrasts markedly with the empirical evidence from macroeconomic forecasting [Armstrong (1985) p. 241, McNown (1986)].

Later studies have presented single equation econometric models mainly in comparison with other techniques. Since 1980, one of six studies on production forecasting and 12 of 16 studies on price forecasting were comparative, mostly with time series methods. The studies represent the state of the art in single equation econometric specification and reveal its present limitations. In short-term forecasting, as will be shown in section 7.2, it performs poorly against time series methods. Since vector autoregression is shown to be much more accurate, the most likely cause of poor performance is insufficient attention to dynamic specification.

Crop production forecasts can reasonably be based on annual series. In particular situations, for example where seasonal prices of stored crops differ from the post-harvest price by only

Table 2
Single equation econometric studies, by commodity, 1952 onwards

<table>
<thead>
<tr>
<th>Type of series</th>
<th>No. of series</th>
<th>Data frequency</th>
<th>Curr. expl. var.</th>
<th>Fcst.</th>
<th>Comp.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A S Q B M D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All crops</td>
<td>25</td>
<td>22</td>
<td>1</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Acres/yield</td>
<td>10</td>
<td>10</td>
<td>-</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>Export quant</td>
<td>7</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>Price</td>
<td>8</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>All livestock</td>
<td>60</td>
<td>8</td>
<td>3</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td>Production</td>
<td>22</td>
<td>6</td>
<td>1</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Price</td>
<td>38</td>
<td>2</td>
<td>2</td>
<td>22</td>
<td>10</td>
</tr>
</tbody>
</table>

12 crop studies, 38 livestock studies, 1 both.

1 Data frequency: A, annual; S, semi-annual; Q, quarterly; B, bimonthly; M, monthly; D, daily.

Curr. var.: number of series with contemporaneous explanatory variables. 31/35 means that four series had known explanatory variables for one step ahead that were unknown for several steps ahead. Fcst.: number of series that produce forecasts after the estimation period. Comp.: number of series that compare forecasts with other non-econometric methods.

Table 3
Single equation econometric studies, by date, 1952 onwards

<table>
<thead>
<tr>
<th>Type of series</th>
<th>No. of series</th>
<th>Data frequency</th>
<th>Curr. expl. var.</th>
<th>Fcst.</th>
<th>Comp.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A S Q B M D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950s</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>0/1</td>
</tr>
<tr>
<td>1960s</td>
<td>10</td>
<td>6</td>
<td>2</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>1970s</td>
<td>29</td>
<td>7</td>
<td>2</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>1980s</td>
<td>38</td>
<td>16</td>
<td>19</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>1990s</td>
<td>6</td>
<td>-</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

1 Different ending dates depending on source. See Appendix A for details.

1 Data frequency: A, annual; S, semi-annual; Q, quarterly; B, bimonthly; M, monthly; D, daily.

Curr. var.: number of series with contemporaneous explanatory variables. 49/53 means that four series had known explanatory variables for one step ahead that were unknown for several steps ahead. Fcst.: number of series that produce forecasts after the estimation period. Comp.: number of series that compare forecasts with other non-econometric methods.
the storage cost, annual crop price forecasts can be useful. Recent livestock studies seem to have settled on quarterly forecasts as being the most appropriate, even when monthly or more frequent data are available, especially for prices. The usefulness in decision making of such aggregate forecasts is doubtful, as has been admitted by one of the few studies to put forecasting in a decision making context [Brandt (1985)].

We lack information on the usefulness of single equation econometric forecasts at longer horizons. In a path-breaking study, Gold (1974) compared annual series for 15 agricultural commodities and 13 non-agricultural commodities or services. Each series was broken into a succession of mutually exclusive sets of length 10 years. The process was repeated for 15-year and 20-year periods. In each case, Gold made forecasts to different horizons. For 10, 15 and 20 years ahead forecasts, agricultural series of 20 years gave better results than did series of 15 or 10 years. For non-agricultural series, 10-year data sets were best and 20-year data sets worst. Agricultural series appear to accommodate fixed parameter models better than do non-agricultural series.

Conway, Hrubovcak and LeBlanc (1990) tested six kinds of stochastically varying parameter (SVP) specifications (Swamy–Tinsley, Hildreth–Houck, Kalman filter and Cooley–Prescott) against six others (autoregressive and fixed parameter models) and the naive no change model. At both 5-year and 10-year horizons, net capital investment in agriculture was most accurately forecast by SVP (with the notable exception of Cooley–Prescott), with naive no change next and fixed parameter models last. But with such limited information, nothing definitive can be claimed for the forecasting ability of varying versus fixed parameter econometric approaches.

5. Sectoral models

A sectoral model contains, at a minimum, a supply equation and a demand equation for a single commodity. Given the lag between decision making and output, particularly in crop production, it was common in the early studies to treat the equation as a recursive system and so justify the use of ordinary least squares estimation. When the commodity is storable, an inventory demand equation is required, separate from demand for consumption, at which point a market clearing identity will also be needed. Sometimes there are demands with significantly different characteristics, for example, wheat for human food or animal feed, eggs for consumption in shell or for breaking. Trade between countries or regions adds import supply and export demand equations. Alternatively, trade can be handled in a spatial equilibrium programming model. A further complication with livestock is that the inventory can be used as investment for further production or can be sold. All of these complications can usually be accommodated by a ten equation system, except where many regions are being analysed.

Sometimes, sectors are so intimately linked that a multisector model is called for. The commonest example concerns the livestock and feed grains sectors. At some point, the multisector system becomes large enough to qualify as a large scale model, as described in the next section. The problem in forecasting with a sectoral model is either that linkages with the rest of the economy are ignored, or they are incorporated through contemporaneous explanatory variables, which must themselves be forecast. Stand-alone large scale models and those linked with large scale macromodels of the economy attempt to endogenize all variables. They are equipped to forecast, but at the cost of complexity.

5.1. Econometric models

The development of sectoral models slightly preceded that of large scale econometric models. There are a large number of sectoral models in the agricultural economics literature. The greater proportion are concerned with explanation or policy analysis. As noted from Table 4, the surge of interest in forecasting occurred in the 1970s, as both Canada and the US and, to a much lesser extent, Australia struggled to bring a set of
Table 4
Distribution of forecasting studies over time

<table>
<thead>
<tr>
<th>Time interval</th>
<th>Single equation</th>
<th>Single sector</th>
<th>Multi-sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950–59</td>
<td>2 (2)</td>
<td>2 (3)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>1960–69</td>
<td>10 (12)</td>
<td>9 (15)</td>
<td>4 (17)</td>
</tr>
<tr>
<td>1970–79</td>
<td>29 (34)</td>
<td>34 (57)</td>
<td>9 (38)</td>
</tr>
<tr>
<td>1980–89</td>
<td>38 (45)</td>
<td>14 (23)</td>
<td>8 (33)</td>
</tr>
<tr>
<td>1990–91</td>
<td>6 (7)</td>
<td>1 (2)</td>
<td>2 (8)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>85 (100)</strong></td>
<td><strong>60 (100)</strong></td>
<td><strong>24 (100)</strong></td>
</tr>
</tbody>
</table>

'Different ending dates. See Appendix A for details. 'Numbers in parentheses are percentages. These are number of models. There are 51 studies using single equation methods and 53 sectoral studies.

sectoral models together into a comprehensive system [Agriculture Canada (1978), Salathe et al. (1982), Kingma et al. (1980)].

Of the 60 sector models summarized in Tables 4 and 5, the majority (43 studies) are livestock models, of which 12 concern the poultry sector, ten, the beef or beef-feed grain sector and nine, the hog sector. Practically all the models contain contemporaneous exogenous variables. A handful, usually those relying on one or two macroeconomic variables and aggregate indexes, reported the values of exogenous variables used to make ex ante forecasts. Where forecasts were made beyond the end of the estimation period they were short, typically covering four quarters. Theil’s $U$ statistic was reported irregularly (frequently the incorrect $U_1$; readers of the agricultural economics journals were unaware of the consequences of the different ways of computing $U$ until Leuthold (1975) pointed them out, though Bliemel (1973) had already done so elsewhere). Theil’s $U_2$ allows a comparison with the naive no change model, although, with the typical sample of size four, the test has low power. The most popular assessment technique was validation by dynamic simulation within the sample used for estimation. The process was started, somewhere in the estimation period, with actual values of lagged endogenous variables. The structural system was then solved for each successive time interval using calculated lagged endogenous and actual exogenous variables. As long as the dependent variables predicted in this way gave reasonable forecast accuracy and turning point statistics, the model was regarded as suitable for use in forecasting.

Table 5
Data frequency in sector and aggregate studies

<table>
<thead>
<tr>
<th>Type of model</th>
<th>No. of series</th>
<th>Data frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
</tr>
<tr>
<td><strong>Single sector models</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crop</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Livestock</td>
<td>43</td>
<td>15</td>
</tr>
<tr>
<td>Misc</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>60</strong></td>
<td><strong>25</strong></td>
</tr>
<tr>
<td><strong>Multisector or aggregate models</strong></td>
<td><strong>24</strong></td>
<td><strong>16</strong></td>
</tr>
</tbody>
</table>

Data frequency: A, annual; S, semi-annual; Q, quarterly; B, bimonthly; M, monthly. S includes models with some A equations, Q includes models with some A and S equations, M includes models with some A and Q equations (five of 60 sector models are mixed).
But such nonstochastic simulations do not adequately test dynamic specifications [Shapiro (1973)], nor do they offer much of a guide to forecasting ability.

5.2. Programming approaches

Programming models are not usually thought of as being suitable for either explanation or forecasting. They had a surge of popularity in the mid 1970s, especially in Canada [e.g. MacAulay (1978), Martin and Zwart (1975)]. Spatial equilibrium or interregional competition models have one advantage over econometric methods in multiregional analyses. They can distribute the total quantity of a commodity among regions in an internally consistent manner. Spatial equilibrium models also calculate the commodity price in each region. The market clearing identity is raised to a new position of prominence. Additional constraints prevent prices in different regions exceeding the cost of transporting the commodity between them. For accuracy, when defining transport cost, close attention needs to be paid to historical price differentials between regions where trade has occurred. The differential can consistently exceed the cost of shipping goods between the regions.

For forecasting, supply and demand functions for each region must be updated and the spatial equilibrium found. Quantities produced and consumed in each region are normally specified as functions of commodity price. The functions are estimated econometrically by separate equations. As a simple example, the supply for a region could be a function of commodity price and a time trend variable. In any period, the time variable is collapsed into the intercept term. Alternatively, the supply function can be dynamic. For example, Martin and Zwart (1975) specified supply in each region as a function of lagged price in that region. Quantity supplied in each region was fixed at the start of calculations. The calculated price in each region then updated the quantity supplied in the next time period.

5.3. Other approaches

Ashby (1964) described a balance sheet approach that could easily be worked on a spreadsheet. The approach consists of collecting forecasts for different regions and for aggregates by whatever means available. His example was a forecast of the world sugar market. In countries with good data series and existing quantitative models, these could be used. Countries with limited data would need to have their supplies forecast judgmentally. Reconciliation of direct forecasts of total world supply or consumption and the sum of the individual country forecasts would require further judgment. The balance sheet layout ensured internal consistency. Ashby's study was the only example of the method discovered.

Modeling change as a Markov process has occasionally been suggested. Its typical use is to forecast the number of businesses of different sizes. The first step is to estimate a transition probability matrix based on historical data. The matrix raised to successively higher powers provides a forecast of changes over time. Dean, Johnson and Carter (1963) provided an early example, using the census data for 1950, 1955 and 1960 to predict the size distribution of California cotton farms to 1975. Dairy industry structure, measured as the distribution of dairy farm sizes, was similarly predicted in Canada [Furniss and Gustafsson (1968)] and in Britain [Colman (1967)]. Colman compared the first post-sample prediction with actual census data, as, more recently, did Edwards, Smith and Patterson (1985). No study was found that compared the forecast error of the Markovian transition probability approach with the error of other methods. Of competing approaches, only judgmental and programming methods could work with such a short series (minimally, two periods).

Crom (1975) described a systems approach, being developed by the USDA for the beef sector, in which enterprise budgets and a detailed specification of the structure of the beef producing sector would be used to forecast industry changes. The author provided no details
on how the budgets would be updated, nor on how they would actually be used to make forecasts.

5.4. Evaluation

The forecasting performance of sectoral models remains an unknown quantity. While such models are being used by government agencies as aids to producing official forecasts, the raw forecasts are not reported. Nor is the typical life of such sectoral models known, though it is likely to be short. Models are either abandoned or revised in such a way that a long series of forecasts is unavailable. We are left to rely on the validation process, and the relation between validation results and forecast performance is unknown.

The institutionalization of formal quantitative models appears to have been a struggle. They were developed in departments or by teams separate from those responsible for the outlook reports. Early models apparently were updated once or twice, then dropped. The researchers who developed the models were not responsible for their maintenance and regular use. The expense and effort required to develop a forecasting model, to maintain and update it and to train a new group of people in its use were typically underestimated [Hedley and Huff (1985)]. The formal modelers and the commodity analysts produced different forecasts. Since only one official forecast is released, the different values produced within an agency must be reconciled. Cluff (1990) described the process at Agriculture Canada.

Most comparisons of sectoral models, including those few where forecasts are made, concern the relative performance of different econometric estimators. Ordinary least squares (OLS) is often as good as methods developed to deal with simultaneous equations bias, but there are too few studies to draw any conclusions. Soliman (1971) found that three stage least squares produced the best forecasts (measured by Theil’s $U$) in two equations and OLS in the other two. In year by year comparisons, two stage least squares was best in 2 years and OLS in 1. Using simulated data with known autocorrelated structure, Naik and Dixon (1986) found that OLS was most accurate within-sample but two-stage least squares and reduced form with autocorrelation corrected (by Durbin’s method) were better when forecasting.

Using the broadest definition of sectoral model, six studies have been located that compare econometric sectoral model forecasts with other methods [Leuthold and Hartman (1981), Kulshreshtha et al. (1982), Park et al. (1989), Chen and Bessler (1990), Vere and Griffith (1990), Fanchon and Wendell (1992)]. These constitute about half of the methods labelled ‘other multivariate’ in section 7.2 where comparisons are discussed in more detail. Most of the systems contain from two to four equations; the largest is the 67 equation cotton sector model of Chen and Bessler. Structural sectoral models were more accurate in only ten of 38 pairwise comparisons with other forecasting approaches. (If the models had been as accurate as the other member of the pair, the probability of 10 or less successes is $P = 0.0025$.) For one step ahead comparisons (which are the vast majority), this result is unsurprising. What is surprising is that sectoral models have not been compared at more distant horizons, where the general belief is that they would do better.

6. Aggregate and large scale econometric models

Penson and Hughes (1979) and Freebairn, Raussur and de Gorter (1982) distinguish three classes or generations of agricultural industry models. First generation models treat agriculture as a separate entity and often fail to link factor demands with output. Examples include Egbert (1969), Quance and Tweeten (1972) and Yeh (1976). These were small, highly aggregated, stand-alone models. They used estimates of elasticities and rates of growth and inflation from other studies. Their purpose was projection of agricultural output and prices under different
policy proposals. Forecasting was incidental; one of the potential policies might have been identified as ‘most likely’.

Within the first generation, a second group of studies could be characterized as large scale multisector models, essentially larger versions of the sector models described in section 5 [including Maki (1963) and Crom and Maki (1965a)]. An exception is Cromarty (1959), generally acknowledged to be the first large scale econometric model of agriculture. It had disposable personal income and the US general price level as exogenous variables. These variables were, however, endogenous in the Klein–Goldberger model of the US economy. Cromarty’s model could be solved after Klein–Goldberger to produce forecasts of output and prices for 12 agricultural commodities or commodity groups.

In second generation models, the macromodel is first used to forecast a set of variables exogenous to the agricultural sector, such as personal disposable income, interest rates and the consumer price index. These forecasted variables are used to solve the agricultural system. Solution values, such as total agricultural output, are then transmitted back to the macromodel. The linkage is incomplete. Typically, variables such as capital accumulation are not transmitted back. The models often fail to include explicit variables to represent sector policies, such as acreage diversions or deficiency payments. Examples include Chen (1977) and Roop and Zeitner (1977). Penson and Hughes (1979) describe third generation models as having direct or indirect accounting of capital accumulation and financing, while Freebairn, Rausser and de Gorter (1982) simply characterize them as having better linkages between the domestic macroeconomy and the international economy or the agricultural sector.

Large scale models with from 30 to several hundred variables are intended to describe multiple sectors of the economy. They may be informal in the sense that several models each of a single sector or interrelated sectors (such as feed-livestock) are examined in concert and re-estimated, if necessary to remove inconsistencies. In this situation, a large scale representation is built from the bottom up (see section 6.1). More commonly, formal models are designed from the top down (section 6.2). The models of the macroeconomy constructed by the principal business forecasting units such as Chase Econometrics, Data Resources Inc. and the Wharton Economic Forecasting Unit are the most familiar. Kost (1981) briefly surveyed these and the individual country models of project LINK. The agricultural components of these models are usually small to non-existent.

Tables 4, 5 and 6 summarize the features of multisector and large scale agricultural forecasting models. The tables also show comparisons

<table>
<thead>
<tr>
<th>Type of model</th>
<th>Macro + micro. exog. vars.</th>
<th>Macro exog. vars. only</th>
<th>No current exog. vars.</th>
<th>No exog. vars. linked</th>
<th>Validate by dynamic simulation?</th>
<th>Make post-sample forecasts?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>43</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>19</td>
<td>38b</td>
</tr>
<tr>
<td>Multi or Aggr.</td>
<td>17</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>11</td>
<td>11d</td>
</tr>
</tbody>
</table>

Tests? refers to summary statistics of forecast accuracy and turning point performance in the post-sample period.

Tests? refers to summary statistics of forecast accuracy and turning point performance in the post-sample period.

Three studies provide insufficient evidence on whether or not validation was done.

Only one after 1979.

Two studies provide insufficient evidence on whether or not validation was done. Macroeconomic exogenous variables are those typically available from forecasting services, e.g., personal income, population and consumer price index. Microeconomic variables would need to be forecast specifically for the study. A model designed to be linked to a system may have exogenous variables that are endogenous to the system.
with models described earlier. Publication of both sectoral and multisectoral models peaked in the 1970s. The more recent surge in single equation models is accounted for by their use in forecasting comparisons, a use to which the larger models have rarely been put. In terms of data frequency, sectoral models substantially, and multisectoral models excessively, rely on annual data—too long an interval for most private decision making purposes, though useful for policy analysis. For forecasting purposes, both types of models typically require forecasts of exogenous variables. And their creators seem reluctant to test their model's forecasting performance.

6.1. Informally linked commodity sector models

The philosophies of the agencies responsible for agricultural forecasting in Australia, Canada and the US appear remarkably similar. The development of formal quantitative models also follows similar paths and timing. From the early 1970s, agencies followed a bottom-up approach, building gradually more complex econometric or programming models, from single to multiple sectors. Development of separate commodity models was piecemeal, often overlapping. Perhaps the most comprehensive description of the process is a two volume report that appeared in 1978 [Agriculture Canada (1978)]. It describes ten of the 13 structural commodity models under development by, or on behalf of, Agriculture Canada.

The ultimate goal of the bottom-up approach is either a single forecasting model for all commodities or a set of consistent sectoral models with formal linkages. In the 1970s, the Economic Research Service of the USDA began a two-step process of developing individual sector models which were then linked together. Attempts to link sectors frequently encountered the problem of variable incompatibility. Individual researchers had failed to consult on variable definitions and data sources, so the individual models had to be redefined and reestimated. The persistence of the problem led to the construction in the late 1970s of a common database, T-DAM, or the time series data access method [Bell et al. (1978)]. At the same time, the annual linked crop-livestock model known as the cross commodity forecasting system (CCFS) was made operational. Initially, it consisted of 133 equations for nine livestock and crop sectors, with other sectors still to be added [Boutwell et al. (1976)]. Several of the sectors were being operated separately rather than being linked in a consistent manner. In 1980–1981, the CCFS was updated, respecified, enlarged and given a policy analysis orientation. Equations were added for cotton, several milk products, price indexes and government outlays. The new model, now with 360 equations, was named FAPSIM, the Food and Agricultural Policy Simulator [Salathe et al. (1982)].

In long-term projections and policy analysis, the USDA uses results from a set of econometric models. The mechanical projections from the models are moderated by judgments from a committee of analysts (Paul Westcott, personal communication, 1993).

6.2. Formally (comprehensively) linked models

6.2.1. Multisector models

Two groups of researchers began work on the livestock-feed sectors in 1965 [Crom and Maki (1965a), Egbert and Reutlinger (1965)]. They realized that there were many linkages among the prices and quantities consumed of the different meats. And, since about 70% of corn, barley-grain sorghum and oats produced in the US went for animal feed, demand for these commodities was a derived demand from the livestock producing sectors. Maki's (1963) 44 equation model and the Crom and Maki (1965a) 30 equation model of the beef-pork industries were the first studies on interrelated sectors. They were unusual in being semi-annual models, while most of the succeeding models were annual (see Table 5). The latter model was simulated within-sample to derive operating rules, such as the changing of a coefficient or an equation specification when a particular variable reaches an extreme value [Crom and Maki (1965b)]. Crom (1970) listed the 128 operating rules de-
veloped from the simulations. He also updated the model and showed graphically how the operating rules improved within-sample performance. The improved system was used for forecasting, but no tests of forecasting performance were made.

The USDA has used a variety of large scale model systems intended to simulate various policy options, with detailed regional and commodity breakdowns. As well as FAPSIM, described earlier, systems include the accounting-type POLYSIM [Ray and Moriak (1976)], the econometrically based TECHSIM [Collins and Taylor (1983)] and its successor AGSIM [Taylor (1990)]. Another example is the Food and Agricultural Policy Institute’s (FAPRI) policy-oriented econometric model [Brandt et al. (1991)]. These are annual models with limited application to short-term forecasting.

6.2.2. Linked agricultural–macro models

The model of Cromarty (1959), mentioned earlier, is generally reckoned to be the first large scale agriculturally oriented linked model. Chen (1977), Roop and Zeitner (1977) and Chan (1981) demonstrate second generation models where the agricultural sectors are essentially add-on components to large macromodels and are solved sequentially. Because the agricultural sector is a small part of the total economy, failure to feedback sector solutions to the main model has little impact. A few third generation models have appeared [e.g. Penson et al. (1984)]. The Freebairn, Rausser and de Gorter (1982) model, in addition to linkages with the macroeconomy and international markets, also contains reaction functions that endogenize policy.

The ORANI model of the Australian economy [Dixon et al. (1982)] is a structurally detailed computable general equilibrium model that is unusual in its detailed treatment of agriculture. There are four geographically distinct multiproduct agricultural industries and four type-of-farming industries for a total of ten commodities. The entire model has 103 other sectors, plus non-competing imports for a grand total of 114 sectors [Higgs (1986)]. Imports and exports are determined endogenously and variables can be reclassified between endogenous and exogenous. The model has been used for policy analysis, although forecasting is possible.

6.3. Evaluation of large scale models

In 1961, in reference to both sectoral and large scale models, Cromarty could observe (1961, p. 365), “We are in the infancy stage of estimating the economic interrelationships among agricultural commodities”. With increasing computing power and longer post-war data series, agricultural economists seized on the excitement of more comprehensive modeling of agriculture. Although the early developers of large scale macroeconomic models regarded forecasting as a major objective, much of the work with agricultural models was concerned with structural specification, estimation techniques and policy simulation. Only occasionally did forecasting appear to be a goal. Even less often was anything more than rudimentary testing of forecasting performance considered. Questions of usefulness and comparative performance of large scale models were rarely addressed, at least in print.

Improved specification and model testing are closely linked. Sometimes formal testing is not required. Around 1972–1973, agricultural prices in the US rose dramatically, surprising everyone and dismaying forecasters. Changes in US farm policy, combined with rapid inflation, small crops and rising export demand, converted the US farm economy from a relatively closed system to a relatively open one. Forecasters concluded that their models were inadequate because they failed to consider the impact of international markets on US prices [Rausser (1982)]. Rapid inflation, at different levels in different nations, meant that price forecasts in nominal terms were impossible.

The price shocks of 1972–1973 marked a watershed in the development of international commodity models. Before that time, researchers in the US had made little progress in improving the forecasting performance of international...
commodity models [Labys (1975)]. Between 1976 and 1980, a series of conferences was held among representatives from the USDA, Agriculture Canada, commercial forecasters and academics [Rausser (1982)]. Greater dependence on world markets had forced Canadian researchers to address the issue of international impacts. Competition forced commercial forecasters to improve the performance of their product. The USDA lagged behind.

Without the feedback from striking events, agricultural forecasters continue to suffer from two major handicaps: lack of a common maintained model and inadequate test protocols. These are among the criticisms that Fildes (1985) levelled at quantitative forecasters in general. While economic theory is some guide to specification, it is silent on many practical issues, especially dynamics. For example, supply of a commodity is a function of commodity price and input prices. But what about competing products? What about constraints on rate of adjustment to price changes? There is no reference specification, for example for corn supply, against which possible improvements can be tested. Because there is no accepted cumulation of past research, it is impossible to take all previous work into account.

Soon after sectoral and large scale models became popular, various commentators began to assess the progress that had been made. They found much to be dissatisfied with. Shapiro (1973, p. 255) noted that

almost all model evaluation procedures to date have employed non-stochastic simulation (with respect to both equation error term and the sampling distribution of the estimated parameters) to generate the experimental parameters — a procedure which is inadequate in testing dynamic theories.

The comment is still true. Stochastic simulation would not address the predictive performance where it depends on contemporaneous exogenous variables.

While the last two decades have seen many new model test procedures appear, actual use of the methods does not appear to have greatly improved. In the context of agricultural trade modeling, Thompson and Abbott (1982) noted that, of the few modeling exercises that listed forecasting as an objective, almost none provided any forecasting measures outside the range of the data used to estimate the model. Improvement is still needed, both in agricultural sector models and in large scale models in general. Fildes (1985) recommended that models be tested for ex ante performance relative to extrapolative or judgmental alternatives, a view held by leading forecasters since the late 1970s. Just (1993) in his ‘conclusions and a call for action’ asks for a reduction in the emphasis on standard statistical concepts of fit. He notes (p. 37) “[t]he crucial criterion for forward-looking analysis is the ability to represent out-of-sample phenomena”.

Cromarty and Myers (1975), speaking from the viewpoint of industry, questioned the usefulness of annual price forecasts. As Table 5 shows, about two thirds of aggregate and large scale models are annual. They also found the complex simultaneous equation models of little use in short term decision making. In most cases, they found that an endogenous variable could be better determined within a sub-system rather than simultaneously within a system. Table 6 shows that most large scale models require forecasts of exogenous variables, although in this respect they are relatively better than sectoral models.

Outlook forecasting at the USDA is a complex and diffuse process without an overriding formal approach [see Bell et al. (1978)]. Single and multiple equation econometric models, statistical analyses of survey data and analyst judgment are combined to produce the official forecasts. For longer horizons, forecasters place relatively more emphasis on econometric models (Paul Westcott, personal communication, 1993). A series of US General Accounting Office reports has recommended that the USDA document its forecasting procedures [US General Accounting Office (1988) p. 76, (1991) p. 58]. The recent documentation of farm income forecasting is a first step [Dubman et al. (1993)].

On the other hand, Agriculture Canada claims that its Food and Agriculture Resource Model
(FARM) is “one of the few being operated on a continuing basis within a successful outlook program” [Hedley and Huff (1985)]. It is updated each quarter and used to produce forecasts for up to 6 years ahead. After the results have been reviewed, the model is calibrated, if necessary, by reverse simulation [Cluff (1990)]. This process maintains internal consistency while achieving forecasts that are judged acceptable. Australia has had an aggregate recursive program available since 1976 and the model has been used for 5 year projections [Kingma et al. (1980)]. Few comparative studies of the forecasting performance of large scale models have been located. Just and Rausser (1981) found that futures prices were more accurate predictors of eight agricultural commodity prices one to three quarters ahead than were four major private forecasters and the USDA (in that order). In three studies [Roop and Zeitner (1977), Stillman (1985), Westcott and Hull (1985)], Theil’s $U_2$ statistic was greater than one in more than half of the variables forecast in the post-estimation period (79 of 161). Results within-sample were much better. There is every reason to believe that these findings are typical of large scale models.

Do large scale models have a use? Their makers would argue that they are intended for policy simulations, not for forecasting. But can their conclusions be trusted? Models would have greater credibility if, once a policy was enacted, the forecasts produced by analysts were conditioned on the given policy variables.

7. Time series models

Deterministic trend extrapolation was an early form of time series analysis, probably widely used in institutional forecasting though little reported in the literature. In the 1960s, interest in the hog cycle led to the use of harmonic analysis, in which production and price of hogs were modeled as cosine functions [Larson (1964)].

7.1. Use of time series methods in agricultural forecasting

Jarrett’s (1965) forecast of Australian wool prices using exponential smoothing marked the first application of modern time series methods to agriculture. For agricultural economists in the US, the era of time series analysis began in 1970 with the appearance of an article illustrating the Box–Jenkins and exponential smoothing methods [Schmitz and Watts (1970)]. Although intended as a demonstration, by reporting proper post-sample forecasts the article set a standard that was not followed for many years. Exponential smoothing produced the more accurate forecasts. In contrast to business forecasting practice, exponential smoothing has practically never since been used for agricultural forecasting.

Articles on spectral analysis began to appear at the same time [Rausser and Cargill (1970), broiler cycles, US; Weiss (1970), world cocoa prices; Cargill and Rausser (1972), futures prices; Hinchy (1978), lead–lag relation between export and Australian saleyard prices of beef]. The intent was to explain historical data patterns rather than to forecast. A demonstration of the use of multivariate cross spectral analysis followed [Ahlund et al. (1977), beef price]. In the 1980s, interest in multivariate time series analysis became evident. Shonkwiler and Spreen (1982) used a transfer function to analyse the much studied relation between the number of hogs slaughtered in the US and the hog–corn price ratio. Again, their interest was to confirm what had already been shown by spectral analysis, the existence of a cycle of about 3.4 years.

At about this time, Bessler introduced vector autoregression (VAR) to the agricultural economics profession. His 1984 article [Bessler (1984a)] provided a good explanation of the basic approach without dwelling on the over-parameterization problem. A series of articles introduced various parameter reduction methods [Brandt and Bessler (1984), Tiao’s exclusion of variables method; Bessler and Hopkins (1986), symmetric and non-symmetric random walk priors; Bessler and Kling (1986), general
Bayesian priors. Bessler also pioneered the commendable practice of providing comparisons among methods based on post-sample forecasts. Kaylen (1988) provided a review of parameter reduction methods, including both exclusion of variables and Bayesian techniques. He also introduced a new exclusion of variables approach that is similar to Hsiao's method but allows for deletion of insignificant intermediate lags.

Turning point measures have fascinated agricultural economists, perhaps because microeconomic theory emphasizes direction of effect over quantity of effect. Since the standard $2 \times 2$ contingency table only distinguishes a change in direction from no change, a forecast of a peak (a rise followed by a fall in the series) will be counted as correct when the actual series displays a trough (a fall followed by a rise). Naik and Leuthold (1986) overcame the problem with a $4 \times 4$ contingency table that distinguished between peak and trough turning points. In a one step ahead forecast, the distinction is unnecessary since actual data will reveal whether the series is rising (with a potential peak about to occur) or falling, though a forecaster might be concerned about the error rate of predicting peaks compared with that of predicting troughs. Only for several steps ahead forecasts is the larger table really necessary [Kaylen and Brandt (1988)].

7.2. Evaluation: comparisons

7.2.1. Accuracy

After the pioneering study of Leuthold et al. (1970) no comparisons of agricultural forecasts

<table>
<thead>
<tr>
<th>Table 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success rates of different methods in pairwise comparisons of ex ante forecasting performance</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>OU</td>
</tr>
<tr>
<td>Ex</td>
</tr>
<tr>
<td>Ec</td>
</tr>
<tr>
<td>VA</td>
</tr>
<tr>
<td>OM</td>
</tr>
<tr>
<td>CS</td>
</tr>
<tr>
<td>CO</td>
</tr>
<tr>
<td>Success rate</td>
</tr>
</tbody>
</table>

The upper triangular matrix is a better than/worse than set of pairwise comparisons. For example, the first entry is interpreted as 'the row entry method (Naive) is better than the column entry method (ARIMA) in 15 pairwise comparisons and worse than ARIMA 23 times' (in a total of 38 comparisons from various studies). Better usually means lower one step ahead forecast root mean square error. Sometimes results are only presented with RMSE over a range of forecast horizons. Sometimes only MAPE is reported.

The lower triangular matrix is the success rate, or proportion of total pairwise comparison in which the column entry method is better than the row entry method. The value is found by dividing the number of times with better than comparisons by the total number of comparisons for each pair of methods. A dash indicates that the methods have never been compared. The interpretation is inverted. For example, the first entry, 0.39, which is 15 divided by 38, is the success rate of the column entry method (Naive) over the row entry method (ARIMA). The bottom row is the overall success rate for the column entry in all comparisons (for example, 0.40 for the Naive method).

The better than/worse than comparison is done as follows. Each member of a group is compared with each method not in the group but not with other members of the group. For example, if a study ranks the methods as follows: (1) composite, (2) econometric, (3) ARIMA, (4) different econometric, then the pairwise orderings considered are (1,2), (1,3), (1,4), (3,1), (3,2), (3,4) and (2,1), (2,3), (4,1), (4,3). In terms of ($x, y$) pairing, these are recorded as composite (3,0), ARIMA (1,2), econometric (1,3).

The example is illustrated in detail in Table 7a. Table 7b collects Table 7a information into the form used in Table 7b.
appeared until almost the 1980s. Table 7 summarizes the comparative forecasting accuracy of nine methods or groups of methods. The comparisons are based on forecasts of 129 agricultural series reported in 49 studies. Because studies typically compare only two or three methods, Table 7 summarizes pairwise comparisons of the different methods. The upper triangle shows \((x,y)\) pairs: the method listed on the left of the row was more accurate than the method listed in the column heading in \(x\) comparisons and worse in \(y\) others. Symmetric off-diagonal elements of the matrix contain the same information in reverse order, so the lower triangle would convey no new information. Instead, the lower triangle has been used to report the success rate between pairs of methods. The value in row \(i\), column \(j\) is the proportion of times the method in column \(j\) is better than the method in row \(i\) in terms of the criterion used in the study. For example, ARIMA beats other univariate 26 times in 34 pairwise comparisons, achieving a success rate of 0.76. This information can also be expressed in reverse order. Since ‘other univariate’ beats ARIMA eight times in 34 comparisons, its success rate is 0.24 or \((1 - 0.76)\). The final row of Table 7 is the overall success rate of forecast accuracy of the method listed at the column heading against all others. For example, the naive method is more accurate in 104 out of a total of 259 pairwise comparisons, a success rate of 104/259 or 0.40.

The (root) mean square error ((R)MSE) was used to rank methods, when the study reported it. In the 5% of studies that failed to present the

\[
\begin{array}{cccc}
\text{Pairing} & \text{Composite} & \text{ARIMA} & \text{Econometric} \\
1.2 & 2.1 & 1.0 & 0.1 \\
1.3 & 3.1 & 1.0 & 0.1 \\
1.4 & 4.1 & 1.0 & \\
2.3 & 3.2 & 0.1 & 1.0 \\
3.4 & 4.3 & 1.0 & 0.1 \\
B/W & 3.0 & 1.2 & 1.3 \\
\end{array}
\]

MSE, Theil’s \(U\) statistic or the mean absolute percentage error (MAPE) was used to make the pairwise comparisons. Choice of criterion matters, and RMSE, while the most popular, is not the most consistent criterion for choosing the best method [Armstrong and Collopy (1992)]. Comparisons between pairs of methods need to be made with caution, since the number of observations is frequently small. Very approximately, the 90% confidence interval for ten comparisons is 0.3, for 20 comparisons 0.2, and for 40 comparisons, 0.125.

Even after grouping methods together, there are many gaps. A notable omission is a comparison of vector autoregression with any form of composite forecast. The most accurate method was other composite (with weights calculated adaptively or from ratios of variances or by regression). Next was simple composite (with weights calculated as simple averages). The success of composite or combining methods is a result that conforms with widely held beliefs. More surprising is that complex methods of calculating weights performed relatively better than simple averaging (based on 11 series in seven studies). The previous statement needs some qualification. The typical study compared three to five composite methods, one of which was simple averaging. Simple averaging was typically better than some of the composites, though rarely the best. But the studies provide insufficient cumulation of evidence to favor any particular composite method over simple averaging.

VAR was the best single method, perhaps because it faced relatively weak opposition most
of the time. But 55 of the 125 pairwise VAR comparisons were with ARIMA models, a competition that ARIMA won 64% of the time. This result would be unsurprising if VAR was regarded as a causal method. But one of the criticisms levelled against it by econometricians is that VAR is atheoretical. Econometric (single equation) and other multivariate methods (transfer function, state space or structural equation systems) do slightly worse than the naive no change forecast, with other univariate methods (trend extrapolation and exponential smoothing) coming off worst. The poor showing of econometric methods is in contrast to the finding in Fildes' (1985) survey that compared causal (econometric and transfer function) and extrapolative (ARIMA and exponential smoothing) methods. There, the success rate of causal methods in short-term ex ante forecasting was 0.67, based on ten studies performed between 1974 and 1984.

A number of criticisms can be levelled at the results. A study that compares a large number of methods tends to dominate the ratios, because it permits more pairwise comparisons. The methods may represent minor variations, for example, different combinations of methods in a composite forecast. On the other hand, the single study should lead to the same conclusions as a number of separate studies that use the same series. The same hog price series (over almost the same time interval) was used by about ten of the studies.

The strength of the competition obviously matters. For example, the state space approach in the competition organized by McIntosh and Dorfman (1990) performed rather badly relative to two strong methods: a VAR approach and an alternating conditional expectations approach estimated by Berck and Chalfant (1990). (Alternating conditional expectations is a form of regression analysis that uses a smoothing algorithm to find the best transformation of each variable.) How would the methods have fared if the competition had been equal? That is, if each comparison had consisted of the same total number of pairs? The answer is found by calculating the simple average of all the success rates for a given method. Rates for some pairs are unknown or are based on very few comparisons. The success rate of 0.50 (equal ability) was arbitrarily assigned where less than ten comparisons existed between a pair of methods. The ordering of methods is essentially unchanged. Only other multivariate methods perform worse than naive methods, but only VAR and ARIMA are markedly better.

7.2.2. Turning points
Forecast accuracy was the commonest criterion used in comparing methods. A smaller number of studies reported turning point performance. Table 8 summarizes comparisons from 41 series reported by 13 studies [Freebairn (1975) contains almost half the series but makes only one pairwise comparison of each]. Conclusions are similar to the forecast accuracy comparisons. Again, VAR and the composite methods perform best. Few comparisons with the naive method were located.

Turning point comparisons need to be interpreted cautiously for two reasons. First, the even smaller number of observations than for the forecast accuracy comparisons makes their ability to rank methods even less powerful. Second, the choice of criterion is important here also. The commonest criterion is number of directional changes correctly forecast. For horizons more than one step ahead, the starting direction must be forecast and the forecast may turn out to be incorrect. If so, the criterion counts as correct a forecast directional change the reverse from the actual change, for example, a forecast peak (a move upwards then downwards) where an actual trough occurs.

Three authors have examined an identical data set using the same seven forecasting methods (Table 9). For one step ahead forecasts the key difference between Brandt and Bessler (1981) and the improved method proposed by Kaylen (1986) is that Kaylen compared forecast change with preceding actual change, whereas Brandt and Bessler used the forecast change. At the time of forecast, the preceding actual change is known, so that certain forecast possibilities are eliminated. A previous price increase, for exam-
Table 8
Comparison of ex ante forecasting performance: turning points

<table>
<thead>
<tr>
<th></th>
<th>Naive</th>
<th>ARIMA</th>
<th>Other univar.</th>
<th>Expert</th>
<th>Econ.</th>
<th>VAR</th>
<th>Other multi</th>
<th>Composite</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naive</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>ARIMA</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other Univar.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>0.00</td>
<td>0.40</td>
<td>0.33</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Econ.</td>
<td>0.00</td>
<td>0.53</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.5</td>
<td>3.5</td>
<td>1</td>
</tr>
<tr>
<td>VAR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other multi</td>
<td>0.00</td>
<td>0.69</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>1</td>
<td>2b</td>
</tr>
<tr>
<td>Comp. simple</td>
<td>-</td>
<td>0.38</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.40</td>
<td>0.00</td>
<td>7</td>
</tr>
<tr>
<td>Comp. other</td>
<td>-</td>
<td>0.42</td>
<td></td>
<td>0.55</td>
<td>0.04</td>
<td>0</td>
<td>0.00</td>
<td>0.35</td>
<td>42</td>
</tr>
<tr>
<td>Overall success rate</td>
<td>0</td>
<td>0.46</td>
<td>0.33</td>
<td>0.58</td>
<td>0.38</td>
<td>0.90</td>
<td>0.13</td>
<td>0.38</td>
<td>0.69</td>
</tr>
</tbody>
</table>

See the footnotes to Table 7. The upper triangular matrix is the number of better than/worse than comparisons. The lower triangular matrix is the success rate, read in the same way as in Table 7. Because turning point measures for different methods often score the same value, those methods tie in the rankings. A tie between two methods receives a better than/worse than score of 0.5. Various turning point measures were used in the different studies, the commonest being the ratio of turning point errors to total number of changes.
Table 9
Rankings of different one step ahead turning point measures on the same methods and data series

<table>
<thead>
<tr>
<th>Author</th>
<th>Error measure</th>
<th>(1) Econometric</th>
<th>(2) ARIMA</th>
<th>(3) Expert</th>
<th>Composite</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1) Simple</td>
<td>(2) Simple</td>
<td>(3) Min. var.</td>
<td>Simple (1) (2) (3)</td>
</tr>
<tr>
<td>Brandt, Bessler</td>
<td>Correct/total</td>
<td>7</td>
<td>1</td>
<td>3=</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Separate peak, trough</td>
<td>4=</td>
<td>7</td>
<td>4=</td>
<td>1=</td>
</tr>
<tr>
<td>Kaylen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McIntosh, Dorfman</td>
<td>Ratio accurate forecasts</td>
<td>4=</td>
<td>7</td>
<td>4=</td>
<td>1=</td>
</tr>
<tr>
<td>Henriksson–Merton</td>
<td></td>
<td>6</td>
<td>7</td>
<td>2</td>
<td>3=</td>
</tr>
<tr>
<td></td>
<td>Number correct up</td>
<td>1=</td>
<td>6=</td>
<td>6=</td>
<td>1=</td>
</tr>
<tr>
<td></td>
<td>Number correct down</td>
<td>7</td>
<td>3=</td>
<td>1=</td>
<td>3=</td>
</tr>
</tbody>
</table>

All studies use the data series for hog prices (price of barrows and gilts at seven terminal markets) and forecasts from Brandt and Bessler (1981). An almost identical data set and results are also found in Bessler and Brandt (1981). The original article contains an error (noted by Naik and Leuthold, 1986) that when removed increases by one the number of correct forecasts for the simple composite of two methods. This adjustment is reflected in the rankings.

In Brandt and Bessler, a turning point is forecast when \((F_{t,t-1} - F_{t-1,t}) \times (F_{t+1,t} - F_{t,t-1}) < 0\) and the forecast is correct when \((F_{t} - F_{t-1}) \times (F_{t+1,t} - F_{t,t-1}) < 0\) where \(F_{t,t-1}\) is the (one step ahead) forecast for time \(t\) made at time \(t - 1\). Since the first part of the expression is known at the time of the forecast, Kaylen (1986) modified the formula to \((F_{t} - F_{t-1}) \times (F_{t+1,t} - F_{t,t-1}) < 0\). In a 2 x 2 table of directional change and no change, both forecast and actual, the sum of the diagonal elements is the number of correct forecasts. McIntosh and Dorfman (1992) use the procedure of Naik and Leuthold (1986) to construct a 4 x 4 table that separates peaks, upward movements, downward movements and troughs. For a one step ahead forecast, the known value of \((F_{t} - F_{t-1})\) eliminates certain off-diagonal elements from consideration. For example, a previous upward price movement limits the forecast to a peak turning point or a continuing upward movement. The ratio of accurate forecasts (RAF) is the sum of the diagonal elements divided by the total number of forecasts. For one step ahead forecasts, the RAF has the same value in both 2 x 2 and 4 x 4 contingency tables. The Henriksson–Merton measure is the conditional probability of a correct forecast. McIntosh and Dorfman calculate an exact confidence level based on the hypergeometric distribution as

\[
c = 1 - \sum_{x=n_1}^{\min(N_1, n_1)} \binom{N_1}{x} \binom{N_2}{n-x} \binom{N}{n}/n,
\]

where \(N_1\) is the number of downward observations, \(N_2\) is the number of upward observations, \(N = N_1 + N_2\), \(n_1\) is the number of correct downward forecasts and \(n_2\) the number of incorrect downward forecasts and \(n = n_1 + n_2\). The sum of the number of correct upward and downward forecasts is the same as the values calculated by Kaylen, so that for overall ranking the disaggregation adds to new information.

7.2.3. Vector autoregression methods

Six studies that compared vector autoregression methods are reported in Table 10. As noted by Kaylen (1988), unrestricted VAR models with large numbers of parameters have been found to perform rather poorly. The variable reduction methods that use variable ordering and information criterion statistics produce the most...
Table 10
Pairwise comparison of vector autoregression methods

<table>
<thead>
<tr>
<th>Author</th>
<th>Date</th>
<th>Series</th>
<th>Criterion and horizon</th>
<th>Unrestricted</th>
<th>Variable reduction</th>
<th>Prefilter (Parzen)</th>
<th>Bayesian</th>
<th>VAR overall best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bessler, Babula</td>
<td>1987</td>
<td>4</td>
<td>RMSE</td>
<td>7.8</td>
<td>[4]</td>
<td>[5]</td>
<td>[1]</td>
<td>[3]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tiao-Box</td>
<td>Hsiao, Schwartz</td>
<td>Symmetric</td>
<td>General</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TPE</td>
<td></td>
<td></td>
<td></td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Kling, Bessler</td>
<td>1985</td>
<td>5</td>
<td>RMSE1</td>
<td>1.4</td>
<td>4.4</td>
<td>5.0</td>
<td>0.5</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MAPE1</td>
<td>0.5</td>
<td>3.5</td>
<td>5.0</td>
<td>1.3</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RMSE4</td>
<td>5.0</td>
<td>2.5</td>
<td>2.3</td>
<td>1.4</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MAPE4</td>
<td>4.0</td>
<td>2.6</td>
<td>2.3</td>
<td>1.4</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[1]</td>
</tr>
<tr>
<td>Park, Bessler</td>
<td>1990</td>
<td>4</td>
<td>RMSE1</td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[2]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[2]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[2]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[2]</td>
</tr>
<tr>
<td>Zapata, Garcia</td>
<td>1990</td>
<td>3</td>
<td>RMSE</td>
<td>u0.3</td>
<td>u2.1</td>
<td>d1.2</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>u2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d3.0</td>
<td></td>
</tr>
<tr>
<td>Fanchon, Wendell</td>
<td>1992</td>
<td>4</td>
<td>MSE</td>
<td>5.1</td>
<td>6.0</td>
<td>1.5</td>
<td>2.8</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[2]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[3]</td>
<td>[2]</td>
</tr>
</tbody>
</table>

1 Bessler and Kling (1976) used an almost identical data set (the forecast period was extended a further 12 quarters) and found that the ordering of accuracy was general Bayesian prior, unrestricted, symmetric Bayesian prior.

The values b,w indicate the number of pairwise comparisons that the method listed at the head of the column is better than and worse than other methods examined. Where this information could not be extracted from the study, the values in brackets [thus] indicate the ranking of the methods according to the criterion listed.

The method of Tiao and Box is to delete insignificant variables, with re-estimation and repeated deletion if necessary. Kling and Bessler (1985) use both t-test and F-test statistics to delete insignificant variables. Their performance was similar with one-step ahead forecasts but the t-test method was worse at four-steps ahead. Hsiao and Schwartz methods examine different combinations of lagged variables in each equation, usually after judgementally ranking the series in order of causality. Parameter weighted information criteria (Akaike, Schwartz) are used to find the best set of lags.

The prefiltering method (Parzen) uses the residuals from univariate ARIMA models applied separately to each series.

Zapata and Garcia (1990), u is undifferenced (raw) data and d is first differenced data.

Elitzak and Bilsard (1989) found that Hsiao’s method was generally slightly more accurate than Kaylen’s for five retail meat series at one, four and eight quarters ahead. Both were more accurate than USDA-ERS forecasts.

8. Current developments and conclusions

A much-needed development is to combine forecasting with decision making (the so-called decision support system). Agricultural economists, because of their historical emphasis on the analysis of resource allocation decisions, are perhaps more likely to emphasize this area than are other forecasters. This, and a number of other practices that have been criticized in earlier sections, are areas where improvements need to be made.

8.1. Probability forecasting

Decision makers want information on the probability distribution of a point forecast. The distribution, or the statistics that summarize it,
provides information on the precision the forecaster attaches to the point forecast. Equally, a decision maker can see the risk that attends taking the point forecast as true. Analysts responsible for outlook forecasts have recognized the uncertainty of the forecast in a qualitative way. Outlook reports, from the earliest to the present day, contain verbal descriptors such as ‘should be around’, ‘expected to return to normal levels later in the year’ and so on. Such descriptors convey only part of the information possessed by the commodity analysts. Bottum (1966) and Timm (1966) both recommended that probabilistic outlook forecasts be developed, in the manner of weather forecasts. Nelson (1980) outlined a specific proposal that included a survey of user needs, development of elicitation procedures, training, forecast evaluation and dissemination of results.

Many studies have reported subjective probability distributions elicited from farmers for future prices and yields. Generally, no actual payoffs are made for forecast accuracy. Grisley and Kellogg’s (1983) study is an exception. When actual rewards are used to motivate individuals, choice of scoring rule might be important. The linear scoring rule used by Grisley and Kellogg (1983) is easy to interpret and explain to participants. However, since the amounts of money were relatively large (up to one day’s pay) and the respondents were risk averse, the linear scoring rule was improper (so that the reward structure gave participants incentives to misrepresent to the researcher their true subjective probabilities). A comment on the paper suggested that such a rule could lead to strategic behavior. Respondents could maximize their expected reward by stating more concentrated probability distributions than they actually believed.

In a practical setting, the goal is to obtain subjective probabilities which are as accurate as possible. Payment of cash rewards might matter much more than choice of scoring rule. Nelson and Bessler (1989) provided the same historical information to a sample of participants and asked each to forecast a series of 40 probability distributions. The respondents who were paid according to an improper scoring rule initially stated the same probability forecasts as those paid according to a proper scoring rule, but after about 20 successive forecasts the two groups diverged. While a reward encourages participants to take the task seriously, basing payments on an improper scoring rule ultimately leads to biased assessments.

More common in the general forecasting literature are attempts to mechanically calculate correct (i.e. well calibrated) forecast probability distributions. The problem was recognized in agricultural forecasting in the late 1970s [Teigen and Bell (1978a,b), Spriggs (1978)]. Their confidence intervals on corn price, based on a sector model, did recognize uncertainty in the values of both the parameters and the explanatory variables. Little follow-up work appeared. Prescott and Stengos (1987) demonstrated the bootstrapping approach by constructing confidence intervals for pork production forecasts. No calibration tests were performed. The first testing of this kind to take place in an agricultural context was by Bessler (1984b). A handful of other agricultural applications have appeared since: in a policy context [Bessler and Kling (1989)] and comparing univariate and multivariate [Bessler and Kling (1990)], based on option prices of commodity futures [Fackler and King (1990)]. Knowledge of the option premium and an assumed lognormal distribution enabled Fackler and King to calculate an artificial probability distribution on the futures price 4 weeks and 8 weeks ahead.

Why has probabilistic forecasting received so little attention, despite the fact that the need for it was recognized almost 30 years ago? Outside of weather and sports forecasting, the same question could be asked about any kind of forecasting. Most likely, shortage of data is the problem. Making and testing probabilistic forecasts requires more data than does point forecasting. Compared with daily or more frequent weather forecasts made at many locations, the typical economic price series of monthly, quarterly or even annual frequency offers many fewer opportunities. A reviewer suggested that probabilistic forecasts are more challenging to present
than are point forecasts. This is perhaps indirectly saying that a forecast needs to make sense to the recipient in a decision making context, a second area of current development.

8.2. Forecasting and decision making

Quite early in the development of the theory of decision making under uncertainty, the outcome using perfect information and the outcome using a frequency distribution from a historical series were compared. The difference gave the value of perfect information. The difference between having no predictive information about the outcome as opposed to some predictive information was not often put in the forecasting context, though there were exceptions [Lave (1963), Eidman et al. (1967), Bullock and Logan (1970), DeCanio (1980)].

Some studies focussed explicitly or implicitly on a risk neutral decision maker. Such a person would be unconcerned about the probability distribution around a forecast, except for bias. Consider, for example, the producer of a (fourth quarter) beef calf who must choose between immediate sale, or rearing for sale as a feeder steer (in the second quarter of the following year) [Spreen and Arnade (1984)]. To attain maximum profits, if the feeder steer price exceeds a break-even value, the correct decision is to retain the calf, otherwise it should be sold. Spreen and Arnade compared forecasts of feeder steer prices by five different methods, including one that forecast the probability of feeder steer price exceeding the break-even price. When each season had passed, the correct decision (the one that gave the greater profit) could be seen. In this study, the forecast from an exponential smoothing model gave the fewest wrong decisions. The highest expected profit might carry too much risk for a risk averse producer who might elect to sell calves more frequently than the expected profit maximizer. A risk averse producer would want to know how precise each of the forecasts was.

Even studies that compare decision makers with different risk preferences usually ignore the forecast error distribution. Rister, Skees and Black (1984) examined a similar situation, where a grain producer with a known cost of storage had to choose each month between storage or immediate sale. Whether outlook price information was available or not, the strategy with highest expected return and highest risk was to store for 3 months, then sell. Under these conditions, a risk neutral decision maker would pay nothing for outlook information. For a moderately risk averse decision maker, the strategy with highest expected return and acceptable risk required outlook information. Requiring a payment from the farmer of $450 for the outlook information removed that strategy from the non-dominated set, providing an indication of the value of outlook information to that producer.

When futures markets exist for a commodity, a producer’s hedging decision can be formulated as a portfolio analysis problem. Early researchers simply used historical data as risk measures. Peck (1975) calculated the optimal hedge for egg producers using the forecast error variance of three forecasting methods (one of which was the futures price itself). Forecast information was of little value, since a fully hedged strategy gave almost the same returns and risk reduction as partial hedges based on the forecasts. Brandt (1985) and Park, Garcia and Leuthold (1989) considered how producers and buyers of hogs might execute selective hedging strategies based on prices forecast by different methods. In a selective hedge, the producer (buyer) sells (purchases) a futures contract when the forecast price is below (above) the futures price, otherwise doing nothing. Brandt’s results suggest that the adaptively weighted composite forecast is better than the simple average composite, econometric and ARIMA models, although with quarterly data the analysis is probably too aggregative for useful decision making. Park et al. (1989) used monthly data. They found, perhaps surprisingly, that ARIMA forecasts performed well at all horizons and for all stochastic dominance criteria against simple composite and econometric model competitors.
8.3. Conclusions

Agricultural forecasting uses a wide range of techniques in a wide variety of situations. The largest group are outlook forecasts, mainly of production, at different national and regional aggregations. They have a long history and a detailed and specialized development of leading indicator analysis unique in forecasting. Emphasis, indeed one might say overemphasis, on econometric modeling of ever increasing complexity has been a hallmark of agricultural forecasting. It stems, perhaps, from the desire and the training of agricultural economists to explain phenomena rather that predict them. Often, analysts make conditional predictions or projections based on assumptions that are over-simplifications of any policy that might arise. The ratio of policy analysis to long-term forecasting appears higher in agricultural applications than elsewhere. If this review appears to have concentrated on short-term forecasting, it is because few published long-term forecasts were located.

Results found here conform generally to the beliefs held by forecasters. The conclusions are drawn from published studies and there are many unpublished forecasts. Where conclusions are based on relatively few results, the addition of a small number of studies might reverse the findings. Composite forecasting is best, although (as noted in section 7.2) the case for using simple averaging over other composite methods is less clear. There are too few comparisons to single out a particular composite method as best. For short-term production forecasts, producer intentions are good indicators. Structural econometric methods do less well than their proponents might have expected, perhaps because in most of the comparisons their specifications are not developed in any systematic way. A typical econometric modeller performs a limited amount of testing of dynamic structure (often only the Durbin–Watson test, though matters are improving). There is also no widely agreed belief about the best specification to build from.

Although the types of forecasts required are similar to those found in business, agricultural forecasters have made little use of univariate time series methods. On the other hand, for pure forecasting applications, vector autoregression is replacing simultaneous equations systems. The limited evidence of a number of careful VAR studies supports the generally held view that unrestricted VAR models are not efficient. But there are too few comparisons of variable reduction and Bayesian approaches to favor one over the other.

There is an impression that econometric practices have improved over the years though concrete evidence is hard to produce. Dividing the econometric models in Table 7 somewhat arbitrarily into the 12 studies (22 series) before 1985 and the 11 studies (23 series) since 1985, no significant improvement is revealed in the success rate for short-term forecasts of econometric models against other methods. More studies are needed on the relative success of causal models for long-term forecasts, where they would be expected to do better. Fildes (1985) reported success rates for causal models in ex ante forecasts of 0.76 for medium and long-term ex ante forecasts (11 studies) and of 0.67 for short-term forecasts (ten studies). However, the difference is not statistically significant.

Fascination with the subtleties of different econometric methods has produced numerous articles but has had little influence on performance. While the sensitivity of parameter estimates to choice of specification is now widely acknowledged, papers rarely quantify the fragility of their estimates. Nor has much systematic investigation into dynamic specification occurred. Exceptions are the relative handful of studies that use VAR or that investigate cointegration among variables where more detailed testing is a necessity. The best that can be said is that, since about the late 197Os, subjecting systems of simultaneous equations to within-sample validation has become common practice. It is usually carried out deterministically to assure the stability of the estimated model over time, a test which should be carried out stochastically anyway. Exogenous variables are taken as given, but lagged endogenous variables are usually the
calculated values. Aside from the need to examine dynamic properties, there is no reason to use calculated values of lagged endogenous variables that would be known at the time the forecast is made (i.e., variables whose lags are longer than the forecast horizon). If current exogenous variables were treated as unknown, a much better assessment of forecasting ability would result. Post-sample forecasting is better than within-sample, but using actual values of exogenous variables that would be unknown at the time of forecast (or conditional forecasting) defeats the objective of finding the best forecasting model.

A current unresolved question is that of whether to build more or less structure into a model. Econometricians generally favor more structure, arguing for using all available data and for conforming with economic theory. Forecasters take the opposite view, to avoid adding layers of variability to the forecast calculation. An argument in favor of more structure is that although forecasts may have higher variance, this may truly reflect the level of ignorance about the future. A counter-argument is that data and the relationships of different reliabilities are mechanically aggregated, creating larger forecast errors than would suitable (and probably judgmental) weighting.

Reporting of empirical studies in economics journals is improving. Many journals now require as part of the guidelines for manuscript submission that the data be available and clearly documented and that details of computations sufficient to permit replication be provided. (The *International Journal of Forecasting* was a pioneer in requiring disclosure, revising its guidelines in Fall, 1988, over 1 year ahead of the *American Economic Review* (September, 1989) and the *American Journal of Agricultural Economics* (February, 1990)). Few go as far as the *Journal of Agricultural and Resource Economics* in demanding documentation of model specifications estimated but not reported in the submitted manuscript. Since there are no standard procedures for testing forecasting ability, details that can affect the outcome of a comparison must be given. Is the model re-estimated? In what way? How often? How many steps ahead? (Does one take the omission of this detail to indicate one step ahead?) Are forecast accuracy statistics based on aggregation of all the $h$ steps ahead forecasts or the one through $h$ steps ahead forecast?

More comparisons are needed, if only with the performance of a naive model. RMSE in particular provides little information on relative performance. Theil’s $U_2$ at least gives a comparison against a naive no change model. One suspects that Theil’s statistic is often not reported because of the poor light it would cast on the model’s forecasting performance. Large scale econometric models seem particularly remiss in making comparisons. A good comparison would be with a single reduced form equation with, in both cases, forecast rather than actual current exogenous variables.

Progress depends on researchers following a proper validation procedure. All applied economists (and not just agricultural forecasters) should ask themselves three questions. How well does the model perform out-of-sample (in a holdout sample for cross section analysis, or in the post-estimation period in time series analysis)? How does the model perform compared with the one it is intended to replace? And why the difference? Finally, we should admit that we know less than we claim. Forecasters should, as a matter of course, make quantitative probability statements. Doing so would shift the focus from the point estimate, which will be wrong anyway, to the information content of the forecast.

9. Acknowledgments

I thank David Bessler, Mark Nerlove, the editor, two anonymous reviewers and other colleagues for comments, help and inspiration. I am particularly grateful to Scott Armstrong and Bill Tomek. Without their help, I would have finished this paper a long time ago, but it would not have been as good.


10. References

Only references not found in the bibliography are listed below. Articles in the bibliography referenced in the main text are denoted by R.


11. Appendix A: Publications searched

Quarterly Review of Agricultural Economics, 1948–1978 (Vols. 3–33) (continues with different subject matter emphasis as Agriculture and Resources Quarterly).
The Dialog database files 10 and 110 (Agricola) were searched over titles and descriptors using the keyword roots ‘economic’ and ‘agricultur’ and either ‘forecast’, ‘predict’ or ‘projection’ for the years 1970–1992. This produced 245 citations of which six had already been identified through searching the above journals. Sixty-six citations were identified as possible additional references, although most were cost of production projections or regional crop and livestock supply–demand projections published by various USDA branches. Two articles out of the 245 were added to the bibliography.

The Dialog database file 139 (Journal of Economic Literature) was searched over descriptor codes 7100, 7110, 7120, 7150 (agriculture; general, supply and demand, situation and outlook, markets and marketing) and the keyword root ‘forecast’ over titles and abstracts for the years 1969–1992. This produced 70 citations of which 27 had already been identified through searching the above journals. A further 21 citations were identified as possible references and eight were added to the bibliography.

12. Appendix B: Bibliography

Bibliography key: CO, comparison of forecasting methods; EV, evaluative or critical assessments; MK, market efficiency; IA, large scale (see 6); OU, outlook (see 3); PG, programming model; PR, probability, also includes value of information and Bayesian decision making (see 8); R, article is referenced in main paper; SN, single equation econometric model; ST, single sector model (see 5); TS, time series (see 7).

Agriculture Canada, 1978, Commodity Forecasting Models for Canadian Agriculture, 2 Vols., coordinated by Z.A. Hassan and H.B. Huff (Ottawa, Canada), publication nos. 78/2 and 78/3, October and December 1978. R SN ST Describes ten of the 13 structural commodity models under development by or on behalf of Agriculture Canada. Includes hog supply, hog marketing, pork, feed grains, eggs, dairy, beef, broiler, wheat and farm inputs. “The primary goal of the research program . . . is to improve the forecasts of basic market relationships to complement the work of commodity specialists.” (H.B. Huff, ‘Introduction and Overview’, p. 1). Several of the models are projections rather than forecasting models and model simulation is illustrated rather than forecasting.

Allen, P.G., 1984, A note on forecasting with econometric models, Northeastern Journal of Agricultural and Resource Economics, 13, 264–267. TS Compares within-sample cranberry yield forecasts from univariate ARIMA and from econometric models with various combinations of actual and forecast monthly average temperatures. ARIMA worst MSE, but econometric with forecast temperatures worse than trend regression.


Aradhyla, S.V. and M.T. Holt, 1988, GARCH time-series models: an application to retail livestock prices, Western Journal of Agricultural Economics, 13, 365–374. R TS CO For beef, pork and chicken prices a within-sample test showed that a GARCH model fitted better than an AR model with time trend.


Australian Bureau of Agricultural and Resource Economics, Outlook 1992, Annual. Also references other ABARE publications. OU.


Beenstock, M. and R.J. Bhansali, 1980, Analysis of cocoa price series by autoregressive model fitting techniques, Journal of Agricultural Economics, 31, 237–242. TS CO AR(p) model fitted to monthly data. P = 2 was best fit according to several final prediction error criteria. AR model more accurate than naive no change in one step ahead post-sample forecasting.

Bell, T.M. et al., 1978, OASIS—an overview, Agricultural Economic Research, 30, 1–7. R OU Describes the process of generating and disseminating forecasts in USDA—Economics, Statistics and Cooperative Service (ESCS) using the outlook and situation information system (OASIS).


Bessler, D.A., 1984b, Subjective probability, in P.J. Barry (editor), Risk Management in Agriculture (Iowa State University Press, Ames, IA), Chapter 4, pp. 43–52. R PR Review and possible applications in agricultural economics, including forecasting.

Bessler, D.A., 1990, Forecasting multiple series with little prior information, American Journal of Agricultural Economics, 72, 788–792. TS CO See McIntosh and Dorfman (1990) competition. Uses VAR.


Bessler, D.A. and J.A. Brandt, 1979, Composite Forecasting of Livestock Prices: An Analysis of Combining Alternative Forecasting Methods Purdue University. Agricultural Experiment Station Bulletin 265, West Lafayette, IN. TS CO The study is condensed and reported in Bessler and Brandt (1981).


and Organization, 18, 249–263. TS CO Compares futures price and expert as forecasts of one quarter ahead cash prices of steer cattle and hogs. Fits three variable VAR model using Hsiao's method to find best specification (according to Akaike final prediction error criterion). Futures price contains all information for forecasting hog price, but not for cattle price where expert forecast contains additional information.


Bessler, D.A. and J.L. Kling, 1990, Frequential analysis of cattle prices, Applied Statistician, 39, 95–106. R TS CO PR Compares univariate and two-variable VAR models of daily cash and futures prices of live cattle. VAR has significantly smaller RMSE for out-of-sample one day ahead cash price forecasts. Generates by simulation daily forecast distributions. VAR has tighter distribution and is well calibrated. Univariate cash price forecasts are not. Distributions cannot be recalibrated to later time periods.

Bessler, D.A. and C.V. Moore, 1979, Use of probability assessments and scoring rules for agricultural forecasts, Agricultural Economic Research, 31, 44–47. PR Demonstrates the use of the logarithmic scoring rule (a proper rule that gives the expert assessor the highest payoff for setting stated beliefs equal to true beliefs). Based on the work of R.L. Winkler and A.H. Murphy, 1968, "Good probability assessors", Journal of Applied Meteorology, 7, 751–758.


Blake, M.J. and I. Clevenger, 1984, A linked annual and monthly model for forecasting alfalfa hay prices, Western Journal of Agricultural Economics, 9, 195–199. ST A four equation sector model, two stages. New Mexico.


Brandt, J.A., 1985, Forecasting and hedging: an illustration of risk reduction in the hog industry, American Journal of Agricultural Economics, 67, 24–31. R SN TS CO Compares same methods as Brandt and Bessler (1981). Focus is on use of forecast by producer or buyer to make hedging decision. Selective hedging based on a price forecast is some advantage over routine hedging or no hedging. Little difference in the means and standard deviations of the prices resulting from each strategy.


Brandt, J.A. and D.A. Bessler, 1984, Forecasting with vector autoregressions versus a univariate ARIMA process: an empirical example with U.S. hog prices, North Central Journal of Agricultural Economics, 6, 29–36. R TS CO Quarterly hog prices forecast better by ARIMA model than by four equation VAR (using Tiao’s method of variable reduction) with first-differenced data.

Brandt, J.A., R.E. Young, III and A.W. Womack, 1991, Modeling the impact of two agricultural policies on the US livestock sector: a systems approach, Agricultural Systems,


Carter, C.A. and C.A. Galopin, 1993. Informational content of government hogs and pigs reports. American Journal of Agricultural Economics, 75, 711–718. MK A hypothetical futures trader was assumed to receive the hogs and pigs report one day in advance and buy or sell a futures contract if the report contained unanticipated information. However, the information content of the report is low. Only a risk-neutral trader with low trading expenses would be willing to pay for such advance information.

Cavin, J.P., 1952. Forecasting the demand for agricultural products, Agricultural Economics Research, 4, 65–76. SN An appraisal of the forecasting method used by the USDA; level of economic activity, aggregate agricultural income and prices, individual farm commodities (pork as example). Also tabulates actual annual changes and forecast changes for 5 years.


Cigno, A., 1971, Production and investment response to changing market conditions, technical know-how and government policies, Review of Economic Studies, 38, 63–94. LA LP formulation with dynamics included through capital stock adjustment, adaptation to risk and technical change applied to forecast five crop and one (?) livestock outputs and prices for N.E. Italy. Projections for 1965–1970 compared with actual, usually over 1965–1968 by means of graph only.

(based on Food and Agricultural Regional Model--large scale econometric, including macroeconomic and trade links). See Barichello, Downey, Jones, Owen, Womack for other articles in same issue.


Colman, D. et al., 1975, Forecasting and Projection in the Agricultural Sector Department of Agricultural Economics, University of Manchester, Bulletin No. 151. EV Review of agricultural forecasting.

Colman, D.K., 1967, The application of Markov chain analysis to structural change in the northwest dairy industry, Journal of Agricultural Economics, 18, 351–361. R PR 1958–1965 sample of 236 farms (not for all years) classed into five sizes (number of cows) plus entry/exit group. Actual 1960 population used to predict 1965 distribution, which had error range of 3.4–28.1% compared with actual.


Crowder, R.T., 1972. Statistical vs judgement and audience considerations in the formulation and use of econometric models, American Journal of Agricultural Economics, 54, 779–783. **LA ST EV** Describes the forecasting situation in the commodity industry.


DeCanio, S.J., 1980. Economic losses from forecasting error in agriculture, Journal of Political Economy, 88, 234–258. **R PR** Assumes that farmers are profit maximizers and their observed production is based on an incorrectly predicted product price ratio. Using an assumed product-transformation curve the difference between gross revenue based on the incorrectly predicted prices and that which would result from using perfectly forecast prices is the value of a perfect price forecast.


Dubman, R., R. McElroy and C. Dodson, 1993. Forecasting Farm Income: Documenting USDA’s Economic Model USDA-ERS Technical Bulletin Number 1825, 48 pp. **R LA** The accounting-type model of approximately 1000 equations (listed in the appendix) forecasts cash receipts for 21 crops and 11 livestock commodities, CCC loans for nine crops and values of inventory change for 17 crops and four livestock commodities. It is updated monthly and published quarterly using expected prices and production provided by USDA analysts.

Eales, J.S., B.K. Engel, R.J. Hauser and S.R. Thompson, 1990. Grain price expectations of Illinois farmers and grain merchandisers, American Journal of Agricultural Economics, 72, 701–708. **PR** In most instances the futures price (of soybeans and corn) is an appropriate proxy for expected price. However, volatilities implied by option premia usually overestimate the subjective variances of farmers and merchants, a finding of overconfidence consistent with the psychology literature.


Egbert, A.C., 1969, An aggregative model of agriculture:
empirical estimates and some policy implications, American Journal of Agricultural Economics, 51, 71–86. R LA Annual four equation model treating agriculture as a separate sector with no feedbacks (first generation model). Validated with sample and used for long-term projections.


Eidman, V.R., G.W. Dean and H.O. Carter, 1967, An application of statistical decision theory to commercial turkey production, Journal of Farm Economics, 49, 852–868. R PR Bayesian decision theory used to aid turkey farmer deciding among: independent production, guaranteed payment per turkey (contract A) and payment per pound (contract B). Value of price predictor (from single econometric equation) $600 per year on expected return of approximately $4500. Value of perfect price forecast $2700 per year.


Ezekiel, M., 1927, Two methods of forecasting hog prices, Journal of the American Statistical Association, 22, 22–30. R SN Based on graphs of forecasts 1–6 months ahead, 'empirical' regression (using lagged explanatory variables known at the time of the forecast) was more accurate than a 'synthetic' method of setting expected supply (estimated from leading indicators such as pig crop survey) against a supply function. Forecasts prepared 12 months later suggest opposite conclusion (based on footnoted actual values). Notes (p. 29) “... eventually the most satisfactory results may be obtained by some combination of the... methods”.


Eckler, P.L. and R.P. King, 1990, Calibration of option-based probability assessments in agricultural commodity markets, American Journal of Agricultural Economics, 72, 73–83. R PR Option price premia used to calculate probability distribution around the closing price of the futures contract of the same commodity 4 and 8 weeks ahead. Calibration tests performed on corn, cattle, soybeans and hogs contracts.

Fanaron, P. and J. Wendell, 1992, Estimating VAR models under non-stationarity and cointegration: alternative approaches to forecasting cattle prices, Applied Economics, 24, 207–217. R TS CO Compares restricted VAR and VEC, unrestricted and Bayesian VAR and univariate models of three cattle prices (for different animal weights) and corn price. Over 1 to 58 months ahead, restricted VAR has least MSE, then VEC, with VEC better only at the longer horizons.

Feather, P.M. and M.S. Kaylen, 1989, Conditional qualitative forecasting, American Journal of Agricultural Economics, 71, 195–201. SN TS CO For qualitative forecasting (e.g. change of direction of price) constructs a composite based on Bayesian updating of probabilities. For quarterly hog price forecasts at various horizons up to 44 steps ahead, expert and ARIMA methods made more correct turning point forecasts than composite, econometric worst.

Findlay, J.R., 1968, Farm practice adoption: a predictive model, Rural Sociology, 33, 518. PR A segmentation (classification or configural) method. Uses four observable binary farm or farmer characteristics (e.g. farm size more than or less than 400 labor hours) to discriminate between early and late adopters on a calibration sample of farmers. Tested on three hold-out samples with correct predictions about 70% of the time.

Fisher, M.R., 1958, A sector model—the poultry industry of the U.S.A., Econometrica, 26, 37–66. ST Annual 12 or 11 equation models estimated from 1915–1940 data by OLS, LISE and Cochrane–Orcutt methods. Sought to determine demand-supply simultaneity. Poor price prediction in backcasts to 1913–1914 suggested (p. 62) “... that a number of naive models would be able to do better.”

Foote, R.J. and H. Weingarten, 1958, Alternative methods for estimating changes in production from data on acreage and condition, Agricultural Economics Research, 10, 20–26. R OU CO Compares forecasting methods that use intentions to plant data with methods that use only past or projected acreage and yield. Generally, use of intentions data explains about 60–80% of actual variation in production; methods that do not use intentions data explain about 20–50% of variation.


Foote, R.J., J.A. Craven and R.P. Williams, Jr., 1972, Quarterly models to predict cash prices of pork bellies, American Journal of Agricultural Economics, 54, 603–610. ST A three equation recursive sectoral model, estimated by 2SLS.

Foote, R.J., R.R. Williams, Jr. and J. Craven, 1973, Quarterly and Shorter-term Price Forecasting Models Relating to Cash and Futures Quotations for Pork Bellies USDA-ERS Technical Bulletin Number 1482, 71 pp. SN ST CO Single equations for liveweight and number of barrow/gilts and sows both by quarter and by month (each time period estimated separately).


Gallagher, P., 1986, U.S. Corn yield capacity and probability: estimation and forecasting with nonsymmetric disturbances, North Central Journal of Agricultural Economics, 8, 109–122. SN Shows difference in point and interval forecasts when nonsymmetric ($\gamma$) distribution used compared with normal.


Garcia, P., R.M. Leuthold, T.R. Fortenberry and G.F. Sarassoro, 1988, Pricing efficiency in the live cattle futures market: further interpretation and measurement, American Journal of Agricultural Economics, 70, 162–169. SN TS CO Compares econometric, ARIMA, composite and futures price forecasts 1 to 6 months ahead, with updating of models. Simple average composite of ARIMA and econometric with latest 72 observations, generally has lowest MSE and futures price generally highest. However, simulated trading in the market using the best-to-date model gave small profit relative to variance.


Gertel, K. and L. Atkinson, 1993, Structured Models and Automated Alternatives For Forecasting Farmland Prices USDA ERS Technical Bulletin Number 1821, 22 pp. TS CO Compares ability of OLS and univariate methods to detect trend reversals and the performance of several univariate methods at 1 year and 2 year ahead forecasts in the 1973–
show that composite formed by unrestricted regression is most accurate.


Believed to be the first economic study to use substantial monetary rewards (up to 1 day's pay) as incentives to elicit accurate subjective probability assessments. Used the visual impact (visual counter) method on a sample of 30 small-scale farmers to get distributions on future prices and yields of their rice, tobacco, soybean and peanut crops.


In reply to Knight et al. admits that the linear scoring rule used can be improper but (1) is easy to communicate (2) not necessarily improper for risk averse individuals (as these were) and (3) is less important than getting the individuals to take the elicitation process seriously.

Gruen, F.H. et al., 1967, Long Term Projections of Agricultural Supply and Demand, Australia 1965 to 1980, Department of Economics, Monash University, Clayton, Australia. LA.


OU CO Analyses 1100 crop production forecasts for barley, corn, oats, potatoes, soybeans, spring wheat and winter wheat for 1929–1970. Makes successive pairwise comparisons of naive no change forecast, initial and revised USDA forecasts. Uses three criteria: accuracy improvement (Theil's revision ratio statistic, R), absolute forecasting error and bias.


Most price analysis models not used for forecasting.


Harlow, A.A., 1962, Factors Affecting the Price and Supply of Hogs USDA-ERS Technical Bulletin Number 12/4, 89 pp. ST A six equation quarterly model of US hog industry. Eight post-sample one step ahead forecasts had ratio of unexplained to total variation of from 0.08–0.93.

Harns, H.M., Jr., 1976, University outlook programs: a review and some suggestions, Southern Journal of Agricultural Economics, 8, 139–149. OU EV Based on survey of 15 agricultural economics departments.


SN TS CO Same data and results as Harris and Leuthold (1985).
alternative forecasting techniques for livestock prices: a case study, North Central Journal of Agricultural Economics, 7, 40–50. SN TS CO Compares single equation OLS, ARIMA, composite, GLS and multivariate ARMA models of cattle and hog prices.


Hayami, Y. and W. Peterson, 1972. Social returns to public information services: statistical reporting of U.S. farm commodities. American Economic Review, 62, 119–130. PR Uses inventory adjustment model. If decision makers adopted USDA production forecasts, then reduction in average forecast error from 3% to 1% would increase aggregate economic surplus about 6% of gross value of crops and about 1% of gross value of livestock products.


Headly, E.O. and D.R. Kaldor, 1954. Expectations and errors in forecasting agricultural prices, Journal of Political Economy, 62, 34–47. OU Forecasts of eight commodity prices 5 to 12 months ahead (depending on commodity) collected from about 200 Iowa farmers. Reports mean forecast, actual price, mean error and distribution of individual forecasts. Also probability forecasts for corn and hog prices.


Holz, M.T. and J.A. Brandt, 1985. Combining price forecasting with hedging of hogs: an evaluation using alternative measures of risk, Journal of Futures Markets, 5, 297–309. SN TS CO Similar to Brandt (1985) but with monthly data. Compares econometric, ARIMA, composite and seasonal index forecasts of hog price 2 to 10 months ahead. Risk neutral and risk averse producers are better off using the forecasts to create or liquidate selective hedges compared with routine and no hedging strategies.


Huang, K.S., 1989. A forecasting model for food and other expenditures, Applied Economics. 21, 1235–1246. LA.


Huddleston, H.F., 1958. Objective methods in forecasting components of corn yield, Agricultural Economics Research, 10, 49–53. R OU On 1 August, number of ears forecast from stalk count. Weight of grain forecast from number of ears in 60 feet of row.


Inge, M. and J. Ferris, 1983. An evaluation of a combination of quarterly and annual models in predicting cattle and hog prices, presented at the American Agricultural Economics Association annual meetings, West Lafayette, IN. (Abstract in American Journal of Agricultural Economics, 65 (December 1983), 1185). SN TS CO Annual demand equations for cattle and hogs combined with quarterly ratio models to produce quarterly models. These were more accurate than standard quarterly models and ARIMA models.

Irwin, S.H., M.E. Gerlow and T.-R. Liu, 1991. The market timing value of outlook price forecasts, presented at the annual meeting of the American Agricultural Economics Association, Manhattan, Kansas, 18 pp. OU CO Compares four different hog price outlooks and three cattle price outlooks one, two and three quarters ahead for ability to predict turning points using regression-based Merton test. Three hog and one cattle price outlooks at one quarter ahead have value in directional indication, but only the hog price outlooks are significantly better than the forecast from a trend plus seasonal dummy variable regression.


Just, R.E., 1993. Discovering production and supply relationships: present status and future opportunities, Review of Marketing and Agricultural Economics, 61, 11–40. R EV The profession has been too occupied with flexible functional forms and duality analysis. Recommends structured representation: incorporating economic theory, imposing in esti-
mation all of the economic principles and practical information that is otherwise considered in evaluating plausibility of results. The result is models with globally plausible functional forms and implications. The key criterion for success is the ability to represent out-of-sample phenomena. Just, R.E. and G.C. Rausser, 1981. Commodity price forecasting with large scale econometric models and the futures market. American Journal of Agricultural Economics, 63, 197–208. R OU LA CO Compares futures prices of eight commodities with USDA and four commercial forecasts.


Kelly, B.W., 1957. Preliminary report on objective procedures for soybean yield forecasts. Agricultural Economics Research, 9, 139–141. R OU Study in which objective was solely to predict number of pods.


Kingma, O.T., J.L. Longmire and A.B. Stoeckel, 1980. A review of three research programs in quantitative modeling in the Bureau of Agricultural Economics, Australian Journal of Agricultural Economics, 24, 224–247. R EV Covers modeling production systems (LP); modeling commodity markets: annual, quarterly, medium term (5 year) projections; progress: 1960s, single equation aggregate State or national models, mid 1970s (wool, livestock models) multi-equation, multi-enterprise system, (Freebairn, 1973); forecasting ability (Freebairn (1975), Bourke (1979), Gellatly (1979)); modeling macroeconomic systems, structural OE systems: the ORANI module of the IMPACI project. Approximately 90 references.


Konyar, K. and K. Knapp, 1990, Dynamic regional analysis of the California alfalfa market with government policy impacts, Western Journal of Agricultural Economics, 157, 22–32. ST 25 county/region demand and acreage regressions used to forecast directly and in spatial equilibrium model.


Koontz, S.R., M.A. Hudson and M.W. Hughes, 1992, Livestock futures markets and rational price formation: evidence for live cattle and live hogs, Southern Journal of Agricultural Economics, 24, 233–249. MK Rational price formation (where prices for contracts reflect average cost of production) is generally supported by distant live cattle and live hog futures. But after feeding commitments are made, market prices reflect expected market conditions.

Kost, W.E., 1981, The agricultural component in macroeconomic models, Agricultural Economics Research, 33, 1–10. R LA A survey of the individual country models in project LINK, including a tabulation of the number and type of equations and variables in the 25 overall models and in the agricultural sectors. Also reviews the international models of several commercial macroeconomic forecasters.


Kulshreshtha, S.N., J.D. Spriggs and A. Akinfemiwa, 1982, A Comparison of Alternative Approaches to Forecasting Cattle Prices in Canada, Department of Agricultural Economics Technical Bulletin 82-01, University of Saskatchewan, 69 pp. R TS CO Compares five approaches each with and without variable updating for forecasting monthly slaughter steer and feeder steer prices. Composite method most accurate and has best turning point performance for all horizons (6, 12 and 36 months) when variables updated, but generally poor when variables not updated.


Kutish, F.A., 1955, Needed changes in state and local crop and livestock reports, Journal of Farm Economics, 37, 1050–1055. OU To increase accuracy of livestock slaughter forecasts, quarterly reports should contain breeding intentions data by month.

Labys, W.C., 1975, The problems and challenges for international commodity models and model builders, American Journal of Agricultural Economics, 57, 873–878. R LA ST EV ST Naive or no change models superior to econometric.

Labys, W.C., 1987, Primary commodity markets and models: an international bibliography (Gower Press, Aldershot, UK), 290 pp. ST.


Ladd, G.W. and Y. Kongtong, 1979, Use of planting intentions to predict actual plantings, North Central Journal of Agricultural Economics, 1, 97–104. R OU SN EV CO For six grain crops, use of intentions data combined with 'objective' data (price, expected yield of crop and competing crops) better than either alone.


model with 18 supply and demand regions, inventory demands for six major countries and policy intervention modelled explicitly for five regions (US, Canada, EEC6, EEC3, Japan).

Lave, L.B., 1963. The value of better weather information to the raisin industry. Econometrica, 31, 151–164. OU A perfect weather forecast over the drying interval (two periods, about 3 weeks) could increase a California raisin grape grower’s expected profit by $91 per acre compared with the best harvest strategy followed in the absence of a weather forecast. Providing a forecast may raise total raisin production, lowering the per acre value of information.


Leuthold, R.M., 1974. The price performance on the futures markets of a nonstoreable commodity: live beef cattle. American Journal of Agricultural Economics, 55, 271–279. MK Tests efficiency and bias of futures price as forecast of closing price by regression using monthly data. Calculates MSE treating the delivery price as actual and the futures price or cash price up to 36 weeks prior as forecast. MSE futures larger than MSE cash for about 15 or more weeks prior to delivery date.


MacDonald, S., 1992. The Accuracy of USDA’s Export Forecasts USDA-FRS-CFD Staff Report Number AGES 9224, 46 pp. OU Reports MAPE for forecasts of annual value of exports of 13 commodity groups and volume of nine groups. Regressions of change in actual as function of change in forecast find some significant bias (mostly upwards) and some inconsistency (slope significantly different from one) in various commodities and regions.


McFarquhar, A.M.M. and M.C. Evans, 1971, Projection models for U.K. food and agriculture, Journal of Agricultural Economics, 22, 321–345. ST Six models: (1) consumer expenditure (linear expenditure system) for 27 food plus nonfood used as final demands in (2) input–output model for 39 agricultural and non-agricultural products, (3) three equation wheat, (4) two equation barley, (5) 20 equation annual cattle and (6) six equation sheep models.


Menikaus, D.J. and R.M. Adams, 1981, Forecasting price movements: an application of discriminant analysis, Western Journal of Agricultural Economics, 6, 229–238. SN CO An 18 equation binary dependent variable model of feeder/yearling cattle price movement as function of other prices and inventory change.


Midmore, P., 1993, Input–output forecasting of regional agricultural policy impacts, Journal of Agricultural Economics, 44, 284–300. LA Accuracy of forecasts of the Welsh agricultural sector from eight input–output tables rapidly declines as horizon increases because of insufficient attention to final demand forecasts. Spill-over effects into other sectors and functional relations between labor and output.

Miller, B.R. and R. Jelinek, 1982, Relative accuracy of price expectations held by Georgia farmers and by other forecast sources in 1980, University of Georgia College of Agriculture Experiment Station Research Bulletin Number 286, 33 pp. LA CO Compares forecasts 3 months ahead for October prices of corn, soybeans, feeder and steer cattle and hogs. Forecasts by experts (farmer survey), futures price, naive, simple average composite of the above and four large scale econometric models (Chase, DRI, Wharton and OASIS). No significant difference in forecasting performance according to Mann–Whitney U test on either RMSE, MAPE or Theil’s U criterion.


Miller, S., 1979, The response of futures prices to new market information: the case of live hogs, Southern Journal of Agricultural Economics, 2, 67–70. MK

571–589. MK For wheat, corn and the soybean complex (beans, oil, meal) results showed that market participants did await USDA announcements to make trading decisions. The first forecasts of the crop year are more important than later ones.

Moffitt, L.J., R.L. Farnsworth, L.R. Zavaleta and M. Kogan, 1986, Economic impact of public pest information: soybean insect forecasts in Illinois, American Journal of Agricultural Economics. 68, 274–279. PR Proposed system for forecasting pest damage compared with existing commercial scouting system. Unless the forecast leads to the same decision as information from scouting at least 90% of the time, scouting will continue to be used. Forecast reliability of 91% is worth 2 cents per acre and a perfect forecast is worth 66.7 cents per acre.

Moore, H., 1917. Forecasting the Yield and Price of Cotton (MacMillan, New York). R SN Regression of cotton yield on May rainfall (for May forecast), on June temperature (for June, July forecasts) and also on August temperature (for August forecast) for 1892–1914 data for Georgia, Alabama and South Carolina had smaller RMSE than USDA forecasts based on condition indexes made 1 or 2 months later.

Myer, G.L. and J.F. Yanagida, 1984, Combining annual econometric forecasts with quarterly ARIMA forecasts: a heuristic approach, Western Journal of Agricultural Economics, 9, 200–206. SN CO ARIMA model used to get quarterly weights of alfalfa price (Petaluma, CA) that are combined with forecasts from econometric model.


Naik, N. and B.L. Dixon. 1986. A Monte-Carlo comparison of alternative estimators of autocorrelated simultaneous systems using a U.S. pork sector model as the true structure, Western Journal of Agricultural Economics, 11, 134–145. R ST CO For a three equation monthly sector model with assumed autocorrelations of 0 to 0.8, OLS reduced form estimation is more accurate within sample, but 2SLS and autocorrelation correction techniques better post-sample.


Nelson, A.G., 1980. The case for and components of a probabilistic agricultural outlook program, Western Journal of Agricultural Economics, 5, 185–193. R OU EV A proposal covering the requirements for a probabilistic outlook program, survey of user needs, development of elicitation procedures, training, evaluation and dissemination.


Nelson, R.G. and D.A. Bessler, 1989, Subjective probabilities and scoring rules: experimental evidence, American Journal of Agricultural Economics, 71, 363–369. R PR Tests the importance of a proper versus improper scoring rule (see comment by Knight et al., 1985). Each subject made a succession of 40 probabilistic forecasts, each from the same series. After about 19 forecasts, those subjects paid according to the linear (improper) scoring rule behaved strategically and no longer stated their believed subjective distributions.

Nerlove, M., 1958, The Dynamics of Supply: Estimation of Farmers’ Response to Price (Johns Hopkins Press, Baltimore). 268 pp. R SN Describes the partial adjustment schemes for price and quantity that Nerlove and others developed in the 1950s. Chapter 3 (pp. 66–86) is a history of dynamic supply response analysis in agriculture from 1917.


Palmer, C.D. and E.O. Schlozhauer, 1950. Methods of forecasting production of fruit, Agricultural Economics Research 2–1, 10–19. R OU Compares standard "par method" with two other methods. All require an estimate of condition of the crop by crop reporters and use various correlations (or bivariate regressions).


Park, W.I., P. Garcia and R.M. Leuthold, 1989. Using a decision support framework to evaluate forecasts, North Central Journal of Agricultural Economics, 11, 233–242. R ST TS CO PR Compares hog price forecasts from econometric (two equation), ARIMA, composite and naive models when forecasts are used to guide producers, buyers or speculators in trading futures contracts. Risk efficiency criteria (FSD, SSD, E-V, risk neutral) used. ARIMA is always in dominant set and for risk neutral decision makers is always best.


Peck, A.E., 1975. Hedging and income stability: concepts, implications and an example. American Journal of Agricultural Economics, 57, 410–419. R PR Calculates and compares hedging strategies for egg producers using regression, expert and futures prices as forecasts 1–5 months ahead. Generally, the futures price was the more accurate forecast. Hedging reduced risk exposure, with little difference between total hedging and partial hedging based on the different forecasts.


Prescott, D.M. and T. Stengos, 1987, Bootstrapping confidence intervals: an application to forecasting the supply of pork. American Journal of Agricultural Economics, 9, 266–273. R TS Demonstrates bootstrapping for both fixed and random explanatory variables at four forecast dates, but conducts no tests. Forecast distributions have positive skew.

Price, J.M., R. Seeley and C.K. Tucker, 1992, The Food and Agricultural Policy Simulator, Estimation of Farm Production Expenditures USDA-ERS Technical Bulletin Number 1803, 32 pp. LA Annual 15 equation submodel for 15 farm expense categories estimated by OLS or (where necessary) by maximum likelihood to correct for first order autocorrelation. Within-sample validation using mean absolute relative error (MAPE/100), Theil's U2 (none over one) and relative turning point error. Re-estimation of same specification after dropping last year of data allows limited 'post-sample' testing.

Ryland, G.J., 1975, Forecasting crop quality, Review of Marketing and Agricultural Economics, 43, 88–102. TS Quadratic spline and ARIMA models applied to weekly data on sugar content of cane sugar.


Sapsford, D. and Y. Varoufakis, 1990, Forecasting coffee prices: ARIMA vs. econometric approaches, Rivista Internazionale di Scienze Economiche e Commerciali, 37, 551–563. SN TS CO Monthly econometric model with all lagged explanatory variables more accurate than seasonal ARIMA for 36 (one step ahead?) ex ante forecasts.

Sarle, C.F., 1925, The forecasting of the price of hogs, American Economic Review 15 Number 3 Supplement Number 2, 1–22. R SN The essay awarded the Balson prize by the AEA. Regression to predict monthly change in hog price (seasonally adjusted) as a function of detrended lagged prices of industrial stocks, corn (annual average) and hogs. Predicts hog prices (equivalent to three steps ahead forecasts) both within and post-sample.


Schluter, G., 1974, Combining input–output and regression analysis in projection models: an application to agriculture, Agricultural Economics Research, 26, 95–105. LA Gross farm product predicted from gnp and output per dollar of final demand for each of ten agricultural sectors, based on 1963 US I-O table. Prediction error regressed on ratio of gfp (gross farm product) and gnp implicit price deflators, ratio of farm output to input and time trend, which improved the post-sample forecast.


Schroeder, T., J.B. Blair and J. Mintert, 1990, Abnormal returns in livestock futures prices around USDA inventory report releases, North Central Journal of Agricultural Economics, 12, 293–304. R MK EV Reviews previous literature on impacts of outlook information on prices. Finds few significant abnormal returns in futures markets. Outlook reports do provide new information, but less information available for hogs than for cattle.

Selzer, R.E. and R.J. Eggert, 1949, Accuracy of livestock price forecasts at Kansas State College, Journal of Farm Economics, 31, 342–345. OU Developed scoring system to rate qualitative forecasts. Monthly forecasts of hog prices 1925–1940 were 64% accurate and cattle prices were 62.7% accurate. Cannot be compared with standard accuracy measures.

Shafer, C.E., 1989, Price and value effects of pecan crop forecast 1971–87, Southern Journal of Agricultural Economics, 21, 97–103. OU CO Pecan prices are more accurately explained by early season crop forecasts than by post-season final estimates, according to the accuracy of within-sample predictions of five different specifications of average farm price.


and production of upland and American Pima cotton for
different months of the crop year (August–July) from 1965–
1990. Tabulated forecast and final estimate values permit
calculation of standard forecast accuracy statistics. No com-
parisons with other forecast methods.

Smallwood, D.M. and J.R. Blalock, 1986, Forecasting
performance of models using the Box–Cox transformation,
Agricultural Economics Research, 38, 14–24. SN Monte
Carlo analysis of known three variable Box–Cox equation
with different values of λ parameter and sample sizes of 30
and 60. For λ greater than zero, forecast RMSE is smaller
in the bigger sample, while this is not true with λ less than zero.

Smith, B.B., 1925, Forecasting the acreage of cotton,
SN First differences of acreage regressed on first differences
of average spot prices for 5 separate months (known at time
of forecast) and time trend using 1907–1921 data. Probable
error (within-sample 50% confidence interval) was 2%.

Smith, B.B., 1927, Forecasting the volume and value of
the cotton crop. Journal of the American Statistical Associa-
tion, 22, 442–459. R SN Uses correlation analysis on eight
variables (including time and production) to forecast cotton
price.

Smyth, D., 1973, Effect of public price forecasts on market
price variation: a stochastic cobweb example, American
Journal of Agricultural Economics, 55, 83–88. PR Variance
shown to be theoretically always less with a forecast than
without.

Soliman, M.A., 1971, Econometric model of the turkey
industry in the United States, Canadian Journal of Agricul-
tural Economics, 19, 47–60. R ST CO A five equation annual
sector model estimated by OLS. 2SLS, 3SLS and LISE. No
method emerged as clearly the most accurate forecaster.

Spilka, Jr., W., 1983. An overview of the USDA crop and
livestock information system, Journal of Futures Markets, 3,
167–176. OU Lists and discusses the USDA reports of
various crop and livestock inventories and production in
process that act as leading indicators of agricultural pro-
duction and prices.

Spreen, T.H. and C.A. Arnade, 1984, Use of forecasts in
decisionmaking: the case of stocker cattle in Florida, South-
ern Journal of Agricultural Economics, 16, 145–150. R SN
TS CO PR Compares five methods (including naive no
change) of forecasting second quarter feeder steer price as
guide to producer of fourth quarter calves when considering
whether or not to raise them to feeders. Single equation
regression most accurate, but exponential smoothing most
useful for decision making.

Spreen, T.H., R.E. Mayer, J.R. Simpson and J.T.
McClave, 1979, Forecasting monthly slaughter cow prices
with a subset autoregression model, Southern Journal of
Agricultural Economics, 11, 127–131. TS ARIMA model of
monthly Florida cow prices, all grades.

Spriggs, J., 1978, A note on confidence intervals for corn
price and utilization forecasts, Agricultural Economics Re-
search, 30, 32–33. R PR Shows that the standard errors of
forecast from Goldberger method are about ten times larger
than those from the approximate method of Teigen and Bell
(1978).

Spriggs, J., 1981, Forecasts of Indiana monthly farm prices
using univariate Box–Jenkins analysis and corn futures
prices, North Central Journal of Agricultural Economics, 3,
81–87. TS CO PR Compares ARIMA, futures price and four
composites. Equal weights combinina most accurate for
composite, but except for one step ahead, futures price is
best.

Stillman, R.P., 1985. A quarterly model of the livestock
industry USDA-ERS Technical Bulletin Number 1711, 40
pp. R ST A 29 equation (eight annual) cattle, hog and
chicken model.

Stillman, R.P., 1987, A quarterly forecasting model of the
U.S. egg sector USDA-ERS Technical Bulletin Number
1729, 23 pp. ST A ten equation model validated by dynamic
simulation both within and post-sample. A part of the model
described by Westcott and Hull (1985).

Stonehouse, D.P., D.H. Harrington and R.K. Sahi, 1978,
An econometric forecasting and policy analysis model of the
Canadian dairy industry, in Agriculture Canada, Commodity
Forecasting Models for Canadian Agriculture, Vol. 1, coordi-
nated by Z.A. Hassan and H.B. Huff, Ottawa, Canada.
Publication no. 78/2, pp. 77–110. ST Basic 42 equation
annual model, with some identities functioning as operating
rules in forecasting/projecting, plus other accounting equa-
tions to calculate producer revenues, government support
costs, producer and consumer surpluses, totalling 72 expres-
sions in all.

Subotnik, A. and J.P. Houck, 1982, A quarterly econo-
metric model for corn: a simultaneous approach to cash and
futures markets, in G.C. Rausser (editor), New Directions in
Econometric Modeling and Forecasting in U.S. Agriculture
(North-Holland, New York). Chapter 8, pp. 225–255. ST A
seven equation plus annual production equation model
validated within-sample by static and dynamic simulation.
Farm price is a function of futures price and quarterly stock
carryover.

forecasts news?, American Journal of Agricultural Eco-
nomics, 71, 1–8. MK The relative change of corn and soybean
futures prices on the day of a USDA announcement is
significantly different from the change ≥ days before and
after. August, September and October announcements ap-
pear to have a stronger impact than July and November.

Surls, F. and G. Gajewski, 1990, How accurate are
USDA’s forecasts?, Agricultural Outlook, USDA-ERS, A0-
164, pp. 2, 4–5. R OU Compares average forecasting error 41
tion and exports for wheat, coarse grains and soybeans.

Swamy, P.A.V.B., R.K. Conway and M.R. LeBlanc, 1989,
The stochastic coefficients approach to econometric model-
ing, part III: estimation, stability testing and prediction,
Journal of Agricultural Economics Research, 41, 4–20. SN
CO Table 1 contains 20 out-of-sample comparisons of root
mean square errors from fixed coefficient and stochastic
coefficient models, including three livestock price models.
from Conway, Hallahan, Stillman and Prentice (1987). The pork model of Conway et al. is one of the two where the fixed coefficient model is better.


Taylor, P.D. and W.G. Tomek, 1984. Forecasting the basis for corn in western New York, Journal of the Northeastern Agricultural Economics Council, 13, 97–102. SN Basis (futures price minus cash price) is predicted from an annual single equation. Auxiliary equations needed to predict values of explanatory variables show large forecast errors. These lead to large forecast errors in basis, as shown by a single within-sample prediction.


Teigen, L.D. and T.M. Bell, 1978a, Confidence intervals for corn price and utilization forecasts, Agricultural Economics Research, 30, 23–29. R PR Forecast variance is approximated by standard error of estimate squared (often obtained judgmentally) times a correction factor \((n + k)/n\), where \(k\) is the average number of parameters per equation and \(n\) is the average sample size of each estimated equation. See comment by Spriggs (1978) and reply.

Teigen, L.D. and T.M. Bell, 1978b, Confidence intervals for corn price and utilization forecasts: a reply, Agricultural Economics Research, 30, 34–35. R PR Compares USDA forecast RMSE with standard error from Goldberger method and authors’ approximation. Their approximation is closer to USDA values than Goldberger is (and smaller). No comparative tests against actual variability.

Thompson, R.L. and P.C. Abbott, 1982, New developments in agricultural trade analysis and forecasting, in Gordon C. Rausser (editor), New Directions in Econometric Modeling and Forecasting in U.S. Agriculture (North-Holland, New York), Chapter 12, pp. 345–387. R LA Review with over 100 references. Notes (p. 371) “Of the few modeling exercises that did list forecasting as an objective, almost none provide any forecasting performance measures out of the range of data used to estimate the model.”

Thomson, J.M., 1974, Analysis of the accuracy of USDA hog farrowings statistics, American Journal of Agricultural Economics, 56, 1213–1217. OU Compares accuracy 42 (using MAPF) and improvements (using Theil’s \(R\) statistic) of first and second revisions (in quarters two and three) to original estimates (in quarter one) for non-probability mail (rural carrier, September 1959 March 1963), probability mail (March 1963–March 1970) and multiple frame (list frame and probability area frame March 1970–March 1973) surveys. Non-probability was more accurate, though perhaps it refers to a period of limited variability.


ST A ten equation model with beef supply, demand, price and exports and 12 months of forecasts.


Tolley, H.R., 1931, The history and objectives of outlook work, Journal of Farm Economics, 13, 523–534. OU In 1929, 40 states produced outlook reports. Reports and meetings viewed by the Extension Service as ‘wedges’ (foot in the door) to get farmers to think about economic issues in making their plans. Referenced in Kunze (1990).

Tomek, W.G. and R.W. Gray, 1970, Temporal relationships among prices on commodity futures markets: their allocative and stabilizing roles, American Journal of Agricultural Economics, 52, 372–380. MK Compares futures price behavior for continuously storable commodities (corn, soybeans) where inventory hedging is important and discontinuously storable commodities (Maine potatoes) where forward pricing is important. Perhaps surprisingly, corn and soybean futures had best price forecasting ability and potato futures best price stabilizing ability.


U.S. Department of Agriculture. Statistical Reporting Service. 1977. Hog reports and market prices. Agricultural Situation. OU Over a 4 year period, weekly average price after Hog and Pigs released was about equally higher and lower than weekly average price in week when report was released. Referenced in Hoffman (1980).


U.S. General Accounting Office. 1991a, Short-Term Forecasting: Accuracy of USDA's Meat Forecasts and Estimates GAO/PEMD-91-16, 76 pp. R OU CO Calculates bias error and total error (MAPE) of annual USDA forecasts of production and price of beef, hogs and broilers. Total errors averaged less than 6% with small underestimates of production and broiler price and small overestimates for beef and hog prices in the 1983–1989 period. Errors were similar to those made by private analysts.

U.S. General Accounting Office. 1991b, USDA Commodity Forecasts: Inaccuracies May Lead to Underestimates of Budget Outlays GAO/PEMD-91-24, 88 pp. OU Examines USDA forecasts 1–5 years ahead for production, price, exports and stocks of four major crops and dairy products. During 1981–1988, all prices and most other variables were overestimated (negative bias). Suggests improvements to USDA's forecasting process.


Vroomen, H., 1991. Forecasting retail fertilizer prices: a combined time series regression analysis approach USDA-ERS Technical Bulletin Number 1789, 14 pp. SN Monthly wholesale price and transportation cost used as explanatory variables in regression equation for retail price of each fertilizer constituent (anhydrous ammonia, phosphoric acid and potassium chloride). Regression equations for retail prices of 14 fertilizer products based on retail prices of the three main constituents. Six steps ahead forecasts of wholesale prices and transport costs from ARIMA models and autoregression respectively.

Vukina, T., 1992, Hedging with forecasting: a state-space approach to modeling vector-valued time series, Journal of Futures Markets, 12, 307–327. TS Model that forecasts both cash and futures prices (of daily soybean meal prices) slightly more accurate than model of futures price alone or cash price plus basis. Using any model as signal is more profitable than either routine hedge or no hedge over a 3 month period, with practically no difference among models.


Wells, G.J., 1980. Forecasting South Carolina tomato prices prior to planting, Southern Journal of Agricultural Economics, 12, 109–112. SN Annual single equation, based on data available in February.

Westcott, P.C., 1986. Aggregate indicators in the quarterly agricultural forecasting model: retail food prices USDA-ERS Staff Report AGES 860916, 32 pp. SN CPIs for 21 food groups estimated by OLS and 3SLS.

Westcott, P.C., 1986. Aggregate indicators in the quarterly agricultural forecasting model: retail food prices USDA-ERS Staff Report AGES 860916, 32 pp. SN CPIs for 21 food groups estimated by OLS and 3SLS.

Westcott, P.C., 1986. Aggregate indicators in the quarterly agricultural forecasting model: retail food prices USDA-ERS Staff Report AGES 860916, 32 pp. SN CPIs for 21 food groups estimated by OLS and 3SLS.

Westcott, P.C., 1986. Aggregate indicators in the quarterly agricultural forecasting model: retail food prices USDA-ERS Staff Report AGES 860916, 32 pp. SN CPIs for 21 food groups estimated by OLS and 3SLS.
Technical Bulletin Number 1717, 34 pp. ST A nine equation
model tested by dynamic simulation (using actual exogenous
variables) one to four steps ahead both within-sample and
post-sample. Milk stocks, imports, support price and price
deductions are all contemporaneous exogenous variables.
Westcott, P.C. and D.B. Hull, 1985, A Quarterly Forecasting
Model for U.S. Agriculture USDA-ERS Technical Bulletin
Number 1700. R LA A 160 equation model of the corn,
wheat, soybean complex, beef, hog and poultry sectors (the
livestock sectors based on Stillman (1985)). Within-sample
and post-sample dynamic simulations one to four steps ahead
using actual values of exogenous variables.
Yu, F.C. and P. Orazem, 1990, The rationality and value of USDA crop forecasts, presented at the American Agricultural Economics Association annual meetings, Vancouver, Canada. (Abstract in: American Journal of Agricultural Economics, 72 (December 1990): 1353.) MK Several forecasts of planted acreage and harvest size of barley, corn, oats, soybeans and spring wheat are found to be inefficient and/or biased. Early forecasts more valuable than later for market supply information.
Zapata, H.O. and P. Garcia, 1990, Price forecasting with time-series methods and nonstationary data: an application to monthly U.S. cattle prices. Western Journal of Agricultural Economics, 15, 123–132. TS CO Compares ARIMA, VAR (raw and differenced data), error correction (cointegration) model and asymmetric Bayesian VAR (raw and differenced data) for forecasting monthly slaughter steer prices. Differ-
enced VAR was most accurate for one to six steps ahead forecasts.
Biography: P. Geoffrey ALLEN is a Professor of Resource Economics at the University of Massachusetts, Amherst. He has a Ph.D. in agricultural economics from the University of California, Davis. His research interests are in analysing production decisions under uncertainty.