University of California, Davis

Department of Agricultural Economics

Forecasting Farmland Prices: A Comparison of Econometric and Time Series Methods

by

Rulon Pope

Randall Kramer

Richard Green

B. Delworth Gardner

Paper presented at American Agricultural Economics Association Annual Meetings, Blacksburg, Virginia, August 7, 1978.

I. Introduction

Recently, much concern has been shown by virtually all connected with agriculture over the rising prices of farmland (e.g. Crowley, Jundt, U.S. News and World Report, Harris). For each of the 40 years between 1937 and 1977 the value of agricultural real estate was higher than the year before, a phenomenon unmatched in the entire history of our country (Gardner). Since 1973, the rate of increase has been particularly rapid, although there is some speculation that the 1978 data will indicate some leveling off from 1977.

The concern is that current and expected net farm income cannot "support" even present land prices, particularly under the price-cost squeeze conditions that seem to lay immediately ahead. Between 1973 and 1976, real U.S. farm real estate values increased at an annual rate of 9.1 percent, while real net farm income fell at an annual rate of approximately 19.3 percent. Even if net farm income recovers in aggregate, the "income-solvency" problem will be severe for those agricultural producers who have invested heavily in new land purchases these past few years.

We would argue that the farmland price issue is the prime ingredient in the highly uncertain and worrisome economic situation in which agriculture finds itself, best exemplified in the frenzied posture of the American Agricultural Movement. High land prices get translated into heavy mortgage payments for those entering agriculture and those expanding their equity base of operations. On the other hand, rapidly shifting land prices mean large changes in the wealth positions of land owners, even those who have not made recent exchanges in the land market. Thus, all agricultural producers as well as those who are affected by agriculture have a vital stake in what happens to land prices. If accurate forecasts of land prices were available, more efficient decisions would be made by agricultural producers and those who supply services to the agricultural sector of the economy.

A number of recent studies have sought to explain movements in land prices (e.g., Klinefelter, Morris). We question, however, whether previous published models of the land market offer insights that can yield forecasts that are sufficiently reliable. The approach followed is not to critically evaluate or revamp earlier models, but to examine their plausibility as explanations of recent market events, and study their predictive ability. The Box-Jenkins forecasts are considered as benchmark results in comparing various models.

II. Some Econometric Models of Farmland Prices

Several simultaneous equation models of the U.S. farm real estate market have been developed. Three of the best known models are those presented by Reynolds and Timmons, Tweeten and Martin, and Herdt and Cochrane. All of the models did a reasonable job of explaining the variations in land prices during the period for which they were originally estimated. To determine how well the models might perform now, they were reestimated utilizing more recent data. These models are briefly reviewed and the results of the reestimation are discussed below.

Reynolds and Timmons used a two-equation recursive model for identifying the principal determinants of agricultural land prices. They found that much of the variation in land prices could be explained by expected capital gains, predicted voluntary transfers of farmland, government payments for land diversion, conservation payments, farm enlargement, and the rate of return on common stock. When the model was reestimated with more recent data, however, there were a number of changes in the signs and magnitudes of the coefficients. In fact, for the price equation four of the nine signs reversed.

Tweeten and Martin presented a five-equation model for explaining changes in farm land values over time. They found that the two major determinants of farm real estate price increases between 1950 and 1963 were capitalized benefits from government programs tied to land and pressures for farm enlargement. A reestimation of the model for more recent years revealed that the parameters of the model had changed considerably. The results were similar to those of White, et al., who applied the model to Georgia data for 1960-1974. They found that in the price equation only one variable, net farm income, had a significant coefficient, and the signs of the coefficients for land in farms and number of transfers were unexpected.

The final simultaneous equation model considered is one presented by Herdt and Cochrane. They concluded that technological progress in conjunction with government supported output prices led to rising farmland prices. As their model was considerably more robust with respect to the time change than the other models, it will be discussed in greater detail. Their model is:

$$Ns = f(P, R, U, Lf) \quad \text{(supply equation)}$$

$$Nd = f(P, R, T, Pr/Pp, G) \quad \text{(demand equation)}$$

$$Ns = Nd \quad \text{(identity)}$$

where Ns is the number of farms supplied;

Nd is the number of farms demanded;

P is the average value per acre of U.S. agricultural real estate in current dollars;

R is the rate of return on nonfarm investment;

U is the unemployment rate;

Lf is amount of land in farms;

T is the USDA productivity index;

Pr/Pp is the ratio of the index of prices received by farmers to the index of prices paid by farmers;

G is the wholesale price index.

Herdt and Cochrane estimated this model for 1913-1962 using two-stage least squares (2SLS). The model has been reestimated for the post-war years 1946-1972. In addition, the model has been estimated for 1913-1972. Both 2SLS and three-stage least squares (3SLS) estimates for the two time periods, as well as the Herdt and Cochrane original estimates, are presented in Table 1.

There is one sign change in each of the four sets of new estimates. In two of the sets of estimates, the sign of the coefficient for number of transfers in the price equation reversed. This sign change is not particularly meaningful since not one of the estimates for this coefficient, including the original estimate of Herdt and Cochrane, is significantly different from zero at conventional levels of type I errors. Also, the sign of the coefficient for the wholesale price index changed. Of the four new estimates for this coefficient, two are positive, two are negative, and all are statistically insignificant. It is seen that the addition of ten more years of data makes it difficult to argue that this coefficient is nonzero. Because the Herdt and Cochrane model withstood the time change considerably better than the other models, it was selected as the representative simultaneous equation model to be used for forecasting comparisons.

In addition to the econometric models discussed above, a recent single equation model was tested which has less structural validity than the other

Table 1. ESTIMATION RESULTS FOR THE HERDT AND COCHRANE MODEL[a]

	Original estimates	New estimates			
	2SLS 1913-1962	2SLS 1913-1972	2SLS 1946-1972	3SLS 1913-1972	3SLS 1946-1972
Transfers Equation					
P	.064 (.119)	.244 (.213)	1.29 (.212)	1.106 (.326)	.239 (.182)
R	-5.672 (1.224)	-1.33 (3.95)	-19.42 (2.71)	-17.04 (4.18)	-1.363 (3.366)
U	-.789 (.188)	-.597 (1.10)	-.337 (.158)	-.54 (.23)	-.892 (.936)
Nf	.004 (.003)	.0001 (.000005)	.00004 (.000006)	.00003 (.00001)	.00001 (.0000...)
Price Equation					
Nd	-1.043 (.697)	[b].729 (1.78)	-1.17 (.517)	-1.36 (.633)	[b].550 (1.423)
R	8.315 (2.191)	16.38 (6.25)	18.94 (1.13)	19.90 (1.42)	17.166 (4.994)
T	1.699 (.321)	2.22 (.684)	2.35 (.321)	2.64 (.392)	2.296 (.550)
Pr/Pp	.757 (.372)	.3995 (.639)	1.00 (.357)	1.22 (.431)	.377 (.512)
G	.379 (.158)	.669 (1.22)	[b]-.0335 (1.164)	[b]-.213 (.200)	.417 (.978)

[a] Standard errors are shown in parentheses

[b] denotes sign change

models, but it fits the data well. Although the model is quite simple, the results have generated professional interest (Brake and Melichar). Klinefelter found that 97 percent of the variation in Illinois land prices between 1951 and 1970 could be explained by net returns, average farm size, number of transfers, and expected capital gains. A model similar to Klinefelter's was estimated for U.S. data for the periods 1946-1972 and 1913-1972.1/ The results are presented in Table 2.

For the 1913-1972 estimates, the coefficients for net farm income and average farm size have unexpected signs. The coefficients for average farm size and number of transfers are not significantly different from zero. When the model was estimated for 1946-1972, the expected signs for net farm income and average farm size were obtained, but the sign for number of transfers reversed. Despite several implausible signs, this model was utilized for forecasting due to the high percentage of price variation explained by the variable ($\overline{R}^2$ of .952 (1913-1972) and .989 (1946-1972)) and for the reasons mentioned above.2/

III. Forecast Results - Econometric Models

In order to forecast with the Herdt and Cochrane model, the reduced form equation for price was calculated and then the values of the exogenous variables were substituted to solve for price. Thus, the forecasts are expost in the sense that actual values of the exogenous variables are used. The results are presented in Table 3. On the basis of root mean square error (RMSE), the various versions of the Herdt and Cochrane model can be compared. It is apparent that for within sample forecasting, both sets of 2SLS estimates outperformed the 3SLS estimates. This is a rather surprising result since econometricians generally prefer 3SLS over 2SLS due to a presumption of the latter's lack of asymptotic efficiency. However, the better forecasting

Table 2. ESTIMATION RESULTS FOR THE MODIFIED KLINEFELTER MODEL[a]

Variables[b]	1913-1972	1946-1972
net farm income	-.0047 (.0008)	.0036 (.0013)
average farm size	-.0536 (.0786)	.5683 (.0805)
number of transfers	-.0250 (.0195)	.9526 (.1669)
expected capital gains	2.4099 (.0583)	.2203 (.3831)
GNP deflator	2.6843 (.0350)	1.1363 (.3066)
$\overline{R}^2$	.952	.989
Durbin-Watson statistic	2.581	.706

[a] Standard errors are given in parentheses

[b] The dependent variable is the average value of U.S. farm real estate per acre.

Table 3. ECONOMETRIC FORECASTS OF FARMLAND PRICES[a]

	1973	1974	1975	RMSE[b] within sample	RMSE beyond sample
Actual	238.14	297.80	340.48	--	--
Herdt Cochrane					
2SLS 1913-72	196.84	212.20	222.07	7.14	87.73
2SLS 1946-72	218.65	246.85	269.4	10.49	51.0
3SLS 1913-72	198.58	214.69	224.06	7.74	85.81
3SLS 1946-72	204.68	228.84	244.12	20.25	68.91
Modified Klinefelter					
1913-72	212.52	238.81	298.04	10.36	44.49
1946-72	224.44	257.72	284.88	4.73	40.35

[a] Forecasts of undeflated value of U.S. agricultural land and buildings per acre.

[b] Root mean square error.

performance of the 2SLS estimates may result from the fact that full information estimation methods, such as 3SLS, are more sensitive to specification error than are the k-class estimators such as 2SLS. Since 3SLS takes into account the correlation between the disturbances of all the structural equations, a specification error in one equation will affect all of the coefficient estimates of the system.

None of the Herdt and Cochrane reduced form equations forecasts well beyond the sample. For example, the actual undeflated value of farm real estate per acre was $340.48 in 1975. The highest forecast for that year was $269.40, and the lowest was $222.07. On the basis of RMSE beyond the sample, the 2SLS 1946-1972 estimates performed the best, followed by the 3SLS 1946-1972 estimates.

The forecast results for the modified Klinefelter model are also presented in Table 3. For within sample forecasts, the 1946-1972 estimates did better than any of the Herdt and Cochrane reduced forms. For both time periods the Klinefelter model forecast better beyond sample than each of the Herdt and Cochrane versions. It is apparent from these results that a simple model with implausible signs can still forecast quite well.

It should be noted that in the Klinefelter model expected capital gains (a 3 year moving average) includes lagged values of the dependent variable price. One is curious whether time series models based solely on the lag structure of the dependent variable plus more general error structures might possess as great a predictive power as the economic models. In the following section, Box-Jenkins forecasts are presented and later compared with the econometric forecasts. The Box-Jenkin's results are viewed as benchmark forecasts, since it is generally hoped that econometric models perform at least as well as naive statistical models.

IV. Forecast Results - Time Series Model

As an alternative to the econometric models, time series models of an integrated autoregressive-moving-average form are used to obtain forecasts of land prices (Box and Jenkins). These are statistical models of the form

$$Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \delta + U_t - \Theta_1 U_{t-1} - \cdots - \Theta_q U_{t-q}$$

where the Z's are observations generated by a stochastic process, the U's are independently distributed random variables with mean zero and constant variance, and δ, ϕ_i, and Θ_i are unknown parameters. The first part of the model is referred to as the autoregressive portion and the latter part as the moving average portion. If the observations are in difference form, then the process is called an integrated autoregressive-moving average process (ARIMA). Differencing of the data is often necessary in order to convert the process into a stationary one.

The first stage in selecting an appropriate time series model is to properly identify the process generating the observations. This is done by examining the estimated autocorrelation and partial autocorrelation functions. Box and Jenkins (pp. 176-77) provide tables describing the nature of the autocorrelation functions for various ARIMA processes. From the estimated autocorrelation and partial autocorrelation functions based on 1913-1972 observations, the model was identified as an ARI(2,2) or possibly a IMA(2,2), that is, an integrated autoregressive process of order 2,2 or an integrated moving average process of order 2,2. Based on these identifications, the models were estimated. The results are (1) for the ARI(2,2) model

$$W_t = \underset{(0.146)}{1.071}\, W_{t-1} - \underset{(0.153)}{0.215}\, W_{t-2} + U_t,$$

and (2) for the IMA(2,2) model

$$W_t = U_t + \underset{(0.065)}{1.104}\, U_{t-1} - \underset{(0.073)}{0.023}\, U_{t-2},$$

where W_t represents the second difference of the Z_t, and the values in parentheses are the standard errors of the estimates.

The forecasting performances of the above estimated models were examined by predicting land prices within and outside the sample. The within sample forecasts were based on one period ahead forecasts for the period 1913-1972, and the forecasts outside the sample were for the years 1973, 1974 and 1975. The results are presented in Table 4. For comparative purposes, results from a logarithmic model for the years 1913-1972 are also presented. Though substantially reducing the degrees of freedom, the post war years were also estimated separately because of the land price spiral during this period.[3/]

From the empirical results it can be seen that all of the models performed much better within than outside of the sampling period. However, the outside forecasts are one, two, and three period ahead forecasts, whereas the within sample forecasts are all one period ahead forecasts. Futhermore, it may appear that the forecasts obtained from the estimated model based on 1946-1972 data appear superior to those of other models. However, the estimated standard errors of the coefficients were high due to the relatively small number of observations used. The rule of thumb in estimating time-series models is that at least 50 observations are needed to adequately estimate a model.

The logarithmic model performs relatively well. On the basis of RMSE, it outperforms all other time series models including the post war model.

TABLE 4

Box-Jenkins Forecasts of Farmland Prices

	1973	1974	1975	RMSE[a/] within sample	RMSE beyond sample
Actual	238.14	297.8	340.48	--	--
IAR(2,2)	214.54	223.14	231.73	6.73	77.37
upper limits[b/]	227.02	240.17	253.39		
lower limits	202.06	206.10	210.06		
IMA(2,2)	218.05	225.74	233.42	6.25	75.41
upper limits	229.88	241.61	251.98		
lower limits	206.23	209.86	214.85		
ARIMA(2,2,2)[c/]	228.49	242.17	257.51	6.01	57.94
upper limits	264.08	313.55	362.12		
lower limits	197.69	187.04	183.12		
ARIMA(1,1,1)[d/]	243.20	249.95	251.13	7.38	58.59
upper limits	259.49	293.52	314.32		
lower limits	226.91	206.38	188.40		

a/ Root mean square error.

b/ Upper and lower limits for 95 percent confidence intervals.

c/ Model based on data in logarithmic form.

d/ Model based on 1946-1972 sample period.

The implication of the logarithmic model is that percentage changes (rather than the level of changes) have remained relatively stable through time.

V. A Brief Comparison of Econometric and Time-Series Forecasts

The simultaneous equation econometric model used in this study yielded forecasts about as accurate as the benchmark forecasts of the Box-Jenkins method for the post war years when land prices were rapidly escalating. For the longer time period (1913-1972), the time series models performed better than the simultaneous equation model on the basis of RMSE, both within and beyond sample. For this same period the Klinefelter model had the lowest beyond sample RMSE. Further, the Klinefelter model performed better than either time series or the simultaneous equation econometric model for the post war years. Overall, the poorest predictors appear to be generated by the simultaneous equation models.

VI. Conclusions

It is not uncommon when comparing time series and econometric forecasts to discover that time series models provide as good or better short term forecasts than econometric models.[4/] The above results are suggestive of this conclusion. However, one notes that the single equation model predicted well, and the case may be made that this model generated the best predictors. This result is surprising - particularly since the model appears to be very sensitive to the sample period used to obtain parameter estimates. Also, although the model may have microeconomic foundations, as a market model it explains little. Since expected capital gains are functionally related to lagged land prices, it appears that more effort must be devoted to explaining, in a causal sense, the recent rise in farm prices and capital gains.

The simultaneous equation models presumably possess greater causal foundations reflecting market behavior of sellers and buyers. Yet, it appears that attempts to incorporate greater structural detail in the econometric models of the land market have not enhanced the forecasting ability of such models. One would expect the magnitude of parameter estimates to be sensitive to the sample period. However, note that for all of the simultaneous equation models considered here, there was an abundance of unexpected sign changes when recent data were added to the sample. This is suggestive of more than a mere structural change in the market, but reflects the need for further model development of the land market in order to determine the salient causal mechanisms. Therefore, if one is concerned with both predictive ability and econometric structure, it would seem on the basis of the empirical performance of earlier econometric models that new research is needed to explain recent movements of farmland prices.

mh 5/22/73

FOOTNOTES

1/ The model used differed from Klinefelter's model as follows: (a) net farm income was used in place of net returns to landlords; (b) instead of deflating variables, the GNP deflator was entered as an explicit variable; (c) in the calculation of capital gains, capital improvements were subtracted out.

2/ Of course a high $\bar{R}^2$ does not necessarily imply that a model will forecast well. It is also noted that the Durbin-Watson statistic for the period 1946-1972 suggests evidence of positive autocorrelation. However, since expected capital gains contains tranformations of lagged values of the dependent variable, the Durbin-Watson may not be appropriate. If autocorrelation is present, parameters estimates will be inconsistent.

3/ Diagnostic checks were made on the above models to test for their adequacy based on data used. This was done by overfitting the models by adding an extra autoregressive or moving average term and checking the significance of the estimated coefficient. These results are not reported.

4/ For an interesting discussion of the relative merits of Box-Jenkins versus econometric models see Naylor.

References

Box, G. E. P. and G. M. Jenkins. Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, 1970.

Brake, J. R. and E. Melidiar. "Agricultural Finance and Capital Markets," in L. R. Martin, ed., A Survey of Agricultural Economics Literature, Volume 1, University of Minnesota Press, Minneapolis, 1977.

Crowley, W. "Is Now a Good Time to Buy Land?" Paper presented, Michigan Farmers Week, Michigan State University East Lansing, Michigan, March 25, 1976.

Gardner, B. D. "Issues Affecting the Availability and Price of Land for Agriculture." Paper presented, California Chapter Meeting of the American Society of Agronomy, Fresno, January 26, 1978.

Harris, D. G. "Inflation-Indexed Price Supports and Land Values." American Journal of Agricultural Economics 59(1977):489-95.

Herdt, R. W. and W. W. Cochrane. "Farmland Prices and Technological Advance." Journal of Farm Economics 48(1966):243-63.

Jundt, D. "The Land Price Rise, Will It Continue?" Doannes Agricultural Report, January 8, 1965.

Klinefelter, D. A. "Factors Affecting Farmland Values in Illinois." Illinois Agricultural Economics 13(1973):27-33.

Naylor, T. H., T. G. Seaks, and D. W. Wichern. "Box-Jenkins Methods: An Alternative to Econometric Models." International Statistical Review 40(1972):123-37.

Reynolds, T. E. and J. F. Timmons. Factors Affecting Farmland Values in the United States. Iowa Agricultural Experiment Station Research Bulletin 566, 1969.

Tweeten, L. G. and J. E. Martin. "A Methodology for Predicting U.S. Farm Real Estate Price Variation." Journal of Farm Economics 48(1966):378-93.

U.S. News and World Report. "An Endless Boom in Farmland Prices." April 29, 1974, p. 49.

White, F. C., W. N. Musser, and A. R. Sheffield. "Some Evidence on the Declining Effect of Farm Consolidation on Farm Real Estate Prices." Southern Journal of Agricultural Economics 9(1977):99-104.