THE AUSTRALIAN JOURNAL OF AGRICULTURAL ECONOMICS

VOL. 12 JUNE, 1968 NO. 1

A MODEL OF THE NEW ZEALAND SHEEP INDUSTRY*

A. C. RAYNER

University of Canterbury

The paper presents a model of the New Zealand Sheep Industry which predicts animal numbers in various sex/age categories. The explanatory variables used are the prices of the end products of the industry and time, to represent technological change. Prices are found to have a significant, though delayed, effect on farmers' stock decisions. Furthermore, the significance of time demonstrates improvements in some forms of animal husbandry. The relative failure of the model's latest predictions emphasizes the importance of irrational optimism in the industry.

Introduction

The purpose of this paper is to present a model of the New Zealand Sheep Industry which is capable of predicting the numbers of animals in various categories. These categories are based on a sex and age split, rather than on breed and geographical area. In order to make the predictions, a system of ten equations, describing various relations in the industry, are estimated. The explanatory variables used in the equations are the various prices of the end products of the industry, and time, to represent technological progress. After presenting the equations, the paper concludes by comparing the predictions of the model with the latest data available, re-estimating the equations and giving a set of predictions for the following year.

The estimated equations show the influence of prices on farmers' decisions. Of particular note is the lag involved between price variations and the implementation of the resultant changes in stock numbers. It is also of interest to see how price relatives affect the numbers of sheep in different categories. Thus the number of wethers depends on the wool/lamb price ratio while the number of wether hoggets on the mutton/lamb price ratio. An encouraging sign is the significant improvement over time of the lamb death and tailing ratios and the ram/ewe ratio, all suggesting improvements in animal husbandry.

The actual comparison of the model's predictions with reality is not very satisfactory, but does emphasize the importance of irrational price optimism, or government exhortation, in the industry.

* The research on which this paper is based was financed by a grant from the Vernon Willey Trust.

Data and Definitions

The study is based on the period of about fourteen years between the atypical, high prices of the Korean war and the recent very low wool price. The homogeneity of the period allows one to expect some success in estimating behaviouristic relations, but throws doubt on the predictive ability of the model when confronted with the present market situation. Even if one thought that the experience of a much longer and more varied period would improve prediction, in spite of probable changes in the structure of the industry over time, the lack of data would make it impossible to estimate the model used here. The source used for the stocks of sheep is the Annual Sheep Returns [1]. This gives data for the stocks of sheep in various categories at a certain date each year. This date has been June 30th since 1951 and was April 30th before that year. To correct for this change in date would be difficult, particularly in the case of hoggets (lambs at 30th June, i.e. immature sheep). Moreover, before 1952 the form of disaggregation of the data makes it very difficult to obtain comparable figures.

The source of data for lambs tailed and lambs and sheep slaughtered is the Statistics of Farm Production [2]. The first of these three series has year end January 31st, but because of the seasonality of lamb births it is clear that the lambs are related to the numbers of ewes the preceding June 30th and to the numbers of hoggets the following June 30th. In the case of numbers slaughtered, the year end is September 30th. To make this comparable with the Sheep Return statistics it was necessary to adjust to the different year end, using the series from the Monthly Abstract of Statistics [3], taking care to use the same definition of hogget. Although these series are not identically defined, the difference is not large and should not lead to any consistent bias. It is assumed throughout this paper that net imports and exports of live animals are small enough to be ignored. The annual average price for all qualities of wool is taken from the Annual Review of the Sheep Industry [4]. For the meat price series I have to thank Mr J. W. B. Guise, formerly of this University and now of the University of New England, Armidale, N.S.W., for the use of part of some data extracted by him. The prices are taken from the South Island Freezing Works Schedule [5] for the last week in February each year, this being considered the most typical slaughtering period. While an average price might be preferable, considerable labour would be involved in the calculation. Moreover, since it is not known what actual price the farmer himself is concerned with in his decision making process, the extent of the error cannot be evaluated. The use of the wrong price as an exogenous variable may lead to bias in our estimates, but this is hard to overcome, not knowing the 'correct' price and is not so serious when the same price is used for prediction.

The following are the symbols used in this study:

Stocks at 30th June in year t

Mature animals, i.e. those almost two years old and over, more technically called two-tooth and over.

$E^b{}_t$ Breeding ewes
$E^{dr}{}_t$ Dry ewes (i.e. non-breeding ewes)
E_t Total ewes

W_t Wethers (i.e. emasculated male sheep)
R_t Rams

Immature animals, i.e. those just under one year old, hoggets.

$H^e{}_t$ Ewe hoggets
$H^w{}_t$ Wether hoggets
$H^r{}_t$ Ram hoggets
Total hoggets is written as H_t
Total sheep is written as S_t

Annual flow variables

$L^t{}_t$ Lambs tailed in the year ended 30th June year t, i.e. those born and surviving for a few weeks, until they are tailed.
$L^{t\%}{}_t$ The tailing rate in year ended 30th June year $t = L^t{}_t/E^b{}_{t-1}$.
$L^{sl}{}_t$ Lambs slaughtered during the year.
$S^{sl}{}_t$ Sheep slaughtered during the same period. This includes all mature and immature animals other than lambs.

The addition of a per cent superscript to either of the last two variables is used to mean that the variable is divided by lambs tailed during the period in the first case, or by the stock of sheep at the beginning of the period in the second. The slaughter variables and rates for other series are similarly defined, i.e. $W^{sl\%}{}_t$ or $E^{sl}{}_t$.

A superscript of d instead of sl in these variables indicates the natural deaths of the animals or group of animals, i.e. $E^{d\%}{}_t$.

Prices

$P^w{}_t$ Average price for the season t for all qualities of wool in pence per lb. greasy.
$P^l{}_t$ Price of lamb in pence per lb. for 29/36 lb. animals. This is the South Island Freezing Works price for animals in wool for the last week in February, year t.
$P^m{}_t$ Similar price for ewes of 49/56 lb., wool not included.

In general, estimated relationships are assumed to contain a random term having all the desirable properties required for the use of ordinary least squares. The use of $'$ shows an estimated series.

It should be noted that no variable outside the industry, other than time, has been included in this list. In fact no account is taken of any competing types of farming in this paper. This is partly justified by the fact that there is only competition, given the ranges of prices experienced in this period, over a marginal section of the industry. A further justification rests on the degree of success obtained by confining attention to the sheep industry alone.

The Model and its Estimation

While it is the aim of this paper to predict the majority of the quantity variables introduced in the previous section, it should be noted that these are not all independent, and further, that they can be subdivided into those of major decision significance and those of a more residual nature. Thus, there are some identities connecting the variables, notably over time, some mainly technical relationships, as those between lambs tailed and breeding ewes, some minor relationships, as between dry and breeding ewes and some of major decision importance. The main decision variable amongst mature sheep is clearly the number of ewes E_t, since they directly produce all the final products, lamb, mutton and wool, and also indirectly the investment good, hoggets. The second variable in order of importance is the number of wethers, W_t. These are subsidiary, since they produce only mutton and wool. Rams are not directly productive, apart from their wool.

E_t, W_t and S^{sl}_t are not independent, as we shall see and it is on these three variables that we shall start this investigation. First, it should be noted that for most purposes the figure for total ewes, rather than breeding ewes, is taken as the decision variable, on the grounds that the number of dry ewes is a residual largely beyond the control of the farmer, in the sense that it is dependent on such factors as weather. It is assumed that the farmer, in making decisions about flock size, is concerned with the *total* number of ewes.

On inspection of the data, it is immediately apparent that, while E_t continuously increased throughout the period, W_t fluctuated much more, and in fact showed a slight decrease. Moreover, the ratio of E_t to H^e_t was of the order of four to one, while W_t and H^w_t were both of the same order of magnitude. This implies that the number of wethers could be changed much more dramatically than ewes, at least in the upward direction. Therefore it seems probable that we can expect these variables to exhibit different behaviour patterns.

Firstly, let us consider E_t. Regressing this on time yielded an R^2 of 0·9797, and on its own past value:

(1A) $E'_t = 0{\cdot}9780 E_{t-1} + 1{,}762{,}000, \quad R^2 = 0{\cdot}9895 \quad d = 1{\cdot}68.$

These high values of R^2 bear out the casual impression of the smoothness of the growth of ewe numbers, but they do raise problems of economic interpretation. The continued growth suggests continuing profitability which allows either substitution of sheep for other farming, or an increased intensification of sheep farming by top-dressing and other means. As will be discussed shortly, the ewe stock can be altered in two ways: either through varying the number of ewe hoggets, i.e. changing the investment rate, or by holding back mature ewes from slaughter for another year or so, thus altering the life of the capital. The former represents the long-term stock adjustment, the latter more of an immediate adjustment to profitability stimuli. The main variables affecting profits will be prices of final products, costs, technical progress and government stimulation. The government acts either directly by altering such things as tax incentives, or indirectly by changing farmers' expectations and outlook.

Over the period covered, prices of final products did not rise significantly, while costs did and so the continuing profitability must have been a result of technical progress and government intervention. Any sudden changes in these latter or fluctuations in the prices of the products will be smoothed by the time lags involved in obtaining breeding ewes, and in new techniques spreading over the country. Therefore in the long term, growth *via* investment in increased numbers of hoggets will be a smooth process, provided the profitability caused by technical progress and government intervention continues.

Although we could try to measure these profitability factors directly, the use of the lagged value of ewes, as in equation (1A), adequately acts as a summary of their past history and explains a very high proportion of the variance in the ewe series. If we were able to measure some true profitability variable, this could be extended to a full Koyck distributed lag model [6]. However, this would increase the econometric problems already present in equation (1A). Although the lack of serial correlation as shown by the value of d, the Durbin Watson statistic, suggests the assumption that E_{t-1} and the error term are independent, the use of the lagged endogenous variable still leads to small sample bias in the estimate, which could be as high as 20% for samples of the size used here. Moreover, since we are estimating an autoregressive equation, the value of d obtained is itself biased and does not in fact preclude the presence of serial correlation and an even greater bias in the estimate.

The danger of this bias and the smoothness of the growth suggest that it would be better to look at year to year changes in numbers and to assume that the underlying profitability, caused by such things as technical progress, will continue to lead to the steady growth observed over the period. Year to year movements around the growth trend show varying slaughtering policy reacting to variations in product prices, irrespective of the overall underlying profitability, which mainly affects the continued investment in new stock. Government policy, particularly as it affects farmers' expectations, can also cause year to year changes which are difficult to predict.

It was found, and this result will be given in detail later, that the sheep slaughter rate was best explained by the lagged value of the sum of the prices of wool and lamb. The rationale for this is that, as suggested earlier, an unusually high price will cause farmers to hold back animals from slaughter for an extra year. Therefore, regressing the absolute change figures on lagged prices we obtain:

$$\text{(1B)} \quad (E_t - E_{t-1})' = 39{,}739(P^l + P^w)_{t-1} - 1{,}465{,}500$$
$$R^2 = 0{\cdot}5320 \quad d = 2{\cdot}69$$

From this equation we can obtain an estimated series for E_t and from this estimated series the residuals about the correct figures can be found. The residual, or unexplained, variance can thus be derived and therefore also the explained variance. The ratio of this latter to the variance of the E_t series gives a quasi R^2, which is 0·9948 in this case. This equation is statistically and economically more appealing than the original ewe equation, (1A), and furthermore, increases the value of R^2 from 0·9895 to 0·9948, thus halving the unexplained variance. It is therefore the final ewe equation used in the model.

The second main decision variable to be considered is the number of wethers. While ewes exhibited steady growth resulting from continued profitability, wethers in contrast showed something of a decline in numbers and greater variation around this trend. Thus the regression of wethers on their own past value yielded an R^2 of only 0·3702, compared with that for ewes of 0·9895. The implication of this is that either there has been a price trend against the outputs of wethers compared with ewes, or that the technical progress has favoured ewes. The technological trend can best be represented by the direct use of time as an exogenous variable. There is obviously no case for having $(P^w + P^l)$ as a second variable, since wethers do not produce lambs. As an alternative the price of wool by itself, or its price relative to that of lamb would seem more reasonable. The latter proved more satisfactory[1]:

$$(2) \quad (W_t)' = 2{,}008{,}000 + \underset{(100{,}200)}{294{,}200}\, P^{w}_{t-1}/P^{l}_{t-1} - \underset{(8{,}640)}{37{,}740}t$$

$$R^2 = 0{\cdot}7068 \quad d = 1{\cdot}84$$

($t = 0$ in 1952. The price ratio is significant at the 2% level and time at 1%.)

In order to complete the investigation of adult sheep numbers, rams and dry ewes must be considered. Rams can be thought of as being linked to ewes by a technical production function. Moreover, they themselves have no end product but wool and dog meat. This implies that farmers can be expected to minimize the stock of rams that they hold to serve a fixed number of ewes. It presumably is possible to change the number slightly as a result of such things as better fencing of land, and so, though the ratio of rams to ewes should remain fairly constant, it may also show a decreasing trend over time.

A further fact known about rams is that their production is more specialized than other branches of the industry. This has certain implications:

(a) Since the ram breeder is trying to predict the market a few years in advance, there will be some fluctuations in the ram/ewe ratio, assuming it is not always corrected by slaughtering adult rams to bring the ratio back to the ideal level.

(b) An increase in the number of ewes that was not expected by the breeders will cause a decrease in the ram/ewe ratio and a negative residual about a time trend. Given that there is a lag in breeding the necessary increase in rams, the shortage is liable to continue for at least another season, even with only average further ewe growth, and therefore the negative residual will continue. In other words, some positive autocorrelation of the residuals can be expected.

(c) It is not clear whether the ratio of rams to ewes in the same year, or of rams in year t to ewes in $t-1$ should be considered. The farmer, who decides on the slaughter of rams, will presumably be concerned with the former, but the breeder may look at the number of ewes in a year to decide how many rams to breed for the future, and thus affect the lagged form of the ratio. Whether the actual ratio results more from the slaughter of rams, or their breeding, will decide which is the correct form of the ratio to use.

[1] Although the use of the ratio of wethers to ewes as the endogenous variable led to a higher R^2, this was caused by the growth of E_t making the time variable more significant. A derived wether series from the ratio regression did not fit the data as well as equation (2). This use of the ratio raises the question of the legitimacy of having independent equations for ewes and wethers, even though they must be to some extent competitors. There is especial doubt when a price ratio is introduced significantly in the absolute wethers equation. The only economic justification rests on the smallness of relative numbers of wethers and thus their insignificant effect on ewes, though not necessarily *vice versa*. Also the difference in behaviour of the two series suggests that they cannot be closely related. Another empirical argument rests on the fact that, though some time was spent in attempting to deal with the two series together, the separate treatment always proved superior.

The simple regression of current values on time yields the following:

(3) $(R_t/E_t)' = 0{\cdot}0026829 - 0{\cdot}0001307\, t$

$R^2 = 0{\cdot}8142 \quad d = 1{\cdot}35$

($t = 0$ in 1952.)

This shows the expected small negative trend in the ratio, and also some suggestion of positive autocorrelation as discussed in note (b) above, though this is not significant. Deriving a series for rams from this, we find the proportion of the variance explained to be 0·9934. Regressing the number of rams on the number of ewes directly yields an R^2 of 0·9907.[2]

A second exogenous variable, $(E_t - E_{t-1})/E_{t-1}$, was introduced to allow for the lag in adjustment to changes in ewe stocks. However, though its estimated coefficient had the expected sign, its level of significance was too small to allow it to be accepted as being different from zero.

A problem faced here will be returned to several times. This is that in order to predict the ram series we need to know the ewe series. In other words, the series E'_t predicted from the final ewe model must be used in order to tell how well the ram ratio equation will predict, unconditionally, the ram series. This lowers the proportion of variance explained to 0·9906. It was noted in point (c) above that it was not certain whether a lagged value of E should be used in the ram ratio. Also, the regression of R_t/E_{t-1} on time overcomes the problem of having to use predicted values for E_t. However, estimation with this ratio led to a slightly lower R^2 than the one above and it was therefore not accepted.

Each year, a small number of ewes at 30th June have not gone to ram and are thus classified as dry ewes, E^{dr}_t. Although this series is of small absolute number and of little direct significance in the predictions, it is still needed to obtain the number of breeding ewes from the total prediction $(E_t)'$ and to complete the estimation of stocks of mature sheep.

Dry ewes are those considered not fit to breed and thus their numbers will be influenced by the weather and by the general level of animal husbandry. For this latter reason we can expect a decreasing proportion of dry ewes to breeding ewes. In fact a regression of the ratio of these two variables on time yields an R^2 of 0·7631. The significance of this regression stems largely from the close relationship between $1/E^b_t$ and time. To avoid this effect the absolute number of dry ewes was considered. Since there is not such a clear expectation, as there is in the case of rams, that there is a technical relationship between their numbers and those of breeding ewes, this is a justifiable procedure.

Although the number of dry ewes has not increased much, there is some indication of a time trend, so that time was the first choice as an exogenous variable. A second one is the price of lamb. It seems a

[2] These two results are similar for the following reason: If $R'_t = a + bE_t$ then, dividing by E_t, we get $R'_t/E_t = a\,1/E_t + b \simeq f(t)$, since E_t is closely related to time, and so $1/E_t$ will also be closely related to time over the range of values of E_t covered during the period, though this will be a negative relation. Thus, in fact, a and b estimated by the first non-ratio equation both have positive signs and so for the regression on time, we would expect, as found, that the estimated coefficient of the time variable has a negative sign. The ratio regression does, however, make more intuitive economic sense and therefore is the one accepted.

reasonable hypothesis that, other things being equal, the higher the lamb price, the greater the number of ewes that are sent to ram. Even if their general health involves some risk in breeding, this risk is partly overcome by the higher reward expected. This regression yields the following result:

$$(4) \quad (E^{dr}_t)' = 559{,}150 - \underset{(2{,}343)}{5{,}546}\, P^l_t + \underset{(1{,}765)}{4{,}220}\, t$$

$$R^2 = 0{\cdot}5795 \quad d = 1{\cdot}57$$

($t = 0$ in 1952. Both coefficients are significant at the 5% level.)

Considering the probable significance of weather, this seems a satisfactory result. For comparison, a regression of dry ewes on their own past value gave an R^2 of 0·4104.

It should be noted that in this case the price variable is not lagged. Therefore we can only use this equation to predict a few months in advance. To avoid this difficulty price would have to be dropped from the equation. The increased error that this would involve would not have much effect on the rest of the model, due to the smallness in absolute size of the dry ewe series.

After considering the numbers of adult sheep, the question of how many of them are slaughtered each year will now be examined. This is the first of the two final products of the industry to be discussed. It is also the least important, both in terms of numbers and value. Unless there is some very dramatic change in the relative price of mutton and lamb, it seems clear that the number of sheep slaughtered will depend mainly on when they are considered to have reached the end of their productive life and not on their value as meat. It was suggested earlier that the slaughter rate was used to vary stock numbers from year to year as an adjustment to short run price movements. Thus, a high price of wool and lamb will make farmers want to increase their flocks, at least in the short run and so will cut down the numbers slaughtered.

The problem remains of deciding on the lag involved and on the weights to be attached to the two prices. It was found that merely adding the two prices together and lagging by one year led to the best results. Attempts were made using different lags and different weights, but they did not prove significantly superior. Although it might seem strange to suggest that farmers ignore current prices in deciding on the number of sheep to be killed, this does appear to be the case. A possible explanation of this lag is that the season is coming to an end before the current price trend becomes clear, by which time most of the slaughtering decisions for that season will have been made. The price of mutton was not used, partly because it did not appear to be theoretically or actually of value as an exogenous variable, and partly because it would involve problems of simultaneity if its current value was used. The final equation was:

$$(5) \quad (S^{sl\%}_t)' = 0{\cdot}24095 - 0{\cdot}001359\ (P^w + P^l)_{t-1}$$

$$R^2 = 0{\cdot}6559 \quad d = 2{\cdot}19$$

From this it is possible to derive a series for sheep slaughtered and compare this to the actual figures. The derived $R^2 = 0{\cdot}9158$. Regressing the absolute number slaughtered on total sheep and the price variable was not so satisfactory.[3]

Proceeding from mature sheep, lambs and specifically the prediction of $L^t{}_t$ follows. This must clearly bear a very close technical relationship to $E^b{}_{t-1}$, a relationship that will vary only due to weather factors affecting the health of the ewes and gradual change caused by improvements in husbandry. The simple regression of $L^t{}_t$ on $E^b{}_{t-1}$ gives an R^2 of 0·9925. As noted earlier, when discussing the ram/ewe ratio, regressing in this way is similar to regressing $L^t{}_t/E^b{}_{t-1}$ on time. This latter makes greater interpretative sense, since it allows a direct demonstration of the improving tailing ratio. Thus, we get:

$$(6) \quad (L^t{}_t/E^b{}_{t-1})' = (L^{t\%}{}_t)' = 0{\cdot}96082 + 0{\cdot}002530\, t$$
$$R^2 = 0{\cdot}3120 \quad d = 1{\cdot}28$$

(R^2 is significant at the 5% level, though the positive autocorrelation throws some doubt on this significance.)

A derived L^t series, using actual E^b values, leads to a quasi R^2 of 0·9928, which is close to the value above, as expected.[4]

Lambs which have been tailed in any season have five possible destinations; they can die a natural death, be slaughtered, or be alive at the end of the season as ewe, wether or ram hoggets. While it is possible to investigate each of these series separately, this is open to the likelihood of the five independent estimates not being consistent with each other. They may add up to more, or less, than the number of lambs tailed, thus producing a problem similar to that observed in the case of mature sheep.

Lamb deaths are simplest to deal with. They cannot be considered a decision variable, other than in the sense of the farmer causing a fall in the death rate over time by improved methods of husbandry. Apart from this, weather is likely to be the most significant variable, both by directly killing the lambs and indirectly by affecting their food supply. Unfortunately, it proved too difficult in the aggregate to measure this

[3] Predicting both stocks at the end of the season and numbers slaughtered during the season can lead to a certain degree of inconsistency. This is because these are related to the stocks at the beginning of the season and to deaths during the season, by an identity. Admittedly we are not predicting the deaths series, but there are clear restrictions on the fluctuations of this series, which we can observe and which are not sufficient to cover the inconsistencies seen in our predictions. Although there is no clear time trend in the case of sheep deaths, we can assume a constant average death rate, and thus indirectly predict a second slaughter series via predicted end of season stocks and the number expected to die naturally. However, in fact this did not prove to be as effective in terms of variance explained as the direct $S^{sl\%}$ regression. Therefore we are left with a conflict between a desire to explain as much of the variance as possible, i.e. to choose the best estimator, and a desire not to produce two alternative predictions, one explicit, the other implicit, which may be inconsistent with each other. In this case, since the $S^{d\%}$ cannot be predicted except as a constant, it was decided to choose the best estimation procedure and to ignore the possibility of implicitly predicting nonsense figures for S^d. As we shall see later, L^{s1} did not cause this conflict.

[4] In order to predict L^t unconditionally we need to use the predicted values of total ewes and dry ewes and of the $L^{t\%}$. This, then, gives us a ratio of explained to total variance of 0·9785, the value being pulled down considerably by the coincidence of the greatest overprediction of $L^{t\%}$ and E^b in the same year. For comparison, the regression of L^t on its own past value gives an R^2 of 0·9632. There is normally a considerable lag between the publication of statistics on stocks for year t and flows for year $t+1$, the latter appearing much later. Therefore in practice there is little value in an unconditional prediction of L^t.

weather factor in some way that can be entered into the regression as a significant variable. A regression of $L^{d\%}$ on time gives the following result:

$$(7)\quad (L^{d\%}{}_t)' = 0{\cdot}05080 - 0{\cdot}001736\,t$$
$$R^2 = 0{\cdot}4517 \quad d = 2{\cdot}60$$
$$(t = 0 \text{ in } 1952.)$$

This has the expected sign and a reasonable degree of significance, considering the probable impact of weather as a random variable.

The surviving lambs are split between those slaughtered and those that become hoggets. It is possible to consider first this split and then to estimate the way that total hoggets are distributed between the three classes. However, this did not produce such satisfactory results as when the three hogget series were considered independently and lambs slaughtered was treated as a residual.

In the case of ewe hoggets, the best exogenous variable proved to be the same as that used in the adult ewe model. In fact the best results were obtained by using the identical form:

$$(8)\quad (H^e{}_t - H^e{}_{t-1})' = 20{,}240(P^w + P^l)_{t-1} - 1{,}016{,}000$$
$$R^2 = 0{\cdot}4185 \quad d = 1{\cdot}87$$

This led to a derived ewe hogget series with a proportion of variance explained of 0·9670. Other forms of price, including the current lamb price, or other variables giving some indication of the availability of lambs, did not prove significant. This is consistent with the idea that the decision taken on ewe hogget numbers is of prime importance, while lambs slaughtered is a residual.

The mature wether equation used time and the wool/lamb price ratio as exogenous variables. However, it is clear that these are not satisfactory in the case of hoggets, firstly because there has not been any trend over time and secondly because a large proportion of wether hoggets are not held to become mature wethers. This latter point can be seen from the fact mentioned earlier that the number of wether hoggets is of similar size to the number of mature wethers, with the implication that there is a high slaughter rate for hoggets.[5] In fact the wool/lamb price ratio is not significant in this case, while the mutton/lamb price is. The mutton price is taken to be a close approximation to the hogget price, so that the higher this ratio is, the more worthwhile it is to keep the wether lambs a few months longer and to slaughter them as hoggets rather than lambs.

The use of this ratio involves the current price instead of the lagged one, so that again, as in the case of dry ewes, prediction can only be made a short time ahead. However, in this case there is a further problem, that of simultaneity, since the number of wether lambs supplied to the market may affect the current price of lamb. In earlier equations it was always possible to assume that price was a predetermined variable, either because of a time lag, or because it was not the price of the dependent variable.

Simultaneity can lead to many econometric problems, including the possibility of bias in the estimates. It can only be overcome by building

[5] A second possibility, of a short life for the mature animal, does not seem to be borne out by the data.

demand equations into the model, which would seem undesirable just to handle this one equation. Luckily it seems likely that the problem may not be too serious in this case for the following reasons:

(a) the actual variation in the number of wether hoggets from year to year represents a small proportion of the total number of lambs slaughtered and thus will presumably have a small effect on price, and

(b) the prices used in the regression are for the last week in February, while we can assume that a certain proportion of wether lambs held back from slaughter are held back after this date and so have no effect on the price.

For these reasons, it was considered reasonable to continue using the equation as it stands.

As in the case of ewe hoggets, a first difference form proved better for prediction:

$$(9) \quad (H^w{}_t - H^w{}_{t-1})' = 2{,}716{,}000\, P^m{}_t/P^l{}_t - 790{,}300$$
$$R^2 = 0{\cdot}6621 \quad d = 2{\cdot}26$$

The derived wether hogget series had a ratio of variance explained of 0·4111.[6]

Ram hoggets were treated in the same way as mature rams; their ratio to ewe hoggets being investigated. Here a strange phenomenon was discovered, which was that, for the first seven years the ratio increased linearly, and for the second seven it decreased with an even closer fit to a straight line. The two R^2 values were 0·8697 and 0·9834. Ram hoggets are bred, to a large extent, by specialists and thus it would seem likely that this behaviour can be explained by the lagged response of the breeders to price—a situation similar to that of the so-called hog-cycle. If series of ram prices were available, this hypothesis could be directly investigated, or if data for a longer period were used, the cycle's length could be estimated. As it is, we can only assume that the current downswing will continue. At least this is consistent with the decrease in the ratio for mature animals. Therefore the equation for the last seven years is:

$$(10) \quad (H^r{}_t/H^e{}_t)' = 0{\cdot}03446 - 0{\cdot}0009073\, t$$
$$R^2 = 0{\cdot}9834$$
$$(t = 0 \text{ in } 1959.)$$

Using this equation, and the predicted value of $H^e{}_t$, a series for $H^r{}_t$ can be obtained, thus completing the hogget coverage.

Using our predicted $L^{t\%}{}_t$ and actual $E^b{}_t$ we can predict $L^t{}_t$. We can also predict $L^d{}_t$, and H_t, leaving $L^{sl}{}_t$ as a residual. This derived series for lambs slaughtered has an R^2 of 0·9825 with the actual data. Attempts to predict this series directly did not prove as successful and so this indirect method was adopted. This then avoided the conflict between consistency and best estimation noted in the case of the sheep slaughtered equation.

[6] This value is less than the R^2 of the first difference regression since the absolute number series had a smaller variance than the change series. This was the result of a combination of no trend in the wether hogget series with some small degree of negative autocorrelation over time.

Prediction and the Re-estimated Model

The endogenous variables of the ten equations making up the model of the industry are not all in terms of absolute sheep numbers. Therefore, to compare the predictions with the actual figures, the absolute numbers must be derived from the model. The initial predictions shown here are:

(a) stocks of animals as at 30th June 1966, and
(b) industry flows for the season 1964/65.

The delay in the publication of the Statistics of Farm Production [2] means that the variables in group (b) cannot be predicted and tested for the most recent (1965/6) season. However, after re-estimation using the 1964/5 data, further predictions are made, for both the 1965/6 and 1966/7 seasons.

The predictions from this model will be compared with those from two simple models:

(I) assuming there is no change from the previous year for the variable being considered, and
(II) assuming the absolute change from year $t - 1$ is the same as from $t - 2$ to $t - 1$.[7]

TABLE 1

Comparison of Predictions

Predicted Variable	Actual Value	Prediction of Full Model	Prediction of Simple Model I	Prediction of Simple Model II
		(a) (b)	(a)	(a)
E_{66}	40,003,500	38,759,400 (−3·11)	37,656,200 (−5·87)	39,142,300 (−2·15)
W_{66}	2,600,700	2,030,600 (−21·92)	2,553,300 (−1·82)	2,755,100 (+5·94)
R_{66}	956,500	968,900 (+1·30)	931,300 (−2·63)	955,900 (−0·06)
E^{dr}_{66}	335,800	499,000 (+48·60)	478,100 (+42·38)	488,000 (+45·33)
H^{e}_{66}	10,353,200	9,647,700 (−6·81)	9,355,400 (−9·64)	9,778,300 (−5·55)
H^{w}_{66}	3,138,600	3,265,400 (+4·04)	2,982,100 (−4·99)	3,301,400 (+5·19)
H^{r}_{66}	290,700	271,200 (−6·71)	269,483 (−7·30)	270,600 (−6·91)
$L^{t}_{64/5}$	34,791,600	35,477,400 (+1·97)	34,751,800 (−0·12)	35,807,800 (+2·92)
$L^{d}_{64/5}$	733,000	940,000 (+28·24)	948,000 (+29·33)	732,000 (−0·14)
$L^{st}_{64/5}$	21,451,000	21,798,000 (+1·62)	21,940,000 (+2·28)	22,874,000 (+6·63)
$S^{st}_{64/5}$	7,776,000	7,196,000 (−7·46)	8,133,000 (+4·59)	8,411,000 (+8·17)

(a) The percentage error is shown in parenthesis below each predicted value.
(b) The predicted absolute values were determined using the ten equations of the model, plus the identity relating lambs slaughtered to the other relevant variables. Where endogenous variables from another part of the model were required to obtain a prediction, their predicted value was used if they were of the same season or date, or dated at the end of the season. Otherwise their actual value was used.

[7] As stated in the previous section, equal absolute change seemed to be more the pattern in the industry than equal percentage change.

Before discussing these rather unsatisfactory results, the re-estimated equations, using the extra year's data, should be compared with those given in the previous section. The level of significance is given unless it is 1% or better.

(1B) $(E_t - E_{t-1})' = 39{,}999(P^l + P^w)_{t-1} - 1{,}393{,}400$
$R^2 = 0{\cdot}3430 \quad d = 1{\cdot}81$

(R^2 is significant at the 5% level.)

(2) $(W_t)' = 2{,}264{,}800 + 164{,}677\, P^w_{t-1}/P^l_{t-1} - 24{,}824\, t$
$\qquad (131{,}000) \qquad (10{,}990)$
$R^2 = 0{\cdot}3843 \quad d = 1{\cdot}00$

(The price variable is significant at the 25% level and the time variable at the 5% level.)

(3) $(R_t/E_t)' = 0{\cdot}026948 - 0{\cdot}0001580\, t$
$R^2 = 0{\cdot}7963 \quad d = 0{\cdot}91$

(4A) $(E^{dr}_t)' = 642{,}800 - 8{,}775 P^l_t - 202\, t$
$\qquad (4{,}103) \quad (1{,}552)$
$R^2 = 0{\cdot}2790$

Since the sign of t has changed from the original estimation and since it is clearly not significant at any acceptable level, the equation was re-estimated using price alone as an exogenous variable.

(4B) $(E^{dr}_t)' = 640{,}500 - 8{,}729\, P^l_t$
$R^2 = 0{\cdot}2786 \quad d = 1{\cdot}41$

(R^2 is significant at the 5% level.)

(5) $(S^{sl\%}_t)' = 0{\cdot}23364 - 0{\cdot}001232(P^w + P^l)_{t-1}$
$R^2 = 0{\cdot}6084 \quad d = 1{\cdot}97$

(6) $(L^{t\%}_t)' = 0{\cdot}96302 + 0{\cdot}001981\, t$
$R^2 = 0{\cdot}2390 \quad d = 1{\cdot}31$

(R^2 is significant at the 10% level.)

(7) $(L^{d\%}_t)' = 0{\cdot}05161 - 0{\cdot}001940\, t$
$R^2 = 0{\cdot}5489 \quad d = 2{\cdot}49$

(8) $(H^e_t - H^e_{t-1})' = 20{,}390(P^w + P^l)_{t-1} - 975{,}000$
$R^2 = 0{\cdot}2729 \quad d = 1{\cdot}58$

(R^2 is almost significant at the 5% level.)

(9) $(H^w_t - H^w_{t-1})' = 2{,}506{,}800\, P^m_t/P^l_t - 735{,}200$
$R^2 = 0{\cdot}6473 \quad d = 2{\cdot}28$

(10) $(H^r_t/H^e_t)' = 0{\cdot}034463 - 0{\cdot}0009096\, t$
$R^2 = 0{\cdot}9890 \quad d = 3{\cdot}23$

Comparing these re-estimated equations with the originals, it can be seen that in 8 out of the 10, the value of R^2, measuring the goodness of fit, has declined. Bound up with this is a considerable decrease in the significance of the estimated parameters. Although in the original estimation a significance level of 5% was chosen as a minimum standard, it seemed reasonable here to continue to keep the same explanatory variables even if the significance level had fallen considerably. In other words, the original period was used to develop the model, while the subsequent year was used to test the predictions and to get more accurate estimates of the parameters. The only exception to this approach was where the actual sign of the parameter changed, in the case of t in equation 4, where the variable was dropped.

As well as the decrease in R^2, many of the re-estimated equations also showed a marked decrease in the d statistic, indicating positive serial correlation of the residuals.[8]

Both in terms of re-estimated equations and predictive efficiency, the flow equations come out better than the stock ones.[9] The reason for this is that the flow equations refer to the 1964/65 season, while it seems to be during the 1965/66 season that the model failed. This has the implication that the flow predictions for 1965/66 can be expected to be bad, at least for sheep and lambs slaughtered, whatever their efficiency for 1966/67. Amongst the stock variables, rams and ram hoggets did best in terms of the re-estimated equations, though not so well in prediction, since the predictions involved the use of the predicted values of ewes and ewe hoggets. Thus the breakdown of the model did not involve these two technological equations.

Also, wether hoggets performed well. It will be remembered that it was suggested in the previous section that the change in the number of wether hoggets was a function of whether the farmer considered it better to slaughter them as lambs, or carry them over the winter for slaughtering the following year as hoggets. This decision does not involve any long term expectation of price movements and apparently for this reason the model proved satisfactory.

In the case of ewes and ewe hoggets, longer term expectations are important and there appears to have been some factor, not covered by the model, making the farmer more optimistic about the future than observed prices warranted. A month by month observation of price movements does not give any indication of anything that could have caused this optimism, which is further demonstrated by the extraordinarily small number of dry ewes and the large number of wethers. If the optimism was not generated by actual prices, nor apparently by falling costs, its source must be found in either technological progress to the extent of some break-through of farming technique, or some factor outside the industry; perhaps government exhortation or tax incentives. The sudden break-though suggestion does not seem to fit in with actual developments in the industry, nor with the combination of unusually large increases in ewes *and* wethers combined with a *decrease* in dry ewes. This suggests more of an across the board expectation of both lamb and wool price rises. The experience of dramatically falling prices over the last few months must have dispelled this optimism.

[8] It seems probable that this apparent serial correlation should be interpreted in a different way. Where there is a situation, as seen here, with observations for $t - 1$ years approximating to a straight line, and then one final observation far off the line, this will shift the fitted regression line in such a way that the estimated residuals will show positive autocorrelation regardless of whether there was, or was not, autocorrelation in the first $t - 1$ years. Therefore the observed shift to positive autocorrelation can be interpreted as showing that the last observation is markedly different from the earlier ones, or that the relationships found in the first period no longer hold completely for the last year. This, then, merely confirms the sudden drop in R^2.

[9] The apparent efficiency of the second simple model's predictions of lamb deaths would imply negative deaths in about three years' time and thus cannot be considered a good long term model.

This then leaves a problem concerning which estimates should be used in predicting for the current year. If, for some relationships, the experience of the last year did not fit in with that of the earlier years and if it seems likely that the cause of this failure has now passed, would it not be better to use the original estimated equations in predicting the coming year? However, if this position were taken, the meaning of significance levels for prediction purposes would be difficult to evaluate. Any year's predictions could be dismissed on the grounds that the year was atypical and not covered by the model. Therefore, this type of external optimism or pessimism has to be thought of as a random factor worsening the predictive efficiency of the model, but, hopefully, to be included in the model in some direct way as an exogenous variable once enough observations have accumulated.

TABLE 2

Predictions of Stocks for 1967
and Industry Flows for 1965/66 and 1966/67

Variable		Predicted Value
E_{67}		41,135,200
W_{67}		2,211,300
R_{67}		1,011,000
$E^{dr}{}_{67}$		500,800
$H^{e}{}_{67}$		10,665,400
$H^{w}{}_{67}$		3,578,600
$H^{r}{}_{67}$		289,900
L^{t}	1965/66	36,834,200
	1966/67	39,379,700
L^{d}	1965/66	900,600
	1966/67	886,400
L^{sl}	1965/66	22,749,300
	1966/67	23,959,400
S^{sl}	1965/66	8,277,200
	1966/67	8,939,800

In conclusion, Table 2 shows the predictions of stocks for 1967 and industry flows for the seasons 1965/66 and 1966/67. These are based on the re-estimated equations for the reasons given in the preceding paragraph, and the final predictions are derived in the same way as in Table 1. The only exception to this is that equation (4B) is used instead of (4) in the prediction of $E^{dr}{}_{67}$.

References

[1] *Sheep Returns*, Department of Statistics, New Zealand.
[2] *Statistics of Farm Production*, Department of Statistics, New Zealand.
[3] *Monthly Abstract of Statistics*, Department of Statistics, New Zealand.
[4] *Annual Review of the Sheep Industry*, New Zealand Meat and Wool Board's Economic Service.
[5] See, for instance, *Meat*, The New Zealand Meat Producers Board.
[6] Koyck, L. M., *Distributed Lags and Investment Analysis*, North Holland, 1954.