



United States  
Department of  
Agriculture

National  
Agricultural  
Statistics  
Service

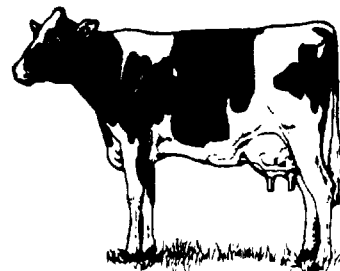
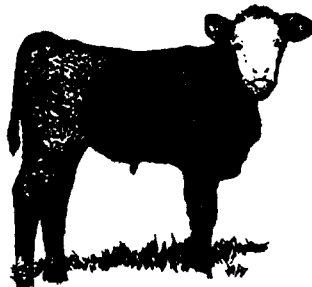
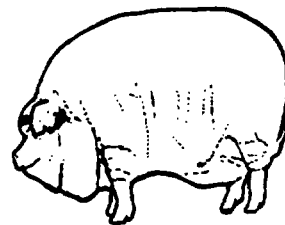
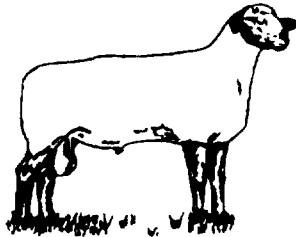
Research Division

RD Research Report  
Number RD-97-05

September 1997

# Improving the Editing and Imputation of Live Weights in the Weekly Slaughter Survey

Dale Atkinson



**IMPROVING THE EDITING AND IMPUTATION OF LIVE WEIGHTS IN THE WEEKLY SLAUGHTER SURVEY**, by Dale Atkinson, Research Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C., September 1997. NASS Research Report No. RD-97-05.

### **ABSTRACT**

The National Agricultural Statistics Service (NASS) summarizes weekly surveys of federally inspected slaughter plants in the United States. These surveys are censuses of all such plants, and serve as the basis for monthly reports of slaughter statistics at the State and U.S. levels. Some supporting data items for the published statistics are readily available for all plants for any particular week, while others are not. Data that are missing or incomplete must be estimated or imputed.

In September 1995 the Livestock and Economics Branch of NASS requested that its Research Division review the imputation procedures for missing and incomplete weight data in the weekly survey. They felt that, since the weight models had been in place for about three years, a tune-up might be in order. This report documents the results of the ensuing study.

It presents the new modeling approach for live weights, and compares the results of the old and new models on the 1996 weekly data. The author recommends updated model parameters (incorporating the 1996 data) and refinements to the way the models are introduced into the statistical editing system. The relatively minor adjustments required to implement these recommendations will provide significant benefits in facilitating periodic updates to the regression models and in allowing the stronger statistical basis of the new approach for missing data to improve the editing of existing data.

### **KEY WORDS**

Editing, Imputation, Nonresponse, Live weight, Dressed weight.

<p>This paper was prepared for limited distribution to the research community outside the U.S. Department of Agriculture.</p>
---

### **ACKNOWLEDGMENTS**

The author would like to thank Linda Simpson for her help in understanding the weekly slaughter survey procedures and Charles Perry for his statistical modeling advice. Thanks also to Mitch Graham for his time and effort in explaining the statistical editing system and in recommending how the new models could be integrated into the existing structure. Finally, thanks to Henry Chiang for his diligence in implementing and testing the system changes.

## TABLE OF CONTENTS

SUMMARY .....	iii
INTRODUCTION .....	1
Table 1: Average Percentages of Dressed and Live Weight Imputed 1994-1996 .....	2
Table 2: The Lower Data Table in the Weekly Slaughter Form .....	3
THE LIVE WEIGHT PROCEDURES USED DURING 1992-1996 .....	3
Description .....	3
Indicated Performance .....	4
THE NEW MODELING APPROACH .....	5
THE RESULTS OF USING THE MODELS .....	8
MODEL IMPLEMENTATION STRATEGY .....	8
CREATING EDIT BOUNDS .....	8
RECOMMENDATIONS .....	11
REFERENCES .....	11
APPENDIX A. THE WEEKLY SLAUGHTER QUESTIONNAIRE .....	12
Exhibit A: The Upper Data Table .....	13
Exhibit B: The Lower Data Table .....	14
APPENDIX B. RESIDUAL PLOTS FROM PREDICTING 1996 DATA .....	15
APPENDIX C. NEW MODEL IMPACT ON U.S. AVERAGE LIVE WEIGHTS .....	20

## SUMMARY

The National Agricultural Statistics Service (NASS) summarizes weekly surveys of federally inspected slaughter plants in the United States. These are censuses of all such plants, and serve as the basis for monthly reports of slaughter statistics at the State and U.S. levels. The reports include counts and average weights of cattle, calves, hogs, pigs, sheep and lambs slaughtered during the previous month. Some supporting data items for the published statistics are readily available for all plants for any particular week, while others are not. Data that are missing or incomplete must be estimated or imputed.

In September 1995 the Livestock and Economics Branch (L&EB) of NASS requested that its Research Division review the imputation procedures for missing and incomplete weight data in the weekly survey. In particular, L&EB staff had noted unusual live weight imputations for some of the large, influential slaughter plants. They felt that, since the weight models had been in place for about three years, a tune-up might be in order. This report documents the results of the ensuing study.

It motivates the discussion by describing the typical amount of imputation for missing data and explains the approaches used for imputing dressed and live weights (and why they differ). It presents the new modeling approach for live weights, and compares the results of the former and new models on the 1996 weekly data. Results of the two modeling approaches are compared at the micro and macro levels, through residual plots and a chart showing impact on the monthly average live weights.

The concept of a tight relationship between editing and imputation is typical of statistical editing systems, and provides the underpinning of the edit and imputation system used for the livestock slaughter data. The implementation of the relationship in this system and the resulting impact of the former and new imputation models on the editing process is discussed.

The author recommends both updated model parameters (incorporating the 1996 data) and refinements to the way the models are introduced into the statistical editing system. The relatively minor adjustments required to implement these recommendations will provide significant benefits in facilitating periodic updates to the regression models and in allowing the stronger statistical basis of the new modeling approach for missing data to improve the editing of existing data.

## INTRODUCTION

The NASS survey forms for the weekly survey of federally inspected slaughter plants are generally filled out by federal plant inspectors from the Food Safety and Inspection Service (FSIS), another agency of the U.S. Department of Agriculture.

The questionnaire used in the collection is designed in two parts, with upper and lower data tables (see Appendix A). The upper table is used to collect the daily numbers of head slaughtered for the reporting week. Under federal regulations this information must be maintained by each slaughter plant and made available to plant inspectors on an ongoing basis. Supported by the mandated availability of these data, the inspectors have historically provided very complete and accurate information on the daily slaughter counts. As a result there are very few missing data in the top portion of the questionnaire and, therefore, no issue with problematic imputations there.

The bottom portion of the questionnaire consists of a table for collecting the weekly number of animals slaughtered (the sum of the daily totals less post-mortem condemnations) and the associated weekly total live and dressed weights<sup>1</sup>. Since

---

<sup>1</sup> Total live weight as discussed here is the weekly total gross weight of live animals when received at the slaughter plant, while total dressed weight is the weekly total weight of the resulting animal carcasses.

federal regulations do not require plants to maintain the information for this portion of the questionnaire, it is not as readily available to inspectors as are the daily counts. As a result, the data in the bottom table are not always reported. Missing weekly head counts of slaughtered animals (less condemnations) do not create too great a problem, however, as these can be accurately estimated from the virtually complete daily entries in the upper portion of the form, since numbers of condemned carcasses are generally small.

Unfortunately, the weekly totals of dressed and live weights are more difficult to estimate, and are subject to more problematic imputation procedures. Also, making these procedures as effective as possible is especially important since (as indicated in Table 1) there are significant missing weight data. Consequently, improving the imputation of these data for missing and incomplete reports was the focus of this study.

The current imputation approach used for plants with missing weekly totals of dressed and live weight is thoroughly documented in an October 1990 research report (Mazur, 1990). The only major deviation from the recommendations of this report was in the implementation of live weight imputation for calves. Since there is only one class of calves, calf live weights could have been imputed based on an historical ratio of live weight to dressed weight, as recommended. Instead, the live weight imputation procedure is the same for calves as it is for all other species.

**Table 1: Average Percentages of Dressed and Live Weight Imputed 1994-1996**

Species and Class	% of Live Weight Imputed	% of Dressed Weight Imputed
Steers		20.4
Heifers		18.8
Total Cows		39.8
Bulls & Stags		37.1
CATTLE-Total	18.7	23.0
CALVES-Total	60.5	39.1
Barrows & Gilts		8.0
Sows		28.6
Stags & Boars		17.8
HOGS-Total	8.8	9.4
Mature Sheep		40.4
Lambs & Yearlings		38.4
SHEEP-Total	42.7	38.5

Note: Data for shaded cells are not collected.

In the current report the author will refrain from rehashing the minute details of the procedures, but instead will highlight the basic differences between those used for dressed and live weights, explain why they differ, and discuss the implications of these differences for the imputation of missing weight data.

The editing and imputation procedures for dressed and live weights are largely dictated by the level of data collected for these items. As indicated in Table 2 (which shows the bottom table of the survey form), total dressed weights are collected at the class level within species, while total live weights are collected only

at the species level. This lower level of detail for live weights is due to the perceived reporting difficulty in obtaining these at the class level. It presents a unique problem for live weight imputation, since each class of animal characteristically has a different dressing percentage (i.e., the ratio of dressed weight to live weight expressed as a percentage) and the mix of slaughtered animals by class varies from week to week.

The differing levels of the collected data, combined with the need to maintain internal consistency in the dressed and live weights imputed for a particular firm (i.e., they can't be imputed independently or dressed weight could exceed live weight), led to the different missing data procedures for dressed and live weights.

In the current livestock slaughter system, all required dressed weight imputation is done first, using a sophisticated statistical profiling approach. A plant's missing dressed weights are estimated based on its reporting history. Specifically, the automated routine imputes missing dressed weights using a weighted biweight of the plant's most recent 13 weeks of reported dressed weights *for that class*. Only when the reporting history for a plant is inadequate (or totally absent) does the system resort to imputing dressed weights based on responses of "similar" plants.

This dressed weight procedure requires no parameter updates and the system continuously refreshes itself with each week's reports. Short of obtaining the current week's report in the first place, this procedure seems to provide dressed weights about as effectively as anything that could

**Table 2: The Lower Data Table in the Weekly Slaughter Form**

Weekly Total Head, Total Live Weight and Total Dressed Weight (Exclude Post-Mortem Condemnations)			
Species and Class	Number of Head	Total Live Weight	Total Dressed Weight
Steers	701		801
Heifers	702		802
Total Cows	703		803
Bulls & Stags	704		804
CATTLE-Total	700	705	800
CALVES-Total	711	715	811
Barrows & Gilts	721		821
Sows	722		822
Stags & Boars	723		823
HOGS-Total	720	725	820
Mature Sheep	731		831
Lambs & Yearlings	732		832
SHEEP-Total	730	735	830

Note: Data for the shaded cells (i.e., class-level live weights) are not collected.

be designed.

Once all missing dressed weights are imputed, live weight imputation begins. This is based on functional equations that relate dressed weight (DW) by class to live weight (LW) for the species. Compared to the one used for dressed weights, the live weight edit/imputation procedure is something of a step-child that warrants a bit more periodic scrutiny and maintenance. Since the live weight procedures seem to be the “weak link” in the livestock slaughter edit/imputation system, they are the primary focus of this report.

### THE LIVE WEIGHT PROCEDURES USED DURING 1992-1996

**Description:** In constructing the statistical editing system for livestock slaughter, the designers followed the third principle of Fellegi-Holt (Fellegi-Holt, 1976). This principle states that the imputation rules should be derived from the editing rules. Accordingly, dual-purpose functional relationships between dressed weights and live weights were designed to serve as both reasonable edit bounds for reported live weights and the basis for imputing missing live weights. Specifically, the procedure used from 1992-1996 imputed live weights for cattle, calves, hogs, and sheep by

averaging the predictions from two functional equations which related live weight to dressed weight through dressing proportions.

As used here, a dressing proportion refers to the ratio of dressed weight to live weight. Often this ratio is expressed in percentage form and referred to as the dressing percentage or yield. In the present context the reciprocals of the dressing proportions for each class of animal are used as coefficients in the functional equations, in a ratio estimation context.

The highest "reasonable" dressing proportion for a particular class of animal will be referred to as its upper dressing proportion (UDP). Similarly, the lowest reasonable value will be referred to as its lower dressing proportion (LDP). The values of UDP and LDP to use in the functional equations were derived from a combination of expert opinion on reasonable ranges and data analysis performed in 1991. The values of LDP and UDP established at that time for each class of animal were then used from 1992-1996.

For example, with the hog classes of barrows and gilts (B&G), sows, and stags and boars (S&B) the functional equations defining the upper and lower editing limits for existing total live weights were as follows, with any reported live weight outside these bounds flagged as an error.

$$\begin{aligned} LW \text{ (Upper)} = & (1/LDP \text{ for B\&G}) * DW \text{ (B\&G)} \\ & + (1/LDP \text{ for Sows}) * DW \text{ (Sows)} \\ & + (1/LDP \text{ for S\&B}) * DW \text{ (S\&B)} \end{aligned}$$

and

$$\begin{aligned} LW \text{ (Lower)} = & (1/UDP \text{ for B\&G}) * DW \text{ (B\&G)} \\ & + (1/UDP \text{ for Sows}) * DW \text{ (Sows)} \\ & + (1/UDP \text{ for S\&B}) * DW \text{ (S\&B)}. \end{aligned}$$

Missing live weights were imputed as

$$LW = (LW \text{ (Upper)} + LW \text{ (Lower)}) / 2.$$

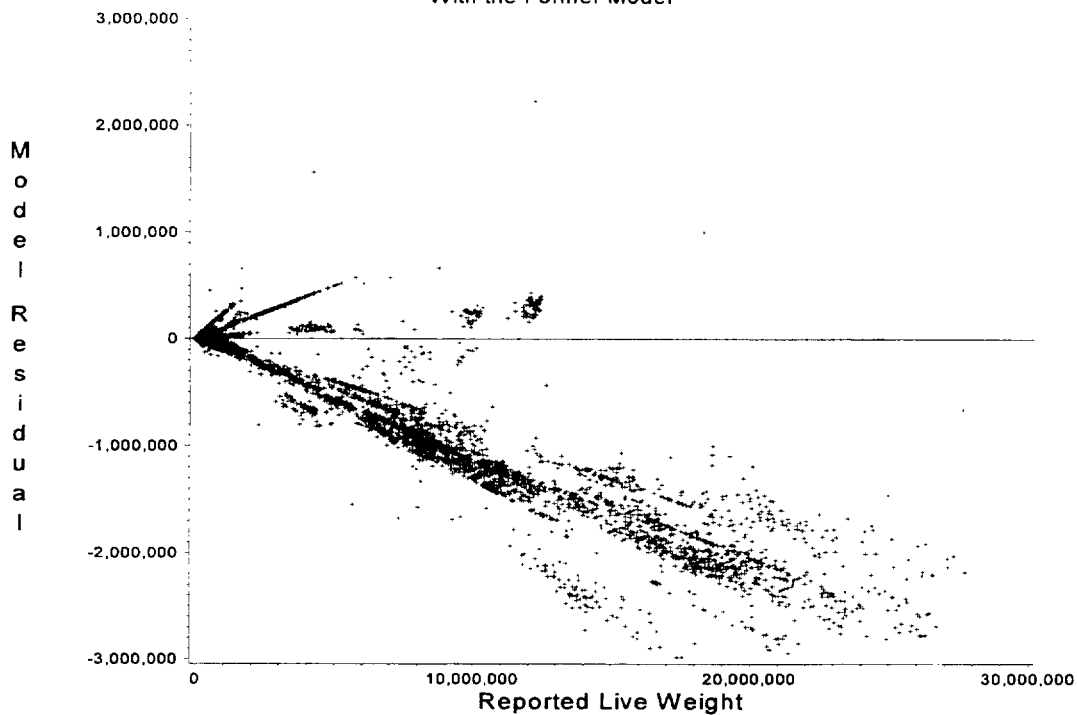
Since missing dressed weights are imputed prior to any processing of live weights, all variables on the right side of these equations were well defined for every record for both editing and imputing live weights.

**Indicated Performance:** A review of differences (residuals) between imputation model results and reported live weights indicated significant room for improvement for all species. Figure 1 shows a residual plot of 1994-95 reported live weights for hogs using the former model. This plot and all the succeeding ones in this report compare the actual reported live weight for good usable records with what the model would have imputed if the live weight had been missing and all other information in the record was unchanged. The residual plots should be indicative of the performance of the models for actual missing data.

The plot in Figure 1 was typical of those observed for other species. Among other anomalies, it shows a clear linear trend with nonzero slope in the residuals, indicating a need for model coefficient adjustment. This type of relationship, in which imputed live weights are progressively further from the reported data as the total slaughter weight increases, explains why the models were noticeably off target for some of the larger plants.



Figure 1: Residual Plot from Predicting Reported 1994-95 Hog Live Weights With the Former Model



### THE NEW MODELING APPROACH

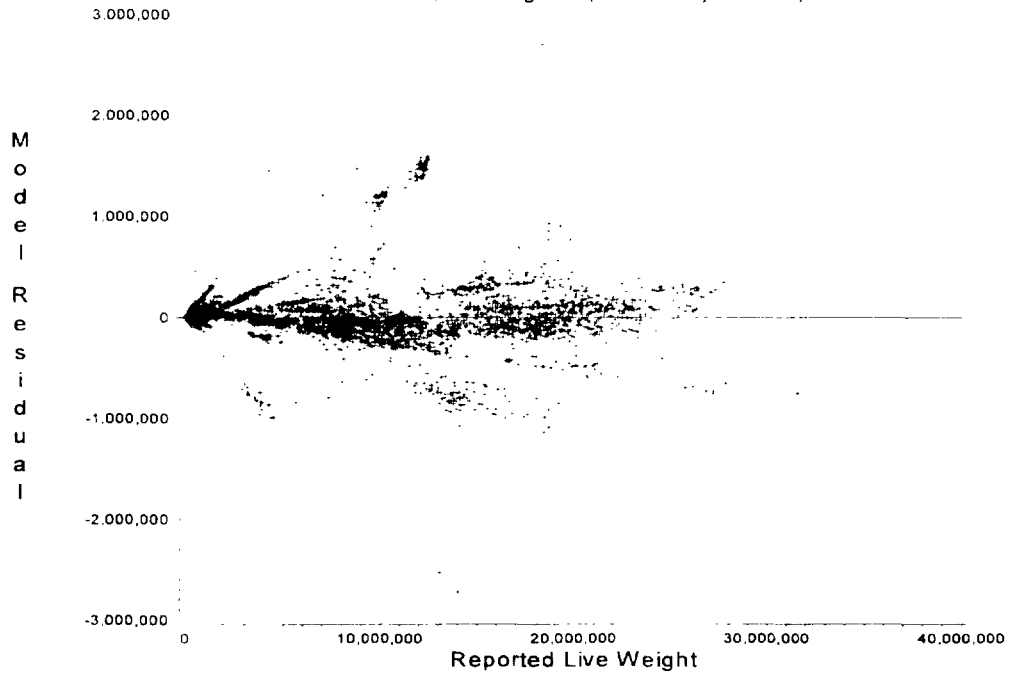
The new approach generates the ratio adjustments relating dressed weight and live weight through regression analysis. It also augments the basic additive model of dressed weight on live weight with multiplicative adjustments for differential effects in dressing proportions due to region, season, size of firm, and individual firm peculiarities. The general form of the models is as follows:

$$L\hat{W} = (\beta_1 * DW_1 + \dots + \beta_k * DW_k) \\ *(multiplicative\ adjustments).$$

The nonlinear models were fit using SAS' "PROC NLIN," with the multiplicative adjustments made through dummy variables representing ten slaughter regions, four seasons, three size breaks on the current week's slaughter totals (to represent the size of the slaughter plant), and individual plant identification. Initially, individual plant adjustments were not considered; however, excluding these resulted in residual plots with a "spikey" appearance. See Figure 2.

Some of the spikes represent the successive weekly reports of an individual plant. As significant individual plant effects were noticed in residual plots and "large residual" listings from early modeling efforts, they were explicitly included in subsequent models.

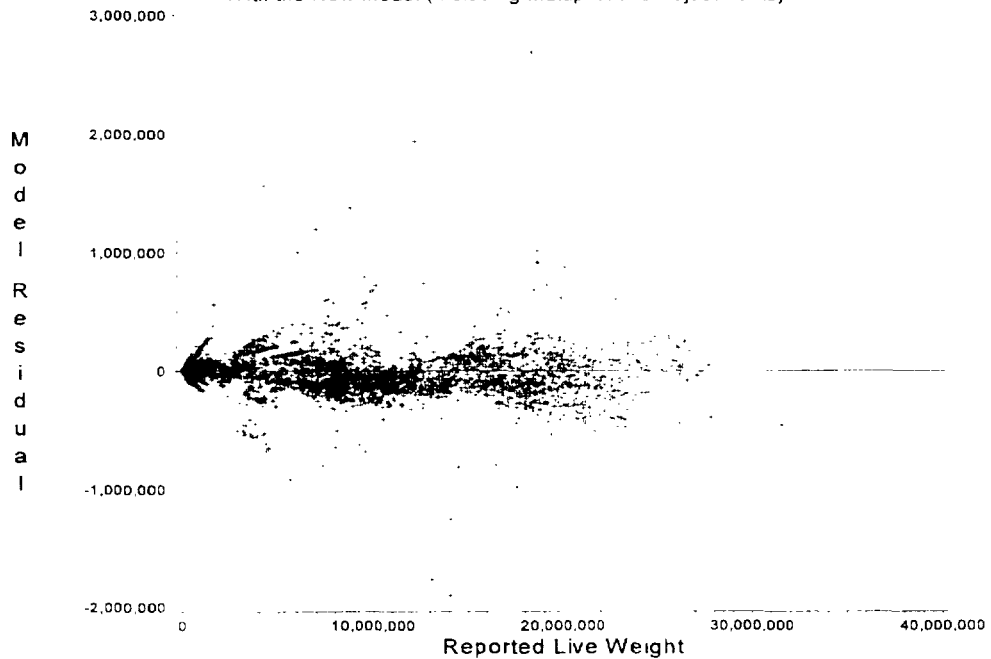
Figure 2 Residual Plot from Predicting Reported 1994-95 Hog Live Weights With the New Model (Excluding Multiplicative Adjustments)



Fitting the full model (including multiplicative adjustments for size of firm, region and individual firm idiosyncrasies)

resulted in the residual plot in Figure 3. Notice that most of the spikey appearance has disappeared, leaving a reasonable looking residual pattern.

Figure 3: Residual Plot from Predicting Reported 1994-95 Hog Live Weights With the New Model (Including Multiplicative Adjustments)



In general, regression modeling is an interactive and somewhat artistic process. It is interactive in that the results of initial trial models often dictate variables (and sometimes observations) to include in subsequent trial models. The artistry comes into play in the identification and handling of outliers. In this application a fairly conservative approach was taken to handling outliers in the modeling process. Only observations that appeared to be very unrepresentative or in error were deleted. While dealing with outliers is generally unsatisfying, some accommodation ultimately must be made for them. Even when large numbers of observations are available for modeling, the underlying least squares methodology makes regression modeling very sensitive to outliers.

While robust regression and iteratively reweighted least squares are possibilities for handling the outlier problem, these are difficult to implement for a nonlinear multiplicative model. Also, most outliers encountered in this analysis for which action was deemed necessary fell in the category of gross error. Erroneous data may be just as effectively removed as down weighted, especially when the overall pool of modeling data is large.

The removed data consisted largely of records that appeared to be key-entry decimal errors that were potentially very detrimental to the modeling effort. These resulted in dressing percentages clustered around 5 percent (i.e., one-tenth of the typical 50 percent) or greater than 100 percent.

There were many potential multiplicative

adjustments for region, season, size of firm, and individual plant. Also, with two years of weekly data available, there were many observations to model. As a result, many indicated adjustments were statistically significant. However, to avoid seriously over-fitting the models, Ockham's Razor was chosen as the guiding principle in the modeling activity. In the loosely translated words of William of Ockham, "Entities are not to be multiplied beyond necessity."

The chosen implementation of this principle specified that only terms that were both statistically and practically significant (resulting in an adjustment of at least 1 percent) were retained in the model. Some artistry (and more trial and error) was required to determine the best-fitting, parsimonious models for each species. The following is an example of one of the final selected models:

$$\begin{aligned} \text{Live Weight} = & 1.47*(\text{DW for B\&G}) \\ & + 1.65*(\text{DW for Sows}) \\ & + 1.88*(\text{DW for S\&B}), \end{aligned}$$

multiplied by

- 0.93 if the slaughter plant is in region rr,
- 0.94 if the plant is in region ss or tt,
- 0.95 if the plant is in region uu or vv,
- 0.97 if the plant is in region ww,
- 1.12 if the plant is in region xx,
- 0.98 if the total slaughter > 4999 head,
- 0.94 if the plant is # www in State aa,
- 1.14 if the plant is # xxx in State bb,
- 0.90 if the plant is # yyy in State cc,
- and
- 0.99 if the plant is # zzz in State dd.

Notice that several of the adjustments are substantial, even some of those for individual slaughter plants.

Anticipated impact in the estimates was assessed at the State and U.S. level and presented first to L&EB staff, and separately to NASS' Senior Management at an Administrator's Staff Meeting in December 1995. The aggregate effect of the new models on average live weights at the U.S. level was expected to be less than 1 percent for cattle, hogs, and sheep, and about 5-6 percent for calves.

Mitch Graham (now in the California SSO), who initially programmed the edit and imputation system for livestock slaughter, provided invaluable assistance in specifying how it could be modified to accommodate the new models. The combination of a document written by him titled "Suggestions for Implementation of Live Weight Imputation Improvements" (12/15/95) and one written by Dale Atkinson titled "Final Slaughter Models and Implementation Parameters" (01/05/96) provided the specifications for implementing the new models.

## **THE RESULTS OF USING THE MODELS**

There's always some uncertainty involved in applying models to a new data set. This modeling exercise was certainly no exception. In particular, individual plant adjustments were specified with some fear and trepidation, wondering how consistent the relationships would be from year to year. Happily, the results in predicting the 1996 data with the models built on 1994-95 data were gratifying. The

relationships were consistent and the adjustments specified, including those for individual slaughter plants, were reflected in the 1996 data. The residual plots for the former models and new models for all four species are included in Appendix B. The new cattle, calf, and hog models performed substantially better than their former counterparts. The sheep models, which seem to be the most difficult to specify well, showed some improvement. Notice the spikes in the calf and hog residual plots for new plants in 1996. These indicate that the models do need to be updated periodically.

## **MODEL IMPLEMENTATION STRATEGY**

For 1996 the former models continued to be used operationally, but the new models were tested in parallel each month. After a complete year of parallel testing, the new models were implemented operationally in January 1997 (for the February slaughter release). The January release contained a comparison table showing the 1996 monthly average live weights from both the former and new models, to help data users understand the impact of the procedural shift. The graphic in Appendix C indicates the actual month-to-month percentage differences in the average live weights obtained from the two procedures. These were very much in line with what had been anticipated.

## **CREATING EDIT BOUNDS**

Probably the most challenging part of fitting the new imputation models into the old structure was in creating upper and lower editing bounds. Recall that in the

former procedure, functional equations that provided reasonable editing limits were established, and their mid-points (i.e., their means) formed the values for any required imputations. With the new modeling approach, the procedure was reversed. The target imputation models were generated and corresponding upper and lower editing bounds (i.e., models) were needed to bracket them.

The initial implementation of the new models called for creating the upper and lower bound editing equations by starting with the coefficients of the target imputation equation and adding and subtracting suitable multiples of their standard errors.

Creating regression equations that appropriately bound a multiplicative regression model in this manner is a nontrivial task. There may not be a unique solution, or any solution for that matter. Even if a solution does exist, it is often a difficult and inelegant process to find it. It's even more difficult to provide guidance for anyone else to do it. If not simplified, this activity could be an obstacle to future model parameter updates and impede the transfer of the modeling effort to other potential modelers.

Fortunately there *is* an easier (and probably better) way to achieve appropriate editing bounds. What we really want to do in creating them is to form a confidence interval around the predicted value from the imputation model. The true bounds for this are as follows:

$$\hat{y}_i \pm c * s.e(\hat{y}_i)$$

where *c* is the confidence coefficient required for an appropriate edit flag rate and

$$s.e(\hat{y}_i) = \text{root mean squared error} + \text{fcn}(\partial \text{ of the model variables})$$

The root mean squared error (RMSE) from the fitted model is constant for all observations, while the function of partial derivatives makes non-negative individual observation adjustments to the prediction intervals. This second term on the right side of the equation is mathematically complex, but in the present modeling situation it contributes relatively little to the overall standard error -- virtually always less than 2 percent. Therefore, a good approximation of our desired edit bounds is simply

$$\hat{y}_i \pm c * RMSE.$$

Edit bounds using this approximation and *c*=3 were formed with good results in identifying potential error records in the 1994-1996 data. The records flagged for each species were identical to those flagged using the more cumbersome (and difficult to maintain) true bounds. The resulting error file also subjectively seemed to be preferable to error files from both the former and new models (as implemented in 1997).

With this revised approach the modeler need only specify the target model parameters, the coefficient "c" and the RMSE. The upper editing bound is then generated as simply the target model prediction plus "c" times the RMSE.

Similarly, the lower bound is generated as the target model prediction minus “c” times the RMSE. The only program code modification necessary will be in the basic definitions of these upper and lower bounds.

One other problem with the current editing of live weights that should be addressed is that there are no critical errors built into the system. In every year’s data reviewed in this modeling exercise (i.e., 1994 through 1996), there were several “reported” dressed weights that exceeded “reported” live weights. There were clusters of these in the ranges of 500-700 percent and 5-7 percent. Since dressing percentages are generally in the 50-70 percent range, these situations look suspiciously like decimal problems in the key-entry process. [Fortunately, they were relatively small reports that had no impact on the summary statistics.]

The former editing process flagged these records, but they were probably missed since far too many other records were also flagged. Table 3 shows the indicated edit flag rates in the “clean” 1994-96 reported live weights.

**Table 3: Indicated Error Rates in Reported 1994-96 Live Weights**

Species	Former Proced.	New Proced.	Revised 1/ Proced.
Cattle	17.6	6.5	2.1
Calves	73.2	23.9	1.6
Hogs	25.1	7.2	7.2
Sheep	27.4	6.6	2.2

1/ With a c value of 3.

With the revised procedure the error rate can be adjusted (by changing the coefficient c of the RMSE in creating the upper and lower bounds), but there is no guarantee that decimal problems such as those indicated above will be flagged. In fact, in experimenting with this approach it was found that these particular problems were not flagged with any reasonable value of the confidence coefficient!

Erroneous records such as these can escape unflagged with the revised editing bounds *where slaughter weights are small*, since for these records the constant confidence interval widths are substantial relative to the reported live weights. While these errors are not likely to have an appreciable effect in summary results, the edit should be designed to flag them. Considering how few real errors there are in the slaughter data (which in general appear very clean), there is no reason to miss gross errors in even small slaughter reports.

However, the only fool-proof way to protect against these small record anomalies is to supplement the statistical bounds with “critical” range edits to flag unrealistic dressing percentages. Values of, say, 10 percent for the lower bound and 90 percent for the upper bound would ensure that clearly erroneous values are not passed. Table 4 shows the distribution quantiles for dressing percentage by species for the 1994-96 data. Notice that yields of less than 40 percent or greater than 80 percent are rare for every species.

**Table 4: Dressing Percentage Distribution  
Quantiles in the 1994-96 Data**

Species	1%	5%	50%	95%	99%
Cattle	42.6	46.0	57.1	64.0	71.4
Calves	45.2	50.0	59.8	70.0	74.2
Hogs	50.0	55.0	70.8	79.2	83.1
Sheep	40.0	41.3	50.0	64.1	80.0

Finally, the models were revised to include the 1996 data. In most cases the changes in model coefficients, where indicated at all, were very small. This was expected, since the models built on 1994-95 data had predicted the 1996 data very well. However, incorporating the 1996 data will add stability to some of the individual coefficients and better reflect new plants which began operation in 1996. To protect the confidentiality of individual plant data, the new model coefficients will be conveyed to the Livestock and Economics and Systems Services Branches outside of this report.

## RECOMMENDATIONS

The author makes the following recommendations:

- 1) Define the upper and lower editing bounds as the target imputation model  $\pm$  constant "c" times the model root mean squared error.

Here "c" and the appropriate root mean

squared errors will be specified by the modeler.

- 2) Add critical range edits to preclude impossible relationships between dressed and live weight (i.e., dressed weight exceeding live weight) that may otherwise pass the edit for small slaughter totals.

and

- 3) Update the model parameters to include the 1996 data, as specified under separate cover.

None of these recommended changes will have any significant impact on the aggregate estimates. Therefore, additional (formal) parallel testing should not be necessary. The changes should be implemented operationally as soon as they are made and tested. The author is willing to assist with the testing process.

## REFERENCES

Fellegi, I.P. and Holt, D, "A Systematic Approach to Automatic Edit and Imputation," Journal of the American Statistical Association, March 1976, Vol. 71, No. 353.

Mazur, Cathy, "A Statistical Edit for Livestock Slaughter Data", SRB Research Report Number SRB-90-01.

**APPENDIX A. THE WEEKLY SLAUGHTER QUESTIONNAIRE**



# Exhibit A: The Upper Data Table

NASS:AMS:FSIS  
Form LS-149

UNITED STATES DEPARTMENT OF AGRICULTURE

## WEEKLY LIVESTOCK SLAUGHTER REPORT

Form Approved  
OMB Number 0535-0005  
Expiration Date 01/31/96  
QID 110040  
Project Code 254

Plant Name \_\_\_\_\_ Establishment No. \_\_\_\_\_ State \_\_\_\_\_

Plant Location (City) \_\_\_\_\_ Inspector \_\_\_\_\_

Response to this form is voluntary and not required by law. However, cooperation is very important in order to fulfill responsibilities mandated by the Meat Inspection Act and to provide statistical information to maintain an orderly flow of red meat throughout the livestock industry.

**INSTRUCTIONS:** *Include all species slaughtered in each plant including custom slaughter. Complete a separate Form LS-149 each week for each Federally inspected plant. See the back of this form for detailed instructions.*

WEEK ENDING (SATURDAY) \_\_\_\_ / \_\_\_\_ / \_\_\_\_

IF NO KILL CHECK HERE

Mo. Day Yr.

NUMBER HEAD SLAUGHTERED DAILY (Including Post-Mortem Condemnations)							
Species and Class		Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Steers -----		101	201	301	401	501	601
Heifers -----		102	202	302	402	502	602
Cows: Dairy -----		105	205	305	405	505	605
All Other -----		106	206	306	406	506	606
Bulls & Stags -----		104	204	304	404	504	604
<b>CATTLE Total</b> 010		100	200	300	400	500	600
<b>CALVES Total</b> 011		110	210	310	410	510	610
Barrows & Gilts -----		121	221	321	421	521	621
Sows -----		122	222	322	422	522	622
Stags & Boars -----		123	223	323	423	523	623
<b>HOGS Total</b> 020		120	220	320	420	520	620
Mature Sheep -----		131	231	331	431	531	631
Lambs & Yearlings -----		132	232	332	432	532	632
<b>SHEEP Total</b> 030		130	230	330	430	530	630
<b>GOATS Total</b> →		140	240	340	440	540	640
<b>EQUINE Total</b> →		150	250	350	450	550	650

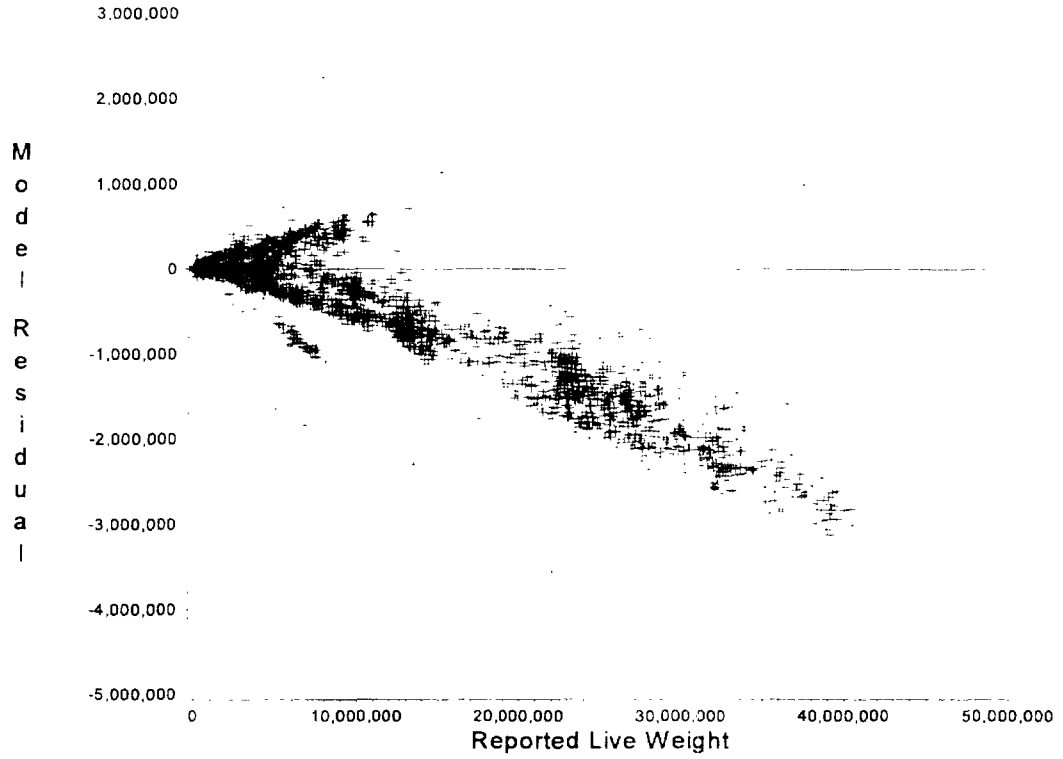
**Exhibit B: The Lower Data Table**

WEEK ENDING (SATURDAY)      /      /       
 Mo. Day Yr.

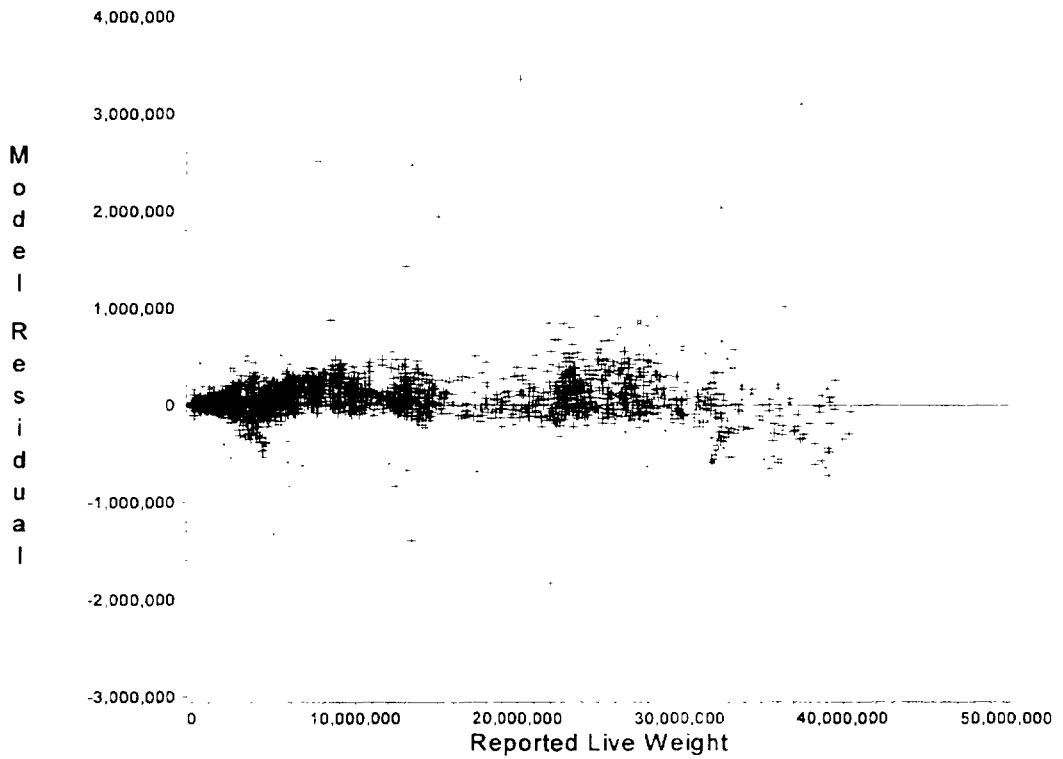
WEEKLY TOTAL HEAD, TOTAL LIVE WEIGHT AND TOTAL DRESSED WEIGHT (Exclude Post-Mortem Condemnations)			
Species and Class	Number of Head	Total Live Weight	Total Dressed Weight
Steers -----	701		801
Heifers -----	702		802
Total Cows -----	703		803
Bulls & Stags -----	704		804
<b>CATTLE - Total</b> →	700	705	800
<b>CALVES - Total</b> →	711	715	811
Barrows & Gilts -----	721		821
Sows -----	722		822
Stags & Boars -----	723		823
<b>HOGS - Total</b> →	720	725	820
Mature Sheep -----	731		831
Lambs & Yearlings -----	732		832
<b>SHEEP - Total</b> →	730	735	830

**APPENDIX B. RESIDUAL PLOTS FROM PREDICTING 1996 DATA**

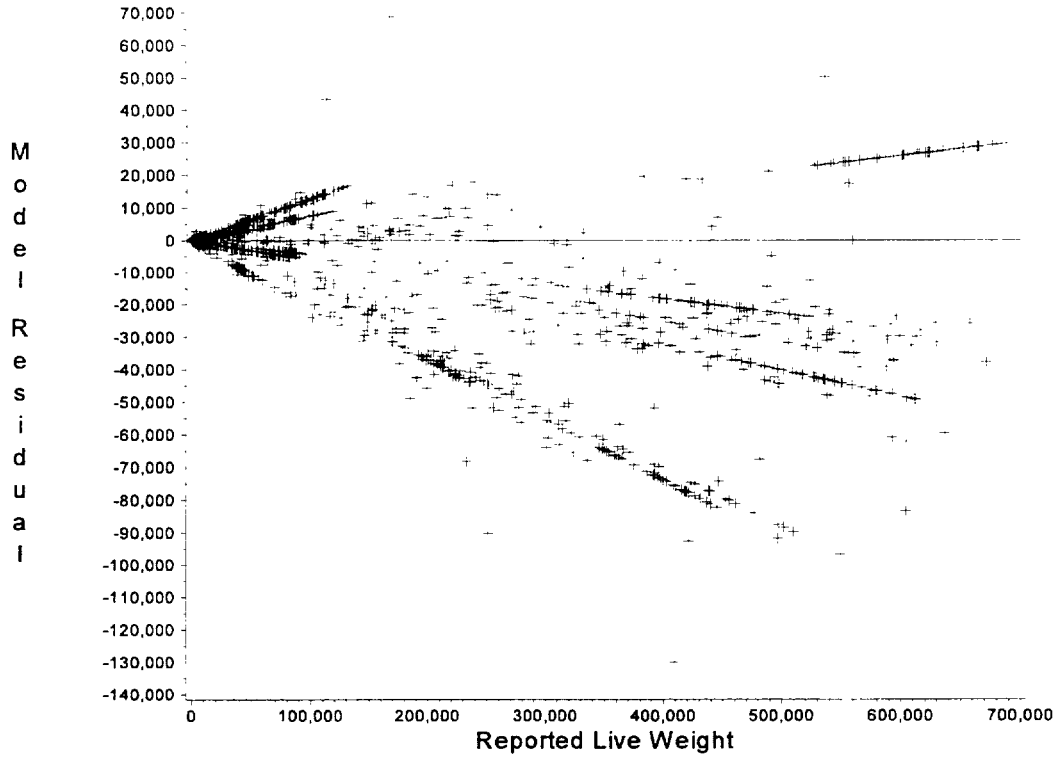
Residual Plot from Predicting Reported 1996 Live Weights of Cattle with the Former Model



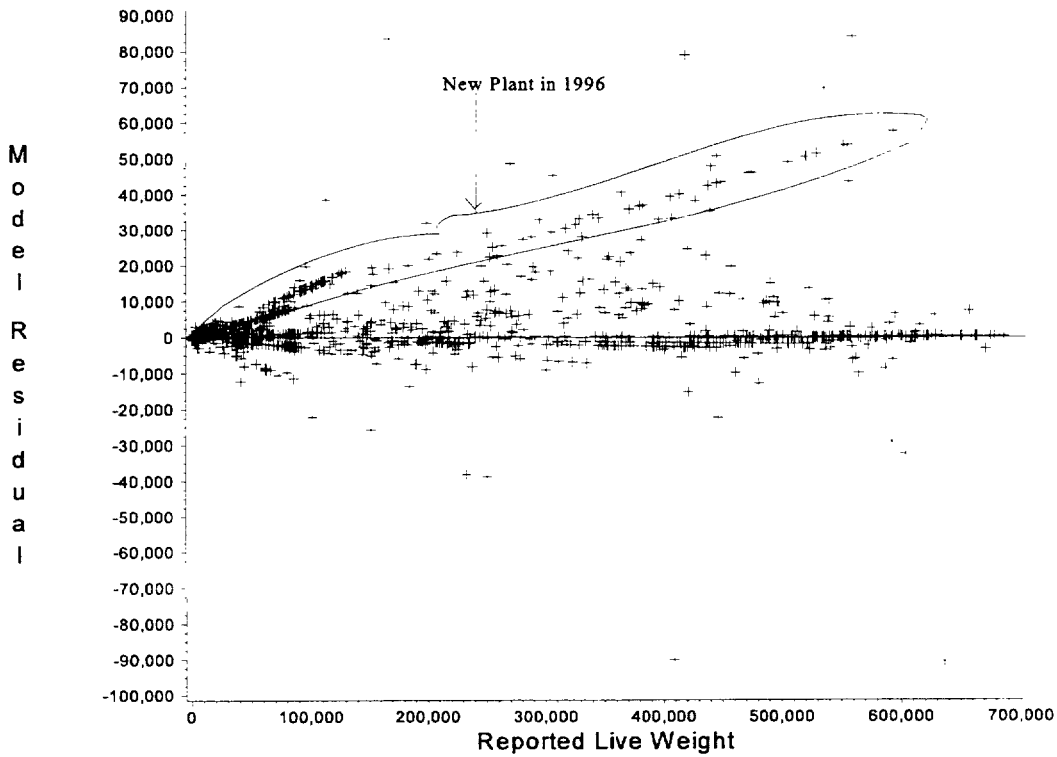
Residual Plot from Predicting Reported 1996 Live Weights of Cattle with the New Model



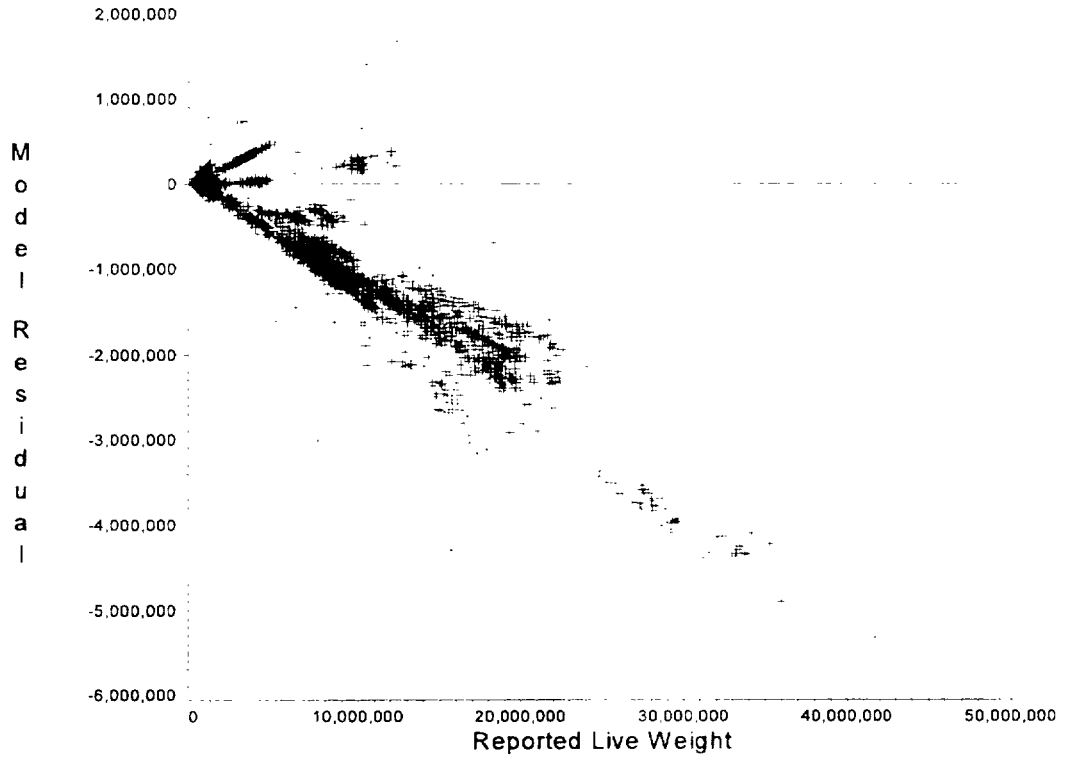
Residual Plot from Predicting Reported 1996 Live Weights of Calves with the Former Model



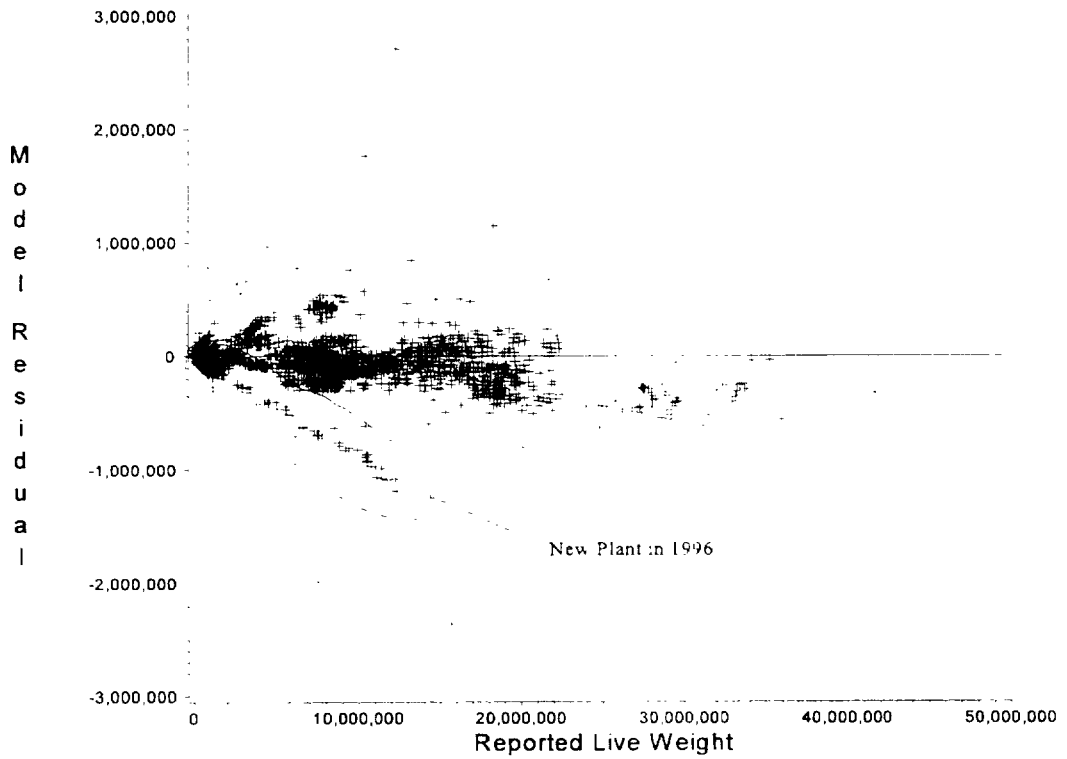
Residual Plot from Predicting Reported 1996 Live Weights of Calves with the New Model



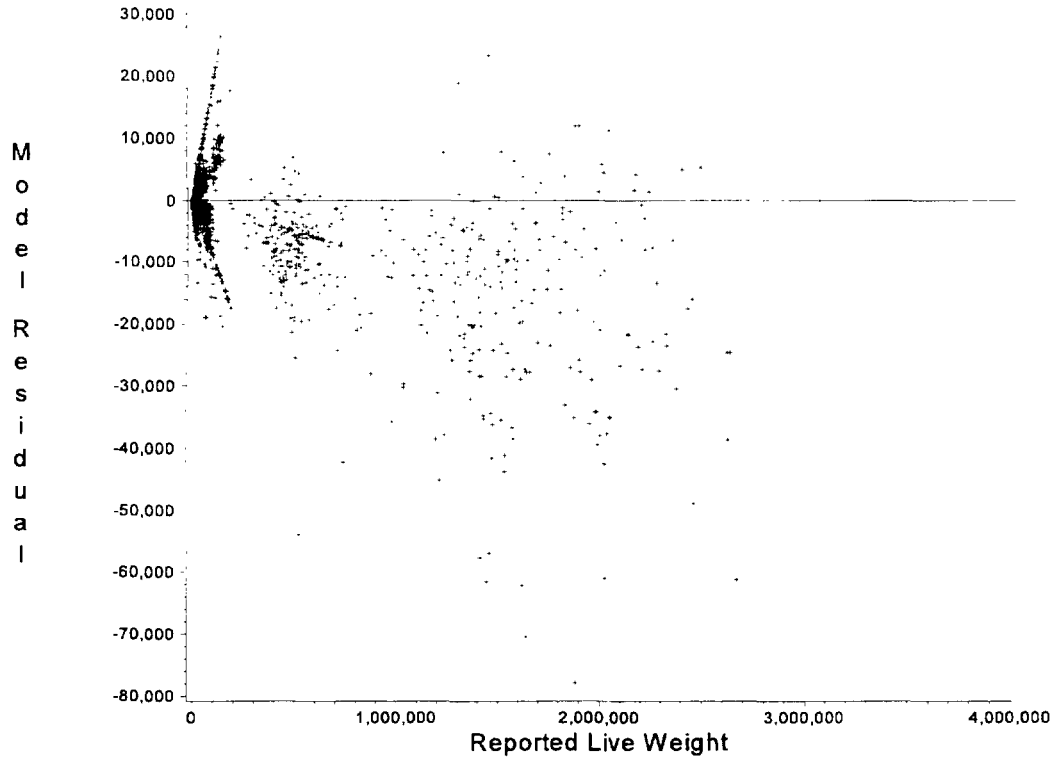
Residual Plot from Predicting Reported 1996 Live Weights of Hogs with the Former Model



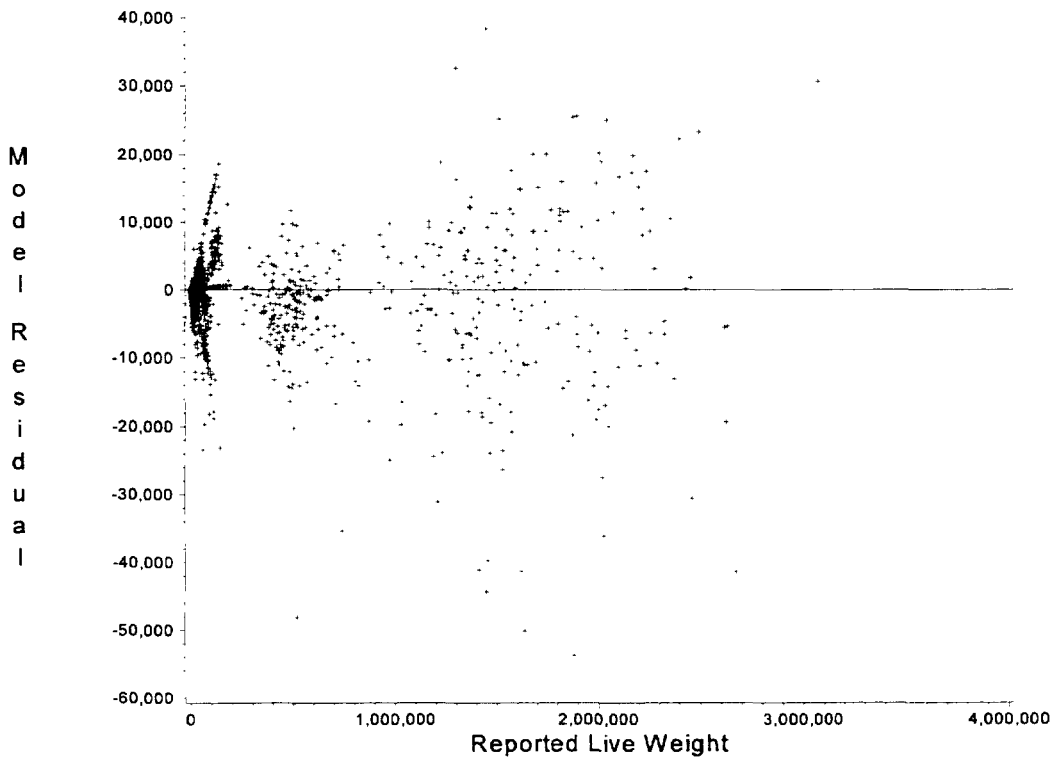
Residual Plot from Predicting Reported 1996 Live Weights of Hogs with the New Model



Residual Plot from Predicting Reported 1996 Live Weights  
of Sheep with the Former Model



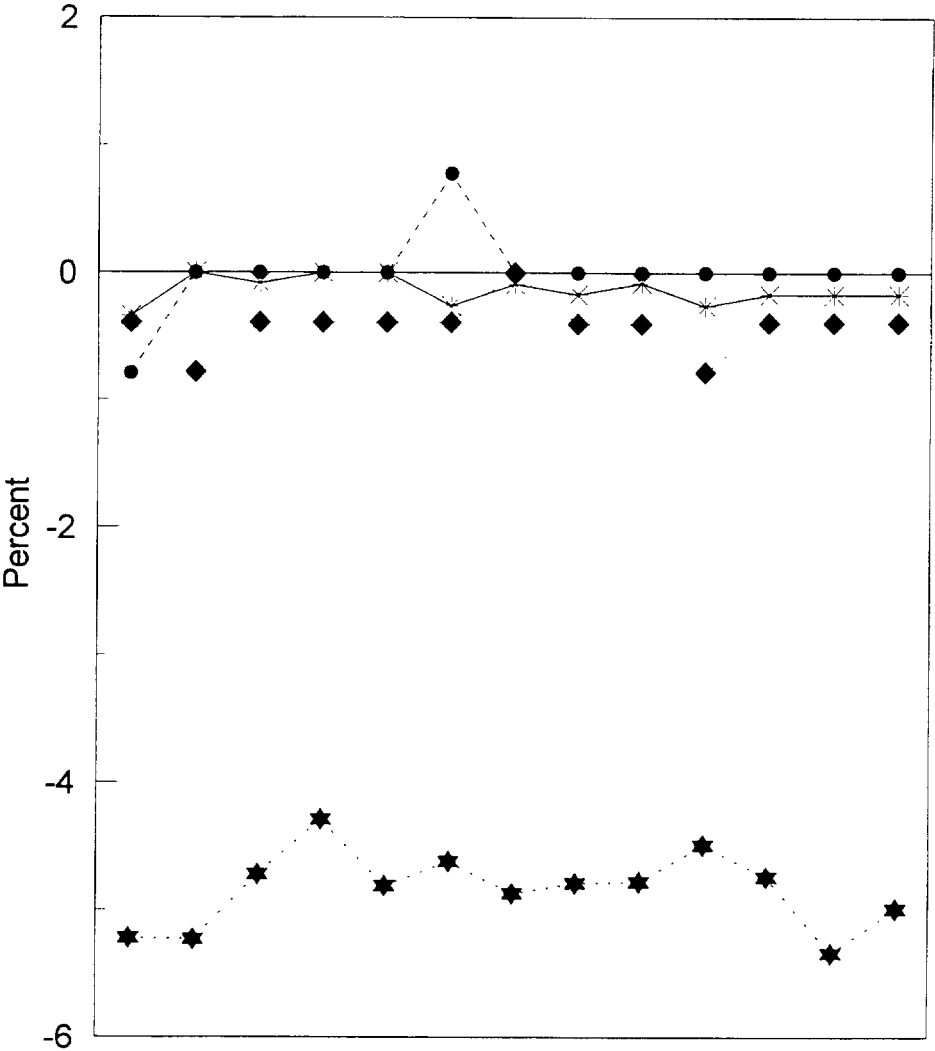
Residual Plot from Predicting Reported 1996 Live Weights  
of Sheep with the New Model



## APPENDIX C. NEW MODEL IMPACT ON U.S. AVERAGE LIVE WEIGHTS



## Percentage Changes in 1996 U.S. Live Weight Averages From Using the New Models



Month	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	All
Cattle *	-0.34	0.00	-0.08	0.00	0.00	-0.26	-0.09	-0.17	-0.08	-0.26	-0.17	-0.17	-0.17
Hogs ◆	-0.39	-0.78	-0.39	-0.39	-0.39	-0.39	0.00	-0.40	-0.40	-0.78	-0.39	-0.39	-0.39
Calves ★	-5.22	-5.23	-4.72	-4.29	-4.81	-4.62	-4.87	-4.79	-4.78	-4.49	-4.74	-5.34	-4.99
Sheep ●	-0.79	0.00	0.00	0.00	0.00	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00