AN ECONOMETRIC ANALYSIS OF NURSING HOME COSTS USING COMPLEX SAMPLE SURVEY DATA

Mark R. Meiners
National Center for Health Statistics

Introduction and Background

In 1973 the National Center for Health Statistics expanded its National Nursing Home Survey (NNHS) to collect data on the cost to facilities of providing long-term health care. The sample design for the cost data and the other facility characteristics was stratified random sampling. As with most large sample surveys of this sort, the principal goal of the sample design was to achieve a stated degree of precision on the estimation of various descriptive statistics (e.g., percentages, means, totals) for a minimum cost. There is, however, considerable interest in using these data in research which is orientated toward making analytical inferences either to suggest or confirm specific hypotheses.

One of the tools most widely used to accomplish analytical research is the multiple regression technique. This is particularly true of econometric research. Problems arise, however, in the direct application of the regression technique to data generated by a complex sample survey procedure because the randomness assumptions are violated. The 1973-74 NNHS, for example, was stratified into 26 certification (e.g., Both Medicare and Medicaid or Medicare only; Medicaid only; and Not Certified), bedsize groups.[1] Further stratification within each of these primary strata was accomplished by ordering by type of ownership, geographic region, State and county. The sample was then selected systematically after a random start within each primary stratum with the sampling fraction varying in an approximate inverse relation to the expected standard deviation of estimates for the stratum. Regression models which cross these primary stratum cannot, therefore, be assumed to have been generated from a simple random sample and additional estimation procedures must be considered.

Judging from what little can be found in the econometric literature to provide guidance on this subject, opinions vary as to what the problems are and, therefore, how to deal with them.[2-7] Central to the discussion is the use of weights in the estimation of econometric models. Weights are discussed in the context of two possible sources of bias which arise when stratified sample survey data are used in a regression analysis.

The first source of bias has to do with the estimate of the regression coefficient. In this context weights are used to compensate for the difference in sampling rates across strata as well as to adjust for nonresponse and post-stratification. The weights are generated as part of the sampling process and are used with the sample data to provide unbiased estimates of the characteristics of a finite population such as nursing homes as defined in the NNHS. In the literature there is little agreement on the need for this type of weighting when the objective is to estimate a regression model and most econometric texts do not discuss this issue. I suspect that the decision depends on whether the objective of the model is to estimate the parameters of some underlying universal law or alternatively to estimate the parameters of some finite population relationships. Further development of these ideas is beyond the scope of this paper. Both weighted and unweighted models, however, were estimated to provide some empirical evidence on the results of the alternative procedures.

The second source of bias has to do with the estimated variance of the regression coefficient. In econometric texts the most widely accepted need for weights with regression is the generalized least squares adjustments suggested to overcome the problems of heteroscedasticity and autocorrelation. These problems arise when the variance of the error terms are not all equal and/or the error terms are correlated. In these cases the observation in the model are transformed by weights which are designed to produce the desired properties of homogeneous variances and zero covariances in the error structure. Heteroscedasticity and autocorrelation are frequently encountered in dealing with cross-sectional data such as the NNHS where the sample has been chosen from groups stratified to reflect the different degrees of variability in the characteristics of interest. In practice, however, it is often quite difficult to obtain the information needed to properly transform the error matrix and at best approximate procedures are used. Fortunately, an alternative approach has been developed.

McCarthy, Kish, Frankel and others have suggested the technique of balanced repeated replications (BRR) as a viable approach to estimating the variance of regression coefficients when the data are from a complex sample survey design.[2, 7, 8-12] To implement this approach subsamples which replicate the sample design are formed by randomly selecting observations from the total sample, with each replicate subsample having approximately half the observations. The essence of the approach is quite straightforward. The regression model of interest is run on the total sample and on each half sample replicate. The estimated variance of a regression coefficient can then be calculated by the following formula:[13]

$$S^2_{\hat{B}} = \frac{1}{r} \sum_{i=1}^{r} (B_i - \hat{B})^2$$

$\hat{B}$ is the regression coefficient estimated from the total sample

B_i is the regression coefficient estimated using half sample replicate i r is the number of half-sample replicates utilized

A detailed exposition on features of the BRR approach is found in a NCHS (1966) publication written by McCarthy.[8] He has found that under quite general conditions analytical unbiased estimates of the standard errors of regression coefficients can be generated using this method.

Outline of the Paper

The purpose of this paper is to compare the empirical results of using BRR regression in analyzing the cost data from the 1973-74 NNHS to the simpler approach of ignoring all or part of the sample design and using ordinary regression. Two basic cost models are estimated using what amounts to four alternative approaches. Approach 1 takes into account all the survey design features by using the sample weights with BRR regression. This is considered the most complete approach and will serve as a point of reference. Approach 2 uses the sample weights but applies ordinary regression. Approach 3 uses BRR regression but does not use the sample weights. Approach 4 disregards the complex survey design and estimation procedures, and treats the data as though they were from a simple random sample, i.e., ordinary regression is applied with no sample weights.

The differences in results between approach 1 and the alternative approaches will be discussed with respect to the various estimation procedures and their potential policy implications. Similar results between approaches 1 and 2 would tend to show that the design effect is small and not a major concern in the estimation process. Similar results between approaches 1 and 3 would indicate that by applying the sample weights with ordinary regression both the type of bias mentioned in the introduction have been eliminated and it would not be necessary to use BRR regression. The comparison of approach 1 and 4 will show the potential differences which occur when the alternative assumptions behind the need for weights (to overcome the first type of bias mentioned in the introduction) are made.

Nursing Home Cost Models

Two major studies of cost functions for nursing homes have appeared in the literature and these serve as the basis for the two models estimated here. The first model is a version of the stock-flow model suggested by Skinner and Yett.[14] It is a multiplicative model estimated in log linear form, its dependent variable is total cost and it features two dimensional output. The second model is the hyperbolic version from the set of "classical" cost functions estimated by Ruckline and Levey.[15] It is an additive model estimated lineraly, its dependent variable is average cost (total cost per resident day) and it features the inverse of the number of beds in the facility as the size/capacity measure. The cost models are estimated for the subset of nursing homes certified by Medicare, i.e., those facilities in the 1973-74 NNHS which were certified by both Medicare and Medicaid or by Medicare only. The complete list of independent variables used in the cost models and a brief description of their purpose is given in table 1. It should be noted that because of BRR program limitations each model is limited to 12 independent variables. While this limitation points up an obvious area for future work it does not seriously hinder the model specifications. Each model includes the variable categories generally considered to be important in the estimation of cost functions for health related facilities, i.e., output or capacity, occupancy rate, scope of facility services, resident characteristics, quality indicator, factor prices, and standardizing measures for ownership and location.

Table 1. Variables used in cost model estimates and their description.

Variable	Description
ADM	ADM is the number of persons admitted to the facility in 1972. This variable is used as one of the two output measures in the total cost model. For a U-shaped average cost curve, the coefficient estimate must be positive.
LNRD	LNRD is the natural log of the total number of resident days of care provided in 1972. This variable is one of the two output measures in the total cost model. For a U-shaped average cost curve, the coefficient estimate must be greater than zero and less than 1.
BEDSINV	BEDSINV is the inverse of the number of beds in the facility in 1972. This variable is used as the size/capacity measure in the average cost model. A positive coefficient estimate indicates possible economies to scale.
OR	OR is the average occupancy rate for the facility in 1972. This variable is included in both the total and average cost model to account for capacity utilization. For both models a negative coefficient estimate is expected.
PROFIT	PROFIT is 1 if the facility is a proprietary facility and 0 if it is a voluntary or government facility. This variable is included in both the total and average cost model to account for the effect of proprietary vs. nonproprietary control.
NE NC S	NE, NC, S are 1 or 0 dummy variables used to represent the four geographic regions--North East, North Central, South and West (this last region is left out as is required when dummy vari-

ables are used). These variables are included in both the total and average cost model to account for regional variation in factor prices otherwise unaccounted for.

PBCERTMR
PBCERTSN
PBCERTIC — PBCERTMR, PBCERTSN, PBCERTIC are the proportion of beds in the facility certified as Medicare beds, Medicaid skilled nursing care beds, Medicaid intermediate care beds, respectively. These variables are included to account variation in the scope of services the facility provides. Because of BRR program limitation, only PBCERTMR and PBCERTIC were included in the total cost model.

KATZAB — KATZAB is the percent of residents in the facility who are totally independent in various activities of daily living or are dependent only in bathing as measured by the Katz patient assessment scale.[16,17] This variable is used in both the total and average cost model to account for resident mix characteristics. For both models a negative coefficient estimate is expected.

AHSLPN — AHSLPN is the average hourly salary of licensed practical nurses working in the facility. This variable is used in both the total and average cost model to account for variation due to wage levels. For both models a positive coefficient estimate is expected.

RNSHF — RNSHF is the number of shifts per day that the facility has a registered nurse on duty. This variable can take a value of 0 to 3. It is included in both the total and average cost model as a quality indicator. For both models a positive coefficient estimate is expected.

Analysis of Results

The results from estimating the total cost model and the average cost model using the alternative estimation approaches are presented in tables 2 and 3 respectively. Columns I and II of these tables comprise approach 1. Columns I and III comprise approach 2, columns IV and V comprise approach 3 and columns IV and VI comprise approach 4. These results are used in tables 2-A, 2-B, 3-A and 3-B to compare the effect of the alternative estimation approaches on the coefficient estimates, the standard error estimates, and the resulting inferential statistics.

Column I of tables 2-A and 3-A gives the ratio of weighted to unweighted coefficient estimates for the two models. While on average the ratio for each model indicates only about a 3 percent difference between the weighted and unweighted coefficient estimates, the variation of these ratios is considerable and in neither model is the weighted coefficient always higher or lower than the unweighted coefficient. For the total cost model the weighted coefficients ranged from 28 percent higher to 25 percent lower than the unweighted coefficients (table 2-A). For the average cost model the range was even wider with the weighted coefficients ranging from 74 percent higher to 50 percent lower than the unweighted coefficients (table 3-A). From a policy point of view it seems likely that such differences would be of concern. For example, using the total cost model with weights implies nearly a 2 percent greater increase in total costs with an increase of one in the number of shifts with a registerd nurse on duty than would be expected using the unweighted version. Also for the total cost model the weighted coefficients imply that being a proprietary facility results in about a 4 percent smaller decrease in total costs than would be indicated by the unweighted coefficient. Similar examples are available from the average cost model. For the weighted version a $1 increase in the average hourly salary of the LPNs increases average costs by $.93 less than for the unweighted version and being in the Northeast implies a $1.75 difference in impact between the weighted and unweighted versions.

The impact of the weights on the standard error estimates of the two models is also shown in tables 2-A and 3-A. Column II shows the weighted to unweighted standard error ratios when the BRR approach was used and column III shows these ratios when the simple random sample (SRS) approach was used. For the total cost model the weighted BRR standard errors average 22 percent higher than the unweighted BRR standard errors with the weighted estimates ranging from 105 percent larger to 11 percent lower than the unweighted estimates. The ratios of SRS standard errors for this model were considerably less variable. The weighted SRS standard errors ranged from 11 percent larger to 9 percent smaller than the unweighted SRS standard errors and were on average only 1 percent larger than the unweighted estimates. For these comparisons the results from the average cost model are similar to those for the total cost model in that the variability of the BRR standard error ratios is considerable larger than that for the SRS standard error ratios. The weighted BRR standard errors averaged 18 percent higher than the unweighted BRR standard errors with the weighted estimates ranging from 165 percent higher to 63 percent lower than the unweighted estimates. The weighted SRS standard errors averaged 5 percent lower than the unweighted SRS with the weighted estimates ranging from 52 percent higher to 44 percent lower than the unweighted estimates.

Other than the smaller differences between weighted and unweighted standard error estimates using the SRS approach rather than the BRR approach, there appears to be no systematic indication of whether weighted or unweighted estimates will be larger. For variables in both models examples can be found where the weighted standard error estimate exceeded the unweighted estimate when the BRR approach was used and was less than the unweighted estimate when the SRS approach was used. It is not possible, therefore, to infer that either weighted or unweighted standard errors would always be larger regardless of the estimation approach nor is it possible to infer that the BRR standard error ratio will be greater than one if the SRS standard error ratio is greater than one.

The design effects (the ratio of the BRR standard error to the SRS standard error) for the variables in the two models are presented in columns IV (weighted) and columns V (unweighted) of tables 2-A and 3-A. For both models the variability of the design effect is considerable. For the total cost model the weighted standard error design effect averaged 1.35 but ranged from 2.47 to .51. The unweighted design effect averaged 1.18 and range from 2.03 to .47. For the average cost model the weighted standard error design effect averaged 1.18 and ranged from 2.11 to .63 while for the unweighted version it averaged 1.05 and range from 1.75 to .61.

Tables 2-B and 3-B give the "t" statistics for the independent variables of the two cost models under the alternative estimation approaches. Column I is calculated by dividing the weighted coefficient by the weighted BRR standard error. Column II is calculated by dividing the weighted coefficient by the weighted SRS standard error. Column III is calculated by dividing the unweighted coefficient by the unweighted BRR standard error. Column IV is calculated by dividing the unweighted coefficient by the unweighted SRS standard error. The "t" statistics are marked to signify their level of significance. Those coefficients not marked can be considered significant only at levels below those normally acceptable. It should be noted that the degrees of freedom for these tests of significance depend on the estimation approach used. For the BRR approach the degrees of freedom are equal to the number of half samples (20) used in the analysis while for the SRS approach the degrees of freedom are the number of cases less the number of independent variables and the constant (603).

For the total cost model (table 2-B) the effect of the alternative estimation approaches on the significance tests is small. Each coefficient is significant at least at the .90 level regardless of the approach used and in many cases the level is much higher. Shifts do occur, however, in the level of confidence at which variables are considered significant when the alternative approaches to estimating standard errors used. The effect of these shifts becomes more noteworthy when the results for the average cost model (table 3-B) are considered. Taking into account both the weighted and unweighted versions of this model there are examples where the effect was such that when the SRS "t" statistic was significant at the 1 percent level the BRR "t" statistic was not significant and, alternatively, when the SRS "t" was not significant the BRR "t" was significant at the 5 percent level. These results indicate that to ignore the effect of the sample design on the standard error estimates can lead to erroneous inferences. The most striking example from a policy point of view is the coefficient (weighted or unweighted) for BEDSINV in the average cost model. Using SRS estimation procedures the positive sign of the coefficient and its "t" statistic indicate economies to scale, a result which, in this model, is shown to be questionable when the effect of the survey design is accounted for by using the BRR technique.

Summary and Conclusions

This paper compared the empirical results of using BRR regression in estimating two representative nursing home cost models to the simpler approach of ignoring all or part of the 1973-74 NNHS sample design and using ordinary regression. The effects of the alternative estimation procedures on the coefficient and standard error estimates and some potential policy implications were discussed.

The results showed that ignoring the sample design when using the 1973-74 NNHS with regression to analyze nursing home costs will effect both the coefficient and standard error estimates. None of the alternative short-cut methods gave results that consistently approximated the results gotten by using information about the sample design.

The effects of the sample weights on the coefficients and standard errors were analyzed by comparing ratios of weighted to unweighted estimates. For the coefficients and their standard errors, the ratios showed considerable variation with no apparent systematic way of predicting the magnitude of the ratio or whether it would be greater or less than one.

The design effect on the standard errors was analized by comparing the ratios of BRR standard errors to the SRS standard errors for both the weighted and unweighted versions of the two models. In all cases the individual design effects varied to such a large degree that the average design effect for a model could not be used to reliable adjust the SRS standard errors.

Finally, the effect of the sample design on inferences was considered by calculating the "t" statistics for the coefficients of the two models estimated under the alternative approaches. The results showed that the potential policy inferences were in some important instances effected by the estimation approached used. Therefore, if the goal is to estimate the cost function parameters for the finite population of nursing homes as defined in the 1973-74 NNHS the full sample design features must be considered in the estimation process.

Table 2. Summary of selected regression results for total cost model using estimation approaches 1 through 4, Medicare certified nursing homes, 1973-74 NNHS data.

Variable	I Wtd. Coef. Est.	II Wtd. BRR S.E. Est.	III Wtd. SRS S.E. Est.	IV Unwtd. Coef. Est.	V Unwtd. BRR S.E. Est.	VI Unwtd. SRS S.E. Est.
ADM	.00071	.00015	.00010	.00057	.00016	.00009
LNRD	.87923	.06109	.02693	.91416	.05247	.02580
OR	-.00736	.00121	.00091	-.00777	.00113	.00086
PROFIT	-.12937	.05571	.02980	-.17340	.03839	.02849
NE	.35548	.01727	.03404	.42311	.01640	.03488
NC	.19772	.04003	.03569	.20273	.04058	.03561
S	.08006	.03049	.03315	.07891	.02385	.03459
PBCERTMR	.09482	.03481	.02693	.08818	.03336	.02739
PBCERTIC	-.08252	.04020	.03674	-.07797	.02830	.03567
KATZAB	-.00335	.00163	.00066	-.00282	.00127	.00066
AHSLPN	.03770	.01005	.00920	.03726	.00490	.00916
RNSHF	.08007	.01245	.01457	.06241	.01402	.01603

Constant = 4.4107   Constant = 4.1706
R^2 = .80339 R^2 = .82339
N = 616   N = 616

Table 3. Summary of selected regression results for average cost model using estimation approaches 1 through 4, Medicare certified nursing homes, 1973-74 NNHS data.

Variable	I Wtd. Coef. Est.	II Wtd. BRR S.E. Est.	III Wtd. SRS S.E. Est.	IV Unwtd. Coef. Est.	V Unwtd. BRR S.E. Est.	VI Unwtd. SRS S.E. Est.
BEDSINV	127.45	99.111	47.763	162.49	115.84	66.057
OR	- .2826	.0136	.0217	- .2895	.0302	.0213
PROFIT	-3.8678	1.5368	.8557	-4.1790	1.2582	.8427
NE	7.5954	.7556	.9708	9.3447	.6263	1.0258
NC	4.6790	.8715	1.0100	4.4155	.7598	1.0414
S	2.6089	.8755	.9394	2.5209	.6893	1.0141
PBCERTMR	2.3062	.7417	.8166	2.2627	.7574	.8201
PBCERTSN	1.8714	1.0347	.9867	1.3383	1.0250	1.0411
PBCERTIC	-1.9005	.9501	1.0915	-1.7406	.7742	1.1162
KATZAB	- .0576	.0393	.0187	- .0534	.0193	.0226
AHSLPN	.9149	.2168	.2605	1.8434	.5816	.4660
RNSHF	2.0497	.6297	.4050	1.1756	.2374	.2673

Constant = 34.561   Constant = 35.163
R^2 = .35934 R^2 = .36959
N = 616   N = 616

Table 2-A. Ratio comparisons of weighted to unweighted coefficient and standard error estimates and design effects on standard errors (s.e. DEFF) for weighted and unweighted regressions, total cost model.

Variable	I Wtd./ Unwtd. Coef. Est.	II Wtd./ Unwtd. BRR S.E. Est.	III Wtd./ Unwtd. SRS S.E. Est.	IV Wtd. S.E. DEFF	V Unwtd. S.E. DEFF
ADM	1.25	.94	1.11	1.48	1.88
LNRD	.96	1.16	1.04	2.27	2.03
OR	.94	1.07	1.06	1.32	1.31
PROFIT	.75	1.45	1.05	1.87	1.34
NE	.84	1.05	.98	.51	.47
NC	.97	.99	1.00	1.12	1.14
S	1.01	1.27	.96	.92	.69
PBCERTMR	1.08	1.04	.98	1.23	1.22
PBCERTIC	1.06	1.42	1.03	1.09	.79
KATZAB	1.19	1.28	1.00	2.47	1.92
AHSLPN	1.01	2.05	1.00	1.09	.53
RNSHF	1.28	.89	.91	.85	.88
Average	1.03	1.22	1.01	1.35	1.18

Table 3-A. Ratio comparisons of weighted to unweighted coefficient and standard error estimates and design effects on standard errors (s.e. DEFF) for weighted and unweighted regressions, total cost model.

Variable	I Wtd./ Unwtd. Coef. Est.	II Wtd./ Unwtd. BRR S.E. Est.	III Wtd./ Unwtd. SRS S.E. Est.	IV Wtd. S.E. DEFF	V Unwtd. S.E. DEFF
BEDSINV	.78	.60	.72	2.08	1.75
OR	.97	.45	1.02	.63	1.42
PROFIT	.93	1.22	1.02	1.80	1.49
NE	.81	1.21	.95	.78	.61
NC	1.06	1.15	.97	.86	.73
S	1.03	1.27	.93	.93	.68
PBCERTMR	1.02	.98	1.00	.91	.92
PBCERTSN	1.39	1.01	.95	1.05	.99
PBCERTIC	1.09	1.23	.98	.87	.69
KATZAB	1.08	2.04	.83	2.11	1.17
AHSLPN	.50	.37	.56	.83	1.25
RNSHF	1.74	2.65	1.52	1.56	.89
Average	1.03	1.18	.95	1.18	1.05

Table 2-B. "t" statistics using weighted and unweighted BRR and SRS approaches total cost model.

Variable	I Wtd. BRR t Stat.	II Wtd. SRS t Stat.	III Unwtd. BRR t Stat.	IV Unwtd. SRS t Stat.
ADM	4.73***	7.10***	3.56***	6.33***
LNRD	14.39***	32.65***	17.42***	35.43***
OR	6.08***	8.09***	6.88***	9.03***
PROFIT	2.32**	4.34***	4.52***	6.09***
NE	20.58***	10.44***	25.80***	12.13***
NC	4.94***	5.54***	5.00***	5.69***
S	2.63**	2.42**	3.31***	2.28**
PBCERTMR	2.72**	3.52***	2.64**	3.22***
PBCERTIC	2.05*	2.25**	2.76**	2.19**
KATZAB	2.06*	5.06***	2.22**	4.27***
AHSLPN	3.75***	4.10***	7.60***	4.07***
RNSHF	6.43***	5.50***	4.45***	3.89***

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 3-B. "t" statistics using weighted and unweighted BRR and SRS approaches, average cost model.

Variable	I Wtd. BRR t Stat.	II Wtd. SRS t Stat.	III Unwtd. BRR t Stat.	IV Unwtd. SRS t Stat.
BEDSINV	1.29	2.67***	1.40	2.46**
OR	20.78***	13.01***	9.58***	13.60***
PROFIT	2.52**	4.52***	3.32***	4.96***
NE	10.05***	7.82***	14.92***	9.11***
NC	5.37***	4.63***	5.81***	4.24***
S	2.98***	2.78***	3.66***	2.49**
PBCERTMR	3.11***	2.82***	2.99***	2.76***
PBCERTSN	1.81*	1.90*	1.31	1.29
PBCERTIC	2.00*	1.74*	2.25**	1.56
KATZAB	1.47	3.09***	2.77**	2.37**
AHSLPN	4.22***	3.51***	3.17***	3.96***
RNSHF	3.26***	5.06***	4.95***	4.40***

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Notes and References

1. For a more detailed description of the survey design see pp. 16-18 of National Center for Health Statistics: Selected Operating and Financial Characteristics of Nursing Homes, United States: 1973-74 National Nursing Home Survey, Vital and Health Statistics, Series 13, No. 22, DHEW Pub. No. (HRA) 76-1773. Washington, U.S. Government Printing Office, Dec. 1975.

2. Cohen, M.S., Rea, A. and Lerman, R.I. A Micro Model of Labor Supply, BLS Staff Paper 4, Government Printing Office, 1970.

3. Cramer, J.S. Empirical Econometrics, Amsterdam: North-Holland, 1971, 141-143.

4. Hu, T.W. and Stromsdorfer, E.W. A Problem of Weighting Bias in Estimating the Weighted Regression Model, Proceedings of the Business and Economics Section of the American Statistical Association, 1970, 513-516.

5. Klein, L.R. and Morgan, J. Results of Alternative Statistical Treatments of Sample Survey Data. Journal of the American Statistical Association, 46, 1951, 442-460.

6. Klein, L.R. A Textbook of Econometrics. Evanston: Row, Peterson and Company, 1953, 406-415.

7. Roth, J. Application of Regression Analysis to Sample Survey Data: Unpublished memorandum, Social Security Administration, 1971.

8. National Center for Health Statistics: Replication: An Approach to the Analysis of Data From Complex Surveys. Vital and Health Statistics, Series 2, No. 14, Public Health Service Pub. No. 1000, Washington, D.C., U.S. Government Printing Office, April 1966.

9. National Center for Health Statistics: Pseudoreplication: Further Evaluation and Application of the Balances Half-Sample Technique, Vital and Health Statistics, Series 2, No. 31, Public Health Service Pub. No. 1000, Washington, D.C., U.S. Government Printing Office, January 1969.

10. Kish, Leslie, and Frankel, Martin. Balanced Repeated Replications for Standard Errors. Journal of the American Statistical Association, Sept. 1970, Vol. 65, No. 331, 1071-1094.

11. Frankel, Martin. Inferences From Sample Surveys: An Empirical Investigation, Ann Arbor, Michigan: Litho Crafters, Inc., 1971.

12. Simmons, Walt R. and Baird, James T. Jr. Pseudo-Replication In the NCHS Health Examination Survey. Proceedings of the Social Statistics Section of the American Statistical Association, 1968, 19-30.

13. To yield more information and improve the variance estimate the alternative variance formula $S_B^2 = \frac{1}{4r} \sum_{i=1}^{r} (B_i - B_i')^2$ was used in this paper. For this formulation the regression coefficient for each compliment half-sample (B_i') is also calculated.

14. Skinner, D.E. and Yett, D.E. Estimation of Cost Functions for Health Services: The Nursing Home Case. Paper presented at the 40th Annual Meeting of the Southern Economic Association, Atlanta, Georgia, November 1970.

15. Ruchlin, H.S. and Levey, S. Nursing Home Cost Analysis: A Case Study. Inquiry, IX, September, 1972, 3-15.

16. Katz, S., et. al. Studies of Illness in the Aged. The Journal of the American Medical Association, Vol. 185, No. 12, September, 1963, 914-919.

17. Katz, S., et. al. Progress in Development of the Index of ADL. The Gerontologist, Vol. 10, Spring, 1970, Part 1, 20-30.