© Journal of Business & Economic Statistics, Vol. 2, No. 2, April 1984

Grouping Tests for Misspecification: An Application to Housing Demand

John S. Greenlees and Kimberly D. Zieschang
Division of Price and Index Number Research, U.S. Bureau of Labor Statistics, Washington, DC 20212

We consider several grouping tests for regression misspecification, with reference to housing-demand function estimation. We compare three existing test procedures, demonstrate modifications necessary in most applications, and propose a fourth test to distinguish between two categories of potential specification error. The test procedures are evaluated in artificial simulations of alternative errors. Finally, we apply the tests to FHA home purchase data. We reject the hypothesis that household and grouped regressions differ only by sampling error or random mismeasurement of household income or price. Our results have implications for choices among test procedures and interpretations of previous housing-demand analysis.

KEY WORDS: Specification error tests; Housing demand; Aggregation bias; Errors-in-variables.

1. INTRODUCTION

As noted by a number of recent authors, empirical analyses of the demand for housing have failed to agree on the extent of consumer responsiveness to price and income changes. In particular, "micro" studies using household data have produced lower income-elasticity estimates than "grouped" studies based on metropolitan-level averages. Although several explanations have been offered, the paradox remains unresolved.

Also in recent years, econometricians have developed statistics to test for specification error through the comparison of coefficients obtained from ungrouped and grouped regressions. Statistically significant deviations between the two estimated coefficient vectors are taken as evidence of either a misspecified regression relationship (caused by, e.g., an explanatory variable excluded or measured with error) or an inappropriate aggregation method (such as grouping according to the value of the dependent variable). Two alternative test statistics were presented by Feige and Watts (1972) and Farebrother (1979). Polinsky and Ellwood (1979) derived a third statistic by adapting the specification test procedure of Hausman (1978) to the grouping case.

This article is an empirical investigation into the housing-demand paradox, making use of several grouping tests for misspecification. We begin in Section 2 with a brief review of our demand-function specification, which was chosen to approximate those employed by previous researchers, and of the micro/grouped controversy that surrounds it. We follow in Section 3 with a review of the Farebrother, Feige-Watts, and Hausman tests and a discussion of the relationships among them. In the process we derive an important practical restriction, namely that, in applying any of the three tests, the researcher must confine his attention to parameter subvectors corresponding to variables that exhibit both within-group and between-group variation. This restriction and the consequent test modifications necessary in most applications have not been noted previously.

The aforementioned grouping tests are concerned with the null hypothesis that the micro and grouped models are both specified correctly. Under this hypothesis, the two coefficient vectors differ only as a result of random sampling error, and rejections of the null hypothesis yield no direct information about which set of estimates is more nearly correct. However, one commonly offered explanation for the micro/grouped paradox in housing demand has been a household-level errors-in-variables problem caused by imprecise measurement of income, price, or both. This type of misspecification has straightforward and well-known results. As long as the measurement errors are uncorrelated across observations and are independent of both the regression disturbances and the true values of the explanatory variables, only the micro-parameter estimator is inconsistent. Therefore, in Section 4 we propose an asymptotic test of this explanation; that is, of the hypothesis that the grouped regression asymptotically satisfies the conditions of the standard linear model, and that the relative magnitude of the micro and grouped squared residual vectors can be explained by a combination of sampling error and micro-level misspecification. When this null hypothesis (which we will set forth more rigorously in Section 4) is also rejected, there can be no grounds for concluding that the grouped coefficients are consistent or are more accurate than the micro estimates. Consequently, while the paradox still remains, other explanations—notably "aggregation bias" or other grouped-level misspecification—should be given relatively greater credence.

Following these clarifications and extensions of grouping test procedures, in Section 5 we employ the tests in an analysis of a nationwide sample of home purchases insured by the Federal Housing Administration (FHA). Using Farebrother's grouping test, we demonstrate that household-level and city-level estimates of housing-demand parameters differ to a statistically significant degree. Furthermore, based on our own asymptotic test of residuals, and contrary to the arguments of some previous authors, we conclude that micro-level misspecification is insufficient to explain the differences between the micro and grouped regressions. Our results have implications both for the choice among alternative test procedures and for the interpretation of previous analyses of housing demand.

2. THE MICRO/GROUPED PARADOX IN HOUSING-DEMAND ESTIMATION

Our specification of the housing-demand function closely follows that used in Polinsky (1977, 1979) and in Polinsky and Ellwood (1979). Total expenditure on housing services C is assumed to be a log-linear function of income I, the price of housing services p_H, and an index p_o of the prices of all other goods. In the form to be estimated, we have, for the ith household,

$$\log(C_i/p_{oi}) = b_o + (1 + b_p)\log(p_{Hi}/p_{oi}) + b_I\log(I_i/p_{oi}) + \epsilon_i, \qquad i = 1, \ldots, M, \quad (1)$$

where b_p and b_I are, respectively, the price and income elasticities of housing demand. The random disturbances ϵ_i are independently distributed with zero mean and constant variance.

Again following Polinsky and Ellwood, we will apply this specification to a sample of new house purchases insured by FHA during calendar year 1969. Appendix A contains a summary of the derivation of (1), a description of the FHA data base, and a review of the production-function-based methodology used to estimate p_H. We emphasize that the specification is obviously a highly simplified one. Aside from possible errors in measurement and functional form, Equation (1) may suffer from a failure to acknowledge the effects of demographic factors, the simultaneity of housing consumption and tenure choice, or the implicit design of the FHA mortgage sample. (A review of the literature on many of these issues can be found in Mayo 1981.) We choose this formulation because it has provided the empirical context for much of the debate on the relative merits of micro and grouped estimates of housing-demand parameters. Economists have offered two primary explanations for the lower income elasticities yielded by household-level regressions. On the one hand, if aggregation by city produces an implicit grouping by the dependent variable, the metropolitan-level coefficients will tend to be biased away from zero. Although the aggregation bias argument has most often been used as a criticism of grouping by census tract, city-level grouping may also cause problems, particularly if the regression function suffers from any additional specification errors. Discussions of aggregation bias in housing demand can be found in Maisel, Burnham, and Austin (1971), Smith and Campbell (1978), and Gillingham and Hagemann (1983).

Micro and grouped regressions can also differ because of measurement error at the household level. Most obviously, if reported or "current" income is a poor proxy for "permanent" income, the income elasticity of demand is likely to be underestimated. This problem has been discussed by numerous authors, including Lee (1968), Rosen (1979), and Goodman and Kawai (1982). By contrast with aggregation bias, the errors-in-variables bias is alleviated by grouping observations.

The study by Polinsky and Ellwood (1979) focuses on household-level measurement error as it affects the estimation of Equation (1). It is argued there that mismeasurement of housing price, through the use of a city-level price term, biases both the b_p and b_I coefficients in household-level regressions. The authors demonstrate that the divergence between micro estimates and grouped estimates is reduced by inclusion of their household-specific, production-function-based estimate of p_H. The remaining difference is attributed to mismeasurement of permanent income in the FHA sample. Therefore, they conclude that their metropolitan-level regressions are the preferred source of information on the size of income and price elasticities.

The primary purpose of this article is to reevaluate Polinsky and Ellwood's conclusions, using test procedures developed in Sections 3 and 4 here. We first consider the hypothesis that the micro and grouped coefficients obtained from our FHA sample differ only as a result of random sampling error. We then ask whether any statistically significant divergence can be explained by household-level measurement error or by any other error that distorts the micro regression while leaving the grouped regressions in (asymptotic) compliance with the standard linear regression model.

3. THREE GROUPING TESTS FOR MISSPECIFICATION

In this section we discuss three specification tests that have appeared in the econometric literature. The null hypothesis underlying each test is summarized by the general micro regression specification,

$$y = X\beta + \epsilon, \qquad \epsilon \sim N(0, \sigma^2 I_n), \quad (2)$$

where y and X are $n \times 1$ and $n \times k$ matrices of observations, β is a $k \times 1$ vector of parameters, and ϵ is an $n \times 1$ vector of independent normal disturbances. The matrix X is assumed to have full column rank k.

The grouped version of the model is obtained through an $m \times n$ matrix G^*, which transforms the micro observations into means of m groups, each multiplied by the square root of the group sample size. That is,

$$G^* = \begin{bmatrix} g_1^* & 0 & \cdots & 0 \\ 0 & g_2^* & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & g_m^* \end{bmatrix}. \quad (3)$$

Each element of the row vector $g_j^*(j = 1, \ldots, m)$ equals $1/\sqrt{n_j}$, where n_j is the number of observations in the jth group.

The grouped model is then given by

$$y_2 = X_2\beta + \epsilon_2, \qquad \epsilon_2 \sim N(0, \sigma^2 I_m), \quad (4)$$

where the subscript 2 indicates premultiplication by G^* and under the crucial assumption that the grouping process is independent of the disturbance term ϵ.

Under the null hypothesis, the micro and grouped least squares estimators of β, given by

$$\hat{\beta} = (X'X)^{-1}X'y \quad (5)$$

and

$$\hat{\beta}_2 = (X_2'X_2)^{-1}X_2'y_2, \quad (6)$$

respectively, are both unbiased and consistent, and the grouping tests are based on the divergence beween these two estimated vectors. We will briefly review the three test procedures in turn.

3.1 The Farebrother Test

Farebrother (1979) shows that a Chow test for misspecification of the micro model is available in the form of a test of the restriction $\gamma = 0$ in

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ X_2 & X_2 \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}, \quad (7)$$

with $y_1 = Fy$, $X_1 = FX$, and $\epsilon_1 = F\epsilon$. The rows of the $(n - m) \times n$ matrix F are defined to be a set of orthonormal solutions to the equation system $G^*c = 0$; that is, we have

$$FF' = I_{n-m}, \qquad G^*F' = 0, \qquad X'X = X_1'X_1 + X_2'X_2. \quad (8)$$

Farebrother uses a result derived earlier (Farebrother 1976) to show that the test statistic can be written in the form

$$B_F = [\delta' D^{-1}\delta/k]/[(Q - \delta' D^{-1}\delta)/(n - 2k)] \sim F(k, n - 2k), \quad (9)$$

under the null hypothesis of no misspecification, where

$$\delta = \hat{\beta}_2 - \hat{\beta}, \quad (10)$$

$$Q = (y - X\hat{\beta})'(y - X\hat{\beta}), \quad (11)$$

and

$$D = [(X_2'X_2)^{-1} - (X'X)^{-1}]. \quad (12)$$

In (9), the test statistic does not require construction of the F matrix; the test may be applied by estimating the micro and grouped models (2) and (4). Notice that the denominator of B_F is equal to the mean squared error (MSE) or estimate of σ^2, which would be obtained from the unconstrained model (7). We will henceforth refer to this estimate as $\hat{\sigma}_U^2$.

3.2 The Feige-Watts Test

Feige and Watts (1972) demonstrate that under the null hypothesis, δ, the difference between the estimated parameter vectors, has zero mean and covariance matrix $\sigma^2 D$. They then construct the statistic

$$B_{FW} = [\delta' D^{-1}\delta/k]/[Q_2/(m - k)] \sim F(k, m - k), \quad (13)$$

where

$$Q_2 = (y_2 - X_2\hat{\beta}_2)'(y_2 - X_2\hat{\beta}_2), \quad (14)$$

the sum of squared residuals from the grouped regression model (4). Feige and Watts show (1972, p. 347) that the quadratic forms $\delta' D^{-1}\delta$ and Q_2 are independently distributed; hence, the statistic B_{FW}, like B_F, follows an exact F distribution.

3.3 The Hausman Test

Hausman (1978) has proposed an asymptotically chi-squared statistic that can be used to test for specification error in a broad range of situations. Polinsky and Ellwood adapt and apply the test, with inconclusive results, in the housing-demand context. The statistic uses the mean squared errors

$$\hat{\sigma}^2 = Q/(n - k) \quad (15)$$

and

$$\hat{\sigma}_2^2 = Q_2/(m - k), \quad (16)$$

and the statistic is

$$B_H = \delta' D_H^{-1}\delta \overset{A}{\sim} \chi^2(k), \quad (17)$$

where the symbol $\overset{A}{\sim}$ denotes "is asymptotically distributed as," and

$$D_H = \hat{\sigma}_2^2(X_2'X_2)^{-1} - \hat{\sigma}^2(X'X)^{-1} \quad (18)$$

is a consistent estimator of the difference between the covariance matrices of $\hat{\beta}_2$ and $\hat{\beta}$. (We wish to thank A. Mitchell Polinsky for graciously providing unpublished computational details.)

3.4 A Restriction in Application

In Section 5 and Appendix B, we will apply the three grouping tests in the context of Equation (1). Before we do so, however, it must be noted that in most econometric applications the test statistics cannot be computed in the forms given by Equations (9), (13), and (17). The presence of certain common classes of explanatory variables in the matrix X requires a restructuring of the matrices in B_F, B_{FW}, and B_H, and a reduction in the implied degrees of freedom. Since no previous author has pointed out the restriction, and since it was overlooked in the one previous application to housing demand equations, we will briefly outline the derivation of the modified test statistics. Of central importance to all three of the foregoing tests is the matrix D. Using Equations (8) and (12) we can write

$$D = (X_2'X_2)^{-1}(X_1'X_1)(X'X)^{-1}, \qquad (19)$$

so that X_1 must have rank k in order that D be nonsingular. This restriction is most obviously violated if any column of $X_1 = FX$ consists entirely of zeros—that is, if any column of X is orthogonal to all rows of F. But by the conditions (8), all rows of F are orthogonal to each row of G and hence to any linear combination of the rows of G^*. A necessary condition, therefore, for the nonsingularity of D is that no column of the microregressor matrix X be expressible as a linear combination of the rows of the grouping matrix.

The implication is that grouping tests cannot be applied to variables that are constant within each group. An obvious example of such a variable is a constant term, although in that case the problem can be avoided by normalizing all regression variables to equal deviations from means. (This normalization apparently explains the identifiability of the B_{FW} values reported by Feige and Watts.) Other examples would be group identifiers, along with any economic or demographic variables measured at the group level.

Fortunately, it remains possible to test coefficients corresponding to variables which do exhibit within-group variation. (We are indebted to Christopher Sims for pointing this out to us.) Assume that X is arranged in such a way that only the first $k - j$ variables are constant within groups. Define the $k \times j$ matrix S as

$$S = \begin{bmatrix} 0 \\ I \end{bmatrix}. \qquad (20)$$

Let

$$\bar{\delta} = S'\delta \qquad (21)$$

and

$$\bar{D} = S'DS. \qquad (22)$$

Then it can be shown that a modified Farebrother statistic is given by

$$B_F' = [\bar{\delta}'\bar{D}^{-1}\bar{\delta}/j]/[(Q - \bar{\delta}'\bar{D}^{-1}\bar{\delta})/(n - k - j)] \sim F(j, n - k - j). \qquad (23)$$

Similarly, we can derive

$$B_{FW}' = [\bar{\delta}'\bar{D}^{-1}\bar{\delta}/j]/[Q_2/(m - k)] \sim F(j, m - k). \qquad (24)$$

It remains to consider the statistic B_H. In finite samples, $\hat{\sigma}_2^2$ and $\hat{\sigma}^2$ will be equal only by chance, so D_H will not generally be a singular matrix. Thus, Polinsky and Ellwood were able to compute and report values of B_H derived from their FHA sample. (We emphasize that it is for reasons of completeness and clarification that we discuss their Hausman test application, which they report in a footnote (1979, p. 203). Polinsky and Ellwood do not base any of the conclusions of their paper on their Hausman test results.) However, the Hausman test is valid only asymptotically, and as m and n become large, $\hat{\sigma}_2^2$ and $\hat{\sigma}^2$ will each converge to σ^2, and D_H will approach the matrix $\sigma^2 D$. Again, therefore, the grouping test must be confined to those coefficients associated with variables that are not constant within groups. The appropriately modified Hausman test statistic is given by

$$B_H' = \bar{\delta}'[\hat{\sigma}_2^2 S'(X_2'X_2)^{-1}S - \hat{\sigma}^2 S'(X'X)^{-1}S]\bar{\delta} \overset{A}{\sim} \chi^2(j), \qquad (25)$$

where once more the matrix S is used to select rows and columns corresponding to testable coefficients.

3.5 Comparison of the Three Procedures

The F statistics, B_F and B_{FW} have the same numerator, but in most applications the Feige-Watts statistic will have many fewer denominator degrees of freedom. We therefore expect that the Farebrother test will generally be the more powerful of the two. The exceptions would occur where the null hypothesis is violated in such a way as to primarily distort the micro regression, possibly leaving $\hat{\sigma}_2^2$ smaller than $\hat{\sigma}_U^2$ and B_{FW} greater than B_F.

The relationship between B_H and the previous two test statistics may be seen by noting that the consistency of D_H is unaffected by the replacement of either $\hat{\sigma}_2^2$ or $\hat{\sigma}^2$ by another consistent estimator of σ^2. (This point is made in another context by Hausman 1978, p. 1,267). For example, if $\hat{\sigma}^2$ is replaced by $\hat{\sigma}_2^2$ in (18), the Hausman statistic reduces to the product of B_{FW} and k. If both $\hat{\sigma}_2^2$ and $\hat{\sigma}^2$ are replaced by $\hat{\sigma}_U^2$, B_H becomes equal to kB_F. Still another Hausman statistic can be obtained if $\hat{\sigma}^2$ is used in both terms of (18). In any of these forms, B_H will be asymptotically $\chi^2(k)$ under the null hypothesis. In finite samples, of course, the distribution of B_H is unknown. Particularly when $m - k$ is small, $\hat{\sigma}_2^2$ may differ greatly from $\hat{\sigma}^2$, and D_H may not even be a positive semidefinite matrix.

In Appendix B we present the results of an exploration into the finite sample performance of the three grouping tests. Several types of misspecification considered in the housing demand literature are simulated using Equation (1) and our FHA data base discussed in Appendix A. Our conclusions are, first, that the Hausman statistic B_H' does not reach its asymptotic $\chi^2(2)$ distribution under the null hypothesis of no misspecification with 34 groups and 11,978 observations. In addition, we found that the Farebrother test was generally more powerful than the Feige-Watts test against the alternative hypotheses considered. Also, in accord with our expectations, the Farebrother test's margin of superiority was especially noticeable in situations where grouping the data aggravated, rather than mitigated, the specification error. Interestingly, Feige and Watts developed their test statistic specifically in order to identify possible grouping bias. It is reasonable to expect that some of their conclusions would have been altered had they used the more powerful Farebrother procedure.

4. A SPECIFICATION TEST FOR THE GROUPED MODEL

When one of the foregoing grouping tests indicates a significant difference between micro and grouped regression coefficients, the source of the misspecification is not identified. In particular, it remains undetermined which of the two estimates of β is more nearly accurate. To approach this question, we present a test of the null hypothesis that the grouped regression model satisfies the specification (4). When this null hypothesis is accepted, we conclude that the parameter divergence is explainable by random sampling error in combination with a broad class of micro-level specification errors that bias $\hat{\beta}$ but leave $\hat{\beta}_2$ consistent. Examples are random errors-in-variables and excluded regressor variables that average zero at the group level. Conversely, rejection suggests the presence of measurement errors that are correlated within groups or other problems leading to aggregation bias in $\hat{\beta}_2$.

To proceed, we utilize the notation of Section 3 but replace the conditions (2) and (4) with the assumption that the distributions of the micro disturbances ϵ_i obey sufficient regularity conditions that

(i) $\operatorname*{plim}_{n\to\infty} \epsilon'\epsilon/n = \sigma^2$

and

(ii) $\epsilon_2 \overset{A}{\sim} N(0, \sigma^2 I_m)$,

where σ^2 is a finite value. We also assume

(iii) $\operatorname*{plim}_{n\to\infty} X_2'\epsilon_2/n = 0$,

and

(iv) $\operatorname*{plim}_{n\to\infty} X_2/\sqrt{n} = A$,

where A is a finite matrix of rank k.

Assumptions (ii), (iii), and (iv) together ensure that $\hat{\beta}_2$ is consistent:

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_2 = \beta + \operatorname{plim}(X_2'X_2/n)^{-1}(X_2'\epsilon_2/n) = \beta. \quad (26)$$

We propose to examine the size of the sum of squared grouped regression residuals Q_2. We can write

$$Q_2 = (y_2 - X_2\hat{\beta}_2)'(y_2 - X_2\hat{\beta}_2) = \epsilon_2'M_2\epsilon_2, \quad (27)$$

where $M_2 = I - X_2(X_2'X_2)^{-1}X_2'$. Using (iv) we see that M_2 converges to a constant idempotent matrix with rank $m - k$, and using (ii) we obtain

$$Q_2/\sigma^2 \overset{A}{\sim} \chi^2(m - k). \quad (28)$$

It remains to consistently estimate σ^2. Using (11) and (15),

$$\begin{aligned}\hat{\sigma}^2 &= (y - X\hat{\beta})'(y - X\hat{\beta})/(n - k) \\ &= \epsilon'\epsilon/(n - k) \\ &\quad - (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta)/(n - k). \quad (29)\end{aligned}$$

By assumption (i), the first term above converges to σ^2 as $n \to \infty$, and we know that $\hat{\beta}_2$ is a consistent estimator of β. We will also assume that $X'X/n$ converges to a constant matrix. Then

$$\operatorname*{plim}_{n\to\infty}[\hat{\sigma}^2 + (\hat{\beta} - \hat{\beta}_2)'(X'X/n)(\hat{\beta} - \hat{\beta}_2)] = \sigma^2, \quad (30)$$

and the statistic

$$E = Q_2/[\hat{\sigma}^2 + \delta'(X'X/n)\delta] \overset{A}{\sim} \chi^2(m - k), \quad (31)$$

under the null hypothesis that (i) through (iv) hold.

Assumptions (i) and (ii) state that the asymptotic variance of each grouped disturbance is equal to the limiting value of the mean squared micro disturbance. The range of conditions under which these assumptions hold, and thus the information content of the test statistic, can best be indicated through several examples.

First, consider a situation where the observed value of the jth regressor X^j is equal to the true value Z^j plus an error V^j. Furthermore, micro values of V^j are independently normally distributed with variance σ_v^2, so that the micro residual $\epsilon = \eta - V^j\beta^j$, where β^j is the jth element of β and η is the usual well-behaved regression disturbance with mean zero and variance σ_η^2. Since ϵ is correlated with X_j, $\hat{\beta}$ is biased and inconsistent. However, $\hat{\beta}_2$ is consistent, and $\epsilon'\epsilon/n$ converges to $\sigma_\eta^2 + (\beta^j)^2\epsilon_v^2$, which is also the variance of the grouped disturbances ϵ_2. The null hypothesis underlying (31) is thus satisfied. This is a typical framework assumed for the analysis of errors-in-variables problems such as the permanent income issue in housing demand estimation.

By contrast, we expect our null hypothesis to be violated when the disturbances within a group are correlated. This could occur because of an implicit grouping by the value of the dependent variable, because of a random group-specific disturbance in an error-components model, or because ϵ includes an unobserved regressor whose value is constant or correlated within groups. For example, if an incorrect inter-area price index p_o' is used in the place of p_o in equation (1) above, the micro disturbance will include the term $(b_p + b_I)\log(p_o'/p_o)$, which is constant within each group. In this situation, the expectation of E can be shown to increase without bound as $n \rightarrow \infty$. Both $\hat{\beta}_2$ and $\hat{\beta}$ will be inconsistent and there is no rigorous basis on which to choose between them.

Two qualifying comments are in order with respect to our null hypothesis. First, if we relax (ii) to allow the grouped disturbances ϵ_2 to be heteroscedastic, $\hat{\beta}_2$ remains consistent, although the standard errors estimated by applying OLS to (4) are biased. In this case the E statistic will not be chi-squared even asymptotically. Modifications to our test, presumably incorporating two-stage procedures to estimate the variances of the m grouped residuals, are left to further research. In Simulation III of Appendix B, we examine such a heteroscedastic model, where the variance of each city-level disturbance depends on the dispersion of housing prices within the city. The results of the simulation indicate that this particular case of heteroscedasticity has little effect on the distribution of E.

It is also true that certain patterns of within-group correlation of the disturbances ϵ_i could violate our hypotheses (i) and (ii) without necessarily destroying any desirable asymptotic qualities of the grouped model. Notice that the variance of the ith element of ϵ_2 depends on the variances of the associated n_i micro disturbances and also on the $n_i(n_i - 1)$ covariances among them. If the sum of these disturbance covariances is of order less than n_i (as in the earlier example of random measurement error, where the covariances are all zero), our null hypothesis is satisfied. If the sum has order greater than n_i (as in the case of an excluded group-specific price variable, where the disturbances in group i have a common covariance and the sum has order n_i^2), our null hypothesis is violated, $\hat{\beta}_2$ is inconsistent, and the variance of ϵ_2^i has no finite limit. However, if the micro disturbances in group i have covariances whose sum is of order n_i, while disturbances in different groups are uncorrelated, the E statistic will not be distributed as in (31), although the grouped model satisfies the ideal OLS conditions asymptotically. Such a situation could arise, for example, from first-order serial correlation processes, which are unlikely to be important in cross-section data such as ours. Alternatively, it could arise as a result of a type of cluster sampling design in which, as the $n_i \rightarrow \infty$, the number of clusters also increases without limit, the average cluster size converging to a constant value. (Our data base, of course, is derived from a census of FHA-insured sales, and so does not follow a conscious sample design.)

Subject to these two qualifications, the E test may be viewed as a test of the consistency of β_2, while allowing specification bias in the micro regression. The finite sample behavior of the test and its power against various alternative hypotheses such as those discussed above are explored in Appendix B. Our simulations using the FHA sample show that the distribution of E approximated the χ^2 in three specifications which did not violate assumptions (i)–(iv). The power of the E test in identifying error in the grouped model appeared to be close to the power of the Farebrother test.

5. EMPIRICAL RESULTS

Table 1 presents two sets of regressions using observed FHA sample values of housing expenditure, income, and price, along with the results of grouping tests applied to each specification. Equation (1) was estimated using first the translog price index p_H and then substituting p_B, the BLS index of home ownership cost for high-income families. (Values of p_B are drawn from U.S. Bureau of Labor Statistics 1972, Table B-1). For clarity it should be noted that the estimation of alternative models, as presented in Table 1, corresponds to Polinsky and Ellwood's use of the term "simulation," whereas we use simulation here to denote the analysis of models by means of artificial data, as in our Appendix B and Table 2.) Our coefficient results are similar to those obtained in previous studies using FHA data. For example, our micro-level estimates of the income and price elasticities are .35 and −.64, respectively; these compare closely to Polinsky and Ellwood's estimates of .39 and −.67, as expected, given the similarity of functional form and variable definition. The agreement is less close at the grouped level. This may arise from the fact that our sample includes three more metropolitan areas than theirs.

Using either definition of the price of housing ser-

Table 1. Demand Function Estimates

	Price Variable	
Estimated Values	*Translog*	BLS
Micro Parameters		
$1 + b_p$	.359	.256
b_I	.349	.397
σ	.156	.164
Grouped Parameters		
$1 + b_p$	.398	.408
b_I	.471	.761
σ	.945	1.208
Farebrother Test		
B_F'	79.23[a]	491.78[a]
Degrees of Freedom	2,11973	1,11974
Feige-Watts Test		
B_{FW}'	2.13	8.73[b]
Degrees of Freedom	2,31	1,31
Micro Misspecification Test		
E	1093.34[a]	1282.22[a]
Degrees of Freedom	31	31

[a] Significant at .001 level.
[b] Significant at .01 level.

Table 2. Simulation Results

Simulated Values	I Mismeasured Income (50%)	II Mismeasured Income (10%)	III Metro Mean Price	IV No Deflation	V BLS Price
Mean Parameter Estimates					
Micro $1 + b_p$	.364	.315	.294	.324	.239
b_I	.395	.545	.617	.600	.640
σ	.179	.163	.159	.156	.162
Grouped $1 + b_p$	.303	.298	.299	.334	.313
b_I	.594	.599	.600	.595	.818
σ	.183	.162	.160	.177	.582
Farebrother Test					
Median B_F'	55.97	6.01	1.20	1.89	118.15
Degrees of Freedom	2,11973	2,11973	1,11974	2,11973	1,11974
Type II Errors					
.1	0	12	78	59	0
.01	0	38	95	84	0
.001	0	63	99	95	0
Feige-Watts Test					
Median B_{FW}'	50.52	5.64	1.22	1.49	8.88
Degrees of Freedom	2,31	2,31	1,31	2.31	1,31
Type II Errors					
.1	0	15	77	72	0
.01	0	47	95	97	23
.001	0	75	99	100	96
Micro Misspecification Test					
Median E	29.01	30.82	31.37	38.93	370.61
Degrees of Freedom	31	31	31	31	31
Type I Errors					
.1	9	9	10	—	—
.01	0	0	0	—	—
.001	0	0	0	—	—
Type II Errors					
.1	—	—	—	57	0
.01	—	—	—	86	0
.001	—	—	—	97	0

vices, we observe the traditional result that grouped regressions produce higher income elasticities. The divergence is less pronounced, and the explanatory power of the regressions is greater, when the household-specific translog price is used. Nevertheless, by means of the Farebrother test with the modifications discussed in Section 3, we can easily reject the null hypothesis of no specification or grouping bias. In the notation of Sections 3 and 4, we have $n = 11{,}978$, $m = 34$, and $k = 3$. Using the translog price, we have $j = 2$; income and price exhibit within group variation. The BLS prices are measured at the metropolitan area level, so $j = 1$ when p_B is used. Therefore, under the null hypothesis, B_F' is distributed as $F(2, 11{,}973)$ and $F(1, 11{,}974)$, respectively, under the two specifications. Table 1 shows that in both cases the statistic is significant at the .001 level.

Based on the value of E, we can also reject the hypothesis that the metropolitan-level regressions asymptotically satisfy the conditions of the standard linear model. The ratios of $\hat{\sigma}_2^2$ to $\hat{\sigma}^2$ are much too high to be explained by any household-level misspecification permitted by assumptions (i)–(iv). The low explanatory power of the grouped regressions also causes the Feige-Watts test to perform relatively poorly. The value of B_{FW}' is insignificant in the translog price regression, and it is significant only at the .01 level when the BLS home ownership price index is used.

6. CONCLUSIONS

In this article we have considered several grouping tests for misspecification. We have noted that three test procedures presented by previous authors must be modified in order to be generally applicable in regression settings. In Appendix B we have presented the results of artificial simulations as evidence that the relative power of these three tests depends not only on the respective degrees of freedom available but also on the nature of the alternative hypothesis. For the data set and demand model analyzed here, the Farebrother (1979) test was slightly more successful than the Feige-Watts (1972) test in identifying random household-level measurement error, and it was much more successful in identifying metropolitan-level misspecification. The sample size of 34 groups and 11,978 households was not sufficient for the Hausman (1978) test to satisfactorily approach its asymptotic properties.

We have also presented new evidence concerning the relative merits of micro and grouped analyses of housing demand. Through application of the modified Farebrother test we were able for the first time to demonstrate that random sampling error is insufficient to explain the divergence between the two coefficient vectors. In addition, using an asymptotic test of regression residuals developed in this article, we showed that random errors-in-variables problems cannot explain the high MSE's of the grouped regressions.

Our primary empirical conclusion is that, contrary to the reconciliation presented in Polinsky and Ellwood (1979), the micro/grouped paradox cannot be attributed to some combination of mismeasurement of household permanent income and the use of the metropolitan mean rather than the micro price of housing. The explanation lies more likely in an error in measurement of the interarea housing price index or in some other misspecification that causes the regression residuals to be correlated within metropolitan areas. Our results point to continued efforts to improve data bases and continued research on model specification, but they cast doubt on the efficacy of grouping by city to reduce or eliminate the measurement and sample design problems inherent in FHA data. In particular, we find no evidence that grouped estimates of demand parameters are consistent and, on that basis, preferable to micro estimates.

APPENDIX A: MODEL AND DATA

As indicated in Section 2, our specification of the housing demand regression equation (1) is drawn from Polinsky and Ellwood (1979). Closely similar specifications have been employed by Rosen (1978) and Gillingham and Greenlees (1981), and the general approach of estimating a logarithmic regression of expenditure on income and price (all deflated by a price index) has formed the basis for numerous articles in the housing-demand literature. Thus, although we make no claims here for the theoretical appropriateness or econometric robustness of (1), we believe that application of our statistical tests in this context has the potential to yield economic as well as methodological implications. In this Appendix we briefly review the derivation of (1) and the sources of data used in its estimation.

We begin by assuming that for a set of households the demand for housing quantity q is a logarithmic function of income I, housing services price p_H, and the index p_o of the prices of other goods and services,

$$\log q_i = b_o + b_p \log p_{Hi} + b_I \log I_i - (b_p + b_I)\log p_{oi} + \epsilon_i, \qquad i = 1, \ldots, n \quad (32)$$

where the coefficient on $\log p_{oi}$ reflects the requirement that demand be invariant to proportional changes in income and all prices. We then add $\log p_H$ and subtract $\log p_o$ from both sides of (32). Defining housing expenditure $C = p_H q$, we arrive at Equation (1), repeated here:

$$\log(C_i/p_{oi}) = b_o + (1 + b_p)\log(p_{Hi}/p_{oi}) + b_I \log(I_i/p_{oi}) + \epsilon_i, \qquad i = 1, \ldots, n. \quad (1)$$

Our data base is identical to that used and discussed in detail by Gillingham and Greenlees (1981), and was constructed as an approximation to that used by Polinsky and Ellwood (1979) and Rosen (1978). It consists of a sample of 11,978 new house purchases insured by the FHA during calendar year 1969. A total of 34 metropolitan areas are represented; the city samples range in size from 7 (Milwaukee) to 1,330 (Seattle). Housing expenditure is defined to equal the sales price of the home. Income is given by the FHA's estimate of the household's annual after-tax income likely to prevail during approximately the first third of the mortgage term. The measure of p_o is the total annual budget cost for higher-income homeowners, less the cost of housing, as estimated by the Bureau of Labor Statistics at the metropolitan area level (U.S. Bureau of Labor Statistics 1972, Table B-1).

The price index for housing services is specific to the individual home purchase. Extending a procedure used by Muth (1971), it is assumed that units of housing services are produced from land and capital (structures) according to a production function that is uniform nationally. For estimation purposes, we use a translog approximation to the indirect production function:

$$\log q = a_o + a_l \log v_l + a_s \log v_s + \tfrac{1}{2}c_{ll}(\log v_l)^2 + \tfrac{1}{2}c_{ss}(\log v_s)^2 + c_{ls}\log v_l \log v_s, \quad (33)$$

where q again is the quantity of services provided by the house, and v_l and v_s are the input prices for land and structures divided by total house cost. Assuming profit maximization by housing producers and perfectly competitive land and structures markets, we can identify the a and c parameters by estimating (33) in budget share form, using nonlinear least squares applied to the same FHA sample described above. Land prices and cost shares are taken from the FHA sales records. A metropolitan-level structure price index is taken from the Boeckh Building Cost Modifier series. Finally, given the parameter estimates, which are those reported in Gillingham and Greenlees (1981), we measure p_H as the index of the total cost required to produce a house with the sample mean value of $\log q$.

Aside from minor differences in data bases used, the above procedure differs from that of earlier authors in that they assume a homogeneous production function and hence the existence of a unit cost function for housing (CES for Polinsky and Ellwood and translog for Rosen). Gillingham and Greenlees (1981) were able to reject the homogeneity assumption econometrically. However, as noted in Gillingham and Greenlees (1983), the alternative models produce indexes that are all almost perfectly correlated, at least at the metropolitan level.

APPENDIX B: GROUPING TEST SIMULATIONS

Having derived appropriate forms of the B'_F, B'_{FW}, B'_H, and E test statistics in Sections 3 and 4, here we compare their performance in the context of econometric models of housing demand. In Section 5 we present micro and grouped housing demand regressions estimated using the FHA mortgage data described in Appendix A. The simulations in this Appendix employ the same sample base and the same basic specification given in Equation (1). However, in the place of actual housing consumption levels we have substituted values generated from an assumed regression model with a stochastic disturbance term. Repeated estimation of misspecified regressions enables us to examine the effects of the specification errors on the parameter estimates as well as on the grouping test statistics.

Our simulations employ the observed values of price and income in our FHA sample, while the values of housing expenditure were generated under the assumption that Equation (1) is correct. We assume that the true values of the parameters in (1) are $b_p = -.7$, $b_I = .6$, and $b_o = 2.02$. The disturbance term ϵ is assumed to be normally distributed with a standard deviation of .156. The above price and income elasticities were chosen to approximate the results obtained in earlier studies. (For example, Polinsky 1977 suggests that the true income and price elasticities are .75 and −.75. The elasticities in the "correctly specified" metro-level regression of Polinsky 1979 are .57 and −.72. Smith and Campbell 1978 argue for an income elasticity of between .50 and .70.) The values of σ^2 and b_o were based on the mean and variance of house price in our sample.

As in Section 5, we have $n = 11{,}978$, $m = 34$, and $k = 3$. For the specification (1), we have $j = 2$, since only income and price exhibit within-group variation. Differences in the micro and grouped estimates of b_I and b_p convey all necessary information about possible specification or grouping bias.

Under this null hypothesis, the Farebrother and Feige-Watts statistics will follow the F distributions given in (23) and (24). That is, B'_F is distributed as $F(2, 11{,}973)$ and B'_{FW} as $F(2, 31)$. However, B'_H is distributed as $\chi^2(2)$ only for sufficiently large values of m. In order to determine whether its asymptotic properties are attained in our sample of 34 cities, we simulated its distribution under the null hypothesis. For each of the 11,978 observations in our data set, we generated 100 values of housing expenditure using Equation (1) and our assumed parameter values. A normal random number generator was used to select values of ϵ. We then estimated micro and grouped regressions using each of the 100 simulated expenditure vectors.

The results of the regressions were not favorable for the Hausman test in the form (25). In four of the 100 simulations the computed value of B'_H was negative. The null hypothesis was incorrectly rejected at the 10% significance level in 17 of the remaining 96 trials. Finally, in two cases B'_H took on values of 37.5 and 82.0; the .999 point of the $\chi^2(2)$ distribution is only 13.8. We conclude that, at least in the tails, our finite-sample distribution of B'_H is not a satisfactory approximation to the chi-square, and we will not include the Hausman test in the error simulations which follow.

B.1 Simulations of Household-Level Misspecification

Simulations I, II, and III reported in Table 2 were designed to compare the performance of grouping test statistics in identifying micro-level measurement errors. Each simulation was performed by generating 100 values of the dependent variable for each observation according to the "correct" model (1), then estimating 100 sets of misspecified micro and grouped regressions. The first two simulations use "current income" as one regressor, obtained by adding a normal disturbance with zero mean to the income variable $\log(I/p_o)$ of Equation (1). Polinsky and Ellwood argue that the measurement error in income is sufficiently large to result in a permanent income elasticity 50% higher than the current income elasticity. This bias is accomplished in Simulation I by assigning a variance to the additive error equal to one-half of the sample variance in the true income term (Johnston 1972, p. 282). This may overstate the likely measurement problem in FHA data. Maisel, Burnham, and Austin (1971), for example, argue that the FHA value is, in fact, an estimate of permanent income, and that the error due to transitory components is small. Therefore, Simulation II applies a smaller disturbance term sufficient to cause a 10% divergence between the current and permanent income elasticities. A random number generator was used to add a separate transitory income component for each observation in each regression in Simulations I and II. In Simulation III, household income is assumed to be measured without error, but the price term $\log(p_H/p_o)$ is replaced by the metropolitan area mean of that term. Since there is no longer any measured intragroup price variation in Simulation III, only the income term is tested using the Farebrother and Feige-Watts statistics. Unfortunately, we do not observe the metropolitan means of $\log(p_H/p_o)$, only the means of the presumably random samples. Therefore, for each of the trials in Simulation III, we defined area means by adding to a group's sample mean a random normal variate with mean zero and standard deviation equal to the standard error of the mean—that group's standard deviation in price divided by the square root of the group sample size.

Table 2 displays mean coefficient values and estimates of β obtained from the micro and grouped regressions. The table also presents the median values of B_F' and B_{FW}', along with the respective numbers of type II errors—that is, acceptances of the misspecified model when (1) is true—at three arbitrarily chosen significance levels.

In each of the first three simulations, the success rates of the Farebrother and Feige-Watts tests are approximately equal. In Simulation I, both tests reject the hypothesis of no misspecification at the .001 significance level in all 100 trials. In Simulation II, where the income measurement error is less extreme, the Farebrother test correctly rejects 88 times at the .1 level, 62 times at the .01 level, and 37 times at the .001 level. The median value of B_{FW}' is lower than that of B_F', and the critical values are higher for the test with fewer degrees of freedom. Therefore, the number of type II errors is somewhat larger using the Feige-Watts test. Neither test is very successful in identifying the incorrect use of the metro-mean housing price. The general failure to achieve significance in Simulation III may result from the sample design used here. Intra-city variation in land prices is likely to be understated in a data set comprising only new, FHA-insured houses, as noted by Polinsky (1979).

The mismeasurement of income in Simulations I and II satisfies the null hypothesis of random observational error that underlies the E test. Therefore, the degree to which the statistic approximates the $\chi^2(31)$ distribution should indicate whether its asymptotic properties are met in our sample of 11,978 households. The median of the predicted distribution is 30.34; as shown in Table 2, the median simulated values are 29.01 and 30.82. The mean values of E are 29.61 and 30.88 in Simulations I and II, respectively—slighly lower than the chi-squared mean of 31. In each simulation the E test produces nine rejections (type I errors) at the 10% significance level and none at the 1% level. We compared the observed and predicted distributions of E using the D and V Kolmogorov-Smirnov goodness-of-fit statistics (Stephens 1970). At conventional significance levels, the null hypothesis that E was distributed as $\chi^2(31)$ could be accepted for both simulations.

Simulation III represents another test of the distribution of E under random micro-level mismeasurement, although the assumptions of the test are not entirely satisfied. Because the variation in price is not the same in each metropolitan area, the observational errors in the grouped regression are heteroscedastic. The within-group standard deviation in reported $\log(p_H/p_o)$ ranges in our sample from .0351 (Atlanta) to .1795 (Chicago). However, based on the results of Simulation III, the effect of this heteroscedasticity is minor. The E statistic has a median of 31.37 and rejects the null hypothesis 10 times at the 10% level. Again, Kolmogorov-Smirnov tests do not lead to rejection of the chi-squared distribution for the observed values. On balance, our first three simulations indicate that the asymptotic properties of E are approximately achieved in our sample.

B.2 Models with Metropolitan-Level Specification Errors

In Simulation IV the dollar values are not deflated by the interarea price index p_o. Simulation V replaces p_H by p_B, the BLS index of home ownership cost for higher-income families, which (in the context of these simulations, at least) is the "wrong" housing services price measure.

The likely effects of these misspecifications can be seen by rewriting Equation (1) in the following two forms:

$$\log(C_i) = b_o + (1 + b_p)\log p_{Hi} + b_I\log I_i + \epsilon_i - (b_p + b_I)\log p_{oi}, \quad (34)$$

$$\log(C_i/p_{oi}) = b_o + (1 + b_p)\log(p_{Bi}/p_{oi}) + b_I\log(I_i/p_{oi}) + \epsilon_i + (1 + b_p)\log(p_{Hi}/p_{Bi}). \quad (35)$$

The last terms in these equations are excluded regressors in Simulations IV and V, respectively. Since p_o and p_B are measured at the metropolitan level, and since much of the variation in p_H is between cities, the expected errors have nonzero means and will be highly correlated for observations within the same city. In the terminology used by Feige and Watts, this destroys the independence between the disturbance term and the grouping matrix G^*, leading to bias in the grouped regression coefficients. The effects should be similar to those resulting from grouping by the value of the dependent variable.

The quantitative impacts of Simulations IV and V are very different, as indicated in Table 2. The effect of failing to deflate the nominal variables is small, although to some extent this results from our particular choice of parameter values. For example, had we chosen to assume $b_I = .5$ and $b_p = -.8$, the coefficient on log p_{oi} in (34) would have tripled from $-.1$ to $-.3$. However, the incorrect use of the BLS price variable severely distorts the parameter estimates, particularly in the grouped regressions. In both simulations the Farebrother test identifies misspecification much more successfully than the Feige-Watts test. For example, at the 1% significance level, B_F' correctly rejects the null hypothesis in 16 Simulation IV trials, compared with three rejections using B_{FW}'. In Simulation V, again at the 1% level, the Feige-Watts test yields 23 type II errors, the Farebrother test none.

The weaker relative performance of the Feige-Watts test in the simulations of metropolitan-level misspecification follows from the presence of $\hat{\sigma}_2$ in the B_{FW}'

computational formula. Using Equations (23) and (24), we can write

$$B'_F/B'_{FW} = \hat{\sigma}_2^2/\hat{\sigma}_U^2. \qquad (36)$$

The errors in Simulations IV and V have a greater effect on the grouped than on the ungrouped regressions, as measured either by the degree of parameter bias or by the goodness of fit as reflected in MSE. Most obviously, the use of the BLS price results in a median $\hat{\sigma}_2$ of .582, which overstates the true σ by 273%. Consequently, the median value of B'_F is more than 13 times as large as the median B'_{FW}.

The high values of $\hat{\sigma}_2^2$ in Simulations IV and V also lead to rejections of the hypothesis that the only specification error is at the micro level. The E statistic is significant at any conventional level in all 100 Simulation V trials. In Simulation IV metropolitan-level misspecification is successfully identified in 43 trials at the 10% level. In both simulations the success rates of the E and B'_F statistics are comparable.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Steven Cobb, Robert Gillingham, Richard McDonald, and a referee for helpful comments and suggestions. The views expressed are those of the authors and do not reflect the policies of the BLS or the views of other staff members.

[Received June 1983. Revised November 1983.]

REFERENCES

FAREBROTHER, R. W. (1976), "A Remark on the Wu Test," *Econometrica*, 44, 475–477.

——— (1979), "A Grouping Test for Misspecification," *Econometrica*, 47, 209–210.

FEIGE, E. L., and WATTS, H. W. (1972), "An Investigation of the Consequences of Partial Aggregation of Micro-Economic Data," *Econometrica*, 40, 343–360.

GILLINGHAM, R., and GREENLEES, J. S. (1981), "Estimating Inter-city Differences in the Price of Housing Services: Further Evidence," *Urban Studies*, 18, 365–369.

——— (1983), "Estimating Inter-city Differences in the Price of Housing Services: Rejoinder," *Urban Studies*, 20, 95–97.

GILLINGHAM, R., and HAGEMANN, R. P. (1983), "Cross-Sectional Estimation of a Simultaneous Model of Tenure Choice and Housing Services Demand," *Journal of Urban Economics*, 14, 16–39.

GOODMAN, A. C., and KAWAI, M. (1982), "Permanent Income, Hedonic Prices and Demand for Housing: New Evidence," *Journal of Urban Economics*, 12, 214–231.

HAUSMAN, J. A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46, 1251–1271.

JOHNSTON, J. (1972), *Econometric Methods*, New York: McGraw-Hill.

LEE, T. H. (1968), "Housing and Permanent Income: Tests Based on a Three-Year Reinterview Survey," *Review of Economics and Statistics*, 50, 480–490.

MAISEL, S. J., BURNHAM, J. B., and AUSTIN, J. S. (1971), "The Demand for Housing: A Comment," *Review of Economics and Statistics*, 53, 410–413.

MAYO, S. K. (1981), "Theory and Estimation in the Economics of Housing Demand," *Journal of Urban Economics*, 10, 95–116.

MUTH, R. (1971), "The Derived Demand for Urban Residential Land," *Urban Studies*, 8, 243–254.

POLINSKY, A. M. (1977), "The Demand for Housing: A Study in Specification and Grouping," *Econometrica*, 45, 447–461.

——— (1979), "The Demand for Housing: An Empirical Postscript," *Econometrica*, 47, 521–523.

POLINSKY, A. M., and ELLWOOD, D. T. (1979), "An Empirical Reconciliation of Micro and Grouped Estimates of the Demand for Housing," *Review of Economics and Statistics*, 61, 199–205.

ROSEN, H. S. (1978), "Estimating Inter-city Differences in the Price of Housing Services," *Urban Studies*, 15, 351–355.

——— (1979), "Owner Occupied Housing and the Federal Income Tax: Estimates and Simulations," *Journal of Urban Economics*, 6, 247–266.

SMITH, B., and CAMPBELL, J. M. (1978), "Aggregation Bias and the Demand for Housing," *International Economic Review*, 19, 495–505.

STEPHENS, M. A. (1970), "Use of the Kolmogorov-Smirnov, Cramer-Von Mises and Related Statistics without Extensive Tables," *Journal of the Royal Statistical Society*, Ser. B, 1, 115–122.

U.S. BUREAU OF LABOR STATISTICS (1972), *Three Budgets for an Urban Family of Four Persons, 1969–70*, Supplement to Bulletin 1570-5, Washington, D.C.: U.S. Government Printing Office.