Econometric Theory, **19**, 2003, 78–99. Printed in the United States of America.
DOI: 10.1017/S0266466603191049

THE RISE AND FALL OF EXTRANEOUS ESTIMATION: LESSONS FROM ECONOMETRIC HISTORY?

A. BUSE
University of Alberta

It is possible to view econometric history as the retrospective assessment of intellectual resource allocations. This paper uses such a perspective to examine the history of extraneous estimation. Independently proposed by Marschak (1939, *Econometrica* 7, 332–335) and Wold (1940, *Efterfrågan på jordbruksprodukter och dess känslighet för prisoch inkomstförändringar* (The demand for agricultural products and its sensitivity to price and income changes)), extraneous estimation was widely used as a collinearity solution in demand analysis from 1950 to 1980. We give a detailed account of the origins of this procedure and the theoretical work that clarified the statistical structure that underpinned it. The subsequent literature is then assessed in terms of the optimal use of available econometric tools and the interaction between theory and applications. We find that optimal econometric procedures were rarely used and that theoretical refinements continued to be published well after extraneous estimation had ceased to be a useful tool for demand analysis.

1. INTRODUCTION

Econometric history can serve a variety of purposes, straight narrative being the most obvious. Although this paper will not avoid narrative, it will attempt to bring a resource allocation perspective to the econometric history of our cho-

An earlier version of this paper titled "To Pool or Not to Pool? Why Is That Still a Question?" was presented at the Thirteenth Annual Meeting of the Canadian Econometrics Study Group, University of Waterloo, September 21–22, 1996. I thank Robin Carter for his comments at that presentation. Another version titled "The Optimum Amount of Econometrics" was presented on October 31, 1998, at Concordia University on the occasion of Gordon Fisher's seventieth birthday. I am indebted to Lady Giovanna Stone and Peter Jones of King's College Library for arranging access to Sir Richard Stone's papers and to the staff of the UCLA Research Library for access to the Jacob Marschak Archive. James Tobin was kind enough to share some of his recollections of the "extraneous estimation" period. George Szava-Kovats and Chris Minns provided able research assistance, and Charlene Hill typed the manuscript with her customary efficiency. Colin and Sally Ash, Reinhold Bergstrom, Jim Hamilton, Balder von Hohenbalken, Joe Ostroy, and Martin Weale provided help of the sort that does not lend itself to detailed elaboration. An earlier version of the paper had the benefit of a close and perceptive reading by John Aldrich. The referees' and editor's comments have also helped shape the paper into its present form, and I thank them for their contributions. Address correspondence to: A. Buse, Department of Economics, University of Alberta, Edmonton, Alberta, Canada T6G 2H4; e-mail: buse@econ.ualberta.ca.

 © 2003 Cambridge University Press 0266-4666/03 $9.50

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

sen topic, extraneous estimation. Issues of resource allocation are, of course, implicit whenever retrospective assessments are made of research programs because every research program has an opportunity cost. We make the resource allocation implications explicit through a case study of the relationship between econometric theory and practice. In this context we are unaware of any study that has tried to determine whether, and to what extent, the best available econometric techniques were used in applied work. Nor has any attempt been made to assess to what extent theoretical work evolves autonomously, independent of the concerns and practices of applied workers. We also ask the normative question as to whether the observed resource allocations were efficient or not. The notion of efficiency in this context is undoubtedly elusive, perhaps illusory, but note that judgments about efficiency are frequently made when contemporary econometricians discuss the state of the discipline (see, e.g., the special issue of the *Journal of Econometrics*, vol. 100, 2000, and also Leamer, 1997, pp. 559–560). We believe that historical assessments such as ours can add useful conditioning information to these contemporary discussions.

The history of extraneous estimation spans 60 years, but the technique was most frequently used to pool time series and cross section data in demand analysis from 1950 to 1980. Neither the theory nor its application has left a significant legacy for contemporary econometrics. The absence of a legacy does not imply, however, a low profile in its own time; major figures such as Marschak, Wold, Stone, Tobin, Klein, and Durbin made contributions. More recently, Tobin's (1950) classic paper on pooling was the benchmark for the experiment in applied econometrics conducted by Magnus and Morgan (1997). Our enquiry and its results can be considered a complement to that experiment.

As a prelude to our main theme we give an account of the origins of extraneous estimation in Section 2 because a historically accurate account is missing from the literature. Section 3 sets out the statistical structure of extraneous estimation to establish a framework for the assessment of the empirical literature. That assessment is made in Section 4. Sections 5 and 6 address the resource allocation question.

2. ORIGINS

It is best to start at the beginning. But where the beginning starts is not an entirely unproblematic matter. Questions of priority can often not be resolved unambiguously, and so it is in the case of pooling under review here. Although nothing quite as momentous as Legendre's complaint against Gauss concerning least squares is at stake in this case, a historically accurate account, insofar as it can be determined, has not yet been put in place. The available evidence suggests that both Marschak and Wold can claim to have originated the pooling procedure in economics, and we consider their respective claims in turn.

The first published proposals to combine cross section budget data with time series market data in demand analysis were made by Marschak (1939a, 1939b,

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

1939c). To facilitate what follows, we present the gist of his proposal as given in (1939c) (with some minor adjustments in the notation to speed the exposition). Let the budget function of a family be defined by

$$q = \alpha + \beta y, \tag{1}$$

where q is the quantity consumed and y is family income. If p is the price, take

$$\alpha = a + cp,$$
$$\beta = b + dp,$$

so that the family's demand function is given by

$$q = q(p, y) = a + by + cp + dpy. \tag{2}$$

At the time of the budget survey $p = p_0$ and the parameters

$$\alpha_0 = a + cp_0, \tag{3}$$
$$\beta_0 = b + dp_0$$

can be calculated from the budget data. Average market demand (i.e., per family) at time t, $E(q_t)$, is obtained by integrating across the distribution of income, $f(y_t)$, i.e.,

$$E(q_t) = \int q(p_t, y_t) f(y_t) dy_t = a + bE(y_t) + cp_t + dp_t E(y_t), \tag{4}$$

where $E(y_t)$ is average income per family. Substituting for a and b from (3) yields

$$E(q_t) - \alpha_0 - \beta_0 E(y_t) = c(p_t - p_0) + dE(y_t)(p_t - p_0), \tag{5}$$

which is to be estimated by least squares.

There is written evidence that Marschak had this idea as early as 1937. In 1931 Marschak published *Elastizität der Nachfrage* (Marschak, 1931b), in which he attempted to derive the (Marshallian) demand curve from budget data. Henry Schultz (1933) reviewed this book in *Weltwirtschaftliches Archiv* and concluded that Marschak's method was not successful, pointing out that what Marschak had estimated in his empirical work was nothing more than the income elasticity of consumption. Schultz (1938, pp. 117–120) reproduces this critique in *The Theory and Measurement of Demand* and in a footnote (p. 120) records Marschak's prepublication response to this critique. The relevant portion of Marschak's letter, dated November 23, 1937, reads as follows:

> I am therefore even more sceptical than R. Frisch as to the usefulness of my old method for a direct derivation of demand curves. On the other hand, income-elasticities can be very useful for the analysis of demand from time series, as they permit the elimination of influences on demand of known changes in the amount and distribution of incomes and thus to account properly for what probably is the most important cause of shifts of a demand curve.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

The phrase "permit the elimination of influences" suggests that Marschak had already developed a procedure for combining cross section and time series data, but we do not know whether the implied procedure is the same as that which he published in 1939. Nor do we have any evidence as to the date when this implied procedure was formulated.[1] Morgan (1990, p. 153) notes that "Marschak also tried to combine market demand analysis with budget studies in two papers (1931 and 1939)" (our references are (1931b) and (1939c)), but this is misleading because Marschak did not deal with time series data in his 1931 elasticity study (Marschak, 1931b). What he did do in that study was to aggregate up to the market level from the budget level data by integrating across the income distribution as illustrated in (4).

Wold's claim rests on the following facts. In 1940 he published (in Swedish) a monograph on the price and income sensitivity of the demand for agricultural products (Wold, 1940).[2] In that monograph he introduced the notion of "conditional regression," which in application consisted of fixing the income elasticity (or the ratio of the income to price elasticity) in the time series equation at a preassigned value to overcome the problem of multicollinearity. A typical equation is thus defined by

$$\log q_t - \eta_y^* \log y_t = \lambda + \eta_p \log(p_t/\pi_t) + u_t, \tag{6}$$

where q_t is per capita consumption, y_t per capita real income, p_t the price, π_t a general price level, and η_y^* a value of the income elasticity determined from cross section information, usually an ordinary least squares (OLS) regression. Conditional regression is thus equivalent to what is now known as extraneous estimation.

Although we know that Wold had worked out the pooling procedure by 1940 it seems unlikely that he would have done so before 1938 because the Agricultural Research Commission, which sponsored his research, was established in 1938 (Wold, 1941, p. 263; 1953, Preface x; 1982). The available evidence suggests, therefore, that Marschak conceived the idea first. It should be noted, however, that Wold (1940) was the first to implement the pooling procedure empirically because Marschak did not do so until 1943 (Marschak, 1943).

This account of the origins of pooling allows us to correct a number of inaccuracies that have appeared in the literature. First, Buse (1981, p. 229) fails to credit Wold along with Marschak as an originator. Second, Gilbert (1991, p. 298) writes as follows about Stone's (1954a) use of extraneous estimators, "So far as I am aware, this is the first example of mixed time series and cross section estimation." Our chronology shows that both Wold and Marschak had combined time series and cross section data prior to Stone. Marschak's (1943) application is cited by Stone (1945, p. 288) with the comment that pooling could not be implemented due to the lack of budget data. By May 1947 Stone was analyzing budget data,[3] and his first published use of extraneous estimation appeared in *Review of the International Statistical Institute* (Stone, 1948). Sim-

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

ilarly, Anderson and Vahid (1998, p. 291) are mistaken when they assert that Tobin (1950) "pioneered the pooling of cross-sectional and time series data." Tobin (1997, p. 648) states that he first pooled time series and cross section data in his 1947 Harvard Ph.D. thesis.

Although there is evidence (set out in Buse, 2000) to suggest that Wold was at odds with Marschak over priority, it is more instructive to explore the differences in Wold's and Marschak's stated motivations for pooling data and the solutions they adopted. Wold's position is the more transparent of the two. For him the problem is a numerical/statistical one. The numerical indeterminacy of the least squares estimator under perfect collinearity becomes the problem of large standard errors and numerically unstable estimates ("nonsense elasticities" in Wold, 1944, p. 117) under near collinearity. An underdetermined system of equations, the normal equations in this instance, can always be solved for a subset of the unknowns by giving an appropriate number of unknowns fixed values. The solution is then conditional on these prespecified values. This is the description given in Wold (1953, pp. 46–48), and in Wold (1944, p. 118) he notes that "the method suggested is only an application of the familiar device for handling side conditions in the method of least squares,"[4] hence the characterization "conditional regression." The side conditions are extraneous information that may be empirical or theoretical, and in Wold's (1953) *Demand Analysis* both types of information are used. Wold does not allow for any uncertainty associated with the extra information and treats the reduction in the standard errors as "equivalent to a decrease in multicollinearity" (1953, p. 237).

Turning now to Marschak, the most explicit statement of his motivation is given in the previously cited letter to Schultz; namely, he thinks that the budget data can be used to eliminate shifts in the demand function due to shifts in the level of income and its distribution. Variations on this theme appear in the three papers (1939a, 1939b, 1939c) that elaborate on the mechanics of pooling, but the importance of the distribution of income is emphasized in all of them.[5] It is the explicit concern with the distribution of income (and implicitly, aggregation) that distinguishes Marschak's demand analysis from Wold's. It is also the distributional concern that ultimately determines how the budget and market data are pooled. Rather than follow through with least squares estimation of the Engel curve as suggested by (1), his 1943 application followed a "purely numerical" approach he mentions in passing in his 1939 papers. Marschak constructed a nonparametric Engel curve by plotting, for each income bracket, mean per capita consumption against mean per capita income. This plot allows him to calculate per capita consumption in the budget year. Assuming proportionate changes in the income distribution, per capita consumption in other years can then be calculated, and these consumption estimates are then used to create the income adjusted time series of per capita consumption. In all of this collinearity plays only an incidental role. It is not mentioned in the 1939 papers, and removing the collinear income variable from the time series appears to be viewed

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

as a serendipitous by-product (1943, pp. 44, 46). The principal advantage Marschak sees for his procedure is "the use of more information," including the income distribution characteristics.

That Marschak and Wold should independently develop the pooling procedure at roughly the same time is not especially surprising.[6] It represents a logical evolution of the research results of the previous four decades. Recall that the principal objective of applied demand analysis for much of this period was to estimate the Marshall–Cournot demand curve, $q = f(p)$, and the associated price elasticity of demand. This project was pursued by both time series and cross section investigations. The time series approach was pioneered by Lehfeldt, Moore, and others, and the details of their achievements and failures are chronicled in Morgan (1990, Chs. 5, 6). The cross section studies designed to extract price elasticities have not received the same attention, but the principal attempts by Pigou, Frisch, Marschak, and Roy are reviewed by Schultz (1938, Ch. 3). With the exception of Pigou, all these cross section attempts were made in the 1930s, and it must have occurred to more than one person to speculate about the relationship between the cross section and time series investigations. These speculations were likely heightened by the rediscovery by Hicks and Allen (1934) of Slutsky's (1915) results that made explicit the importance of the budget constraint in the optimizing process. This focused attention on the role of the income variable in empirical specifications of time series demand functions that prior to this time had typically included own price and possibly one substitute or complementary good. Once it was realized that income responses could be measured in both time series and cross sections the last step was simple enough. The characterization "simple" does not, however, imply that it was obvious, except ex post, because it required the necessary statistical sophistication and imagination to realize that improved accuracy in estimation might follow from combining two sources of data.

3. RISE

The decade following World War II can be considered the heyday of extraneous estimation. The indisputable center of research on this topic, both applied and theoretical, was the Department of Applied Economics at Cambridge (UK), directed by Richard Stone. Of the nine applications published during the period 1946–1955 only two, Wold (1953) and Klein and Goldberger (1955), did not have any connection with Cambridge (see Table 1). On the theoretical side, Durbin (1953), at the prompting of Farrell and Stone, clarified the statistical structure of extraneous estimation. Indeed, so central are Durbin's results for an assessment of subsequent theoretical and applied work that a short exposition of his analysis is unavoidable. We do so with a simple scalar example; a general matrix account is available in Buse (2000).

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

TABLE 1. A summary of the pooling literature by decades, 1939–2000

Applications	Theory and methods
1939–1949	
Wold (1940), Marschak (1943), Staehle (1945), Stone (1948),[a] Prest (1949)[a]	Marschak (1939c), Wold (1940)
1950–1959	
Tobin (1950),[a] Stone (1951),[a] Houthakker and Tobin (1952),[a] Wold (1953), Stone (1954a),[a] Farrell (1954),[a] Klein and Goldberger (1955), Stone and Rowe (1958)[a]	Durbin (1953), Klein (1953), Kuh and Meyer (1957)
1960–1969	
Tomek and Cochrane (1962),[b,c] Solow (1964), Fox (1965), Moore (1966), Holmes (1968),[b] Lee (1969)[b]	Kaliski (1960), Theil and Goldberger (1961), Chetty (1968)
1970–1979	
Laughhunn (1970), Manhertz (1970), Carman (1972),[c] Chen (1972), Hamilton (1972), Saito (1977),[c] Grabowski, Vernon, and Thomas (1978), Konandreas and Schmitz (1978)[b]	Aigner and Simon (1970), Maddala (1971), Desai (1974), Ghali (1977), Jeong (1978), Passi (1979)
1980–1989	
Schneider, Klein, and Murphy (1981), Rossi (1988)	Izan (1980), Pudney (1980), Buse (1981), Havenner and Craine (1981), Bodkin and Raj (1985)
1990–2000	
Fuglie (1994)[b]	Mikhail and Ghazal (1991), Srivastava and Ullah (1995), Buse (1997), Anderson and Vahid (1998)

[a]Associated with Department of Applied Economics, Cambridge.
[b]Appeared in agricultural journal or model.
[c]Nondemand application.

Let the cross section and time series equations be, respectively,

$$y_{1i} = z_i \alpha_1 + \mu_{1i}, \qquad i = 1, \ldots, n, \qquad \mu_{1i} \sim N(0, \sigma_1^2), \tag{7}$$

$$y_{2t} = x_{1t} \beta_1 + x_{2t} \beta_2 + \mu_{2t}, \qquad t = 1, \ldots, T, \qquad \mu_{2t} \sim N(0, \sigma_2^2), \tag{8}$$

where the disturbances are uncorrelated across the equations. OLS estimation of (8) yields the estimators $\hat{\beta}_1^0$ and $\hat{\beta}_2^0$ and their variances,

$$V_1^0 = \sigma_2^2 / (1 - r_{12}^2) T s_1^2, \tag{9a}$$

$$V_2^0 = \sigma_2^2 / T s_2^2 + r_{12}^2 (s_1^2 / s_2^2) V_1^0, \tag{9b}$$

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

where $Ts_j^2 = \sum_t x_{jt}^2$. The correlation between x_{12} and x_{2t}, r_{12}, is large enough to preclude reliable OLS estimation of β_1 and β_2. Extraneous information about β_1 is available from (7) in the form of the OLS estimator $\hat{\alpha}_1^e$ and its variance V_1^e. The substitution of $\hat{\alpha}_1^e$ into (8) to break the variance inflating collinearity between x_{1t} and x_{2t} gives

$$y_{2t} - x_{1t}\hat{\alpha}_1^e = x_{2t}\beta_2 + \mu_{2t}^e, \tag{10}$$

where $u_{2t}^e = u_{2t} + x_{1t}(\beta_1 - \hat{\alpha}_1^e)$. The OLS extraneous estimator of β_2, $\hat{\beta}_2^e = \sum(y_{2t} - x_{1t}\hat{\alpha}_1^e)x_{2t}/\sum x_{2t}^2$, has unconditional variance

$$V_2^e = V(\hat{\beta}_2^e|\hat{\alpha}_1^e) + r_{12}^2(s_1^2/s_2^2)V_1^e, \tag{11}$$

where $V(\hat{\beta}_2^e|\hat{\alpha}_1^e) = \sigma_2^2/Ts_2^2$ is the conditional variance of $\hat{\beta}_2^e$. Comparison of (9b) and (11) reveals that extraneous estimation will have been successful in reducing the standard error of $\hat{\beta}_2^0$ only if $V_1^e < V_1^0$, that is, if the extraneous estimator of β_1 has smaller variance than the OLS estimator. Note also that the unconditional variance, V_2^e, is always larger than the conditional variance, $V(\hat{\beta}_2^e|\hat{\alpha}_1^e)$, and it is the latter that was so often mistakenly calculated in the applications to give overoptimistic results for significance tests.

Although extraneous estimation is a solution to collinearity (subject to $V_1^e < V_1^0$), Durbin also provided the optimal solution for the pooling of the data. He did not attack the problem directly by estimating the two equations jointly with an across equation restriction.[7] Instead, he posed the problem in terms of the optimal linear combination of the OLS estimators from (7) and (8). The solution is given by a generalized least squares regression applied to the equations

$$\begin{pmatrix} \hat{\alpha}_1^e \\ \hat{\beta}_1^0 \\ \hat{\beta}_2^0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}, \tag{12}$$

where $\varepsilon_0 = \hat{\alpha}_1^e - \beta_1$, $\varepsilon_j = \hat{\beta}_j^0 - \beta_j$ and $\mathrm{Var}(\varepsilon_0) = V_1^e, \mathrm{Var}(\varepsilon_j) = V_j^0 (j = 1,2)$. Not surprisingly, the optimal pooled estimator for β_1 takes the form of a weighted average

$$\hat{\beta}_1^p = (w_1 + w_2)^{-1}(w_1\hat{\alpha}_1^e + w_2\hat{\beta}_1^0), \tag{13}$$

where the weights, $w_1 = 1/V_1^e$ and $w_2 = 1/V_1^0$, are the reciprocals of the OLS variances. The pooled estimator for β_2 is given by $\hat{\beta}_2^p = \sum x_{2t}(y_{2t} - x_{1t}\hat{\beta}_1^p)/\sum x_{2t}^2$ with variance

$$V_2^p = \sigma_2^2/Ts_2^2 + r_{12}^2(s_1^2/s_2^2)V_1^e(1 + (V_1^e/V_1^0))^{-1}. \tag{14}$$

Although Durbin did not do so, it is instructive to examine the determinants of the efficiency losses, $V_2^e - V_2^p$, of the extraneous estimator for β_2. A comparison of (11) and (14) reveals that the efficiency losses will decline as V_1^e/V_1^0

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

declines; that is, the losses are smaller the greater the reliability of the extraneous estimator of β_1 relative to the OLS estimator. There are no efficiency losses if the extraneous information about β_1 is exact ($V_1^e = 0$). The losses are also zero if $r_{12}^2 = 0$ or as $r_{12}^2 \to 1$. When the correlation is zero the OLS estimators in (8) are independent, so that the extraneous information about β_1 has no bearing on our ability to estimate β_2. When $r_{12}^2 \to 1$, $V_1^0 \to \infty$ and from (13) we see that $\hat{\beta}_2^p \to \hat{\beta}_2^e$, so that efficiency losses from extraneous estimation tend to zero. It can also be shown that the relative efficiency losses are quadratic in r_{12} with a maximum loss at $r_{12}^2 = 0.5$. Under plausible assumptions the maximum loss for $\hat{\beta}_2^e$ is 12.5%, whereas for $\hat{\alpha}_1^e$ it is 50% (details appear in Buse, 2000).

Stone's (1954a) *Measurement of Consumers' Expenditure and Behaviour in the U.K.* was the capstone to the work at the Department of Applied Economics; he subsequently published only one paper (Stone and Rowe, 1958) using pooled data. Although Stone had access to Durbin's efficient estimator, "ease of calculation" persuaded him to work exclusively with the extraneous estimator in *Measurement*. He had, however, noted that the efficiency losses for β_1 would be small if r_{12} were large.[8] There is one result in Stone's (p. 305) discussion of efficiency that has not received the attention it deserves. For (8) and the extraneous information $\hat{\alpha}_1^e$ and V_1^e he specifies the log-likelihood (in our notation)

$$L = \text{constant} - \tfrac{1}{2}\left\{(\hat{\alpha}_1^e - \beta_1)^2/V_1^e + \sum (y_{2t} - x_{1t}\beta_1 - x_{2t}\beta_2)^2/\sigma_2^2\right\}. \qquad \textbf{(15)}$$

Maximization of (15) yields the efficient pooled estimators. The implied linear structure for this likelihood is

$$\hat{\alpha}_1^e = \beta_1 + \varepsilon_0,$$

$$y_{2t} = x_{1t}\beta_1 + x_{2t}\beta_2 + \mu_{2t}, \qquad \begin{pmatrix} \varepsilon_0 \\ \mu_{2t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_1^e & 0 \\ 0 & \sigma_2^2 \end{pmatrix}\right). \qquad \textbf{(16)}$$

The system (16) is the mixed estimation format for uncertain prior information introduced by Theil and Goldberger (1961).

It would be remiss not to examine the other two seminal studies that helped shape the development of extraneous estimation. Unlike Stone, Tobin (1950) and Wold (1953) did not have Durbin's result for the unconditional standard errors, so their reported inferences are biased toward rejection of the null. All three of them addressed a set of common problems, and their modeling strategies have common elements and also some interesting variations. Aggregation was a common theme, and both Tobin and Stone assume that income changes proportionally, which allows them to construct an aggregate demand curve without having to worry about higher moments of the income distributions (cf. Marschak's formulation). Wold tackles aggregation only insofar as he shows that a market elasticity can always be written as a share weighted average of individual elasticities, a result that is similar to Stone's on the market Slutsky substi-

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

tution matrix. Related to aggregation is the choice of functional form, and all three made the pragmatic decision to use the double logarithmic form even though the constant income elasticity in the cross section was known to be at variance with the facts.

Stone differenced his data, and the implicit trend was expected to capture a variety of evolving variables such as tastes and income distribution. Wold argued against trend removal on the grounds that it threw away statistical information inasmuch as deviations from trend would only measure short-run responses whereas the long-run response was the primary objective. Tobin treated trends as a minor issue and focused instead on dynamics via lagged income variables. Both Stone and Wold discussed the need for dynamics, but neither incorporated lags in their empirical specifications.

Diagnostic tools to evaluate models were few in this period, focusing principally on the serial independence of the errors. Stone had the Durbin–Watson statistic at his disposal, whereas Tobin had to rely on the Von Neuman ratio. Wold, on the other hand, uses a large sample test suggested by Whittle (1952), although most of the time the correlogram of the OLS residuals is used as an informal diagnostic. Worth noting is Wold's (1953, pp. 210–213, 234–240) development and use of autocorrelation consistent standard errors for the OLS estimates, a first not widely known or acknowledged.

Although the Cowles approach represented the frontier of econometric research during this decade, it failed to displace OLS in extraneous estimation.[9] Wold had developed his causal chain approach as early as 1946, and it should come as no surprise that he remained an unrepentant advocate of OLS. Stone reviews the arguments for structural estimation but rejects it because of the extra computational burden, the larger variances, and the likely loss of consistency under measurement error. He also argues that autocorrelation has a much greater impact on his inferential procedures than the systems complications. Tobin gives the most credence to the possibility of simultaneity. He estimates the reduced form of a two equation system by maximum likelihood and concludes that supply is predetermined so that with independent disturbances across equations the model is recursive and OLS optimal.

4. NORMAL SCIENCE

The work of Durbin, Stone, Tobin, and Wold provided a template that would be used by others to shape their own work. Inasmuch as Durbin had made an almost exhaustive theoretical analysis of the problem, applied researchers were in a position to improve on the earlier empirical work. Progress is also made when underlying assumptions are subject to critical scrutiny. Kuh and Meyer (1957) addressed the legitimacy of the key assumption in extraneous estimation, namely, that cross section and time series income responses measure the same thing. They emphasized the long-run–short-run distinction and claimed that there was an observable and systematic larger long-run response. This was

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

evident in Tobin's results, and Kuh and Meyer developed an informal diagnostic based on alternative R^2 statistics that allowed them to confirm indirectly this regularity in Stone's *Measurement*. The validity and impact of this critique are discussed in the next section.

That Kuh and Meyer (and also Durbin) did not address the compatibility question by casting it in the form of a testable hypothesis undoubtedly reflects the emphasis on estimation, relative to testing, that was prevalent at the time. Given the model defined by (7) and (8), the relevant null is $H_0: \alpha_1 = \beta_1$, and the associated Wald test is given by

$$\tau_w = (\hat{\alpha}_1^e - \hat{\beta}_1^0)^2/(\hat{V}_1^e + \hat{V}_1^0) \overset{a}{\sim} \chi^2(1). \tag{17}$$

Working with an analysis of variance framework set out by Cochran (1954), Stone and Rowe (1958) were the first to test for equality across data sets, but their test went unnoticed and unused in the subsequent literature.[10]

The importance of the question notwithstanding, an explicit response to the validity of pooling question had to wait until Maddala (1971) addressed the problem in his comparison of the likelihood and Bayesian approach to pooling, the latter having been explored by Chetty (1968). Maddala emphasized the need for a pooling check, be it a classical likelihood ratio (LR) test or posterior odds. In addition to a formal test, the first part of Maddala's paper provides a fully general statement of traditional approaches and derives various estimators and their efficiency rankings. His work can be considered the completion of the classical theoretical framework for the original extraneous estimation problem.

Space limitations preclude a detailed examination of the applied literature, so we resort to a summary statement for the period 1939–2000;[11] the detail is available in Buse (2000). Reference to Table 1 shows a steady flow of applied studies during the 1960s and 1970s, but this flow all but disappears in the 1980s. The bulk of the applications (27 out of 30) were in demand analysis, and by and large they deviated but little from the original Stone–Tobin–Wold format. Indeed, insofar as the aggregation problem disappears from sight these studies are less sophisticated than the seminal papers. Applications of extraneous estimation to systems of demand equations by Saito (1977) and Rossi (1988) were the only noteworthy extensions. With one exception, Moore (1966), OLS continued to be the estimator of choice, but this was usually modified to a Cochrane–Orcutt procedure to allow for AR(1) disturbances. Dynamics, via lagged dependent variables and distributed lags, appear with some regularity, but the long-run–short-run distinction that troubled Kuh and Meyer was never properly addressed even though a potential solution was buried in the dynamics (for one simple solution, see Tobin's (1950) two period analysis). Aigner and Simon (1970) demonstrated that omitted dynamics in the cross section equation would produce cross section income elasticities larger than the time series, but their critique did not receive the same exposure as Kuh and Meyer.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

5. SKEWED SCIENCE

The narrative of the previous section might well be read as an account of the process of normal science as practiced by economists. A technique is found to solve a particular problem; it gains prominence and acceptance through the work of some leading practitioners, which in turn draws the interest of the theorists, who provide analytical improvements for subsequent applied work. Eventually one of two possible outcomes occurs. One possibility is survival in that the technique is incorporated permanently into the body of accepted procedures because it has wider reach and relevance than the original application. Alternatively, the technique falls into disuse because better techniques are developed for dealing with the original problem or the problem itself ceases to be of interest.

Given the historical pattern described in the previous section, it seems safe to conclude that extraneous estimation has failed to survive as a standard procedure, and this conclusion can readily be confirmed by reference to contemporary econometrics textbooks. A definitive account of the demise of extraneous estimation is not a realistic option insofar as researchers seldom record why they chose not to do something and referees' reports in journal archives are not accessible. Nonetheless, the following points can be made. First, later approaches to collinearity such as principal components, ridge regression, and Bayesian methods were almost never used in demand analysis, so extraneous estimation was not displaced by alternative techniques. Nor can one state categorically that the critique of Kuh and Meyer had a significant deterrent effect. The divergence between cross section and time series estimates was central to the Kuh and Meyer critique, but their evidence for this claim is fragile. Kuh and Meyer focus on Tobin (1950) and Stone (1954a), the evidence from Wold (1953) having been judged by them to be inconclusive although they fail to cite Wold's argument (p. 230) that budget income elasticities would "on the whole tend to be" smaller than the time series elasticities. They note correctly that Tobin's cross section estimate is twice as large as his time series estimate, but they fail to note that Stone's discussion of Tobin's paper (p. 142) shows that the time series–cross section divergence disappears when an alternative (Commerce Department) time series of food expenditure is used in place of Tobin's quantity series. In assessing Stone's (1954) results Kuh and Meyer convinced themselves, using their R^2 diagnostic, that 9 out of 10 cross section elasticities must have been at least twice as large as the time series. Yet, Stone (1951, pp. 21–22) (not cited by Kuh and Meyer), reporting early versions of the 1954 results, discusses the differences between his time series and cross section estimates, and that discussion does not suggest that 90% of Stone's budget elasticities were twice the time series value. The evidence after 1957 is just as fragile, if not more so. Five of the post-1957 applied papers (Stone and Rowe, 1958; Tomek and Cochrane, 1962; Moore, 1966; Lee, 1969; Manhertz, 1970) report both time series and cross section estimates. In 20 cases the time series

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

estimate is larger than the cross section, and in 24 the converse holds. If we add the instances in Burk (1958) and Crockett and Friend (1960) the count becomes 28 and 26, respectively. We do not know whether demand analysts took their cue from Kuh and Meyer or from all the evidence available in the published literature. Agnosticism would seem to be the most sensible position.

Other factors were probably at work. First, extraneous estimation failed to prove useful outside the realm of demand analysis insofar as only 3 (out of 30) nondemand applications were made (see notes to Table 1). Second, demand analysis underwent a transformation from the single equation tradition to the systems approach, a transformation initiated by Stone (1954b) with the Linear Expenditure System. In demand systems, across equation restrictions will mitigate collinearity problems as will the large number of observations available in panel data sets. Finally, even though the single equation tradition has survived, the collinearity problem has ceased to be as critical as longer data sets became available. The second point is, in our view, the most important factor in the disappearance of extraneous estimation.

That an econometric procedure should fall into disuse is, of course, not news. What may be newsworthy is the surprising disjunction of theory and practice that emerges when the extraneous estimation story is given closer scrutiny. Recall that Durbin's theoretical analysis provided both the correct standard errors for the extraneous estimator (11) and a new efficient estimator (13). Stone (1954a) used the correct standard errors, and, correspondingly, his inferences are reliable (in the large sample sense because the distribution of the standard test statistics is no longer exact, even under normality). Such reliability cannot be attributed to the literature that followed Stone, for we have been unable to find a single instance in which there was an explicit statement that appropriate standard errors had been calculated.[12] Furthermore, Rossi (1988) is the only study using joint efficient estimation among the 21 applications papers published after Durbin's results became available.

Fuglie (1994) is the last known application of extraneous estimation, so we need to enquire why, over a 40 year span, the applied literature failed to use optimal econometric procedures. We consider efficiency first. Computational complexity comes to mind as a likely deterrent against efficient estimation, but this fails as an explanation because Durbin makes it clear that the efficient estimator requires the modification of only two sample moments in the normal equations; i.e., the matrix inversion is of the same order as the original OLS problem. In any case, what would have been a computational challenge in 1953 would hardly be so in 1994. A plausible explanation comes from the conjunction of two facts. First, Durbin gave an illustrative example of the efficient estimator and found there were no efficiency gains relative to the extraneous estimator (see the pork numbers in note 8). Second, Stone, having taken note of Durbin's example, may have concluded that the minimal gains from efficient estimation could not justify the extra effort of reestimating all his equations (196 in total). Others then found it easy to follow the example set by

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

Stone although the issue of efficiency is not discussed in any of the applied papers, citations of Durbin notwithstanding.

If Stone set the example on efficiency, he also set the example on the correct standard errors. But in this instance the literature ignored his good example, reporting instead upward biased *t*-statistics. A justification based on computational complexity cannot be sustained as the calculation of correct standard errors in the prototypical study with one extraneous coefficient requires only one extra regression.

When the issue of testing the compatibility of time series and cross section data is examined a similar story emerges. The compatibility question had been given a thorough airing by Kuh and Meyer, and, though widely acknowledged in citations, it had little impact on the practice of extraneous estimation, in part no doubt because Kuh and Meyer failed to provide a clear cut test procedure. As noted earlier, Stone and Rowe had used Cochran's analysis of variance setup to test for the equality of time series and cross section income elasticities, but their lead found no followers. Maddala's (1971) LR test for pooling revived the question but again without apparent impact on applied work. Konandreas and Schmitz (1978, p. 80) test for compatibility with a χ^2 test, but because neither Stone and Rowe nor Maddala is cited we do not know whether a Wald or an LR test was used. A Wald test for the compatibility of time series and cross section elasticities is given by Pudney (1980, p. 889), but the only application of a compatibility test in a system is an LR test used by Rossi (1988).

What we have documented here is a pattern, barring the noted exceptions, in which theoretical developments appeared to have no noticeable effect on how applied studies were executed. The typical study did not test for the compatibility of the time series and cross section data, and the extraneous estimates were treated as certain information. Thus, the typical study did not differ from the format laid down in the early days of Wold and Tobin.

An alternative perspective on the divergence between theory and practice can be had by examining Figure 1. In this figure the cumulative distribution of theory and applied papers is plotted against time. With the exception of the pre-1940 period when the pooling method entered the literature, the theory and methods distribution always lies to the right of the applications distribution, and the median paper in theory (mid-1975) was published a decade after the median applied paper (mid-1964). To make the point slightly differently, we can note that by the time the median theory paper was written 80% of the applications had been published. We can conclude from this that as the applications were winding down half of the theoretical refinements were yet to come. This would not be a matter for concern if the remaining 20% of applied papers had been influenced substantively by the results made available in the postmedian theory papers, but such was not the case. The theorists and applied researchers proceeded independently of each other. We will explore the implications in the next section.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

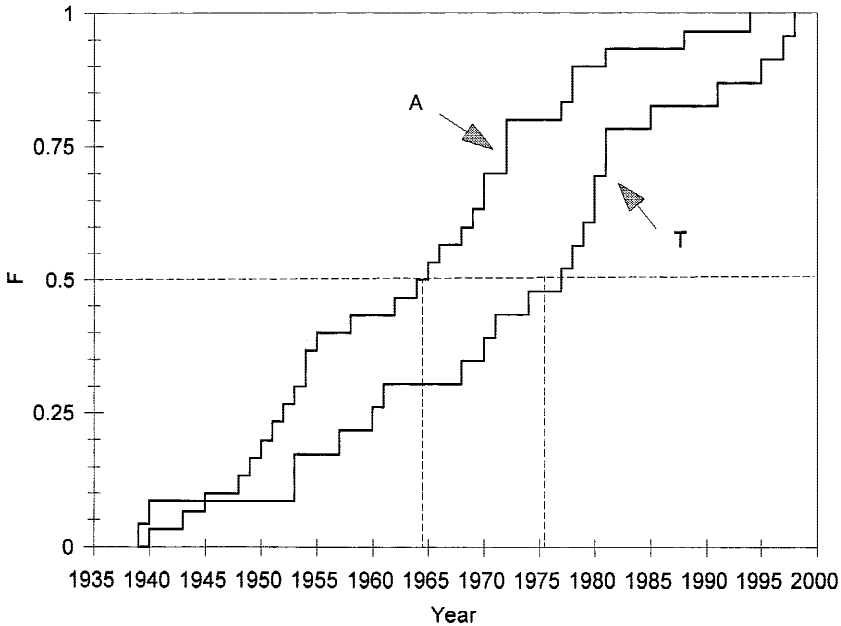


FIGURE 1. Cumulative frequency across time of applied (A) and theory (T) papers.

6. LESSONS?

Our enquiry has raised two issues relevant to the question of intellectual resource allocations. The first, arising out of the failure of applied workers to use optimal econometric procedures, can be characterized as a question of resource underutilization. The second arises out of the demonstrated disjunction between theory and practice. It is a resource allocation issue because it asks whether the marginal improvements in the theory, after applications had effectively ceased, utilized scarce resources that might have been more productively deployed elsewhere.

There are a number of factors that might account for the failure to use optimal econometric procedures. Much of the applied work was executed in the postwar econometrics boom. Many of the applied economists who turned to econometric tools would have had only limited training, if any, in statistics and mathematics, so that simple methods would invariably be preferred to complex. Furthermore, the referees of journal submissions would have been drawn from the same cohort of skills, so they would not be likely agents of change. But this argument begins to lose purchase when applied to the 1970s and 1980s, for by this time the professionalization of econometrics had occurred and explicit accounts of extraneous estimation had been incorporated into the textbooks (see, e.g., Goldberger, 1964; Johnston, 1972). Two possibilities might

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

account for the failure of the optimal procedures to be incorporated into later applications. First, the process of referee selections did not extend beyond the cohort that published the early work where the optimal procedures were ignored because the classic studies had done so. Second, the postmedian applications migrated to field journals,[13] especially agricultural economics, where referees and editors might be less interested in theoretical refinements. None of this, however, leads one away from the conclusion that editors and referees failed to monitor effectively.

It is relevant to ask whether this failure matters. It does when empirical results filter down into the policy domain. Take, for example, Hamilton's (1972) paper that attempted to determine the effects of advertising and antismoking information on cigarette demand, effects with obvious public policy implications. Hamilton concludes that banning advertising is less effective than antismoking health scare information. His statistically significant health scare dummies are based on conditional standard errors, and, because many of his *t*-ratios are only marginally above two, we are inclined to believe that statistical significance would disappear if the correct unconditional standard errors had been calculated.[14] In this example there is a link to policy decisions that should be decided on the basis of reliable empirical evidence about likely consequences of a policy choice. In the case of evidence based on extraneous estimation, the historical record suggests that such reliability is in doubt.

Given a single case study we cannot assert that the failure to use best practice technique is a systemic problem, but it does provide us with a cautionary signal that initial conditions are very important. It would appear that the strong path dependence of research will tend to propagate any questionable initial choices through the system for quite a long time.

To address the question of resource misallocation we return to the time profile of theory papers summarized in Figure 1. We can note that slightly over one-half (13/23) of the theory papers appeared after Maddala's (1971) paper completed the basic analytic framework for estimation and inference with extraneous estimators. Apart from the tenuous link between Pudney (1980) and Rossi (1988), none of the post-Maddala papers played any role in the execution of applied work after 1971. The theory had, in effect, become autonomous. The appealing idea that there is a symbiotic relationship between theory and practice is simply not confirmed by the facts, and we take this conclusion as a major finding of our enquiry. Contrary to what might be expected, much of the theory continued to find a home in top tier journals. Eight of the 13 post-Maddala papers appeared in *Econometrica*, *International Economic Review*, *Journal of the American Statistical Association*, *Journal of Econometrics*, and *Review of Economic Studies*.[15] When leading journals allocate scarce journal space to an econometric technique that is in perceptible decline, the optimality of resource allocation becomes a legitimate question.

It can be argued that to make this charge is to miss the point about the nature and function of theory. The potential benefits of any theoretical work are not

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

predictable or without beneficial spillovers into other areas, so what may appear as irrelevant at the moment may well prove to be productive in the future. Thus the observed autonomy of theory reflects accurately a reasonably efficient long-run allocation. We acknowledge the general merits of the argument, but that should not blind us to the possibility that in a specific context this argument may fail, even in the ex ante sense. One could reasonably expect that theory papers with the characteristics cited previously would have innovative features such as a new conceptual framework or technique for modeling. Examination of the post-Maddala theory papers reveals that they were concerned with the elaboration and fine-tuning of the existing theory. Among the questions addressed were diagnostics (Desai, 1974; Ghali, 1977), estimation of disturbance variance (Jeong, 1978; Bodkin and Raj, 1985), tests for pooling (Izan, 1980; Pudney, 1980; Buse, 1981; Anderson and Vahid, 1998), equivalencies among existing estimators (Passi, 1979; Havenner and Craine, 1981; Buse, 1997), and finite sample distributions of two-stage least squares with extraneous estimation (Mikhail and Ghazal, 1991; Srivastava and Ullah, 1995). It does not require the benefit of hindsight to see that these papers do not have the potential for the kind of payoff that the conventional argument presupposes. These papers are, as suggested earlier, marginal refinements to a technique that was being abandoned by applied workers. Unless one is willing to argue that a zero marginal product is the appropriate stopping rule in the refinement of a theoretical technique one has to entertain the possibility that in the case of extraneous estimation some resource misallocation occurred.

Whether this is a systemic failure in the professional culture of peer review is certainly a debatable question. Being in favor of better stopping rules is easier than implementing them. The interests of referees, the interests of authors, and the larger public good (generating a better understanding of the economy) are not automatically aligned. Nonetheless, a useful first step would be a wider recognition that the problem may exist.

NOTES

1. I have been unable to find either a copy of the November 23, 1937, letter or any other documents related to the pooling question in the Marschak Archive at UCLA. There is a sliver of evidence that Marschak was thinking about the pooling question as early as 1931. In an article (1931a, p. 300) on consumption in the *International Encyclopedia of the Social Sciences* reviewing the evidence on demand elasticities he writes, "Further progress may be looked for in a mutual control of the two methods, that of family budget statistics and that of price statistics."

2. I have not seen this monograph. The account given here is based on the statements made by Wold (1944, pp. 117–118; 1953, pp. 47, 326). Wold (1941) summarizes in a nontechnical format the main points of the 1940 study. He indicates that conditional regression was successfully used in the determination of demand elasticities.

3. In a letter dated May 5, 1947, Stone wrote as follows to T. Harrison of Mass Observation:

> I am particularly interested myself at the present time in trying to complete some sort of bridge between the market demand studies of the above kind [Stone, 1945], which relate to the behaviour of a whole body of consumers, and demand studies based on the experiences

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

of individual families or consumer units, and have been doing some work on the analysis of such behaviour insofar as it can be studied from budget investigations.

4. Wold's Lagrangian characterizations of extraneous estimation have historical antecedents worth noting. Gauss ([1823] 1995, Part II, Sect. 30) solved the problem of least squares with one restriction using the Lagrange multiplier. Aldrich (1999) has noted that Chauvenet's (1863) astronomy text of 1863 described a procedure for solving indeterminate equations by using additional information; details are given in Buse (2000).

5. In his review (1939b) Marschak chides Schultz for failing to include a role for the distribution of income. It may be worth noting for contextual purposes that Marshall (1920, pp. 87–88) constructed his illustrative aggregate market demand curve for peas by adding together the three different demand curves of three income classes. I owe this point to John Aldrich.

6. Staehle (1945, p. 248) makes a halfhearted claim to independent discovery when he writes,

> The present writer is not certain to what extent he was influenced by Marschak's suggestions in the Economic Journal when he first applied that method in 1940–41, but he wishes to draw the reader's attention to those suggestions of Marschak's and to the fact that they were published in 1939.

7. Marschak (1943, pp. 47–48) had proposed such joint estimation and even provided the sums of squares to be minimized but did not solve for the estimators. Theil and Goldberger (1961, p. 71) introduced the two equation format with an across equation restriction.

8. Because Stone differenced his data prior to estimation the correlations between regressors in his empirical work are actually quite small. Six of his preferred equations satisfy the simple framework laid out here, and the correlations between own price and real income are (1) flour (0.40), (2) cakes and biscuits (0.69), (3) eggs (0.41), (4) lard (0.50), (5) dried legumes (0.39), (6) pork (analyzed by Durbin) (0.37). None of these correlations justify a characterization of severe collinearity. Using the efficiency loss measures given in the text, losses in percentages for β_1 and β_2 are, respectively, (1) (3.82, 0.02), (2) (44.01, 9.65), (3) (7.99, 0.12), (4) (22.80, 1.32), (5) (2.16, 0.01), (6) (2.93, 1.02). The low correlations have generated substantial efficiency losses for some β_1, whereas for β_2 Stone's preference for extraneous estimators is vindicated, although not for the reasons advanced by Stone.

9. The Cowles approach to pooling time series and cross section data is given by Klein (1953). See also Klein's (1954) review of Stone's *Measurement*.

10. The problem of testing for equality of coefficients had been addressed in the statistical literature in the 1930s; see, e.g., the references in Cochran (1954). Parallel to the earlier result that mixed estimation is just another format for efficient pooling, Theil's (1963) compatibility test statistic is identical to (17); a proof is given in Buse (1997). Applying (17) to the six foods discussed in note 10 yields not a single rejection of the null, whereas the Kuh and Meyer R^2 diagnostic indicates that all six cross section and time series estimates are incompatible. The reliability of the R^2 diagnostic is very much in doubt, as is their conclusion that 90% of Stone's estimates are incompatible.

11. For the period 1965–2000 the *Social Sciences Citation Index* was used to track down applied papers in all fields, including demand analysis. Prior to 1965 we relied on year by year examination of the major journals, including the *American Journal of Agricultural Economics* (formerly the *Journal of Farm Economics*), which has a long tradition of publishing applied demand papers. Review articles on demand analysis in these journals also provided leads; details are available from the author on request.

12. The absence of an explanation as to how the standard errors were calculated does not necessarily imply that the reported standard errors are incorrect. Nonetheless, if someone has made the effort to get it right, it is only natural that they would announce this commendable effort to the reader, and no such announcements can be found in the post-Stone literature.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

13. To document this claim let the postmedian (1964) applied papers be the "later studies." Let the *American* (and *European*) *Journal of Agricultural Economics, Economic Development and Cultural Change, the Journal of Money, Credit, and Banking, the Journal of Marketing Research, the Journal of Law and Economics*, and *Regional Studies* define the set of field journals and take books and book chapters as equivalent to field publications. Then pre- and postmedian field applications were 6/15 and 11/15, respectively. There was one agricultural field publication prior to 1964 and five after 1964.

14. The difference between the unconditional and conditional standard errors can be substantial even without severe collinearity. For example, in the simple model of Section 3 if $r_{12}^2 = 0.83$ and $V_1^e/V_1^0 = 0.20$, then the unconditional errors are twice as large as the conditional. Furthermore, it is unlikely that inference decisions would be reversed with efficient estimation because our analysis in Section 3 suggests the efficiency gains are typically small.

15. We need to add that Pudney (1980) and Havenner and Craine (1981) do not have extraneous estimation as their primary focus.

REFERENCES

Aigner, D.J. & J.L. Simon (1970) A specification bias interpretation of cross section vs. time series parameter estimates. *Western Economic Journal* 8, 144–161.

Aldrich, J. (1999) Determinancy in the linear model: Gauss to Bose and Koopmans. *International Statistical Review* 67, 211–219.

Anderson, H.M. & F. Vahid (1998) On the pooling of cross-sectional and time-series data in the presence of heteroskedasticity. *Economics Letters* 60, 291–296.

Bodkin, R.G. & B. Raj (1985) An alternative estimator of the variance of the disturbance term in extraneous and mixed estimation. *International Economic Review* 26, 67–69.

Burk, M.C. (1958) Some analyses of income-food relationships. *Journal of the American Statistical Association* 53, 905–927.

Buse, A. (1981) Pooling time series and cross-section data: A minimum chi-squared approach. *Manchester School* 49, 229–244.

Buse, A. (1997) On the equivalence of pooled and mixed estimation. *Statistical Papers* 38, 231–241.

Buse, A. (2000) The History of Extraneous Estimation. Discussion paper 2000-01, Department of Economics, University of Alberta.

Carman, H.F. (1972) Improving sales forecasts for appliances. *Journal of Marketing Research* 9, 214–218.

Chauvenet, W. (1863) *A Manual of Spherical and Practical Astronomy*, vols. 1 and 2. Philadelphia: Lippincott.

Chen, N. (1972) Agricultural productivity in a newly settled region: The case of Manchuria. *Economic Development and Cultural Change* 21, 87–95.

Chetty, V.K. (1968) Pooling time series and cross section data. *Econometrica* 36, 279–290.

Cochran, W.G. (1954) The combination of estimates from different experiments. *Biometrics* 10, 101–129.

Crockett, J. & I. Friend (1960) A complete set of consumer demand relationships. In I. Friend & R. Jones (ed.), *Proceedings of Conference on Consumption and Savings* 1, pp. 1–92. Philadelphia: University of Pennsylvania.

Desai, M. (1974) Pooling as a specification error: A note. *Econometrica* 42, 389–391.

Durbin, J. (1953) A note on regression when there is extraneous information about one of the coefficients. *Journal of the American Statistical Association* 48, 799–808.

Farrell, M.J. (1954) The demand for motor cars in the United States. *Journal of the Royal Statistical Society Series A*, 117, 171–201.

Fox, K.L. (1965) A submodel of the agricultural sector. In J.S. Duesenberry, G. Fromm, L.R. Klein, & E. Kuh (eds.), *The Brookings Quarterly Econometric Model of the United States*, pp. 409–464. Chicago: Rand-McNally.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

Fuglie, K.O. (1994) The demand for potatoes in Tunisia: Are they a cereal substitute? *European Review of Agricultural Economics* 21, 277–286.
Gauss, C.F. [1823] (1995) *Theoria Combinationis Observationum Erroribus Minimus Obnoxiae, Part Two*. Translated by G.W. Stewart. Philadelphia: SIAM.
Ghali, M.A. (1977) Pooling as a specification error: A comment. *Econometrica* 45, 755–757.
Gilbert, C.L. (1991) Richard Stone, Demand theory, and the emergence of modern econometrics. *Economic Journal* 101, 288–302.
Goldberger, A.S. (1964) *Econometric Theory*. New York: Wiley.
Grabowski, H.G., J.M. Vernon, & L.G. Thomas (1978) Estimating the effects of regulation on innovation: An international comparative analysis of the pharmaceutical industry. *Journal of Law and Economics* 21, 133–163.
Hamilton, J. (1972) The demand for cigarettes: Advertising, the health scare, and the cigarette advertising ban. *Review of Economics and Statistics* 54, 401–411.
Havenner, A. & A. Craine (1981) Estimation analogies in control. *Journal of the American Statistical Association* 76, 850–859.
Hicks, J.R. & R.G.D. Allen (1934) A reconsideration of the theory of value. *Economica*, n.s., 1, 52–76.
Holmes, R.A. (1968) Combining cross section and time series information on demand relationships for substitute goods. *American Journal of Agricultural Economics* 50, 56–65.
Houthakker, H.S. & J. Tobin (1952) Estimates of the free demand for rationed foodstuffs. *Economic Journal* 62, 103–118.
Izan, H.Y. (1980) To pool or not to pool? A re-examination of Tobin's food demand problem. *Journal of Econometrics* 13, 391–402.
Jeong, K.J. (1978) Estimating and testing a linear model when extraneous information exists. *International Economic Review* 19, 541–543.
Johnston, J. (1972) *Econometric Methods*, 2nd ed. New York: McGraw-Hill.
Kaliski, S.F. (1960) Extraneous estimates and goodness of fit: A problem in empirical economics. *Manchester School* 28, 59–71.
Klein, L.R. (1953) *A Textbook of Econometrics*. Evanston: Row, Peterson.
Klein, L.R. (1954) Review of R. Stone, *The Measurement of Consumers' Expenditure in the United Kingdom*, 1920–1938. *American Economic Review* 44, 426–433.
Klein, L.R. & A.S. Goldberger (1955) *An Econometric Model of the United States 1929–1952*. Amsterdam: North-Holland.
Konandreas, P.M. & A. Schmitz (1978) Welfare implications of grain price stabilization: Some empirical evidence for the United States. *American Journal of Agricultural Economics* 1978, 74–84.
Kuh, E. & J.R. Meyer (1957) How extraneous are extraneous estimates? *Review of Economics and Statistics* 39, 380–393.
Laughhunn, D.J. (1970) Distributional effects in demand analysis: Observations and predictive tests. *Journal of the American Statistical Association* 65, 576–585.
Leamer, E.E. (1997) Revisiting Tobin's 1950 study of food expenditure. *Journal of Applied Econometrics* 12, 533–561.
Lee, F.Y. (1969) An analysis of farm household expenditures on basic living materials in Japan. *American Journal of Agricultural Economics* 51, 650–661.
Maddala, G.S. (1971) The likelihood approach to pooling cross section and time series data. *Econometrica* 39, 939–954.
Magnus, J.R. & M. Morgan (1997) Design of the experiment. *Journal of Applied Econometrics* 12, 459–466.
Manhertz, H.G. (1970) Some econometric considerations in studying regional demand patterns in the United States. *Regional Studies* 4, 107–115.
Marschak, J. (1931a) Consumption: Problems of measurement. In Edwin R.A. Seligman (ed.), *The Encyclopedia of the Social Sciences*, pp. 295–301. New York: MacMillan.
Marschak, J. (1931b) *Elastizität der Nachfrage*. Tübingen: J.C.B. Mohr.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

Marschak, J. (1939a) Sources for demand analysis: Market, budget, and income data. *Cowles Commission for Research in Economics*, Report of the Fifth Annual Conference in Economics and Statistics held at Colorado Springs, July 3–28.
Marschak, J. (1939b) Review of *Theory and Measurement of Demand*, by H. Schultz. *Economic Journal* 49, 485–487.
Marschak, J. (1939c) On combining market and budget data in demand studies: A suggestion. *Econometrica* 7, 332–335.
Marschak, J. (1943) Money illusion and demand analysis. *Review of Economic Statistics* 25, 40–48.
Marshall, A. (1920) *Principles of Economics*, 8th ed. London: Macmillan.
Mikhail, W.M. & G.A. Ghazal (1991) On a pooled estimator and its finite sample moments. *Journal of Econometrics* 48, 195–214.
Moore, T.G. (1966) The demand for Broadway theatre tickets. *Review of Economics and Statistics* 48, 79–87.
Morgan, M. (1990) *The History of Econometric Ideas*. Cambridge: Cambridge University Press.
Passi, R.M. (1979) Linear model estimation with extraneous estimates of a parameter subset, with an application to aircraft tracking. *Technometrics* 21, 511–514.
Prest, A.R. (1949) Some experiments in demand analysis. *Review of Economics and Statistics* 31, 33–49.
Pudney, S.E. (1980) Disaggregated demand analysis: The estimation of a class of non-linear demand systems. *Review of Economic Studies* 47, 875–892.
Rossi, N. (1988) Budget share demographic translation and the aggregate almost ideal demand system. *European Economic Review* 31, 1301–1318.
Saito, M. (1977) Household flow of funds equations. *Journal of Money, Credit, and Banking* 9, 1–20.
Schneider, L., B. Klein, & K.H. Murphy (1981) Governmental regulation of cigarette health information. *Journal of Law and Economics* 24, 575–612.
Schultz, H. (1933) Review of *Elastizität der Nachfrage*, by J. Marschak. *Weltwirtschaftliches Archiv* 33, 28–39.
Schultz, H. (1938) *The Theory and Measurement of Demand*. Chicago: University of Chicago Press.
Slutsky, E. (1915) Sulia teoria del bilancio des consomatore. *Giornale degli Economisti* 51, 1–26. English translation in G.J. Stigler & K.E. Boulding (eds.), *Readings in Price Theory*, pp. 27–56 (Chicago: Richard D. Irwin, 1952).
Solow, R.M. (1964) Capital, labor, and income in manufacturing. In C.L. Schultze & L. Weiner (eds.), *Behavior of Income Shares Studies in Income and Wealth*, vol. 27, pp. 101–128. Princeton: Princeton University Press.
Srivastava, V.K. & A. Ullah (1995) Efficiency properties of some estimators in pooling time-series and cross-section data. *Journal of Quantitative Economics* 11, 63–76.
Staehle, H. (1945) Relative prices and post-war markets for animal food products. *Quarterly Journal of Economics* 59, 237–270.
Stone, J.R.N. (1945) The analysis of market demand. *Journal of the Royal Statistical Society* 108, 286–382.
Stone, J.R.N. (1948) The analysis of market demand: An outline of methods and results. *Review of the International Statistical Institute* 16, 23–35.
Stone, J.R.N. (1951) The demand for food in the United Kingdom before the war. *Metroeconomica* 3, 8–27.
Theil, H. (1963) On the use of incomplete prior information in regression analysis. *Journal of the American Statistical Society* 58, 401–414.
Theil, H. & A.S. Goldberger (1961) On pure and mixed statistical estimation in economics. *International Economic Review* 2, 65–78.
Tobin, J. (1950) A statistical demand function for food in the U.S.A. *Journal of the Royal Statistical Society, Series A* 113, 113–141.
Tobin, J. (1997) Comments by Professor James Tobin. *Journal of Applied Econometrics* 12, 647–650.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press

Tomek, W.G. & W.W. Cochrane (1962) Long-run demand: A concept and elasticity estimates for meats. *American Journal of Agricultural Economics* 44, 717–730.
Whittle, P. (1952) Tests of fit in time series. *Biometrika* 9, 309–318.
Wold, H. (1940) *Efterfrågan på jordbruksprodukter och dess känslighet för prisoch inkomstförändringar*. (The demand for agricultural products and its sensitivity to price and income changes). Stockholm: Statens Offentliga Utredningar.
Wold, H. (1941) Eine Schwedische Untersuchung über die Elastizität der Nachfrage. *Metron* 14, 263–285.
Wold, H. (1944) A synthesis of pure demand analysis, part III. *Skandinavisk Aktuarietidskrift* 27, 69–120.
Wold, H., in association with L. Jureen (1953). *Demand analysis*. New York: Wiley.
Wold, H. (1982) Models for knowledge. In J. Gani (ed.), *The Making of Statisticians*, pp. 189–212. New York: Springer-Verlag.

https://doi.org/10.1017/S0266466603191049 Published online by Cambridge University Press