12-1-1985

Identification by Disaggregation

Matthew J. Cushing
University of Nebraska-Lincoln, MCUSHING1@UNL.EDU

Mary G. McGarvey
Univerisity of Nebraska-Lincoln, MMCGARVEY1@UNL.EDU

Published in *The American Economic Review,* Vol. 75, No. 5 (Dec., 1985), pp. 1165-1167.

Published by the American Economic Association.

Identification by Disaggregation

By MATTHEW J. CUSHING AND MARY G. MCGARVEY*

Standard economic theory predicts that the actions of individual participants in competitive markets have negligible effects on market-determined aggregates. Applied researchers,[1] and even some econometric textbooks,[2] incorrectly infer from this that market prices can be modeled as econometrically exogenous with respect to the quantity demanded of an individual consumer.[3] This faulty inference has even led some researchers (for example, Robert Engle, 1978; Nicholas Kiefer, 1984; Roger Waud, 1974) to employ an estimation strategy we call identification by disaggregation (*IBD*). This procedure attempts to circumvent the simultaneity problem in a macro regression by disaggregating the dependent variable and estimating the relationship for individual agents or sectors. This note provides a simple proof that estimates using disaggregated dependent variables suffer, on average, from the same degree of simultaneity bias as the estimates using aggregate data.

Let Y be a $T \times 1$ vector of T observations on a macro variable and X be a $T \times k$ matrix of T observations on k macro variables. For example, Y and X could be aggregate output growth and money growth. Let the aggregate relationship between Y and X be

$$(1) \qquad Y = XB + e,$$

where all variables are deviations from their means and $\text{plim}(X'X/T)^{-1} = Q$ and $E(e) = 0$. Assume that there is some feedback from X to Y so that $\text{plim}(X'e/T) = P \neq 0$. The ordinary least square (*OLS*) estimation of B from (1) yields the inconsistent estimator $\hat{B} = (X'X)^{-1}X'Y$, where the inconsistency is

$$(2) \qquad \text{plim}(\hat{B} - B) = QP.$$

It is often argued that the inconsistency in sectoral regressions is smaller. For example, Waud argues that since the feedback from employment in a particular industry to aggregate money growth should be minimal, the simultaneous equations bias in sectoral regressions should be smaller than that in the aggregate regression.[4]

From (1) we can write the disaggregated relationships between the Y_i and X as

$$(1') \qquad Y_i = XB_i + e_i$$

where $Y = \sum_{i=1}^{n} Y_i/n, \quad B = \sum_{i=1}^{n} B_i/n,$

$$e = \sum_{i=1}^{n} e_i/n$$

and n is the number of sectors considered. The *OLS* estimator of B_i from (1′) is $\hat{B}_i = (X'X)^{-1}X'Y_i$. It is immediately obvious that

*Departments of Economics, Emory University, Atlanta, GA 30322, and Georgia State University, Atlanta, GA 30303, respectively. This note has benefited from helpful comments from Mark Meador and an anonymous referee.

[1] Nicholas Kiefer uses household data to estimate the Rotterdam demand model, arguing "...according to the usual arguments the simultaneity problem is not present. Surely supply to an individual is perfectly elastic..." (1984, p. 288).

[2] "While studying the demand for gasoline by households, we can treat the quantity demanded as endogenous and income and price as exogenous, arguing that the household does not have control over these" G. S. Maddala (1977, p. 5).

[3] The premise is that individual agents or sectors do not control the aggregate variables. However, econometric exogeneity can fail if the error term of the individual behavioral equation is merely correlated with the aggregate variables. For a careful discussion of alternative definitions of econometric exogeneity and their usefulness for inference, see Robert Engle, David Hendry, and Jean-Francois Richard (1983).

[4] However, Waud admits "...it is very difficult to establish unequivocally that the reduced form approach used here definitely reduces single-equation least squares bias, even though it seems likely that it does" (p. 186).

the *OLS* estimator $\hat{B}$ from (1) is simply the average of the *OLS* estimators $\hat{B}_i$ from (1′). Defining $\text{plim}(X'e_i/T) = P_i$, the inconsistency of any $\hat{B}_i$ is

$$(2') \qquad \text{plim}(\hat{B}_i - B_i) = QP_i.$$

Clearly, the inconsistency of the aggregate estimator, (2), is simply the average inconsistency of the sectoral estimators, (2′), over the n sectors. Of course, estimates from sectors with lower than average P_i's exhibit smaller inconsistencies than the aggregate estimator. However, some prior information is required in order to identify these sectors. If sectors are chosen at random, the expected inconsistency of the sectoral estimator is identical to the inconsistency of the aggregate estimator.

The above result is a consequence of the linearity of ordinary least squares. The result holds for any linear estimator (for example, *OLS*, *GLS* with a known covariance matrix, and estimation under linear restrictions). Let $\hat{B} = A'Y$, where A is some $T \times k$ linear transformation matrix. The corresponding sectoral estimator is $\hat{B}_i = A'Y_i$. It is clear that $\hat{B}$ is simply the average of the $\hat{B}_i$'s so that the inconsistency of $\hat{B}$ is the average of the $\hat{B}_i$'s inconsistencies.[5]

The results of this paper allow us to reevaluate the discussion of identification by disaggregation contained in Thomas Cooley and Stephen LeRoy (1981). By working through a specific example we can see how the intuition behind *IBD* fails.

Consider a simple, aggregate money demand function (in deviations from means),

$$(3) \qquad m = ar + e,$$

where m is defined as average money balances, $m = \sum_{i=1}^{n} m_i/n$, and n is the number of sectors. Let the money supply function follow a feedback rule of the form

$$(4) \qquad r = bm + v.$$

[5]Although we consider only linear estimation techniques, there is no reason to believe that disaggregating a nonlinear macro relationship and estimating the sectoral relationships would reduce the simultaneity bias.

To avoid the obvious simultaneity bias in estimating a with aggregate data, Cooley and LeRoy explore the possibility of estimating (3) using sectoral money stock data. The demand for money in the ith sector can be modeled as

$$(3') \qquad m_i = a_i r + e_i.$$

Assume that the sectoral shock e_i is the sum of two mutually uncorrelated components: w, a shock common to all sectors with variance σ_w^2, and u_i, a shock specific to sector i and uncorrelated across sectors with variance σ_i^2. The inconsistency of the sectoral *OLS* estimator is (Cooley-LeRoy, p. 840)

$$(5') \quad \text{plim}(\hat{a}_i - a_i) = \text{plim}(r'e_i/T)/\text{plim}(r'r/T) = b(\sigma_w^2 + \sigma_i^2/n)(1 - ba)/(b^2\sigma_e^2 + \sigma_v^2).$$

Cooley and LeRoy observe that the inconsistency of the sectoral estimator will be small if there are a large number of sectors (n is large) and the error terms (e_i) are mutually uncorrelated (i.e., the common factor is zero). Cooley and LeRoy argue that, in this particular case, *IBD* will not resolve the identification problem because money stock data is available for only four sectors and the assumption of a zero common component is implausible. However, their discussion leaves the impression that, although *IBD* is not useful in this case, it may be useful in other circumstances.

The results of this paper provide a much stronger and more general criticism of this attempt at identification by disaggregation. It is true that if the common shock is zero ($\sigma_w^2 = 0$) and n, the number of sectors, is large (holding σ_i^2 constant), the inconsistency of the sectoral estimator will be small. However, consider the inconsistency of the aggregate estimator

$$(5) \quad \text{plim}(\hat{a} - a) = \text{plim}(r'e/T)/\text{plim}(r'r/T) = (1 - ba)b\sigma_e^2/(b^2\sigma_e^2 + \sigma_v^w).$$

Recall that the aggregate shock *e* is the average of the sectoral shocks. If the number of sectors is large and the shocks are independent, the variance of the aggregate shock must be small. As such, the inconsistency of the aggregate estimator, (5), is clearly small. The conditions under which the inconsistencies of the sectoral estimators are small guarantee that the inconsistency of the aggregate estimator is small. Identification by disaggregation is a strategy that succeeds only when it is unnecessary.

REFERENCES

Cooley, Thomas F. and LeRoy, Stephen F., "Identification and Estimation of Money Demand," *American Economic Review*, December 1981, *71*, 825–44.

Engle, Robert F., "Testing Price Equations for Stability Across Spectral Frequency Bands," *Econometrica*, July 1978, *46*, 869–81.

______, Hendry, David F. and Richard, Jean-Francois, "Exogeneity," *Econometrica*, March 1983, *51*, 277–304.

Kiefer, Nicholas M., "Microeconometric Evidence on the Neoclassical Model of Demand," *Journal of Econometrics*, July 1984, *25*, 285–302.

Maddala, G. S., *Econometrics*, New York: McGraw-Hill, 1977.

Waud, Roger N., "Monetary and Fiscal Effects on Economic Activity: A Reduced Form Examination of their Relative Importance," *Review of Economics and Statistics*, May 1974, *56*, 177–87.