A Note on Supply and Demand Functions.

By R. C. Geary

In a recent very useful paper* A.S. Goldberger calls attention to the fundamental, but largely ignored, work of Sewall Wright in the difficult and contentious field of identification in demand and supply equations.

In 1928 it appears that Wright and his father P.G. Wright took the classical static supply-demand model –

$$(1) \quad q^D = \alpha p + u, \quad q^S = \beta p + \tilde{v}, \quad q^D = q^S \ (= q).$$

They state that "the elasticities of supply and demand cannot be computed from price, output and consumption data alone (ESRI practitioners please note !) ... Elasticity of supply (demand) can be computed only when assurance is obtained that the cost (demand) curve remains fixed while the demand (cost) curve is changing its position...." The latter language, even in the light of later work by S. Wright and explanation by Goldberger, is somewhat obscure. The object of this note is to try to make it clearer. The Wright's algebra helps to show what is meant. Suppose we have another variable z – we may assume that we are dealing with time series for all variables – for which $C(z, v)$ (C = covariance) may be assumed zero but $C(z, u)$ is not zero. We have another variable x with the opposite properties $C(x, u)$ is zero but $C(x, v)$ is not zero. Then –

$$(2) \quad \beta = C(q, z)/C(p, z); \quad \alpha = C(q, x)/C(p, x).$$

We need not trouble with estimation symbols, $\hat{\beta}$ etc. The point is that z and x are revealed as instrumental variables in which field O Reiersøl and R. C. Geary were early researchers. (Though irrelevant Geary may

recall a result: if, in the notation of (2), variables (p, q, z) are distributed on the 3-dimensional normal surface, $f(\beta)$, the function f being known in form, is distributed as Student-Fisher t.) Of course, in (2), the covariances involved in the denominators are deemed to be significantly different from zero. The Wrights applied their method to estimation of price elasticities of demand and supply of butter anf flaxseed to find for (α , β), (-0. 6, 1. 4) and (-0. 8, 2. 4), plausible as having the "right" signs.

Suppose that, instead of (1), our system is –

Demand : $q^D = \alpha \ p + \underset{\sim}{\Upsilon} '\underset{\sim}{z} + u'$

(3) Supply : $q^S = \beta \ p + \underset{\sim}{\delta} '\underset{\sim}{x} + v \ '$

$\qquad q^D = q^S = q$

The scalar $\underset{\sim}{\Upsilon} '\underset{\sim}{z}$ will contain as many as required of demand exogenous variables (exos). Similarly as regards the $\underset{\sim}{\delta} '\underset{\sim}{x}$ of the supply equation. The vectors $\underset{\sim}{z}$ and $\underset{\sim}{x}$ may have some elements in common, i.e. variables which can be regarded as both demand and supply (like p itself). However, each equation is assumed to be identifiable.

If model (3) be set up having full regard to economic theory and the coefficients $\underset{\sim}{\Upsilon}$, $\underset{\sim}{\beta}$, $\underset{\sim}{\delta}$ estimated by FIML there are tests to show whether the model is a satisfactory representation of the data: these bear on the new residuals, u' and v '. Having regard to (1) we may set –

$\qquad u = \underset{\sim}{\Upsilon} '\underset{\sim}{z} + u'$

(4) $\qquad v = \underset{\sim}{\delta} '\underset{\sim}{x} + v \ '$

If we have done our work properly we may plausibly assume that u' and v ' are completely random variables, algebraically interpretable as that

their covariance with one another and with all other variables is zero.

With Goldberger, from now on we regard $\Upsilon'\underset{\sim}{z}$ and $\underset{\sim}{\delta}'\underset{\sim}{x}$

as single terms, Υz and δx respectively: z and x may be regarded as

genuinely single specific variables related to p and q, or proxies (perhaps

principal components) for _sets_ of variables. Goldberger's interpretation

of Sewall Wright's 1934 position contemplates the model –

$$(i) \quad q = \alpha\, p + \Upsilon z + u'$$

$$(5) \qquad (ii) \quad q = \beta\, p + v$$

(We bypass for the moment the difficulty that if (5) is a system of 2 equations

with 2 endos p and q and u' and v disturbances, the first equation is formally

not identified.) The instrument z, we are told, might be "the price of a sub-

stitute or an index of prosperity". The property of non-shifting of the supply

curve is algebraically interpreted as $C(z, v) = 0$.

As Goldberger points out, in (5) there are 5 estimable moments,

$V(q)$ [V = variance] , $V(p)$, $C(q,p)$, $C(z,q)$, $C(z,p)$ and 6 parameters α ,

β , Υ , $V(u')$, $V(v)$, $C(u', v)$. The system is formally under-identified.

However, as already indicated, we have no difficulty about assuming that

$C(u', v)$ is zero so we are left with 5 parameters and (in effect) 5 linear

equations for which, in general, there is an unique solution – set, perfectly iden-

tifiable.

So far Goldberger. As a comment: The coefficients α and

β in (5) are estimated as follows, using the data and properties specified.

Multiply (5) (ii) across by z, sum and average. Hence $\beta = C(z,q)/C(z,p)$

since $C(z, v) = 0$. Hence $v = q - \beta\, p$ is known. Multiplying (5) (i)

across by v , so calculated, summing and averaging $C(q, v) = \alpha\, C(p, v)$

since $C(z, v)$ and $C(u', v)$ are zero. Compare these formula with those

in the text for model (1).

I do not see how to estimate Υ without a further assumption. The most natural is perhaps that, in (5) (i), $C(z, u')$ is zero, whence -

$$C(q, z) = \alpha C(p, z) + \Upsilon V(z),$$

giving Υ , since α is already known.

According to Goldberger, Wright states that "if the assumption [that $C(u', v) = 0$] is not justified" we use the fuller model -

$$(6) \qquad q = \alpha p + \Upsilon z + u'$$
$$q = \beta p + \delta x + v'$$

with 7 estimable moments, $V(q)$, $V(p)$, $C(q,p)$, $C(z,q)$, $C(z,p)$, $C(x,q)$, and to determine 7 parameters, α , β , Υ , δ , $V(u')$, $V(v')$ and $C(u', v)$. Goldberger, of course, indicates the difficulty that $C(u', v')$ almost by definition, is likely to be zero, so that there are, in reality, only 6 parameters. The system is overidentified.

At the time (ca. 1940) Reiersøl and Geary, independently and almost simultaneously, produced a purely algebraic (i. e. with no regard to economics) solution there was little consciousness then of systems of equations (e.g. 2 in the supply-demand case) and their special problems, identification in particular. Our problem was the estimation of coefficients in the single equation model. Nor were we troubled about the distinction between endos and exos, and causation. The single linear model is -

$$(7) \qquad \sum_{i=1}^{k} \beta_i x_{it} + u_t = 0, \ t = 1, 2 \ldots, T,$$

t being time or cross-section. We assumed that we were given a large number of variables X_i, with $x_{it} = X_{it} - \overline{X}_i$. One β_i in (7) (the _numeraire_ is supposed

known, say unity. The problem was to estimate the remaining (k-1) variables.

The k variables in (7) are called the <u>equation</u> set.
We assume also available k' additional variables, where k' ⩾ k - 1, say
x_{k+j}, j = 1, 2, ..., k', the <u>instruments.</u> Pick any (k - 1) of these, numbering them consecutively from j = 1. Multiply (7) across by $x_{k+j,t}$ and
sum for t -

(8) $\qquad \sum_{i=1}^{k} \beta_i C(x_i, x_{k+j}) + C(u, x_{k+j}) = 0$

We now assume that the disturbance u in (7) is a random variable, (as random
as we can make it !) so that the last C on the left of (8) may be assumed to be
not significantly different from zero so that the equation system is -

(9) $\qquad \sum_{i=1}^{k} \beta_i C(x_i, x_{k+j}) = 0$, j = 1, 2, ..., k - 1,

from which the (k - 1) values of β_i will be estimated.

The method is ruthlessly empirical. We assume
given (k + k')= K variables to start with. We would then try out (7) for
k = 2, 3 etc. the test of <u>completeness</u> of relationship being that disturbance
u was random. Geary even evolved large sample significance tests for the
estimates, unfortunately inoperable in the pre-computer age, but perhaps
worthy of a glance, as to practability, nowadays. There was even an
asymptotic theory for determining how many linear equations of the type
required in the K sets, using an approximate chi-squared test but unfortunately not identifying the variables in each equation.

Clearly Sewall Wright would have used equation
systems like (9) in his theory.

The Reiersøl-Geary theory was evolved mainly to deal with the errors in variables situation: all equation variables are regarded as subject to error. Regression is a particular case of the R-G theory, namely that of one variable being subject to error (i. e. the depvar) all the rest (i. e. the indvars) being accurately observed. The standard equations of OLS regression are those of (9) in which the (k - 1) instrumental variables are the indvars.

I do not think that there is any need for recourse to Reiersøl-Geary to evolve demand and supply equations, for here we can regard ourselves as in a cause-effect situation. The equations at (3) are to be regarded as explaining q in two entirely different ways. If economic theory be properly applied there is no very compelling reason why each equation should not be solved separately. No doubt a considerable degree of experimentation (i. e. what indvars to include in $\underset{\sim}{z}$ and $\underset{\sim}{x}$) will normally be required. Only if asymptotic efficiency of estimation is required need recourse be had to FIML applied to the 2-equation system, for forecasting, for example. Truth to say, there is little to be gained, and some lack of clarity to be lost, by such rigid statistical rectitude.

Most of the trouble that arose in connection with estimation of demand and supply equations was due to trying to deal with form (1), instead of form (3). As we now recognise, neither equation in (1) is identified. We had only one classical way of tackling the problem, namely OLS of q on p (or their logs). Only one equation transpired and there was vast disputation as to whether it was a demand equation or a supply equation: in general it was neither, if with a leaning towards demand as frequently yielding a negative regression coefficient (i. e. price increases associated with quantity decreases and _vice_ versa).

I agree with Goldberger that the ignoration of Sewall Wright was tragic for its postponement of the application of correct econometric practice in this field. Henry Schultz is shown to be particularly at fault in this regard, for he was well aware of Wright's approach, as Goldberger shows. Although Schultz's work The Theory and Measurement of Demand (1938) is always dubbed "classical", I recall finding it turgid and generally unsatisfactory.

3 October 1973   R. C. Geary.