Excerpt from: "Structural Relations, Cointegration and Identification: Some Simple Results and Their Application" (Working Paper version, 1996)

This appeared as Section 2.4 of the working paper, with Section 2.4 of the published version (*Journal of Econometrics* 87, pp87-113 (1998)) as Section 2.5.

2.4 Missing Variables

An important question is the role of these results under mis-specification. It has been assumed thus far that the variables in the investigation are linked by a closed dynamic system of form (2.5), in which β has a structural interpretation. We should also know how to interpret regressions when some variables that appear in the long-run solution of (2.5) are missing from the data set. Of course, those identified structural relations whose included variables are present still appear in the IC set. The complicating factor is that the IC vectors may include solved relations derived from structural relations that do not themselves appear in the set. Take example (2.7) again, under the assumptions $\beta_{41} = \beta_{32} = 0$, but suppose that q_t is unobserved. Among the remaining three variables we find a single IC relation that, as we know, is a reduced form, but is derived from a structural relation that does not exist for the included data. This means that Theorems 4 and 5 do not always apply to incomplete models, although note that Theorems 1, 2 and 3 are true in all circumstances.

The omission of variables is innocuous from the point of view of applying Theorems 4 and 5 if β has a block-diagonal form such that no structural vector combines missing and observed variables. A second possibility is that an omitted variable appears in a single structural relation only. In this case, there can be no solved forms involving this particular relation (see the proof of Theorem 5 for verification) and its omission is once again innocuous from the point of view of applying Theorems 4 and 5. Again take (2.7) with $\beta_{41} = \beta_{32} = 0$ as the example, but now suppose r_t is unobserved. This appears only in the supply relation, and hence is in both the solved forms. Amongst the observed variables there is only a single IC relation, and this is the demand relation.

A third possibility is that a missing variable appears in two or more structural relations involving observed variables. While this invalidates Theorems 4 and 5, it presents the possibility of using one or more of these relations to generate a proxy for the missing variable. In this case, certain solved relations are elevated to the status of quasi-structural forms, in which the missing variable is substituted out. We can decide to treat such relations as structural for the purposes of the exercise, and redefine β accordingly. Take the example of (2.7) with $\beta_{22} = \beta_{41} = \beta_{32} = 0$ as the example, and q_t missing. Let r_t proxy for q_t, and note that one solved relation takes the form

$$\beta_{21} p_t + \beta_{31} w_t - \beta_{42} r_t \sim I(0)$$

Note that the coefficients of p_t and w_t are the same as in the demand relation (in any normalization their relative magnitudes are the same) and in this sense the designation 'quasi-structural' is appropriate. This property will hold whenever the observed variables in the relation do not also appear in the proxy. Note that if the relations containing the missing variable are identified, and hence IC in the complete system, the solved forms containing the proxy are also IC, and hence, might be thought of as

identified structural relations in the redefined system. Theorems 4 and 5 now apply to this system as before.

The same argument applies to any number of missing variables that can be individually proxied. Needless to say, the resulting quasi-structures are not always interesting, and the first of the cases given above (missing q_t in the just-identified model) is a case in point. One can conceive of more difficult cases, such as two or more missing variables only appearing in conjunction, with different weights. The general principle remains that with missing variables, one has the option of designating some solved forms as structural, and in the context of such a reparameterization, the results of Section 2.1 remain valid.