

(1990). A skeletal form of the model can be written as

$$y_{1i} = \alpha_0 + \beta_1 y_{2i} + \beta_2 z_i + u_{1i} + \eta_i \quad (32)$$

$$y_{2i} = \gamma_0 + \gamma_1 y_{1i} + \gamma_2 z_i + u_{2i}$$

where  $y_1$ ,  $y_2$ , and  $y_a$  are endogenous variables,  $z$ , and  $z_2$  are vectors of exogenous variables,  $u_1$  and  $u_2$  are error terms, and  $\eta_i$  is unobserved heterogeneity. In the Rosenzweig and Schultz papers, the first equation explains the number of births in terms of the endogenous contraceptive effort  $y_2$ , so that  $g_{i,t}$  is couple-specific fecundity. The second equation is used to explain various characteristics of child health which are also influenced by latent fecundity. In the Pitt, Rosenzweig and Hassan paper, which is concerned with nutritional status and consumption, the first equation relates weight for height to calorie consumption (the two endogenous variables) and an individual "endowment"  $A_i$ . In this case,  $y_3 = y_2$ , which is calorie consumption, and the parameter  $\delta$  measures the extent to which the household reinforces ( $\delta > 0$ ) or offsets ( $\delta < 0$ ) natural endowments in the intrahousehold allocation of food.

As always with MIMIC models, the major issue is identification, and strong assumptions are required to be able to recover  $\theta$ . Provided that the  $p$ 's and  $y$ 's are identified — which poses no non-standard issues —  $\theta$  is identified from the covariance matrix of the residuals provided that  $u_1$  and  $u_2$  are orthogonal — which requires that there be no common omitted variables in the two equations — and provided the instruments are orthogonal to the unobservable  $\eta_i$ , a set of conditions that would seem to be indefensible in any real application. In practice, (32) is usually estimated by applying instrumental variables to the first equation and then using the residuals as a regressor in the second equation, with some allowance for the "measurement error" that comes from the presence of  $u_1$  in the first equation. (Note also that without correction, such a two-step procedure will not generally lead to valid standard errors.) An alternative (and more direct) procedure would be to substitute for

$y_{2i}$  in the second equation from the first, and to estimate the resulting equation by instrumental variables.

A second example comes from my own work on estimating price elasticities of demand using the spatial price variation that is revealed in cross-sectional household surveys when respondents report, not only how much they have spent, but also the physical quantity bought, so that for each household we can construct a unit value index. This unit value index reflects both local prices and the quality choices of individuals, with unit values usually higher for better-off households who purchase more expensive varieties, even of relatively homogeneous goods. A stripped-down version of the model can be written as follows,