sociological science

Supplement to:

Breen, Richard, Seongsoo Choi and Anders Holm. 2015. "Heterogeneous Causal Effects and Sample Selection Bias." Sociological Science 2: 351-369.

Derivation of Equation (15)

Starting with Equation (14),

$$\frac{\int_{-\infty}^{\infty}\int_{-\gamma Z/\sigma_u}^{\infty} \varepsilon f(\varepsilon,u,\theta)\partial u\partial\varepsilon}{\int_{-\infty}^{\infty}\int_{-\gamma Z/\sigma_u}^{\infty} f(\varepsilon,u,\theta)\partial u\partial\varepsilon} - \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{-\gamma Z/\sigma_u} \varepsilon f(\varepsilon,u,\theta)\partial u\partial\varepsilon}{\int_{-\infty}^{\infty}\int_{-\infty}^{-\gamma Z/\sigma_u} f(\varepsilon,u,\theta)\partial u\partial\varepsilon}$$

the first term in the denominator can be written

$$\int_{-\infty}^{\infty}\int_{-\gamma Z/\sigma_u}^{\infty} f(\varepsilon,u,\theta)\partial u\partial\varepsilon$$
$$= \int_{-\gamma Z/\sigma_u}^{\infty} f(u)du = \Pr\left[u > -\frac{\gamma Z}{\sigma_u}\right] = 1 - F\left[-\frac{\gamma Z}{\sigma_u}\right] = p(Z)$$

and similarly for the other part of the denominator:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{-\gamma Z/\sigma_u} f(\varepsilon,u,\theta)\partial u\partial\varepsilon = 1 - p(Z)$$

Then we have, for the whole expression,

$$\frac{\int_{-\infty}^{\infty}\int_{-\gamma Z/\sigma_u}^{\infty} \varepsilon f(\varepsilon,u,\theta)\partial u\partial\varepsilon \times (1-p(Z)) - \int_{-\infty}^{\infty}\int_{-\infty}^{-\gamma Z/\sigma_u} \varepsilon f(\varepsilon,u,\theta)\partial u\partial\varepsilon \times p(Z)}{p(Z)(1-p(Z))}$$

and, because $E(\varepsilon) = 0$ this reduces to

$$\frac{\int_{-\infty}^{\infty} \int_{-\gamma Z/\sigma_u}^{\infty} \varepsilon f(\varepsilon, u, \theta) \partial u \partial \varepsilon}{p(Z)(1 - p(Z))} = \frac{\int_{-\infty}^{\infty} E(\varepsilon|u) f(u) du}{(1 - p(Z))}$$

as required. In the case where ε and u are bivariate normal, we can write the conditional expectation of ε given u as

$$\rho_{\varepsilon,u} \frac{\sigma_\varepsilon}{\sigma_u} \left[\frac{\phi\left(-\frac{\gamma Z}{\sigma_u}\right)}{\Phi\left(-\frac{\gamma Z}{\sigma_u}\right)} \right] = \rho_{\varepsilon,u} \frac{\sigma_\varepsilon}{\sigma_u} \left[\frac{\phi\left(\Phi^{-1}\left(p(Z)\right)\right)}{p(Z)} \right]$$

Substituting this into the previous expression gives us Equation (13) (under the usual assumption that u has a standard deviation of 1).