Marginal Effects for Probit and Tobit with Endogeneity

Kirill S. Evdokimov* Ilze Kalnina† Andrei Zeleneev‡§

This version: February 28, 2023

Abstract

We consider estimation of the partial effects for the Instrumental Variable (IV) Probit and Tobit estimators. It is important to distinguish between structural endogeneity and measurement errors: two sources of endogeneity that affect the partial effects differently. We show that failing to account for this difference leads to either attenuation or overstatement of the magnitude of the partial effects by these IV estimators. We develop simple estimators of the bounds on the partial effects allowing for both types of endogeneity.

Keywords: (Average) Partial Effects, Instrumental Variables, Control Variables

1 Introduction

Probit and Tobit are some of the most popular nonlinear models in applied economics. When a covariate is endogenous, IV-Probit and IV-Tobit models are used for instrumental variable (IV) estimation of the coefficients (Smith and Blundell, 1986, Rivers and Vuong, 1988).[1]

A covariate can be endogenous for two reasons. First, the covariate can be correlated with the individual's unobserved characteristics (Unobserved Heterogeneity, UH). Second, mismeasurement of the covariate (Errors-in-Variable, EiV) also results in endogeneity. We will refer to these two types of endogeneity as the "structural" endogeneity and the EiV. In many empirical settings both sources of endogeneity need to be addressed simultaneously.

*Universitat Pompeu Fabra and Barcelona School of Economics: kirill.evdokimov@upf.edu.

†North Carolina State University: ikalnin@ncsu.edu.

‡University College London: a.zeleneev@ucl.ac.uk.

§Evdokimov gratefully acknowledges the support from the Spanish MCIN/AEI via grants RYC2020-030623-I, PID2019-107352GB-I00, and Severo Ochoa Programme CEX2019-000915-S.

[1] For example, in Stata, these estimators are *ivprobit* and *ivtobit*.

The goal of this note is to characterize the partial effects in IV-Probit and IV-Tobit models, allowing for both types of endogeneity, and to emphasize the importance of distinguishing between the two types. We provide the expressions for the partial effects and average partial effects that correctly account for the two kinds of endogeneity. Although the two sources of endogeneity cannot be precisely distinguished using the observed data, we use the constraints of the model to obtain bounds on the amounts of endogeneity that can be attributed to each source. We use these bounds to provide simple estimators of the bounds on the partial and average partial effects.

The need to differentiate between the two kinds of endogeneity arises because in nonlinear models, structural endogeneity and EiV play different roles. In particular, partial effects of covariates are averaged with respect to the distribution of the individual unobserved heterogeneity. On the other hand, one aims to remove the impact of the measurement errors, since they are not properties of individuals but a deficiency in the measurement process. Thus, even though IV-Probit and IV-Tobit models consistently estimate the coefficients on all regressors regardless of the source of endogeneity, the effects of the covariates on the outcomes are only partially identified, since the distribution of the unobserved heterogeneity is partially identified. We find that ignoring the distinction between the two types of endogeneity leads to either attenuation or overstatement of the magnitude of the partial effects by these IV estimators, depending on the values of the parameters.[2]

IV-Probit and IV-Tobit are control variable estimators. Partial effects in general control variable models were considered by Blundell and Powell (2003), Chesher (2003), Imbens and Newey (2009), and Wooldridge (2005, 2015), among others. In nonlinear models, accounting for both types of endogeneity is difficult, see, e.g., Schennach (2022), and it is not clear how to characterize the bounds on the partial effects in practice. The assumption of gaussianity of the unobservables simplifies the analysis of IV-Probit and IV-Tobit models and allows us to obtain simple bounds on the partial effects that are easy to implement.

This note is organized as follows. We first present the analysis for the Tobit model in Sections 2-4. Then we explain how the methods apply to the Probit model in Section 5.

2 The Model

Tobit model is often used for estimation of economic models with a "corner solution," i.e., models where the outcome variable Y_i is forced to be non-negative. The examples of such

[2] Wooldridge (2010), page 586, alludes to the potential importance of the sources of endogeneity for the partial effects in IV-Probit, but does not elaborate.

dependent variables Y_i include the amounts of charitable contributions, hours worked, or monthly consumption of cigarettes.

First, consider the standard Tobit model with exogenous covariates and without EiV:

$$Y_i = m\left(\theta_{01} X_i^* + \theta_{02}' W_i + U_i^*\right), \quad \text{where } m\left(s\right) = \max(s, 0), \tag{1}$$

the individual unobserved heterogeneity U_i^* has a normal distribution $N\left(0, \sigma_{U^*}^2\right)$ and is independent from the covariates X_i^* and W_i. We use the asterisk to denote variables that will be affected by the EiV, as we explain in detail below.

We collect the covariates in a vector $H_i^* = (X_i^*, W_i')'$, so (1) can be written as $Y_i = m\left(\theta_0' H_i^* + U_i^*\right)$. The standard normal cumulative distribution and density functions are denoted by Φ and ϕ, respectively.

In Tobit model, one is usually interested in the partial effects (marginal effects) of covariates H_i^* on $\mathrm{E}\left(Y_i|H_i^*\right)$ and $\mathrm{P}\left(Y_i > 0|H_i^*\right)$. For brevity, we focus on the partial effects of the continuously distributed covariates.

The partial effect of the j^{th} covariate on the mean $\mathrm{E}\left(Y_i|H_i^* = h\right)$ at a given h is

$$PE_{h_j}^{\mathrm{Tob}}\left(h\right) = \frac{\partial}{\partial h_j} \mathrm{E}\left(Y_i|H_i^* = h\right) = \Phi\left(\frac{\theta_0' h}{\sigma_{U^*}}\right)\theta_{0j}. \tag{2}$$

The partial effect of the j^{th} covariate on the probability $P\left(Y_i > 0|H_i^* = h\right)$ is

$$PE_{h_j}^{\mathrm{Pr}}\left(h\right) = \frac{\partial}{\partial h_j} \mathrm{P}\left(Y_i > 0|H_i^* = h\right) = \phi\left(\frac{\theta_0' h}{\sigma_{U^*}}\right)\frac{\theta_{0j}}{\sigma_{U^*}}. \tag{3}$$

These formulas for the PE_{h_j} are standard, see, e.g., Wooldridge (2010), for detailed calculations. Most often one considers the partial effects at the means of the covariates $h = E\left[H_i^*\right]$.

In addition, researchers are often interested in the Average Partial Effects

$$APE_{h_j}^{\mathrm{Tob}} \equiv \mathrm{E}\left[PE_{h_j}^{\mathrm{Tob}}\left(H_i^*\right)\right] \text{ and } APE_{h_j}^{\mathrm{Pr}} \equiv \mathrm{E}\left[PE_{h_j}^{\mathrm{Pr}}\left(H_i^*\right)\right], \tag{4}$$

which are the pointwise partial effects $PE_{h_j}\left(h\right)$ averaged with respect to the distribution of H_i^*.

When X_i^* is correlated with U_i^* and we observe data (Y_i, X_i^*, W_i, Z_i), IV-Tobit model can be estimated using instrumental variables Z_i, as proposed by Smith and Blundell (1986) and

Rivers and Vuong (1988). Assume that

$$
\begin{aligned}
Y_i &= m\left(\theta_{01} X_i^* + \theta_{02}' W_i + U_i^*\right), \quad m(s) = \max(s, 0), && (5)\\
X_i^* &= \pi_{01}' Z_i + \pi_{02}' W_i + V_i^*, && (6)
\end{aligned}
$$

where V_i^* is a normal random variable, possibly correlated with U_i^*:

$$
\begin{pmatrix} U_i^* \\ V_i^* \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{U^*}^2 & \sigma_{U^*V^*} \\ \sigma_{U^*V^*} & \sigma_{V^*}^2 \end{pmatrix}\right). \tag{7}
$$

IV-Tobit model in (5)-(7) is typically estimated using a random sample of (Y_i, X_i^*, W_i, Z_i) in two steps, see, e.g., Wooldridge (2010). First, one estimates V_i^* in equation (6) by the the residuals $\hat{V}_i^*$ in the regression of X_i^* on (W_i, Z_i). Note that we can write $U_i^* = e_i^* + \theta_{V^*} V_i^*$, where $\theta_{V^*} \equiv \sigma_{V^*U^*}/\sigma_{V^*}^2$, and e_i^* is independent of Z_i, W_i, and V_i^* (and hence of X_i^*). Then, one estimates the standard Tobit model

$$
Y_i = m\left(\theta_{01} X_i^* + \theta_{02}' W_i + \theta_{V^*} V_i^* + e_i^*\right),
$$

where V_i^* are replaced by their estimates $\hat{V}_i^*$. (The two steps can also be combined and all of the parameters can be estimated simultaneously by the Maximum Likelihood Estimator.) The reason this approach works is that equation (6) creates a control variable V_i^*, and the inclusion of V_i^* in the above equation makes X_i^* exogenous.

To estimate the partial effects, one would plug the estimates $\widehat{\theta}$ and $\widehat{\sigma}_{U^*}^2$ into equations (2)-(3) in place of θ_0 and $\sigma_{U^*}^2$.

So far we were assuming that the data had no measurement errors. We now allow X_i^* to be mismeasured, i.e., that instead of X_i^* we observe its noisy measurement X_i:

$$
X_i = X_i^* + \varepsilon_i, \qquad \varepsilon_i \sim N\left(0, \sigma_\varepsilon^2\right). \tag{8}
$$

We assume that $\varepsilon_i \perp (U_i^*, \eta_i, W_i, Z_i)$, i.e., the measurement error is classical.

Note that the objects of researchers' interest do not change: the goal is to estimate the partial effects (2)-(3). The structural endogeneity and measurement errors are difficulties that an estimation procedure needs to overcome. In particular, note that we are interested in estimation of the effect of X_i^* on Y_i, and not in the effect of the mismeasured X_i.[3]

[3] This is similar to the linear regression settings, where one would be interested in the effect of X_i^* on Y_i. The slope coefficient in the OLS regression of Y_i on X_i is not the object of interest because it is subject to the attenuation bias due to the EiV (and also possibly due to the endogeneity of X_i^*).

3 Analysis of the Model

First, we use the model in equations (5)-(8) to obtain the model in terms of the observable X_i. Since $X_i^* = X_i - \varepsilon_i$, we can rewrite (5) as

$$\begin{aligned} Y_i &= m\left(\theta_{01} X_i^* + \theta_{02}' W_i + U_i^*\right) = m\left(\theta_{01} X_i + \theta_{02}' W_i - \theta_{01}\varepsilon_i + U_i^*\right) \\ &= m\left(\theta_{01} X_i + \theta_{02}' W_i + U_i\right), \end{aligned}$$

where $U_i \equiv U_i^* - \theta_{01}\varepsilon_i$. Let $V_i \equiv V_i^* + \varepsilon_i$. The model in equations (5)-(8) can be written as

$$\begin{aligned} Y_i &= m\left(\theta_{01} X_i + \theta_{02}' W_i + U_i\right), &(9)\\ X_i &= \pi_{01}' Z_i + \pi_{02}' W_i + V_i, &(10)\\ \begin{pmatrix} U_i \\ V_i \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_U^2 & \sigma_{UV} \\ \sigma_{UV} & \sigma_V^2 \end{pmatrix}\right). &(11) \end{aligned}$$

The definitions of U_i and V_i imply that

$$\sigma_U^2 = \sigma_{U^*}^2 + \theta_{01}^2 \sigma_\varepsilon^2, \quad \sigma_V^2 = \sigma_{V^*}^2 + \sigma_\varepsilon^2, \quad \sigma_{UV} = \sigma_{U^*V^*} - \theta_{01}\sigma_\varepsilon^2. \tag{12}$$

Note that variables X_i, U_i, V_i are the analogs of the true variables X_i^*, U_i^*, V_i^* that arise due to the measurement errors ε_i. In the absence of measurement errors, i.e., when $\varepsilon_i = 0$, we have $X_i = X_i^*$, $U_i = U_i^*$, $V_i = V_i^*$.

The model in equations (9)-(11) can be estimated using IV-Tobit estimator described earlier. Specifically, this will consistently estimate parameters θ_0 and the covariance matrix of the unobservables in equation (11), i.e., σ_U^2, σ_{UV}, and σ_V^2. Note that because the model is nonlinear, the marginal effects (2)-(3) depend not only on θ_0 but also on $\sigma_{U^*}^2$. Thus, even though the available data (Y_i, X_i, W_i, Z_i) allows immediately estimating θ_0 using IV-Tobit, we cannot obtain the marginal effects because we do not know $\sigma_{U^*}^2$. Naively using an estimate of σ_U^2 in place of $\sigma_{U^*}^2$ would lead to a biased estimate of the partial effects, since $\sigma_U^2 \geq \sigma_{U^*}^2$, as implied by equation (12).

The problem with identifying $\sigma_{U^*}^2$ is that the data only allows identification of the 3 parameters σ_U^2, σ_{UV}, and σ_V^2. However, the distribution of the true $(U_i^*, V_i^*, \varepsilon_i)$ is governed by 4 parameters: $\sigma_{U^*}^2$, $\sigma_{U^*V^*}$, $\sigma_{V^*}^2$, and σ_ε^2. Thus, one cannot uniquely determine these 4 parameters from the 3 equations (12). In other words, models with different values of σ_ε^2 are observationally equivalent: they correspond to identical distributions of the observables (Y_i, X_i, W_i, Z_i) even though they imply different values of true $\sigma_{U^*}^2$. Thus, one cannot uniquely determine (i.e., point-identify) $\sigma_{U^*}^2$ from the data (Y_i, X_i, W_i, Z_i). Correspondingly, one cannot point-identify the partial effects, which depend on $\sigma^2_{U^*}$.

Equations (12) provide restrictions on $\sigma^2_{U^*}$, which we will use to provide bounds on the possible values of true $\sigma^2_{U^*}$, and hence on the values of the partial effects.

Bounds on $\sigma^2_{U^*}$ From equations (12) the upper bound on $\sigma^2_{U^*}$ is $\sigma^2_{U^*} \leq \sigma^2_U$. We now obtain the lower bound on $\sigma^2_{U^*}$. In particular, we look to find the smallest $\sigma^2_{U^*}$ that satisfies equations (12), Cauchy-Schwarz inequality $\sigma^2_{U^*V^*} \leq \sigma^2_{U^*}\sigma^2_{V^*}$, and the non-negativity constraints $\sigma^2_{U^*} \geq 0$, $\sigma^2_{V^*} \geq 0$, and $\sigma^2_\varepsilon \geq 0$. Let $\rho_{UV} = \text{corr}\,(U_i, V_i)$.

Proposition 1 *Suppose $|\rho_{UV}| < 1$ in model (9)-(11). Then*

$$\sigma^2_{U^*} \in \left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right],$$

where

$$\underline{\sigma}^2_{U^*} \equiv \max\left\{\frac{\left(\theta_{01}\sigma_{UV} + \sigma^2_U\right)^2}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U},\ \sigma^2_U - \theta^2_{01}\sigma^2_V\right\}. \tag{13}$$

Proposition 1 provides the bounds in terms of the quantities that can be estimated using the data (Y_i, X_i, W_i, Z_i). Condition $|\rho_{UV}| < 1$ guarantees that the denominator in the fraction above is positive. The proof of Proposition 1 also provides bounds on $\sigma_{U^*V^*}$ and σ^2_ε.

Correct Partial Effects We now use the bounds on $\sigma^2_{U^*}$ from Proposition **1** to obtain the bounds on the partial effects, in terms of the parameters that can be recovered from data.

For a given $\sigma^2_{U^*}$, the partial effects for the j^{th} covariate are, as in equations (2)-(3),

$$PE^{\text{Tob}}_{h_j}\left(h, \sigma^2_{U^*}\right) = \Phi\left(\frac{\theta_0' h}{\sigma_{U^*}}\right)\theta_{0j} \quad \text{and} \quad PE^{\text{Pr}}_{h_j}\left(h, \sigma^2_{U^*}\right) = \phi\left(\frac{\theta_0' h}{\sigma_{U^*}}\right)\frac{\theta_{0j}}{\sigma_{U^*}}. \tag{14}$$

Notice that $PE^{\text{Tob}}_{h_j}(h, \sigma^2_{U^*})$ in (14) is a monotone function of $\sigma^2_{U^*}$, so the bounds on the PE follow immediately from the bounds on $\sigma^2_{U^*}$ in Proposition **1**. The partial effect of the j^{th} covariate is bounded between $PE^{\text{Tob}}_{h_j}(h, \sigma^2_U)$ and $PE^{\text{Tob}}_{h_j}(h, \underline{\sigma}^2_{U^*})$. Specifically,

$$PE^{\text{Tob}}_{h_j}(h) \in \begin{cases} \left[PE^{\text{Tob}}_{h_j}\left(h, \sigma^2_U\right), PE^{\text{Tob}}_{h_j}\left(h, \underline{\sigma}^2_{U^*}\right)\right] \text{ if } \theta_0' h \times \theta_{0j} \geq 0; \\ \left[PE^{\text{Tob}}_{h_j}\left(h, \underline{\sigma}^2_{U^*}\right), PE^{\text{Tob}}_{h_j}\left(h, \sigma^2_U\right)\right] \text{ if } \theta_0' h \times \theta_{0j} < 0. \end{cases} \tag{15}$$

For the j^{th} covariate, the partial effect $PE^{\text{Pr}}_{h_j}(h)$ on the probability is bounded between

$$\min_{\sigma^2_{U^*} \in \left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right]} PE^{\text{Pr}}_{h_j}\left(h, \underline{\sigma}^2_{U^*}\right) \quad \text{and} \quad \max_{\sigma^2_{U^*} \in \left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right]} PE^{\text{Pr}}_{h_j}\left(h, \sigma^2_{U^*}\right). \tag{16}$$

Function $PE^{\text{Pr}}_{h_j}(h, \sigma^2_{U^*})$ in equation (14) is not monotone in $\sigma^2_{U^*}$, but it is still easy easy to obtain the bounds. The minimum and maximum over $\sigma^2_{U^*} \in [\underline{\sigma}^2_{U^*}, \sigma^2_U]$ can be attained only at $\sigma^2_{U^*} = \underline{\sigma}^2_{U^*}$, at $\sigma^2_{U^*} = \sigma^2_U$, and, when $(\theta_0' h)^2 \in [\underline{\sigma}^2_{U^*}, \sigma^2_U]$, at $\sigma^2_{U^*} = (\theta_0' h)^2$. Thus, one only needs to evaluate $PE^{\text{Pr}}_{h_j}(h, \sigma^2_{U^*})$ at these 2 or 3 points to calculate the minimum and maximum in equation (16).

Since $\sigma_U \geq \sigma_{U^*}$, naively using σ_U instead of σ_{U^*} when calculating $PE^{\text{Tob}}_{h_j}(h)$, would lead to attenuation bias when $\theta_0' h > 0$, but would bias $PE^{\text{Tob}}_{h_j}(h)$ away from zero when $\theta_0' h < 0$, i.e., the EiV would make naive $PE^{\text{Tob}}_{h_j}(h, \sigma^2_U)$ over-estimate the partial effects $PE^{\text{Tob}}_{h_j}(h)$ in the latter case. Likewise, for the probability, naively using $PE^{\text{Pr}}_{h_j}(h, \sigma^2_U)$ can both under- and over-estimate the true value of the partial effect $PE^{\text{Pr}}_{h_j}(h)$.

Estimation From the estimation procedure in Section 2, one obtains the estimates of θ_0, σ^2_U, σ^2_V, and σ^2_{UV}. Then, from equation (13) one obtains the estimate of $\underline{\sigma}^2_{U^*}$.

Given a value of $\sigma^2_{U^*}$, the estimated partial effects are

$$\widehat{PE}^{\text{Tob}}_{h_j}\left(h, \sigma^2_{U^*}\right) = \Phi\left(\frac{\widehat{\theta}' h}{\sigma_{U^*}}\right)\widehat{\theta}_j \quad \text{and} \quad \widehat{PE}^{\text{Pr}}_{h_j}\left(h, \sigma^2_{U^*}\right) = \phi\left(\frac{\widehat{\theta}' h}{\sigma_{U^*}}\right)\frac{\widehat{\theta}_j}{\sigma_{U^*}}. \tag{17}$$

Then, the estimated bounds on $PE^{\text{Tob}}_{h_j}(h)$ are $\widehat{PE}^{\text{Tob}}_{h_j}\left(h, \widehat{\sigma}^2_U\right)$ and $\widehat{PE}^{\text{Tob}}_{h_j}\left(h, \widehat{\underline{\sigma}}^2_{U^*}\right)$, with equation (15) determining which of the two bounds is the upper and which is the lower bound, depending on the sign of $\widehat{\theta}' h \times \widehat{\theta}_j$.

The estimated bounds on $PE^{\text{Pr}}_{h_j}(h)$ are

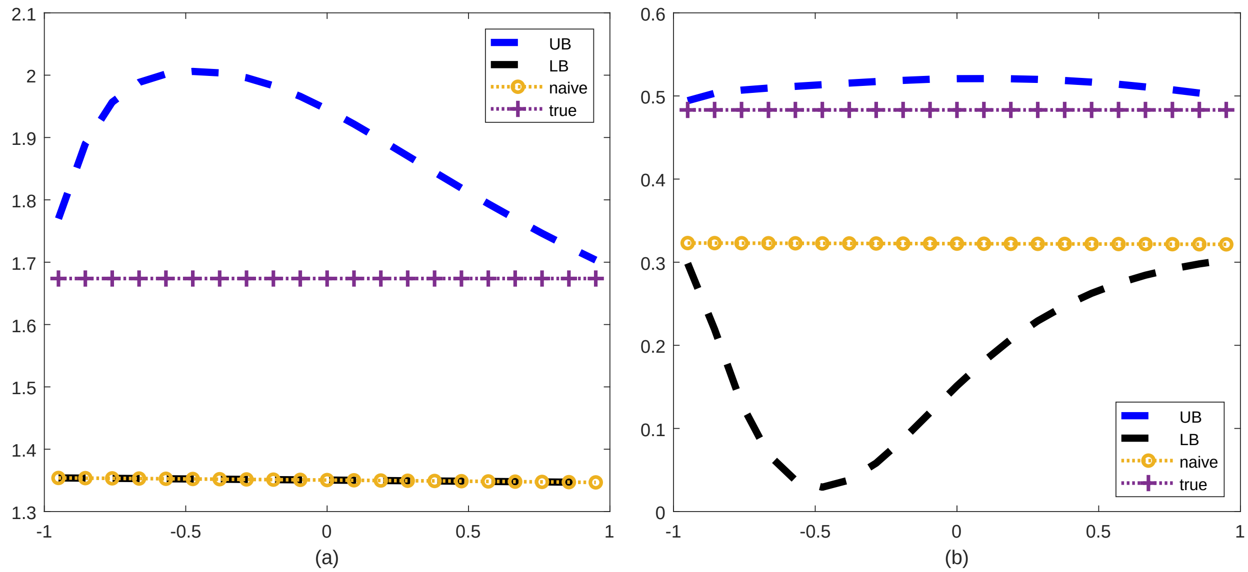
$$\min_{v \in \left[\widehat{\underline{\sigma}}^2_{U^*}, \widehat{\sigma}^2_U\right]} \widehat{PE}^{\text{Pr}}_{h_j}(h, v) \quad \text{and} \quad \max_{v \in \left[\widehat{\underline{\sigma}}^2_{U^*}, \widehat{\sigma}^2_U\right]} \widehat{PE}^{\text{Pr}}_{h_j}(h, v), \tag{18}$$

where the minimum and maximum are easily computed as described earlier.

For example, one often considers the partial effects at the mean values of covariates taking $h = \left(\overline{X}, \overline{W}'\right)'$, where $\overline{X}$ and $\overline{W}$ are the sample averages. Note that $\mathrm{E}[X_i] = \mathrm{E}[X_i^*]$.

Numerical Illustration We simulate a Tobit model with endogenous and mismeasured X_i^*, as in equations (5)-(8), with $W_i = 1$, $Z_i \sim N(0,1)$, $(\theta_{01}, \theta_{02}, \sigma_{V^*}, \sigma_{U^*}, \pi_{01}, \pi_{02}) = (2, 1, 1, 1, 1, 0)$, and $n = 200$.

The figure below plots the results for the Partial Effects (PEs) of X_i^* at the mean values of covariates, $h = (\overline{X}, 1)'$. We consider a range of designs corresponding to the true values of $\rho_{U^*V^*} \in [-0.95, 0.95]$ on the horizontal axis. The figure shows the true PE ("true"), as well as the averages over Monte Carlo replications of the estimated lower and upper bounds on the PE ("LB" and "UB"), and of the naive IV-Tobit estimator for the PE ("naive"). The naive estimators of the partial effects are $\widehat{PE}_{h_1}^{\text{Tob}}\left(h, \widehat{\sigma}_U^2\right)$ and $\widehat{PE}_{h_1}^{\text{Pr}}\left(h, \widehat{\sigma}_U^2\right)$, i.e., they replace $\sigma_{U^*}^2$ with $\widehat{\sigma}_U^2$ in equation (17). By construction, the naive estimator of $PE_{h_1}^{\text{Tob}}(h)$ coincides with one of the bounds.



Panel (a) shows the bounds on $PE_{h_1}^{\text{Tob}}(h)$. The naive IV-Tobit estimator of the PE coincides with the lower bound at ≈ 1.35, while the true PE is ≈ 1.68. As expected, the true PE is between the lower and upper bounds for all values of $\rho_{U^*V^*}$.

Panel (b) considers $PE_{h_1}^{\text{Pr}}(h)$. The naive estimator of this PE is ≈ 0.32, significantly underestimating the true PE (≈ 0.48), and the bounds include the true PE. In both panels, for larger positive $\rho_{U^*V^*}$, the bounds are relatively narrow and are approaching the naive estimator (from below) and the true PE (from above).

4 Average Partial Effects

In addition to the partial effects at a given h, one may also be interested in calculating the Average Partial Effects,

$$
\begin{aligned}
APE_{h_j}^{\text{Tob}}\left(\sigma_{U^*}^2\right) &= \mathrm{E}\left[PE_{h_j}^{\text{Tob}}\left(H_i^*, \sigma_{U^*}^2\right)\right] \text{ and} \\
APE_{h_j}^{\text{Pr}}\left(\sigma_{U^*}^2\right) &= \mathrm{E}\left[PE_{h_j}^{\text{Pr}}\left(H_i^*, \sigma_{U^*}^2\right)\right].
\end{aligned}
$$

APEs are expectations over the distribution of $H_i^* = (X_i^*, W_i')'$, but X_i^* is not observable due to the EiV. In the Appendix, we show that these APEs can be calculated as

$$APE_{h_j}^{\mathrm{Tob}}\left(\sigma_{U^*}^2\right) = \mathrm{E}\left[\Phi\left(\frac{\theta_{01}\pi_{01}' Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)' W_i}{\sqrt{2\sigma_{U^*}^2 - \sigma_U^2 + \theta_{01}^2\sigma_V^2}}\right)\right]\theta_{0j}, \tag{19}$$

$$APE_{h_j}^{\mathrm{Pr}}\left(\sigma_{U^*}^2\right) = \mathrm{E}\left[\phi\left(\frac{\theta_{01}\pi_{01}' Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)' W_i}{\sqrt{2\sigma_{U^*}^2 - \sigma_U^2 + \theta_{01}^2\sigma_V^2}}\right)\right]\frac{\theta_{0j}}{\sqrt{2\sigma_{U^*}^2 - \sigma_U^2 + \theta_{01}^2\sigma_V^2}}. \tag{20}$$

Hence, for any given value of $\sigma_{U^*}^2$, these APEs can be estimated by

$$\widehat{APE}_{h_j}^{\mathrm{Tob}}\left(\sigma_{U^*}^2\right) = \frac{1}{n}\sum_{i=1}^{n}\Phi\left(\frac{\hat\theta_1\hat\pi_1' Z_i + \left(\hat\theta_1\hat\pi_2 + \hat\theta_2\right)' W_i}{\sqrt{2\sigma_{U^*}^2 - \hat\sigma_U^2 + \hat\theta_1^2\hat\sigma_V^2}}\right)\hat\theta_j,$$

$$\widehat{APE}_{h_j}^{\mathrm{Pr}}\left(\sigma_{U^*}^2\right) = \frac{1}{n}\sum_{i=1}^{n}\phi\left(\frac{\hat\theta_1\hat\pi_1' Z_i + \left(\hat\theta_1\hat\pi_2 + \hat\theta_2\right)' W_i}{\sqrt{2\sigma_{U^*}^2 - \hat\sigma_U^2 + \hat\theta_1^2\hat\sigma_V^2}}\right)\frac{\hat\theta_j}{\sqrt{2\sigma_{U^*}^2 - \hat\sigma_U^2 + \hat\theta_1^2\hat\sigma_V^2}}.$$

Finally, the estimated bounds on APEs are obtained by finding the maximum and minimum over $\sigma_{U^*}^2 \in \left[\underline{\hat\sigma}_{U^*}^2, \hat\sigma_U^2\right]$. These can be easily computed numerically, since $\widehat{APE}_{h_j}(v)$ are smooth functions of a scalar argument v.

5 Probit

IV-Probit model is the same as IV-Tobit model except $m(s) = 1\{s > 0\}$ in equation (5). Since Probit is a binary outcome model, in equation (11) one imposes the standard normalization $\sigma_U = 1$. For Probit, we are interested in the partial effects of covariates on the probability of $Y_i = 1$, which coincides with $PE_{h_j}^{\mathrm{Pr}}(h)$ in equation (3). The estimation procedure of IV-Probit is identical to that of IV-Tobit described in Section 2, except the second step uses the standard Probit estimator in place of Tobit estimator. Then the bounds on $PE_{h_j}^{\mathrm{Pr}}(h)$ are estimated as in equation (18), and the bounds on the $APE_{h_j}^{\mathrm{Pr}}$ are estimated as in Section 4.

References

Blundell, R. and J. L. Powell (2003): "Endogeneity in Nonparametric and Semiparametric Regression Models," in *Advances in Economics and Econometrics*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, 312–357.

Chesher, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71, 1405–1441.

Imbens, G. and W. Newey (2009): "Identification and Estimation of Triangular Simultaneous Equations Models without Additivity," *Econometrica*, 77, 1481–1512.

Rivers, D. and Q. H. Vuong (1988): "Limited Information Estimators And Exogeneity Tests For Simultaneous Probit Models," *Journal of Econometrics*, 39, 347–366.

Schennach, S. (2022): "Measurement Systems," *Journal of Economic Literature*, 60, 1223–1263.

Smith, R. J. and R. W. Blundell (1986): "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply," *Econometrica*, 54, 679.

Wooldridge, J. M. (2005): "Unobserved Heterogeneity and Estimation of Average Partial Effects," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. by D. Andrews and J. Stock, Cambridge University Press, 27–55.

——— (2010): *Econometric Analysis of Cross Section and Panel Data, Second Edition*, The MIT Press.

——— (2015): "Control Function Methods in Applied Econometrics," *Journal of Human Resources*, 50, 420–445.

Appendix

Proof of Proposition 1. Note that $\sigma^2_{U^*V^*} \leq \sigma^2_{U^*}\sigma^2_{V^*}$ combined with equation (12) implies

$$\begin{aligned} 0 &\leq \sigma^2_{U^*}\sigma^2_{V^*} - \sigma^2_{U^*V^*} = \left(\sigma^2_U - \theta^2_{01}\sigma^2_\varepsilon\right)\left(\sigma^2_V - \sigma^2_\varepsilon\right) - \left(\sigma_{UV} + \theta_{01}\sigma^2_\varepsilon\right)^2 \\ &= \sigma^2_U\sigma^2_V - \sigma^2_{UV} - \sigma^2_\varepsilon\left(\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U\right), \text{ and hence} \\ 0 &\leq \sigma^2_\varepsilon \leq \frac{\sigma^2_U\sigma^2_V - \sigma^2_{UV}}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U}. \end{aligned}$$

Since $|\rho_{UV}| < 1$, the denominator in this fraction is positive. Since $\sigma^2_\varepsilon \leq \sigma^2_V$, let

$$\overline{\sigma}^2_\varepsilon \equiv \min\left\{\frac{\sigma^2_U\sigma^2_V - \sigma^2_{UV}}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U}, \sigma^2_V\right\}.$$

Then

$$\underline{\sigma}^2_{U^*} \leq \sigma^2_{U^*} \leq \sigma^2_U, \qquad \underline{\sigma}^2_{U^*} \equiv \max\left\{\sigma^2_U - \theta^2_{01}\overline{\sigma}^2_{\varepsilon}, 0\right\}. \tag{21}$$

Note that

$$\sigma^2_U - \theta^2_{01}\frac{\sigma^2_U\sigma^2_V - \sigma^2_{UV}}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U} = \frac{\sigma^4_U + 2\sigma^2_U\sigma_{UV}\theta_{01} + \theta^2_{01}\sigma^2_{UV}}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U} = \frac{\left(\theta_{01}\sigma_{UV} + \sigma^2_U\right)^2}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U}.$$

Thus, $\underline{\sigma}^2_{U^*}$ in equation (21) can be equivalently written as

$$\underline{\sigma}^2_{U^*} \equiv \max\left\{\frac{\left(\theta_{01}\sigma_{UV} + \sigma^2_U\right)^2}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U}, \sigma^2_U - \theta^2_{01}\sigma^2_V\right\},$$

where the fraction is always non-negative, ensuring that $\underline{\sigma}^2_{U^*} \geq 0$. Finally, equation (12) implies that $\sigma_{U^*V^*}$ is bounded between σ_{UV} and $\sigma_{UV} + \theta_{01}\overline{\sigma}^2_{\varepsilon}$. ■

Derivation of equation (19):

$$\begin{aligned}
APE^{\text{Tob}}_{h_j}\left(\sigma^2_{U^*}\right) &= \mathrm{E}\left[\Phi\left(\frac{\theta_0'H_i^*}{\sigma_{U^*}}\right)\theta_{0j}\right] = \mathrm{E}\left[\Phi\left(\frac{\theta_{01}X_i^* + \theta_{02}'W_i}{\sigma_{U^*}}\right)\right]\theta_{0j} \\
&= \mathrm{E}\left[\Phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)'W_i}{\sigma_{U^*}} + \frac{\theta_{01}V_i^*}{\sigma_{U^*}}\right)\right]\theta_{0j} = \mathrm{E}\left[\Phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)'W_i}{\sigma_{U^*}\sqrt{1 + \theta^2_{01}\sigma^2_{V^*}/\sigma^2_{U^*}}}\right)\right]\theta_{0j} \\
&= \mathrm{E}\left[\Phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)'W_i}{\sqrt{2\sigma^2_{U^*} - \sigma^2_U + \theta^2_{01}\sigma^2_V}}\right)\right]\theta_{0j},
\end{aligned}$$

where the penultimate equality comes from taking expectation with respect to V_i^*, and the last equality follows from equation (12). ■

Derivation of equation (20) is similar:

$$\begin{aligned}
APE^{\text{Pr}}_{h_j}\left(\sigma^2_{U^*}\right) &= \mathrm{E}\left[\phi\left(\frac{\theta_0'H_i^*}{\sigma_{U^*}}\right)\frac{\theta_{0j}}{\sigma_{U^*}}\right] = \mathrm{E}\left[\phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)'W_i}{\sigma_{U^*}} + \frac{\theta_{01}V_i^*}{\sigma_{U^*}}\right)\right]\frac{\theta_{0j}}{\sigma_{U^*}} \\
&= \frac{1}{\sqrt{1 + \theta^2_{01}\sigma^2_{V^*}/\sigma^2_{U^*}}}\mathrm{E}\left[\phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)'W_i}{\sigma_{U^*}\sqrt{1 + \theta^2_{01}\sigma^2_{V^*}/\sigma^2_{U^*}}}\right)\right]\frac{\theta_{0j}}{\sigma_{U^*}} \\
&= \mathrm{E}\left[\phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02} + \theta_{02}\right)'W_i}{\sqrt{2\sigma^2_{U^*} - \sigma^2_U + \theta^2_{01}\sigma^2_V}}\right)\right]\frac{\theta_{0j}}{\sqrt{2\sigma^2_{U^*} - \sigma^2_U + \theta^2_{01}\sigma^2_V}}.
\end{aligned}$$

■