Marginal Effects for Probit and Tobit with Endogeneity

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

This version: June 22, 2023

Abstract

When evaluating partial effects, it is important to distinguish between structural endogeneity and measurement errors. In contrast to linear models, these two sources of endogeneity affect partial effects differently in nonlinear models. We study this issue focusing on the Instrumental Variable (IV) Probit and Tobit models. We show that even when a valid IV is available, failing to differentiate between the two types of endogeneity can lead to either under- or over-estimation of the partial effects. We develop simple estimators of the bounds on the partial effects and provide easy to implement confidence intervals that correctly account for both types of endogeneity. We illustrate the methods in a Monte Carlo simulation and an empirical application.

Keywords: (Average) Partial Effects, Instrumental Variable, Control Variable, Errors-in-Variables, Counterfactuals

*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 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). Second, mismeasurement of the covariate (Errors-in-Variables, 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.

The goal of this paper is to characterize the partial effects in the classic 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, allowing for both types of endogeneity. We also provide corresponding valid confidence intervals that are easy to calculate.

In nonlinear models, the need to differentiate between the two kinds of endogeneity arises because 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 the IV-Probit and IV-Tobit methods 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. The width of the identified set depends on how hard it is to disentangle structural endogeneity and EiV for the data at hand. Importantly, we find that naively ignoring the distinction between the two types of endogeneity can result in both under- and over-estimation of the magnitude of the partial effects by these IV estimators.[2]

IV-Probit and IV-Tobit can be interpreted as 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. Recently, Chesher, Kim, and Rosen (2023) studied IV methods for the Tobit models relaxing the first stage specification

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

[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.

assumptions, but did not consider EiV.

The assumption of gaussianity of the unobservables simplifies the analysis of the IV-Probit and IV-Tobit models and allows us to obtain simple bounds and confidence intervals on the true partial effects that are easy to implement. It would be important to relax the gaussianity assumption in the future. At the same time, the simplicity of the proposed approach makes it a convenient starting point in empirical analysis. In particular, it allows researchers to gauge the importance of properly accounting for both types of endogeneity, which is essential given the ubiquity of both in economics.

The rest of the paper is organized as follows. The analysis of partial effects in the Probit and Tobit models is virtually identical, thus we first consider Tobit in Sections 2-3. In Section 4, we explain how the methods apply to the Probit model, and extend the analysis to cover the average partial effects and other counterfactuals. Section 5 provides some Monte Carlo simulation results. Section 6 presents an illustrative empirical application.

2 The Model

The 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 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:

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

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.[3]

In the 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 concreteness we consider 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

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

[3] Most of the analysis in Sections 2 and 3 equally applies to the Probit model. For simplicity of exposition, we focus on the Tobit model for the moment, and then discuss Probit in Section 4.1.

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

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

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

When X_i^* is correlated with U_i^* and we observe data (Y_i, X_i^*, W_i, Z_i), the IV-Tobit model can be estimated using instrumental variables Z_i, as proposed by Smith and Blundell (1986), Newey (1987), and Rivers and Vuong (1988). Assume that

$$Y_i = m\left(\theta_{01} X_i^* + \theta_{02}' W_i + U_i^*\right), \quad m\left(s\right) = \max(s, 0), \tag{4}$$

$$X_i^* = \pi_{01}' Z_i + \pi_{02}' W_i + V_i^*, \tag{5}$$

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

$$\begin{pmatrix} U_i^* \\ V_i^* \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{U^*}^2 & \sigma_{U^*V^*} \\ \sigma_{U^*V^*} & \sigma_{V^*}^2 \end{pmatrix}\right). \tag{6}$$

The IV-Tobit model in (4)-(6) can be 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 (5) 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 $\hat{V}_i^*$. (Alternatively, the two steps can be combined and all of the parameters can be estimated simultaneously by the Maximum Likelihood Estimator.) The reason this approach works is that equation (5) 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 has 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:

$$X_i = X_i^* + \varepsilon_i, \qquad \varepsilon_i \sim N\left(0, \sigma_\varepsilon^2\right). \tag{7}$$

We assume that $\varepsilon_i \perp (U_i^*, V_i^*, W_i, Z_i)$, 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.[4]

3 Analysis of the Model

First, we use the model in equations (4)-(7) to obtain the model in terms of the observable X_i. Since $X_i^* = X_i - \varepsilon_i$, we can rewrite (4) 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 (4)-(7) can be written as

$$Y_i = m\left(\theta_{01} X_i + \theta_{02}' W_i + U_i\right), \tag{8}$$

$$X_i = \pi_{01}' Z_i + \pi_{02}' W_i + V_i, \tag{9}$$

$$\begin{pmatrix} U_i \\ V_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_U^2 & \sigma_{UV} \\ \sigma_{UV} & \sigma_V^2 \end{pmatrix}\right). \tag{10}$$

The definitions of U_i and V_i imply that

$$\sigma_U^2 = \sigma_{U^*}^2 + \theta_{01}^2\sigma_\varepsilon^2, \quad \sigma_V^2 = \sigma_{V^*}^2 + \sigma_\varepsilon^2, \quad \sigma_{UV} = \sigma_{U^*V^*} - \theta_{01}\sigma_\varepsilon^2. \tag{11}$$

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 (8)-(10) can be estimated by MLE or using the control variable two-step approach described earlier. Specifically, both approaches will consistently estimate parameters θ_0 and the covariance matrix of the unobservables in equation (10), 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, 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 (11).

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 (11). 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

[4] 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^*).

different values of true $\sigma^2_{U^*}$. Thus, one cannot uniquely determine (i.e., point-identify) $\sigma^2_{U^*}$ 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 (11) 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 (11) 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 (11), 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 (8)-(10). Then*

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

where

$$\underline{\sigma}^2_{U^*} \equiv \max\left\{\frac{\left(\theta_{01}\sigma_{UV} + \sigma^2_U\right)^2}{\sigma^2_V\theta^2_{01} + 2\sigma_{UV}\theta_{01} + \sigma^2_U},\ \sigma^2_U - \theta^2_{01}\sigma^2_V\right\}. \tag{12}$$

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),

$$PE^{\text{Tob}}_j\left(h, \sigma^2_{U^*}\right) = \Phi\left(\frac{\theta_0' h}{\sigma_{U^*}}\right)\theta_{0j} \quad \text{and} \quad PE^{\text{Pr}}_j\left(h, \sigma^2_{U^*}\right) = \phi\left(\frac{\theta_0' h}{\sigma_{U^*}}\right)\frac{\theta_{0j}}{\sigma_{U^*}}. \tag{13}$$

The lower and upper bounds for partial effects for the j^{th} covariate, $PE_j(h)$, are computed as

$$\min_{\sigma^2_{U^*} \in \left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right]} PE_j\left(h, \underline{\sigma}^2_{U^*}\right) \quad \text{and} \quad \max_{\sigma^2_{U^*} \in \left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right]} PE_j\left(h, \sigma^2_{U^*}\right). \tag{14}$$

Function $PE^{\text{Tob}}_j\left(h, \sigma^2_{U^*}\right)$ in (13) is a monotone function of $\sigma^2_{U^*}$, so the minimum and maximum in equation (14) are achieved on the boundaries of interval $\left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right]$.

Function $PE^{\text{Pr}}_j\left(h, \sigma^2_{U^*}\right)$ in equation (13) is not monotone in $\sigma^2_{U^*}$, but the bounds in equation (14) for $PE^{\text{Pr}}_j(h)$ can also be simplified. The minimum and maximum over $\sigma^2_{U^*} \in \left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right]$ can be attained only at $\sigma^2_{U^*} = \underline{\sigma}^2_{U^*}$, at $\sigma^2_{U^*} = \sigma^2_U$, and, when $\left(\theta_0' h\right)^2 \in \left[\underline{\sigma}^2_{U^*}, \sigma^2_U\right]$, at $\sigma^2_{U^*} = \left(\theta_0' h\right)^2$. Thus, one only needs to evaluate $PE^{\text{Pr}}_j\left(h, \sigma^2_{U^*}\right)$ at these 2 or 3 points to calculate the minimum and maximum in equation (14).

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

Estimation Using the standard two-step or MLE approaches described in Section 2, one obtains the estimates of θ_0, σ_U^2, σ_V^2, and σ_{UV} (and of their variance-covariance matrix for inference). Then, from equation (12) one obtains the estimate of $\underline{\sigma}_{U^*}^2$.

For a given value of $\sigma_{U^*}^2$, the estimated partial effects would be

$$\widehat{PE}_j^{\text{Tob}}\left(h, \sigma_{U^*}^2\right) = \Phi\left(\frac{\widehat{\theta}'h}{\sigma_{U^*}}\right)\widehat{\theta}_j \quad \text{and} \quad \widehat{PE}_j^{\text{Pr}}\left(h, \sigma_{U^*}^2\right) = \phi\left(\frac{\widehat{\theta}'h}{\sigma_{U^*}}\right)\frac{\widehat{\theta}_j}{\sigma_{U^*}}. \tag{15}$$

Then, the estimated bounds on $PE_j(h)$ are

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

where the minimum and maximum are easily computed using univariate numerical optimization. For the partial effects in equation (15), these extrema can also be computed as described under equation (14).

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

Inference To provide a simple method for inference about the partial effects, we adopt a Bonferroni approach (e.g., McCloskey, 2017). This approach allows us to avoid computational challenges that often arise in the context of subvector inference in partially identified models. The construction of a $1-\alpha$ confidence interval for a partial effect $PE_j(h)$ proceeds in two steps:

1. Pick $\alpha_1 \in (0, \alpha)$ and construct $CI_{1-\alpha_1}^{\sigma_{U^*}^2}$, a $1-\alpha_1$ confidence interval for $\sigma_{U^*}^2$, based on the bounds provided in Proposition **1**.

2. Construct a $1-\alpha$ confidence interval for $PE_j(h)$ as the union $CI_{1-\alpha}^{PE_j(h)} = \bigcup_{\sigma_{U^*}^2 \in CI_{1-\alpha_1}^{\sigma_{U^*}^2}} CI_{1-(\alpha-\alpha_1)}^{PE_j(h)}\left(\sigma_{U^*}^2\right)$, where $CI_{1-(\alpha-\alpha_1)}^{PE_j(h)}\left(\sigma_{U^*}^2\right)$ is a standard $1-(\alpha-\alpha_1)$ confidence interval for $PE_j(h)$ based on $\widehat{PE}_j\left(h, \sigma_{U^*}^2\right)$ in equation (15) for a given $\sigma_{U^*}^2$.

We now provide the implementation details for each step.

Step 1. The confidence interval for $\sigma_{U^*}^2$ is constructed based on the bounds given in Proposition **1**. As the upper bound, we take $\hat{\sigma}_U^2 + z_{1-\alpha_1/2} \times s_{\hat{\sigma}_U^2}$, where $s_{\hat{\sigma}_U^2}$ is the standard error of $\hat{\sigma}_U^2$, and $z_{1-\alpha_1/2}$ is the $1-\alpha_1/2$ quantile of the standard normal distribution. The lower bound is based on $\underline{\hat{\sigma}}_{U^*}^2 = \max\{\hat{\xi}_1, \hat{\xi}_2\}$, where $\hat{\xi}_1$ and $\hat{\xi}_2$ are the plug-in estimators of the two terms on the right hand side of equation (12). Note that $\hat{\xi}_1$ and $\hat{\xi}_2$ are (generally) jointly asymptotically normal and their asymptotic variance-covariance matrix can be computed using the delta method. Then, as the lower bound of $CI_{1-\alpha_1}^{\sigma_{U^*}^2}$, we take $\max\{\hat{\xi}_1 - c_{1-\alpha_1/2} \times s_{\hat{\xi}_1}, \hat{\xi}_2 - c_{1-\alpha_1/2} \times s_{\hat{\xi}_2}\}$. Here $s_{\hat{\xi}_1}$ and $s_{\hat{\xi}_2}$ are the standard errors of $\hat{\xi}_1$ and $\hat{\xi}_2$, and $c_{1-\alpha_1/2}$ is the $1-\alpha_1/2$ quantile of $\max\{\eta_1, \eta_2\}$, where (η_1, η_2) are jointly normal with unit variances and correlation $\hat{\rho}_{\hat{\xi}_1,\hat{\xi}_2}$, and $\hat{\rho}_{\hat{\xi}_1,\hat{\xi}_2}$ is an estimator of the correlation between $\hat{\xi}_1$ and $\hat{\xi}_2$ (e.g., see Romano and Wolf, 2005). By a standard argument, the confidence interval for $\sigma_{U^*}^2$ given by

$$CI_{1-\alpha_1}^{\sigma_{U^*}^2} = \left[\max\left\{\hat{\xi}_1 - c_{1-\alpha_1/2} \times s_{\hat{\xi}_1}, \hat{\xi}_2 - c_{1-\alpha_1/2} \times s_{\hat{\xi}_2}\right\}, \hat{\sigma}_U^2 + z_{1-\alpha_1/2} \times s_{\hat{\sigma}_U^2}\right]$$

has asymptotic coverage at least $1-\alpha_1$ for the true $\sigma_{U^*}^2$. In the numerical illustrations we take $\alpha_1 = \alpha/10$.

Step 2. First, the standard $CI_{1-(\alpha-\alpha_1)}^{PE_j(h)}\left(\sigma_{U^*}^2\right)$ is $\left[l_{1-(\alpha-\alpha_1)}^{PE_j(h)}\left(\sigma_{U^*}^2\right), u_{1-(\alpha-\alpha_1)}^{PE_j(h)}\left(\sigma_{U^*}^2\right)\right]$ constructed by adding and subtracting $z_{1-(\alpha-\alpha_1)/2} \times s_{\widehat{PE}_j(h,\sigma_{U^*}^2)}$ from $\widehat{PE}_j(h, \sigma_{U^*}^2)$. The standard error $s_{\widehat{PE}_j(h,\sigma_{U^*}^2)}$ of $\widehat{PE}_j(h, \sigma_{U^*}^2)$ can be computed using the delta method. Then we can construct $CI_{1-\alpha}^{PE_j(h)}$ as

$$CI_{1-\alpha}^{PE_j(h)} = \left[\min_{\sigma_{U^*}^2 \in CI_{1-\alpha_1}^{\sigma_{U^*}^2}} l_{1-(\alpha-\alpha_1)}^{PE_j(h)}\left(\sigma_{U^*}^2\right), \max_{\sigma_{U^*}^2 \in CI_{1-\alpha_1}^{\sigma_{U^*}^2}} u_{1-(\alpha-\alpha_1)}^{PE_j(h)}\left(\sigma_{U^*}^2\right)\right],$$

where the minimum and maximum are easily calculated using univariate numerical optimization over $\sigma_{U^*}^2$. By the standard Bonferroni argument, the confidence interval $CI_{1-\alpha}^{PE_j(h)}$ has asymptotic coverage of at least $1-\alpha$ for the true partial effect $PE_j(h)$.

The constructed confidence interval is asymptotically valid as long as (i) the first step confidence interval $CI_{1-\alpha_1}^{\sigma_{U^*}^2}$ covers the true $\sigma_{U^*}^2$ with probability at least $1-\alpha_1$ asymptotically, and (ii) the delta method applies to $\widehat{PE}_j(h, \sigma_{U^*}^2)$ for the true $\sigma_{U^*}^2$. Both conditions are satisfied provided that the true $\sigma_{U^*}^2$ is bounded away from zero. Note that in this case $CI_{1-\alpha_1}^{\sigma_{U^*}^2}$ is valid even if $\theta_{01}\sigma_{UV} + \sigma_U^2$ is equal to (or local to) zero, which implies that $CI_{1-\alpha}^{PE_j(h)}$ is also valid.

4 Extensions

4.1 Probit

IV-Probit is the same as IV-Tobit except $m(s) = 1\{s > 0\}$ in equation (4). Since Probit is a binary outcome model, in equation (10) 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 are given by $PE_j^{\text{Pr}}(h)$ in equation (3). Similarly to the IV-Tobit model, the IV-Probit model can be estimated by MLE or by the two-step approach identical to the one described in Section 2, except the second step uses the standard Probit estimator in place of the Tobit estimator. Then the bounds on $PE_j^{\text{Pr}}(h)$ are estimated as in equation (16). Confidence intervals for $PE_j^{\text{Pr}}(h)$ can be computed exactly as described above.

4.2 Average Partial Effects and other Other Counterfactuals

In addition to the partial effects at a given h, researchers are often interested in the Average Partial Effects

$$APE_j^{\text{Tob}} \equiv \mathrm{E}\left[PE_j^{\text{Tob}}\left(H_i^*\right)\right] \text{ and } APE_j^{\text{Pr}} \equiv \mathrm{E}\left[PE_j^{\text{Pr}}\left(H_i^*\right)\right], \tag{17}$$

which are the partial effects $PE_j\left(h\right)$ averaged with respect to the distribution of $H_i^* = (X_i^*, W_i')'$. Define

$$APE_j^{\text{Tob}}\left(\sigma_{U^*}^2\right) \equiv \mathrm{E}\left[PE_j^{\text{Tob}}\left(H_i^*, \sigma_{U^*}^2\right)\right] \text{ and } APE_j^{\text{Pr}}\left(\sigma_{U^*}^2\right) \equiv \mathrm{E}\left[PE_j^{\text{Pr}}\left(H_i^*, \sigma_{U^*}^2\right)\right].$$

Note that the distribution of X_i^* is not directly observable due to the EiV. Averaging $PE_j\left(h\right)$ with respect to the distribution of the observed $H_i = (X_i, W_i')'$ would result in biased estimators of the APEs. To account for this, in the Appendix we show that these APEs can be calculated as

$$APE_j^{\text{Tob}}\left(\sigma_{U^*}^2\right) = \mathrm{E}\left[\Phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02}+\theta_{02}\right)'W_i}{\sqrt{2\sigma_{U^*}^2 - \sigma_U^2 + \theta_{01}^2\sigma_V^2}}\right)\right]\theta_{0j}, \tag{18}$$

$$APE_j^{\text{Pr}}\left(\sigma_{U^*}^2\right) = \mathrm{E}\left[\phi\left(\frac{\theta_{01}\pi_{01}'Z_i + \left(\theta_{01}\pi_{02}+\theta_{02}\right)'W_i}{\sqrt{2\sigma_{U^*}^2 - \sigma_U^2 + \theta_{01}^2\sigma_V^2}}\right)\right]\frac{\theta_{0j}}{\sqrt{2\sigma_{U^*}^2 - \sigma_U^2 + \theta_{01}^2\sigma_V^2}}. \tag{19}$$

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

$$\widehat{APE}_j^{\text{Tob}}\left(\sigma_{U^*}^2\right) = \frac{1}{n}\sum_{i=1}^n \Phi\left(\frac{\hat{\theta}_1\hat{\pi}_1'Z_i + \left(\hat{\theta}_1\hat{\pi}_2+\hat{\theta}_2\right)'W_i}{\sqrt{2\sigma_{U^*}^2 - \hat{\sigma}_U^2 + \hat{\theta}_1^2\hat{\sigma}_V^2}}\right)\hat{\theta}_j,$$

$$\widehat{APE}_j^{\text{Pr}}\left(\sigma_{U^*}^2\right) = \frac{1}{n}\sum_{i=1}^n \phi\left(\frac{\hat{\theta}_1\hat{\pi}_1'Z_i + \left(\hat{\theta}_1\hat{\pi}_2+\hat{\theta}_2\right)'W_i}{\sqrt{2\sigma_{U^*}^2 - \hat{\sigma}_U^2 + \hat{\theta}_1^2\hat{\sigma}_V^2}}\right)\frac{\hat{\theta}_j}{\sqrt{2\sigma_{U^*}^2 - \hat{\sigma}_U^2 + \hat{\theta}_1^2\hat{\sigma}_V^2}}.$$

Finally, the estimated bounds on the APEs are obtained by finding the minimum and maximum over $\sigma_{U^*}^2 \in \left[\underline{\hat{\sigma}}_{U^*}^2, \hat{\sigma}_U^2\right]$. These can be easily computed numerically, since $\widehat{APE}_j\left(v\right)$ are smooth functions of a scalar argument v. Our two-step approach to inference also applies to the APEs with a minimal modification. The only difference is that in Step 2 the construction of the standard error $s_{\widehat{APE}_j}\left(\sigma_{U^*}^2\right)$ as usual needs to account for the sampling variability in both the parameter estimators and the data entering the expressions for the APEs directly.

It is also straightforward to apply our analysis to other counterfactuals, including partial effects and APEs of discrete covariates, as well as to the ordered Probit and two-sided Tobit models. Proposition 1 and the bounds on $\sigma_{U^*}^2$ remain the same, and hence the estimation and inference procedures remain unchanged, except for different formulas in equations (13)-(16) corresponding to the counterfactuals of interest.

5 Numerical Illustration

We simulate a Tobit model with endogenous and mismeasured X_i^*, as in equations (4)-(7), with $W_i = 1$, $Z_i \sim N(0,1)$, $(\theta_{01}, \theta_{02}, \sigma_{V^*}, \sigma_{U^*}, \sigma_\varepsilon, \pi_{01}, \pi_{02}) = (2,1,1,1,1,1,0)$, and $n = 200$. Figure 1 plots the results for the Partial Effects (PEs) of X_i^* at the population mean values of the covariates.

We consider a range of designs corresponding to the true values of $\rho_{U^*V^*} \in [-0.95, 0.95]$ on the horizontal axis. For each $\rho_{U^*V^*}$, the figure shows the true PE ("true"), the true (population) bounds for the PE obtained using Proposition 1 ("true bounds"), as well as the medians over the Monte Carlo replications of the estimated lower and upper bounds on the PE ("LB" and "UB") and the corresponding 95% confidence intervals ("CI") based on the two-step IV-Tobit estimator. The true bounds for the PE are calculated using the point identified parameters θ_0, σ_U^2, σ_{UV}, and σ_V^2, see equations (8)-(10). For comparison, we also include the results for the PE calculated using the standard naive IV-Tobit estimator ("naive") and the corresponding confidence intervals ("CI naive"). The "naive" estimators of the partial effects are $\widehat{PE}_1^{\text{Tob}}\left(h, \widehat{\sigma}_U^2\right)$ and $\widehat{PE}_1^{\text{Pr}}\left(h, \widehat{\sigma}_U^2\right)$, i.e., they replace $\sigma_{U^*}^2$ with $\widehat{\sigma}_U^2$ in equation (15).

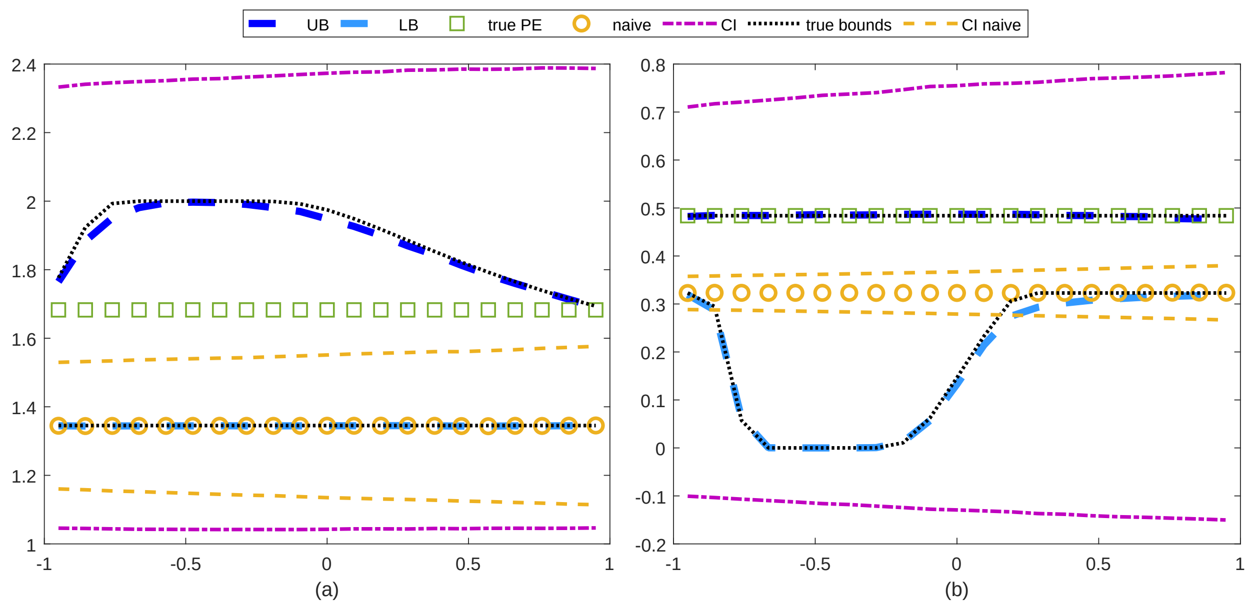

Figure 1: Simulation results for partial effects on: (a) expectations, $PE_{h_1}^{\text{Tob}}(h)$, and (b) probability, $PE_1^{\text{Pr}}(h)$. Values of $\rho_{U^*V^*}$ are on the horizontal axis.

Figure 1 (a) shows the bounds on $PE_{h_1}^{\text{Tob}}(h)$, while Figure 1 (b) considers $PE_1^{\text{Pr}}(h)$. As expected, the true PE is between the lower and upper bounds for all values of $\rho_{U^*V^*}$. By construction, the "naive" IV-Tobit estimator of $PE_1^{\text{Tob}}(h)$ coincides with one of the bounds. In both panels, the "naive" estimates are below the true values for every $\rho_{U^*V^*}$, and the "naive" IV-Tobit confidence intervals do not include the true PE. In this design, the identified set and the confidence intervals for the true partial effects are much wider than those of the "naive" estimator. The relative width of the identified set and the confidence intervals depends on the specific parameter values. In contrast to these simulations, in the example of the next section, the identified set is very narrow and the confidence intervals for the partial effects have width similar to those of the naive estimator.

To gain some intuition about the shape of the bounds in Figure 1, notice that when $\theta_{01} > 0$, measurement error in X_i^* introduces a negative correlation between X_i and U_i. Thus, observing $\rho_{UV} > 0$ is only consistent with $\rho_{U^*V^*} > 0$, but not with $\rho_{U^*V^*} \leq 0$, i.e., implies positive correlation due to the structural endogeneity. On the other hand, observing $\rho_{UV} < 0$ can be explained both by the effect of EiV combined with $\rho_{U^*V^*} \gtrless 0$, and by $\rho_{U^*V^*} < 0$ without any EiV. Thus, when $\rho_{UV} < 0$ it is harder to disentangle structural endogeneity and EiV. As a consequence, the true and the estimated correct bounds on the PEs in Figure 1 are wider for the negative values of $\rho_{U^*V^*}$.

6 Empirical Illustration

We illustrate the proposed methods in the estimation of the Tobit and Probit models for women's labor force participation. We use the classic data set from Mroz (1987) who estimates Tobit and related models to explain married women's hours of work. The data contains 753 women, 428 of which report working non-zero hours. For Tobit, the dependent variable is the number of hours worked (*hours*), and for Probit, the dependent variable is working at some point during the year ($1\{hours > 0\}$). In both models, the covariates are age, education, experience, experience squared, nonwife income in thousands (*nwifeinc*), number of children less than six years of age, number of children between 6 and 18 inclusive, and an intercept. Husband's years of schooling, *huseduc*, are used as an instrument for *nwifeinc*. This specification is used in, e.g., Wooldridge (2010). Nonwife income could be correlated with the unobserved characteristics (structural endogeneity), and income variables are also known to be frequently mismeasured.

	Tobit	IV-Tobit	CI for IV-Tobit	[LB, UB]	CI
nwifeinc	-5.33	-19.0	[-39.6, 1.68]	[-19.1, -19.0]	[-41.6, 2.44]
educ	48.7	70.3	[29.0, 112]	[70.3, 70.8]	[26.9, 117]
exper	79.5	74.9	[51.6, 98.2]	[74.9, 75.4]	[50.3, 102]
exper2	-1.13	-1.14	[-1.82, -0.468]	[-1.15, -1.14]	[-1.89, -0.444]
age	-32.9	-28.2	[-39.3, -17.2]	[-28.4, -28.2]	[-40.6, -16.8]

Table 1: Tobit. Partial Effects on Expectation.

Tables 1-2 contain the results on partial effects for Tobit. Table 1 contains the results on partial effects on expectation, PE^{Tob}, while Table 2 contains partial effects on probability, PE^{Pr}. All partial effects are evaluated at the mean values of covariates. In both tables, the first column ("Tobit") provides the partial effects for different covariates in the standard Tobit MLE where all covariates are assumed to be exogenous. The remaining columns are based on the two-step IV-Tobit estimator, where *huseduc* is used to instrument for the endogenous *nwifeinc.* The second column

	Tobit	IV-Tobit	CI for IV-Tobit	[LB, UB]	CI
nwifeinc	-0.303	-1.06	[-2.16, 0.043]	[-1.10, -1.06]	[-2.65, 0.157]
educ	2.77	3.92	[1.75, 6.10]	[3.92, 4.08]	[1.33, 7.48]
exper	4.52	4.18	[2.77, 5.59]	[4.18, 4.34]	[2.51, 6.51]
exper2	-0.064	-0.064	[-0.102, -0.026]	[-0.066, -0.064]	[-0.121, -0.022]
age	-1.87	-1.58	[-2.26, -0.890]	[-1.64, -1.58]	[-2.60, -0.834]

Table 2: Tobit. Partial Effects on Probability. All numbers are multiplied by 100.

	Probit	IV-Probit	CI for IV-Probit	[LB, UB]	CI
nwifeinc	-0.470	-1.39	[-2.67, -0.104]	[-1.49, -1.39]	[-3.29, 0.079]
educ	5.11	6.41	[3.96, 8.86]	[6.41, 6.87]	[2.98, 10.8]
exper	4.82	4.38	[2.68, 6.08]	[4.38, 4.70]	[2.49, 6.82]
exper2	-0.074	-0.073	[-0.118, -0.028]	[-0.079, -0.073]	[-0.137, -0.024]
age	-2.06	-1.69	[-2.58, -0.804]	[-1.81, -1.69]	[-2.87, -0.784]

Table 3: Probit. Partial Effects on Probability. All numbers are multiplied by 100.

("IV-Tobit") contains the naive estimators of the partial effects, followed by the 95% confidence intervals (column "CI for IV-Tobit"). Column "[LB, UB]" provides the proposed estimated bounds for the partial effects that account for both types of endogeneity. The last column contains the corresponding confidence intervals for the partial effects.

In both Tables 1 and 2, we observe that the confidence intervals for the correct partial effects at the mean are only slightly wider than the naive ones of IV-Tobit. In particular, using the correct inference approach does not change any of the conclusions about the effects of the variables being statistically significant.

Table 3 contains the corresponding results for Probit. Again, the confidence intervals for the correct partial effects are not much wider than the naive ones. Unlike the Tobit case, not all conclusions about the statistical significance of the partial effects are preserved, as the partial effect of nonwife income becomes statistically insignificant. Hence, if we base our analysis only on the binary outcome model, properly accounting for the roles of measurement error and structural endogeneity can reverse the conclusion that the partial effect of nonwife income is statistically significantly different from zero.

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.

Chesher, A., D. Kim, and A. M. Rosen (2023): "IV methods for Tobit models," Tech. rep., University College London.

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

McCloskey, A. (2017): "Bonferroni-based size-correction for nonstandard testing problems," *Journal of Econometrics*, 200, 17–35.

Mroz, T. A. (1987): "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica*, 55, 765.

Newey, W. K. (1987): "Efficient estimation of limited dependent variable models with endogenous explanatory variables," *Journal of Econometrics*, 36, 231–250.

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

Romano, J. P. and M. Wolf (2005): "Stepwise Multiple Testing as Formalized Data Snooping," *Econometrica*, 73, 1237–1282.

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 (11) 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 $\sigma_\varepsilon^2 \le \sigma_V^2$, let

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

Then

$$\underline{\sigma}_{U^*}^2 \le \sigma_{U^*}^2 \le \sigma_U^2, \qquad \underline{\sigma}_{U^*}^2 \equiv \max\left\{\sigma_U^2 - \theta_{01}^2\overline{\sigma}_\varepsilon^2, 0\right\}. \tag{20}$$

Note that

$$\sigma_U^2 - \theta_{01}^2 \frac{\sigma_U^2\sigma_V^2 - \sigma_{UV}^2}{\sigma_V^2\theta_{01}^2 + 2\sigma_{UV}\theta_{01} + \sigma_U^2} = \frac{\sigma_U^4 + 2\sigma_U^2\sigma_{UV}\theta_{01} + \theta_{01}^2\sigma_{UV}^2}{\sigma_V^2\theta_{01}^2 + 2\sigma_{UV}\theta_{01} + \sigma_U^2} = \frac{\left(\theta_{01}\sigma_{UV} + \sigma_U^2\right)^2}{\sigma_V^2\theta_{01}^2 + 2\sigma_{UV}\theta_{01} + \sigma_U^2}.$$

Thus, $\underline{\sigma}_{U^*}^2$ in equation (20) can be equivalently written as

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

where the fraction is always non-negative, ensuring that $\underline{\sigma}_{U^*}^2 \ge 0$. Finally, equation (11) implies that $\sigma_{U^*V^*}$ is bounded between σ_{UV} and $\sigma_{UV} + \theta_{01}\overline{\sigma}_\varepsilon^2$. ■

Derivation of equation (18):

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

where the penultimate equality comes from taking expectation with respect to V_i^*, and the last equality follows from equation (11). ■

Derivation of equation (19) is similar:

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

■