Gutenberg School of Management and Economics

& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

Estimating Causal Effects in Binary Response Models with Binary Endogenous Explanatory Variables

A Comparison of Possible Estimators

Manuel Denzer

December 10, 2019

Discussion paper number 1916

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
https://wiwi.uni-mainz.de/

Contact Details:

Manuel Denzer
Chair of Applied Statistics and Econometrics
Gutenberg School of Management and Economics
Johannes Gutenberg University
Jakob-Welder-Weg 4, D-55128 Mainz, Germany
manuel.denzer@uni-mainz.de

Estimating Causal Effects in Binary Response Models with Binary Endogenous Explanatory Variables

A Comparison of Possible Estimators*

Manuel Denzer †
Johannes Gutenberg-University Mainz

December 10, 2019

Abstract

This paper reviews and compares different estimators used in the past to estimate a binary response model (BRM) with a binary endogenous explanatory variable (EEV) to give practical insights to applied econometricians. It also gives a guidance how the average structural function (ASF) can be used in such a setting to estimate average partial effects (APEs). In total, the (relative) performance of six different linear parametric, non-linear parametric as well as non-linear semi-parametric estimators is compared in specific scenarios like the prevalence of weak instruments. A simulation study shows that the non-linear parametric estimator dominates in a majority of scenarios even when the corresponding parametric assumptions are not fulfilled. Moreover, while the semi-parametric non-linear estimator might be seen as a suitable alternative for estimating coefficients, it suffers from weaknesses in estimating partial effects. These insights are confirmed by an empirical illustration of the individual decision to supply labor.

JEL classification: C25, C26

Keywords: Binary choice, Binomial response, Binary Endogenous Explanatory Variable, Average Structural Function

*I thank David Card as Director of the Econometrics Laboratory at UC Berkeley allowing me to use the EML's infrastructure to run my simulations. I gratefully acknowledge comments on my paper by Constantin Weiser, Thorsten Schank, Christoph M. Schmidt, Florian Hett and seminar participants of the University of Mainz. All remaining errors are mine.

†Gutenberg School of Management and Economics, Jakob-Welder-Weg 4, 55128 Mainz, ✉ manuel.denzer@uni-mainz.de

Contents

1 Introduction

In different fields of economics the estimation of binary response models (BRMs) where the effect of interest comes from a variable which is dichotomous itself and additionally endogenous is often needed. Examples are from the fields of education (Evans & Schwab, 1995; Angrist, Bettinger, Bloom, King, & Kremer, 2002; Altonji, Elder, & Taber, 2005), health (Angrist & Evans, 1998), labor (Sasaki, 2002) or migration (Dong & Lewbel, 2015) and endogeneity may be not exclusively due to one specific reason although the case of a bias due to an omitted variable might be the most prominent.[1] The setting of a model with a binary endogenous explanatory variable (EEV) is closely related to the so-called treatment literature where the treatment indicator is represented by a binary variable, which is equal to unity for the part of the population which receives a treatment and for the rest it equals to zero. Although experimental economics has grown in importance in recent years, most of applied economist are still working with observational data where treatment decisions are often not determined by fully random draws and are therefore prone to endogeneity. Presumably, the most prominent study in this respect is the one of Angrist (1990) on individual earning outcomes where the treatment of serving in the US military during Vietnam war cannot be considered as random.

In general, taking care of endogeneity in non-linear models like BRMs is more challenging than in linear models. For instance, mimicking two-stage-least-squares (2SLS) in non-linear models by substituting endogenous variables with their fitted values leads to the so-called *forbidden regression* and inconsistent estimates (Hausman, 1975). The complexity for non-linear models is in particular pronounced when the EEV is binary itself (Wooldridge, 2010). Even in the case when there is no endogeneity present, many applied researchers rather tend to use a linear probability model (LPM) instead of a probit or logit model in the setting of a BRM due to its simplicity and by arguing that it delivers adequate estimates (Wooldridge, 2010). However, it is obvious that making use of the the information that the dependent variable is limited should be not harmful or disadvantageous and should theoretically outperform linear models due to their drawbacks.[2] One of those drawbacks as outlined by Imbens and Angrist (1994) is that a LPM could deliver estimates of an average treatment effect for a population which is not observed. Typically, using non-linear instead of linear models comes with the cost of making stronger assumptions (Imbens & Angrist, 1994) and this behavior can be amplified in an EEV setting due to the fact that it is necessary to deal with the endogeneity in order to get unbiased estimates.

This paper reviews and compares different estimators which have been used in previous studies in order to estimate the effect of a dichotomous EEV in a BRM. To be more precise, this study considers six estimators which differ in terms of their assumption on the non-linearity of the model and the parametric specification which jointly results in the complexity of the model to estimate. From the class of linear parametric models, the most popular, i.e. the most frequent

[1] Other reasons for endogeneity are measurement error, simultaneity, functional form misspecification or sample selection.

[2] Horrace and Oaxaca (2006) present conditions when the approximation of a non-linear model through a LPM yields consistent estimates for the average partial effects (APEs). However, the authors admit that this would require fortuitous circumstances.

used, 2SLS approach as well as a less known alternative to this one suggested by Angrist and Pischke (2009) using generated instruments is considered. In terms of non-linear parametric models the focus is on the performance of the maximum likelihood (ML) estimator of a *recursive bivariate probit* model (Heckman, 1978; Amemiya, 1978). With respect to the class of (non-linear) semi-parametric models, the performance of two different types of the estimator of the *special regressor* approach according to Lewbel (2000) is investigated. In order to be able to compare the performance of those estimators which are taking into account the endogeneity of the binary variable of interest with the one of an estimator who is ignoring it, the performance of a regular probit estimator is also presented. Given the fact that we concentrate on the impact of a binary instead of a continuous EEV in a BRM, we are unable to consider the prominent *control function* approach also known as two-stage-residual-inclusion (2SRI) approach as suggested by Rivers and Vuong (1988) and further developed by Blundell and Powell (2004).[3]

Our contribution to the existing literature is straightforward. According to the best of our knowledge, there exists no exhaustive analysis comparing the estimators mentioned above in the setting of a BRM with a binary EEV although all of those estimators have been frequently used in the past. Our goal is to critically discuss and answer the question which estimator should be used in this setting given specific scenarios within this setting. By making use of Monte Carlo simulations, we explicitly test the performance of the estimators given different scenarios. First, we analyze the performance given different second moments of one specific exogenous explanatory variable which is crucial for the *special regressor* approach. Second, we test the impact of weak instruments on the estimators as suggested by Lewbel, Dong, and Yang (2012). Third and finally, we investigate how different assumptions on a specific dependence structure reflecting the endogeneity in our proposed model can affect the predicting power of the estimators. Our reference for the evaluation of the performance of the estimators is always the *true* APE, i.e. the one which we set by our simulation setup. Since the calculation of partial effects given endogeneity is not as simple as in the case of its absence, we draw on insights of the average structural function (ASF) according to Blundell and Powell (2004) to get unbiased estimates of the partial effects. To the best of our knowledge, we are the first estimating APEs for the *special regressor* approach via the the ASF in a (simulated) setting of a BRM with a binary EEV.

Our findings are that the ML estimator is the best performing estimator in terms of a low root mean squared error (RMSE) in general but even in scenarios where the assumptions raised by the *recursive bivariate probit* approach are not fulfilled. Estimators of linear models turn out to outperform both *special regressor* approaches in terms of a smaller spread of their estimates

[3]This approach which delivers identical estimates as the 2SLS approach in a linear model is only valid in a setting with continuous EEVs. To be more precise, in the setting of a binary EEV the error term of the equation projecting the endogenous variable on the exogenous variables - implied by the *control function* approach - suffers by definition from heteroskedasticity and therefore can not be independent of the exogenous variables. However, this is a necessary condition as pointed out by Blundell and Powell (2004). This limitation has been overlooked in a couple of studies in the past. For instance Terza, Basu, and Rathouz (2008) advocating for the 2SRI approach in different classes of non-linear models. However, Wan, Small, and Mitra (2018) show that the conclusions made by Terza et al. (2008), i.e. that the 2SRI approach delivers consistent estimates for a variety of non-linear models, is only true under some unrealistic assumptions.

in most of the tested scenarios. One of the two *special regressor* approaches always dominates the other by delivering more accurate point estimates for the APEs.

The rest of the paper is organized as follows: Section 2 presents a short review of other studies investigating the performance of different estimators in similar settings to the one we investigate. Section 3 introduces the model. Section 4 discusses the different estimators under investigation and explains how we estimate the APEs. In Section 5 we present the results of our Monte Carlo simulations. An application to demonstrate the dominance of the *recursive bivariate probit* approach is given in Section 6 by the reestimation of results of the study of Angrist and Evans (1998). Finally, Section 7 briefly discusses our findings and concludes.

2 Literature

The number of studies comparing the performance of different estimators in settings of non-linear models with EEVs is rather small. For the specific setting of a BRM with only one binary EEV we are not aware of an exhaustive one. The studies which are briefly outlined in the following are similar to our study in particular in terms of the estimators under investigation.

Kang and Lee (2014) analyze the performance of six different estimators in a setting which is similar to ours, i.e. a BRM with a binary EEV, but also in another one where the EEV is continuous. They use a "real-data-based" simulation to investigate the performance of different parametric and semi-parametric approaches.[4] For the case where the EEV is binary they compare estimators of a 2SLS, a *control function*, an *artificial instrumental regressor*, a *special regressor* and *recursive bivariate probit* approach. For the *special regressor* approach they limit their analysis to the one of both types which uses a sorted data density and which we will identify to be dominated by the other. Moreover, it remains open if one of the crucial assumptions of this approach namely the conditional independence of the chosen special regressor is fulfilled in their setting. In addition to these debatable points, their comparison of the estimators is to some extent based on a measure of an APE which does not account for the prevalent endogeneity. In total, the authors recommend to use the 2SLS and the *control function* approach - even in scenarios where they are theoretically invalid - due to their analytic and computational simplicity.

Lin and Wooldridge (2015a) present an estimator for a setting of a BRM with one binary EEV and an unlimited number of continuous EEVs. Actually, their proposed estimator is a combination of the *control function* approach and the *recursive bivariate probit* approach: The errors of linear projections of all continuous EEVs on the set of exogenous variables are plugged into the structural equation of the *recursive bivariate probit* model before the ML estimator is applied. By running Monte Carlo simulations with reducing complexity to one binary and one continuous EEV each, the authors show that both types of their proposed estimator, i.e. the

[4] By the expression "real-data-based" simulation, the authors mean that the regressors they use in their model are drawn form real data instead of being artificially generated. However, they set the parameters of their equation system and the distribution of the errors manually.

one with sequentially averaged APEs as well as the one with jointly averaged APEs, outperform all other estimators namely the 2SLS, the original *control function* and both types of the *special regressor* approach.[5] However, it should be noted that maybe due to the fact that the *special regressor* approach is not in focus of this study there exists an inconsistency in this discussion paper complemented by a minor misconception regarding the construction of the *special regressor*.[6] Maybe due to this and other reasons the focus of this study has considerably changed in an updated version (Lin & Wooldridge, 2018) where the *special regressor* approach is no longer part of the investigation.[7]

By following the DGP setup described in Lewbel (2000), Bontemps and Nauges (2017) present a simulation study which focuses on a model of a two equation system where the structural equation is a BRM but where the EEV is continuous. Due to this characteristic - which is different to our analyzed setting - the reduced form equation is assumed to have a linear form. In their study, the authors compare the performance of an estimator of the *special regressor* approach in contrast to a ML estimator of a *control function* approach.[8] Their main findings are that the *large support* condition of the *special regressor* approach is a necessary one to get unbiased estimates and that the behavior of trimming data during the estimation process as suggested by Lewbel et al. (2012) can lead to a severe bias in the estimates. In contrast to the studies mentioned above, the authors evaluate the two different approaches based on the estimation of the coefficients of the structural equation instead of partial effects or alternatively coefficient ratios. However, it is obvious that besides other minor weaknesses such as mixing up the parameters of a uniform distribution or a misleading conclusion with respect to the trimming of data, this behavior is irritating since those coefficients are only identified up to scale.[9]

3 Model

In order to compare the different estimators with respect to their performance in a BRM with a binary EEV, we set up the following two equation model:

$$y_1 = 1[y_1^* > 0] = 1[\mathbf{x}\boldsymbol{\psi} + u_1 > 0] = 1[\gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + \alpha v + u_1 > 0] \tag{1}$$

$$y_2 = 1[y_2^* > 0] = 1[\mathbf{z}\boldsymbol{\delta} + \zeta v + u_2 > 0] \tag{2}$$

[5] The APEs of a joint ML estimation always serve as benchmark.

[6] To be more precise, although Lin and Wooldridge (2015a, p. 14) emphasize to use a continuous independent variable as the *special regressor* in the structural equation, its distribution is a Bernoulli one in the stated data generating process (DGP). Moreover, the statement of the authors that the *special regressor* method puts unrealistic assumptions on the reduced form equation turns out to be too strong.

[7] In the newer version of this study the model is amplified to allow for additional unobserved heterogeneity through a switching regime. Moreover, the application of the proposed model to fractional instead of binary response models is discussed and tests to assess the degree of endogeneity of the binary and the continuous EEVs are developed.

[8] The authors employ Stata's `ivprobit` command which is a *control function* estimator despite its name supposing an instrumental variable (IV) estimator.

[9] See Wooldridge (2005) for a detailed discussion why focusing on parameters in non-linear models can be misguiding.

Equation 1 is the structural equation which describes the relationship of interest. Equation 2 is the reduced form equation for the binary EEV y_2. The indicator function $1[\cdot]$ takes the value one when the statement in brackets is true and zero otherwise. The expressions y_1^* and y_2^* represent the unobserved dependent variables in terms of an underlying latent variable model, which are assumed to be linear functions of the explanatory variables. The vector $\mathbf{x} \in \mathbb{R}^{d_x = 2 + d_{z_1}}$ - where $d_.$ denotes the dimension of the respective vector - contains all explanatory variables of the structural equation, i.e. the endogenous as well as the exogenous ones. The vector $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2) \in \mathbb{R}^{d_z = d_{z_1} + d_{z_2}}$ is a vector of all exogenous variables in the model with the exemption of a special regressor denoted as $v \in \mathbb{R}$ with a normalized coefficient α equal to one.[10] In terms of the exclusion restriction $\mathbf{z}_1$ has to be a strict subset of $\mathbf{z}$, i.e. $d_{z_1} < d_z$. In all cases, the first element of the vector $\mathbf{z}_1$ is unity which given that the following minimal conditions $\mathbb{E}(\mathbf{z}_1\, u_1) = 0$ and $\mathbb{E}(\mathbf{z}\, u_2) = 0$ hold leads to the result of $\mathbb{E}(u_1) = \mathbb{E}(u_2) = 0$.[11]

This system of equations applies to the most prominent reason of endogeneity namely the case of an omitted variable. Other reasons for endogeneity such as simultaneity or measurement error are not captured by this system. In case of the first, y_2 would have to be a function of y_1, i.e. y_1 would have to be part of the right hand side (RHS) of the reduced form equation, i.e. Equation 2. In case of endogeneity due to measurement error, y_2 could be expressed as $y_2 = y_2^\circ + \epsilon_{y_2}$ where y_2° represents the true value of y_2 and ϵ_{y_2} the corresponding measurement error. Hence, we would have to have an additional error in the structural equation, i.e. $-\gamma\, \epsilon_{y_2}$.

Formally, endogeneity in the model described by both equations is prevalent when $\mathbb{E}(y_2\, u_1) \neq 0$ which means that y_2 and u_1 are correlated. This relationship can be equivalently expressed by a linear projection of one of the errors of the structural equation and the reduced form equation on the other:

$$u_1 = \chi\, u_2 + \nu_1 \tag{3}$$

where $\chi \equiv \mathbb{E}(u_2'\, u_2)^{-1}\, \mathbb{E}(u_2'\, u_1)$ has to be different from zero with $\mathbb{E}(u_2\, \nu_1) = 0$ by definition of a linear prediction.

In our analysis we do not set up the dependence structure between both errors u_1 and u_2 by means of a linear projection as described in Equation 3, which is besides directly parameterizing the correlation between both errors the most popular way of doing this (cf. Kang and Lee (2014) Lin and Wooldridge (2015a), Lin and Wooldridge (2018)). Instead we use a copula function approach. A copula function approach uses the marginal cumulative distribution functions (CDFs) of a set of variables and a specific copula function to construct a joined CDF for those variables (Cameron & Trivedi, 2005). Let $\mathbb{C}_\theta(\ddot{u}_1, \ddot{u}_2)$ be a copula function with parameter θ governing the strength of dependence and $\ddot{u}_1$ and $\ddot{u}_2$ being equal to the marginal CDFs of u_1 and u_2 respectively, i.e. $\ddot{u}_1 = F_{u_1}$ and $\ddot{u}_2 = F_{u_2}$ where $F_.$ stands for the marginal

[10] In Subsection 4.1.4 we will explain why we separate v from the other exogenous variables contained in $\mathbf{z}$. The normalization of the coefficient α is harmless and equivalent to normalizing the variance of the error term to be equal to one like in a probit model.

[11] Formally, those two minimal conditions $\mathbb{E}(\mathbf{z}_1\, u_1) = 0$ and $\mathbb{E}(\mathbf{z}\, u_2) = 0$ have to hold because otherwise it is not ruled out that elements of the vector $\mathbf{z}$ are endogenous as well.

CDF. Then, $\mathbb{C}_\theta(\ddot{u}_1, \ddot{u}_2) = \mathrm{P}\{\ddot{U}_1 \leq \ddot{u}_1, \ddot{U}_2 \leq \ddot{u}_2\} = \mathbb{C}_\theta(F_{u_1}, F_{u_2}) = F_{U_1,U_2}(u_1, u_2)$ is the most general expression of a copula function for the bivariate case which is equal to the joint CDF of both components $\ddot{u}_1, \ddot{u}_2$ where $\ddot{U}_1, \ddot{U}_2$ are realizations of those components (Nelsen, 2006). The advantage of using a copula in order to describe the exact dependence structure between both errors of the single equations of our model is that it allows to be more flexible in comparison to just defining the Spearman correlation coefficient or the parameter α in Equation 3. The copula function does not pose any distribution restrictions on the marginal CDFs. Therefore, the generation of a flexible multivariate distribution is simple by just choosing both marginal CDFs and the copula function which binds them together (Cameron & Trivedi, 2005). In Section 5 we will use copulas of different classes namely the Elliptical and Archimedian class for our simulations. However, our baseline for all simulations will be that the marginal CDFs of both errors u_1 and u_2 follow a normal distribution and that the copula used is a Gaussian one which is part of the Elliptical class. In other words, we set $F_{U_1,U_2}(u_1, u_2) = \Phi_2(u_1, u_2, \theta)$ where Φ_2 represents a bivariate normal distribution. Obviously, we rule out the product / independence copula in all the following since in this case our model would not suffer from any endogeneity problems.

4 Estimation

4.1 Estimators

In this subsection we describe the different estimators used which comprise specific parametric as well as semi-parametric approaches. We refrain from applying non-parametric approaches which have been suggested by Vytlacil and Yildiz (2007) or Chesher and Rosen (2013) and also other semi-parametric approaches like the ones of Yildiz (2013), Mu and Zhang (2018) or most recently Han and Lee (2019) which are relatively more complex. Indeed, we focus on those approaches which have been frequently used in the previous literature in order to estimate a causal effect in the setting of a BRM with a binary EEV.[12] Hence, we also consider two common estimators for linear models in our analysis.

4.1.1 2SLS

Originally proposed by Wright (1928) to estimate the elasticities of supply and demand for butter and linseed oil, the 2SLS approach is the most frequently used approach in the class of linear instrumental variable regressions. It identifies the part of the variation of the endogenous variable in the structural equation which is not prone to endogeneity by the means of an ordinary-least-squares (OLS) regression of the reduced form equation and plugs the fitted values of the

[12] In fact, we checked the number of citations of those above mentioned publications and found that although they have been published in journals such as *American Economic Review* or *Econometric Theory* the number of citations is considerably low for Yildiz (2013) and Chesher and Rosen (2013). It amounts to not more than five and fifteen citations respectively according to Google Scholar when correcting for double counts. Vytlacil and Yildiz (2007) is published in *Econometrica* and cited much more often. Mu and Zhang (2018) is published in the *Econometrics Journal* and has not be cited until now.

reduced form equation as substitutes for the endogenous variable into the structural equation (Wooldridge, 2010). Given that the relevance and exclusion conditions hold and the instrument variable does not suffer from endogeneity itself, the 2SLS approach allows for consistent estimates by regressing the adjusted structural equation with OLS whereby the estimates of the standard errors have to account for the estimation error in the reduced form equation itself. Whereas this method is an adequate approach for continuous dependent variables it has some drawbacks in the case of explaining a limited dependent variable like a binary response. Indeed, the 2SLS method shares the same drawbacks a LPM has in the case of explaining a binary choice outcome in the absence of any endogeneity. First, it can yield to predicted probabilities smaller than zero or larger than one respectively. Second, the error term suffers from heteroskedasticity by definition. Third and finally, this method assumes constant marginal effects (Wooldridge, 2010). In contrast, the advantages of the 2SLS approach are that it is simplistic - in particular when using professional statistical software which allows to directly estimate the standard errors properly - and that it sets no assumptions on the distribution of the error terms of both the structural equation and the reduced form equation. Despite the obvious deficiencies, the 2SLS method has often been used in non-linear models settings like in Angrist and Evans (1998), Evans, Farrelly, and Montgomery (1999), Angrist et al. (2002), Angrist, Lavy, and Schlosser (2010), Conley and Heerwig (2011), Islam and Raschky (2015) or Farbmacher, Guber, and Vikström (2018).

4.1.2 Alternative 2SLS

Angrist and Pischke (2009) propose an alternative to the classic 2SLS estimator in the context of the discussion of the so-called *forbidden regression*, i.e. the view that mimicking 2SLS in non-linear models by substituting endogenous variables with their fitted values yields to inconsistent estimates. Their idea is to combine a linear model approach with the information that the EEV itself is binary. Therefore, they propose a three-step estimation approach where in the first step the BRM described by the reduced form equation is estimated by a non-linear model, for instance a probit model. In the second step predictions of this regression are calculated. In a third step those fitted values are used as substitutes for the original instruments in a classic 2SLS approach. According to Angrist and Pischke (2009), this estimator could be more efficient in comparison to the classical 2SLS estimator if the assumed non-linear model in the first step delivers a better approximation of the reduced form equation's conditional expectation function (CEF) as a linear model would do. Although, this estimator might be more efficient than the classical 2SLS one it shares the same drawbacks (Newey, 1990). Wooldridge (2010) implicitly mentions this three-step estimation procedure in the context of *Two-Stage Least Squares with Generated Instruments*. Therefore, we call this estimator in the following the *Alternative 2SLS* or the *Generated Instrument 2SLS*.

4.1.3 Recursive Bivariate Probit

The *recursive bivariate probit* approach relies on the work of Heckman (1978) and Amemiya (1978). In contrast to both types of the 2SLS estimator, it explicitly addresses the non-linearity present in the Equations 1 and 2. For that reason, it raises assumptions on the marginal and joint distribution of the errors of the structural and the reduced form equation. In fact, this approach assumes that (u_1, u_2) follows a bivariate normal distribution Φ_2 which implies that it is assumed that both the structural and the reduced form equation are probit models each. Moreover, since it is a fully parametrized approach it requires that the exact index functions - in our case y_1^* and y_2^* - are known to the researcher (Wooldridge, 2010). Given this information a ML estimator is applied to estimate the simultaneous equation model and all parameters are jointly identified.

The likelihood of the four different choice probabilities of the four different states of (y_1, y_2) for the joint identification of the parameters of this model can be written as

$$\mathcal{L}(\boldsymbol{\kappa}) = \prod P(y_1 = 1, y_2 = 1|\boldsymbol{\kappa})^{y_1 y_2} \cdot P(y_1 = 0, y_2 = 1|\boldsymbol{\kappa})^{(1-y_1)y_2} \cdot P(y_1 = 1, y_2 = 0|\boldsymbol{\kappa})^{y_1(1-y_2)} \cdot P(y_1 = 0, y_2 = 0|\boldsymbol{\kappa})^{(1-y_1)(1-y_2)}$$

where $\boldsymbol{\kappa}$ represents a vector of all parameters to be estimated.

As outlined in Greene (2017), the four choice probabilities in the *recursive bivariate probit* approach are equal to

$$\begin{aligned}
P(y_1 = 1, y_2 = 1) &= \Phi_2(\gamma + \mathbf{z}_1\boldsymbol{\beta} + \alpha v, \mathbf{z}\boldsymbol{\delta} + \zeta v, \rho) \\
P(y_1 = 0, y_2 = 1) &= \Phi_2(-(\gamma + \mathbf{z}_1\boldsymbol{\beta} + \alpha v), \mathbf{z}\boldsymbol{\delta} + \zeta v, -\rho) \\
P(y_1 = 1, y_2 = 0) &= \Phi_2(\mathbf{z}_1\boldsymbol{\beta} + \alpha v, -(\mathbf{z}\boldsymbol{\delta} + \zeta v), -\rho) \\
P(y_1 = 0, y_2 = 0) &= \Phi_2(-(\mathbf{z}_1\boldsymbol{\beta} + \alpha v), -(\mathbf{z}\boldsymbol{\delta} + \zeta v), \rho)
\end{aligned}$$

where one advantage of the *recursive bivariate probit* approach immediately appears: The endogenous nature of y_2 has no impact on the formulation of the likelihood.[13] As a result the log-likelihood simplifies to

$$\ln\mathcal{L} = \sum \ln\Phi_2[q_{y_1}(\gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + \alpha v)\,,\, q_{y_2}(\mathbf{z}\boldsymbol{\delta} + \zeta v)\,,\, q_{y_1} \cdot q_{y_2} \cdot \rho]$$

where $q_{y_1} = (2y_1 - 1)$ and $q_{y_2} = (2y_2 - 1)$.

It is well known that in the case when the densities of the ML function are correctly specified - which implies that the index functions and the distribution of the error terms are correctly specified - , the ML estimator is the most asymptotically efficient estimator (Wooldridge, 2010). However, this requirement of a complete and correct specification is a severe burden because

[13] To be more precise, as discussed in Greene (2017) and Wooldridge (2010), the likelihood function of a *regular bivariate probit* model where two probit models have correlated disturbances but where in contrast to the *recursive bivariate probit* model y_1 is not a function of y_2 is the same as the one shown above.

misspecification typically leads to inconsistent estimates in the absence of quasi-maximum likelihood (QML) (Wooldridge, 2010). Moreover, given its iterative search the *recursive bivariate probit* approach is prone to computational convergence problems when estimating nuisance parameters, for instance due to local (multiple) maxima of the likelihood function.

4.1.4 Special Regressor

Lewbel (2000) proposed another approach which can be used to estimate the causal effect of a binary EEV in a BRM. In contrast to the approaches just described it is a semi-parametric approach, i.e. it is not as much requiring on distributional assumptions as the other approaches. Moreover, this estimator is not based on an iterative search such as a ML-estimator and it is also not limited to the specific setting of a BRM but can and has been used in a universe of EEV settings such as ordered or multinomial choice models, censored regression models, selection and dynamic choice models in the past.[14] While this approach relaxes assumptions such that a specification of the reduced form equation of the endogenous variable is needed, it requires certain properties of one exogenous regressor in the structural equation which will be represented by v in the following.[15] Moreover, it prevents to have polynomials of the special regressor v as additional regressors due to a key conditional distribution assumption. On the other hand, it can be extremely helpful for the case of multiple endogenous regressors since only one valid *special regressor* is needed for all of them.

According to Dong and Lewbel (2015), the special regressor v has to fulfill three requirements: First, it has to appear additively to u_1 in the structural equation. Second, the special regressor v is required to be continuously distributed with a sufficient large support and third it is required to be conditionally independent of the error term of the structural equation, i.e. $u_1 \perp v|\mathbf{z}, y_2$.

The first requirement is always fulfilled in BRMs as it can be seen by the latent variable form expression of Equation 1. With respect to the second requirement, the *special regressor* requires that the condition $supp(\gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + u_1) \subseteq supp(-v)$ holds. However, as it as been shown by Magnac and Maurin (2007) the part of this requirement referring to the continuous distribution can be relaxed to a distribution with tail symmetry. The third requirement can be simplified by means of a linear projection of v on $(y_2, \mathbf{z})$ in order to be able to assess only the unconditional independence.

Given a special regressor fulfilling those requirements as well as the standard assumptions of a 2SLS estimation hold, i.e. $\mathbb{E}(\mathbf{z}u_1) = 0$ and $rank(\mathbb{E}(\mathbf{z}\mathbf{z}_1')) = d_z$, the virtue of the semi-parametric *special regressor* approach is that the conditional expectation $\mathbb{E}(y_1|y_2, \mathbf{z}, v)$ will equal the conditional distribution of $\gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + u_1$ conditioning on $(y_2, \mathbf{z})$ and evaluated at v.[16] This equality can then be used to get consistent estimates of the coefficients of the structural equation, i.e. Equation 1, by the following multi-step estimation procedure proposed by Dong and Lewbel

[14]For a full list, see Dong and Lewbel (2015).

[15]Due to this circumstance, this methods got its name the *special regressor* approach.

[16]See Dong and Lewbel (2015) for a simple example with just one parameter to estimate.

(2015) based on earlier work of Lewbel et al. (2012), Lewbel (2007) and Lewbel (2000):

Step 1 Demean the special regressor v and estimate a linear projection of v on $\mathbf{z}$ and y_2. For each observation i calculate the residuals $\hat{w}$ of this regression, i.e. $\hat{w}_i = v_i - \mathbf{z}_i \hat{\boldsymbol{\tau}} - \hat{\xi}\, y_{2i}$

Step 2 Estimate the density of w (f_i) at each observation $i = 1 \ldots n$ by one of the two possibilities:

2.1 *Kernel Density (KeDe) Estimator*: Apply a standard one-dimensional kernel estimator

$$\hat{f}_i = \frac{1}{nh} \sum_{j=1}^{n} K\left(\frac{\hat{w}_i - \hat{w}_j}{h}\right) \tag{4}$$

to each observation where h is the bandwidth and K is a symmetric kernel density function.

2.2 *Sorted Data (SoDa) Estimator* according to Lewbel (2007): Sort the observations in terms of $\hat{w}$ from lowest to highest. Treat $\hat{w}^+$ as the next largest value of $\hat{w}$ (after removing any ties) and similarly $\hat{w}^-$ the next smallest value. Define the density as

$$\hat{f}_i = \frac{2/n}{\hat{w}^+ - \hat{w}^-} \tag{5}$$

Step 3 For each observation i construct $\hat{t}_i$ where

$$\hat{t}_i = \frac{y_{1i} - 1[v_i \geq 0]}{\hat{f}_i}$$

Step 4 Regress $\hat{t}_i$ on $(y_2, \mathbf{z}_1)$ by classical 2SLS where y_2 is instrumented by $\mathbf{z}_2$. According to Theorem 1 in the Appendix this regression yields consistent estimates of $(\gamma, \boldsymbol{\beta})$.[17]

Standard errors of this multi-step estimation procedure for $\hat{\gamma}, \hat{\boldsymbol{\beta}}$ can be obtained by bootstrap mechanisms.

4.2 Partial Effects

In the case of comparing estimators of different non-linear models one has to know that comparing their coefficients is inappropriate since those parameters are only identified up to scale (Greene, 2017). In order to compare different estimators regarding their performance one can either compare coefficient ratios by randomly choosing one coefficient as the base category or by estimating partial effects (Wooldridge, 2005). Since we are interested in the direct effect of y_2 on y_1 in our BRM, we use partial effects resting on the respective estimated coefficients to evaluate the performance of the different estimators described in the previous subsection.

[17] Dong and Lewbel (2015) as well as Lewbel et al. (2012) point out that it might be necessary to discard outliers in this final step. However, based on our results we agree with Bontemps and Nauges (2017) that discarding outliers by trimming data in this step is disadvantageous in the sense that it leads to biased estimates. Therefore, we do not discard outliers in the following simulations.

Whereas the estimation of partial effects in linear models is trivial, since they are identical to the parameters itself or a combination of those, estimating partial effects in non-linear models is on average hardly more difficult. There exist two common measures to obtain a single number for an average partial effect both relying on the choice probability $CP(\mathbf{x}) = P(y_1 = 1|\mathbf{x}) = \mathbb{E}(y_1|\mathbf{x})$ also synonymously known as propensity score.[18] Given the model introduced in Section 3, the choice probability can be written as

$$\mathbb{E}(y_{1i}|\mathbf{x}_i) = CP(\mathbf{x}_i) = F_{-u_1|\mathbf{x}}(\mathbf{x}_i\boldsymbol{\psi}|\mathbf{x}_i) \tag{6}$$

where $F_.$ still represents the respective marginal CDF, and in terms of a better understanding of the following derivations we stop suppressing the index i indicating the individual observations. The first measure known as the APE averages individual partial effects across the distribution of the respective explanatory variable, whereas the second measure, i.e. the partial effect of the average (PEA), typically plugs the mean of the fitted values of the latent variable y_1^* into the marginal CDF of the error term also known as link function (Wooldridge, 2010). Since a major drawback of the PEA is that it may not represent the partial effect of any unit in the population, which is in particular in the case of discrete explanatory variables (Wooldridge, 2010), we stick to the APE definition to describe an average partial effect. Based on the choice probability this is defined as

$$APE_x = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial\, CP(\mathbf{x}_i)}{\partial\, x} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial\, F_{-u_1|\mathbf{x}}(\mathbf{x}_i\boldsymbol{\psi}|\mathbf{x}_i)}{\partial\, x} \tag{7}$$

for all continuous $x \in \mathbf{x}$ and

$$APE_x = \frac{1}{n}\sum_{i=1}^{n} CP(\mathbf{x}_i^{(1)}) - CP(\mathbf{x}_i^{(0)}) = \frac{1}{n}\sum_{i=1}^{n} F_{-u_1|\mathbf{x}}(\mathbf{x}_i\boldsymbol{\psi}|\mathbf{x}_i^{(1)}) - F_{-u_1|\mathbf{x}}(\mathbf{x}_i\boldsymbol{\psi}|\mathbf{x}_i^{(0)}) \tag{8}$$

for all discrete $x \in \mathbf{x}$ where the superscripts (1) and (0) denote that the respective discrete explanatory variable is being fixed at particular values.

In the presence of EEVs the calculation of partial effects is more ambitious than when the model only contains exogenous explanatory variables. The choice probability $\mathbb{E}(y_1|\mathbf{x})$ is no longer a suitable measure since it is affected by the correlation between the error term of the structural equation and the endogenous regressors (Lin & Wooldridge, 2015b).[19] Lewbel et al. (2012) propose a function called average index function (AIF) as a substitute for the choice probability in the presence of EEVs. The AIF reads:

$$AIF(\mathbf{x}_i) = \mathbb{E}(y_{1i}|\mathbf{x}_i\boldsymbol{\psi}) = F_{-u_1|\mathbf{x}\boldsymbol{\psi}}(\mathbf{x}_i\boldsymbol{\psi}|\mathbf{x}_i\boldsymbol{\psi}) \tag{9}$$

In their point of view this measure reflects the quantity of interest since it can be regarded as a counterfactual propensity score, i.e. a propensity score if EEVs would not be present.

[18] This term is in particular used in the potential outcomes literature, see for instance Rosenbaum and Rubin (1983).

[19] Only in the case when the marginal CDF of the error term is parametrized itself, the choice probability remains to be a meaningful quantity of interest.

Moreover, it offers the advantage that is relatively easy to estimate by a one-dimensional non-linear regression due to the fact that $\mathbf{x}_i\boldsymbol{\psi}$ is just a scalar.[20] However, Lin and Wooldridge (2015b) show that the AIF suffers from the same weaknesses as the choice probability in terms of estimating APEs when endogeneity is present. By providing examples of a simple linear model as well as a BRM, the authors derive the bias the AIF exhibits in estimating a meaningful response probability. Due to this insight, we make use of the concept of the ASF proposed by Blundell and Powell (2003, 2004) which is widely accepted to be able to estimate counterfactual propensity scores although it is computationally more cumbersome (Wooldridge, 2010). Given our model described in Section 3 the ASF can be denoted as:

$$\begin{aligned} ASF(\mathbf{x}_i) &= \mathbb{E}_{u_{1i}}(y_{1i}) = \mathbb{E}_{u_{1i}}\big[\,1[\mathbf{x}_i\boldsymbol{\psi} + u_{1i} > 0]\,\big] \\ &= \int 1[\mathbf{x}_i\boldsymbol{\psi} > -u_{1i}]\, f_{u_{1i}}(u_{1i})d(u_{1i}) = F_{-u_{1i}}(\mathbf{x}_i\boldsymbol{\psi}) \end{aligned} \tag{10}$$

where the elements of vector $\mathbf{x}$ are fixed and where $\mathbb{E}_{u_{1i}}(\cdot)$ denotes the expected value with respect to u_{1i}.

In simple words, the ASF is a function of specific values of the observed covariates which breaks the correlation between the error term and the endogenous variables in the structural equation by averaging out the unobservables contained in the error term across the population without conditioning on the explanatory variables.[21] Once properly estimated, the ASF replaces the choice probability in the formulas defining the calculation of the APEs valid for the case when no endogeneity is present.[22]

In terms of estimating partial effects for the *recursive bivariate probit* approach in the following simulations, we make use of the choice probability in Equation 6. As already explained before, this is appropriate despite the presence of endogenous variables since this ML method explicitly assumes the marginal CDFs of the error terms of the structural and the reduced form equation to follow a bivariate normal distribution (see Subsection 4.1.3).[23]

In terms of estimating partial effects for both types of the *special regressor* approach we take up the idea by Lee and Li (2018). In fact, we introduce the concept of conditional ASFs to be able to estimate the unconditional ASF for each individual i. To be more precise, we condition on the values of the endogenous, the exogenous as well as the instrumental variables such that Equation 10 reads as follows

$$\begin{aligned} ASF(\mathbf{x}_i) &= \int_{Supp(y_2,\mathbf{z})} \mathbb{E}_{u_{1i}|y2,\mathbf{z}}(y_{1i}|y_2 = y_{2i}, \mathbf{z} = \mathbf{z}_i, v = v_i)\, dF_{y_2,\mathbf{z}}(y_{2i}, \mathbf{z}_i) \\ &= \int_{Supp(y_2,\mathbf{z})} \Big[\, F_{-u_{1i}|y2,\mathbf{z}}(y_{2i}, \mathbf{z}_i, v_i) \,\Big]\, f_{y_2,\mathbf{z}}(y_{2i}, \mathbf{z}_i)\, d(y_{2i}, \mathbf{z}_i) \end{aligned} \tag{11}$$

[20]See Derivation 2 in the Appendix for a derivation of the corresponding non-linear estimator.
[21]For more details, see Blundell and Powell (2004).
[22]Given the definitions of the choice probability, the AIF and the ASF, one can immediately see that in the absence of endogeneity, i.e. $u_1 \perp \mathbf{x}$, all those measures are identical.
[23]For more details, see Greene (2017)

where the conditional marginal CDF $F_{-u_{1i}|y2,\mathbf{z}}(y_{2i}, v_i, \mathbf{z}_i)$ is assumed to be differentiable with respect to v. Instead of using the means of the variables contained in $\mathbf{x}$ as supposed by Lee and Li (2018) to construct the index value $\hat{a} = \mathbf{x}\hat{\boldsymbol{\psi}}$ which is then be used to redefine v as $v = \hat{a} - \mathbf{z}_1\hat{\boldsymbol{\beta}}$, we construct $\hat{a}_i$ and hence the redefined v_i for each observation i individually.[24] The partial effect of v for observation i can then be obtained by taking the derivative of Equation 11, with respect to v:

$$\begin{aligned}\frac{\partial ASF(\mathbf{x}_i)}{\partial v} &= \int_{Supp(y_2,\mathbf{z})} \frac{\partial}{\partial v}\mathbb{E}_{u_{1i}|y2,\mathbf{z}}(y_{1i}|y_2 = y_{2i}, \mathbf{z} = \mathbf{z}_i, v = v_i)\, dF_{y_2,\mathbf{z}}(y_{2i}, \mathbf{z}_i) \\ &= \int_{Supp(y_2,\mathbf{z})} \left[\frac{\partial}{\partial v} F_{-u_{1i}|y2,\mathbf{z}}(y_{2i}, \mathbf{z}_i, v_i)\right] f_{y_2,\mathbf{z}}(y_{2i}, \mathbf{z}_i)\, d(y_{2i}, \mathbf{z}_i) \\ &= PE_v(\mathbf{x}_i)\end{aligned} \tag{12}$$

The individual partial effects of the variables y_2 and those contained in $\mathbf{z}_1$ are obtained by multiplying the respective coefficients of the structural equation with the expression of Equation 12, for instance $PE_{y_2}(\mathbf{x}_i) = \gamma PE_v(\mathbf{x}_i)$. APEs are obtained in analogy to Equation 7 as explained before.

In terms of the estimation of the conditional marginal CDF $F_{-u_{1i}|y2,\mathbf{z}}(y_{2i}, v_i, \mathbf{z}_i)$ we follow Lee and Li (2018) and use a non-parametric logistic series estimator.[25] Let $R^K(v_i, \mathbf{z}_i)$ be a K-dimensional vector of basis functions and $\pi_K \in \mathbb{R}^K$ the argument which minimizes the log-likehood function

$$\mathcal{L}_n(\pi) = \sum_{i=1}^{n} y_{1i}\, ln\, \Lambda\big[R^K(y_{2i}, \mathbf{z}_i, v_i)'\pi\big] + (1 - y_{1i})\, ln\Big[1 - \Lambda\big[R^K(y_{2i}, \mathbf{z}_i, v_i)'\pi\big]\Big] \tag{13}$$

with $\Lambda(c) = exp^c/(1 + exp^c)$ being the logistic link function. Then $F_{-u_{1i}|y2,\mathbf{z}}(y_{2i}, \mathbf{z}_i, v_i)$ becomes $\Lambda(R^K(y_{2i}, \mathbf{z}_i, v_i)'\hat{\pi}_K)$ and this as well as the fact that each observation i is not a duplicate of another observation can be used to simplify Equation 12 to

$$PE_v(\mathbf{x}_i) = \frac{1}{n}\sum_{i=1}^{n}\left[\Lambda'\big[R^K(y_{2i}, \mathbf{z}_i, v_i)'\hat{\pi}_K\big] * \frac{\partial}{\partial v} R^K(y_{2i}, \mathbf{z}_i, v_i)'\hat{\pi}_K\, \Big|_{v_i = \hat{a}_i - \mathbf{z}_{1i}\hat{\boldsymbol{\beta}}}\right] \tag{14}$$

where Λ' denotes the logistic probability density function (PDF) $\Lambda'(c) = \Lambda(c)(1 - \Lambda(c))$.[26]

[24] If we would follow Lee and Li (2018) by using the means of $\mathbf{x}$, we would obtain estimates for the partial effects which can be compared to other PEA estimates. However, as explained previously, those are less meaningful and not comparable to estimates of partial effects via the AIF.

[25] Due the specific conditional setting a non-parametric kernel regression is not an alternative to estimate the conditional marginal CDF.

[26] Based on the idea of Lee and Li (2018) and the program *sspecialreg* of Baum (2012) which delivers estimates of the parameters of the structural equation (as well as partial effects via the AIF), we wrote an own Stata program called *mspecialreg* which estimates partial effects using the ASF constructed by the mentioned non-parametric regression. The ado-file is available upon request.

5 Monte Carlo Simulations

In this section the results of the Monte Carlo simulations are presented to assess the performance of the different estimators introduced in the last section in the setting of a BRM with a binary EEV. First, the design of the baseline setting is defined, which is then followed by the presentation of the corresponding results. Afterwards, the results of different scenarios are discussed when some of the assumptions made in the baseline design are altered in order to test (1) the impact of the distribution of the special regressor, (2) the prevalence of weak instruments, and (3) different assumptions regarding the endogeneity structure causing the omitted variable bias when estimating Equation 1 without controlling for endogeneity.

5.1 Baseline DGP

For all subsequent simulations, the baseline DGP where the individual observations $i = 1, \ldots, n$ are independently drawn from is assumed as follows:

The error terms (u_1, u_2) are chosen to be both normally distributed and linked to each other by the means of a Gaussian copula such that they are jointly following a bivariate normal distribution $\Phi_2(u_1, u_2, \theta)$:

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \sim Normal \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix} \right]$$

The special regressor v is also assumed to be normally distributed:

$$v \sim Normal(0, 2)$$

Without loss of generality, the set of exogenous variables $\mathbf{z}$ consists of one variable each for both $\mathbf{z}_1$ and $\mathbf{z}_2$. The exogenous covariate $z_{11} \in \mathbf{z}_1$ and the instrument $z_{21} \in \mathbf{z}_2$ are generated as follows:[27]

$$z_{11} \sim Normal(0, 1) + 0.1 * v$$
$$z_{21} \sim Normal(0, 1)$$

[27] A correlation between the instrument and the special regressor would not cause a violation of any assumption. However, since this correlation would have a major impact on the strength of the instrument in the parametric linear in comparison to the semi-parametric non-linear approaches, we set the correlation to zero in favor of a more meaningful comparison.

The binary EEV and the binary dependent variable are generated as:

$$
\begin{aligned}
y_2^* &= 0.1 + 1.5 * z_{21} - 0.8 * z_{11} + 0 * v + u_2 \\
y_2 &= 1[y_2^* > 0] \\
y_1^* &= 0.1 + 1.0 * y_2 - 0.5 * z_{11} + 1.0 * v + u_1 \\
y_1 &= 1[y_1^* > 0]
\end{aligned}
$$

where the coefficients are chosen such that the share of zeros and ones is close to 50% for y_1 as well as y_2.[28]

Hence, the true ASF in the baseline setting for observation i which is identical to the choice probability is:

$$ASF(\mathbf{x}_i) = CP(\mathbf{x}_i) = \Phi(0.1 + 1.0 * y_{2i} - 0.5 * z_{11i} + 1.0 * v_i)$$

and therefore the true APEs for y_2 and z_{11} respectively are:

$$
\begin{aligned}
APE_{y_2}^{true} &= \frac{1}{n}\sum_{i=1}^{n}\Big[\Phi(0.1 + 1.0 * 1 - 0.5 * z_{11i} + 1.0 * v_i) - \Phi(0.1 + 1.0 * 0 - 0.5 * z_{11i} + 1.0 * v_i)\Big] \\
APE_{z_{11}}^{true} &= \frac{1}{n}\sum_{i=1}^{n}\phi(0.1 + 1.0 * y_{2i} - 0.5 * z_{11i} + 1.0 * v_i) * -0.5
\end{aligned}
$$

The estimated APEs of the different estimators described in Section 3 are contrasted to their true values by the measures of the mean difference (BIAS) and the root mean squared error (RMSE):

$$BIAS = \frac{1}{R}\sum_{r=1}^{R}(\widehat{APE}_r - APE_r^{true}) \qquad RMSE = \sqrt{\frac{1}{R}\sum_{r=1}^{R}(\widehat{APE}_r - APE_r^{true})^2}$$

where R is the number of replications which is set to 1000 for all simulations.

5.2 Results

5.2.1 Baseline

Table 1 shows the results of Monte Carlo simulations for different numbers of observations in a setting described in the previous subsection. In this setting the instrument's power is strong and the second and third requirement of the *special regressor* approach are fulfilled as well.[29] Besides the estimators mentioned in Section 4, results of naive probit estimations are shown in column (6) to contrast the estimators taken into account the endogeneity of y_2.

[28] The coefficient of the special regressor v in the equation determining y_2^* is set to zero without the loss of generality.

[29] Note that the first requirement of the *special regressor* approach is fulfilled by definition, see Subsection 4.1.4.

Among the different estimators listed in the table, there does not exist a clear pattern with respect to the measure of the BIAS introduced in the previous subsection. While the probit estimator has, as expected, always the largest bias, i.e. it is upward biased by 54%, for four out of the five different sample sizes the *special regressor* sorted data estimator has the second largest bias. For the estimators displayed in column (1) until (4) the order of the estimators in terms of the smallest bias depends on the sample size, and the bias does not shrink consistently with the increase of the sample size among each estimator. As it can explicitly be seen by the last column of Table A1, this is due to the fact that the bias bounces around for all estimators except the probit one with an increase in the sample size.

With respect to the measure of the RMSE, the *recursive bivariate probit approach* always dominates the other estimators. Hence, it is the most efficient one. The GI2SLS estimator always performs better than the 2SLS estimator which in turn is dominated by the *special regressor* kernel density estimator. Similar to the finding in terms of the BIAS, the *special regressor* sorted data estimator as well as the probit estimator always have the relative largest RMSEs.

Table 1: Simulation Results for APE of y_2 - Baseline

	(1) 2SLS	(2) Generated Instr. 2SLS	(3) Recursive Biprobit	(4) Special Reg. Kernel Dens.	(5) Special Reg. Sorted Data	(6) Probit
Baseline			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.0023	0.0028	-0.0001	-0.0027	-0.0039	0.0934
RMSE	0.0117	0.0109	0.0091	0.0130	0.0163	0.0936
N = 5000						
BIAS	0.0026	0.0029	0.0005	-0.0016	-0.0033	0.0933
RMSE	0.0165	0.0152	0.0131	0.0182	0.0229	0.0938
N = 1000						
BIAS	0.0040	0.0055	0.0016	-0.0013	-0.0051	0.0937
RMSE	0.0368	0.0343	0.0292	0.0419	0.0513	0.0959
N = 500						
BIAS	0.0009	0.0018	-0.0013	-0.0027	-0.0068	0.0930
RMSE	0.0507	0.0475	0.0414	0.0582	0.0716	0.0975
N = 250						
BIAS	-0.0024	0.0002	-0.0047	-0.0070	-0.0129	0.0927
RMSE	0.0718	0.0658	0.0591	0.0825	0.0981	0.1016

Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $corr(u_1, u_2) = 0.6$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$; The BIAS is defined as the average difference between the estimated and the true APE. RMSE is the root mean squared error. Detailed statistics can be found in Table A1. A visual overview of the empirical distributions of the different estimates is given by Figure A1.

Some of these findings of the baseline setting could have been anticipated without any simulation like the ones that (1) the non-linear ML estimator should be the best performing estimator given this setting, (2) that an estimator based on a sorted data density should be less precise than a comparable one based on a kernel density, or (3) that each semi-parametric approach should be less efficient than a fully parametrized one. However, this baseline setting offers two new main insights: In terms of the BIAS, the flexible semi-parametric *special regressor approach* does not perform better than both parametric linear estimators (2SLS and GI2SLS). Moreover, although the difference in the performance of the two parametric linear estimators is rather small - which can also be seen in Figure A1 sketching the different empirical distributions of the APEs for the case of N = 10.000 - the 2SLS estimator is relatively less biased while the GI2SLS estimator is relatively more efficient.

5.2.2 Special Regressor

In order to test the relevance of the large support requirement of the *special regressor* approach (cf. Subsection 4.1.4), the variance of the special regressor v is altered. While the standard deviations of v and $(y_1 - v)$ have been equal to 2 and 1.38 respectively in the baseline (cf. Subsection 5.1), Table 2 presents three different cases where the standard deviation of $(y_1 - v)$ is either approximately equal, lower or considerably higher than the standard deviation of v. The first panel shows the case when the standard deviations of v and $(y_1 - v)$ are both set to approximately 1.38. The second panel reflects the case when the standard deviation of v is set to 0.5 while the standard deviation of $(y_1 - v)$ is kept constant at 1.38. Finally, the last panel reports the case when the standard deviation of the special regressor considerably exceeds the one of the difference between the structural equation's regressand and the special regressor. To be more precise, the standard deviation of v is set to the value of 5 while the standard deviation of $(y_1 - v)$ is kept constant at the value of 1.38.[30]

Table 2: Simulation Results for APE of y_2 - Different Support for Special Regressor

	(1) 2SLS	(2) Generated Instr. 2SLS	(3) Recursive Biprobit	(4) Special Reg. Kernel Dens.	(5) Special Reg. Sorted Data	(6) Probit
$SD(v) \approx 1.38$ $SD(y_1 - v) \approx 1.38$			*true* $APE_{y_2} = 0.215$			
N = 10000						
BIAS	0.0066	0.0075	-0.0004	-0.0090	-0.0095	0.1171
RMSE	0.0134	0.0131	0.0099	0.0180	0.0230	0.1173
N = 5000						
BIAS	0.0074	0.0081	0.0005	-0.0075	-0.0091	0.1169
RMSE	0.0186	0.0178	0.0147	0.0218	0.0277	0.1173
N = 1000						
BIAS	0.0085	0.0102	0.0014	-0.0088	-0.0113	0.1169
RMSE	0.0397	0.0378	0.0331	0.0456	0.0570	0.1191
N = 500						
BIAS	0.0070	0.0083	-0.0001	-0.0091	-0.0151	0.1169
RMSE	0.0545	0.0512	0.0469	0.0637	0.0819	0.1214
N = 250						
BIAS	0.0052	0.0066	-0.0037	-0.0128	-0.0159	0.1177
RMSE	0.0756	0.0705	0.0661	0.0931	0.1104	0.1265
$SD(v) \approx 0.5$ $SD(y_1 - v) \approx 1.38$			*true* $APE_{y_2} = 0.285$			

-continued on next page-

[30] A relatively larger increase in the standard deviation of v is not meaningful since the variation of y_1 will be disproportionally affected.

- *Table 2 continued* -

N = 10000						
BIAS	0.0222	0.0248	-0.0002	-0.0520	-0.0489	0.1588
RMSE	0.0260	0.0279	0.0125	0.0820	0.0822	0.1591
N = 5000						
BIAS	0.0230	0.0256	0.0007	-0.0601	-0.0564	0.1590
RMSE	0.0297	0.0310	0.0170	0.0902	0.0907	0.1594
N = 1000						
BIAS	0.0223	0.0253	0.0004	-0.0646	-0.0649	0.1583
RMSE	0.0472	0.0463	0.0380	0.1057	0.1130	0.1605
N = 500						
BIAS	0.0222	0.0249	-0.0021	-0.0827	-0.0843	0.1574
RMSE	0.0634	0.0609	0.0552	0.1275	0.1348	0.1619
N = 250						
BIAS	0.0212	0.0247	-0.0043	-0.0910	-0.0911	0.1586
RMSE	0.0849	0.0814	0.0786	0.1426	0.1494	0.1674
$SD(v) \approx 5.0$ $SD(y_1 - v) \approx 1.38$			*true* $APE_{y_2} = 0.081$			
N = 10000						
BIAS	0.0001	0.0002	-0.0001	0.0014	0.0009	0.0431
RMSE	0.0104	0.0097	0.0062	0.0103	0.0124	0.0434
N = 5000						
BIAS	-0.0002	-0.0000	0.0001	0.0019	0.0011	0.0430
RMSE	0.0153	0.0138	0.0087	0.0146	0.0174	0.0435
N = 1000						
BIAS	0.0015	0.0012	-0.0005	0.0027	0.0002	0.0431
RMSE	0.0338	0.0307	0.0198	0.0325	0.0402	0.0453
N = 500						
BIAS	-0.0015	-0.0019	-0.0011	0.0020	0.0005	0.0432
RMSE	0.0478	0.0441	0.0291	0.0461	0.0558	0.0478
N = 250						
BIAS	-0.0028	-0.0019	-0.0016	-0.0005	-0.0050	0.0422
RMSE	0.0646	0.0598	0.0390	0.0724	0.0788	0.0502

Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $corr(u_1, u_2) = 0.6$, $F[\delta_{z_{21}}] \geq 100.00$; The BIAS is defined as the average difference between the estimated and the true APE. RMSE is the root mean squared error. Detailed statistics can be found in Table A2. A visual overview of the empirical distributions of the different estimates is given by Figure A2.

In contrast to the baseline setting which is defined by a standard deviation of v which is larger than the one used in the regressions of the first panel but smaller than the one of the regressions displayed in the third panel, the *recursive bivariate probit* estimator always has the smallest bias in the scenarios described by the first two panels. For the other estimators and the scenario depicted by the third panel of Table 2, there does not seem to exist a clear order in terms of the smallest bias, even for the different types of the *special regressor* approach. Only the naive probit estimator is always biased most. With respect to the RMSE, the order of the estimators remains the same as in the baseline with the exception that in the third scenario when the standard deviation of the special regressor largely exceeds the one of the difference between the structural equation's regressand and the special regressor, the relative rank of the *special regressor* kernel density estimator changes with the rank of the 2SLS estimator (for all but the smallest sample size).

In line with Bontemps and Nauges (2017), Table 2 shows the importance of the large support requirement in terms of the property of relatively unbiased and efficient estimates for the *special regressor* approach. The drop in the standard deviation of the special regressor as displayed in panel two leads to a relative increase of the bias of 18% in the largest sample, and the RMSE increases by up to the factor of fourteen in comparison to the case when the above mentioned standard deviations are of equivalent size.[31] On the opposite, an increase in the standard deviation of the special regressor in particular collapses the variation of the *special regressor* estimator.[32] However, this effect is absolutely modest in terms of the relative performance of the different estimators.

5.2.3 Weak Instruments

The relevance condition of any IV strategy ensures that the variation in the instruments sufficiently explains the exogenous part of the variation of the endogenous variable. If this condition is not fulfilled, the parameter estimates are biased and experience a large spread (Angrist & Pischke, 2009). For linear models it is widely accepted to assess this condition by the rule of thumb suggested by Staiger and Stock (1997). To be precise, a set of instruments is regarded as relevant in terms of "strong enough" if the F-statistic of the test of joint significance of the instruments' first stage coefficients exceeds the critical value of ten. However, as noted by Stock, Yogo, and Wright (2002), this critical value is only meaningful in some specific cases since it heavily depends on the number of instruments. Stock et al. propose two alternative definitions of weak instruments which have been used in the field of econometrics. Based on simulations, Stock and Yogo (2005) tabulate critical values which depend on certain parameters like the number of endogenous regressors and the number of instruments. In the case of a single instrument only their definition which is based on the size of a Wald test on the structural parameter of interest $\hat{\gamma} = \gamma_0$, where γ_0 is the true value, is applicable. It delivers a critical value of 16.38 for the F-statistic of the test of joint significance of the instruments' first stage coefficients when the share of wrong rejections is set to a maximum of 10% given a nominal level of 5%. For non-linear models, there does not exist any rule of thumb or formal definition when a set of instruments has to be considered as too weak to resolve endogeneity. We therefore follow the recommendation given by Nichols (2011) and use both critical value definitions mentioned above as benchmarks for non-linear models.[33]

The three different panels of Table 3 report the results of the scenarios when (1) the instrument's strength is not sufficient according to both critical values $[F[\delta_{z_{21}} = 0] \approx 5]$, (2) the instrument's strength is not sufficient according to the definition of Stock and Yogo (2005) but sufficient according to the rule of thumb of Staiger and Stock (1997) $[F[\delta_{z_{21}} = 0] \approx 10]$, and (3) the instruments' strength is sufficient according to both critical values $[F[\delta_{z_{21}} = 0] \approx 20]$ but not as strong as in the baseline where it exceeds the value of 100. The value of the F-statistic al-

[31] For the other estimators the RMSE just increases by the half of this factor at most.

[32] This reduction amounts up to 40% relative to the scenario displayed in the first panel.

[33] Jeffrey Wooldridge gives the same recommendation in the thread No. 361582 at www.statalist.org.

ways refers to the test of the intrument's power of the 2SLS estimator and the *special regressor* estimators which are similar due to the chosen DGP as outlined in footnote 27. Due to the specific three-step estimation procedure of the GI2SLS estimator, the power of the instrument has a self-reinforcing character for this estimator and therefore the F-statistic of the corresponding weak instrument test is always larger in comparison to the other ones.[34]

The first panel of Table 3 shows that in the case of weak instruments the 2SLS estimator as well as both *special regressor* estimators are heavily biased. The bias of the 2SLS estimator is actually up to 80%. Both *special regressor* estimators always underestimate by ca. 29%. The bias of the *recursive bivariate probit* approach varies "only" between 2% and 25%, which to some extent may be due to the fact that the assumed non-linearity helps in predicting the endogenous variable (cf. Wooldridge (2010)). However in terms of efficiency, the bivariate probit estimator as well as the GI2SLS estimator with its self-reinforced power of the instrument are poor performing.[35] In comparison to the baseline, the RMSE increases by the minimum factor of 5.3 and 2.2 respectively.

Table 3: Simulation Results for APE of y_2 - Weak instrument setting

	(1) 2SLS	(2) Generated Instr. 2SLS	(3) Recursive Biprobit	(4) Special Reg. Kernel Dens.	(5) Special Reg. Sorted Data	(6) Probit
$F[\delta_{z_{21}}] \approx 5$			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.1985	0.0090	0.0051	-0.0506	-0.0666	0.1724
RMSE	5.6721	0.0778	0.0389	0.3611	0.3998	0.1726
N = 5000						
BIAS	-0.7742	0.0104	0.0094	-0.0471	-0.0635	0.1722
RMSE	15.6767	0.1079	0.0601	0.3630	0.4009	0.1725
N = 1000						
BIAS	-0.0377	0.0166	0.0321	-0.0351	-0.0456	0.1711
RMSE	1.5652	0.2166	0.1290	0.3562	0.3906	0.1723
N = 500						
BIAS	0.1200	-0.0006	0.0425	-0.0285	-0.0427	0.1709
RMSE	3.2988	0.3550	0.1629	0.3583	0.3903	0.1736
N = 250						
BIAS	-0.0229	0.0070	0.0437	-0.0563	-0.0599	0.1686
RMSE	2.9437	0.5209	0.1894	0.3699	0.4025	0.1740
$F[\delta_{z_{21}}] \approx 10$			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	-0.0227	0.0090	0.0051	-0.0352	-0.0463	0.1723
RMSE	0.3526	0.0760	0.0384	0.2731	0.3141	0.1725
N = 5000						
BIAS	-0.0498	0.0098	0.0080	-0.0228	-0.0309	0.1720
RMSE	0.9749	0.1031	0.0552	0.2737	0.3117	0.1722

-continued on next page-

[34]For instance, in the case of the first panel, i.e. a F-statistic of five for the test using the estimates of the 2SLS and the *special regressor* estimator, the F-statistic for the same test using the estimates of the GI2SLS estimator varies between 8.59 (N=250) and 109.39 (N=10.000).

[35]Since the other estimators are clearly biased, it does not make sense to compare their performance in terms of the RMSE.

- Table 3 continued -

N = 1000						
BIAS	-0.0220	0.0137	0.0259	-0.0284	-0.0467	0.1701
RMSE	0.4344	0.1886	0.1202	0.2766	0.3218	0.1713
N = 500						
BIAS	-0.0039	0.0007	0.0283	-0.0227	-0.0412	0.1690
RMSE	0.6164	0.2483	0.1450	0.2779	0.3192	0.1717
N = 250						
BIAS	-0.0466	0.0092	0.0299	-0.0387	-0.0423	0.1655
RMSE	0.7158	0.2611	0.1712	0.2798	0.3122	0.1708
$F[\delta_{z_{21}}] \approx 20$			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	-0.0046	0.0089	0.0050	-0.0193	-0.0182	0.1721
RMSE	0.2041	0.0730	0.0379	0.2019	0.2371	0.1722
N = 5000						
BIAS	-0.0028	0.0088	0.0071	-0.0125	-0.0215	0.1717
RMSE	0.1974	0.0950	0.0519	0.2025	0.2447	0.1719
N = 1000						
BIAS	0.0029	0.0119	0.0159	-0.0173	-0.0250	0.1685
RMSE	0.2027	0.1493	0.1027	0.2028	0.2399	0.1697
N = 500						
BIAS	-0.0076	-0.0005	0.0114	-0.0159	-0.0221	0.1660
RMSE	0.1977	0.1706	0.1195	0.2043	0.2416	0.1688
N = 250						
BIAS	-0.0127	0.0017	0.0096	-0.0256	-0.0298	0.1592
RMSE	0.1988	0.1747	0.1385	0.2081	0.2391	0.1647

Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $corr(u_1, u_2) = 0.6$, $SD(v) = 2.0$; The BIAS is defined as the average difference between the estimated and the true APE. RMSE is the root mean squared error. Detailed statistics can be found in Table A3. A visual overview of the empirical distributions of the different estimates is given by Figure A3.

When the instrument's power is fixed at a value which is still regarded as the global critical value by many applied econometricians, i.e. the value of ten stemming from the rule of thumb of Staiger and Stock (1997), the performance of the 2SLS as well as both *special regressor* estimators improve while the performance of the GI2SLS, the bivariate probit as well as the naive probit estimator does not change considerably in comparison to the previous scenario. The *recursive bivariate probit* approach still exhibits the smallest RMSE for all different sample sizes and is the relatively less biased for larger samples. The 2SLS estimator is for some simulations relatively more or less biased than both *special regressor* estimators. However, for each sample size it has the largest spread. The spread of the 2SLS estimator is even larger than the one of the probit estimator which is expected to be biased.

In the scenario shown in the third panel of Table 3 the Stock and Yogo (2005) definition of the prevalence of strong instruments is met but the instrument is not as strong as in the baseline setting. For all estimators displayed in the first five columns the bias is moderate. The maximum amounts to 17% and the 2SLS estimator actually has the smallest bias for the majority of simulations which can be explained by the bouncing behavior of the bias as reported in the last column of Table A3. From the estimators taking into account the endogenity of y_2 the two types of the *special regressor* estimator exhibit always the largest bias. In terms of the RMSE the pattern is relatively more similar to the baseline in comparison to those described by the other panels of Table 3.

To sum up, by taking the definition of weak instruments from linear models as reference to assess the strength of the instrument in the model described by Equations 1 & 2, the simulations show that all estimators are reliant on a sufficiently strong instrument. If the instrument is assumed to be rather weak, the *recursive bivariate probit* as well as the GI2SLS estimator are the ones who are relatively best performing. The *special regressor* approach offers no advantage in the case of weak instruments. Even when the definition of strong instruments by Stock and Yogo (2005) is met, the estimates of the 2SLS and the *special regressor* estimators should be seen with caution because of their large spread.

5.2.4 Endogeneity Structure

In the following, the outcomes of different scenarios are discussed when different parameters that are setting the endogeneity structure in the DGP are altered. First, the effect of the degree of endogeneity on the performance of the estimators is investigated. Second, the impact of altering the joint CDF of the error terms (u_1, u_2) is described and finally it is analyzed how the distributional assumptions on the marginal CDFs of the error terms are influencing the ranking of the estimators in terms of a low bias as well as a low RMSE.

The two panels of Table 4 report the results of scenarios when the degree of endogeneity reflected by the Spearman correlation coefficient of the error terms is either relatively weak (first panel) or relatively strong (second panel). In the former case the bias is obviously low for all estimators even for the probit one. The bias of both linear parametric estimators is of similar magnitude as the one of the ML estimator while the bias of the two *special regressor* estimators is always larger in absolute values and negative, i.e. the *special regressor* estimators always underestimate the true effect as it was the case in the majority of the previously discussed scenarios. In terms of a low RMSE, the pattern is the same as in the baseline ($corr(u_1, u_2) = 0.6$), with the exception that the probit estimator exhibits the lowest RMSE until the case of the sample size of N = 5000.

In the simulations shown in the second panel of Table 4 the correlation coefficient of the error terms has been set to a value which is close to the one describing a perfect positive relationship. For all sample sizes except the smallest one, the *recursive bivariate probit* estimator is the relatively most unbiased while the naive probit estimator is the relatively most biased. Among the other estimators there does not exist a clear pattern which is due to the bouncing behaviour reflected by the fact that the bias does not shrink with the increase of the sample size. With respect to the spread of the estimates, the pattern remains the same as in the baseline with the *bivariate probit* estimator being the relative most efficient one. Moreover, the increase of the RMSE of each estimator in comparison to the baseline is absolutely modest.

In total, the simulations of Table 4 show that the *recursive bivariate probit* estimator is the most robust. In the case of a weak degree of endogeneity all estimators under investigation might be meaningful to use even the one ignoring endogeneity which would mean to accept a (small) bias but to gain precision.

Table 4: Simulation Results for APE of y_2 - Different degree of endogeneity

	(1) 2SLS	(2) Generated Instr. 2SLS	(3) Recursive Biprobit	(4) Special Reg. Kernel Dens.	(5) Special Reg. Sorted Data	(6) Probit
corr(u_1*,*u_2*) = 0.1*			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.0002	0.0002	0.0000	-0.0027	-0.0040	0.0156
RMSE	0.0117	0.0107	0.0096	0.0126	0.0166	0.0169
N = 5000						
BIAS	0.0000	0.0001	0.0001	-0.0020	-0.0036	0.0153
RMSE	0.0165	0.0151	0.0135	0.0181	0.0223	0.0178
N = 1000						
BIAS	0.0009	0.0017	0.0003	-0.0027	-0.0061	0.0153
RMSE	0.0366	0.0338	0.0299	0.0406	0.0502	0.0259
N = 500						
BIAS	-0.0002	-0.0012	-0.0017	-0.0020	-0.0080	0.0145
RMSE	0.0517	0.0486	0.0442	0.0566	0.0678	0.0335
N = 250						
BIAS	-0.0028	-0.0020	-0.0036	-0.0060	-0.0101	0.0140
RMSE	0.0712	0.0663	0.0628	0.0797	0.0972	0.0453
corr(u_1*,*u_2*) = 0.9*			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.0046	0.0058	-0.0002	0.0206	0.0194	0.1406
RMSE	0.0122	0.0119	0.0077	0.0272	0.0289	0.1408
N = 5000						
BIAS	0.0046	0.0058	-0.0000	-0.0011	-0.0029	0.1404
RMSE	0.0171	0.0165	0.0118	0.0193	0.0245	0.1407
N = 1000						
BIAS	0.0064	0.0082	0.0008	-0.0008	-0.0019	0.1414
RMSE	0.0377	0.0353	0.0257	0.0429	0.0534	0.1429
N = 500						
BIAS	0.0027	0.0043	-0.0012	-0.0023	-0.0048	0.1408
RMSE	0.0507	0.0478	0.0371	0.0583	0.0726	0.1436
N = 250						
BIAS	0.0001	0.0035	0.0029	-0.0066	-0.0122	0.1406
RMSE	0.0709	0.0647	0.0497	0.0843	0.1018	0.1462

Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$; The BIAS is defined as the average difference between the estimated and the true APE. RMSE is the root mean squared error. Detailed statistics can be found in Table A4. A visual overview of the empirical distributions of the different estimates is given by Figure A4.

Table 5 contains results from four different scenarios testing if the joint distribution of the error terms has an impact on the performance of the estimators.[36] As explained in Section 3, different copula functions are used to alternate the joint distribution. From the Archimedean copula class the Clayton, Frank and Gumbel copula are chosen (panel one to three of Table 5) while from the elliptical copula class the t copula (panel four) is used as an alternative to the

[36] The marginal distributions of the error terms are kept to be normal distributed as it was the case in the baseline setting, cf. Section 5.1. They are altered in the scenarios shown in Table 6.

normal copula of the baseline.[37] In principle, copula functions are fully flexible in defining joint distributions in comparison to assuming a joint normal CDF. The latter one is often assumed just due to the lack of information although it does not have to represent the reality. In particular in the field of finance different copula functions have been used to set up realistic dependence structures.[38]

For almost all scenarios shown in Table 5 the assessment of the estimators' performance with respect to the RMSE does not change from the one of the baseline. Only for the case of a joint error distribution realized by a t copula in large samples the *recursive bivariate probit* estimator has not the relatively lowest RMSE. In fact, both parametric linear estimators slightly perform better in such a scenario.

Table 5: Simulation Results for APE of y_2 - Different joint CDF

	(1) 2SLS	(2) Generated Instr. 2SLS	(3) Recursive Biprobit	(4) Special Reg. Kernel Dens.	(5) Special Reg. Sorted Data	(6) Probit
$\mathbb{C}_\theta$ = *Clayton*			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.0022	0.0023	0.0018	-0.0045	-0.0058	0.0974
RMSE	0.0112	0.0104	0.0090	0.0140	0.0175	0.0977
N = 5000						
BIAS	0.0014	0.0017	0.0016	-0.0030	-0.0040	0.0976
RMSE	0.0157	0.0144	0.0126	0.0184	0.0245	0.0980
N = 1000						
BIAS	0.0019	0.0022	0.0006	-0.0034	-0.0059	0.0973
RMSE	0.0383	0.0348	0.0290	0.0414	0.0510	0.0994
N = 500						
BIAS	0.0002	0.0011	-0.0011	-0.0024	-0.0063	0.0970
RMSE	0.0510	0.0479	0.0420	0.0565	0.0728	0.1016
N = 250						
BIAS	0.0036	0.0052	0.0011	-0.0011	-0.0028	0.0971
RMSE	0.0687	0.0658	0.0586	0.0807	0.0988	0.1055
$\mathbb{C}_\theta$ = *Frank*			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.0024	0.0025	0.0014	-0.0025	-0.0040	0.0958
RMSE	0.0111	0.0103	0.0086	0.0126	0.0161	0.0960
N = 5000						
BIAS	0.0018	0.0023	0.0012	-0.0021	-0.0040	0.0957
RMSE	0.0157	0.0145	0.0129	0.0182	0.0232	0.0961
N = 1000						
BIAS	0.0023	0.0028	0.0011	-0.0026	-0.0050	0.0954

-continued on next page-

[37]Visualizations of the realized (joint) error term distributions by the different copula functions are represented by Figures A7, A8, A9, A10 and A11.

[38]For instance, Longin and Solnik (2001) show that for describing the dependence structure of international equity markets, copula functions derived from the extreme value theory are better in describing the reality as multivariate normal distributions. In the same vein, Ang and Chen (2002) describe that the data speaks against a joint normal distribution of U.S. stock and U.S. aggregated market performance but that the dependence structure could be captured by copula functions. MacKenzie and Spears (2014) explain how the usage of the inadequately used normal copula contributed to the credit crisis in 2007. Low, Alcock, Faff, and Brailsford (2013) show that Clayton copulas are the best in describing downside correlations of returns for portfolios with 3 - 12 constituents. For a field not related to finance, Wu (2014) shows that asymmetric copulas help to explain car warranty claims.

- Table 5 continued -

RMSE	0.0381	0.0347	0.0294	0.0414	0.0517	0.0976
N = 500						
BIAS	0.0010	0.0023	-0.0006	-0.0017	-0.0058	0.0961
RMSE	0.0514	0.0483	0.0424	0.0566	0.0702	0.1004
N = 250						
BIAS	0.0030	0.0046	0.0016	-0.0019	-0.0073	0.0959
RMSE	0.0687	0.0653	0.0574	0.0801	0.0950	0.1043
$\mathbb{C}_\theta$ = *Gumbel*			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.0030	0.0037	0.0025	-0.0021	-0.0031	0.0912
RMSE	0.0113	0.0109	0.0091	0.0128	0.0166	0.0914
N = 5000						
BIAS	0.0029	0.0038	0.0026	-0.0016	-0.0030	0.0910
RMSE	0.0166	0.0154	0.0136	0.0190	0.0241	0.0915
N = 1000						
BIAS	0.0048	0.0053	0.0031	-0.0015	-0.0047	0.0913
RMSE	0.0351	0.0326	0.0287	0.0397	0.0493	0.0933
N = 500						
BIAS	0.0064	0.0063	0.0037	0.0030	-0.0015	0.0922
RMSE	0.0528	0.0484	0.0419	0.0584	0.0716	0.0966
N = 250						
BIAS	-0.0005	0.0028	0.0005	-0.0041	-0.0088	0.0912
RMSE	0.0694	0.0637	0.0573	0.0833	0.1015	0.0996
$\mathbb{C}_\theta$ = *t*			*true* $APE_{y_2} = 0.173$			
N = 10000						
BIAS	0.0032	0.0040	0.0107	-0.0028	-0.0037	0.0885
RMSE	0.0119	0.0113	0.0139	0.0130	0.0168	0.0888
N = 5000						
BIAS	0.0033	0.0042	0.0107	-0.0017	-0.0037	0.0883
RMSE	0.0162	0.0150	0.0166	0.0180	0.0230	0.0888
N = 1000						
BIAS	0.0047	0.0065	0.0112	-0.0015	-0.0035	0.0892
RMSE	0.0372	0.0353	0.0314	0.0418	0.0508	0.0916
N = 500						
BIAS	0.0018	0.0029	0.0073	-0.0027	-0.0086	0.0879
RMSE	0.0502	0.0478	0.0423	0.0573	0.0694	0.0928
N = 250						
BIAS	0.0002	0.0035	0.0059	-0.0066	-0.0090	0.0894
RMSE	0.0725	0.0665	0.0600	0.0826	0.0980	0.0989

Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $corr(u_1,u_2) = 0.6$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$; The BIAS is defined as the average difference between the estimated and the true APE. RMSE is the root mean squared error. Detailed statistics can be found in Table A5. A visual overview of the empirical distributions of the different estimates is given by Figure A5.

With respect to the property of an unbiased estimator, the *recursive bivariate probit* estimator strictly dominates the other estimators when the joint CDF of the error terms is set by a Frank copula, i.e. panel two of Table 5. For all other scenarios a bouncing behavior of the bias is visible for all estimators where the GI2SLS estimator always overestimates and both *special regressor* estimators always underestimate the true APE. In absolute terms the biases are rather comparable to the ones of the baseline setting and the maximum value for all scenarios shown in Table 5 amounts to a bias of 6.5%.

In general, Table 5 shows that all estimators discussed in Section 4 are quite robust to different joint distributions of (u_1, u_2). While this might be anticipated for the semi-parametric estimators (cf. Section 4.1.4), it is a key insight that even when the bivariate normality assumption of the *recursive bivariate probit* approach is not met, its estimator is the best performing among all tested ones and both parametric linear estimators are qualitatively good approximations.[39]

Instead of focusing on the impact of the joint CDF of the error terms, Table 6 presents the results when the baseline DGP's assumption on the normal marginal distribution of the error terms is altered. In the first panel the results of the scenario when both error terms follow a F-distribution with ten and six degrees of freedoms each are displayed. The second panel reflects the scenario when both u_1 and u_2 are logistically distributed with a scale of 0.9. In the third panel, the distribution of the error terms is set to follow a t-distribution with three degrees of freedoms. In all scenarios the joint CDF is kept to be normal distributed as it has been in the baseline.

Table 6: Simulation Results for APE of y_2 - Different marginal CDFs

	(1) 2SLS	(2) Generated Instr. 2SLS	(3) Recursive Biprobit	(4) Special Reg. Kernel Dens.	(5) Special Reg. Sorted Data	(6) Probit
$u_1 \sim F(10,6)$ $u_2 \sim F(10,6)$			*true* $APE_{y_2} = 0.172$			
N = 10000						
BIAS	0.0093	0.0099	0.0055	0.0360	0.0346	0.0906
RMSE	0.0149	0.0145	0.0109	0.0601	0.0575	0.0908
N = 5000						
BIAS	0.0092	0.0097	0.0056	0.0380	0.0364	0.0902
RMSE	0.0186	0.0176	0.0145	0.0680	0.0675	0.0907
N = 1000						
BIAS	0.0109	0.0116	0.0053	0.0387	0.0349	0.0910
RMSE	0.0373	0.0345	0.0298	0.0886	0.0979	0.0935
N = 500						
BIAS	0.0072	0.0067	0.0013	0.0185	0.0142	0.0878
RMSE	0.0533	0.0485	0.0432	0.1047	0.1096	0.0930
N = 250						
BIAS	0.0061	0.0079	0.0015	0.0098	0.0008	0.0896
RMSE	0.0693	0.0642	0.0588	0.1151	0.1320	0.0990
$u_1 \sim log(0,0.9)$ $u_2 \sim log(0,0.9)$			*true* $APE_{y_2} = 0.154$			
N = 10000						
BIAS	0.0064	0.0071	-0.0011	-0.0005	-0.0010	0.1675
RMSE	0.0160	0.0158	0.0131	0.0246	0.0282	0.1676
N = 5000						
BIAS	0.0057	0.0066	-0.0016	-0.0016	-0.0021	0.1672
RMSE	0.0221	0.0215	0.0192	0.0328	0.0385	0.1675
N = 1000						
BIAS	0.0077	0.0088	-0.0002	-0.0016	-0.0024	0.1673
RMSE	0.0474	0.0457	0.0413	0.0630	0.0771	0.1689

-continued on next page-

[39] A most recent study by Han and Lee (2019) solely focusing on a ML estimation in the class of the *recursive bivariate probit* approach comes to the same result, i.e. that this estimator is quite robust to misspecification with respect to different joint distributions of the error terms.

- Table 6 continued -

N = 500						
BIAS	0.0028	0.0036	-0.0043	-0.0063	-0.0093	0.1664
RMSE	0.0662	0.0638	0.0594	0.0842	0.1029	0.1698
N = 250						
BIAS	0.0016	0.0031	-0.0060	-0.0088	-0.0156	0.1675
RMSE	0.0931	0.0899	0.0860	0.1137	0.1327	0.1743
$u_1 \sim t(3)$ $u_2 \sim t(3)$			*true* $APE_{y_2} = 0.161$			
N = 10000						
BIAS	0.0056	0.0063	-0.0026	-0.0008	-0.0024	0.1387
RMSE	0.0143	0.0138	0.0116	0.0285	0.0332	0.1389
N = 5000						
BIAS	0.0056	0.0062	-0.0028	-0.0018	-0.0027	0.1380
RMSE	0.0199	0.0189	0.0171	0.0454	0.0465	0.1384
N = 1000						
BIAS	0.0064	0.0079	-0.0023	-0.0018	-0.0054	0.1383
RMSE	0.0430	0.0407	0.0372	0.0612	0.0769	0.1401
N = 500						
BIAS	0.0034	0.0042	-0.0054	-0.0068	-0.0077	0.1370
RMSE	0.0595	0.0560	0.0525	0.0822	0.0951	0.1407
N = 250						
BIAS	0.0000	0.0016	-0.0082	-0.0100	-0.0156	0.1380
RMSE	0.0831	0.0786	0.0761	0.1088	0.1317	0.1460

Scenario characteristics: $\mathbb{C}_\theta = Gaussian$, $corr(u_1, u_2) = 0.6$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$; The BIAS is defined as the average difference between the estimated and the true APE. RMSE is the root mean squared error. Detailed statistics can be found in Table A6. A visual overview of the empirical distributions of the different estimates is given by Figure A6.

When the error terms follow a F-distribution which in contrast to the previously used marginal CDFs has no symmetric character but is positively skewed, the performance pattern of the estimators remains the same as in the baseline, i.e. the *recursive bivariate probit* estimator exhibits the smallest RMSE but also the smallest bias.[40] In terms of the latter measure, it is followed by the 2SLS estimator while in terms of a small RMSE the GI2SLS estimator is at second rank. Both non-linear semi-parametric estimators are dominated by the linear parametric ones for both the bias and the RMSE with the exemption of the smallest sample size. In comparison to the baseline setting, the absolute bias as well as the RMSE increases almost for each sample size and estimator combination displayed in the columns one to five of Table 6. The best performing estimator, i.e. the *bivariate probit* one, is biased by 3.2% and the RMSE is 20% larger than in the baseline for the largest sample.

In the case of u_1 and u_2 being both logistically distributed with a scale of 0.9, i.e. the marginal distributions of the error terms exhibit a relatively larger spread with relatively larger tails in comparison to a normal distribution, the performance of the estimators in terms of a low RMSE remains exactly the same as in the previous scenario or the baseline setting. For all sample sizes the non-linear parametric estimator dominates the linear parametric ones which in

[40] It should be noted that the correlation of both errors has to be slightly reduced to a value of 0.48 in order to use this type of marginal distribution for the error terms while parallelly maintaining the large support requirement of the *special regressor*.

turn dominate the non-linear semi-parametric ones. In comparison to the baseline, the RMSE relatively increases for all estimators but in particular for the *recursive bivariate probit* estimator which is meaningful since the assumptions made by this parametric estimator are not (fully) met anymore. With respect to the bias, there does not exists a clear pattern since the ones of the estimators displayed in the first three columns exhibit a bouncing behavior. On the opposite, the *special regressor* estimators are again always downwards biased but the bias shrinks consistently with an increase in sample size.

In the last panel of Table 6 the results of simulations where the marginal CDFs of u_1 and u_2 have been altered from normal to t-distributions (with three degrees of freedom) are displayed.[41] Also in this scenario where the error term distributions exhibit a relatively more dense distribution but which is still relatively comparable to the baseline one, the *recursive bivariate probit* estimator is the best performing in terms of a low RMSE. Moreover, the pattern of the estimators in terms of a small RMSE does not change in comparison to the baseline, and the relative increase in the RMSE amounts on average to a value of 42.1% for all sample size estimator combinations which is mainly driven by both *special regressor* estimators. In terms of the bias, the linear as well as the non-linear parametric estimators are prone to bouncing behavior while the *special regressor* estimators are again relatively downward biased but the bias shrinks with the increase of the sample size.

In summary, Table 6 shows that even when assumptions on the marginal distributions of the error terms raised by the fully parametric ML approach are not met, none of the other tested estimators, like the semi-parametric ones, is able to deliver quantitatively better results of the true APE. Similar to the case of the non-normality of the joint CDF of the error terms (cf. Table 5) the *recursive bivariate probit* estimator seems to be surprisingly robust to misspecification of the marginal CDFs of the error terms regarding the measure of the RMSE.[42]

6 Empirical Illustration

As an empirical illustration, we revisit the prominent study of Angrist and Evans (1998) where a BRM with a dichotomous EEV is estimated by estimators suited for linear models. Angrist and Evans study the determinants of labor supply for women with data from the US Census 1980. Among others, they explicitly model the binary decision of having worked for pay in the time when the survey was conducted and investigate the effect of having more than two children on this individual outcome. Since the parenthood measure is endogenous they use two different types of instruments in a 2SLS estimation ignoring the non-linearity of the dependent variable. First, they use the information if the first two born children share the same sex to construct a binary instrument. Second and alternatively, the authors use the similar but slightly different

[41] As in the case when the error values are sampled from a F-distribution, the correlation between u_1 and u_2 is slightly reduced to 0.54 in order to fulfill the large support requirement of the special regressor.

[42] Han and Lee (2019) show in their recent study that given certain assumptions hold their proposed semi-parametric estimator can outperform the classical *recursive bivariate probit* approach in the case when the marginal distributions of the error terms are set by a mixture of two normal distributions.

Table 7: Estimates of Labor Supply Models using 1980 Census data

	All women							Married women						
	(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)
Estimation method	OLS	2SLS	BP	SRK	2SLS	BP	SRK	OLS	2SLS	BP	SRK	2SLS	BP	SRK
Instrument for *More than 2 children*	—	Same sex			Two boys, Two girls			—	Same sex			Two boys, Two girls		
Original sample														
Dependent Variable:														
Worked for pay	-0.176	-0.117			-0.110			-0.167	-0.117			-0.109		
	(0.002)	(0.025)			(0.025)			(0.002)	(0.028)			(0.028)		
Observations	394840	394840			394840			254652	254652			254652		
Reduced sample														
Dependent Variable:														
Worked for pay	-0.170	-0.110	-0.228	-0.137	-0.101	-0.228	-0.140	-0.168	-0.144	-0.193	0.246	-0.137	-0.193	0.251
	(0.004)	(0.066)	(0.053)	(0.128)	(0.066)	(0.053)	(0.123)	(0.004)	(0.058)	(0.048)	(0.099)	(0.057)	(0.048)	(0.105)
Observations	65000	65000	65000	65000	65000	65000	65000	65000	65000	65000	65000	65000	65000	65000

BP = Recursive Bivariate Probit; SRK = Special Regressor Kernel Density

Note: The first panel, i.e. the original sample, is a replication of the first row of Table 7 of Angrist and Evans (1998). Numbers differ from the published text version but are consistent with the published log-file. The second panel, i.e. the reduced sample, consists of a random selection of 65.000 observations each for both original samples. For all estimations the displayed estimates are equivalent to the APE and the set of covariates consists of controls for *individual age*, *individual age at first birth* and indicators for *First-born boy*, *Second-born boy*, *Black*, *Hispanic* and *other race*. For the *special regressor* estimations the negative demeaned age of the women in the sample was taken as the special regressor. Standard errors of the estimated APEs of the *special regressor* approach are obtained by bootstrapping with 49 bootstrap samples.

information if the two first born are two boys or two girls (cf. Angrist and Evans (1998, Table 7)). Both identification strategies are applied on two different samples, consisting of women with (1) different relationship statuses and (2) just married women.

Instead of using the full sample of 394.840 and 254.652 observations respectively, we shrink the sample sizes to 65.000 randomly picked observations each due to computation issues for the *special regressor* approach. However, the comparison of both panels of Table 7 - which reports the original estimates and those for the randomly reduced sample size - shows that the difference in the estimates is considerably small.[43] Besides presenting the estimates of the original OLS as well as 2SLS estimations, Table 7 also provides estimates from estimating the same model with the *recursive bivariate probit* approach as well as the *special regressor kernel density* approach for the reduced sample. For the latter estimation approach the negative demeaned age of the women in the sample was taken as the special regressor.[44]

There are four main insights which can be concluded from Table 7. First, in the *All women* sample the effect of having more than two children estimated by the *recursive bivariate probit* approach is quite different to the one of the 2SLS estimation approach for both instruments estimations. While the linear parametric estimator suggests a negative effect of ten or eleven percentage points respectively, the non-linear parametric estimator reveals an estimate of close to minus twenty-three percentage points. For the *Married women* sample the difference between those estimators is smaller, i.e. the estimates of the 2SLS estimations are ca. 25% smaller than

[43] To be more precise, the difference in the estimates for the *All women* sample amounts to 0.009 at maximum for all different estimators, while it is 0.028 at maximum for the *Married women* sample.

[44] The variable of individual age has been taken as special regressor in the majority of previous studies, for instance Dong and Lewbel (2015).

the ones of the *recursive bivariate probit* approach.[45]

Second, the ML estimator always has the smallest standard error. However, the standard errors of the 2SLS estimations are not much bigger. For the estimations of the *All women* sample the APE estimated by the 2SLS estimator looses its significance in comparison to the full sample. However, this fact only partly points to the property of the *recursive bivariate probit* estimator to be the most efficient one given the considerable difference in the size of the coefficients of both estimators.

Third, for the sample of *All women* the *special regresssor* estimates shown in column (2.3) and (3.3) suggest a negative effect of fourteen percentage points. This is in the middle of the estimates of the other approaches. The fact that the *special regressor* estimator delivers relatively more similar estimates to the 2SLS approach in comparison to the ML approach fits to the simulation results of Section 5. For the sample of *Married women* the *special regressor* approach suggests a considerable positive effect of having at least three children. This is in contrast to the other estimators which still reveal a negative effect of similar magnitude as they did before. That might be explained by the relatively large spread of this estimator (cf. Section 5).

In total, this empirical illustration shows that when using real data the different approaches can deliver considerably different estimates and thus that there is need for some guidance which estimator should be used in which scenario as it has been done in the previous section. For the key aspects like the relative size of the standard errors, this empirical illustration is in line with findings of the simulations.

7 Conclusion

This paper has compared six different estimators of the classes of linear parametric (2SLS, GI2SLS), non-linear parametric (*recursive bivariate probit*) as well non-linear semi-parametric approaches in order to estimate the effect of a binary EEV in a BRM. Regarding the last one, i.e. the *special regressor* approach, it has also offered a guidance how the ASF can be used to estimate APEs in such a setting. While in theory there are advantages as well as disadvantages for each of the different estimators, the practical review in terms of simulations shows the dominance of one specific estimator. In particular with regards to efficiency, the simulation study shows that in all tested scenarios where the effect of the spread of the *special regressor*, the prevalence of weak instruments, the degreee of endogeneity and the assumed joint as well as marginal distributions of the error terms of the model has been investigated respectively, the ML estimator of the *recursive bivariate probit* approach performs relatively best in comparison to two 2SLS as well as two *special regressor* approaches. This finding which is based on the comparison of APEs is

[45] To rule out that the considerable difference in the estimates of the 2SLS and the *recursive bivariate probit* approach is driven by the process of the random sample reduction, we also estimated all sample specification combinations with the *recursive bivariate probit* approach for the original sample. For the *All women* sample the point estimate differs only by 0.17 to those displayed in Table 7 while for the *Married women* sample it differs by 2.6.

confirmed by an empirical illustration of the effect of parenthood on the decision to supply labor.

Taking up the findings, this paper can be seen as a strong recommendation to primarily rely on the ML estimator of the *recursive bivariate probit* approach in a setting of a binary dependent as well as a binary endogenous variable. In particular since the estimation of non-linear models has become relatively easy meanwhile in terms of the computational dimension, this method should not be ignored any longer by the argument that linear models could deliver adequate approximations. In fact, this study shows that the fully parametrized ML approach delivers relatively better estimates in each tested scenario and outperforms the other estimators by its approximations for the true APE even in settings where its assumptions are not fulfilled. In the case that the ML estimator of the *recursive bivariate probit* approach is not an option, for instance due to its iterative search on a globally flat likelihood function, the proposed alternative 2SLS estimator (GI2SLS) could be an attractive alternative in particular due to its reinforcing character in terms of the instrument's strength.

Although this study does not provide any obvious support to promote the semi-parametric *special regressor* approach, it would be questionable to disregard this method completely. In fact, a closer look on the results of the *special regressor* approach yields an interesting insight. Both types of the *special regressor* approach are able to deliver relatively better estimates of the coefficients of the structural equation in comparison their estimates of the APEs. This in particular true when the errors of the equations describing the model are not marginal normally distributed. In other words, these semi-parametric approaches lose in performance power when the logistic series estimator is used to calculate the ASF for the estimation of APEs which can be interpreted as reflecting causal effects. This circumstance fits to the observation which can be retraced by the tabulated results in the appendix, that although the usage of the AIF (instead of the ASF) leads to biased estimates in theory, it sometimes outperforms the ASF in practice in terms of the property of relatively unbiased and efficient estimates of the APEs. Moreover, it always clearly outperforms the ASF in terms of computation time.

Future research could concentrate on three different issues. First, a reliable test on the strength of the instruments in non-linear models could push the usage of non-linear models forward when endogeneity is present. At the moment, the insights from linear models serve as reference. However, these benchmarks could be quite misleading since estimation assumptions on non-linearity typically strengthen the prediction power on the endogenous regressor itself. Second, research focusing on the improvement of linear models could explore to which extent the usage of the alternative 2SLS estimator allows the usage of relatively weaker instruments. Third, research on the improvement of semi-parametric or non-parametric estimation procedures could help to close the performance gap of the *special regressor* approach to the other estimators and could make its advantages, such as the possibility of an unlimited number of endogenous regressors, more attractive in complex models.

References

Altonji, J., Elder, T., & Taber, C. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, *113*(1), 151-184.

Amemiya, T. (1978). The Estimation of a Simultaneous Equation Generalized Probit Model. *Econometrica*, *46*(5), 1193–1205.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, *63*(3), 443 - 494.

Angrist, J. (1990). Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records. *The American Economic Review*, *80*(3), 313–336.

Angrist, J., Bettinger, E., Bloom, E., King, E., & Kremer, M. (2002). Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment . *American Economic Review*, *92*(5), 1535-1558.

Angrist, J., & Evans, W. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review*, *88*(3), 450–477.

Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics*, *28*(4), 773-824.

Angrist, J., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Baum, C. (2012). *SSPECIALREG: Stata module to estimate binary choice model with discrete endogenous regressor via special regressor method.* Statistical Software Components, Boston College Department of Economics.

Blundell, R., & Powell, J. (2003). Endogeneity in Nonparametric and Semiparametric Regression Models. In M. Dewatripont, L. Hansen, & S. Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress* (Vol. 2, p. 312-357). Cambridge University Press.

Blundell, R., & Powell, J. (2004). Endogeneity in Semiparametric Binary Response Models. *The Review of Economic Studies*, *71*(3), 655–679.

Bontemps, C., & Nauges, C. (2017). *Endogenous Variables in Binary Choice Models: Some Insights for Practitioners.*

Cameron, A., & Trivedi, P. (2005). *Microeconometrics.* Cambridge University Press.

Chesher, A., & Rosen, A. (2013). What Do Instrumental Variable Models Deliver with Discrete Dependent Variables? *American Economic Review*, *103*(3), 557-62.

Conley, D., & Heerwig, J. (2011). The War at Home: Effects of Vietnam-Era Military Service on Postwar Household Stability. *American Economic Review*, *101*(3), 350-54.

Dong, Y., & Lewbel, A. (2015). A Simple Estimator for Binary Choice Models with Endogenous Regressors. *Econometric Reviews*, *34*(1-2), 82-105.

Evans, W., Farrelly, M., & Montgomery, E. (1999). Do Workplace Smoking Bans Reduce Smoking? *American Economic Review*, *89*(4), 728-747.

Evans, W., & Schwab, R. (1995). Finishing High School and Starting College: Do Catholic Schools Make a Difference? *The Quarterly Journal of Economics*, *110*(4), 941-974.

Farbmacher, H., Guber, R., & Vikström, J. (2018). Increasing the credibility of the twin birth instrument. *Journal of Applied Econometrics*, *33*(3), 457-472.

Greene, W. (2017). *Econometric Analysis* (8th ed.). Pearson Education.

Han, S., & Lee, S. (2019). Estimation in a generalization of bivariate probit models with dummy endogenous regressors. *Journal of Applied Econometrics*, *34*(6), 994-1015.

Hausman, J. (1975). An Instrumental Variable Approach to Full Information Estimators for Linear and Certain Nonlinear Econometric Models. *Econometrica*, *43*(4), 727–738.

Heckman, J. (1978). Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica*, *46*(4), 931–959.

Horrace, W., & Oaxaca, R. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, *90*(3), 321 - 327.

Imbens, G., & Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475.

Islam, A., & Raschky, P. (2015). Genetic distance, immigrants' identity, and labor market outcomes. *Journal of Population Economics*, *28*(3), 845–868.

Kang, C., & Lee, M.-J. (2014). Estimation of Binary Response Models With Endogenous Regressors. *Pacific Economic Review*, *19*(4), 502-530.

Klein, R., & Spady, R. (1993). An Efficient Semiparametric Estimator for Binary Response Models. *Econometrica*, *61*(2), 387-421.

Lee, Y.-Y., & Li, H.-H. (2018). Partial effects in binary response models using a special regressor. *Economics Letters*, *169*, 15 - 19.

Lewbel, A. (2000). Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, *97*(1), 145 - 177.

Lewbel, A. (2007). Endogenous selection or treatment model estimation. *Journal of Econometrics*, *141*(2), 777 - 806.

Lewbel, A., Dong, Y., & Yang, T. (2012). Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics/Revue canadienne d'économique*, *45*(3), 809-829.

Lin, W., & Wooldridge, J. (2015a). *Estimating Binary Response Models with Endogenous Explanatory Variables, Using Control Functions in Quasi-LIML.*

Lin, W., & Wooldridge, J. (2015b). On different approaches to obtaining partial effects in binary response models with endogenous regressors. *Economics Letters*, *134*, 58 - 61.

Lin, W., & Wooldridge, J. (2018). *Binary and Fractional Response Models with Continuous and Binary Endogenous Explanatory Variables.*

Longin, F., & Solnik, B. (2001). Extreme Correlation of International Equity Markets. *The Journal of Finance*, *56*(2), 649-676.

Low, R., Alcock, J., Faff, R., & Brailsford, T. (2013). Canonical vine copulas in the context of modern portfolio management: Are they worth it? *Journal of Banking & Finance*, *37*(8), 3085 - 3099.

MacKenzie, D., & Spears, T. (2014). 'The formula that killed Wall Street': The Gaussian copula and modelling practices in investment banking. *Social Studies of Science*, *44*(3), 393–417.

Magnac, T., & Maurin, E. (2007). Identification and information in monotone binary models. *Journal of Econometrics*, *139*(1), 76 - 104.

Mu, B., & Zhang, Z. (2018). Identification and estimation of heteroscedastic binary choice models with endogenous dummy regressors. *The Econometrics Journal*, *21*(2), 218-246.

Nelsen, R. (2006). *An Introduction to Copulas.* Springer.

Newey, W. (1990). Semiparametric Efficiency Bounds. *Journal of Applied Econometrics*, *5*(2), 99–135.

Nichols, A. (2011). *Causal inference for binary regression with observational data* (CHI11 Stata Conference No. 6). Stata Users Group.

Rivers, D., & Vuong, Q. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics*, *39*(3), 347 - 366.

Rosenbaum, P., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*(1), 41–55.

Sasaki, M. (2002). The Causal Effect of Family Structure on Labor Force Participation among Japanese Married Women. *The Journal of Human Resources*, *37*(2), 429–440.

Staiger, D., & Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586.

Stock, J., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In D. Andrews (Ed.), *Identification and Inference for Econometric Models* (p. 80-108). New

York: Cambridge University Press.
Stock, J., Yogo, M., & Wright, J. (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics*, *20*, 518 – 529.
Terza, J., Basu, A., & Rathouz, P. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531 - 543.
Vytlacil, E., & Yildiz, N. (2007). Dummy Endogenous Variables in Weakly Separable Models. *Econometrica*, *75*(3), 757-779.
Wan, F., Small, D., & Mitra, N. (2018). A general approach to evaluating the bias of 2-stage instrumental variable estimators. *Statistics in Medicine*, *37*(12), 1997-2015.
Wooldridge, J. (2005). Unobserved Heterogeneity and Estimation of Average Partial Effects. In D. Andrews & J. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (p. 27-55). Cambridge University Press.
Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). The MIT Press.
Wright, P. (1928). *The Tariff on Animal and Vegetable Oils.* Macmillan Company.
Wu, S. (2014). Construction of asymmetric copulas and its application in two-dimensional reliability modelling. *European Journal of Operational Research*, *238*(2), 476 - 485.
Yildiz, N. (2013). Estimation of binary choice models with linear index and dummy endogenous variables. *Econometric Theory*, *29*(2), 354-392.

Appendix

Derivations:

Derivation 1

The following derivation is a modified version of the one of Dong and Lewbel (2015).

Theorem 1. *Assume* $y_1 = 1[\gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + v + u_1]$, $\mathbb{E}(\mathbf{z}_2 u_1) = 0$, *and* $v = g(y_2, \mathbf{z}, w)$. *Assume* $supp(\gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + u_1) \subseteq supp(-v)$, $\mathbb{E}(v) = 0$, *g is differentiable and strictly monotonically increasing in its last element,* $w \perp (y_2, \mathbf{z}, u_1)$ *and w is continuously distributed. Let f(w) be the probability density function of w. Let* $M(v)$ *be any mean zero distribution function on supp(v) such that* $M(v_0) = 0$ *and* $M(v_1) = 1$ *for some points* v_0 *and* v_1 *that are in the interior of supp(v).*

Define t by

$$t = \frac{y_1 - M(v)}{f(w)} \frac{\partial g(y_2, \mathbf{z}, w)}{\partial w} \tag{15}$$

Then $t = \gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + \tilde{u}_1$, *where* $\mathbb{E}(\mathbf{z}_2 \tilde{u}_1) = 0$

Proof. Define $\breve{y}_1 = \gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + u_1$ and let $M(v) = 1[v \leq 0]$ for simplicity. Then by the definition of the conditional expectation it follows:

$$\begin{aligned}
\mathbb{E}(t|y_2, \mathbf{z}, u_1) &= \int_{supp(w|y_2,\mathbf{z},u_1)} \frac{1[\breve{y}_1 + g(y_2, \mathbf{z}, w) \geq 0] - 1[g(y_2, \mathbf{z}, w) \geq 0]}{f(w)} \\
&\quad * \frac{\partial g(y_2, \mathbf{z}, w)}{\partial w} \, f(w|y_2, \mathbf{z}, u_1)\, dw \\
&= \int_{supp(w|y_2,\mathbf{z},u_1)} \big[1[\breve{y}_1 + g(y_2, \mathbf{z}, w) \geq 0] - 1[g(y_2, \mathbf{z}, w) \geq 0]\big] \, \frac{\partial g(y_2, \mathbf{z}, w)}{\partial w} \, dw \\
&= \int_{supp(v|y_2,\mathbf{z},u_1)} \big[1[\breve{y}_1 + v \geq 0] - 1[v \geq 0]\big] \, dv
\end{aligned}$$

The second equality follows from $w \perp (y_2, \mathbf{z}, u_1)$ or expressed differently as $f(w) = f(w|y_2, \mathbf{z}, u_1)$. The third equality uses a change of variables from w to v.

If $\breve{y}_1 \geq 0$, then

$$\mathbb{E}(t|y_2, \mathbf{z}, u_1) = \int_{supp(v|y_2,\mathbf{z},u_1)} 1[-\breve{y}_1 \leq v \leq 0] \, dv = \int_{-\breve{y}_1}^{0} 1 \, dv = \breve{y}_1$$

and if $\breve{y}_1 \leq 0$, then

$$\mathbb{E}(t|y_2, \mathbf{z}, u_1) = \int_{supp(v|y_2,\mathbf{z},u_1)} -1[0 \leq v \leq -\breve{y}_1] \, dv = -\int_{0}^{-\breve{y}_1} 1 \, dv = \breve{y}_1$$

which both together proves that $\mathbb{E}(t|y_2, \mathbf{z}, u_1) = \gamma y_2 + \mathbf{z}_1\boldsymbol{\beta} + u_1$.

Defining $\tilde{u}_1 = t - \mathbf{z}_1\boldsymbol{\beta} - \gamma y_2$, it follows

$$\begin{aligned}\mathbb{E}[\mathbf{z}_2\tilde{u}_1] &= \mathbb{E}[\mathbf{z}_2(t - \mathbf{z}_1\boldsymbol{\beta} - \gamma y_2)] = \mathbb{E}\big[\mathbb{E}[\mathbf{z}_2(t - \mathbf{z}_1\boldsymbol{\beta} - \gamma y_2)] \,|\, y_2, \mathbf{z}, u_1\big] \\ &= \mathbb{E}[\mathbf{z}_2(\mathbb{E}(t|y_2, \mathbf{z}, u_1) - \mathbf{z}_1\boldsymbol{\beta} - \gamma y_2)] = \mathbb{E}(\mathbf{z}_2 u_1) = 0\end{aligned}$$

To show that the theorem holds for other choices of $M(v)$, replace $y_1 - M(v)$ of Equation 15 with $[y_1 - 1(v \geq 0)] + [1(v \geq 0) - M(v)]$. Then $\mathbb{E}(t|y_2, \mathbf{z}, u_1)$ equals the sum of the term given above and $\int_{supp(w|y_2,\mathbf{z},u_1)}[1(v \geq 0) - M(v)]\, dv$. Applying an integration by parts to this term gives

$$[1(v \geq 0) - M(v)]v|_{supp(v|y_2,\mathbf{z},u_1)} - \int_{supp(v|y_2,\mathbf{z},u_1)} -\frac{\partial M(v)}{\partial v} v\, dv$$

The first term of this expression is zero because $M(v)$ is a distribution function that equals zero and one strictly inside the support of v, and the second term is zero because $M(v)$ is a mean zero distribution function. So $\mathbb{E}(t|y_2, \mathbf{z}, u_1)$ is unchanged by replacing $1[v \geq 0]$ with $M(v)$.

□

Derivation 2

The following derivation is a detailed version of the one of Lewbel et al. (2012).

Given estimates of the coefficients of the structural equation $\boldsymbol{\psi}$, the AIF approach suggests to condition on the index of y_{1i}, i.e. to use Equation 9 instead of $\mathbb{E}(y_{1i}|\mathbf{x}_i)$.

Hence, an estimator is needed for this expression which is called M in the following with m being its derivative, i.e. $\hat{M}_i = \hat{\mathbb{E}}(y_{1i}|\mathbf{x}_i\boldsymbol{\psi})$. Partial effects can then be estimated analogous to Equation 7 by $\hat{m}\gamma$.

Lewbel et al. (2012) propose to estimate a standard one-dimensional kernel regression of y_1 on $x\hat{\boldsymbol{\psi}}$ to receive an estimate of M:

$$\hat{M}_i = \frac{\sum_{j=1}^n y_{1j}\, K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\sum_{j=1}^n K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)} \quad \text{for } \ i = 1, \dots, n$$

where K is a standard one-dimensional kernel function and h denotes a bandwidth.[46]

The derivative $\hat{m}_i = \frac{\partial \hat{M}_i}{\partial(\mathbf{x}_i\hat{\boldsymbol{\psi}})}$ can be derived by applying a combination of the quotient and chain rule:

$$\hat{m}_i = \frac{\partial \hat{M}_i}{\partial(\mathbf{x}_i\hat{\boldsymbol{\psi}})} = \frac{\sum_{j=1}^n y_{1j}\, K'\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big) * \frac{1}{h} * \sum_{j=1}^n K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\Big[\sum_{j=1}^n K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)\Big]^2} - \frac{\sum_{j=1}^n y_{1j}\, K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big) * \sum_{j=1}^n K'\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big) * \frac{1}{h}}{\Big[\sum_{j=1}^n K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)\Big]^2}$$

[46] $\hat{M}_i$ equals the propensity score used in the Klein and Spady (1993) estimator.

Define for ease of notation $\omega = \sum_{j=1}^{n} K(\cdot)$ in the following which is constant. By rearranging the expression from above on gets

$$= \frac{\frac{1}{h}\sum_{j=1}^{n} y_{1j}\, K'\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\omega} - \frac{\frac{1}{h}\sum_{j=1}^{n} \frac{Y_{1j}\, K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\omega} * \sum_{j=1}^{n} K'\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\omega}$$

$$= \frac{\frac{1}{h}\sum_{j=1}^{n} y_{1j}\, K'\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\omega} - \frac{\frac{1}{h}\hat{M} * \sum_{j=1}^{n} K'\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\omega}$$

where $\hat{M}$ is also constant. Thus, by plugging in ω back again and rearranging the whole expression is equal to:

$$\hat{m}_i = \frac{\partial \hat{M}_i}{\partial(\mathbf{x}_i\hat{\boldsymbol{\psi}})} = \frac{\frac{1}{h}\sum_{j=1}^{n}(y_{1j} - \hat{M}_j)\, K'\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}{\sum_{j=1}^{n} K\big(\frac{(\mathbf{x}_i\hat{\boldsymbol{\psi}})-(\mathbf{x}_j\hat{\boldsymbol{\psi}})}{h}\big)}$$

which multiplied by the respective coefficient of $x \in \mathbf{x}$ gives the partial effect of x for individual i.

Figures:

Distribution of APEs

Figure A1: Baseline

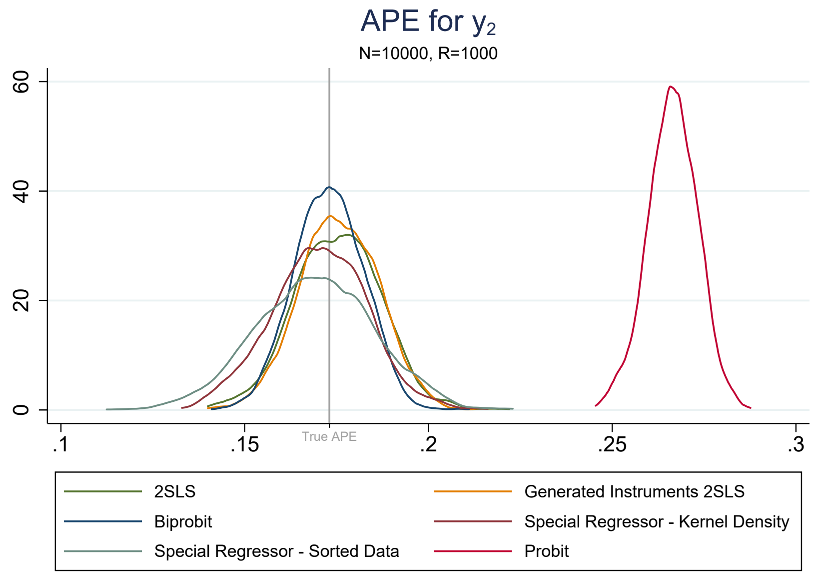
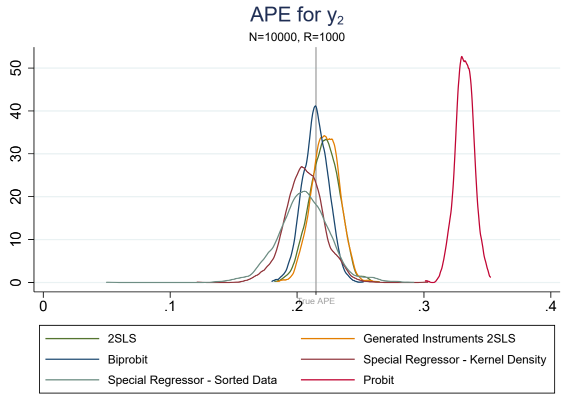
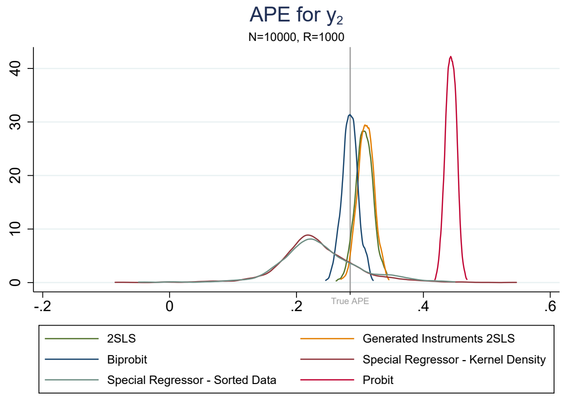


Figure A2: Different Support for Special Regressor

(a) $SD(v) = 1.385 \approx SD(y_1 - v) = 1.382$



(b) $SD(v) = 0.5 < SD(y_1 - v) = 1.38$



(c) $SD(v) = 5 \gg SD(y_1 - v) = 1.413$

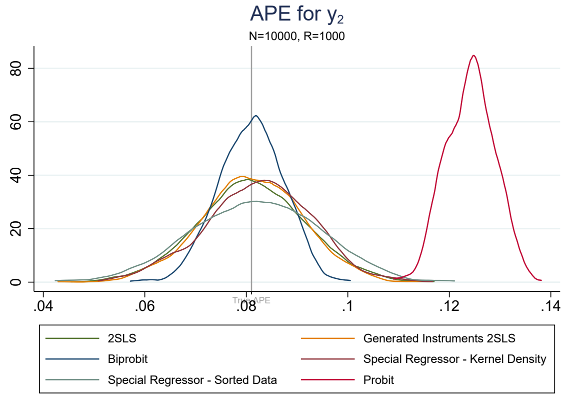
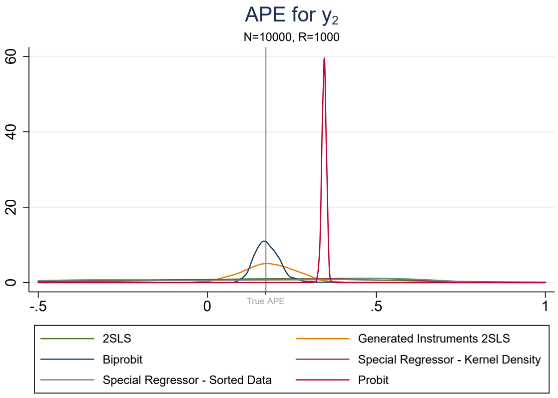
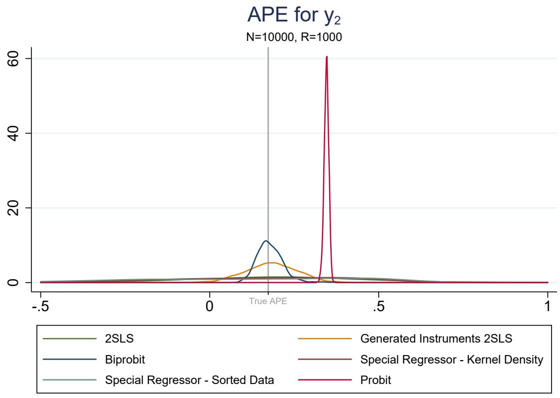


Figure A3: Weak Instruments

(a) $F[\delta_{z_{21}} = 0] \approx 5$



(b) $F[\delta_{z_{21}} = 0] \approx 10$



(c) $F[\delta_{z_{21}} = 0] \approx 20$

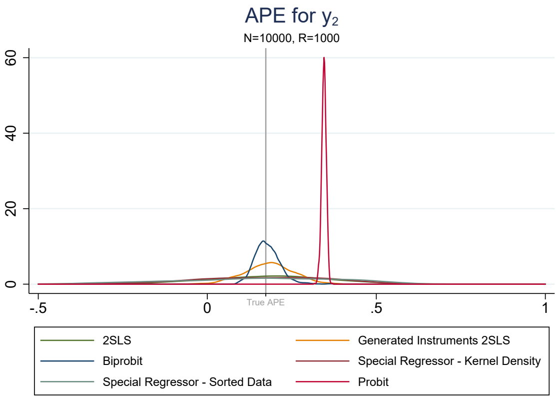
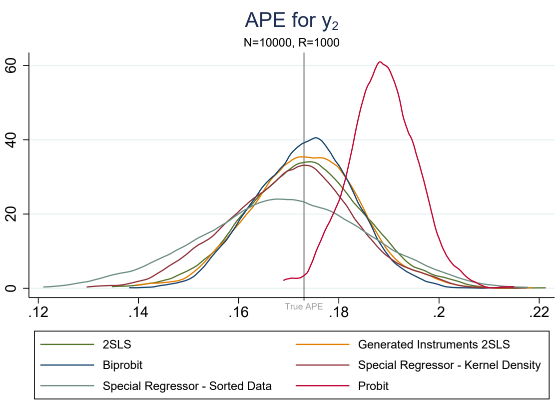


Figure A4: Different degree of endogeneity

(a) $\text{corr}(u_1,u_2) = 0.1$



(b) $\text{corr}(u_1,u_2) = 0.9$

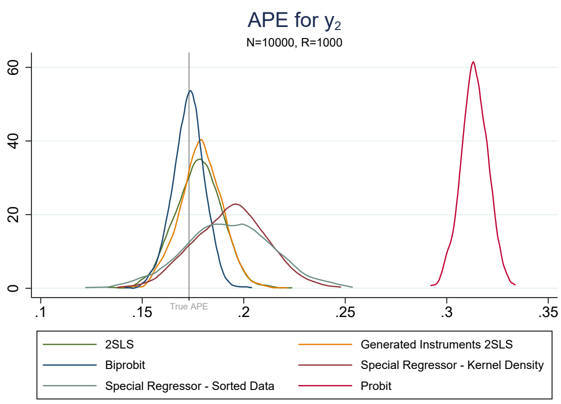
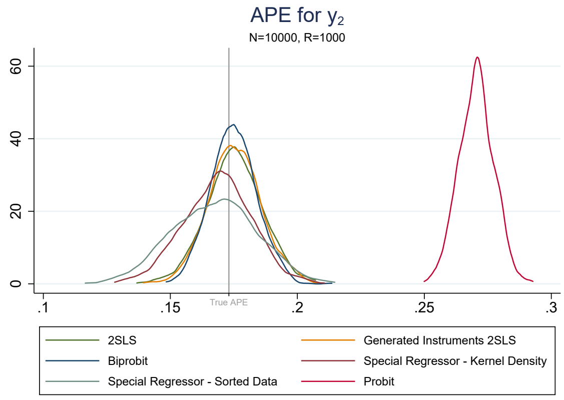
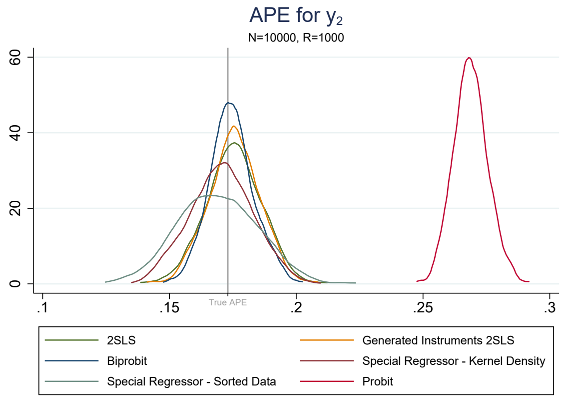


Figure A5: Different joint CDF

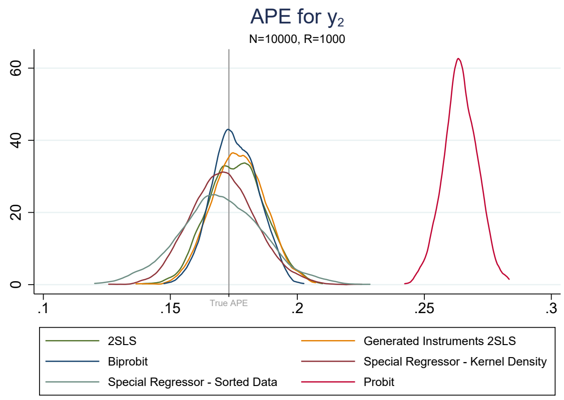
(a) $F_{U_1,U_2}(u_1, u_2)$ constructed by Clayton Copula $\mathbb{C}_\theta^C$



(b) $F_{U_1,U_2}(u_1, u_2)$ constructed by Frank Copula $\mathbb{C}_\theta^F$



(c) $F_{U_1,U_2}(u_1, u_2)$ constructed by Gumbel Copula $\mathbb{C}_\theta^G$



(d) $F_{U_1,U_2}(u_1, u_2)$ constructed by t Copula $\mathbb{C}_\theta^t$

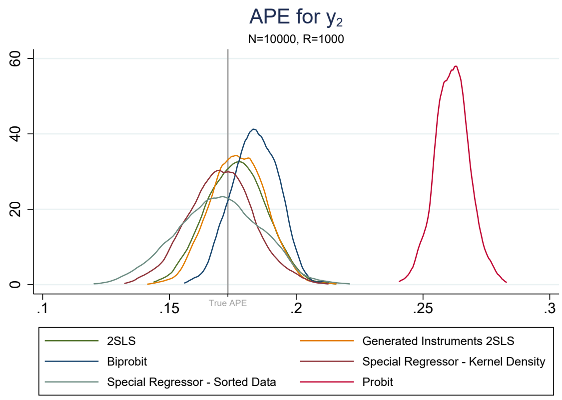
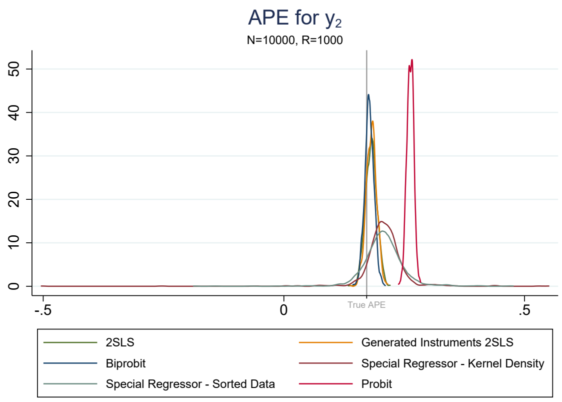
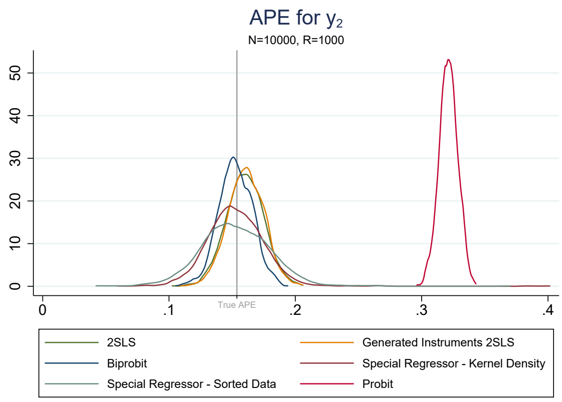


Figure A6: Different marginal CDF

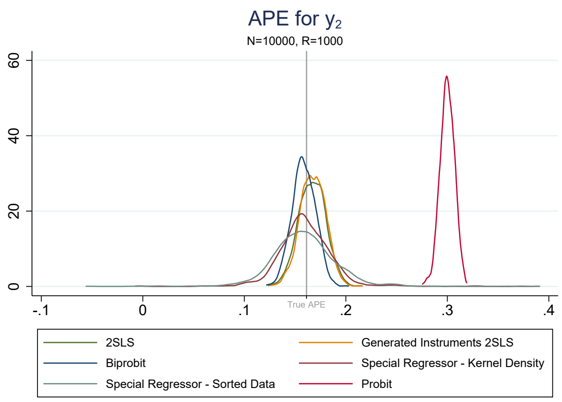
(a) $u_1 \sim F(10,6)$, $u_2 \sim F(10,6)$



(b) $u_1 \sim log(10,6)$, $u_2 \sim log(10,6)$



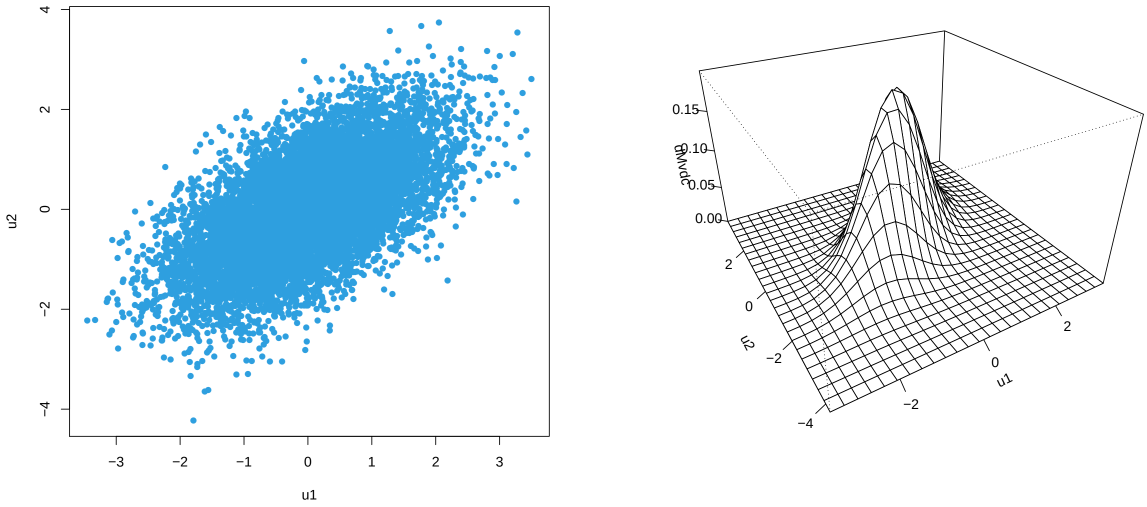
(c) $u_1 \sim t(3)$, $u_2 \sim t(3)$



Visualization of (joint) CDF(s)

Figure A7: (Joint) CDF(s) of a Normal Copula

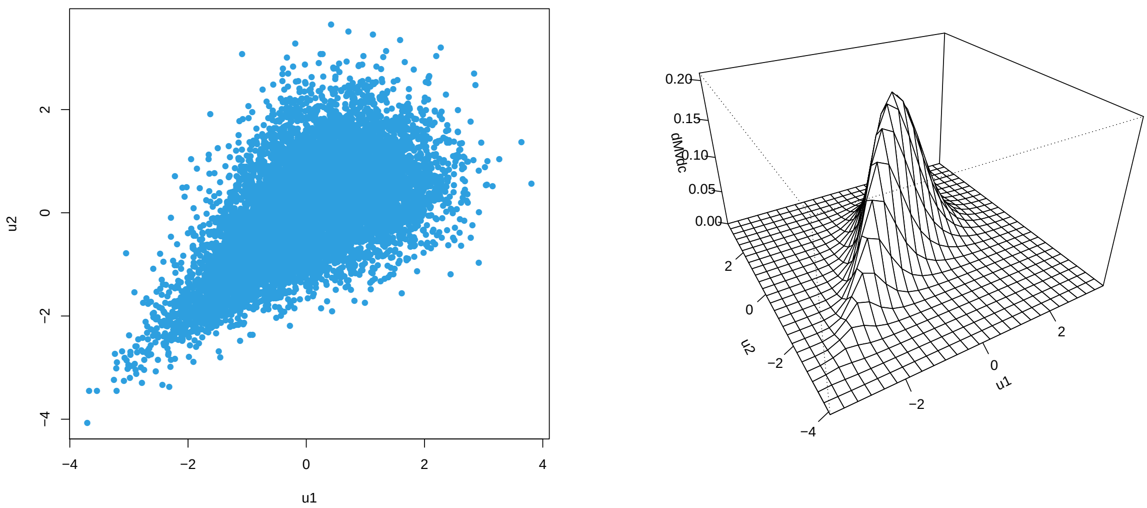
(a) Scatter plot (b) Joint PDF



Note: $u_1 \sim N(0,1)$, $u_2 \sim N(0,1)$

Figure A8: (Joint) CDF(s) of a Clayton Copula

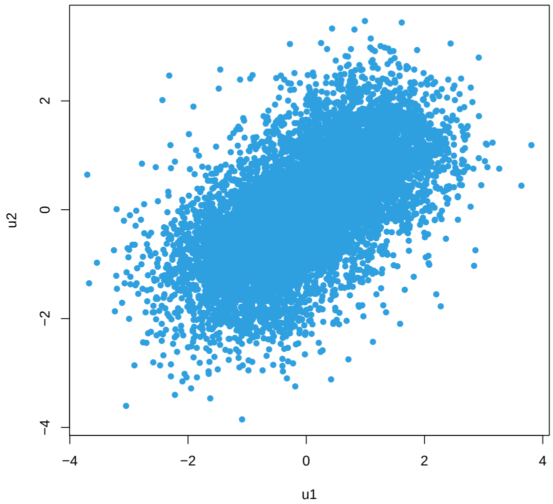
(a) Scatter plot (b) Joint PDF



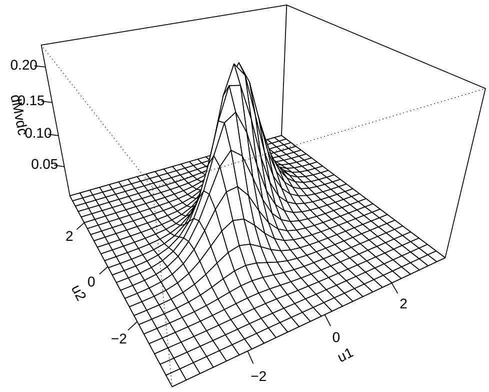
Note: $u_1 \sim N(0,1)$, $u_2 \sim N(0,1)$

Figure A9: (Joint) CDF(s) of a Frank Copula

(a) Scatter plot



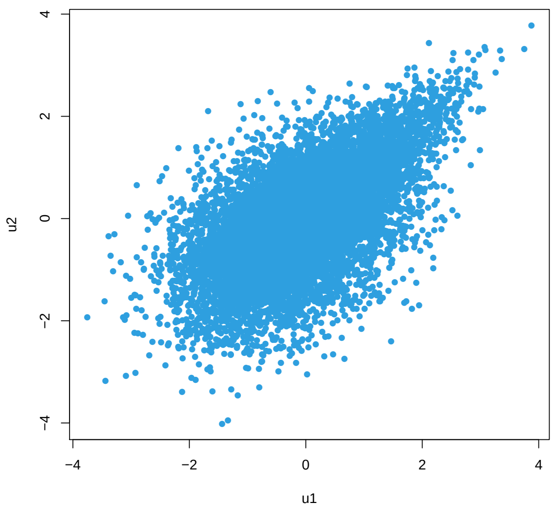
(b) Joint PDF



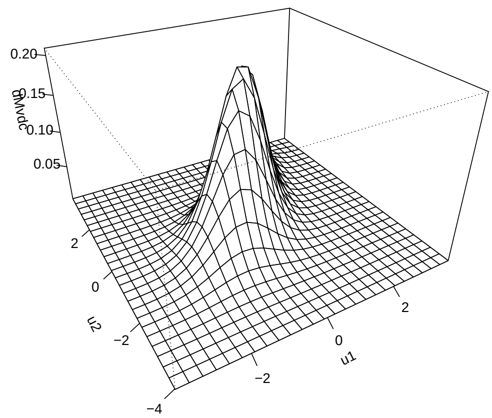
Note: $u_1 \sim N(0,1)$, $u_2 \sim N(0,1)$

Figure A10: (Joint) CDF(s) of a Gumbel Copula

(a) Scatter plot



(b) Joint PDF

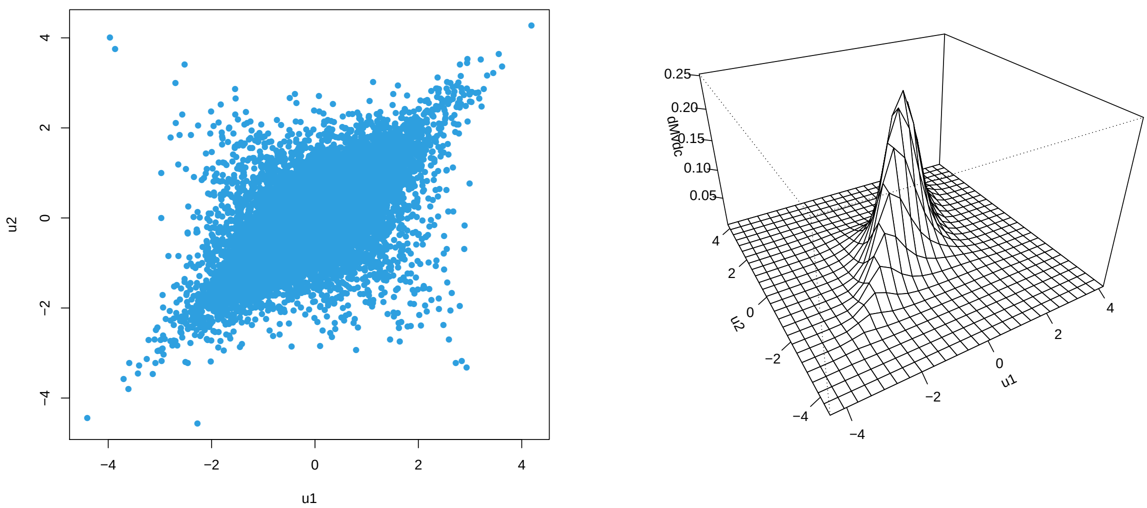


Note: $u_1 \sim N(0,1)$, $u_2 \sim N(0,1)$

Figure A11: (Joint) CDF(s) of a t Copula

(a) Scatter plot

(b) Joint PDF



Note: $u_1 \sim N(0,1)$, $u_2 \sim N(0,1)$

Tables:

Table A1: Simulation Results for APE of y_2 - Baseline (All statistics)

Baseline				*true* $APE_{y_2} = 0.173$					
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1729	-0.0001	0.0069	0.0066	0.1684	0.1733	0.1777	-	50-50
2SLS	0.1753	0.0023	0.0117	0.0117	0.1675	0.1752	0.1833	-	43-57
Generated Instr. 2SLS	0.1758	0.0028	0.0108	0.0109	0.1687	0.1757	0.1833	-	40-60
Recursive Biprobit	0.1729	-0.0001	0.0093	0.0091	0.1665	0.1729	0.1792	-	51-49
Special Reg. KeDe AIF	0.1708	-0.0022	0.0129	0.0129	0.1625	0.1710	0.1799	-	57-43
Special Reg. SoDa AIF	0.1697	-0.0034	0.0159	0.0161	0.1590	0.1695	0.1802	-	58-42
Special Reg. KeDe ASF	0.1703	-0.0027	0.0129	0.0130	0.1620	0.1705	0.1794	-	58-42
Special Reg. SoDa ASF	0.1691	-0.0039	0.0160	0.0163	0.1583	0.1691	0.1796	-	59-41
Probit	0.2665	0.0934	0.0068	0.0936	0.2620	0.2666	0.2710	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1731	0.0000	0.0097	0.0095	0.1670	0.1733	0.1794	-	50-50
2SLS	0.1757	0.0026	0.0165	0.0165	0.1648	0.1762	0.1866	-	43-57
Generated Instr. 2SLS	0.1760	0.0029	0.0151	0.0152	0.1663	0.1763	0.1858	-	41-59
Recursive Biprobit	0.1736	0.0005	0.0133	0.0131	0.1647	0.1743	0.1825	-	47-53
Special Reg. KeDe AIF	0.1716	-0.0015	0.0182	0.0181	0.1592	0.1716	0.1842	-	52-48
Special Reg. SoDa AIF	0.1700	-0.0031	0.0226	0.0227	0.1545	0.1703	0.1863	-	55-45
Special Reg. KeDe ASF	0.1715	-0.0016	0.0183	0.0182	0.1591	0.1714	0.1842	-	53-47
Special Reg. SoDa ASF	0.1698	-0.0033	0.0227	0.0229	0.1542	0.1700	0.1860	-	55-45
Probit	0.2664	0.0933	0.0096	0.0938	0.2600	0.2665	0.2727	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1739	0.0009	0.0218	0.0214	0.1596	0.1739	0.1886	-	48-52
2SLS	0.1770	0.0040	0.0370	0.0368	0.1522	0.1782	0.2015	-	44-56
Generated Instr. 2SLS	0.1785	0.0055	0.0344	0.0343	0.1569	0.1795	0.2026	-	41-59
Recursive Biprobit	0.1746	0.0016	0.0296	0.0292	0.1545	0.1757	0.1947	-	47-53
Special Reg. KeDe AIF	0.1700	-0.0030	0.0414	0.0413	0.1440	0.1709	0.1991	-	52-48
Special Reg. SoDa AIF	0.1662	-0.0069	0.0507	0.0508	0.1328	0.1676	0.2016	-	55-45
Special Reg. KeDe ASF	0.1718	-0.0013	0.0421	0.0419	0.1447	0.1721	0.2007	-	51-49
Special Reg. SoDa ASF	0.1679	-0.0051	0.0514	0.0513	0.1339	0.1695	0.2035	-	53-47
Probit	0.2667	0.0937	0.0212	0.0959	0.2526	0.2666	0.2807	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1727	-0.0005	0.0314	0.0306	0.1515	0.1728	0.1931	-	51-49
2SLS	0.1742	0.0009	0.0513	0.0507	0.1405	0.1752	0.2071	-	49-51
Generated Instr. 2SLS	0.1751	0.0018	0.0482	0.0475	0.1421	0.1752	0.2046	-	49-51
Recursive Biprobit	0.1720	-0.0013	0.0418	0.0414	0.1411	0.1718	0.2000	-	52-48
Special Reg. KeDe AIF	0.1671	-0.0062	0.0572	0.0572	0.1279	0.1686	0.2085	-	52-48
Special Reg. SoDa AIF	0.1629	-0.0104	0.0697	0.0702	0.1142	0.1657	0.2117	-	54-46
Special Reg. KeDe ASF	0.1706	-0.0027	0.0585	0.0582	0.1309	0.1722	0.2134	-	50-50
Special Reg. SoDa ASF	0.1665	-0.0068	0.0716	0.0716	0.1157	0.1688	0.2172	-	53-47
Probit	0.2663	0.0930	0.0306	0.0975	0.2442	0.2672	0.2876	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1727	-0.0005	0.0433	0.0425	0.1434	0.1725	0.2015	-	51-49
2SLS	0.1709	-0.0024	0.0724	0.0718	0.1216	0.1712	0.2194	-	51-49
Generated Instr. 2SLS	0.1734	0.0002	0.0667	0.0658	0.1285	0.1764	0.2178	-	48-52
Recursive Biprobit	0.1685	-0.0047	0.0596	0.0591	0.1260	0.1706	0.2061	-	51-49
Special Reg. KeDe AIF	0.1608	-0.0124	0.0798	0.0806	0.1067	0.1603	0.2169	-	56-44
Special Reg. SoDa AIF	0.1546	-0.0187	0.0928	0.0946	0.0956	0.1532	0.2215	-	59-41
Special Reg. KeDe ASF	0.1663	-0.0070	0.0824	0.0825	0.1108	0.1635	0.2245	-	54-46
Special Reg. SoDa ASF	0.1603	-0.0129	0.0973	0.0981	0.1002	0.1578	0.2273	-	56-44
Probit	0.2659	0.0927	0.0433	0.1016	0.2365	0.2658	0.2958	**	1-99

"Reflecting endogeneity" plugs u_2 as additional covariate in the structural equation.
BIAS = average difference between the estimated and the true APE; SD = standard deviation; RMSE = root mean squared error; LQ = lower quartile; UP = upper quartile; Difftest = significance of the t-test of $H_0 : \mathbb{E}[\widehat{APE}] = \mathbb{E}[APE^{true}]$ with - : $p \leq 0.1$, * : $p < 0.1$, ** : $p < 0.05$, *** : $p < 0.01$, Biasratio describes the ratio of positive vs. negative bias
Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $corr(u_1, u_2) = 0.6$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$

Table A2: Simulation Results for APE of y_2 - Different Support for Special Regressor (All statistics)

$SD(v) < SD(y_1 - v)$ — *true* $APE_{y_2} = 0.215$

N = 10000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2146	-0.0003	0.0073	0.0072	0.2099	0.2149	0.2195	-	51-49
2SLS	0.2215	0.0066	0.0118	0.0134	0.2136	0.2218	0.2296	-	28-72
Generated Instr. 2SLS	0.2224	0.0075	0.0109	0.0131	0.2150	0.2225	0.2299	-	24-76
Recursive Biprobit	0.2145	-0.0004	0.0101	0.0099	0.2076	0.2145	0.2213	-	51-49
Special Reg. KeDe AIF	0.2092	-0.0057	0.0157	0.0166	0.1995	0.2093	0.2190	-	65-35
Special Reg. SoDa AIF	0.2086	-0.0063	0.0210	0.0218	0.1964	0.2085	0.2215	-	63-37
Special Reg. KeDe ASF	0.2059	-0.0090	0.0157	0.0180	0.1962	0.2059	0.2157	-	73-27
Special Reg. SoDa ASF	0.2054	-0.0095	0.0210	0.0230	0.1929	0.2052	0.2178	-	70-30
Probit	0.3320	0.1171	0.0073	0.1173	0.3272	0.3320	0.3370	***	100

N = 5000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2149	-0.0000	0.0108	0.0106	0.2075	0.2149	0.2219	-	50-50
2SLS	0.2223	0.0074	0.0171	0.0186	0.2110	0.2224	0.2337	-	33-67
Generated Instr. 2SLS	0.2230	0.0081	0.0161	0.0178	0.2126	0.2228	0.2334	-	29-71
Recursive Biprobit	0.2155	0.0005	0.0149	0.0147	0.2060	0.2157	0.2251	-	47-53
Special Reg. KeDe AIF	0.2101	-0.0048	0.0205	0.0210	0.1967	0.2097	0.2249	-	60-40
Special Reg. SoDa AIF	0.2086	-0.0064	0.0262	0.0269	0.1903	0.2095	0.2245	-	58-42
Special Reg. KeDe ASF	0.2074	-0.0075	0.0206	0.0218	0.1944	0.2068	0.2215	-	65-35
Special Reg. SoDa ASF	0.2058	-0.0091	0.0263	0.0277	0.1878	0.2066	0.2219	-	63-37
Probit	0.3318	0.1169	0.0106	0.1173	0.3248	0.3315	0.3386	***	100

N = 1000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2155	0.0006	0.0241	0.0238	0.1989	0.2159	0.2321	-	49-51
2SLS	0.2234	0.0085	0.0391	0.0397	0.1968	0.2239	0.2512	-	40-60
Generated Instr. 2SLS	0.2252	0.0102	0.0368	0.0378	0.2017	0.2256	0.2511	-	38-62
Recursive Biprobit	0.2163	0.0014	0.0334	0.0331	0.1941	0.2165	0.2392	-	48-52
Special Reg. KeDe AIF	0.2067	-0.0082	0.0444	0.0450	0.1777	0.2085	0.2364	-	56-44
Special Reg. SoDa AIF	0.2038	-0.0112	0.0549	0.0560	0.1697	0.2063	0.2412	-	57-43
Special Reg. KeDe ASF	0.2061	-0.0088	0.0448	0.0456	0.1767	0.2075	0.2355	-	57-43
Special Reg. SoDa ASF	0.2037	-0.0113	0.0559	0.0570	0.1685	0.2044	0.2401	-	58-42
Probit	0.3319	0.1169	0.0235	0.1191	0.3164	0.3330	0.3479	***	100

N = 500

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2155	0.0003	0.0349	0.0342	0.1911	0.2164	0.2389	-	49-51
2SLS	0.2222	0.0070	0.0547	0.0545	0.1837	0.2209	0.2611	-	45-55
Generated Instr. 2SLS	0.2234	0.0083	0.0512	0.0512	0.1876	0.2242	0.2583	-	43-57
Recursive Biprobit	0.2151	-0.0001	0.0474	0.0469	0.1820	0.2146	0.2468	-	51-49
Special Reg. KeDe AIF	0.2039	-0.0113	0.0623	0.0630	0.1615	0.2066	0.2489	-	55-45
Special Reg. SoDa AIF	0.1974	-0.0178	0.0790	0.0807	0.1456	0.1988	0.2545	-	57-43
Special Reg. KeDe ASF	0.2061	-0.0091	0.0633	0.0637	0.1628	0.2077	0.2522	-	54-46
Special Reg. SoDa ASF	0.2000	-0.0151	0.0807	0.0819	0.1467	0.2019	0.2584	-	56-44
Probit	0.3320	0.1169	0.0341	0.1214	0.3085	0.3327	0.3571	***	100

N = 250

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2173	0.0020	0.0486	0.0477	0.1838	0.2139	0.2505	-	50-50
2SLS	0.2204	0.0052	0.0761	0.0756	0.1682	0.2190	0.2727	-	48-52
Generated Instr. 2SLS	0.2219	0.0066	0.0712	0.0705	0.1748	0.2204	0.2727	-	46-54
Recursive Biprobit	0.2116	-0.0037	0.0669	0.0661	0.1686	0.2117	0.2566	-	51-49
Special Reg. KeDe AIF	0.1964	-0.0188	0.0871	0.0887	0.1413	0.2006	0.2542	-	56-43
Special Reg. SoDa AIF	0.1928	-0.0225	0.1024	0.1044	0.1271	0.2000	0.2662	-	56-44
Special Reg. KeDe ASF	0.2024	-0.0128	0.0927	0.0931	0.1451	0.2049	0.2631	-	54-46
Special Reg. SoDa ASF	0.1994	-0.0159	0.1098	0.1104	0.1275	0.2057	0.2751	-	53-47
Probit	0.3330	0.1177	0.0479	0.1265	0.2988	0.3322	0.3637	**	100

$SD(v) < SD(y_1 - v)$ — *true* $APE_{y_2} = 0.285$

N = 10000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2844	-0.0002	0.0089	0.0088	0.2787	0.2847	0.2905	-	50-50
2SLS	0.3068	0.0222	0.0136	0.0260	0.2980	0.3067	0.3160	-	5-95
Generated Instr. 2SLS	0.3094	0.0248	0.0128	0.0279	0.3009	0.3097	0.3186	*	3-97

-continued on next page-

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Recursive Biprobit	0.2844	-0.0002	0.0125	0.0125	0.2762	0.2846	0.2925	-	51-49
Special Reg. KeDe AIF	0.2721	-0.0125	0.0622	0.0634	0.2452	0.2754	0.3060	-	59-41
Special Reg. SoDa AIF	0.2736	-0.0110	0.0644	0.0654	0.2425	0.2762	0.3108	-	56-44
Special Reg. KeDe ASF	0.2327	-0.0520	0.0635	0.0820	0.1989	0.2258	0.2644	-	84-16
Special Reg. SoDa ASF	0.2357	-0.0489	0.0661	0.0822	0.1998	0.2296	0.2695	-	81-19
Probit	0.4435	0.1588	0.0086	0.1591	0.4375	0.4437	0.4496	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2852	0.0005	0.0126	0.0125	0.2769	0.2849	0.2943	-	49-51
2SLS	0.3076	0.0230	0.0188	0.0297	0.2955	0.3071	0.3199	-	10-90
Generated Instr. 2SLS	0.3102	0.0256	0.0175	0.0310	0.2990	0.3094	0.3215	-	7-93
Recursive Biprobit	0.2853	0.0007	0.0171	0.0170	0.2738	0.2844	0.2964	-	51-49
Special Reg. KeDe AIF	0.2674	-0.0172	0.0711	0.0731	0.2379	0.2729	0.3072	-	59-41
Special Reg. SoDa AIF	0.2682	-0.0164	0.0701	0.0720	0.2335	0.2734	0.3125	-	58-42
Special Reg. KeDe ASF	0.2245	-0.0601	0.0673	0.0902	0.1897	0.2224	0.2587	-	85-15
Special Reg. SoDa ASF	0.2282	-0.0564	0.0709	0.0907	0.1874	0.2257	0.2685	-	81-19
Probit	0.4436	0.1590	0.0118	0.1594	0.4359	0.4439	0.4518	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2846	-0.0000	0.0276	0.0276	0.2653	0.2849	0.3023	-	50-50
2SLS	0.3069	0.0223	0.0416	0.0472	0.2788	0.3077	0.3359	-	29-71
Generated Instr. 2SLS	0.3100	0.0253	0.0388	0.0463	0.2835	0.3097	0.3360	-	26-74
Recursive Biprobit	0.2851	0.0004	0.0381	0.0380	0.2590	0.2858	0.3108	-	49-51
Special Reg. KeDe AIF	0.2620	-0.0227	0.0853	0.0883	0.2149	0.2698	0.3142	-	59-41
Special Reg. SoDa AIF	0.2591	-0.0255	0.0949	0.0983	0.2086	0.2664	0.3216	-	57-43
Special Reg. KeDe ASF	0.2201	-0.0646	0.0836	0.1057	0.1711	0.2159	0.2680	-	82-18
Special Reg. SoDa ASF	0.2197	-0.0649	0.0926	0.1130	0.1636	0.2167	0.2760	-	77-23
Probit	0.4430	0.1583	0.0267	0.1605	0.4268	0.4429	0.4591	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2840	-0.0008	0.0406	0.0402	0.2547	0.2835	0.3099	-	51-49
2SLS	0.3070	0.0222	0.0597	0.0634	0.2648	0.3068	0.3457	-	35-65
Generated Instr. 2SLS	0.3097	0.0249	0.0559	0.0609	0.2706	0.3095	0.3458	-	33-67
Recursive Biprobit	0.2827	-0.0021	0.0555	0.0552	0.2434	0.2823	0.3194	-	52-48
Special Reg. KeDe AIF	0.2429	-0.0419	0.1008	0.1091	0.1936	0.2499	0.3062	-	65-35
Special Reg. SoDa AIF	0.2402	-0.0446	0.1088	0.1174	0.1771	0.2522	0.3178	-	62-38
Special Reg. KeDe ASF	0.2021	-0.0827	0.0971	0.1275	0.1487	0.2017	0.2508	-	83-17
Special Reg. SoDa ASF	0.2005	-0.0843	0.1054	0.1348	0.1377	0.2001	0.2657	-	81-19
Probit	0.4422	0.1574	0.0385	0.1619	0.4160	0.4434	0.4690	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.2853	0.0005	0.0586	0.0580	0.2444	0.2862	0.3247	-	49-51
2SLS	0.3060	0.0212	0.0826	0.0849	0.2540	0.3075	0.3607	-	39-61
Generated Instr. 2SLS	0.3095	0.0247	0.0781	0.0814	0.2574	0.3101	0.3620	-	36-64
Recursive Biprobit	0.2804	-0.0043	0.0790	0.0786	0.2301	0.2809	0.3312	-	52-48
Special Reg. KeDe AIF	0.2350	-0.0498	0.1141	0.1243	0.1619	0.2460	0.3147	-	64-36
Special Reg. SoDa AIF	0.2295	-0.0553	0.1217	0.1334	0.1502	0.2452	0.3149	-	64-36
Special Reg. KeDe ASF	0.1938	-0.0910	0.1099	0.1426	0.1284	0.1915	0.2666	-	80-20
Special Reg. SoDa ASF	0.1937	-0.0911	0.1187	0.1494	0.1179	0.1961	0.2747	-	79-21
Probit	0.4434	0.1586	0.0543	0.1674	0.4077	0.4446	0.4810	***	100
$SD(v) \gg SD(y_1 - v)$			*true* $APE_{y_2} = 0.081$						
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.0810	-0.0000	0.0047	0.0045	0.0777	0.0811	0.0841	-	50-50
2SLS	0.0811	0.0001	0.0106	0.0104	0.0743	0.0807	0.0881	-	50-50
Generated Instr. 2SLS	0.0812	0.0002	0.0098	0.0097	0.0748	0.0809	0.0878	-	50-50
Recursive Biprobit	0.0809	-0.0001	0.0064	0.0062	0.0765	0.0811	0.0852	-	50-50
Special Reg. KeDe AIF	0.0821	0.0011	0.0104	0.0102	0.0755	0.0826	0.0893	-	44-56
Special Reg. SoDa AIF	0.0817	0.0006	0.0125	0.0124	0.0731	0.0819	0.0903	-	47-53
Special Reg. KeDe ASF	0.0824	0.0014	0.0104	0.0103	0.0757	0.0829	0.0896	-	43-57
Special Reg. SoDa ASF	0.0819	0.0009	0.0126	0.0124	0.0732	0.0822	0.0904	-	46-54
Probit	0.1242	0.0431	0.0048	0.0434	0.1207	0.1243	0.1274	***	100
N = 5000									

-continued on next page-

- Table A2 continued -

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.0809	-0.0002	0.0069	0.0066	0.0762	0.0809	0.0855	-	51-49
2SLS	0.0809	-0.0002	0.0154	0.0153	0.0704	0.0808	0.0913	-	51-49
Generated Instr. 2SLS	0.0810	-0.0000	0.0140	0.0138	0.0712	0.0809	0.0910	-	51-49
Recursive Biprobit	0.0811	0.0001	0.0089	0.0087	0.0750	0.0810	0.0875	-	50-50
Special Reg. KeDe AIF	0.0826	0.0015	0.0147	0.0145	0.0721	0.0830	0.0925	-	46-54
Special Reg. SoDa AIF	0.0817	0.0007	0.0175	0.0173	0.0699	0.0815	0.0937	-	48-52
Special Reg. KeDe ASF	0.0830	0.0019	0.0148	0.0146	0.0723	0.0833	0.0928	-	45-55
Special Reg. SoDa ASF	0.0821	0.0011	0.0176	0.0174	0.0702	0.0819	0.0941	-	47-53
Probit	0.1241	0.0430	0.0069	0.0435	0.1193	0.1240	0.1287	***	100

N = 1000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.0810	0.0000	0.0151	0.0146	0.0709	0.0815	0.0903	-	51-49
2SLS	0.0825	0.0015	0.0341	0.0338	0.0590	0.0830	0.1055	-	47-53
Generated Instr. 2SLS	0.0822	0.0012	0.0310	0.0307	0.0611	0.0825	0.1034	-	48-52
Recursive Biprobit	0.0805	-0.0005	0.0202	0.0198	0.0674	0.0804	0.0931	-	51-49
Special Reg. KeDe AIF	0.0825	0.0015	0.0323	0.0318	0.0606	0.0823	0.1039	-	49-51
Special Reg. SoDa AIF	0.0801	-0.0009	0.0400	0.0396	0.0521	0.0773	0.1081	-	53-47
Special Reg. KeDe ASF	0.0837	0.0027	0.0330	0.0325	0.0616	0.0838	0.1051	-	47-53
Special Reg. SoDa ASF	0.0812	0.0002	0.0406	0.0402	0.0532	0.0783	0.1086	-	51-49
Probit	0.1241	0.0431	0.0151	0.0453	0.1146	0.1241	0.1339	***	100

N = 500

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.0805	-0.0006	0.0223	0.0214	0.0649	0.0801	0.0958	-	51-49
2SLS	0.0796	-0.0015	0.0482	0.0478	0.0478	0.0775	0.1133	-	52-48
Generated Instr. 2SLS	0.0792	-0.0019	0.0445	0.0441	0.0487	0.0778	0.1093	-	53-47
Recursive Biprobit	0.0800	-0.0011	0.0298	0.0291	0.0610	0.0802	0.0989	-	51-49
Special Reg. KeDe AIF	0.0813	0.0002	0.0457	0.0448	0.0503	0.0805	0.1119	-	50-50
Special Reg. SoDa AIF	0.0799	-0.0012	0.0551	0.0543	0.0417	0.0782	0.1183	-	52-48
Special Reg. KeDe ASF	0.0831	0.0020	0.0470	0.0461	0.0516	0.0821	0.1149	-	48-52
Special Reg. SoDa ASF	0.0816	0.0005	0.0566	0.0558	0.0424	0.0806	0.1198	-	51-49
Probit	0.1243	0.0432	0.0222	0.0478	0.1095	0.1243	0.1386	*	1-99

N = 250

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.0805	-0.0005	0.0298	0.0288	0.0595	0.0810	0.1007	-	49-51
2SLS	0.0782	-0.0028	0.0653	0.0646	0.0346	0.0774	0.1223	-	53-47
Generated Instr. 2SLS	0.0791	-0.0019	0.0605	0.0598	0.0358	0.0790	0.1205	-	52-48
Recursive Biprobit	0.0794	-0.0016	0.0395	0.0390	0.0529	0.0784	0.1057	-	52-48
Special Reg. KeDe AIF	0.0766	-0.0044	0.0644	0.0639	0.0339	0.0772	0.1200	-	52-48
Special Reg. SoDa AIF	0.0728	-0.0082	0.0739	0.0739	0.0231	0.0691	0.1242	-	56-44
Special Reg. KeDe ASF	0.0805	-0.0005	0.0730	0.0724	0.0358	0.0778	0.1231	-	50-50
Special Reg. SoDa ASF	0.0760	-0.0050	0.0790	0.0788	0.0250	0.0717	0.1293	-	55-45
Probit	0.1232	0.0422	0.0292	0.0502	0.1034	0.1226	0.1434	-	7-93

"Reflecting endogeneity" plugs u_2 as additional covariate in the strucutural equation.
BIAS = average difference between the estimated and the true APE; SD = standard deviation; RMSE = root mean squared error; LQ = lower quartile; UP = upper quartile; Difftest = significance of the t-test of $H_0 : \mathbb{E}[\widehat{APE}] = \mathbb{E}[APE^{true}]$ with - : p $\leq$ 0.1 , * : p $<$ 0.1 , ** : p $<$ 0.05, *** : p $<$ 0.01, Biasratio describes the ratio of positive vs. negative bias
Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $corr(u_1,u_2) = 0.6$, $F[\delta_{z_{21}}] \geq 100.00$

Table A3: Simulation Results for APE of y_2 - Weak instrument setting (All statistics)

$F[\delta_{z_{21}} = 0] \approx 5$				*true* $APE_{y_2} = 0.173$					
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1733	0.0003	0.0096	0.0094	0.1673	0.1735	0.1797	-	48-52
2SLS	0.3716	0.1985	5.6715	5.6721	-0.1119	0.1870	0.4434	-	48-52
Generated Instr. 2SLS	0.1820	0.0090	0.0774	0.0778	0.1311	0.1823	0.2363	-	45-55
Recursive Biprobit	0.1782	0.0051	0.0386	0.0389	0.1512	0.1746	0.2014	-	49-51
Special Reg. KeDe AIF	0.1634	-0.0096	0.2768	0.2767	-0.1095	0.1693	0.4255	-	51-49
Special Reg. SoDa AIF	0.1614	-0.0117	0.2981	0.2981	-0.1432	0.1577	0.4584	-	51-49
Special Reg. KeDe ASF	0.1224	-0.0506	0.3578	0.3611	-0.1565	0.1675	0.4434	-	51-49
Special Reg. SoDa ASF	0.1064	-0.0666	0.3944	0.3998	-0.2460	0.1547	0.4821	-	52-48
Probit	0.3455	0.1724	0.0068	0.1726	0.3409	0.3458	0.3500	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1734	0.0003	0.0140	0.0138	0.1639	0.1736	0.1829	-	49-51
2SLS	-0.6011	-0.7742	15.6653	15.6767	-0.1019	0.1967	0.4422	-	48-52
Generated Instr. 2SLS	0.1835	0.0104	0.1074	0.1079	0.1161	0.1858	0.2530	-	44-56
Recursive Biprobit	0.1825	0.0094	0.0594	0.0601	0.1420	0.1744	0.2122	-	49-51
Special Reg. KeDe AIF	0.1674	-0.0057	0.2801	0.2800	-0.1113	0.1907	0.4321	-	49-51
Special Reg. SoDa AIF	0.1631	-0.0100	0.3015	0.3014	-0.1453	0.1472	0.4628	-	52-48
Special Reg. KeDe ASF	0.1260	-0.0471	0.3602	0.3630	-0.1707	0.1852	0.4453	-	49-51
Special Reg. SoDa ASF	0.1096	-0.0635	0.3961	0.4009	-0.2408	0.1430	0.4863	-	52-48
Probit	0.3453	0.1722	0.0100	0.1725	0.3387	0.3453	0.3521	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1728	-0.0002	0.0308	0.0305	0.1529	0.1732	0.1928	-	49-51
2SLS	0.1353	-0.0377	1.5655	1.5652	-0.0833	0.2145	0.4801	-	46-54
Generated Instr. 2SLS	0.1897	0.0166	0.2163	0.2166	0.0647	0.1907	0.3265	-	46-54
Recursive Biprobit	0.2051	0.0321	0.1252	0.1290	0.1172	0.1834	0.2705	-	47-53
Special Reg. KeDe AIF	0.1733	0.0003	0.2721	0.2720	-0.0810	0.1861	0.4258	-	49-51
Special Reg. SoDa AIF	0.1722	-0.0008	0.2941	0.2939	-0.1221	0.1829	0.4539	-	49-51
Special Reg. KeDe ASF	0.1380	-0.0351	0.3546	0.3562	-0.1191	0.1858	0.4476	-	49-51
Special Reg. SoDa ASF	0.1274	-0.0456	0.3880	0.3906	-0.2036	0.1812	0.4816	-	49-51
Probit	0.3441	0.1711	0.0216	0.1723	0.3288	0.3440	0.3593	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1741	0.0008	0.0453	0.0446	0.1444	0.1715	0.2046	-	52-48
2SLS	0.2933	0.1200	3.2981	3.2988	-0.0864	0.2044	0.4654	-	47-53
Generated Instr. 2SLS	0.1727	-0.0006	0.3552	0.3550	0.0159	0.2115	0.3529	-	45-55
Recursive Biprobit	0.2158	0.0425	0.1574	0.1629	0.0926	0.1976	0.3214	-	45-55
Special Reg. KeDe AIF	0.1810	0.0077	0.2641	0.2640	-0.0715	0.2081	0.4123	-	46-54
Special Reg. SoDa AIF	0.1748	0.0015	0.2841	0.2840	-0.1072	0.1779	0.4408	-	49-51
Special Reg. KeDe ASF	0.1448	-0.0285	0.3573	0.3583	-0.1118	0.2100	0.4352	-	47-53
Special Reg. SoDa ASF	0.1306	-0.0427	0.3881	0.3903	-0.1925	0.1765	0.4730	-	50-50
Probit	0.3442	0.1709	0.0325	0.1736	0.3211	0.3429	0.3647	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1721	-0.0011	0.0633	0.0626	0.1308	0.1703	0.2129	-	51-49
2SLS	0.1503	-0.0229	2.9459	2.9437	-0.1107	0.1668	0.4143	-	51-49
Generated Instr. 2SLS	0.1802	0.0070	0.5212	0.5209	0.0034	0.1947	0.3946	-	46-54
Recursive Biprobit	0.2169	0.0437	0.1846	0.1894	0.0652	0.1907	0.3547	-	47-53
Special Reg. KeDe AIF	0.1548	-0.0185	0.2730	0.2734	-0.1087	0.1606	0.4013	-	51-49
Special Reg. SoDa AIF	0.1626	-0.0106	0.2860	0.2862	-0.1153	0.1723	0.4252	-	50-50
Special Reg. KeDe ASF	0.1169	-0.0563	0.3657	0.3699	-0.1791	0.1597	0.4346	-	51-49
Special Reg. SoDa ASF	0.1133	-0.0599	0.3980	0.4025	-0.2422	0.1647	0.4669	-	51-49
Probit	0.3419	0.1686	0.0451	0.1740	0.3117	0.3411	0.3719	***	100

$F[\delta_{z_{21}} = 0] \approx 10$				*true* $APE_{y_2} = 0.173$					
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1733	0.0002	0.0095	0.0094	0.1670	0.1734	0.1798	-	49-51
2SLS	0.1503	-0.0227	0.3521	0.3526	-0.0148	0.1788	0.3513	-	49-51
Generated Instr. 2SLS	0.1820	0.0090	0.0756	0.0760	0.1317	0.1820	0.2331	-	45-55

-continued on next page-

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Recursive Biprobit	0.1782	0.0051	0.0382	0.0384	0.1521	0.1743	0.2004	-	49-51
Special Reg. KeDe AIF	0.1510	-0.0221	0.2374	0.2382	-0.0613	0.1656	0.3585	-	51-49
Special Reg. SoDa AIF	0.1481	-0.0250	0.2655	0.2664	-0.1167	0.1566	0.3935	-	51-49
Special Reg. KeDe ASF	0.1379	-0.0352	0.2711	0.2731	-0.0627	0.1636	0.3612	-	52-48
Special Reg. SoDa ASF	0.1268	-0.0463	0.3109	0.3141	-0.1250	0.1550	0.4000	-	51-49
Probit	0.3454	0.1723	0.0068	0.1725	0.3408	0.3455	0.3498	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1733	0.0002	0.0140	0.0138	0.1637	0.1736	0.1830	-	49-51
2SLS	0.1233	-0.0498	0.9741	0.9749	-0.0167	0.1852	0.3501	-	48-52
Generated Instr. 2SLS	0.1829	0.0098	0.1026	0.1031	0.1167	0.1848	0.2527	-	45-55
Recursive Biprobit	0.1811	0.0080	0.0547	0.0552	0.1436	0.1752	0.2109	-	49-51
Special Reg. KeDe AIF	0.1607	-0.0124	0.2409	0.2411	-0.0594	0.1850	0.3707	-	48-52
Special Reg. SoDa AIF	0.1591	-0.0140	0.2687	0.2689	-0.0996	0.1766	0.4074	-	50-50
Special Reg. KeDe ASF	0.1503	-0.0228	0.2729	0.2737	-0.0608	0.1830	0.3757	-	48-52
Special Reg. SoDa ASF	0.1422	-0.0309	0.3104	0.3117	-0.1063	0.1760	0.4167	-	50-50
Probit	0.3451	0.1720	0.0100	0.1722	0.3385	0.3451	0.3518	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1727	-0.0004	0.0307	0.0304	0.1521	0.1737	0.1922	-	49-51
2SLS	0.1511	-0.0220	0.4343	0.4344	0.0005	0.1998	0.3590	-	46-54
Generated Instr. 2SLS	0.1867	0.0137	0.1885	0.1886	0.0721	0.1885	0.3070	-	47-53
Recursive Biprobit	0.1990	0.0259	0.1177	0.1202	0.1155	0.1814	0.2618	-	47-53
Special Reg. KeDe AIF	0.1591	-0.0139	0.2323	0.2326	-0.0297	0.1772	0.3578	-	49-51
Special Reg. SoDa AIF	0.1496	-0.0234	0.2623	0.2633	-0.1082	0.1579	0.3959	-	52-48
Special Reg. KeDe ASF	0.1446	-0.0284	0.2753	0.2766	-0.0339	0.1781	0.3640	-	49-51
Special Reg. SoDa ASF	0.1264	-0.0467	0.3185	0.3218	-0.1315	0.1585	0.4091	-	51-49
Probit	0.3432	0.1701	0.0215	0.1713	0.3281	0.3428	0.3589	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1734	0.0001	0.0451	0.0444	0.1419	0.1721	0.2056	-	50-50
2SLS	0.1693	-0.0039	0.6167	0.6164	-0.0069	0.1860	0.3619	-	48-52
Generated Instr. 2SLS	0.1740	0.0007	0.2485	0.2483	0.0369	0.1937	0.3155	-	47-53
Recursive Biprobit	0.2016	0.0283	0.1424	0.1450	0.0961	0.1876	0.2883	-	47-53
Special Reg. KeDe AIF	0.1629	-0.0104	0.2321	0.2321	-0.0302	0.1923	0.3611	-	47-53
Special Reg. SoDa AIF	0.1540	-0.0193	0.2590	0.2595	-0.0967	0.1639	0.3849	-	51-49
Special Reg. KeDe ASF	0.1506	-0.0227	0.2773	0.2779	-0.0374	0.1932	0.3718	-	47-53
Special Reg. SoDa ASF	0.1321	-0.0412	0.3168	0.3192	-0.1306	0.1679	0.4012	-	51-49
Probit	0.3423	0.1690	0.0324	0.1717	0.3206	0.3414	0.3640	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1726	-0.0006	0.0613	0.0605	0.1334	0.1686	0.2118	-	52-48
2SLS	0.1266	-0.0466	0.7145	0.7158	-0.0217	0.1543	0.3287	-	53-47
Generated Instr. 2SLS	0.1824	0.0092	0.2615	0.2611	0.0298	0.1804	0.3435	-	49-51
Recursive Biprobit	0.2031	0.0299	0.1691	0.1712	0.0767	0.1780	0.3206	-	49-51
Special Reg. KeDe AIF	0.1441	-0.0291	0.2322	0.2340	-0.0608	0.1636	0.3452	-	52-48
Special Reg. SoDa AIF	0.1475	-0.0258	0.2546	0.2558	-0.0764	0.1489	0.3742	-	52-48
Special Reg. KeDe ASF	0.1345	-0.0387	0.2770	0.2798	-0.0728	0.1658	0.3585	-	51-49
Special Reg. SoDa ASF	0.1309	-0.0423	0.3095	0.3122	-0.1017	0.1522	0.3946	-	52-48
Probit	0.3387	0.1655	0.0449	0.1708	0.3072	0.3371	0.3692	***	100
$F[\delta_{z_{21}} = 0] \approx 20$				*true* $APE_{y_2} = 0.173$					
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1732	0.0001	0.0095	0.0094	0.1671	0.1731	0.1798	-	50-50
2SLS	0.1684	-0.0046	0.2043	0.2041	0.0529	0.1818	0.2974	-	48-52
Generated Instr. 2SLS	0.1820	0.0089	0.0726	0.0730	0.1359	0.1842	0.2298	-	44-56
Recursive Biprobit	0.1781	0.0050	0.0376	0.0379	0.1531	0.1746	0.1999	-	48-52
Special Reg. KeDe AIF	0.1561	-0.0169	0.1902	0.1908	0.0121	0.1680	0.3087	-	50-50
Special Reg. SoDa AIF	0.1591	-0.0139	0.2189	0.2192	-0.0127	0.1747	0.3427	-	50-50
Special Reg. KeDe ASF	0.1538	-0.0193	0.2012	0.2019	0.0122	0.1668	0.3087	-	51-49
Special Reg. SoDa ASF	0.1548	-0.0182	0.2366	0.2371	-0.0128	0.1736	0.3447	-	50-50
Probit	0.3452	0.1721	0.0068	0.1722	0.3408	0.3452	0.3497	***	100
N = 5000									

-continued on next page-

- Table A3 continued -

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1733	0.0002	0.0138	0.0136	0.1641	0.1736	0.1827	-	49-51
2SLS	0.1703	-0.0028	0.1975	0.1974	0.0544	0.1849	0.2970	-	47-53
Generated Instr. 2SLS	0.1819	0.0088	0.0946	0.0950	0.1226	0.1837	0.2446	-	45-55
Recursive Biprobit	0.1802	0.0071	0.0515	0.0519	0.1437	0.1742	0.2105	-	49-51
Special Reg. KeDe AIF	0.1619	-0.0112	0.1920	0.1921	0.0101	0.1880	0.3119	-	49-51
Special Reg. SoDa AIF	0.1551	-0.0180	0.2266	0.2272	-0.0291	0.1652	0.3498	-	51-49
Special Reg. KeDe ASF	0.1606	-0.0125	0.2023	0.2025	0.0102	0.1874	0.3129	-	49-51
Special Reg. SoDa ASF	0.1516	-0.0215	0.2439	0.2447	-0.0294	0.1645	0.3519	-	51-49
Probit	0.3448	0.1717	0.0099	0.1719	0.3380	0.3447	0.3513	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1727	-0.0003	0.0304	0.0300	0.1528	0.1730	0.1929	-	51-49
2SLS	0.1759	0.0029	0.2031	0.2027	0.0538	0.1926	0.3020	-	45-55
Generated Instr. 2SLS	0.1849	0.0119	0.1492	0.1493	0.0945	0.1872	0.2826	-	45-55
Recursive Biprobit	0.1889	0.0159	0.1019	0.1027	0.1210	0.1756	0.2427	-	48-52
Special Reg. KeDe AIF	0.1575	-0.0155	0.1869	0.1874	0.0246	0.1725	0.3038	-	51-49
Special Reg. SoDa AIF	0.1522	-0.0209	0.2160	0.2169	-0.0236	0.1595	0.3369	-	52-48
Special Reg. KeDe ASF	0.1558	-0.0173	0.2023	0.2028	0.0252	0.1733	0.3075	-	50-50
Special Reg. SoDa ASF	0.1480	-0.0250	0.2387	0.2399	-0.0244	0.1589	0.3435	-	52-48
Probit	0.3416	0.1685	0.0216	0.1697	0.3272	0.3413	0.3562	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1734	0.0001	0.0446	0.0439	0.1423	0.1727	0.2031	-	50-50
2SLS	0.1657	-0.0076	0.1978	0.1977	0.0406	0.1812	0.3001	-	49-51
Generated Instr. 2SLS	0.1728	-0.0005	0.1708	0.1706	0.0626	0.1856	0.2900	-	47-53
Recursive Biprobit	0.1847	0.0114	0.1190	0.1195	0.0991	0.1779	0.2610	-	48-52
Special Reg. KeDe AIF	0.1569	-0.0163	0.1893	0.1898	0.0187	0.1811	0.3061	-	49-51
Special Reg. SoDa AIF	0.1538	-0.0195	0.2180	0.2186	-0.0278	0.1795	0.3327	-	49-51
Special Reg. KeDe ASF	0.1574	-0.0159	0.2039	0.2043	0.0192	0.1855	0.3109	-	48-52
Special Reg. SoDa ASF	0.1512	-0.0221	0.2409	0.2416	-0.0293	0.1830	0.3462	-	49-51
Probit	0.3393	0.1660	0.0327	0.1688	0.3173	0.3385	0.3603	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1722	-0.0011	0.0584	0.0577	0.1345	0.1715	0.2114	-	51-49
2SLS	0.1605	-0.0127	0.1987	0.1988	0.0305	0.1590	0.2927	-	54-46
Generated Instr. 2SLS	0.1750	0.0017	0.1752	0.1747	0.0554	0.1750	0.2952	-	49-51
Recursive Biprobit	0.1829	0.0096	0.1387	0.1385	0.0811	0.1642	0.2714	-	52-48
Special Reg. KeDe AIF	0.1456	-0.0276	0.1857	0.1878	0.0004	0.1540	0.2917	-	54-46
Special Reg. SoDa AIF	0.1454	-0.0278	0.2071	0.2088	-0.0130	0.1507	0.3059	-	54-46
Special Reg. KeDe ASF	0.1477	-0.0256	0.2064	0.2081	-0.0006	0.1611	0.3036	-	52-48
Special Reg. SoDa ASF	0.1434	-0.0298	0.2374	0.2391	-0.0134	0.1547	0.3195	-	52-48
Probit	0.3325	0.1592	0.0445	0.1647	0.3020	0.3322	0.3631	***	100

"Reflecting endogeneity" plugs u_2 as additional covariate in the structural equation.
BIAS = average difference between the estimated and the true APE; SD = standard deviation; RMSE = root mean squared error; LQ = lower quartile; UP = upper quartile; Difftest = significance of the t-test of $H_0 : \mathbb{E}[\widehat{APE}] = \mathbb{E}[APE^{true}]$ with - : p $\leq$ 0.1 , * : p < 0.1 , ** : p < 0.05, *** : p < 0.01, Biasratio describes the ratio of positive vs. negative bias
Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $corr(u_1,u_2) = 0.6$, $SD(v) = 2.0$

Table A4: Simulation Results for APE of y_2 - Different degree of endogeneity (All statistics)

corr(u_1,u_2) = 0.1 — *true* APE_{y_2} = 0.173

N = 10000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1731	0.0000	0.0076	0.0073	0.1680	0.1733	0.1780	-	49-51
2SLS	0.1733	0.0002	0.0119	0.0117	0.1652	0.1731	0.1811	-	50-50
Generated Instr. 2SLS	0.1733	0.0002	0.0109	0.0107	0.1662	0.1735	0.1806	-	49-51
Recursive Biprobit	0.1731	0.0000	0.0098	0.0096	0.1664	0.1736	0.1795	-	48-52
Special Reg. KeDe AIF	0.1698	-0.0032	0.0124	0.0126	0.1618	0.1703	0.1781	-	59-41
Special Reg. SoDa AIF	0.1686	-0.0045	0.0161	0.0165	0.1581	0.1687	0.1796	-	61-39
Special Reg. KeDe ASF	0.1704	-0.0027	0.0125	0.0126	0.1619	0.1710	0.1785	-	58-42
Special Reg. SoDa ASF	0.1691	-0.0040	0.0163	0.0166	0.1583	0.1691	0.1804	-	59-41
Probit	0.1887	0.0156	0.0067	0.0169	0.1844	0.1887	0.1932	**	1-99

N = 5000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1732	0.0001	0.0103	0.0101	0.1657	0.1731	0.1806	-	50-50
2SLS	0.1731	0.0000	0.0166	0.0165	0.1621	0.1738	0.1843	-	49-51
Generated Instr. 2SLS	0.1732	0.0001	0.0152	0.0151	0.1629	0.1735	0.1830	-	49-51
Recursive Biprobit	0.1732	0.0001	0.0136	0.0135	0.1638	0.1734	0.1821	-	49-51
Special Reg. KeDe AIF	0.1702	-0.0029	0.0179	0.0180	0.1577	0.1707	0.1821	-	56-44
Special Reg. SoDa AIF	0.1685	-0.0046	0.0218	0.0221	0.1551	0.1685	0.1827	-	59-41
Special Reg. KeDe ASF	0.1711	-0.0020	0.0182	0.0181	0.1585	0.1714	0.1834	-	53-47
Special Reg. SoDa ASF	0.1695	-0.0036	0.0221	0.0223	0.1562	0.1698	0.1837	-	56-44
Probit	0.1884	0.0153	0.0093	0.0178	0.1821	0.1884	0.1948	-	5-95

N = 1000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1729	-0.0001	0.0234	0.0229	0.1572	0.1717	0.1884	-	51-49
2SLS	0.1739	0.0009	0.0371	0.0366	0.1492	0.1744	0.2004	-	47-53
Generated Instr. 2SLS	0.1747	0.0017	0.0343	0.0338	0.1519	0.1761	0.1980	-	46-54
Recursive Biprobit	0.1733	0.0003	0.0304	0.0299	0.1521	0.1740	0.1938	-	49-51
Special Reg. KeDe AIF	0.1675	-0.0056	0.0398	0.0399	0.1413	0.1685	0.1949	-	55-45
Special Reg. SoDa AIF	0.1642	-0.0088	0.0488	0.0494	0.1305	0.1656	0.1976	-	56-44
Special Reg. KeDe ASF	0.1703	-0.0027	0.0408	0.0406	0.1432	0.1714	0.1989	-	51-49
Special Reg. SoDa ASF	0.1670	-0.0061	0.0501	0.0502	0.1333	0.1677	0.2001	-	54-46
Probit	0.1883	0.0153	0.0216	0.0259	0.1741	0.1879	0.2025	-	23-77

N = 500

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1718	-0.0015	0.0347	0.0338	0.1486	0.1711	0.1952	-	53-47
2SLS	0.1731	-0.0002	0.0525	0.0517	0.1382	0.1732	0.2077	-	50-50
Generated Instr. 2SLS	0.1720	-0.0012	0.0494	0.0486	0.1390	0.1728	0.2038	-	50-50
Recursive Biprobit	0.1715	-0.0017	0.0449	0.0442	0.1417	0.1706	0.2022	-	52-48
Special Reg. KeDe AIF	0.1667	-0.0066	0.0553	0.0552	0.1268	0.1678	0.2060	-	52-48
Special Reg. SoDa AIF	0.1608	-0.0125	0.0657	0.0663	0.1130	0.1629	0.2080	-	56-44
Special Reg. KeDe ASF	0.1713	-0.0020	0.0572	0.0566	0.1309	0.1743	0.2133	-	50-50
Special Reg. SoDa ASF	0.1653	-0.0080	0.0680	0.0678	0.1163	0.1662	0.2140	-	54-46
Probit	0.1878	0.0145	0.0313	0.0335	0.1661	0.1874	0.2083	-	31-69

N = 250

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1718	-0.0014	0.0480	0.0473	0.1400	0.1728	0.2055	-	51-49
2SLS	0.1704	-0.0028	0.0718	0.0712	0.1208	0.1695	0.2187	-	53-47
Generated Instr. 2SLS	0.1712	-0.0020	0.0673	0.0663	0.1252	0.1722	0.2185	-	52-48
Recursive Biprobit	0.1697	-0.0036	0.0636	0.0628	0.1248	0.1685	0.2153	-	53-47
Special Reg. KeDe AIF	0.1601	-0.0131	0.0759	0.0767	0.1095	0.1607	0.2162	-	56-44
Special Reg. SoDa AIF	0.1562	-0.0170	0.0923	0.0935	0.0928	0.1587	0.2218	-	55-45
Special Reg. KeDe ASF	0.1673	-0.0060	0.0799	0.0797	0.1135	0.1664	0.2230	-	53-47
Special Reg. SoDa ASF	0.1632	-0.0101	0.0971	0.0972	0.0963	0.1636	0.2326	-	53-47
Probit	0.1872	0.0140	0.0442	0.0453	0.1580	0.1873	0.2166	-	38-62

corr(u_1,u_2) = 0.9 — *true* APE_{y_2} = 0.173

N = 10000

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1728	-0.0003	0.0053	0.0050	0.1693	0.1729	0.1763	-	52-48
2SLS	0.1776	0.0046	0.0115	0.0122	0.1701	0.1779	0.1854	-	33-67
Generated Instr. 2SLS	0.1789	0.0058	0.0106	0.0119	0.1721	0.1790	0.1861	-	27-73

-continued on next page-

- Table A4 continued -

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Recursive Biprobit	0.1728	-0.0002	0.0079	0.0077	0.1678	0.1732	0.1779	-	50-50
Special Reg. KeDe AIF	0.1715	-0.0016	0.0129	0.0128	0.1630	0.1715	0.1805	-	56-44
Special Reg. SoDa AIF	0.1705	-0.0026	0.0165	0.0165	0.1591	0.1707	0.1822	-	56-44
Special Reg. KeDe ASF	0.1936	0.0206	0.0179	0.0272	0.1813	0.1941	0.2056	-	13-87
Special Reg. SoDa ASF	0.1925	0.0194	0.0215	0.0289	0.1783	0.1926	0.2067	-	18-82
Probit	0.3137	0.1406	0.0067	0.1408	0.3093	0.3136	0.3182	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1729	-0.0002	0.0080	0.0077	0.1675	0.1729	0.1781	-	51-49
2SLS	0.1777	0.0046	0.0167	0.0171	0.1671	0.1781	0.1889	-	39-61
Generated Instr. 2SLS	0.1789	0.0058	0.0156	0.0165	0.1684	0.1795	0.1896	-	35-65
Recursive Biprobit	0.1731	-0.0000	0.0120	0.0118	0.1654	0.1737	0.1809	-	48-52
Special Reg. KeDe AIF	0.1719	-0.0012	0.0189	0.0188	0.1589	0.1720	0.1850	-	52-48
Special Reg. SoDa AIF	0.1701	-0.0030	0.0240	0.0242	0.1538	0.1698	0.1861	-	56-44
Special Reg. KeDe ASF	0.1720	-0.0011	0.0193	0.0193	0.1582	0.1722	0.1850	-	52-48
Special Reg. SoDa ASF	0.1702	-0.0029	0.0244	0.0245	0.1537	0.1695	0.1860	-	56-44
Probit	0.3135	0.1404	0.0097	0.1407	0.3071	0.3135	0.3202	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1738	0.0007	0.0179	0.0172	0.1614	0.1736	0.1851	-	49-51
2SLS	0.1794	0.0064	0.0377	0.0377	0.1551	0.1806	0.2055	-	42-58
Generated Instr. 2SLS	0.1812	0.0082	0.0350	0.0353	0.1589	0.1819	0.2053	-	38-62
Recursive Biprobit	0.1738	0.0008	0.0263	0.0257	0.1541	0.1743	0.1919	-	49-51
Special Reg. KeDe AIF	0.1710	-0.0020	0.0428	0.0425	0.1424	0.1718	0.2003	-	51-49
Special Reg. SoDa AIF	0.1698	-0.0032	0.0531	0.0529	0.1359	0.1712	0.2050	-	52-48
Special Reg. KeDe ASF	0.1722	-0.0008	0.0432	0.0429	0.1438	0.1726	0.2019	-	51-49
Special Reg. SoDa ASF	0.1711	-0.0019	0.0537	0.0534	0.1371	0.1721	0.2054	-	51-49
Probit	0.3145	0.1414	0.0214	0.1429	0.2985	0.3142	0.3298	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1724	-0.0009	0.0258	0.0249	0.1558	0.1720	0.1892	-	53-47
2SLS	0.1760	0.0027	0.0513	0.0507	0.1420	0.1764	0.2089	-	48-52
Generated Instr. 2SLS	0.1776	0.0043	0.0484	0.0478	0.1468	0.1770	0.2100	-	47-53
Recursive Biprobit	0.1721	-0.0012	0.0377	0.0371	0.1471	0.1714	0.1968	-	52-48
Special Reg. KeDe AIF	0.1679	-0.0054	0.0576	0.0575	0.1287	0.1699	0.2089	-	53-47
Special Reg. SoDa AIF	0.1654	-0.0079	0.0713	0.0713	0.1171	0.1668	0.2176	-	53-47
Special Reg. KeDe ASF	0.1710	-0.0023	0.0586	0.0583	0.1303	0.1739	0.2136	-	50-50
Special Reg. SoDa ASF	0.1685	-0.0048	0.0728	0.0726	0.1183	0.1704	0.2207	-	51-49
Probit	0.3141	0.1408	0.0304	0.1436	0.2935	0.3156	0.3341	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1728	-0.0004	0.0355	0.0344	0.1478	0.1724	0.1966	-	50-50
2SLS	0.1733	0.0001	0.0716	0.0709	0.1259	0.1719	0.2243	-	50-50
Generated Instr. 2SLS	0.1767	0.0035	0.0658	0.0647	0.1330	0.1783	0.2215	-	47-53
Recursive Biprobit	0.1762	0.0029	0.0503	0.0497	0.1425	0.1738	0.2089	-	48-52
Special Reg. KeDe AIF	0.1610	-0.0123	0.0813	0.0819	0.1066	0.1619	0.2189	-	56-44
Special Reg. SoDa AIF	0.1551	-0.0181	0.0965	0.0981	0.0903	0.1571	0.2209	-	56-44
Special Reg. KeDe ASF	0.1667	-0.0066	0.0843	0.0843	0.1108	0.1658	0.2272	-	53-47
Special Reg. SoDa ASF	0.1610	-0.0122	0.1010	0.1018	0.0954	0.1634	0.2294	-	54-46
Probit	0.3139	0.1406	0.0423	0.1462	0.2857	0.3138	0.3435	***	100

"Reflecting endogeneity" plugs u_2 as additional covariate in the structural equation.
BIAS = average difference between the estimated and the true APE; SD = standard deviation; RMSE = root mean squared error; LQ = lower quartile; UP = upper quartile; Difftest = significance of the t-test of $H_0 : \mathbb{E}[\widehat{APE}] = \mathbb{E}[APE^{true}]$ with - : $p \leq 0.1$, * : $p < 0.1$, ** : $p < 0.05$, *** : $p < 0.01$, Biasratio describes the ratio of positive vs. negative bias
Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $\mathbb{C}_\theta = Gaussian$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$

Table A5: Simulation Results for APE of y_2 - Different joint CDF (All statistics)

$\mathbb{C}_\theta$ = *Clayton*			*true* $APE_{y_2} = 0.173$						
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1727	-0.0004	0.0068	0.0066	0.1681	0.1726	0.1774	-	52-48
2SLS	0.1753	0.0022	0.0111	0.0112	0.1681	0.1752	0.1826	-	41-59
Generated Instr. 2SLS	0.1754	0.0023	0.0104	0.0104	0.1685	0.1753	0.1821	-	42-58
Recursive Biprobit	0.1749	0.0018	0.0090	0.0090	0.1686	0.1746	0.1806	-	43-57
Special Reg. KeDe AIF	0.1711	-0.0020	0.0125	0.0125	0.1626	0.1712	0.1796	-	56-44
Special Reg. SoDa AIF	0.1698	-0.0033	0.0159	0.0162	0.1580	0.1702	0.1804	-	57-43
Special Reg. KeDe ASF	0.1686	-0.0045	0.0134	0.0140	0.1597	0.1693	0.1778	-	62-38
Special Reg. SoDa ASF	0.1673	-0.0058	0.0166	0.0175	0.1555	0.1675	0.1781	-	62-38
Probit	0.2705	0.0974	0.0068	0.0977	0.2656	0.2706	0.2749	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1727	-0.0004	0.0092	0.0091	0.1661	0.1729	0.1790	-	52-48
2SLS	0.1744	0.0014	0.0157	0.0157	0.1635	0.1746	0.1842	-	47-53
Generated Instr. 2SLS	0.1748	0.0017	0.0144	0.0144	0.1652	0.1745	0.1852	-	46-54
Recursive Biprobit	0.1747	0.0016	0.0126	0.0126	0.1659	0.1743	0.1830	-	46-54
Special Reg. KeDe AIF	0.1709	-0.0022	0.0181	0.0182	0.1593	0.1705	0.1827	-	55-45
Special Reg. SoDa AIF	0.1698	-0.0033	0.0241	0.0243	0.1544	0.1702	0.1868	-	55-45
Special Reg. KeDe ASF	0.1701	-0.0030	0.0182	0.0184	0.1583	0.1699	0.1819	-	57-43
Special Reg. SoDa ASF	0.1691	-0.0040	0.0242	0.0245	0.1530	0.1692	0.1859	-	56-44
Probit	0.2707	0.0976	0.0092	0.0980	0.2643	0.2705	0.2768	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1723	-0.0007	0.0219	0.0213	0.1568	0.1722	0.1868	-	52-48
2SLS	0.1749	0.0019	0.0387	0.0383	0.1484	0.1749	0.2014	-	48-52
Generated Instr. 2SLS	0.1752	0.0022	0.0352	0.0348	0.1505	0.1755	0.1997	-	47-53
Recursive Biprobit	0.1736	0.0006	0.0295	0.0290	0.1532	0.1731	0.1939	-	50-50
Special Reg. KeDe AIF	0.1681	-0.0049	0.0410	0.0410	0.1410	0.1681	0.1972	-	55-45
Special Reg. SoDa AIF	0.1656	-0.0074	0.0501	0.0504	0.1329	0.1679	0.2008	-	54-46
Special Reg. KeDe ASF	0.1697	-0.0034	0.0415	0.0414	0.1415	0.1693	0.1992	-	54-46
Special Reg. SoDa ASF	0.1672	-0.0059	0.0509	0.0510	0.1339	0.1684	0.2024	-	53-47
Probit	0.2703	0.0973	0.0217	0.0994	0.2552	0.2701	0.2846	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1719	-0.0014	0.0312	0.0306	0.1513	0.1714	0.1937	-	52-48
2SLS	0.1735	0.0002	0.0514	0.0510	0.1374	0.1747	0.2085	-	49-51
Generated Instr. 2SLS	0.1744	0.0011	0.0484	0.0479	0.1410	0.1767	0.2070	-	48-52
Recursive Biprobit	0.1722	-0.0011	0.0426	0.0420	0.1443	0.1710	0.2023	-	52-48
Special Reg. KeDe AIF	0.1674	-0.0059	0.0553	0.0553	0.1294	0.1694	0.2058	-	53-47
Special Reg. SoDa AIF	0.1633	-0.0100	0.0707	0.0711	0.1172	0.1632	0.2112	-	55-45
Special Reg. KeDe ASF	0.1708	-0.0024	0.0568	0.0565	0.1300	0.1724	0.2077	-	52-48
Special Reg. SoDa ASF	0.1670	-0.0063	0.0729	0.0728	0.1203	0.1667	0.2143	-	54-46
Probit	0.2703	0.0970	0.0314	0.1016	0.2479	0.2702	0.2917	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1734	0.0002	0.0436	0.0424	0.1438	0.1723	0.2011	-	49-51
2SLS	0.1768	0.0036	0.0695	0.0687	0.1317	0.1768	0.2241	-	47-53
Generated Instr. 2SLS	0.1784	0.0052	0.0666	0.0658	0.1324	0.1781	0.2237	-	47-53
Recursive Biprobit	0.1744	0.0011	0.0594	0.0586	0.1337	0.1746	0.2123	-	49-51
Special Reg. KeDe AIF	0.1660	-0.0072	0.0776	0.0776	0.1139	0.1662	0.2210	-	54-46
Special Reg. SoDa AIF	0.1641	-0.0091	0.0940	0.0943	0.1039	0.1666	0.2273	-	54-46
Special Reg. KeDe ASF	0.1722	-0.0011	0.0811	0.0807	0.1162	0.1729	0.2299	-	50-50
Special Reg. SoDa ASF	0.1704	-0.0028	0.0990	0.0988	0.1060	0.1696	0.2353	-	51-49
Probit	0.2703	0.0971	0.0433	0.1055	0.2395	0.2690	0.3013	**	1-99

$\mathbb{C}_\theta$ = *Frank*			*true* $APE_{y_2} = 0.173$						
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1744	0.0014	0.0068	0.0067	0.1699	0.1743	0.1789	-	42-58
2SLS	0.1754	0.0024	0.0110	0.0111	0.1684	0.1755	0.1826	-	41-59
Generated Instr. 2SLS	0.1756	0.0025	0.0102	0.0103	0.1690	0.1757	0.1823	-	39-61

-continued on next page-

Recursive Biprobit	0.1745	0.0014	0.0087	0.0086	0.1687	0.1741	0.1797	-	44-56
Special Reg. KeDe AIF	0.1711	-0.0019	0.0124	0.0124	0.1627	0.1710	0.1796	-	57-43
Special Reg. SoDa AIF	0.1696	-0.0034	0.0157	0.0160	0.1585	0.1694	0.1806	-	59-41
Special Reg. KeDe ASF	0.1705	-0.0025	0.0125	0.0126	0.1620	0.1704	0.1788	-	59-41
Special Reg. SoDa ASF	0.1690	-0.0040	0.0158	0.0161	0.1577	0.1688	0.1798	-	60-40
Probit	0.2688	0.0958	0.0068	0.0960	0.2643	0.2686	0.2731	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1745	0.0014	0.0097	0.0096	0.1676	0.1746	0.1809	-	44-56
2SLS	0.1749	0.0018	0.0156	0.0157	0.1636	0.1747	0.1853	-	47-53
Generated Instr. 2SLS	0.1754	0.0023	0.0144	0.0145	0.1656	0.1752	0.1851	-	44-56
Recursive Biprobit	0.1743	0.0012	0.0129	0.0129	0.1653	0.1744	0.1830	-	46-54
Special Reg. KeDe AIF	0.1713	-0.0018	0.0181	0.0181	0.1594	0.1713	0.1829	-	55-45
Special Reg. SoDa AIF	0.1693	-0.0038	0.0227	0.0230	0.1553	0.1690	0.1854	-	57-43
Special Reg. KeDe ASF	0.1710	-0.0021	0.0182	0.0182	0.1590	0.1710	0.1827	-	55-45
Special Reg. SoDa ASF	0.1691	-0.0040	0.0228	0.0232	0.1551	0.1687	0.1853	-	57-42
Probit	0.2687	0.0957	0.0098	0.0961	0.2621	0.2687	0.2756	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1741	0.0011	0.0226	0.0220	0.1592	0.1737	0.1890	-	49-51
2SLS	0.1753	0.0023	0.0385	0.0381	0.1483	0.1743	0.2031	-	49-51
Generated Instr. 2SLS	0.1758	0.0028	0.0352	0.0347	0.1522	0.1757	0.1996	-	47-53
Recursive Biprobit	0.1741	0.0011	0.0299	0.0294	0.1545	0.1724	0.1938	-	50-50
Special Reg. KeDe AIF	0.1686	-0.0045	0.0411	0.0410	0.1402	0.1688	0.1986	-	53-47
Special Reg. SoDa AIF	0.1662	-0.0068	0.0508	0.0510	0.1324	0.1673	0.2017	-	53-47
Special Reg. KeDe ASF	0.1704	-0.0026	0.0416	0.0414	0.1417	0.1704	0.1998	-	52-48
Special Reg. SoDa ASF	0.1681	-0.0050	0.0517	0.0517	0.1337	0.1692	0.2048	-	52-48
Probit	0.2684	0.0954	0.0218	0.0976	0.2537	0.2682	0.2828	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1744	0.0011	0.0307	0.0300	0.1549	0.1746	0.1950	-	49-51
2SLS	0.1743	0.0010	0.0519	0.0514	0.1395	0.1760	0.2098	-	48-52
Generated Instr. 2SLS	0.1756	0.0023	0.0489	0.0483	0.1431	0.1768	0.2087	-	47-53
Recursive Biprobit	0.1727	-0.0006	0.0430	0.0424	0.1428	0.1734	0.1998	-	49-51
Special Reg. KeDe AIF	0.1682	-0.0051	0.0557	0.0556	0.1298	0.1702	0.2069	-	52-48
Special Reg. SoDa AIF	0.1637	-0.0096	0.0684	0.0688	0.1216	0.1652	0.2099	-	56-44
Special Reg. KeDe ASF	0.1716	-0.0017	0.0569	0.0566	0.1325	0.1724	0.2108	-	50-50
Special Reg. SoDa ASF	0.1674	-0.0058	0.0703	0.0702	0.1242	0.1681	0.2159	-	54-46
Probit	0.2694	0.0961	0.0307	0.1004	0.2487	0.2699	0.2883	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1743	0.0010	0.0437	0.0424	0.1457	0.1732	0.2019	-	50-50
2SLS	0.1763	0.0030	0.0696	0.0687	0.1295	0.1767	0.2250	-	48-52
Generated Instr. 2SLS	0.1778	0.0046	0.0661	0.0653	0.1335	0.1784	0.2210	-	46-54
Recursive Biprobit	0.1748	0.0016	0.0582	0.0574	0.1365	0.1726	0.2131	-	50-50
Special Reg. KeDe AIF	0.1661	-0.0071	0.0777	0.0776	0.1134	0.1675	0.2208	-	53-47
Special Reg. SoDa AIF	0.1600	-0.0133	0.0909	0.0915	0.0997	0.1636	0.2222	-	55-45
Special Reg. KeDe ASF	0.1714	-0.0019	0.0805	0.0801	0.1152	0.1709	0.2285	-	50-50
Special Reg. SoDa ASF	0.1660	-0.0073	0.0951	0.0950	0.1018	0.1679	0.2303	-	53-47
Probit	0.2692	0.0959	0.0434	0.1043	0.2404	0.2689	0.2998	**	1-99
$\mathbb{C}_\theta$ = *Gumbel*				true $APE_{y_2} = 0.173$					
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1741	0.0010	0.0066	0.0066	0.1697	0.1740	0.1787	-	46-54
2SLS	0.1760	0.0030	0.0110	0.0113	0.1685	0.1760	0.1835	-	42-58
Generated Instr. 2SLS	0.1768	0.0037	0.0104	0.0109	0.1699	0.1768	0.1837	-	36-64
Recursive Biprobit	0.1755	0.0025	0.0089	0.0091	0.1692	0.1750	0.1817	-	40-60
Special Reg. KeDe AIF	0.1713	-0.0017	0.0126	0.0127	0.1628	0.1710	0.1798	-	56-44
Special Reg. SoDa AIF	0.1703	-0.0027	0.0162	0.0164	0.1597	0.1698	0.1811	-	58-42
Special Reg. KeDe ASF	0.1709	-0.0021	0.0127	0.0128	0.1623	0.1705	0.1789	-	57-42
Special Reg. SoDa ASF	0.1699	-0.0031	0.0163	0.0166	0.1592	0.1692	0.1806	-	58-42
Probit	0.2642	0.0912	0.0067	0.0914	0.2598	0.2639	0.2687	***	100
N = 5000									

-continued on next page-

- Table A5 continued -

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1741	0.0010	0.0103	0.0101	0.1669	0.1743	0.1811	-	45-55
2SLS	0.1760	0.0029	0.0165	0.0166	0.1649	0.1766	0.1868	-	42-58
Generated Instr. 2SLS	0.1769	0.0038	0.0151	0.0154	0.1672	0.1775	0.1875	-	40-60
Recursive Biprobit	0.1757	0.0026	0.0135	0.0136	0.1672	0.1757	0.1841	-	42-58
Special Reg. KeDe AIF	0.1715	-0.0016	0.0190	0.0189	0.1591	0.1719	0.1839	-	53-47
Special Reg. SoDa AIF	0.1701	-0.0030	0.0239	0.0240	0.1533	0.1696	0.1866	-	55-45
Special Reg. KeDe ASF	0.1715	-0.0016	0.0191	0.0190	0.1590	0.1716	0.1844	-	53-47
Special Reg. SoDa ASF	0.1701	-0.0030	0.0240	0.0241	0.1532	0.1699	0.1868	-	56-44
Probit	0.2641	0.0910	0.0101	0.0915	0.2571	0.2643	0.2709	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1742	0.0011	0.0212	0.0207	0.1596	0.1740	0.1882	-	47-53
2SLS	0.1778	0.0048	0.0352	0.0351	0.1543	0.1796	0.2024	-	43-57
Generated Instr. 2SLS	0.1783	0.0053	0.0327	0.0326	0.1554	0.1800	0.2008	-	42-58
Recursive Biprobit	0.1761	0.0031	0.0290	0.0287	0.1562	0.1758	0.1953	-	46-54
Special Reg. KeDe AIF	0.1698	-0.0032	0.0391	0.0390	0.1431	0.1692	0.1961	-	54-46
Special Reg. SoDa AIF	0.1665	-0.0065	0.0481	0.0484	0.1330	0.1675	0.1989	-	55-45
Special Reg. KeDe ASF	0.1716	-0.0015	0.0399	0.0397	0.1444	0.1706	0.1982	-	53-47
Special Reg. SoDa ASF	0.1684	-0.0047	0.0492	0.0493	0.1341	0.1689	0.2009	-	53-47
Probit	0.2643	0.0913	0.0202	0.0933	0.2510	0.2643	0.2779	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1756	0.0023	0.0311	0.0303	0.1549	0.1763	0.1957	-	46-54
2SLS	0.1797	0.0064	0.0528	0.0528	0.1448	0.1790	0.2164	-	46-54
Generated Instr. 2SLS	0.1796	0.0063	0.0486	0.0484	0.1480	0.1806	0.2130	-	45-55
Recursive Biprobit	0.1769	0.0037	0.0421	0.0419	0.1475	0.1768	0.2055	-	46-54
Special Reg. KeDe AIF	0.1724	-0.0009	0.0571	0.0569	0.1337	0.1735	0.2116	-	49-51
Special Reg. SoDa AIF	0.1677	-0.0056	0.0696	0.0697	0.1258	0.1707	0.2173	-	51-49
Special Reg. KeDe ASF	0.1763	0.0030	0.0586	0.0584	0.1373	0.1766	0.2155	-	47-53
Special Reg. SoDa ASF	0.1718	-0.0015	0.0717	0.0716	0.1274	0.1745	0.2217	-	50-50
Probit	0.2655	0.0922	0.0305	0.0966	0.2447	0.2661	0.2860	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1743	0.0010	0.0437	0.0429	0.1436	0.1717	0.2036	-	50-50
2SLS	0.1727	-0.0005	0.0699	0.0694	0.1245	0.1761	0.2205	-	49-51
Generated Instr. 2SLS	0.1761	0.0028	0.0641	0.0637	0.1316	0.1770	0.2231	-	48-52
Recursive Biprobit	0.1738	0.0005	0.0578	0.0573	0.1349	0.1733	0.2153	-	50-50
Special Reg. KeDe AIF	0.1622	-0.0110	0.0790	0.0796	0.1130	0.1632	0.2135	-	55-45
Special Reg. SoDa AIF	0.1568	-0.0164	0.0928	0.0945	0.0911	0.1555	0.2226	-	57-43
Special Reg. KeDe ASF	0.1691	-0.0041	0.0833	0.0833	0.1169	0.1704	0.2238	-	51-49
Special Reg. SoDa ASF	0.1644	-0.0088	0.1009	0.1015	0.0965	0.1622	0.2320	-	55-45
Probit	0.2644	0.0912	0.0420	0.0996	0.2333	0.2655	0.2936	**	1-99
$\mathbb{C}_\theta = t$				*true* $APE_{y_2} = 0.173$					
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1756	0.0025	0.0069	0.0071	0.1711	0.1753	0.1802	-	35-65
2SLS	0.1763	0.0032	0.0116	0.0119	0.1680	0.1763	0.1841	-	39-61
Generated Instr. 2SLS	0.1771	0.0040	0.0108	0.0113	0.1697	0.1772	0.1847	-	36-64
Recursive Biprobit	0.1837	0.0107	0.0091	0.0139	0.1775	0.1838	0.1903	-	13-87
Special Reg. KeDe AIF	0.1707	-0.0024	0.0128	0.0129	0.1622	0.1708	0.1792	-	57-43
Special Reg. SoDa AIF	0.1698	-0.0032	0.0165	0.0167	0.1591	0.1703	0.1813	-	57-43
Special Reg. KeDe ASF	0.1703	-0.0028	0.0129	0.0130	0.1617	0.1704	0.1786	-	58-42
Special Reg. SoDa ASF	0.1694	-0.0037	0.0165	0.0168	0.1586	0.1698	0.1809	-	58-42
Probit	0.2616	0.0885	0.0068	0.0888	0.2569	0.2614	0.2659	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1760	0.0029	0.0096	0.0098	0.1692	0.1758	0.1827	-	38-62
2SLS	0.1764	0.0033	0.0160	0.0162	0.1661	0.1762	0.1875	-	41-59
Generated Instr. 2SLS	0.1773	0.0042	0.0147	0.0150	0.1682	0.1772	0.1874	-	38-62
Recursive Biprobit	0.1838	0.0107	0.0129	0.0166	0.1753	0.1840	0.1923	-	20-80
Special Reg. KeDe AIF	0.1715	-0.0016	0.0179	0.0178	0.1604	0.1717	0.1836	-	54-46
Special Reg. SoDa AIF	0.1695	-0.0036	0.0227	0.0229	0.1551	0.1698	0.1848	-	56-44
Special Reg. KeDe ASF	0.1714	-0.0017	0.0180	0.0180	0.1600	0.1714	0.1836	-	54-46

-continued on next page-

- Table A5 continued -

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Special Reg. SoDa ASF	0.1694	-0.0037	0.0228	0.0230	0.1549	0.1699	0.1844	-	57-43
Probit	0.2614	0.0883	0.0097	0.0888	0.2546	0.2616	0.2680	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1761	0.0030	0.0224	0.0221	0.1608	0.1769	0.1904	-	44-56
2SLS	0.1778	0.0047	0.0374	0.0372	0.1526	0.1768	0.2039	-	45-55
Generated Instr. 2SLS	0.1795	0.0065	0.0353	0.0353	0.1569	0.1798	0.2035	-	42-58
Recursive Biprobit	0.1842	0.0112	0.0299	0.0314	0.1645	0.1846	0.2057	-	35-65
Special Reg. KeDe AIF	0.1697	-0.0033	0.0414	0.0413	0.1400	0.1712	0.1993	-	52-48
Special Reg. SoDa AIF	0.1676	-0.0054	0.0500	0.0500	0.1319	0.1701	0.2019	-	52-48
Special Reg. KeDe ASF	0.1715	-0.0015	0.0420	0.0418	0.1425	0.1720	0.2013	-	51-49
Special Reg. SoDa ASF	0.1696	-0.0035	0.0510	0.0508	0.1335	0.1721	0.2043	-	51-49
Probit	0.2622	0.0892	0.0219	0.0916	0.2477	0.2619	0.2777	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1758	0.0025	0.0319	0.0308	0.1544	0.1755	0.1951	-	47-53
2SLS	0.1751	0.0018	0.0510	0.0502	0.1410	0.1749	0.2092	-	50-50
Generated Instr. 2SLS	0.1761	0.0029	0.0486	0.0478	0.1428	0.1749	0.2072	-	49-51
Recursive Biprobit	0.1806	0.0073	0.0426	0.0423	0.1520	0.1789	0.2088	-	43-57
Special Reg. KeDe AIF	0.1669	-0.0064	0.0562	0.0560	0.1288	0.1677	0.2067	-	54-46
Special Reg. SoDa AIF	0.1610	-0.0122	0.0673	0.0680	0.1167	0.1619	0.2053	-	56-44
Special Reg. KeDe ASF	0.1706	-0.0027	0.0577	0.0573	0.1324	0.1728	0.2122	-	51-49
Special Reg. SoDa ASF	0.1647	-0.0086	0.0692	0.0694	0.1193	0.1656	0.2123	-	55-45
Probit	0.2612	0.0879	0.0312	0.0928	0.2399	0.2618	0.2825	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1770	0.0038	0.0440	0.0429	0.1475	0.1755	0.2073	-	48-52
2SLS	0.1734	0.0002	0.0733	0.0725	0.1214	0.1729	0.2220	-	50-50
Generated Instr. 2SLS	0.1768	0.0035	0.0675	0.0665	0.1297	0.1775	0.2214	-	47-53
Recursive Biprobit	0.1791	0.0059	0.0607	0.0600	0.1367	0.1777	0.2195	-	47-53
Special Reg. KeDe AIF	0.1608	-0.0124	0.0787	0.0794	0.1065	0.1623	0.2169	-	56-44
Special Reg. SoDa AIF	0.1582	-0.0150	0.0928	0.0939	0.0969	0.1625	0.2249	-	54-46
Special Reg. KeDe ASF	0.1666	-0.0066	0.0826	0.0826	0.1098	0.1687	0.2225	-	53-47
Special Reg. SoDa ASF	0.1642	-0.0090	0.0977	0.0980	0.1003	0.1676	0.2309	-	52-48
Probit	0.2626	0.0894	0.0444	0.0989	0.2327	0.2622	0.2925	**	2-98

"Reflecting endogeneity" plugs u_2 as additional covariate in the structural equation.
BIAS = average difference between the estimated and the true APE; SD = standard deviation; RMSE = root mean squared error; LQ = lower quartile; UP = upper quartile; Difftest = significance of the t-test of $H_0 : \mathbb{E}[\widehat{APE}] = \mathbb{E}[APE^{true}]$ with - : p $\leq$ 0.1 , * : p $<$ 0.1 , ** : p $<$ 0.05, *** : p $<$ 0.01, Biasratio describes the ratio of positive vs. negative bias
Scenario characteristics: $u_1 \sim Normal(0,1)$, $u_2 \sim Normal(0,1)$, $corr(u_1,u_2) = 0.6$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$

Table A6: Simulation Results for APE of y_2 - Different Marginal CDF (All statistics)

$u_1 \sim F(10,6)$ $u_2 \sim F(10,6)$			true $APE_{y_2} = 0.172$						
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1899	0.0179	0.0081	0.0195	0.1840	0.1898	0.1955	**	100
2SLS	0.1813	0.0093	0.0119	0.0149	0.1729	0.1819	0.1891	-	21-79
Generated Instr. 2SLS	0.1819	0.0099	0.0108	0.0145	0.1747	0.1824	0.1888	-	17-83
Recursive Biprobit	0.1775	0.0055	0.0096	0.0109	0.1714	0.1772	0.1839	-	26-74
Special Reg. KeDe AIF	0.1745	0.0025	0.0358	0.0359	0.1608	0.1746	0.1895	-	45-55
Special Reg. SoDa AIF	0.1732	0.0012	0.0363	0.0363	0.1559	0.1734	0.1900	-	48-52
Special Reg. KeDe ASF	0.2080	0.0360	0.0482	0.0601	0.1909	0.2075	0.2253	-	9-91
Special Reg. SoDa ASF	0.2067	0.0346	0.0460	0.0575	0.1846	0.2059	0.2272	-	14-86
Probit	0.2626	0.0906	0.0071	0.0908	0.2579	0.2627	0.2676	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1898	0.0176	0.0116	0.0211	0.1816	0.1901	0.1974	-	5-95
2SLS	0.1814	0.0092	0.0164	0.0186	0.1705	0.1819	0.1927	-	28-72
Generated Instr. 2SLS	0.1818	0.0097	0.0149	0.0176	0.1718	0.1826	0.1916	-	25-75
Recursive Biprobit	0.1778	0.0056	0.0134	0.0145	0.1691	0.1782	0.1873	-	32-68
Special Reg. KeDe AIF	0.1757	0.0035	0.0384	0.0384	0.1568	0.1745	0.1933	-	47-53
Special Reg. SoDa AIF	0.1741	0.0019	0.0443	0.0442	0.1517	0.1734	0.1961	-	49-51
Special Reg. KeDe ASF	0.2102	0.0380	0.0565	0.0680	0.1861	0.2081	0.2311	-	12-88
Special Reg. SoDa ASF	0.2086	0.0364	0.0570	0.0675	0.1797	0.2065	0.2342	-	20-80
Probit	0.2624	0.0902	0.0098	0.0907	0.2556	0.2624	0.2691	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1892	0.0174	0.0240	0.0293	0.1741	0.1892	0.2050	-	23-77
2SLS	0.1828	0.0109	0.0362	0.0373	0.1579	0.1838	0.2077	-	37-63
Generated Instr. 2SLS	0.1834	0.0116	0.0330	0.0345	0.1625	0.1830	0.2072	-	36-64
Recursive Biprobit	0.1772	0.0053	0.0296	0.0298	0.1581	0.1775	0.1963	-	41-59
Special Reg. KeDe AIF	0.1752	0.0034	0.0589	0.0589	0.1403	0.1735	0.2080	-	49-51
Special Reg. SoDa AIF	0.1717	-0.0002	0.0686	0.0686	0.1306	0.1695	0.2139	-	51-49
Special Reg. KeDe ASF	0.2106	0.0387	0.0797	0.0886	0.1654	0.2044	0.2489	-	28-72
Special Reg. SoDa ASF	0.2067	0.0349	0.0914	0.0979	0.1537	0.2000	0.2528	-	34-66
Probit	0.2629	0.0910	0.0219	0.0935	0.2482	0.2621	0.2768	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1832	0.0112	0.0333	0.0346	0.1606	0.1821	0.2049	-	36-64
2SLS	0.1792	0.0072	0.0535	0.0533	0.1421	0.1801	0.2115	-	44-56
Generated Instr. 2SLS	0.1787	0.0067	0.0488	0.0485	0.1456	0.1779	0.2110	-	45-55
Recursive Biprobit	0.1733	0.0013	0.0438	0.0432	0.1427	0.1736	0.2017	-	48-52
Special Reg. KeDe AIF	0.1659	-0.0060	0.0776	0.0775	0.1221	0.1676	0.2165	-	53-47
Special Reg. SoDa AIF	0.1621	-0.0099	0.0877	0.0879	0.1050	0.1648	0.2185	-	53-47
Special Reg. KeDe ASF	0.1905	0.0185	0.1033	0.1047	0.1389	0.1883	0.2431	-	42-58
Special Reg. SoDa ASF	0.1862	0.0142	0.1090	0.1096	0.1206	0.1851	0.2502	-	44-56
Probit	0.2598	0.0878	0.0316	0.0930	0.2389	0.2598	0.2796	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1843	0.0121	0.0457	0.0466	0.1532	0.1847	0.2141	-	40-60
2SLS	0.1784	0.0061	0.0698	0.0693	0.1310	0.1817	0.2260	-	46-54
Generated Instr. 2SLS	0.1801	0.0079	0.0648	0.0642	0.1383	0.1822	0.2244	-	44-56
Recursive Biprobit	0.1737	0.0015	0.0593	0.0588	0.1348	0.1711	0.2140	-	51-49
Special Reg. KeDe AIF	0.1626	-0.0096	0.0892	0.0897	0.1112	0.1689	0.2211	-	52-48
Special Reg. SoDa AIF	0.1541	-0.0182	0.1046	0.1062	0.0888	0.1613	0.2239	-	55-45
Special Reg. KeDe ASF	0.1821	0.0098	0.1146	0.1151	0.1196	0.1825	0.2405	-	45-55
Special Reg. SoDa ASF	0.1730	0.0008	0.1319	0.1320	0.0993	0.1710	0.2507	-	50-50
Probit	0.2618	0.0896	0.0432	0.0990	0.2332	0.2618	0.2916	**	2-98

$u_1 \sim log(0,0.9)$ $u_2 \sim log(0,0.9)$			true $APE_{y_2} = 0.154$						
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio

-continued on next page-

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1565	0.0027	0.0083	0.0087	0.1511	0.1567	0.1619	-	36-64
2SLS	0.1601	0.0064	0.0148	0.0160	0.1505	0.1600	0.1701	-	33-67
Generated Instr. 2SLS	0.1609	0.0071	0.0142	0.0158	0.1514	0.1608	0.1708	-	31-69
Recursive Biprobit	0.1526	-0.0011	0.0132	0.0131	0.1438	0.1526	0.1619	-	55-45
Special Reg. KeDe AIF	0.1560	0.0022	0.0246	0.0246	0.1414	0.1547	0.1696	-	48-52
Special Reg. SoDa AIF	0.1555	0.0018	0.0286	0.0286	0.1373	0.1545	0.1738	-	49-51
Special Reg. KeDe ASF	0.1532	-0.0005	0.0246	0.0246	0.1387	0.1521	0.1664	-	53-47
Special Reg. SoDa ASF	0.1528	-0.0010	0.0283	0.0282	0.1347	0.1517	0.1703	-	53-47
Probit	0.3212	0.1675	0.0073	0.1676	0.3164	0.3212	0.3262	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1563	0.0025	0.0121	0.0123	0.1484	0.1569	0.1642	-	40-60
2SLS	0.1595	0.0057	0.0213	0.0221	0.1450	0.1593	0.1736	-	40-60
Generated Instr. 2SLS	0.1603	0.0066	0.0206	0.0215	0.1463	0.1608	0.1735	-	38-62
Recursive Biprobit	0.1521	-0.0016	0.0192	0.0192	0.1394	0.1516	0.1649	-	54-46
Special Reg. KeDe AIF	0.1545	0.0007	0.0329	0.0328	0.1342	0.1550	0.1734	-	49-51
Special Reg. SoDa AIF	0.1539	0.0001	0.0389	0.0388	0.1302	0.1560	0.1780	-	48-52
Special Reg. KeDe ASF	0.1522	-0.0016	0.0328	0.0328	0.1327	0.1524	0.1709	-	51-49
Special Reg. SoDa ASF	0.1517	-0.0021	0.0386	0.0385	0.1281	0.1534	0.1758	-	51-49
Probit	0.3210	0.1672	0.0108	0.1675	0.3138	0.3211	0.3282	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1570	0.0033	0.0266	0.0266	0.1388	0.1571	0.1754	-	46-54
2SLS	0.1615	0.0077	0.0470	0.0474	0.1295	0.1633	0.1933	-	43-57
Generated Instr. 2SLS	0.1626	0.0088	0.0451	0.0457	0.1311	0.1641	0.1932	-	40-60
Recursive Biprobit	0.1536	-0.0002	0.0416	0.0413	0.1253	0.1531	0.1827	-	51-49
Special Reg. KeDe AIF	0.1524	-0.0014	0.0618	0.0618	0.1123	0.1545	0.1941	-	50-50
Special Reg. SoDa AIF	0.1514	-0.0024	0.0748	0.0747	0.1021	0.1539	0.1988	-	50-50
Special Reg. KeDe ASF	0.1522	-0.0016	0.0630	0.0630	0.1123	0.1546	0.1937	-	50-50
Special Reg. SoDa ASF	0.1514	-0.0024	0.0772	0.0771	0.1022	0.1541	0.1983	-	50-50
Probit	0.3211	0.1673	0.0234	0.1689	0.3049	0.3210	0.3363	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1564	0.0024	0.0398	0.0394	0.1286	0.1563	0.1808	-	48-52
2SLS	0.1567	0.0028	0.0665	0.0662	0.1101	0.1551	0.2030	-	50-50
Generated Instr. 2SLS	0.1575	0.0036	0.0641	0.0638	0.1134	0.1545	0.2006	-	49-51
Recursive Biprobit	0.1496	-0.0043	0.0595	0.0594	0.1092	0.1469	0.1901	-	55-45
Special Reg. KeDe AIF	0.1466	-0.0073	0.0827	0.0828	0.0889	0.1500	0.2029	-	52-48
Special Reg. SoDa AIF	0.1431	-0.0108	0.0993	0.0996	0.0770	0.1476	0.2153	-	54-46
Special Reg. KeDe ASF	0.1476	-0.0063	0.0841	0.0842	0.0880	0.1497	0.2047	-	52-48
Special Reg. SoDa ASF	0.1446	-0.0093	0.1028	0.1029	0.0758	0.1464	0.2177	-	53-47
Probit	0.3203	0.1664	0.0347	0.1698	0.2957	0.3196	0.3441	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1574	0.0035	0.0554	0.0551	0.1165	0.1565	0.1950	-	49-51
2SLS	0.1555	0.0016	0.0935	0.0931	0.0933	0.1516	0.2170	-	51-49
Generated Instr. 2SLS	0.1571	0.0031	0.0904	0.0899	0.0962	0.1538	0.2212	-	51-49
Recursive Biprobit	0.1479	-0.0060	0.0863	0.0860	0.0901	0.1476	0.2026	-	53-47
Special Reg. KeDe AIF	0.1429	-0.0110	0.1079	0.1082	0.0666	0.1434	0.2180	-	54-46
Special Reg. SoDa AIF	0.1364	-0.0176	0.1245	0.1255	0.0522	0.1370	0.2244	-	56-44
Special Reg. KeDe ASF	0.1452	-0.0088	0.1137	0.1137	0.0677	0.1458	0.2276	-	53-47
Special Reg. SoDa ASF	0.1383	-0.0156	0.1321	0.1327	0.0548	0.1414	0.2291	-	55-45
Probit	0.3214	0.1675	0.0493	0.1743	0.2868	0.3199	0.3555	***	100
$u_1 \sim t(3)$ $u_2 \sim t(3)$				*true* $APE_{y_2} = 0.161$					
N = 10000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1724	0.0111	0.0085	0.0139	0.1673	0.1726	0.1779	-	7-93
2SLS	0.1669	0.0056	0.0133	0.0143	0.1576	0.1666	0.1760	-	34-66
Generated Instr. 2SLS	0.1676	0.0063	0.0125	0.0138	0.1592	0.1675	0.1760	-	30-70
Recursive Biprobit	0.1586	-0.0026	0.0115	0.0116	0.1512	0.1586	0.1667	-	60-40
Special Reg. KeDe AIF	0.1624	0.0012	0.0283	0.0282	0.1476	0.1607	0.1769	-	50-50
Special Reg. SoDa AIF	0.1607	-0.0006	0.0330	0.0329	0.1423	0.1601	0.1784	-	52-48
Special Reg. KeDe ASF	0.1605	-0.0008	0.0286	0.0285	0.1454	0.1588	0.1747	-	55-45

-continued on next page-

- Table A6 continued -

	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Special Reg. SoDa ASF	0.1589	-0.0024	0.0333	0.0332	0.1400	0.1580	0.1761	-	55-45
Probit	0.3000	0.1387	0.0071	0.1389	0.2952	0.2998	0.3051	***	100
N = 5000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1724	0.0110	0.0127	0.0167	0.1639	0.1723	0.1804	-	16-84
2SLS	0.1669	0.0056	0.0192	0.0199	0.1544	0.1674	0.1800	-	38-62
Generated Instr. 2SLS	0.1675	0.0062	0.0179	0.0189	0.1562	0.1674	0.1794	-	36-64
Recursive Biprobit	0.1585	-0.0028	0.0170	0.0171	0.1473	0.1579	0.1699	-	56-44
Special Reg. KeDe AIF	0.1611	-0.0002	0.0415	0.0414	0.1419	0.1610	0.1796	-	51-49
Special Reg. SoDa AIF	0.1600	-0.0014	0.0452	0.0452	0.1375	0.1596	0.1836	-	52-48
Special Reg. KeDe ASF	0.1595	-0.0018	0.0454	0.0454	0.1405	0.1591	0.1781	-	54-46
Special Reg. SoDa ASF	0.1586	-0.0027	0.0465	0.0465	0.1359	0.1573	0.1814	-	54-46
Probit	0.2994	0.1380	0.0107	0.1384	0.2919	0.2993	0.3066	***	100
N = 1000									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1731	0.0119	0.0269	0.0291	0.1550	0.1733	0.1911	-	34-66
2SLS	0.1677	0.0064	0.0429	0.0430	0.1387	0.1667	0.1981	-	45-55
Generated Instr. 2SLS	0.1692	0.0079	0.0404	0.0407	0.1416	0.1709	0.1979	-	41-59
Recursive Biprobit	0.1590	-0.0023	0.0375	0.0372	0.1337	0.1595	0.1861	-	52-48
Special Reg. KeDe AIF	0.1587	-0.0026	0.0599	0.0598	0.1239	0.1586	0.1960	-	52-48
Special Reg. SoDa AIF	0.1554	-0.0059	0.0727	0.0729	0.1137	0.1567	0.2023	-	53-47
Special Reg. KeDe ASF	0.1595	-0.0018	0.0613	0.0612	0.1241	0.1595	0.1968	-	52-48
Special Reg. SoDa ASF	0.1558	-0.0054	0.0767	0.0769	0.1139	0.1578	0.2029	-	52-48
Probit	0.2995	0.1383	0.0233	0.1401	0.2840	0.2996	0.3151	***	100
N = 500									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1703	0.0088	0.0371	0.0376	0.1441	0.1698	0.1948	-	41-59
2SLS	0.1649	0.0034	0.0598	0.0595	0.1252	0.1630	0.2071	-	49-51
Generated Instr. 2SLS	0.1657	0.0042	0.0562	0.0560	0.1272	0.1646	0.2043	-	48-52
Recursive Biprobit	0.1561	-0.0054	0.0525	0.0525	0.1201	0.1549	0.1921	-	56-44
Special Reg. KeDe AIF	0.1527	-0.0088	0.0782	0.0786	0.1036	0.1555	0.2076	-	53-47
Special Reg. SoDa AIF	0.1509	-0.0106	0.0911	0.0915	0.0891	0.1542	0.2147	-	53-47
Special Reg. KeDe ASF	0.1547	-0.0068	0.0820	0.0822	0.1057	0.1557	0.2092	-	53-47
Special Reg. SoDa ASF	0.1538	-0.0077	0.0949	0.0951	0.0906	0.1575	0.2219	-	52-48
Probit	0.2985	0.1370	0.0334	0.1407	0.2755	0.2980	0.3217	***	100
N = 250									
	Mean	BIAS	SD	RMSE	LQ	Median	UQ	Difftest	Biasratio
Reflecting Endogeneity	0.1699	0.0085	0.0529	0.0530	0.1333	0.1678	0.2043	-	46-54
2SLS	0.1615	0.0000	0.0837	0.0831	0.1042	0.1586	0.2171	-	51-49
Generated Instr. 2SLS	0.1631	0.0016	0.0793	0.0786	0.1084	0.1624	0.2139	-	49-51
Recursive Biprobit	0.1533	-0.0082	0.0762	0.0761	0.1052	0.1544	0.2025	-	54-46
Special Reg. KeDe AIF	0.1475	-0.0140	0.1005	0.1012	0.0792	0.1481	0.2162	-	54-46
Special Reg. SoDa AIF	0.1407	-0.0207	0.1192	0.1208	0.0586	0.1443	0.2275	-	55-45
Special Reg. KeDe ASF	0.1514	-0.0100	0.1086	0.1088	0.0828	0.1504	0.2230	-	53-47
Special Reg. SoDa ASF	0.1458	-0.0156	0.1309	0.1317	0.0613	0.1490	0.2344	-	53-47
Probit	0.2995	0.1380	0.0491	0.1460	0.2654	0.2982	0.3322	***	100

"Reflecting endogeneity" plugs u_2 as additional covariate in the structural equation.
BIAS = average difference between the estimated and the true APE; SD = standard deviation; RMSE = root mean squared error; LQ = lower quartile; UP = upper quartile; Difftest = significance of the t-test of $H_0 : \mathbb{E}[\widehat{APE}] = \mathbb{E}[APE^{true}]$ with - : p $\leq$ 0.1 , * : p $<$ 0.1 , ** : p $<$ 0.05, *** : p $<$ 0.01, Biasratio describes the ratio of positive vs. negative bias
Scenario characteristics: $\mathbb{C}_\theta = Gaussian$, $corr(u_1, u_2) = 0.6$, $SD(v) = 2.0$, $F[\delta_{z_{21}}] \geq 100.00$