

Weak identification  
in probit models with  
endogenous covariates

Jean-Marie Dufour and Joachim Wilde

Working Paper 95  
February 2013

INSTITUTE OF EMPIRICAL ECONOMIC RESEARCH  
University of Osnabrück  
Rolandstraße 8  
49069 Osnabrück  
Germany

# Weak identification in probit models with endogenous covariates

Jean-Marie Dufour  
*McGill University*

Joachim Wilde<sup>1</sup>  
*Osnabrück University*

## **Abstract**

Weak identification is a well known topic for linear multiple equation models. However, little is known whether this problem also matters for probit models with endogenous covariates. Therefore, the behaviour of the usual z-statistic in case of weak identification is analysed in a simulation study. It shows large size distortions. However, a new puzzle is found: The magnitude of the size distortion depends heavily on the parameter value that is tested. Alternatively the LR-statistic was calculated which is known to be more robust against weak identification in case of linear multiple equation models. The same seems to be true for probit equations. No size distortions are found. However, medium undersizing is observed.

Keywords: probit model, weak identification

JEL classification: C 35

---

<sup>1</sup>Corresponding author. Postal address: Fachbereich Wirtschaftswissenschaften, Rolandstr. 8, 49069 Osnabrueck, Germany; phone: ++49 541 969 2746; e-mail: Joachim.Wilde@uni-osnabrueck.de

## 1. Introduction

Probit models are by now widely used in applied econometrics (in recent years, e.g., Brav 2009, Duryea/ Lam/ Levison 2007, Engelhardt et.al. 2010, Esaka 2010, Fitzenberger/ Kohn/ Wang 2011, Hajivassiliou/ Ioannides 2007, Hao/ Ng 2011, Hlaing/ Pourjalali 2012, Kauppi/ Saikkonen 2008, Litchfield/ Reilly/ Veneziani 2012, Mizutani/ Uranishi 2010, Schneider 2007). As in linear models, one or more explaining variables might be endogenous. This problem can be solved by using instrumental variables (cf. Wilde 2008 for a comparison of different estimation methods using instrumental variables). The resulting estimates can be used to calculate test statistics for the parameters of the model.

However, in linear models it is well known that weak instruments may cause considerable size distortions (cf. Dufour 2003 for an overview). Wald-type tests like the usual t-tests and F-tests are especially vulnerable to this problem (cf. Dufour 1997). In probit models a single parameter hypothesis is usually tested by the so-called z-test, i.e. the ratio of a consistent estimate and its asymptotical standard error. This is a Wald-type test. Therefore, big size distortions can be expected. Nevertheless, the topic seems to be largely a white spot in the literature. Exceptions are the recent theoretical papers of Andrews/ Cheng (2011a) and Andrews/ Cheng (2011b), who address the probit model as an example. However, Andrews/ Cheng (2011a) restrict their numerical analysis to a probit model with a nonlinear regression function and without endogeneity, and Andrews/ Cheng (2011b) don't analyze the probit model numerically.<sup>2</sup>

The paper contributes different new aspects: First, large size distortions in probit models with endogeneity are demonstrated with a simulation study. Second, a new puzzle is shown: Whereas size distortions testing the null of a parameter being equal to zero are of medium size, testing other values of the parameter gives large size distortions like those in linear models. Third, the behaviour of the classical likelihood ratio statistic in that case is analysed. Given the simulation design no size distortions are found. However, undersizing is observed so that the power of the test might be low. Fourth, as a by-product some insights concerning the estimation of probit models with endogenous covariates are given.

Section 2 shows the econometric model and the test statistics. Section 3 describes the simulation design and the estimators that are used. Since a probit equation is part of the model, some formulas become more complicated than in the linear case. They are described detailed because textbook descriptions are missing so far. Section 4 presents the results of the simulation study, and section 5 concludes. For ease of exposition always the binary probit model is considered.

---

<sup>2</sup> A further exception is Magnusson (2007), who considered in an early version of his paper the probit model with endogenous covariates as an example and found medium size distortions. However, in later versions of the working paper and in the published version (Magnusson 2010) the probit example was deleted.

## 2. Model and classical tests

The equation of interest is a structural probit equation. One of the explaining variables is endogenous, and a reduced form equation for this variable is specified. Thus, the following type of a model is considered:

$$\begin{aligned} y_{li}^* &= \gamma_1 y_{2i} + \beta_1 x_{li} + u_{li} \\ y_{2i} &= \pi_{21} x_{li} + \pi_{22} x_{2i} + v_{2i} \end{aligned} \quad y_{li} = \begin{cases} 1, & y_{li}^* > 0 \\ 0, & \text{else} \end{cases}, \quad i = 1, \dots, N, \quad (2.1)$$

where  $y_{li}^*$  is a latent variable,  $y_{li}$  is its observable indicator,  $y_{2i}$  is an endogenous and observable variable,  $x_{li}$  and  $x_{2i}$  are  $(K_1 \times 1)$  and  $(K_2 \times 1)$  vectors of exogenous variables,  $\gamma_1$ ,  $\beta_1$ ,  $\pi_{21}$ ,  $\pi_{22}$  are unknown parameter vectors of dimensions 1 ( $\gamma_1$ ),  $K_1$  ( $\beta_1$ ,  $\pi_{21}$ ) and  $K_2$  ( $\pi_{22}$ ), and  $u_{li}$  and  $v_{2i}$  are error terms with mean zero, variance  $\sigma_{u_1}^2$  and  $\sigma_{v_2}^2$  respectively, and  $\text{Cov}(u_{li}, u_{lj}) = \text{Cov}(v_{2i}, v_{2j}) = 0$  for  $i \neq j$ . The probit model assumes that the  $u_{li}$ 's are normally distributed. Whether the distribution of  $v_{2i}$  must be also specified or not depends on the methods of estimating and testing. If a distribution is needed a joint normal distribution of  $u_{li}$  and  $v_{2i}$  is assumed, i.e.

$$(u_{li}, v_{2i}) \Big| x_{li}, x_{2i} \stackrel{\text{iid}}{\sim} \mathbf{N} \left( \mathbf{0}, \begin{bmatrix} \sigma_{u_1}^2 & \sigma_{u_1 v_2} \\ \sigma_{u_1 v_2} & \sigma_{v_2}^2 \end{bmatrix} \right).$$

The parameter of special interest is  $\gamma_1$ . It is not identified if  $\pi_{22}$  is equal to zero. Therefore, weak identification means that  $\pi_{22}$  is "near" to zero. Sometimes weak identification is quantified by the so-called concentration parameter (cf. Stock/ Wright/ Yogo 2002, p. 519). However, this parameter grows with  $N$ , and hence it suggests that the problem of weak identification is reduced by enlarging the sample size. This is misleading, and therefore it is not considered in the study.

Testing the significance of  $\gamma_1$  in empirical studies is usually done by the  $z$ -statistic (implemented in almost all econometric software packages):

$$z = \frac{\hat{\gamma}_1}{\sqrt{\text{asy}\hat{\text{Var}}(\hat{\gamma}_1)}}. \quad (2.2)$$

Given that  $\hat{\gamma}_1$  is a consistent and asymptotically normal distributed estimator,  $z$  is asymptotically standard normal distributed under the assumption of strong identification. The parameter  $\gamma_1$  can be estimated via two step methods (cf. Blundell/ Smith 1993 for an overview) or via joint GMM or joint Maximum Likelihood (ML) estimation of both equations (cf. Wilde 2008 for a comparison).

The  $z$ -test is a Wald-type test. The classical alternatives, the Likelihood Ratio (LR) and Lagrange Multiplier (LM) test, are based on the ML estimation of the parameters. Here, we focus on the LR test. Given the estimates, the LR-statistic is calculated easily, whereas for the LM test an estimation of the complicated information matrix is needed, and the results may depend on which estimation procedure was chosen.

The loglikelihood function of model (2.1) under the standard assumptions above is given as (cf. Wilde 2008, appendix 2):

$$\begin{aligned} \ln l(\theta) = & \sum_{i=1}^N \left[ -0.5 \ln(2\pi\sigma_{v_2}^2) - 0.5 \left( \frac{y_{2i} - x_i\pi_2}{\sigma_{v_2}} \right)^2 \right. \\ & + y_{1i} \ln \Phi \left( \frac{1}{\sqrt{1-\rho_v^2}} \left( \frac{(\gamma_1\pi_{21} + \beta_1)x_{1i} + \gamma_1\pi_{22}x_{2i}}{\sigma_{v_1}} + \rho_v \left( \frac{y_{2i} - \pi_2x_i}{\sigma_{v_2}} \right) \right) \right) \right] \\ & + (1-y_{1i}) \ln \left[ 1 - \Phi \left( \frac{1}{\sqrt{1-\rho_v^2}} \left( \frac{(\gamma_1\pi_{21} + \beta_1)x_{1i} + \gamma_1\pi_{22}x_{2i}}{\sigma_{v_1}} + \rho_v \left( \frac{y_{2i} - \pi_2x_i}{\sigma_{v_2}} \right) \right) \right) \right] \right], \quad (2.3) \end{aligned}$$

where  $x_i = (x_{1i}', x_{2i}')'$ ,  $v_{1i} = u_{1i} + \gamma_1 v_{2i}$ ,  $\theta = (\gamma_1, \beta_1, \pi_2, \sigma_{v_2}, \rho_v)'$ ,  $\pi_2 = (\pi_{21}, \pi_{22})$ ,  $\rho_v = \text{Corr}(v_{1i}, v_{2i})$ , and  $\sigma_{v_1} = \text{Var}(v_{1i})$ . Since only the ratios of the structural parameters and  $\sigma_{v_1}$  are identified, the latter cannot be estimated separately. Therefore, in the simulation study it is replaced by its true value.

The hypothesis, which is tested, is

$$H_0: \gamma_1 = \tilde{\gamma}_1 \text{ vs. } H_1: \gamma_1 \neq \tilde{\gamma}_1.$$

Denote  $\hat{\theta}_{ML}$  the unrestricted ML estimator of  $\theta$  using (2.3) and  $\hat{\theta}_{RML}$  the restricted ML estimator under the null hypothesis. Then, the LR statistic is defined as

$$LR = 2 \left( \ln l(\hat{\theta}_{ML}) - \ln l(\hat{\theta}_{RML}) \right).$$

Under the usual assumptions (including strong identification) LR is asymptotically  $\chi^2_{(1)}$  distributed.

### 3. Design of the simulation study and estimation methods

The simulation design is as simple as possible to avoid arbitrary choices of unnecessary nuisance parameters. However, ML estimation of this simple model causes numerical problems, i.e. many replications ended with the message "maximum iterations reached" even after 200 and more iterations. Therefore, GMM estimation is used to calculate the z-statistic. To analyse the behaviour of the LR-statistic a slight generalization of the basic model is used that avoided the numerical problems on the one hand and did not change the results concerning the weak identification on the other hand (details see below).

In (2.1), the second equation is a reduced form equation. Endogeneity of  $y_2$  can be caused by two reasons: correlation between the error terms of the structural equations and/ or simultaneity between  $y_1$  and  $y_2$ . Both cause a correlation between  $u_1$  and  $v_2$ . In the simulation design we focus on simultaneity because it can be interpreted more easily. Nevertheless, all

results can be reproduced by assuming correlation between the error terms of the structural equations. For simplicity, in the first design  $\beta_1 = \pi_{21} = 0$  and  $K_2 = 1$ .

Therefore, the basic data generating model is of the following structure:

$$y_{li}^* = \gamma_1 y_{2i} + u_{li}, \quad y_{li} = \begin{cases} 1, & y_{li}^* > 0 \\ 0, & \text{sonst} \end{cases} \quad i = 1, \dots, N \quad (3.1)$$

$$y_{2i} = \gamma_2 y_{li}^* + \beta_2 x_i + u_{2i}.$$

The residuals  $u_{li}$  and  $u_{2i}$  were drawn independently from a  $N(0, 16)$  distribution, i.e. the residual variances are equal for both equations. The exogenous variable  $x_i$  was drawn from a  $N(0.5, 16)$  distribution so that the expected number of ones for  $y_{li}$  differs from the expected number of zeros. Weak identification is equivalent to  $\beta_2$  “near to zero”. In the simulation,  $\beta_2 = 0.01$  is chosen. Smaller values of  $\beta_2$  did not sharpen the results any more. The case of strong identification was simulated by  $\beta_2 = 1$ . The simulations were done for the sample sizes  $N = 400$  (medium size) and  $N = 2000$  (large size) and were replicated 5000 times.

The estimated model was

$$y_{li}^* = \gamma_1 y_{2i} + u_{li}, \quad y_{li} \text{ as defined in (3.1), } i = 1, \dots, N \quad (3.2)$$

$$y_{2i} = \pi_2 x_i + v_{2i},$$

where  $\pi_2 = \frac{\beta_2}{1 - \gamma_1 \gamma_2}$  and  $v_{2i} = \frac{u_{2i} + \gamma_2 u_{li}}{1 - \gamma_1 \gamma_2}$ . Both parameters in (3.2) are exactly identified.

(3.2) is estimated via GMM using the “natural” moment conditions (cf. Wilde 2008)

$$E \begin{bmatrix} x_i (y_{li} - \Phi(\gamma_1 \pi_2 x_i / \sigma_{v_1})) \\ x_i (y_{2i} - \pi_2 x_i) \end{bmatrix} = 0.$$

Denoting

$$\theta = \begin{pmatrix} \gamma_1 \\ \pi_2 \end{pmatrix}, \quad m_i(\theta) = \begin{pmatrix} x_i (y_{li} - \Phi(\gamma_1 \pi_2 x_i / \sigma_{v_1})) \\ x_i (y_{2i} - \pi_2 x_i) \end{pmatrix}, \quad \text{and } \bar{m}(\theta) = \frac{1}{N} \sum_{i=1}^N m_i(\theta)$$

it is calculated

$$\hat{\theta}_{\text{GMM}} = \arg \min_{\theta} \bar{m}(\theta)' W_N \bar{m}(\theta), \quad W_N \text{ a weighting matrix,}$$

with regard to  $\theta$ .<sup>3</sup> Since the number of moment conditions is equal to the number of parameters the weighting matrix in the criterion function of the GMM estimator does not matter theoretically and the same asymptotic covariance matrix of the estimator can be used for all choices of  $W_N$  (Harris/ Mátyás 1999, p. 22). To be more precise, the asymptotic covariance matrix is (cf. Greene 2008, p. 445):

<sup>3</sup> Again,  $\sigma_{v_1}$  is not identified. It is replaced by the standard deviation of  $v_1$  in each replication.

$$\text{asyVar}(\hat{\theta}) = \frac{1}{N} [\mathbf{G}' \Phi^{-1} \mathbf{G}]^{-1},$$

$$\Phi = \text{asyVar}(\sqrt{N}\bar{\mathbf{m}}),$$

$$\mathbf{G} = \frac{\partial \bar{\mathbf{m}}}{\partial \theta'} = \begin{pmatrix} \frac{1}{N} \sum_{i=1}^N \frac{\partial m_i^{(1)}(\theta)}{\partial \gamma_1} & \frac{1}{N} \sum_{i=1}^N \frac{\partial m_i^{(1)}(\theta)}{\partial \pi_2} \\ \frac{1}{N} \sum_{i=1}^N \frac{\partial m_i^{(2)}(\theta)}{\partial \gamma_1} & \frac{1}{N} \sum_{i=1}^N \frac{\partial m_i^{(2)}(\theta)}{\partial \pi_2} \end{pmatrix},$$

$$\frac{\partial m_i^{(1)}(\theta)}{\partial \gamma_1} = \frac{-x_i^2 \pi_2}{\sigma_{v_1}} \varphi\left(\frac{\gamma_1 \pi_2 x_i}{\sigma_{v_1}}\right), \quad \frac{\partial m_i^{(1)}(\theta)}{\partial \pi_2} = \frac{-x_i^2 \gamma_1}{\sigma_{v_1}} \varphi\left(\frac{\gamma_1 \pi_2 x_i}{\sigma_{v_1}}\right),$$

$$\frac{\partial m_i^{(2)}(\theta)}{\partial \gamma_1} = 0, \quad \frac{\partial m_i^{(2)}(\theta)}{\partial \pi_2} = -x_i^2.$$

Given the assumptions above this matrix can be estimated consistently via

$$\text{est asyVar}(\hat{\theta}) = \frac{1}{N} [\hat{\mathbf{G}}' \hat{\Phi}^{-1} \hat{\mathbf{G}}]^{-1},$$

$$\hat{\Phi} = \frac{1}{N} \sum_{i=1}^N \mathbf{m}_i(\hat{\theta}) \mathbf{m}_i(\hat{\theta})', \quad \hat{\mathbf{G}} = \mathbf{G} \text{ after substituting } \hat{\theta} \text{ for } \theta.$$

The square root of the first diagonal element is the denominator of the z-statistic (2.2). The nominator is calculated using the identity matrix and the estimator  $\hat{\Phi}^{-1}$  of the optimal weighting matrix. The results were often the same. However, sometimes the identity matrix produces numerically more stable results, because no inverse needs to be calculated. Therefore, the reported results are based on the identity matrix.

Concerning the ML estimation of the basic model everything works fine for strong identification, whereas for weak identification the algorithm did not find the maximum for nearly half of the replications. This is caused by the following: In the basic model, for instance the second summand of (2.3) reduces to

$$y_{li} \ln \Phi \left( \frac{1}{\sqrt{1-\rho_v^2}} \left( \frac{\gamma_1 \pi_2 x_{2i}}{\sigma_{v_1}} + \rho_v \left( \frac{y_{2i} - \pi_2 x_i}{\sigma_{v_2}} \right) \right) \right).$$

Thus, the parameter  $\gamma_1$  is part of the log likelihood function only in the product  $\gamma_1 \pi_2$ . However, in case of weak identification  $\pi_2$  is close to zero, so that it is very difficult to find the "true" value of  $\gamma_1$ . The log likelihood function is rather flat concerning  $\gamma_1$ . Therefore, a constant is added in both equations and the following data generating process is considered:

$$y_{li}^* = \gamma_1 y_{2i} + \beta_{10} + u_{li} \quad y_{li} = \begin{cases} 1, & y_{li}^* > 0 \\ 0, & \text{else} \end{cases},$$

$$y_{2i} = \gamma_2 y_{li}^* + \beta_{20} + \beta_{21} x_i + u_{2i}$$

where  $\beta_{10}=0.5$ ,  $\beta_{20}=0.25$ ,  $x_i \sim N(0,16)$  for all replications, and  $u_1$  and  $u_2$  are as above.

The estimated model is now

$$\begin{aligned}
 y_{li}^* &= \gamma_1 y_{2i} + \beta_{10} + u_{li} \\
 y_{2i} &= \pi_{20} + \pi_{21} x_i + v_{2i} \\
 \text{with } \pi_{20} &= \frac{\beta_{20} + \gamma_2 \beta_{10}}{1 - \gamma_1 \gamma_2}, \quad \pi_{21} = \frac{\beta_{21}}{1 - \gamma_1 \gamma_2}, \quad v_{2i} = \frac{u_{2i} + \gamma_2 u_{li}}{1 - \gamma_1 \gamma_2}.
 \end{aligned}
 \quad y_{li} = \begin{cases} 1, & y_{li}^* > 0 \\ 0, & \text{else} \end{cases}$$

As expected, this avoids the numerical problems discussed above without changing the results concerning weak identification. The latter aspect was confirmed for those parameter values for which the optimum was found in the basic model.

#### 4. Results of the simulation study<sup>4</sup>

Two cases are distinguished: First, the null hypothesis “ $\gamma_1=0$ ” is tested, i.e. the significance of  $y_2$ . Second, the null hypothesis “ $\gamma_1=c$ ” is tested,  $c$  a constant different from zero, and the results for  $c=2$  are presented. In both cases, the simulations were done for different values of  $\gamma_2$ . Small values of  $\gamma_2$  correspond with a “small” problem of simultaneity. We choose  $\gamma_2 = \pm 0.5$ . A “medium” problem of simultaneity is represented by  $\gamma_2 = \pm 1.5$ , and a “large” problem of simultaneity is represented by  $\gamma_2 = \pm 3$  and  $\gamma_2 = \pm 6$ . Since the differences between a medium and a large sample size were only moderate, the results are presented only for  $N = 2000$ .

##### 4.1 Results for the z-Test

In case of testing significance and under strong identification, no size problems can be observed:

$\gamma_2$	nominal size		
	1 %	5 %	10 %
-3	0.0120	0.0470	0.0982
-0.5	0.0098	0.0502	0.1002
3	0.0094	0.0492	0.0964

Table 1: Percentage of rejections of the z-statistic under strong identification,  $H_0: \gamma_1 = 0$ .

In case of weak identification the picture is mixed.<sup>5</sup> If there is only „weak“ simultaneity, an extreme undersizing is observed. However, “strong” simultaneity causes medium size distortions:

<sup>4</sup> All simulations were done using R. The R codes are available on request.

<sup>5</sup> This result is similar to that of Magnusson (2007).



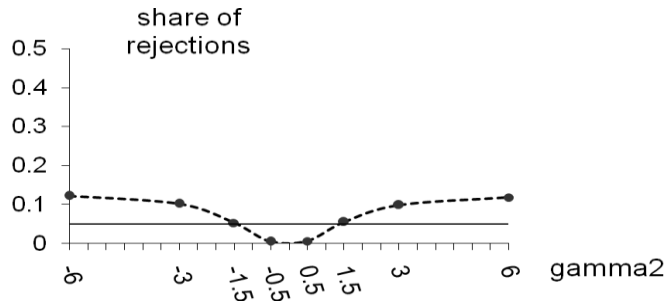


Figure 1: Percentage of rejections of the z-statistic under weak identification,  $H_0: \gamma_1 = 0$ , nominal size 5 %.

Nevertheless, the size distortions are smaller than those in linear simultaneous equations. Therefore, the data were also simulated using other values of  $c$ . Afterwards the results for  $c = 2$  are shown as an example. Again, in case of strong identification, no size distortion appears:

$\gamma_2$	nominal size		
	1 %	5 %	10 %
-3	0.0110	0.0470	0.0910
-0.5	0.0056	0.0356	0.0688
3	0.0086	0.0456	0.0896

Table 2: Percentage of rejections of the z-statistic under strong identification,  $H_0: \gamma_1 = 2$ .

However, in case of weak identification *and* strong simultaneity, the size distortions became very large ( $\gamma_2 = 0.5$  has to be omitted here because  $\gamma_1$  is no longer identified in that case):

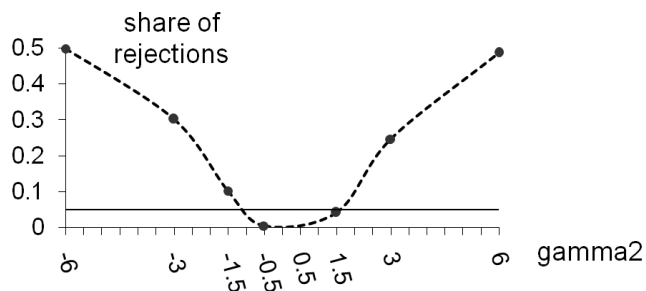


Figure 2: Percentage of rejections of the z-statistic under weak identification,  $H_0: \gamma_1 = 2$ , nominal size 5 %.

The empirical size becomes more than tenfold as high as the nominal size. Thus, in probit models similar size distortions as in linear simultaneous equations models can be observed.

## 4.2 Results for the LR-Test

In case of testing significance and under strong identification, again no size problems can be observed:

$\gamma_2$	nominal size		
	1 %	5 %	10 %
-3	0.01	0.0494	0.101
-0.5	0.012	0.0488	0.0986
3	0.0106	0.0548	0.0998

Table 3: Percentage of rejections of the LR-statistic under strong identification,  $H_0: \gamma_1 = 0$ .

Simulating the case of weak identification,  $\beta_2 = 0.0001$  is used, i.e. a smaller value of  $\beta_2$  is needed to get the strongest results. They differ substantially from those for the z-test:

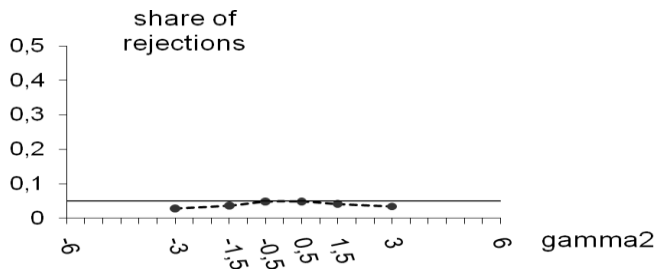


Figure 3: Percentage of rejections of the LR-statistic under weak identification,  $H_0: \gamma_1 = 0$ , nominal size 5 %.

If simultaneity is weak, the observed share of rejections is near to the true size, if it is medium or strong, *undersizing* is observed. Results for  $\gamma_2 = \pm 6$  are missing, because the program stopped with an error message for some replications. This message was caused by the following: Consider for instance  $\gamma_1=0$  and  $\gamma_2=6$ . This implies  $\rho_v=0.9864$ , i.e. the bivariate normal distribution of  $v_1$  and  $v_2$  is near to singularity. Furthermore, in (2.3)  $1/\sqrt{1-\rho_v^2} = 6.08$ ,  $\Phi(\approx 6) = 1$ ,  $1-\Phi(\approx 6) = 0$ , and  $\ln(1-\Phi(\approx 6))$  is not defined. Therefore, ML estimation is less robust against a high correlation of the reduced form errors than GMM estimation.

The results for the LR-test do not change if  $\gamma_1=2$  is tested. In that case  $\gamma_2=6$  is possible:

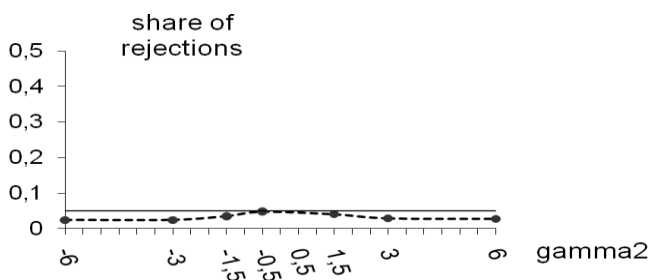


Figure 4: Percentage of rejections of the LR-statistic under weak identification,  $H_0: \gamma_1 = 2$ , nominal size 5 %.

Again, in case of weak identification and strong simultaneity the empirical size is about half of the nominal size.

## 5. Conclusion

The paper is a kind of first serve concerning the analysis of weak identification in probit models with endogenous covariates. It shows remarkable size distortions concerning the usual z-statistic. However, further research is needed to clarify why the magnitude depends heavily on the tested hypothesis. The likelihood ratio statistic could be a conservative alternative. However, its power might be low. Thus, further research is useful to clarify how advanced methods like those of Andrews/Cheng (2011b), Dufour (2006) or Kleibergen (2005) will work for probit models with endogenous covariates.

## References

- Andrews, D.W.K./ Cheng, X.* (2011a), Maximum likelihood estimation and uniform inference with sporadic identification failure, Cowles Foundation Discussion Paper No. 1824, Yale University.
- Andrews, D.W.K./ Cheng, X.* (2011b), GMM estimation and uniform subvector inference with possible identification failure, Cowles Foundation Discussion Paper No. 1828, Yale University.
- Brav, O.* (2009), Access to capital, capital structure, and the funding of the firm, *Journal of Finance* 64, 263-308.
- Dufour, J.M.* (1997), Some impossibility theorems in econometrics with applications to structural and dynamic models, *Econometrica* 65, 1365-1387.
- Dufour, J.M.* (2003), Identification, weak instruments, and statistical inference in econometrics, *Canadian Journal of Economics* 36, 767-808.
- Dufour, J.M.* (2006), Monte carlo tests with nuisance parameters: A general approach to finite sample inference and nonstandard asymptotics, *Journal of Econometrics* 133, 443-477.
- Duryea, S./ Lam, D./ Levison, D.* (2007), Effects of economic shocks on children's employment and schooling in Brazil, *Journal of Development Economics* 84, 188-214.
- Engelhardt, G.V./ Eriksen, M.D./ Gale, W.G./ Mills, G.B.* (2010), What are the social benefits of homeownership? Empirical evidence for low-income households, *Journal of Urban Economics* 67, 249-258.
- Esaka, T.* (2010), De facto exchange rate regimes and currency crises: Are pegged regimes with capital account liberalization really more prone to speculative attacks?, *Journal of Banking and Finance* 34, 1109-1128.

- Fitzenberger, B./ Kohn, K./ Wang, Q.* (2011), The erosion of union membership in Germany: determinants, densities, decompositions, *Journal of Population Economics* 24, 141-165.
- Greene, W.H.* (2008), *Econometric analysis* 6<sup>th</sup> ed., Prentice Hall, Upper Saddle River.
- Hajivassiliou, V.A./ Ioannides, Y.M.* (2007), Unemployment and liquidity constraints, *Journal of Applied Econometrics* 22, 479-510.
- Hao, L./ Ng, E.C.Y.* (2011), Predicting Canadian recessions using dynamic probit modelling approaches, *Canadian Journal of Economics* 44, 1297-1330.
- Harris, D./ Mátyás, L.* (1999), Introduction to the generalized method of moments estimation, in: Mátyás, L. (ed.), *Generalized method of moments estimation*, Cambridge University Press, Cambridge, 3-30.
- Hlaing, K.P./ Pourjalali, H.* (2012), Economic Reasons for Reporting Property, Plant, and Equipment at Fair Market Value by Foreign Cross-Listed Firms in the United States, *Journal of Accounting, Auditing & Finance* 27, 557-576.
- Kauppi, H./ Saikkonen, P.* (2008), Predicting U.S. recessions with dynamic binary response models, *Review of Economics and Statistics* 90, 777-791.
- Kleibergen, F.* (2005), Testing parameters in GMM without assuming that they are identified, *Econometrica* 73, 1103-1123.
- Litchfield, J./ Reilly, B./ Veneziani, M.* (2012), An analysis of life satisfaction in Albania: An heteroscedastic ordered probit model approach, *Journal of Economic Behavior & Organization* 81, 731-741.
- Magnusson, L.M.* (2007), Weak instruments robust tests for limited dependent variable models, Working Paper, Brown University (RI).
- Magnusson, L.M.* (2010), Inference in limited dependent variable models robust to weak identification, *Econometrics Journal* 13, S56-S79.
- Mizutani, F./ Uranishi, S.* (2010), Determinants of privatization of public corporations: Evidence from the Japanese experience, *Empirical Economics* 39, 515-535.
- Schneider, C.J.* (2007), Enlargement processes and distributional conflicts: The politics of discriminatory membership in the European Union, *Public Choice* 132, 85-102.
- Stock, J.H./ Wright, J.H./ Yogo, M.* (2002), A survey of weak instruments and weak identification in Generalized Method of Moments, *Journal of Business and Economic Statistics* 20, 518-529
- Wilde, J.* (2008), A note on GMM estimation of probit models with endogenous regressors, *Statistical Papers* 49, 471-484.