# Weak identification in probit models with endogenous covariates

**Jean-Marie Dufour** · **Joachim Wilde**

April 2018

**Abstract** Weak identification is a well-known issue in the context of linear structural models. However, for probit models with endogenous explanatory variables, this problem has been little explored. In this paper, we study by simulation the behavior of the usual z-test and the LR test in the presence of weak identification. We find that the usual asymptotic z-test exhibits large level distortions (over-rejections under the null hypothesis). The magnitude of the level distortions depends heavily on the parameter value tested. In contrast, asymptotic LR tests do not over-reject and appear to be robust to weak identification.

**Keywords** probit model · weak identification

**JEL Classification** C35

The authors thank Leandro Magnusson and two anonymous referees for several useful comments, and Dietrich Trenkler and Sebastian Veldhuis for valuable assistance. This work was supported by the William Dow Chair in Political Economy (McGill University), the Bank of Canada (Research Fellowship), the Toulouse School of Economics (Pierre-de-Fermat Chair of excellence), the Universitad Carlos III de Madrid (Banco Santander de Madrid Chair of excellence), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec).

Jean-Marie Dufour
William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en conomie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 414, 855 Sherbrooke Street West, Montreal, Québec H3A 2T7, Canada. TEL: (1) 514 398 6071; FAX: (1) 514 398 4800; E-mail: jean-marie.dufour@mcgill.ca, Web page: `http://www.jeanmariedufour.com`

Joachim Wilde
Corresponding author. Mailing address: Fachbereich Wirtschaftswissenschaften, Rolandstr. 8, 49069 Osnabrueck, Germany; Tel: (49) 541 969 2746; E-mail: Joachim.Wilde@uni-osnabrueck.de, Web page: `https://www.wiwi.uni-osnabrueck.de/en/departments_and_institutes/econometrics_and_statistics_prof_wilde.html`

## 1 Introduction

Probit models are widely used in applied econometrics; for some recent examples, see Abramitzky and Lavy (2014), Beck et al (2014), Bijsterbosch and Dahlhaus (2015), Bouoiyour et al (2016), Cornelli et al (2013), Croushore and Marsten (2016), Engelhardt et al (2010), Esaka (2010), Fitzenberger et al (2011), Haider and Jahangir (2017), Hao and Ng (2011), Hlaing and Pourjalali (2012), Horvath and Katuscakova (2016), Khanna et al (2015), Litchfield et al (2012), Massa and Zhang (2013), Wen and Gordon (2014). As in linear models, one or more explanatory variables can be endogenous. This problem can be solved by using instrumental variables; see 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.

In linear models, it is well known that weak instruments can cause considerable level (or size) distortions [see Dufour (2003) for an overview]. Wald-type tests like the usual t-tests and F-tests are especially vulnerable to this problem [see 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 asymptotic standard error. This is a Wald-type test. Therefore, large level 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 and Cheng (2013, 2014), who address the probit model as an example. However, Andrews and Cheng (2013) restrict their numerical analysis to a probit model with a nonlinear regression function and without endogeneity, and Andrews and Cheng (2014) do not analyze the probit model numerically.[1]

This paper makes several contributions to this problem. First, large level distortions in probit models with endogeneity are demonstrated by a simulation study. Second, we show that level distortions depend heavily on which parameter value is tested: whereas level distortions are moderate for the problem of testing the null of a zero parameter, testing other values of the parameter yield large level distortions. Third, the behaviour of the classical likelihood ratio statistic in this case is analyzed. For the simulation design considered, no level distortions are observed. However, the probability of type I error can be notably lower than the nominal level of the test (undersizing). Fourth, some new insights concerning the estimation of probit models with endogenous covariates are provided.

Section 2 describes the econometric model and the test statistics. Section 3 explains the simulation design and the estimators used. Since a probit equation is part of the model, some formulae become more complicated than in the linear case. They are described in detail because textbook descriptions are missing so far. Section 4 presents the results of the simulation study, and Section 5 concludes. For ease of exposition, we focus on the binary probit model.

## 2 Model and classical tests

We study a structural probit model, where one of the explanatory variables is endogenous, and a reduced-form equation for this variable is specified. The specific model considered is:

[1] 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 level distortions. However, in later versions of the working paper and in the published version Magnusson (2010) the probit example was deleted.

$$y_{1i}^* = \gamma_1 y_{2i} + \beta_1 x_{1i} + u_{1i} , \qquad y_{1i} = \begin{cases} 1, & \text{if } y_{1i}^* > 0 \\ 0, & \text{otherwise} \end{cases} , \quad i = 1, \dots, N , \tag{1}$$
$$y_{2i} = \pi_{21} x_{1i} + \pi_{22} x_{2i} + v_{2i} ,$$

where $y_{1i}^*$ is a latent variable, $y_{1i}$ is its observable indicator, $y_{2i}$ is an endogenous (observable) variable, $x_{1i}$ 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, $1 \times K_1$, $1 \times K_1$ and $1 \times K_2$ respectively, and $u_{1i}$ and $v_{2i}$ are error terms with mean zero, variances $\sigma_{u_1}^2$ and $\sigma_{v_2}^2$ respectively, and $\mathrm{Cov}(u_{1i}, u_{1j}) = \mathrm{Cov}(v_{2i}, v_{2j}) = \mathrm{Cov}(u_{1i}, v_{2j}) = 0$ for $i \neq j$. The probit model assumes that the $u_{1i}$'s are normally distributed. Whether the distribution of $v_{2i}$ must also be specified depends on the estimation method. If a distribution is needed, we assume that $u_{1i}$ and $v_{2i}$ follow a joint normal distribution, i.e.

$$(u_{1i}, v_{2i}) | x_{1i}, x_{2i} \overset{\text{iid}}{\sim} \mathrm{N}\left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1 v_2} \\ \sigma_{u_1 v_2} & \sigma_{v_1}^2 \end{pmatrix} \right] .$$

The parameter of special interest is $\gamma_1$. It is not identified when $\pi_{22}$ is equal to zero. Therefore, we say $\gamma_1$ is weakly identified if $\pi_{22}$ is close to zero. Sometimes weak identification is quantified by the so-called concentration parameter; see Stock et al (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 we do not use it as guideline in our study.

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

$$z = \frac{\hat{\gamma}_1}{\sqrt{\hat{V}(\hat{\gamma}_1)}} \tag{2}$$

where $\hat{V}(\hat{\gamma}_1)$ is a consistent estimate (assuming identification) of the asymptotic variance of $\sqrt{N}(\hat{\gamma}_1 - \gamma_1)$. If $\hat{\gamma}_1$ is consistent with a standard normal asymptotic distribution, $z$ is asymptotically standard normal under the assumption of strong identification. The parameter $\gamma_1$ can be estimated by two-step methods [see Blundell and Smith (1993) for an overview] or via joint GMM or maximum likelihood (ML) estimation of both equations [see Wilde (2008) for a comparison].

The z-test is a Wald-type test. The classical alternatives to it — the Likelihood Ratio (LR) and the Lagrange Multiplier (LM) tests — are based on ML estimation of the parameters. In linear models, the latter are less affected by weak identification than Wald-type tests. Here, we focus on the LR test. Given the estimates, the LR statistic is calculated easily, whereas the LM test requires the estimation of a complicated information matrix, so the results may depend on the estimation procedure chosen. The log-likelihood function of model (1) under the above standard assumptions is:

$$\begin{aligned} \ln l(\theta) = \sum_{i=1}^{N} \Bigg[ & -0.5 \ln(2\pi\sigma_{v_2}^2) - 0.5 \left( \frac{y_{2i} - \pi_2 x_i}{\sigma_{v_2}} \right)^2 \\ & + y_{1i} \ln \Phi \left( \frac{1}{\sqrt{1-\rho_v^2}} \left( \frac{(\gamma_1 \pi_{21} + \beta_1) x_{1i} + \gamma_1 \pi_{22} x_{2i}}{\sigma_{v_1}} + \rho_v \left( \frac{y_{2i} - \pi_2 x_i}{\sigma_{v_2}} \right) \right) \right) \\ & + (1 - y_{1i}) \ln \left( 1 - \Phi \left( \frac{1}{\sqrt{1-\rho_v^2}} \left( \frac{(\gamma_1 \pi_{21} + \beta_1) x_{1i} + \gamma_1 \pi_{22} x_{2i}}{\sigma_{v_1}} + \rho_v \left( \frac{y_{2i} - \pi_2 x_i}{\sigma_{v_2}} \right) \right) \right) \right) \Bigg] \end{aligned} \tag{3}$$

where $x_i = (x'_{1i}, x'_{2i})'$ , $v_{1i} = u_{1i} + \gamma_1 v_{2i}$, $\theta = (\gamma_1, \beta_1, \pi_2, \sigma_{v_2}, \rho_v)'$, $\pi_2 = (\pi_{21}, \pi_{22})$, $\sigma_{v_1} = \sqrt{\mathrm{Var}(v_{1i})}$, and $\rho_v = \mathrm{Corr}(v_{1i}, v_{2i})$; see Wilde (2008), appendix 2. Since the structural parameters enter the likelihood only through ratios with a standard deviation and the latter does not appear separately, only these ratios are identifiable. Therefore, in our simulation study, $\sigma_{v_1}$ is taken as known.

We consider the problem of testing

$$\mathrm{H}_0 : \gamma_1 = \tilde{\gamma}_1 \quad \text{vs.} \quad \mathrm{H}_1 : \gamma_1 \neq \tilde{\gamma}_1 \,.$$

We denote by $\hat{\theta}_{\mathrm{ML}}$ the unrestricted ML estimator of $\theta$ [based on (3)] and by $\hat{\theta}_{\mathrm{RML}}$ the restricted ML estimator under the null hypothesis. The LR statistic has the form

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

Under the usual assumptions (including strong identification) LR is asymptotically $\chi^2(1)$ under $\mathrm{H}_0$.

## 3 Simulation design

In order to avoid arbitrary choices of unnecessary nuisance parameters, we consider a simple simulation design with $K_1 \in \{0,1\}$ and $K_2 = 1$. The value $K_1 = 0$ defines a model without constants, whereas $K_1 = 1$ is used to add a constant in both equations. With $K_1 = 0$, ML estimation causes numerical problems, i.e. many replications end with the message "maximum iterations reached" even after 200 and more iterations. Since GMM estimation is a well-known alternative to ML estimation for probit-type models and the efficiency loss in models like ours seems to be small [see Wilde (2008)], we used GMM estimation for the z-test. As will be shown below, the results under weak identification are nearly the same for models with and without constants. Therefore, we focus our comparison of z-test and LR test for models including constants in both equations.

In (1), the second equation is a reduced-form equation. Endogeneity of $y_2$ can occur for at least two reasons: correlation between the error terms of the structural equations and/or simultaneity between $y_1$ and $y_2$. Both yield a correlation between $u_1$ and $v_2$. In our simulation, 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. Therefore, our main data generating model with constants is:

$$\begin{aligned} y^*_{1i} &= \gamma_1 y_{2i} + \beta_{11} + u_{1i}, \\ y_{2i} &= \gamma_2 y^*_{1i} + \beta_{21} + \beta_{22} x_i + u_{2i}, \end{aligned} \qquad y_{1i} = \begin{cases} 1\,, & \text{if } y^*_{1i} > 0 \\ 0\,, & \text{otherwise} \end{cases}, \quad i = 1, \ldots, N\,. \tag{4}$$

The residuals $u_{1i}$ and $u_{2i}$ are drawn independently from the $\mathrm{N}(0,16)$ distribution, so the residual variances are equal for both equations. The exogenous variable $x_i$ is drawn from the $\mathrm{N}(0.5,16)$ distribution, so the expected number of ones for $y_{1i}$ differs from the expected number of zeros also for the model without constants. Alternatively, we draw the residuals independently from the $\mathrm{N}(0,1)$ distribution and $x_i$ from an $\mathrm{N}(0.5,1)$ distribution. The constants are $\beta_{11} = 0.5$ and $\beta_{21} = 0.25$. Weak identification is equivalent to $\beta_{22}$ being close to zero. In our simulations, we took $\beta_{22} = 0.0001$. Smaller values of $\beta_{22}$ yield qualitatively similar results. The case of strong identification is simulated with $\beta_{22} = 1$. The simulations were also run with $\beta_{22} = 0.1$ to see whether weak identification also causes problems for moderate parameter values. The sample sizes are $N = 400$ (medium sample size) and $N = 2000$ (large sample), and the number of replications is 5000.

The estimated model in these simulations is:

$$\begin{aligned} y_{1i}^* &= \gamma_1 y_{2i} + \beta_{11} + u_{1i} \\ y_{2i} &= \pi_{21} + \pi_{22} x_i + v_{2i} \end{aligned} \qquad y_{1i} = \begin{cases} 1\,, & y_{1i}^* > 0 \\ 0\,, & \text{else} \end{cases}\,, \quad i = 1, \dots, N\,, \tag{5}$$

where

$$\pi_{21} = (\gamma_2 \beta_{11} + \beta_{21})/(1 - \gamma_1 \gamma_2), \quad \pi_{22} = \beta_{22}/(1 - \gamma_1 \gamma_2), \quad v_{21} = (\gamma_2 u_{1i} + u_{2i})/(1 - \gamma_1 \gamma_2).$$

In (5), $\gamma_1$ is exactly identified as long as $\pi_{22}$ is different from zero. (5) is estimated by GMM using the 'natural' moment conditions [see Wilde (2008)]:

$$\mathrm{E}\begin{bmatrix} x_i \left( y_{1i} - \Phi\left( \dfrac{\gamma_1(\pi_{21} + \pi_{22} x_i) + \beta_{11}}{\sigma_{v_1}} \right) \right) \\ x_i (y_{2i} - \pi_{21} - \pi_{22} x_i) \\ 1 \left( y_{1i} - \Phi\left( \dfrac{\gamma_1(\pi_{21} + \pi_{22} x_i) + \beta_{11}}{\sigma_{v_1}} \right) \right) \\ 1 (y_{2i} - \pi_{21} - \pi_{22} x_i) \end{bmatrix} = 0\,.$$

Setting

$$\theta = \begin{pmatrix} \gamma_1 \\ \pi_{22} \\ \beta_{11} \\ \pi_{21} \end{pmatrix}, \quad m_i(\theta) = \begin{bmatrix} x_i \left( y_{1i} - \Phi\left( \dfrac{\gamma_1(\pi_{21} + \pi_{22} x_i) + \beta_{11}}{\sigma_{v_1}} \right) \right) \\ x_i (y_{2i} - \pi_{21} - \pi_{22} x_i) \\ 1 \left( y_{1i} - \Phi\left( \dfrac{\gamma_1(\pi_{21} + \pi_{22} x_i) + \beta_{11}}{\sigma_{v_1}} \right) \right) \\ 1 (y_{2i} - \pi_{21} - \pi_{22} x_i) \end{bmatrix}, \quad \bar{m}(\theta) = \frac{1}{N} \sum_{i=1}^{N} m_i(\theta)\,,$$

we calculate

$$\hat{\theta} = \hat{\theta}_{\text{GMM}} = \operatorname*{argmin}_{\theta} \{\bar{m}(\theta)' W_N \, \bar{m}(\theta)$$

where $W_N$ is a weighting matrix. 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 and Mátyás (1999), p. 22]. To be more precise, the asymptotic covariance matrix is [see Greene (2008), p. 445]:

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

$$\Psi = \text{asyVar}(\sqrt{N}\bar{m})\,, \quad \varphi_i := \varphi \left( \frac{\gamma_1(\pi_{21} + \pi_{22} x_i) + \beta_{11}}{\sigma_{v_1}} \right),$$

$$G = \frac{\partial \bar{m}}{\partial \theta'} = \begin{pmatrix} \frac{1}{N} \sum_{i=1}^{N} -\frac{\pi_{21} x_i + \pi_{22} x_i^2}{\sigma_{v_1}} \varphi_i & \frac{1}{N} \sum_{i=1}^{N} -\frac{\gamma_1 x_i^2}{\sigma_{v_1}} \varphi_i & \frac{1}{N} \sum_{i=1}^{N} -\frac{x_i}{\sigma_{v_1}} \varphi_i & \frac{1}{N} \sum_{i=1}^{N} -\frac{\gamma_1 x_i}{\sigma_{v_1}} \varphi_i \\ 0 & \frac{1}{N} \sum_{i=1}^{N} -x_i^2 & 0 & \frac{1}{N} \sum_{i=1}^{N} -x_i \\ \frac{1}{N} \sum_{i=1}^{N} -\frac{\pi_{21} + \pi_{22} x_i}{\sigma_{v_1}} \varphi_i & \frac{1}{N} \sum_{i=1}^{N} -\frac{\gamma_1 x_i}{\sigma_{v_1}} \varphi_i & \frac{1}{N} \sum_{i=1}^{N} -\frac{1}{\sigma_{v_1}} \varphi_i & \frac{1}{N} \sum_{i=1}^{N} -\frac{\gamma_1}{\sigma_{v_1}} \varphi_i \\ 0 & \frac{1}{N} \sum_{i=1}^{N} -x_i & 0 & -1 \end{pmatrix}.$$

Given the above assumptions, this matrix can be estimated consistently by

$$\mathrm{est\,asyVar}(\hat{\theta}) = \frac{1}{N}\left[\hat{G}'\hat{\Psi}^{-1}\hat{G}\right]^{-1},$$

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

The square root of the first diagonal element is the denominator of the z-statistic (2). $\hat{\Psi}^{-1}$ is the optimal weighting matrix and is used to calculate the nominator of (2).[2] In the model without constants only the first two moment conditions are used (setting $\pi_{21}$ and $\beta_{11}$ equal to zero), and $G$ reduces to a $2 \times 2$ matrix.

For ML estimation of the model without constants, everything works fine under strong identification, whereas under weak identification the algorithm did not find the maximum for nearly half of the replications. This is not surprising under weak identification. For example, in the model without constants, the second summand of (3) reduces to

$$y_{1i} \ln \Phi\left(\frac{1}{\sqrt{1-\rho_v^2}}\left(\frac{\gamma_1\pi_{22}x_i}{\sigma_{v_1}} + \rho_v\left(\frac{y_{2i}-\pi_{22}x_i}{\sigma_{v_2}}\right)\right)\right)$$

where the parameter $\gamma_1$ affects the log-likelihood function only through the product $\gamma_1\pi_{22}$. However, under weak identification, $\pi_{22}$ is close to zero, so that it is very difficult to find the 'true' value of $\gamma_1$: the log-likelihood function is relatively flat with respect to $\gamma_1$. However, including constants 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 also found in the model without constants. [3]

## 4 Simulation results

We distinguish two cases. First, we test the null hypothesis $\gamma_1 = 0$, i.e. the significance of $y_2$. Second, we test the null hypothesis $\gamma_1 = c$, where $c$ is a constant different from zero, and we present the results for $c = 2$. In both cases, the simulations consider different values of $\gamma_2$. Small values of $\gamma_2$ correspond to a 'small' simultaneity problem, here $\gamma_2 = \pm 0.5$. 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 $\gamma_2 = \pm 6$.

### 4.1 Results for the z-test

*Main model*

The results for the models with and without constants are similar. Therefore, we focus on the model with constants. The results for the model without constants can be found in the appendix. In case of strong identification and testing $\gamma_1 = 0$, no level problems can be observed at the nominal levels 10%, 5%, and (to a lesser degree) 1%; see Table 1. Under weak

[2] All simulations were done using R; see R Core Team (2016). The GMM estimation was done using the package `GMM`, version 1.6-1. The case of iid observations can be implemented by the option `vcov="iid"`; see Chaussé (2010), p. 13. All R codes are available on request.

[3] For the ML estimation the exogenous variable $x_i$ was drawn from an $\mathrm{N}(0,16)$ distribution.

identification, the picture is mixed.[4] The key result is highlighted in Figure 1 for $N = 2000$ and a nominal level of 5%. If there is only 'weak' simultaneity, an extreme undersizing phenomenon is observed. However, 'strong' simultaneity causes medium level distortions. Nevertheless, the level distortions are smaller than those in linear simultaneous equations. This is surprising because weak identification should cause similar problems in probit models. Furthermore, the results for $\beta_{22} = 0.1$ are close to the results for $\beta_{22} = 0.0001$, i.e. oversizing occurs even for moderate values of the parameter $\beta_{22}$. All results are similar for $N = 2000$ and $N = 400$.

Next, we consider the results concerning data simulated with $\gamma_1 = 2$ and testing $\mathrm{H}_0 : \gamma_1 = 2$. Again, under strong identification, no level distortion appears at the nominal levels 10% and 5%; see Table 2. However, with weak identification and strong simultaneity, the level distortions become very large ($\gamma_2 = 0.5$ is omitted because $\gamma_1$ is no longer identified in this case); see Figure 2. The empirical level becomes more than tenfold as high as the nominal level. Thus, level distortions comparable to those in linear simultaneous equations models can be observed in probit models. With strong simultaneity, level distortions also occur for $\beta_{22} = 0.1$. They are stronger for $N = 400$, i.e. the sample size makes a notable difference in this case.

The differences between testing $\gamma_1 = 0$ and $\gamma_1 = 2$ demonstrate an important feature: in probit models, level distortions depend heavily on the parameter value tested. This property is not easily explained by the concentration parameter. Following the definition of Stock et al (2002), p. 519, the concentration parameter $\mu^2$ in the model without constants is:

$$\mu^2 = \frac{\sum_{i=1}^{N} x_i^2 \pi_{22}^2}{\sigma_{v_2}^2} = \frac{\sum_{i=1}^{N} x_i^2 \left(\frac{\beta_{22}}{1-\gamma_1\gamma_2}\right)^2}{\frac{\sigma_{u_2}^2 + \gamma_2^2 \sigma_{u_1}^2}{(1-\gamma_1\gamma_2)^2}} = \frac{\sum_{i=1}^{N} x_i^2 \beta_{22}^2}{\sigma_{u_2}^2 + \gamma_2^2 \sigma_{u_1}^2} .$$

The concentration parameter does not depend on $\gamma_1$. In the model with constants the formula is more bulky, but the message remains unchanged. Thus, concerning the concentration parameter the problem of weak identification should not depend on the value of $\gamma_1$.

*Sensitivity analysis*

We performed several robustness checks on our results. First, we vary the estimation method, by iterating our two-step GMM estimator, and using $I$ instead of $\Psi^{-1}$ as weighting matrix.[5] Second, we change the variances, i.e. we draw the $u_i$'s from the $\mathrm{N}(0,1)$ distribution and $x_i$ from an $\mathrm{N}(0.5,1)$ distribution. Third, we change from exact identification of $\gamma_1$ to underidentification ($\beta_{22} = 0$) and overidentification (a second exogenous variable is added to the second equation). Except for overidentification, the different variants do not change the results. Thus, we present only examples of the these variants.

Changing the estimation method does not affect the results, i.e. the rejection frequencies are the same as in Tables 1 and 2. Changing the variances in the sample design slightly modifies the results, but the pattern remains the same: we observe only moderate level distortions when testing $\gamma_1 = 0$ and large level distortions when testing $\gamma_1 = 2$ are observed (see Table 3). The results in the case of underidentification are very close to those of $\beta_2 = 0.0001$ (see Table 4).

[4] This result is similar to that of Magnusson (2007).

[5] We also tried the option `"CUE"` for the continuous updating estimator. However, we always got the error message `"node stack overflow"`.

In contrast with the above results, overidentification leads to new insights. In the second equations of (4) and (5), there is an additional exogenous variable drawn from an $\mathrm{N}(0,16)$ distribution. The corresponding structural parameter $\beta_{23}$ varies like $\beta_{22}$: for strong identification, we consider $\beta_{22} = \beta_{23} = 1$, and for weak identification $\beta_{22} = \beta_{23} = 0.0001$ (very weak identification) and $\beta_{22} = \beta_{23} = 0.1$. The second exogenous variable is used as an instrument in the same way as $x_i$ and the ones in the previous simulations. The matrix $G$ now becomes a $(6 \times 5)$-matrix, i.e. the number of moment conditions exceeds the number of parameters. Thus, the choice of the weighting matrix should matter. Therefore, we compare the results using the identity matrix $I$ with those based on the 'optimal' weighting matrix. We focus on a nominal level of $\alpha = 5\%$.

First, the sample size is important. Under strong (over-)identification, everything works fine with $N = 2000$ (see Table 5), but $N = 400$ is too small to meet the nominal level. For testing $\gamma_1 = 0$, the rejection frequencies are in the interval [0.357, 0.3754]; for testing $\gamma_1 = 2$, the rejection frequencies are in the interval [0.2846, 0.3942]. This is an interesting side result and a cautionary note if models like ours are used with macroeconomic or experimental data, where sample sizes can easily be small.

Second, testing $\gamma_1 = 0$ with $N = 2000$ leads to larger level distortions with weak identification and large simultaneity. Now, empirical rejection frequencies reach four times the nominal level (see Table 5). The results are slightly stronger if $I$ is used as weighting matrix. With weak identification and weak simultaneity, *under*sizing is still observable, but it is now less pronounced. With strong simultaneity, the results for $\beta_{22} = \beta_{23} = 0.1$ are again close to those under weak identification, whereas with weak simultaneity, the results with $\beta_{22} = \beta_{23} = 0.1$ are close to those under strong identification.

Third, when testing $\gamma_1 = 2$ with $N = 2000$, the picture is mixed. Using the weighting matrix $I$ strengthens the results in case of strong simultaneity in comparison with the results for model (5). The results are now close to the theoretical expectation. However, using the optimal weighting matrix yields smaller level distortions than under exact identification. The interpretation of this result is difficult: for some parameter combinations, the rejection frequencies for $\beta_{22} = \beta_{23} = 0.0001$ are even lower than for $\beta_{22} = \beta_{23} = 0.1$ . This may be due to the fact that the asymptotic standard errors are a complicated function of $\gamma_1$. If the identity matrix $I$ is used as the weighting matrix, only $G'G$ must be inverted for computing the asymptotic variance-covariance matrix [see Greene (2008), p. 451]. In contrast, using the optimal weighting matrix requires the inversion of $G'\Psi^{-1}G$. Under weak identification, the latter calculation can lead to 'bad' results. However, further research is needed to clarify the reasons for this. Nevertheless, under strong and even medium simultaneity, noteworthy level distortions are observed.

## 4.2 Results for the LR test on the main model

We consider again the main model (5). We use the same parameter values as in Section 4.1, and we calculate the Maximum Likelihood estimator and the LR statistic. In case of strong identification, no level problems can be observed (see Table 6). However, the results under weak identification differ substantially from those for the z-test. If simultaneity is weak, the observed rejection frequencies are close to the nominal level; if simultaneity is medium or strong, undersizing is observed. Thus, the LR test may be a conservative alternative to the z-test.

For testing $\gamma_1 = 0$, the results with $\gamma_2 = 6$ are missing, because the program stopped with an error message for some replications. This can be explained as follows. 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 close to being singular. In the log-likelihood function (3), we then have $1/\sqrt{1-\rho_v^2} = 6.08$, $\Phi(\approx 6) = 1$ , $1-\Phi(\approx 6) = 0$, 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. This is similar to the findings in Wilde (2008), p. 476, and an interesting side finding of this paper.

In contrast with the z-test, the results for the LR test under weak identification do *not* change if $\gamma_1 = 2$ is tested. Furthermore, $\gamma_2 = 6$ does not raise numerical problems. With $\beta_{22} = 0.1$ and $N = 2000$, only strong simultaneity ($\gamma_2 = \pm 6$) leads to undersizing (for $N = 2000$), i.e. the results are more stable for moderate values of $\beta_{22}$ than those for the z-test.

## 5 Conclusion

The paper analyses weak identification in probit models with endogenous covariates. It shows remarkable level distortions concerning the usual z-test. However, further research is needed to clarify why the magnitude depends heavily on the parameter value tested. The likelihood ratio statistic seems to be a conservative alternative which is robust to weak identification. Further research is useful to clarify how advanced methods like those of Andrews and Cheng (2014), Dufour (2006) or Kleibergen (2005) will work for probit models with endogenous covariates.

## References

Abramitzky R, Lavy V (2014) How responsive is investment in schooling to changes in redistributive policies and in returns? Econometrica 82(4):1241–1272

Andrews DW, Cheng X (2013) Maximum likelihood estimation and uniform inference with sporadic identification failure. Journal of Econometrics 173(1):36–56

Andrews DW, Cheng X (2014) GMM estimation and uniform subvector inference with possible identification failure. Econometric Theory 30(2):287–333

Beck T, Lin C, Ma Y (2014) Why do firms evade taxes? The role of information sharing and financial sector outreach. The Journal of Finance 69(2):763–817

Bijsterbosch M, Dahlhaus T (2015) Key features and determinants of credit-less recoveries. Empirical Economics 49(4):1245–1269

Blundell RW, Smith RJ (1993) Simultaneous microeconometric models with censored or qualitative dependent variables, in G.S. Maddala, C.R. Rao and H.D. Vinod (ed.) Handbook of statistics, vol 11, Econometrics, Amsterdam: North Holland, pp 117–143

Bouoiyour J, Miftah A, Mouhoud EM (2016) Education, male gender preference and migrants' remittances: Interactions in rural Morocco. Economic Modelling 57:324–331

Chaussé P (2010) Computing generalized method of moments and generalized empirical likelihood with R. Journal of Statistical Software 34(11):1–35

Cornelli F, Kominek Z, Ljungqvist A (2013) Monitoring managers: Does it matter? The Journal of Finance 68(2):431–481

Croushore D, Marsten K (2016) Reassessing the relative power of the yield spread in forecasting recessions. Journal of Applied Econometrics 31(6):1183–1191

Dufour JM (1997) Some impossibility theorems in econometrics with applications to structural and dynamic models. Econometrica 65(6):1365–1387

Dufour JM (2003) Identification, weak instruments, and statistical inference in econometrics. Canadian Journal of Economics/Revue canadienne d'économique 36(4):767–808

Dufour JM (2006) Monte carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics. Journal of Econometrics 133(2):443–477

Engelhardt GV, Eriksen MD, Gale WG, Mills GB (2010) What are the social benefits of homeownership? Experimental evidence for low-income households. Journal of Urban Economics 67(3):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 & Finance 34(6):1109–1128

Fitzenberger B, Kohn K, Wang Q (2011) The erosion of union membership in germany: determinants, densities, decompositions. Journal of Population Economics 24(1):141–165

Greene WH (2008) Econometric analysis 6th edition. Prentice Hall, Upper Saddle River

Haider A, Jahangir A (2017) La familia – how trust towards family decreases female labor force participation. Journal of Labor Research 38(1):122–144

Hao L, Ng EC (2011) Predicting Canadian recessions using dynamic probit modelling approaches. Canadian Journal of Economics/Revue canadienne d'économique 44(4):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, pp 3–30

Hlaing KP, 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(4):557–576

Horvath R, Katuscakova D (2016) Transparency and trust: the case of the European central bank. Applied Economics 48(57):5625–5638

Khanna V, Kim EH, Lu Y (2015) CEO connectedness and corporate fraud. The Journal of Finance 70(3):1203–1252

Kleibergen F (2005) Testing parameters in GMM without assuming that they are identified. Econometrica 73(4):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(3):731–741

Magnusson LM (2007) Weak instruments robust tests for limited dependent variable models. Tech. rep., Working Paper, Brown University (RI)

Magnusson LM (2010) Inference in limited dependent variable models robust to weak identification. The Econometrics Journal 13(3):S56–S79

Massa M, Zhang L (2013) Monetary policy and regional availability of debt financing. Journal of Monetary Economics 60(4):439–458

R Core Team (2016) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL `https://www.R-project.org/`

Stock JH, Wright JH, Yogo M (2002) A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics 20(4):518–529

Wen JF, Gordon DV (2014) An empirical model of tax convexity and self-employment. Review of Economics and Statistics 96(3):471–482

Wilde J (2008) A note on GMM estimation of probit models with endogenous regressors. Statistical Papers 49(3):471–484

**Table 1** Rejection frequencies of the z-test, $H_0 : \gamma_1 = 0$

| | | Nominal level | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\beta_{22} = 1$ | | | $\beta_{22} = 0.1$ | | | $\beta_{22} = 0.0001$ | | |
| $\gamma_2$ | $N$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| $-6$ | 2000 | 0.0908 | 0.0484 | 0.0178 | 0.1634 | 0.1142 | 0.0512 | 0.1690 | 0.1114 | 0.0464 |
| | 400 | 0.0984 | 0.0688 | 0.0320 | 0.1764 | 0.1206 | 0.0500 | 0.1700 | 0.1196 | 0.0514 |
| $-3$ | 2000 | 0.0950 | 0.0502 | 0.0118 | 0.1310 | 0.0880 | 0.0414 | 0.1424 | 0.0916 | 0.0350 |
| | 400 | 0.0868 | 0.0508 | 0.0206 | 0.1528 | 0.1018 | 0.0394 | 0.1452 | 0.0938 | 0.0382 |
| $-1.5$ | 2000 | 0.0988 | 0.0506 | 0.0114 | 0.0942 | 0.0626 | 0.0242 | 0.0854 | 0.0508 | 0.0124 |
| | 400 | 0.0964 | 0.0482 | 0.0136 | 0.1044 | 0.0640 | 0.0196 | 0.0896 | 0.0528 | 0.0144 |
| $-0.5$ | 2000 | 0.0984 | 0.0500 | 0.0102 | 0.0700 | 0.0340 | 0.0058 | 0.0120 | 0.0030 | 0.0000 |
| | 400 | 0.0996 | 0.0510 | 0.0116 | 0.0432 | 0.0202 | 0.0012 | 0.0148 | 0.0056 | 0.0006 |
| 0.5 | 2000 | 0.0986 | 0.0470 | 0.0110 | 0.0698 | 0.0320 | 0.0080 | 0.0128 | 0.0046 | 0.0004 |
| | 400 | 0.0974 | 0.0500 | 0.0102 | 0.0382 | 0.0154 | 0.0016 | 0.0126 | 0.0038 | 0.0006 |
| 1.5 | 2000 | 0.0970 | 0.0486 | 0.0100 | 0.0914 | 0.0626 | 0.0254 | 0.0836 | 0.0458 | 0.0136 |
| | 400 | 0.0966 | 0.0480 | 0.0102 | 0.0960 | 0.0566 | 0.0178 | 0.0896 | 0.0480 | 0.0128 |
| 3 | 2000 | 0.0960 | 0.0468 | 0.0108 | 0.1212 | 0.0866 | 0.0388 | 0.1430 | 0.0922 | 0.0342 |
| | 400 | 0.0864 | 0.0464 | 0.0160 | 0.1408 | 0.0914 | 0.0332 | 0.1488 | 0.0976 | 0.0340 |
| 6 | 2000 | 0.0886 | 0.0462 | 0.0160 | 0.1548 | 0.1076 | 0.0484 | 0.1674 | 0.1138 | 0.0458 |
| | 400 | 0.0930 | 0.0658 | 0.0276 | 0.1696 | 0.1130 | 0.0454 | 0.1768 | 0.1196 | 0.0482 |

**Table 2** Rejection frequencies of the z-test, $H_0 : \gamma_1 = 2$

| | | Nominal level | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\beta_{22} = 1$ | | | $\beta_{22} = 0.1$ | | | $\beta_{22} = 0.0001$ | | |
| $\gamma_2$ | $N$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| $-6$ | 2000 | 0.0890 | 0.0556 | 0.0234 | 0.2724 | 0.2280 | 0.1696 | 0.5586 | 0.5142 | 0.4366 |
| | 400 | 0.1038 | 0.0726 | 0.0382 | 0.4480 | 0.4078 | 0.3322 | 0.5732 | 0.5260 | 0.4540 |
| $-3$ | 2000 | 0.0996 | 0.0518 | 0.0144 | 0.1586 | 0.1244 | 0.0810 | 0.3686 | 0.3100 | 0.2076 |
| | 400 | 0.0842 | 0.0480 | 0.0200 | 0.2772 | 0.2314 | 0.1634 | 0.3756 | 0.3150 | 0.2180 |
| $-1.5$ | 2000 | 0.0970 | 0.0498 | 0.0108 | 0.1042 | 0.0756 | 0.0320 | 0.1562 | 0.0996 | 0.0404 |
| | 400 | 0.0932 | 0.0490 | 0.0116 | 0.1460 | 0.1002 | 0.0434 | 0.1632 | 0.1040 | 0.0432 |
| $-0.5$ | 2000 | 0.0990 | 0.0468 | 0.0108 | 0.0684 | 0.0316 | 0.0056 | 0.0110 | 0.0026 | 0.0000 |
| | 400 | 0.0928 | 0.0490 | 0.0086 | 0.0418 | 0.0150 | 0.0012 | 0.0126 | 0.0032 | 0.0002 |
| 1.5 | 2000 | 0.0996 | 0.0494 | 0.0102 | 0.0864 | 0.0540 | 0.0218 | 0.0694 | 0.0358 | 0.0078 |
| | 400 | 0.0914 | 0.0424 | 0.0084 | 0.0856 | 0.0522 | 0.0166 | 0.0700 | 0.0400 | 0.0084 |
| 3 | 2000 | 0.0946 | 0.0472 | 0.0116 | 0.1576 | 0.1182 | 0.0708 | 0.3184 | 0.2530 | 0.1554 |
| | 400 | 0.0774 | 0.0426 | 0.0146 | 0.2496 | 0.1968 | 0.1246 | 0.3246 | 0.2608 | 0.1620 |
| 6 | 2000 | 0.0872 | 0.0482 | 0.0198 | 0.2640 | 0.2260 | 0.1674 | 0.5416 | 0.4966 | 0.4124 |
| | 400 | 0.1036 | 0.0696 | 0.0334 | 0.4446 | 0.4004 | 0.3256 | 0.5592 | 0.5152 | 0.4304 |

**Table 3** Rejection frequencies of the z-test, sample design with variances =1

| | | | Nominal level | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\beta_{22}=1$ | | | $\beta_{22}=0.1$ | | | $\beta_{22}=0.0001$ | | |
| $\gamma_1$ | $\gamma_2$ | $N$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 0 | 6 | 2000 | 0.0908 | 0.0474 | 0.0162 | 0.1278 | 0.0834 | 0.0352 | 0.1424 | 0.0898 | 0.0330 |
| | | 400 | 0.0954 | 0.0618 | 0.0252 | 0.1358 | 0.0852 | 0.0340 | 0.1480 | 0.0946 | 0.0332 |
| 2 | 6 | 2000 | 0.0808 | 0.0510 | 0.0176 | 0.3126 | 0.2448 | 0.1708 | 0.5718 | 0.5084 | 0.4078 |
| | | 400 | 0.1054 | 0.0724 | 0.0352 | 0.4376 | 0.3902 | 0.3110 | 0.5464 | 0.5034 | 0.4194 |

**Table 4** Rejection frequencies of the z-test in case of underidentification, i.e. $\beta_{22}=0$

| | | Nominal level | | |
|---|---|---|---|---|
| $\gamma_2$ | $N$ | 10% | 5% | 1% |
| 6 | 2000 | 0.1674 | 0.1138 | 0.0458 |
| | 400 | 0.1770 | 0.1194 | 0.0482 |
| 6 | 2000 | 0.5424 | 0.4964 | 0.4116 |
| | 400 | 0.5592 | 0.5150 | 0.4300 |

**Table 5** Rejection frequencies of the z-test, nominal level 5%, $N=2000$, with overidentification of $\gamma_1$ in the theoretical model

| | | Weighing matrix | | | | | |
|---|---|---|---|---|---|---|---|
| | | $\beta_{22}=\beta_{23}=1$ | | $\beta_{22}=\beta_{23}=0.1$ | | $\beta_{22}=\beta_{23}=0.0001$ | |
| $\gamma_1$ | $\gamma_2$ | $I$ | Optimal | $I$ | Optimal | $I$ | Optimal |
| 0 | −6 | 0.0526 | 0.0450 | 0.2396 | 0.1724 | 0.2682 | 0.2072 |
| 0 | −3 | 0.0484 | 0.0464 | 0.1578 | 0.1100 | 0.2304 | 0.1758 |
| 0 | −1.5 | 0.0466 | 0.0464 | 0.0818 | 0.0686 | 0.1350 | 0.1174 |
| 0 | −0.5 | 0.0466 | 0.0464 | 0.0456 | 0.0488 | 0.0112 | 0.0122 |
| 0 | 0.5 | 0.0470 | 0.0456 | 0.0458 | 0.0456 | 0.0124 | 0.0150 |
| 0 | 1.5 | 0.0474 | 0.0460 | 0.0826 | 0.0710 | 0.1310 | 0.1232 |
| 0 | 3 | 0.0492 | 0.0464 | 0.1482 | 0.1062 | 0.2252 | 0.1904 |
| 0 | 6 | 0.0548 | 0.0496 | 0.2256 | 0.1692 | 0.2686 | 0.2120 |
| 2 | −6 | 0.0376 | 0.0454 | 0.3580 | 0.1574 | 0.7844 | 0.0990 |
| 2 | −3 | 0.0452 | 0.0486 | 0.1952 | 0.1130 | 0.5708 | 0.1126 |
| 2 | −1.5 | 0.0470 | 0.0456 | 0.0888 | 0.0704 | 0.2416 | 0.1156 |
| 2 | −0.5 | 0.0442 | 0.0458 | 0.0440 | 0.0446 | 0.0108 | 0.0154 |
| 2 | 1.5 | 0.0538 | 0.0548 | 0.0714 | 0.0598 | 0.1100 | 0.1104 |
| 2 | 3 | 0.0460 | 0.0504 | 0.2006 | 0.1106 | 0.4968 | 0.2102 |
| 2 | 6 | 0.0470 | 0.0542 | 0.3882 | 0.1606 | 0.7736 | 0.1344 |

**Table 6** Rejection frequencies of the LR test

| | | | Nominal level | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\beta_{22} = 1$ | | | $\beta_{22} = 0.1$ | | | $\beta_{22} = 0.0001$ | | |
| $\gamma_1$ | $\gamma_2$ | $N$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 0 | $-3$ | 2000 | 0.1010 | 0.0494 | 0.0100 | 0.0960 | 0.0464 | 0.0082 | 0.0654 | 0.0283 | 0.0048 |
| | | 400 | 0.0992 | 0.0518 | 0.0128 | 0.0798 | 0.0388 | 0.0090 | 0.0698 | 0.0324 | 0.0066 |
| 0 | $-1.5$ | 2000 | 0.1000 | 0.0492 | 0.0106 | 0.0998 | 0.0490 | 0.0104 | 0.0828 | 0.0366 | 0.0062 |
| | | 400 | 0.1000 | 0.0514 | 0.0118 | 0.0946 | 0.0476 | 0.0108 | 0.0834 | 0.0408 | 0.0072 |
| 0 | $-0.5$ | 2000 | 0.0986 | 0.0488 | 0.0120 | 0.0986 | 0.0488 | 0.0120 | 0.0958 | 0.0478 | 0.0108 |
| | | 400 | 0.0996 | 0.0522 | 0.0112 | 0.0996 | 0.0520 | 0.0110 | 0.0956 | 0.0510 | 0.0110 |
| 0 | 0.5 | 2000 | 0.1000 | 0.0488 | 0.0112 | 0.1000 | 0.0488 | 0.0112 | 0.0966 | 0.0484 | 0.0098 |
| | | 400 | 0.0996 | 0.0534 | 0.0116 | 0.0992 | 0.0530 | 0.0114 | 0.0962 | 0.0518 | 0.0110 |
| 0 | 1.5 | 2000 | 0.0988 | 0.0538 | 0.0110 | 0.0986 | 0.0536 | 0.0106 | 0.0810 | 0.0410 | 0.0080 |
| | | 400 | 0.0990 | 0.0510 | 0.0106 | 0.0948 | 0.0434 | 0.0086 | 0.0806 | 0.0414 | 0.0070 |
| 0 | 3 | 2000 | 0.0998 | 0.0548 | 0.0106 | 0.0956 | 0.0488 | 0.0076 | 0.0664 | 0.0340 | 0.0054 |
| | | 400 | 0.0962 | 0.0496 | 0.0102 | 0.0732 | 0.0324 | 0.0042 | 0.0676 | 0.0298 | 0.0050 |
| 2 | $-6$ | 2000 | 0.1014 | 0.0536 | 0.0100 | 0.0648 | 0.0312 | 0.0066 | 0.0510 | 0.0240 | 0.0042 |
| | | 400 | 0.0956 | 0.0514 | 0.0114 | 0.0482 | 0.0244 | 0.0058 | 0.0456 | 0.0238 | 0.0044 |
| 2 | $-3$ | 2000 | 0.0974 | 0.0520 | 0.0092 | 0.0942 | 0.0482 | 0.0092 | 0.0554 | 0.0240 | 0.0046 |
| | | 400 | 0.0990 | 0.0504 | 0.0106 | 0.0620 | 0.0300 | 0.0068 | 0.0498 | 0.0248 | 0.0050 |
| 2 | $-1.5$ | 2000 | 0.0940 | 0.0466 | 0.0096 | 0.0978 | 0.0514 | 0.0114 | 0.0740 | 0.0350 | 0.0056 |
| | | 400 | 0.0972 | 0.0524 | 0.0078 | 0.0826 | 0.0416 | 0.0070 | 0.0682 | 0.0328 | 0.0066 |
| 2 | 0.5 | 2000 | 0.0974 | 0.0478 | 0.0104 | 0.0968 | 0.0478 | 0.0124 | 0.0936 | 0.0472 | 0.0102 |
| | | 400 | 0.1010 | 0.0462 | 0.0078 | 0.0954 | 0.0486 | 0.0086 | 0.0950 | 0.0456 | 0.0092 |
| 2 | 1.5 | 2000 | 0.0990 | 0.0514 | 0.0100 | 0.0994 | 0.0522 | 0.0096 | 0.0808 | 0.0404 | 0.0084 |
| | | 400 | 0.0922 | 0.0496 | 0.0092 | 0.0914 | 0.0474 | 0.0062 | 0.0814 | 0.0380 | 0.0072 |
| 2 | 3 | 2000 | 0.0996 | 0.0492 | 0.0104 | 0.0926 | 0.0454 | 0.0062 | 0.0588 | 0.0284 | 0.0056 |
| | | 400 | 0.0940 | 0.0476 | 0.0074 | 0.0708 | 0.0302 | 0.0020 | 0.0538 | 0.0258 | 0.0046 |
| 2 | 6 | 2000 | 0.0994 | 0.0492 | 0.0106 | 0.0678 | 0.0276 | 0.0046 | 0.0492 | 0.0268 | 0.0062 |
| | | 400 | 0.0894 | 0.0462 | 0.0084 | 0.0492 | 0.0218 | 0.0036 | 0.0462 | 0.0238 | 0.0052 |

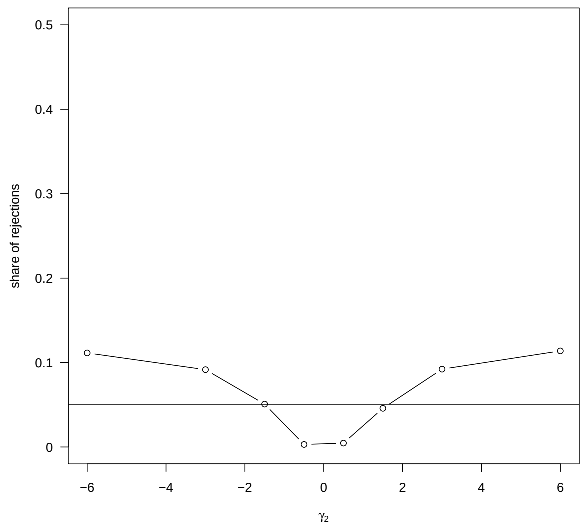


**Fig. 1** Rejection frequencies of the z-test under weak identification, $H_0 : \gamma_1 = 0$, nominal level 5%, N = 2000

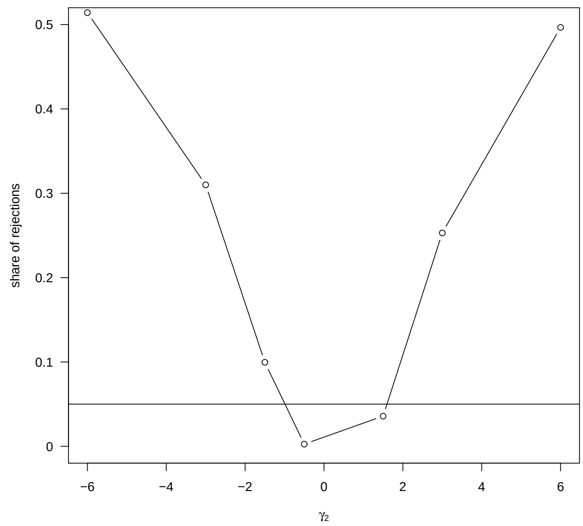


**Fig. 2** Rejection frequencies of the z-test under weak identification, $H_0 : \gamma_1 = 2$, nominal level 5%, N = 2000

## Appendix: Simulation results for model (5) with $\beta_{11} = \pi_{21} = 0$

**Table 7** Rejection frequencies of the z-test, $H_0 : \gamma_1 = 0$

| | | Nominal level | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\beta_{22} = 1$ | | | $\beta_{22} = 0.1$ | | | $\beta_{22} = 0.0001$ | | |
| $\gamma_2$ | $N$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| −6 | 2000 | 0.0888 | 0.0482 | 0.0178 | 0.1586 | 0.1118 | 0.0516 | 0.1724 | 0.1172 | 0.0474 |
| | 400 | 0.0982 | 0.0668 | 0.0308 | 0.1774 | 0.1188 | 0.0544 | 0.1776 | 0.1232 | 0.0498 |
| −3 | 2000 | 0.0982 | 0.0470 | 0.0120 | 0.1248 | 0.0894 | 0.0408 | 0.1508 | 0.0942 | 0.0350 |
| | 400 | 0.0866 | 0.0478 | 0.0190 | 0.1516 | 0.1012 | 0.0396 | 0.1534 | 0.1010 | 0.0350 |
| −1.5 | 2000 | 0.0998 | 0.0484 | 0.0090 | 0.0930 | 0.0636 | 0.0242 | 0.0892 | 0.0490 | 0.0140 |
| | 400 | 0.0960 | 0.0468 | 0.0126 | 0.1042 | 0.0608 | 0.0214 | 0.0940 | 0.0528 | 0.0132 |
| −0.5 | 2000 | 0.1002 | 0.0502 | 0.0098 | 0.0714 | 0.0340 | 0.0078 | 0.0128 | 0.0030 | 0.0000 |
| | 400 | 0.0958 | 0.0470 | 0.0110 | 0.0456 | 0.0188 | 0.0018 | 0.0144 | 0.0050 | 0.0002 |
| 0.5 | 2000 | 0.1018 | 0.0510 | 0.0074 | 0.0712 | 0.0330 | 0.0096 | 0.0116 | 0.0042 | 0.0002 |
| | 400 | 0.0948 | 0.0478 | 0.0096 | 0.0398 | 0.0172 | 0.0026 | 0.0132 | 0.0046 | 0.0004 |
| 1.5 | 2000 | 0.0992 | 0.0492 | 0.0088 | 0.0956 | 0.0618 | 0.0242 | 0.0860 | 0.0488 | 0.0132 |
| | 400 | 0.0916 | 0.0482 | 0.0100 | 0.0972 | 0.0602 | 0.0194 | 0.0904 | 0.0508 | 0.0124 |
| 3 | 2000 | 0.0964 | 0.0492 | 0.0094 | 0.1264 | 0.0888 | 0.0392 | 0.1494 | 0.0942 | 0.0342 |
| | 400 | 0.0862 | 0.0464 | 0.0152 | 0.1438 | 0.0956 | 0.0370 | 0.1556 | 0.1004 | 0.0360 |
| 6 | 2000 | 0.0902 | 0.0482 | 0.0136 | 0.1596 | 0.1108 | 0.0504 | 0.1736 | 0.1192 | 0.0486 |
| | 400 | 0.0944 | 0.0626 | 0.0280 | 0.1706 | 0.1166 | 0.0486 | 0.1800 | 0.1236 | 0.0482 |

**Table 8** Rejection frequencies of the z-test, $H_0 : \gamma_1 = 2$

| | | Nominal level | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\beta_{22} = 1$ | | | $\beta_{22} = 0.1$ | | | $\beta_{22} = 0.0001$ | | |
| $\gamma_2$ | $N$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| −6 | 2000 | 0.0874 | 0.0542 | 0.0216 | 0.2668 | 0.2282 | 0.1668 | 0.5634 | 0.5166 | 0.4388 |
| | 400 | 0.1036 | 0.0702 | 0.0406 | 0.4466 | 0.4058 | 0.3358 | 0.5700 | 0.5276 | 0.4510 |
| −3 | 2000 | 0.0968 | 0.0486 | 0.0128 | 0.1588 | 0.1248 | 0.0802 | 0.3724 | 0.3128 | 0.2164 |
| | 400 | 0.0826 | 0.0520 | 0.0198 | 0.2818 | 0.2362 | 0.1602 | 0.3756 | 0.3162 | 0.2214 |
| −1.5 | 2000 | 0.0922 | 0.0462 | 0.0104 | 0.1058 | 0.0726 | 0.0330 | 0.1584 | 0.1020 | 0.0408 |
| | 400 | 0.0976 | 0.0500 | 0.0110 | 0.1422 | 0.0946 | 0.0416 | 0.1628 | 0.1046 | 0.0408 |
| −0.5 | 2000 | 0.0938 | 0.0464 | 0.0100 | 0.0676 | 0.0318 | 0.0062 | 0.0124 | 0.0034 | 0.0000 |
| | 400 | 0.1004 | 0.0486 | 0.0078 | 0.0412 | 0.0148 | 0.0016 | 0.0142 | 0.0044 | 0.0002 |
| 1.5 | 2000 | 0.0994 | 0.0514 | 0.0102 | 0.0904 | 0.0580 | 0.0230 | 0.0712 | 0.0368 | 0.0078 |
| | 400 | 0.0956 | 0.0412 | 0.0096 | 0.0840 | 0.0510 | 0.0174 | 0.0736 | 0.0406 | 0.0092 |
| 3 | 2000 | 0.0986 | 0.0510 | 0.0106 | 0.1558 | 0.1214 | 0.0758 | 0.3184 | 0.2576 | 0.1604 |
| | 400 | 0.0830 | 0.0436 | 0.0154 | 0.2488 | 0.1966 | 0.1208 | 0.3266 | 0.2670 | 0.1630 |
| 6 | 2000 | 0.0922 | 0.0510 | 0.0188 | 0.2648 | 0.2252 | 0.1654 | 0.5490 | 0.5020 | 0.4210 |
| | 400 | 0.1044 | 0.0702 | 0.0324 | 0.4402 | 0.3968 | 0.3240 | 0.5604 | 0.5144 | 0.4304 |