# Heteroskedasticity in One-Way Error Component Probit Models

**Richard Kouamé Moussa**

ENSEA, Abidjan 08, Cote D'lvoire; richard.moussa@ensea.ed.ci

**Abstract:** This paper introduces an estimation procedure for a random effects probit model in presence of heteroskedasticity and a likelihood ratio test for homoskedasticity. The cases where the heteroskedasticity is due to individual effects or idiosyncratic errors or both are analyzed. Monte Carlo simulations show that the test performs well in the case of high degree of heteroskedasticity. Furthermore, the power of the test increases with larger individual and time dimensions. The robustness analysis shows that applying the wrong approach may generate misleading results except for the case where both individual effects and idiosyncratic errors are modelled as heteroskedastic.

**Keywords:** heteroskedasticity; probit; panel data; Gauss–Hermite quadrature; Monte Carlo simulation

**JEL Classification:** C23; C61; C63

## 1. Introduction

The problem that heteroskedasticity presents for panel data regression has been widely discussed in the literature (Baltagi 2008; Baltagi et al. 2006; Montes-Rojas and Sosa-Escudero 2011). Let us consider the one-way error component model, i.e., with the error term defined as $u_{it} = \mu_i + v_{it}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$ where the individual effects $\mu_i$ and the idiosyncratic errors $v_{it}$ are assumed to be random (i.e., $\mu_i \sim iid(0, \sigma_\mu^2)$ and $v_{it} \sim iid(0, \sigma_v^2)$). Several authors (Baltagi 1988; Mazodier and Trognon 1978; Randolph 1988; Wansbeek 1989, among others) consider different types of heteroskedasticity depending upon whether individual effects ($\mu_i$) or idiosyncratic errors ($v_{it}$) or both are heteroskedastic. Baltagi et al. (2006) and later Montes-Rojas and Sosa-Escudero (2011) proposed Lagrange Multiplier (LM) test procedures to check for the presence of heteroskedasticity in linear models for various cases. However, such test procedures for panel data binary choice models are lacking. In addition, to the best of my knowledge, there is no existing procedure to estimate an heteroskedastic probit model on panel data.

The use of random effects probit models panel data has been popularized due to the problem of incidental parameters (Baltagi 2008; Lancaster 2000). Since this model is generally applied to micro-panels, heteroskedasticity problems are likely to arise. One must account for heteroskedasticity since it could result in misleading conclusions about coefficients and marginal effects interpretation (Greene 2012). Two approaches are used to calculate the marginal effects after probit models in applied works: (i) integrating with respect to individual effects, or (ii) assuming the individual effects to be null (Bland and Cook 2018). In the case (i), the probability of positive outcome is given by $Pr(y_{it} = 1|X_{it}) = \Phi\left(\frac{X_{it}\beta}{\sqrt{\sigma_v^2 + \sigma_\mu^2}}\right)$ while in case (ii), this probability is given by $Pr(y_{it} = 1|X_{it}, \mu_i = 0) = \Phi\left(\frac{X_{it}\beta}{\sigma_v}\right)$, where $\Phi$ and $\phi$ are respectively the standard normal cumulative probability and the standard normal

density functions. Thus, the marginal effect for variable $x_k$ (denoted $me(x_k)$) is given by Equation (1) for case (i) and Equation (2) for case (ii):

$$me(x_k) \quad = \quad \frac{\beta_k}{\sqrt{\sigma_\nu^2 + \sigma_\mu^2}} \phi\left(\frac{X_{it}\beta}{\sqrt{\sigma_\nu^2 + \sigma_\mu^2}}\right) \tag{1}$$

$$me(x_k) \quad = \quad \frac{\beta_k}{\sigma_\nu} \phi\left(\frac{X_{it}\beta}{\sigma_\nu}\right). \tag{2}$$

In Equations (1) and (2), it clearly appears that the marginal effects estimated in both case (i) and (ii) depend on the variance components. Since these variance components are functions of individual characteristics in the presence of heteroskedasticity, thus considering an homoskedastic model yields misestimated marginal effects. This stresses the need for an empirical implementation of an estimation and test procedure to deal with heteroskedasticity on panel probit models.

The aim of this paper is to introduce an estimation procedure that accounts for this heteroskedasticity using the Gauss-Hermite quadrature scheme[1]. In addition, the papers aims at providing a likelihood ratio (LR) test procedure for homoskedasticity in a panel probit model that allows one to investigate various forms of heteroskedasticity under alternative hypothesis. Monte Carlo simulations are conducted to estimate the power and the empirical size of the test and a robustness analysis is completed to ensure that the test and estimation procedures perform well. Results suggest that the estimation procedure has good performances and that it performance also depends on the quadrature parameters. The LR test has excellent power when there is high degree of heteroskedasticity and its performance depends on sample size in the situation of low degree of heteroskedasticity. The contribution of this paper to the literature is twofold. Firstly, it introduces a procedure to estimate a panel probit model with heteroskedasticity. The procedure allows one to deal with different sources of heteroskedasticity. Secondly, based on the power and the empirical size of the test, it shows that the LR test for homoskedasticity has good performance. The robustness of the estimation procedure and the test performance have been assessed using an extensive Monte Carlo simulation.

The rest of this paper is organized as follows. Section 2 presents the different forms of heteroskedasticity encountered in the literature and derives the likelihood estimator in a general setting. Section 3 discusses the estimation requirements and test procedures to deal with heteroskedasticity. In Section 4, the power and the empirical size of the test as well as the bias and the mean square error of the estimated parameters are computed based on Monte Carlo simulations. Section 5 presents robustness analysis. Section 6 presents a case study that illustrates the estimation of the parameters and marginal effects in the presence of heteroskedasticity. Section 7 concludes.

## 2. Heteroskedasticity and Likelihood Function

This section discusses the different types of heteroskedasticity encountered in the literature and specifies the likelihood function.

### 2.1. Different Sources of Heteroskedasticity

Consider the following one-way error components probit model:

$$y_{it} = \mathbb{1}_{\mathbb{R}^+}\left(X_{it}\beta + u_{it}\right) \quad \forall i = 1, \ldots, N \; ; t = 1, \ldots, T, \tag{3}$$

---

[1] A user-written Stata's ado file is provided to deal with these purposes. This ado file is an extension of the existing Stata's *hetprobit* and *xtprobit, re* commands that accounts for each of the types of heteroskedasticity observed in panel one-way error component models in the literature. A Stata code for computing the marginal effects after the proposed estimation procedure is given in the Appendix A.

where $u_{it}$ is decomposed in individual unobserved effects ($\mu_i$) and idiosyncratic errors ($\nu_{it}$). We consider the random effects model. Classical assumptions for the estimation of a random effects model are the following: (i) the individual effects $\mu_i$ are independent from the idiosyncratic errors $\nu_{it}$, and (ii) the explanatory variables $X_{it}$ are independent from the individual effects $\mu_i$ and the idiosyncratic errors $\nu_{it}$. In addition, some assumptions are made on the variance components to deal with heteroskedasticity issues. These assumptions lead to three cases of heteroskedasticity identified in the literature:

- Heteroskedasticity a la Mazodier and Trognon (1978): The heteroskedasticity is due to the individual effects. Thus, $\mu_i \sim iid(0, \sigma^2_{\mu_i})$ and $\nu_{it} \sim iid(0, \sigma^2_\nu)$.
- Heteroskedasticity a la Baltagi (1988) and Wansbeek (1989): the heteroskedasticity is due to the idiosyncratic errors. Thus, $\mu_i \sim iid(0, \sigma^2_\mu)$ and $\nu_{it} \sim iid(0, \sigma^2_{\nu_{it}})$.
- Heteroskedasticity a la Randolph (1988): The heteroskedasticity is due to both the individual effects and the idiosyncratic errors. Thus, $\mu_i \sim iid(0, \sigma^2_{\mu_i})$ and $\nu_{it} \sim iid(0, \sigma^2_{\nu_{it}})$. An alternative specification by Verbon (1980) is to consider that $\mu_i \sim iid(0, \sigma^2_{\mu_i})$ and $\nu_{it} \sim iid(0, \sigma^2_{\nu_i})$.

In this paper, the heteroskedastic component is assumed to a function of some observed variables. More specifically, when the heteroskedasticity is due to the $\mu_i$, the variance depends on time-invariant exogenous variables $Z_{\mu_i}$ and expressed as $\sigma_{\mu_i} = \sigma_\mu h_\mu \left( Z'_{\mu_i} \theta_\mu \right)$. Alternatively, when the heteroskedasticity is due to the $\nu_{it}$, the variance depends on exogenous variables $Z_{\nu_{it}}$ and has the following expression: $\sigma_{\nu_{it}} = \sigma_\nu h_\nu \left( Z'_{\nu_{it}} \theta_\nu \right)$. The approach of Verbon (1980) can be modelled with a variance of idiosyncratic errors that is $\sigma_{\nu_i} = \sigma_\nu h_\nu \left( Z'_{\nu_i} \theta_\nu \right)$. The functions $h_\mu(.)$ and $h_\nu(.)$ are twice continuously differentiable and satisfy $h_\mu(.) > 0$, $h_\nu(.) > 0$, $h_\mu(0) = 1$ and $h_\nu(0) = 1$. $Z_{\mu_i}$ and $Z_{\nu_{it}}$ are vectors of regressors and have no constant term included. Note that the variance of the idiosyncratic errors is set to one ($\sigma_\nu = 1$) in order to avoid identification problems (Greene 2012). This identification problem occurs when a constant term is included since it implies that the variance of idiosyncratic errors will not be 1 when $\theta_\nu$ is null.

In the rest of the paper, as in Montes-Rojas and Sosa-Escudero (2011) and Baltagi et al. (2006), the results are reported for the functions $h_\mu(.)$ and $h_\nu(.)$ set to exponential functions. Then, $h_\mu \left( Z'_{\mu_i} \theta_\mu \right) = exp \left( Z'_{\mu_i} \theta_\mu \right)$ and $h_\nu \left( Z'_{\nu_{it}} \theta_\nu \right) = exp \left( Z'_{\nu_{it}} \theta_\nu \right)$. Thus, the variance of the individual effects can be rewritten as $\sigma_{\mu_i} = \sigma_\mu exp \left( Z'_{\mu_i} \theta_\mu \right) = exp \left( \lambda_0 + Z'_{\mu_i} \theta_\mu \right)$, with $\lambda_0 = log(\sigma_\mu)$.

*2.2. Likelihood Function*

The individual level likelihood is given by:

$$L_i = \int_{\mathbb{R}} \left[ \prod_{t=1}^{T_i} \Phi(\epsilon_{it}) \right] \phi_\mu(\mu_i) d\mu_i \tag{4}$$

where $\Phi$ denotes the standard normal cumulative distribution function, $\phi_\mu$ denotes the density function of a normal distribution with mean 0 and variance equal to the variance of $\mu_i$ that is $\sigma_{\mu_i} = \sigma_\mu h_\mu \left( Z'_{\mu_i} \theta_\mu \right)$, and

$$\epsilon_{it} = \frac{q_{it} \left( X_{it} \beta + \mu_i \right)}{h_\nu \left( Z'_{\nu_{it}} \theta_\nu \right)}$$

$$q_{it} = 2y_{it} - 1.$$

Note that this general form of the likelihood function allows dealing with homoskedasticity and each of the aforementioned heteroskedasticity cases. The homoskedastic model is given by $\theta_\mu = \theta_\nu = 0$. The heteroskedastic model where $\mu_i$ are heteroskedastic is given by $\theta_\mu \neq 0$ and $\theta_\nu = 0$; the heteroskedastic model where $\nu_{it}$ are heteroskedastic is given by $\theta_\mu = 0$ and $\theta_\nu \neq 0$; while the heteroskedastic model where both $\mu_i$ and $\nu_{it}$ are heteroskedastic is given by $\theta_\mu \neq 0$ and $\theta_\nu \neq 0$.

## 3. Estimation and Tests

The estimation procedure was based on a Gauss–Hermite quadrature scheme. This section discusses the requirements for the use of this approach. Then, the procedure to test for homoskedasticity is presented.

### 3.1. Estimation Requirements

Given that the likelihood function (Equation (4)) has an integral form, it is common in the literature to use a numerical integration method. Gauss–Hermite quadrature is used to provide an approximation of the likelihood function that has a tractable form for likelihood maximization algorithms (Liu and Pierce 1994; Naylor and Smith 1982). It consists in approximating the integral $\int_\mathbb{R} g(x)exp(-x^2)dx$ by a weighted sum of the function $g(.)$ taken at some specific points called nodes.

The Gauss–Hermite quadrature scheme used herein is the one proposed by Liu and Pierce (1994). Let $Q$ denotes the number of quadrature points, $x_q$, $q = 1, \ldots, Q$ and $w_q$, $q = 1, \ldots, Q$ denote respectively the quadrature points (nodes) and their corresponding weights. Then, the individual level likelihood is given in Equation (4) can be re-expressed as a sum of functions as follows:

$$L_i = \int_\mathbb{R} \left[ \prod_{t=1}^{T_i} \Phi(\epsilon_{it}) \right] \phi_\mu(\mu_i) d\mu_i = \sum_{q=1}^{Q} w_q^* g(x_q^*). \tag{5}$$

With

$$g(\mu_i) = \left[ \prod_{t=1}^{T_i} \Phi(\epsilon_{it}) \right] \phi_\mu(\mu_i)$$

$$x_q^* = \gamma + \sqrt{2}\sigma x_q$$

$$w_q^* = \sqrt{2}\sigma w_q exp(x_q^2)$$

$$\gamma = Arg\left( \max_{\mu_i} g(\mu_i) \right)$$

$$\sigma = \left( -\frac{\partial^2 log(g(\mu_i))}{\partial \mu_i^2} |_{\mu_i=\gamma} \right)^{-1/2}$$

$$\frac{\partial^2 log(g(\mu_i))}{\partial \mu_i^2} = -\frac{\phi(\epsilon_{it})\Phi(u_{it}) + \left( \phi(\epsilon_{it}) \right)^2}{\left( h_\nu \left( Z'_{\nu_{it}} \theta_\nu \right) \Phi(\epsilon_{it}) \right)^2} - \frac{1}{\left( h_\mu \left( Z'_{\mu_i} \theta_\mu \right) \right)^2},$$

where $\phi$ denotes the density function of the standard normal distribution[2].

The individual level log-likelihood function depends on the selected number of quadrature points $Q$. A discussion based on empirical applications of the effect of the number of quadrature points on the estimation results and the computing time is presented by Moussa and Delattre (2018). Researchers

---

[2] The estimation procedure described above has been implemented as a Stata user-written ado file using the Stata's *d0* procedure for maximum likelihood estimation (see Gould et al. 2010; Moussa and Delattre 2018).

can check the impact of a selected number of quadrature points on the results. A quadrature points check is conducted in Section 5 for the model with heteroskedasticity due to both $\mu_i$ and $\nu_{it}$ that is the most complete case of heteroskedasticity in panel models[3].

### 3.2. Test Procedure

As described by Greene (2012), the issue of heteroskedasticity test can be analyzed using a misspecification test procedure. The homoskedastic panel probit model can be viewed as a restricted model in which $\theta_\mu$ and $\theta_\mu$ are constrained to be null ($Z_{\mu_i}$ and $Z_{\nu_{it}}$ are omitted in the model). Thus, the homoskedastic probit model is nested in the heteroskedastic one. The omitted variables tests in literature are based on the likelihood-ratio (LR), the Lagrange multiplier (LM), and the Wald test. These three tests are asymptotically equivalent. However, this equivalence is valid for probit models only if the error components are homoskedastic and uncorrelated over time (Lechner 1995). The following relationship between test statistics for linear models has been proved ($Wald \geq LR \geq LM$, Johnston and DiNardo 2001). This implies that the LM test is less likely to reject the null hypothesis of homoskedasticity. In the literature, the LM test is mostly used to test for homoskedasticity in linear models even for panel data models (Baltagi et al. 2006; Montes-Rojas and Sosa-Escudero 2011). The LR test is mainly used for nonlinear models. On the cross-section probit model, Davidson and MacKinnon (1984) show that the LR test performs well. The Wald test has poor performance on finite sample when testing for nonlinear hypothesis (Davidson and MacKinnon 1984; Wooldridge 2001). The power of the LM test for homoskedasticity on probit model may be problematic since it fails to distinguish between heteroskedasticity and simple omission of a variable in the index function (Greene 2018; Davidson and MacKinnon 1984).

For the aforementioned reasons, the heteroskedasticity tests used herein are based on the LR test procedure. LR test addresses the issue of the change in model fit when new variables are added (Wooldridge 2001). Thus, it requires the estimation of both the full heteroskedastic and the homoskedastic models. Since the aim of this paper is to propose an estimation procedure of a random effects probit model for panel data in presence of heteroskedasticity, the LR statistics will be easy to compute. The LR test statistics is given by:

$$LR = 2\left( LogL_U - LogL_R \right) \sim \chi^2(p) \tag{6}$$

where $LogL_R$ and $LogL_U$ denote the log-likelihood of the restricted and unrestricted models respectively, and $p$ is the number of parameters that are omitted in the homoskedastic panel probit model, i.e., the dimension (number of column) of $Z_{\mu_i}$ or $Z_{\nu_{it}}$ or the sum of the dimensions of $Z_{\mu_i}$ and $Z_{\nu_{it}}$.

Following the sources of the heteroskedasticity and as specified by Baltagi et al. (2006), three types of hypothesis can be tested. These hypothesis are related to the joint test for homoskedasticity of both individual effects and idiosyncratic errors ($H_0 : \theta_\mu = \theta_\nu = 0$) and to the two marginal tests for homoskedasticity of one of the aforementioned error components assuming the other component homoskedastic (i.e., $H_0 : \theta_\mu = 0 \,|\, \theta_\nu = 0$ and $H_0 : \theta_\nu = 0 \,|\, \theta_\mu = 0$).

Monte Carlo simulations are conducted in Section 4 to check for the robustness of the test by estimating its power and empirical size. The power of the test is defined as the percentage of rejection at 5% significance level of the null hypothesis of homoskedasticity in presence of heteroskedasticity. The empirical size refers to the percentage of false rejection at 5% significance level of the null hypothesis of homoskedasticity.

---

[3] For all others applications presented herein, $Q = 10$ is used as the number of quadrature points.

## 4. Monte Carlo Experiments

The Monte Carlo[4] experiments conducted herein are based on a data generated as follows. For $i = 1, \ldots, N$ and $t = 1, \ldots, T$ the binary dependant variable is generated as:

$$y_{it} = \mathbb{1}_{\mathbb{R}^+}\left(\alpha_0 + \alpha_1 * X_{1_{it}} + \alpha_2 * X_{2_{it}} + \mu_i + \nu_{it}\right), \tag{7}$$

where $X_1$ and $X_2$ are generated from a random uniform distribution. The error components $\mu_i$ and $\nu_{it}$ are generated following a normal data generating process with zero mean and standard deviation $\sigma_{\mu_i} = \sigma_\mu h_\mu\left(Z'_{\mu_i}\theta_\mu\right)$ and $\sigma_{\nu_{it}} = h_\nu\left(Z'_{\nu_{it}}\theta_\nu\right)$ respectively. The time invariant variable $Z_{\mu_i}$ and the variable $Z_{\nu_{it}}$ are generated from a random uniform distribution. The parameters of the index function are set to $\alpha_0 = 1.5$, $\alpha_1 = 0.8$, and $\alpha_2 = -2$. For each type of heteroskedasticity presented in Section 2, nine (9) Monte Carlo experiments in which $N = \{50, 100, 500\}$ and $T = \{5, 10, 20\}$ are conducted with 5000 replications.

To estimate the power of the test, two cases are considered. The first set of experiments consists of a generated dataset with low degree of heteroskedasticity (i.e., setting $\theta_\mu = 0.7$ and $\theta_\nu = 0.6$). A second set of experiments consists of a generated dataset with a high degree of heteroskedasticity (i.e., setting $\theta_\mu = 2.1$ and $\theta_\nu = 1.8$) for each of the 27 models aforementioned. In these experiments, the variance of the individual effects is set to $\sigma_\mu^2 = 0.2$ (i.e., $\lambda_0 = -0.8$). The results of these experiments are presented in Section 4.1. The empirical size of the test is estimated using a generated dataset with no heteroskedasticity (i.e., setting $\theta_\mu = \theta_\nu = 0$). A well performing test would be such that its empirical size does not significantly differ from the nominal size of 5%. The results of these simulations are presented in Section 4.1.

Further Monte Carlo experiments have been conducted to cover several situations. These experiments consist of setting the following parameters: $\sigma_\mu^2 = \{2, 6\}$, $\theta_\mu = \{0, 1, 2, 3\}$ and $\theta_\nu = \{0, 1, 2, 3\}$. The results of this second set of experiments for $N = \{50, 500\}$ and $T = \{5, 20\}$ are reported in Tables A1–A3 in Appendix D.

### 4.1. Power and Empirical Size of the Test

Table 1 shows the power of the LR test for each of the aforementioned experiments. The Monte Carlo experiment for the marginal test $H_0 : \theta_\mu = 0 \,|\, \theta_\nu = 0$ reveals that the test performs well in the case of high degree of heteroskedasticity even for small samples ($N = 50$ and $T = 5$, the power of the LR test is 81.72%). However, in the case of low degree of heteroskedasticity, the test does not perform well on a sample with small $N$. For $N = 50$, the power of the test is 12.04% when $T = 5$ and increasing $T$ to 10 and 20 yields in a power of 19.9% and 27.4% respectively. But, the power of the LR test is very high for a sample with large $N$. For $N = 500$, the power of the test is respectively 69.54% for $T = 5$, 94.14% for $T = 10$ and 99.26% for $T = 20$. Nonetheless, increasing $\sigma_\mu^2$ from 0.2 a larger value, say 2 or 6 results in a decrease in the power of the test when $T$ is large and to an increase of the power of the test when $T$ is low (see Table A2 in Appendix D).

For the marginal test $H_0 : \theta_\nu = 0 \,|\, \theta_\mu = 0$, the Monte Carlo experiment shows that the performance of the test is mitigated for small samples in presence of high heteroskedasticity. With $N = 50$, when $T = 5$ the power of the test is 47.78%. However, increasing the time dimension to $T = 10$ and $T = 20$ the power of the test increases drastically to 93.7% and 99.98% respectively. In the case of low degree of heteroskedasticity, the power of the test is very high when $N$ or $T$ is large. When $N$ is small, the performance is low (22.86% with $N = 50$ and $T = 5$) and increases with $T$ (39.74% with $T = 10$ and

---

[4]　An example of the Stata code for the experiment of the power of the test in presence of heteroskedasticity due to both $\mu_i$ and $\nu_{it}$ with $N = 100$ and $T = 5$ is provided in the Appendix C. The Appendix B reports the Stata code used to generate the data.

67.14% with $T = 20$). Nonetheless, increasing $\sigma_\mu^2$ from 0.2 a larger value, say two or six results in a drastic increase in the power of the test (see Table A3 in Appendix D).

As for the joint test $H_0 : \theta_\mu = \theta_\nu = 0$, the Monte Carlo experiment reveals that the test has good performance in the case of high heteroskedasticity even when $N$ is small. For $N = 50$, the power of the test is 65.16% with $T = 5$ and reaches 100% with $T = 20$. In the case of low degree of heteroskedasticity, the power of the test is excellent when $N$ or $T$ is large. For small $N$, the power of the test is low. It is 14.62% when $N = 50$ and $T = 5$, and when $T$ is fixed at 5, increasing $N$ to 500 yields in a drastic increase in the power of the test (96.74%). Nonetheless, fixing $N$ at 50 and increasing $T$ to 10 and 20 yields in a power of the test of 35.16% and 63.74% respectively. Furthermore, increasing $\sigma_\mu^2$ from 0.2 a larger value, say two or six results in an increase in the power of the test (see Table A1 in Appendix D).

**Table 1.** Power of the likelihood ratio (LR) test for homoskedasticity based on 5000 replications.

| Settings | | $H_0 : \theta_\mu = 0 \mid \theta_\nu = 0$ | $H_0 : \theta_\nu = 0 \mid \theta_\mu = 0$ | $H_0 : \theta_\mu = \theta_\nu = 0$ |
|---|---|---|---|---|
| **Dimensions** | **Obs.** | **%** | **%** | **%** |
| Low degree of heteroskedasticity: $\sigma_\mu^2 = 0.2$, $\theta_\mu = 0.7$ and $\theta_\nu = 0.6$ | | | | |
| $(N,T) = (50,5)$ | 250 | 12.04 | 23.86 | 14.62 |
| $(N,T) = (100,5)$ | 500 | 19.88 | 43.32 | 31.68 |
| $(N,T) = (500,5)$ | 2500 | 69.54 | 97.22 | 96.74 |
| $(N,T) = (50,10)$ | 500 | 19.9 | 39.74 | 35.16 |
| $(N,T) = (100,10)$ | 1000 | 34.04 | 65.62 | 64.72 |
| $(N,T) = (500,10)$ | 5000 | 94.14 | 99.98 | 100 |
| $(N,T) = (50,20)$ | 1000 | 27.4 | 67.14 | 63.74 |
| $(N,T) = (100,20)$ | 2000 | 50.58 | 93.2 | 92.94 |
| $(N,T) = (500,20)$ | 10,000 | 99.26 | 100 | 100 |
| High degree of heteroskedasticity: $\sigma_\mu^2 = 0.2$, $\theta_\mu = 2.1$ and $\theta_\nu = 1.8$ | | | | |
| $(N,T) = (50,5)$ | 250 | 81.72 | 47.78 | 65.16 |
| $(N,T) = (100,5)$ | 500 | 98.44 | 85 | 96.2 |
| $(N,T) = (500,5)$ | 2500 | 100 | 100 | 100 |
| $(N,T) = (50,10)$ | 500 | 94.72 | 93.7 | 98.12 |
| $(N,T) = (100,10)$ | 1000 | 98.88 | 99.9 | 99.96 |
| $(N,T) = (500,10)$ | 5000 | 100 | 100 | 100 |
| $(N,T) = (50,20)$ | 1000 | 98.36 | 99.98 | 100 |
| $(N,T) = (100,20)$ | 2000 | 100 | 100 | 100 |
| $(N,T) = (500,20)$ | 10,000 | 100 | 100 | 100 |

Table 2 presents the empirical size of the test based the Monte Carlo experiments described above. The empirical size of the test $H_0 : \theta_\mu = 0 \mid \theta_\nu = 0$ varies between 4.54% and 5.36%. The empirical size of the test $H_0 : \theta_\nu = 0 \mid \theta_\mu = 0$ varies between 4.48% and 5.54% and between 4.52% and 5.14% in the case of the joint test $H_0 : \theta_\mu = \theta_\nu = 0$. All these empirical sizes do not significantly differ from the nominal size[5] of the test.

---

[5] The empirical size estimated on 5000 replications is significantly different from the nominal size of 5% if it does not range between 4.4% and 5.6%. These thresholds are calculated as $0.05 \pm 1.96\sqrt{\frac{0.05*0.95}{5000}}$.

**Table 2.** Empirical size of the LR test for homoskedasticity based on 5000 replications.

| Settings | | $H_0: \theta_\mu = 0 \mid \theta_\nu = 0$ | $H_0: \theta_\nu = 0 \mid \theta_\mu = 0$ | $H_0: \theta_\mu = \theta_\nu = 0$ |
|---|---|---|---|---|
| **Dimensions** | **Obs.** | **%** | **%** | **%** |
| $(N, T) = (50, 5)$ | 250 | 4.82 | 4.98 | 4.54 |
| $(N, T) = (100, 5)$ | 500 | 4.7 | 5.02 | 4.52 |
| $(N, T) = (500, 5)$ | 2500 | 4.64 | 5.34 | 4.68 |
| $(N, T) = (50, 10)$ | 500 | 5.36 | 5.46 | 4.72 |
| $(N, T) = (100, 10)$ | 1000 | 4.54 | 4.74 | 5.14 |
| $(N, T) = (500, 10)$ | 5000 | 4.62 | 4.48 | 4.68 |
| $(N, T) = (50, 20)$ | 1000 | 5.08 | 5.00 | 4.94 |
| $(N, T) = (100, 20)$ | 2000 | 4.92 | 4.48 | 5.04 |
| $(N, T) = (500, 20)$ | 10,000 | 4.58 | 5.54 | 5.04 |

*4.2. Bias and Mean Square Error of the Estimates*

This subsection aims at evaluating the robustness of the proposed estimation procedure. For this purpose, the modelling approach for the case where both $\mu_i$ and $\nu_{it}$ are heteroskedastic is used. The parameters of the model are those set in Section 4. The bias and the mean square error (MSE) of the estimates are computed based on 5000 replications for $N = \{50, 500\}$ and $T = \{5, 20\}$. The results are presented in Table 3.

The results suggest that the MSEs of both index function and variance parameters decrease with the number of observations. The bias of the index function parameters are lower than 5% regardless of the individual and time dimensions of the panel. The bias for the parameters of the variance of $\mu_i$ and $\nu_{it}$ becomes lower as the time dimension of the panel increases. It reaches 5% for $(N, T) = (500, 20)$.

**Table 3.** Bias and mean square error (MSE) of the estimates based on 5000 replications.

| Settings | | $(N, T) = (50, 5)$ | | $(N, T) = (50, 20)$ | | $(N, T) = (500, 5)$ | | $(N, T) = (500, 20)$ | |
|---|---|---|---|---|---|---|---|---|---|
| *Parameter* | *DGP* | **Bias** | **MSE** | **Bias** | **MSE** | **Bias** | **MSE** | **Bias** | **MSE** |
| Parameters of the index function | | | | | | | | | |
| $\alpha_0$ | 1.5 | 0.0009 | 0.2435 | 0.0814 | 0.0554 | 0.0402 | 0.0225 | 0.0606 | 0.0126 |
| $\alpha_1$ | 0.8 | 0.0040 | 0.2474 | 0.0331 | 0.0474 | 0.0237 | 0.0221 | 0.0400 | 0.0062 |
| $\alpha_2$ | $-2$ | 0.0072 | 0.4323 | 0.0834 | 0.0814 | 0.0530 | 0.0388 | 0.0909 | 0.0172 |
| Parameters of the variances of $\mu_i$ and $\nu_{it}$ | | | | | | | | | |
| $\lambda_0$ | $-0.8$ | 0.1683 | 1.5406 | 0.0609 | 0.2058 | 0.0517 | 0.0846 | 0.0407 | 0.0177 |
| $\theta_\mu$ | 0.7 | 0.1660 | 1.1061 | 0.0232 | 0.4044 | 0.0412 | 0.1204 | 0.0353 | 0.0316 |
| $\theta_\nu$ | 0.6 | 0.0721 | 0.2369 | 0.0618 | 0.0447 | 0.0456 | 0.0225 | 0.0301 | 0.0119 |

*4.3. Robustness of Validity*

In this subsection, the robustness of validity of the test procedure is assessed using the framework described by Montes-Rojas and Sosa-Escudero (2011). The aim is to assess how the departure away from normality of the data generating process (DGP) of the error components might affect the results of the test. For this purpose, the empirical size of the tests ($H_0: \theta_\mu = 0 \mid \theta_\nu = 0$, $H_0: \theta_\nu = 0 \mid \theta_\mu = 0$, and $H_0: \theta_\mu = \theta_\nu = 0$) is computed for $N = 50$ and $T = 5$ using 5000 replications. The empirical sizes for normal, student with three degrees of freedom, exponential, uniform, and chi-square DGP are estimated respectively. The results are presented in Table 4.

The results suggest that a deviation from the normal DGP has heavy consequences on the empirical size on the test. The higher effect on the empirical size of the test is observed for exponential DGP. These results were expected since the estimation procedure, i.e., the Gauss–Hermite quadrature, is accurate only when the integral function has a Gaussian factor.

| *DGP* | $H_0: \theta_\mu = 0 \mid \theta_\nu = 0$ | $H_0: \theta_\nu = 0 \mid \theta_\mu = 0$ | $H_0: \theta_\mu = \theta_\nu = 0$ |
|---|---|---|---|
| Normal | 4.82 | 4.98 | 5.54 |
| Student (3) | 6.38 | 7.86 | 8.32 |
| Exponential | 17.56 | 7.18 | 21.36 |
| Uniform | 5.74 | 7.68 | 8.7 |
| Chi-square | 3.24 | 5.8 | 4.5 |

## 5. Additional Robustness Checks

To further check for the robustness of the proposed approach, three analysis are conducted. Based on data simulated with parameters setted in Section 4, the first analysis consists in checking whether the estimation procedure provides estimates that are consistent with the data generating process. This analysis complements the measure of bias and MSE done in Section 4.2 by focusing of the difference between each estimated parameter and the DGP. The second robustness analysis consists in checking the effect of the number of quadrature points on the estimated parameters. The third analysis focuses on the robustness to misspecified heteroskedasticity, i.e., how the test procedure performs when a researcher applies the wrong test.

### 5.1. Application Examples and Comparisons

For each of the three cases of heteroskedasticity described in Section 2, two applications are provided: (i) the first on a random sample of size $N = 500$ and $T = 5$, (ii) and the second on a random sample of size $N = 500$ and $T = 20$. For each of these applications, comparisons with the homoskedastic panel probit and the heteroskedastic pooled probit models are provided. The log-likelihood and the LR statistics are provided for the heteroskedastic pooled probit and the heteroskedastic panel probit models. Estimates are in the Appendix E for models with heteroskedasticity due to $\mu_i$ are provided in Table A4, those of models with heteroskedasticity due to $\nu_{it}$ are provided in Table A5, and Table A6 provides the estimates of models with heteroskedasticity due to both $\mu_i$ and $\nu_{it}$.

Results suggest that in the presence of heteroskedasticity due to $\mu_i$, the pooled heteroskedastic model underestimates the heteroskedastic factor (coefficient of variable $Z_{\mu_i}$) for both $T = 5$ and $T = 20$ models. The homoskedastic part of the variance is well estimated using the homoskedastic panel probit model. It also appears that the parameters estimated from the homoskedastic and the heteroskedastic panel probit models are not different. However, as expected, the pooled model yields to bias in the estimated parameters especially when $T$ is large.

In the presence of heteroskedasticity due to $\nu_{it}$, the pooled heteroskedastic model gives correct estimates of the heteroskedastic factor (coefficient of variable $Z_{\nu_{it}}$) and the parameters of the model with $T = 20$. However, with $T = 5$, the estimated parameters are different from that of the data generating process (DGP). These estimates are not different from those provided by the pooled heteroskedastic model. The homoskedastic panel probit model yields estimates of parameters and variance components that differ from the DGP.

The estimation of the model with heteroskedasticity due to both $\mu_i$ and $\nu_{it}$ by a homoskedastic panel probit model leads to parameters that are different from the DGP. The heteroskedatic pooled probit model yields to estimated individual effects variance that is different from the DGP.

### 5.2. Quadrature Points Check

The data generated for the examples in Section 5.1 with $N = 500$ and $T = 20$ are used for the quadrature points check. This quadrature check is conducted on the heteroskedastic model where the heteroskedasticity is due to both $\mu_i$ and $\nu_{it}$ which is the more general case of heteroskedasticity.

The quadrature points check shows that using $Q$ under 10, in this example, leads to significant differences in the estimated parameters from the DGP (See details in Table A7 in the Appendix F).

For $Q = 10$ or more quadrature points, the estimated parameters are not significantly different from the DGP. Furthermore, as the number of quadrature points increases, the estimated parameters converge to the DGP's values and become more accurate. This result has also been found in several applications that use Gauss–Hermite quadrature (Baltagi 2008; Moussa and Delattre 2018). Moreover, for $Q = 10$, the relative difference in log-likelihood is around 0.001 while the relative difference in the LR statistics is around 0.1. In terms of computation time, the convergence is generally reached quickly. It takes from 41 seconds for $Q = 6$ to 133 seconds for $Q = 20$ for the model to converge. However, the computation time may vary considerably according to the number of explanatory variables and to the sample size.

*5.3. Misspecified Heteroskedasticity: Effects of Applying the Wrong Approach*

To check for robustness, this subsection analyzes what happens when researchers apply the wrong heteroskedasticity modelling approach to a model with heteroskedasticity. For example, what happens if in the presence of heteroskedasticity due to both $\mu_i$ and $\nu_{it}$, researchers apply the procedure for estimation of heteroskedasticity due to $\mu_i$? A second set of robustness check consists in applying one of the three tests to a homoskedastic model. For this purpose, the data generated in Section 5.1 for the examples with $N = 500$ and $T = 20$ are used. Thus, for each case, the number of quadrature points is set to $Q = 10$. For example, on a dataset generated with $\mu_i$ heteroskedastic, the heteroskedasticity due to $\nu_{it}$ and the heteroskedasticity due to both $\mu_i$ and $\nu_{it}$ modelling approaches are applied.

Table 5 shows the results of LR tests and the variance components for each of the aforementioned cases. Results suggest that when only $\mu_i$ are heteroskedastic, the application of the heteroskedasticity due to $\nu_{it}$ modelling approach results in incorrect estimates for the variance components and the LR test concludes to the presence of heteroskedasticity due to $\nu_{it}$. The results of Monte Carlo simulations presented in Table A3 in Appendix D show that the acceptance rate of the null hypothesis in such a situation varies between 6.6% and 96.4% according to the panel's dimension and the degree of heteroskedasticity. Furthermore, the higher the variance of $\mu_i$, the higher the acceptance rate. Contrary to the latter and as expected, if researchers apply the heteroskedasticity due to both $\mu_i$ and $\nu_{it}$ modelling approach, the results are consistent with that obtained when applying the right approach and the parameter $\theta_\nu$ that indicates the presence of heteroskedasticity due to $\nu_{it}$ is not significantly different from zero. The same results hold for the case where only $\nu_{it}$ are heteroskedastic except that the LR test concludes to no presence of heteroskedasticity due to $\mu_i$ when this modelling approach is used. The results from Monte Carlo simulations presented in Table A1 in Appendix D show that the acceptance rate of the null hypothesis does not significantly differ from 5%. In the case where both $\mu_i$ and $\nu_{it}$ are heteroskedastic, applying the wrong modelling approach yields in identification of the related heteroskedasticity while the others forms are ignored. For example, if researchers apply the heteroskedasticity due to $\nu_{it}$ modelling approach, the LR test concludes to the existence of heteroskedasticity due to $\nu_{it}$ while the heteroskedasticity from the individual effects is ignored. The Monte Carlo simulations conducted (see Tables A2 and A3 in Appendix D) show that the power of the test is close to that of the right test. This result suggest that it is better starting by the heteroskedasticity due to both $\mu_i$ and $\nu_{it}$ modelling approach. Then, if one of the sources has no significant contribution to heteroskedasticity, then researchers can turn to the other source with the use of the specific modelling approach.

**Table 5.** Estimated variance components and LR tests on wrong models.

| Case | $\mu_i$ Heteroskedastic | | $v_{it}$ Heteroskedastic | | $\mu_i$ and $v_{it}$ Heteroskedastic | |
|---|---|---|---|---|---|---|
| **Model** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| $LogL$ | −4499.6355 | −4492.5752 | −5231.2132 | −5189.5929 | −5235.0043 | −5210.8403 |
| $LR\,stat$ | 7.0066 *** | 21.1272 *** | 0.5819 | 83.8224 *** | 23.1831 *** | 71.5111 *** |
| | | | The variance parameters. | | | |
| $Z_{\mu_i}$ | | 0.6952 *** [0.3888;1.0016] | 0.1588 [−0.2497;0.5673] | 0.1392 [−0.2652;0.5436] | 0.8079 *** [0.4655;1.1504] | |
| $\lambda_0$ | | −0.8047 *** [−0.9916;−0.6178] | −1.1308 *** [−1.3659;−0.8957] | −0.8886 *** [−1.1252;−0.6519] | −1.1642 *** [−1.3771;−0.9513] | |
| $\sigma_\mu$ | 0.5985 *** [0.5369;0.6672] | | | | | 0.6012 *** [0.5419;0.6669] |
| $Z_{v_{it}}$ | −0.1974 *** [−0.3455;−0.0493] | −0.0442 [−0.1627;0.0743] | | 0.6025 *** [0.4726;0.7324] | | 0.5479 *** [0.4223;0.6734] |

95% level confident interval in brackets; ***: Significant at the 1%. In columns (1) and (2), the dataset has been generated with $\mu_i$ heteroskedastic. Then, the modelling and test approaches for heteroskedasticity due to $v_{it}$ (column 1) and to both $\mu_i$ and $v_{it}$ (column 2) are applied. For columns (4) and (5) the dataset is generated with $v_i$ heteroskedastic and the modelling and test approaches for heteroskedasticity due to $\mu_i$ (column 3) and to both $\mu_i$ and $v_{it}$ (column 4) are applied. In columns (5) and (6), the dataset is generated with both $\mu_i$ and $v_{it}$ heteroskedastic and the modelling and test approaches for heteroskedasticity due to $\mu_i$ (column 5) and to $v_{it}$ (column 6) are applied.

Table 6 shows the results of the application of one of the three tests to a situation where there is no heteroskedasticity. As expected, the LR tests conclude to homoskedasticity for all of the three modelling approaches. The Monte Carlo simulations conducted show that the acceptance rate of the null hypothesis does not differ significantly from the nominal size of the test (see Tables A1–A3 in Appendix D).

**Table 6.** Estimated variance components and LR tests on homoskedastic model.

| **Model** | **(1)** | **(2)** | **(3)** |
|---|---|---|---|
| $LogL$ | −4536.406 | −4535.2483 | −4535.1225 |
| $LR\,stat$ | 0.2644 | 2.5797 | 2.8313 |
| | The variance parameters. | | |
| $Z_{\mu_i}$ | 0.0856 [−0.2406;0.4118] | | 0.0835 [−0.2428;0.4098] |
| $\lambda_0$ | −0.7031 *** [−0.886;−0.5203] | | −0.7452 *** [−0.9357;−0.5548] |
| $\sigma_\mu$ | | 0.4938 *** [0.4421;0.5514] | |
| $Z_{v_{it}}$ | | −0.0987 [−0.2197;0.0222] | −0.0985 [−0.2194;0.0224] |

95% level confident interval in brackets; ***: Significant at the 1% level; The data used for the results in this Table are generated with no heteroskedasticity. Then, the modelling and test approaches for heteroskedasticity due to $\mu_i$ (column 1), to $v_{it}$ (column 2) and to both $\mu_i$ and $v_{it}$ (column 3) are applied.

## 6. Case Study

In this section, the illustration dataset for panel probit models used by Greene (2012) and refereed as Example 17.11 pp. 274–275 is used. This dataset is related to German health care utilization and contains 26,326 observations with $N = 7293$ and $T$ varying between 1 and 7. The model estimates the effects of socioeconomic variables (age, income, kids, education, and marital status) on the probability to visit a doctor. The results by Greene (2012) are replicated and the marginal effects are calculated with respect to the two approaches aforementioned. Then, the heteroskedastic probit model in the more general setting where both $\mu_i$ and $v_{it}$ are heteroskedastic is estimated and the marginal effects are computed. Table 7 shows the results of the estimates for the two approaches.

**Table 7.** Estimated coefficients and marginal effects.

| Variables | Homoskedastic Model | | | Heteroskedastic Model | | |
|---|---|---|---|---|---|---|
| | *Coef.* | *M.E.*$^{+}$ | *M.E.*$^{++}$ | *Coef.* | *M.E.*$^{+}$ | *M.E.*$^{++}$ |
| age | 0.0201 *** [0.0175;0.0228] | 0.0055 *** [0.0048;0.0062] | 0.0069 *** [0.0061;0.0078] | 0.0027 *** [0.0023;0.0032] | 0.0061 *** [0.0055;0.0066] | 0.0076 *** [0.007;0.0082] |
| income | −0.0032 [−0.1341;0.1278] | −0.0009 [−0.0366;0.0349] | −0.0011 [−0.0463;0.0441] | −0.001 [−0.0256;0.0237] | −0.0212 [−0.0581;0.0157] | −0.0316 [−0.0747;0.0115] |
| kids | −0.1538 *** [−0.2079;−0.0996] | −0.0420 *** [−0.0567;−0.0272] | −0.053 *** [−0.0717;−0.0344] | −0.0336 *** [−0.0433;−0.0238] | −0.0497 *** [−0.064;−0.0353] | −0.0549 *** [−0.0707;−0.039] |
| education | −0.0337 *** [−0.0462;−0.0212] | −0.0092 *** [−0.0126;−0.0058] | −0.0116 *** [−0.0159;−0.0073] | −0.0065 *** [−0.0083;−0.0046] | −0.0038 *** [−0.0066;−0.001] | −0.0018 [−0.0051;0.0014] |
| married | 0.0163 [−0.0477;0.0803] | 0.0045 [−0.013;0.0219] | 0.0056 [−0.0164;0.0277] | 0.0005 [−0.0101;0.011] | 0.0007 [−0.0149;0.0163] | 0.0008 [−0.0164;0.018] |
| intercept | 0.0341 [−0.1591;0.2273] | | | 0.0558 *** [0.0251;0.0864] | | |
| The variance parameters: variance of $\mu_i$ | | | | | | |
| female | | | | −0.0766 *** [−0.1101;−0.0431] | | |
| $\lambda_0$ | | | | −2.1074 *** [−2.1311;−2.0837] | | |
| $\sigma_\mu$ | 0.9007 *** [0.8649;0.9379] | | | | | |
| The variance parameters: variance of $\nu_{it}$ | | | | | | |
| age | | | | −0.0215 *** [−0.0232;−0.0198] | | |
| income | | | | 0.2098 ** [0.028;0.3916] | | |
| education | | | | −0.061 *** [−0.0691;−0.0529] | | |
| LogL | −16,273.964 | | | −14,019.325 | | |
| LR stat | | | | 4509.45 *** | | |

95% level confident interval in brackets; ***: Significant at the 1% level; **: Significant at the 5% level; +: marginal effects by integrating with respect to $\mu_i$; ++: marginal effects assuming $\mu_i = 0$. The coefficients of the homoskedastic model are those reported by Greene (2012).

The LR test of homoskedasticity leads to the rejection of the null hypothesis of homoskedasticity. Thus, there is the presence of heteroskedasticity due to both $\mu_i$ and $\nu_{it}$. The estimated parameters are significantly different between the homoskedastic and the heteroskedastic models. The same result holds for the marginal effects. However, the differences in marginal effects are lower using the marginal effects integrated with respect to $\mu_i$ that the ones computed assuming $\mu_i = 0$. Using the former approach, results suggest that ageing increases by 0.55% the probability to visit a doctor using the homoskedastic model and by 0.61% using the heteroskedastic model. These estimates are respectively 0.69% and 0.76% using the second approach. An increase in the number education years reduces by 0.92% the probability of visiting a doctor using the homoskedastic model and by 0.38% using the heteroskedastic model. Assuming $\mu_i = 0$, an increase in the number of education years reduce by 1.16% the probability of visiting a doctor using the homoskedastic model while the effect is not significant using the heteroskedastic model.

## 7. Conclusions

The use of a random effects probit model has been popularized due to the problem of incidental parameters encountered when dealing with fixed effects models for binary outcomes in panel data. However, researchers do not test for the presence of heteroskedasticity in the error terms and then do not control for that when estimating these models. This paper proposes an estimation procedure that accounts for heteroskedasticity for both individual effects and idiosyncratic errors separately and jointly as well as a LR test for homoskedasticity.

A Monte Carlo experiment was conducted to estimate the power of the test. It shows that the LR test performs well generally. However, on samples with a low degree of heteroskedasticity, the power of the test is around 20% for panels with small *N* and *T* but it increases drastically with larger *N* and *T*. The analysis also show that applying the wrong estimation and test procedures may yield misleading conclusions about heteroskedasticity.

**Conflicts of Interest:** The author declare no conflict of interest.

## Appendix A. STATA Code for Computing the Marginal Effects

```
//model with mu_i heteroskedastics
xthetprobit y_muhet   x1 x2, quad(10) hetmu(xmu) hettype(mu)
//marginal effect for variable x1 (integrated with respect to mu_i)
margins, dydx(x1) expr(normal(xb(#1)/sqrt(1+exp(xb(#2))^2)))
//marginal effect for variable x1 (assuming mu_i=0)
margins, dydx(x1) expr(normal(xb(#1)/exp(xb(#2))))

//model with nu_it heteroskedastics
xthetprobit y_muhet   x1 x2, quad(10) hetnu(xnu) hettype(nu)
//marginal effect for variable x1 (integrated with respect to mu_i)
margins, dydx(x1) expr(normal(xb(#1)/sqrt(exp(xb(#2))^2+exp(xb(#3))^2)))
//marginal effect for variable x1 (assuming mu_i=0)
margins, dydx(x1) expr(normal(xb(#1)/exp(xb(#2))))

//model with both mu_i and nu_it heteroskedastics
xthetprobit y_bothhet x1 x2, quad(10) hetmu(xmu) hetnu(xnu) hettype(both)
//marginal effect for variable x1 (integrated with respect to mu_i)
margins, dydx(x1) expr(normal(xb(#1)/sqrt(exp(xb(#2))^2+exp(xb(#3))^2)))
```

```
//marginal effect for variable x1 (assuming mu_i=0)
margins, dydx(x1) expr(normal(xb(#1)/exp(xb(#3))))
```

**Appendix B. STATA Code for Generating the Dataset**

```
clear
// T=5 is set here for first individual
set obs 5
gen id=1
lab var id "panel id"
forval n = 2/500 { // N=500 is set here
        local k = 5*'n'  // T=5 for others individuals
        qui set obs 'k'
        qui replace id = 'n' if id==.
}
// Time variable is generated here
by id, sort : gen t=_n
lab var t "time id"
// Setting seed to avoid variation for random variables
set seed 1
gen x1 = runiform()
gen x2 = runiform()
// Generate hetmu that is time−invariant
gen zmu = runiform() if t==5
by id, sort: replace zmu = zmu[_N] if _n!=_N
gen znu = runiform()
// Generate the variances
gen sig_mu  = exp(−0.8 + 0.7*zmu)
gen sig_mu1 = exp(−0.8)
gen sig_nu  = exp(0.6*znu)
// Generate mu_i that is time−invariant
gen mu_i = rnormal(0, 1) if t==5
by id, sort: replace mu_i = mu_i[_N] if _n!=_N
// Generating y
gen p_mu   = normal((1.5+0.8*x1−2*x2+mu_i*sig_mu))
gen p_nu   = normal((1.5+0.8*x1−2*x2+mu_i*sig_mu1)/sig_nu)
gen p_both = normal((1.5+0.8*x1−2*x2+mu_i*sig_mu)/sig_nu)
gen y_mu   = cond(runiform()<=p_muhet,1,0)
gen y_nu   = cond(runiform()<=p_nuhet,1,0)
gen y_both = cond(runiform()<=p_bothhet,1,0)
// Setting panel
xtset id t
//Heteroskedastic model: Mazodier and Trognon (1978)
xthetprobit y_mu x1 x2, quad(10) hetmu(zmu) hettype(mu)
xtprobit y_mu x1 x2, intp(10) nolog // Comparison models
hetprobit y_mu x1 x2, het(zmu)
//Heteroskedastic model: Baltagi (1988) and Wansbeek (1989)
xthetprobit y_nu x1 x2, quad(10) hetnu(znu) hettype(nu)
xtprobit y_nu x1 x2, intp(10) nolog // Comparison models
hetprobit y_nu x1 x2, het(znu)
```

```
//Heteroskedastic model: Randolph (1988)
xthetprobit y_both x1 x2, quad(10) hetmu(zmu) hetnu(znu) hettype(both)
xtprobit y_both x1 x2, intp(10) nolog // Comparison models
hetprobit y_both x1 x2, het(zmu znu)
```

## Appendix C. STATA Code for Monte Carlo Experiments

```
clear
set matsize 10000
mat LRS = J(5001,1,0)
set seed 1
set obs 5
gen id=1
lab var id "panel id"
forval n = 2/100 {
        local k = 5*'n'
        set obs 'k'
        replace id = 'n' if id==.
}
by id, sort : gen t=_n
lab var t "time id"
forval kk=1/5000 {
qui {
keep id t
gen x1 = runiform()
gen x2 = runiform()
gen zmu = runiform() if t==5
by id, sort: replace xmu = xmu[_N] if _n!=_N
gen znu = runiform()
gen sig_mu = exp(-0.8 + 0.7*zmu)
gen sig_nu = exp(0.6*znu)
gen mu_i = rnormal(0, 1) if t==5
by id, sort: replace mu_i = mu_i[_N] if _n!=_N
gen p1 = normal((1.5+0.8*x1-2*x2+mu_i*sig_mu)/sig_nu)
gen y1 = cond(runiform()<=p1,1,0)
xtset id t
xthetprobit y1 x1 x2, quad(10) hetmu(zmu) hetnu(znu) hettype(both)
mat LRS['kk',1]=e(lrstat)
        if chiprob(2,e(lrstat)) < 0.05 {
                mat LRS[5001,1] = LRS[5001,1]+1
        }
}
}
```

## Appendix D. Power of the Test for Different Degrees of Heteroskedasticity

*Appendix D.1. Testing for the Joint Hypothesis*

**Table A1.** Power and size of the LR test based on 5000 replications: case of $H_0 : \theta_\mu = \theta_\nu = 0$.

| Setting | | $\sigma_\mu^2 = 2$ | | | | $\sigma_\mu^2 = 6$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $N = 50$ | | $N = 500$ | | $N = 50$ | | $N = 500$ | |
| $\theta_\mu$ | $\theta_\nu$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ |
| 0 | 0 | 4.48 | 4.88 | 4.44 | 5.42 | 5.58 | 5.52 | 5.6 | 5.6 |
| 0 | 1 | 12.68 | 91.24 | 99.48 | 100 | 6.6 | 92.24 | 98.42 | 100 |
| 0 | 2 | 35.88 | 99.92 | 100 | 100 | 31.84 | 100 | 100 | 100 |
| 0 | 3 | 50.22 | 99.94 | 100 | 100 | 38.42 | 100 | 100 | 100 |
| 1 | 0 | 22.08 | 39.08 | 99.48 | 100 | 13.36 | 29.3 | 98.46 | 100 |
| 1 | 1 | 26.56 | 97.56 | 100 | 100 | 14.46 | 90.04 | 99.6 | 100 |
| 1 | 2 | 46.36 | 100 | 100 | 100 | 49.26 | 100 | 100 | 100 |
| 1 | 3 | 45.78 | 100 | 100 | 100 | 62.56 | 100 | 100 | 100 |
| 2 | 0 | 30.79 | 72.34 | 100 | 100 | 21.08 | 43.12 | 100 | 100 |
| 2 | 1 | 54.14 | 99.26 | 100 | 100 | 23.08 | 75.34 | 100 | 100 |
| 2 | 2 | 78.28 | 100 | 100 | 100 | 65.62 | 100 | 100 | 100 |
| 2 | 3 | 79.28 | 100 | 100 | 100 | 87.06 | 100 | 100 | 100 |
| 3 | 0 | 50.92 | 73.8 | 100 | 100 | 30.08 | 58.46 | 100 | 100 |
| 3 | 1 | 59.52 | 96.88 | 100 | 100 | 37.72 | 56.46 | 100 | 100 |
| 3 | 2 | 90.3 | 100 | 100 | 100 | 57.88 | 99.92 | 100 | 100 |
| 3 | 3 | 95.46 | 100 | 100 | 100 | 93.8 | 100 | 100 | 100 |

*Appendix D.2. Testing for the Marginal Hypothesis of No Heteroskedasticity in Individual Effects Given Homoskedastic Idiosyncratic Errors*

**Table A2.** Power and size of the LR test based on 5000 replications: case of $H_0 : \theta_\mu = 0 \,|\, \theta_\nu = 0$.

| | | $\sigma_\mu^2 = 2$ | | | | $\sigma_\mu^2 = 6$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| Setting | | $N = 50$ | | $N = 500$ | | $N = 50$ | | $N = 500$ | |
| | | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ |
| $\theta_\mu$ | | | | | | | | | |
| 0 | | 5.26 | 5.46 | 5.06 | 5.02 | 4.42 | 5.08 | 4.76 | 4.72 |
| 1 | | 24.86 | 47.12 | 99.54 | 100 | 5.88 | 15.76 | 77.84 | 98.68 |
| 2 | | 35.16 | 71.7 | 100 | 100 | 6.22 | 17.68 | 90.52 | 100 |
| 3 | | 29.14 | 64.98 | 100 | 100 | 4.94 | 15.28 | 99.32 | 100 |
| $\theta_\nu$ | | | | | | | | | |
| 0 | | 5.26 | 5.46 | 5.06 | 5.02 | 4.42 | 5.08 | 4.76 | 4.72 |
| 1 | | 5.78 | 5.46 | 5.56 | 4.74 | 5.06 | 4.82 | 4.72 | 4.4 |
| 2 | | 5.26 | 5.06 | 5.58 | 4.46 | 5.56 | 4.62 | 5.02 | 4.42 |
| 3 | | 5.12 | 4.94 | 4.96 | 4.44 | 5.44 | 4.74 | 4.92 | 4.42 |
| $\theta_\mu$ | $\theta_\nu$ | | | | | | | | |
| 0 | 0 | 5.26 | 5.46 | 5.06 | 5.02 | 4.42 | 5.08 | 4.76 | 4.72 |
| 0 | 1 | 5.78 | 5.46 | 5.56 | 4.74 | 5.06 | 4.82 | 4.72 | 4.4 |
| 0 | 2 | 5.26 | 5.06 | 5.58 | 4.46 | 5.56 | 4.62 | 5.02 | 4.42 |
| 0 | 3 | 5.12 | 4.94 | 4.96 | 4.44 | 5.44 | 4.74 | 4.92 | 4.42 |

**Table A2.** *Cont.*

| Setting | | $\sigma_\mu^2 = 2$ | | | | $\sigma_\mu^2 = 6$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $N = 50$ | | $N = 500$ | | $N = 50$ | | $N = 500$ | |
| | | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ |
| 1 | 0 | 24.86 | 47.12 | 99.54 | 100 | 5.88 | 15.76 | 77.84 | 98.68 |
| 1 | 1 | 27.2 | 53.2 | 99.6 | 100 | 15.66 | 34.26 | 96.3 | 100 |
| 1 | 2 | 23.88 | 45.78 | 97.2 | 100 | 22.56 | 42.84 | 98.64 | 100 |
| 1 | 3 | 14 | 33.56 | 79.58 | 100 | 19.08 | 31.72 | 92.44 | 100 |
| 2 | 0 | 35.16 | 71.7 | 100 | 100 | 6.22 | 17.68 | 90.52 | 100 |
| 2 | 1 | 59.46 | 91.98 | 100 | 100 | 24.76 | 59.52 | 100 | 100 |
| 2 | 2 | 67 | 94.98 | 100 | 100 | 55.48 | 89.68 | 100 | 100 |
| 2 | 3 | 52.72 | 89.16 | 100 | 100 | 55.9 | 88.36 | 100 | 100 |
| 3 | 0 | 29.14 | 64.98 | 100 | 100 | 4.94 | 15.28 | 99.32 | 100 |
| 3 | 1 | 66.46 | 95.96 | 100 | 100 | 21.56 | 59.06 | 100 | 100 |
| 3 | 2 | 88.42 | 99.8 | 100 | 100 | 65.1 | 97.06 | 100 | 100 |
| 3 | 3 | 86.9 | 99.66 | 100 | 100 | 83.36 | 99.58 | 100 | 100 |

*Appendix D.3. Testing for the Marginal Hypothesis of no Heteroskedasticity in Idiosyncratic Errors Given Homoskedastic Individual Effects*

**Table A3.** Power and size of the LR test based on 5000 replications: case of $H_0 : \theta_v = 0 \,|\, \theta_\mu = 0$.

| Setting | | $\sigma_\mu^2 = 2$ | | | | $\sigma_\mu^2 = 6$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $N = 50$ | | $N = 500$ | | $N = 50$ | | $N = 500$ | |
| | | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ | $T = 5$ | $T = 20$ |
| $\theta_v$ | | | | | | | | | |
| 0 | | 4.64 | 4.48 | 4.44 | 5.1 | 5.56 | 5.6 | 5.6 | 5.6 |
| 1 | | 20.5 | 95.6 | 99.8 | 100 | 11.12 | 96.62 | 99.5 | 100 |
| 2 | | 54.28 | 99.98 | 100 | 100 | 52.8 | 100 | 100 | 100 |
| 3 | | 66.18 | 99.96 | 100 | 100 | 62.46 | 100 | 100 | 100 |
| $\theta_\mu$ | | | | | | | | | |
| 0 | | 4.64 | 4.48 | 4.44 | 5.1 | 5.56 | 5.6 | 5.6 | 5.6 |
| 1 | | 6.6 | 8.2 | 39.06 | 43.18 | 23.96 | 23.98 | 87.02 | 96.4 |
| 2 | | 16.08 | 23.78 | 77.06 | 96.1 | 15.54 | 34.34 | 99.96 | 99.42 |
| 3 | | 25.28 | 29.19 | 97.98 | 98.72 | 24.58 | 49.7 | 99.96 | 100 |
| $\theta_\mu$ | $\theta_v$ | | | | | | | | |
| 0 | 0 | 4.64 | 4.48 | 4.44 | 5.1 | 5.56 | 5.6 | 5.6 | 5.6 |
| 0 | 1 | 20.5 | 95.6 | 99.8 | 100 | 11.12 | 96.62 | 99.5 | 100 |
| 0 | 2 | 54.28 | 99.98 | 100 | 100 | 52.8 | 100 | 100 | 100 |
| 0 | 3 | 66.18 | 99.96 | 100 | 100 | 62.46 | 100 | 100 | 100 |
| 1 | 0 | 6.6 | 8.2 | 39.06 | 43.18 | 23.96 | 23.98 | 87.02 | 96.4 |
| 1 | 1 | 11.52 | 96.46 | 99.54 | 100 | 34.2 | 87.7 | 84.38 | 100 |
| 1 | 2 | 50.92 | 100 | 100 | 100 | 53.16 | 100 | 100 | 100 |
| 1 | 3 | 60.96 | 100 | 100 | 100 | 76.9 | 100 | 100 | 100 |
| 2 | 0 | 16.08 | 23.78 | 77.06 | 96.1 | 15.54 | 34.34 | 99.96 | 99.42 |
| 2 | 1 | 24.22 | 86.38 | 91.18 | 100 | 25.24 | 50.28 | 100 | 100 |
| 2 | 2 | 45.94 | 100 | 100 | 100 | 34.06 | 99.98 | 100 | 100 |
| 2 | 3 | 70.08 | 100 | 100 | 100 | 78.88 | 100 | 100 | 100 |
| 3 | 0 | 25.28 | 29.19 | 97.98 | 98.72 | 24.58 | 49.7 | 99.96 | 100 |
| 3 | 1 | 33.71 | 56.64 | 98.04 | 99.92 | 32.36 | 59.6 | 100 | 100 |
| 3 | 2 | 48.88 | 99.98 | 100 | 100 | 51.34 | 99.32 | 100 | 100 |
| 3 | 3 | 67.52 | 100 | 100 | 100 | 65.98 | 100 | 100 | 100 |

## Appendix E. Application and Comparisons

*Appendix E.1. Application and Comparison for Data Generated with Individual Effects Heteroskedastic*

**Table A4.** Estimated index function and variance parameters.

| Variables | *DGP* | Homoskedastic Panel Probit | Heteroskedastic Pooled Probit | Heteroskedastic Panel Probit |
|---|---|---|---|---|
| With $(N, T) = (500, 5)$ | | | | |
| *LogL* | | $-1241.7666$ | $-1283.699$ | $-1238.6702$ |
| *LR stat* | | | 9.52 *** | 6.2328 ** |
| The estimated index function parameters. | | | | |
| $X_1$ | 0.8 | 0.8285 *** [0.6008;1.0561] | 0.87 *** [0.6053;1.1348] | 0.8289 *** [0.6021;1.0557] |
| $X_2$ | $-2$ | $-1.9558$ *** [$-2.2014;-1.7102$] | $-2.1149$ *** [$-2.5151;-1.7146$] | $-1.9628$ *** [$-2.2077;-1.7179$] |
| *intercept* | 1.5 | 1.3993 *** [1.2052;1.5922] | 1.4881 *** [1.206;1.7702] | 1.3992 *** [1.2077;1.5907] |
| The variance parameters. | | | | |
| $Z_{\mu_i}$ | 0.7 | | 0.4086 *** [0.1485;0.6686] | 0.7147 ** [0.1192;1.3103] |
| $\lambda_0$ | $-0.8$ | $-0.8713$ *** [$-1.2052;-0.5373$] | | $-0.8449$ *** [$-1.257;-0.4327$] |
| $\sigma_v$ *(assumed)* | 1 | 1 | 1 | 1 |
| With $(N, T) = (500, 20)$ | | | | |
| *LogL* | | $-4500.4005$ | $-4928.991$ | $-4492.8442$ |
| *LR stat* | | | 8.81 *** | 20.5892 *** |
| The estimated index function parameters. | | | | |
| $X_1$ | 0.8 | 0.7235 *** [0.6114;0.8356] | 0.6531 *** [0.5392;0.7671] | 0.719 *** [0.6073;0.8307] |
| $X_2$ | $-2$ | $-2.0158$ *** [$-2.1368;-1.8949$] | $-1.8368$ *** [$-1.9884;-1.6852$] | $-2.0038$ *** [$-2.1238;-1.8838$] |
| *intercept* | 1.5 | 1.5913 *** [1.4813;1.7012] | 1.4502 *** [1.3315;1.5689] | 1.5688 *** [1.4651;1.6725] |
| The variance parameters. | | | | |
| $Z_{\mu_i}$ | 0.7 | | 0.182 *** [0.0613;0.3028] | 0.6953 *** [0.39;1.0016] |
| $\lambda_0$ | $-0.8$ | $-0.8252$ *** [$-1.0042;-0.6462$] | | $-0.7853$ *** [$-0.9646;-0.6059$] |
| $\sigma_v$ *(assumed)* | 1 | 1 | 1 | 1 |

95% level confident interval in brackets; ***: Significant at the 1% level. **: Significant at the 5% level.

*Appendix E.2. Application and Comparison for Data Generated with Idiosyncratic Errors Heteroskedastic*

**Table A5.** Estimated index function and variance parameters.

| Variables | DGP | Homoskedastic Panel Probit | Heteroskedastic Pooled Probit | Heteroskedastic Panel Probit |
|---|---|---|---|---|
| With $(N, T) = (500, 5)$ | | | | |
| $LogL$ | | −1356.3042 | −1363.723 | −1352.2662 |
| $LR\,stat$ | | | 5.99 ** | 8.0787 *** |
| The estimated index function parameters. | | | | |
| $X_1$ | 0.8 | 0.7217 *** <br> [0.515;0.9284] | 0.752 *** <br> [0.511;0.9931] | 0.8878 *** <br> [0.6085;1.167] |
| $X_2$ | −2 | −1.5232 *** <br> [−1.7353;−1.3111] | −1.698 *** <br> [−2.0197;−1.3763] | −1.9064 *** <br> [−2.2846;−1.5283] |
| $intercept$ | 1.5 | 1.0644 *** <br> [0.9026;1.2262] | 1.2019 *** <br> [0.9675;1.4364] | 1.3267 *** <br> [1.0544;1.599] |
| The variance parameters. | | | | |
| $Z_{\nu_{it}}$ | 0.6 | | 0.3358 *** <br> [0.0649;0.6067] | 0.4247 *** <br> [0.1352;0.7143] |
| $\sigma_\mu$ | 0.45 | 0.3857 *** <br> [0.2935;0.507] | | 0.4805 *** <br> [0.3619;0.638] |
| With $(N, T) = (500, 20)$ | | | | |
| $LogL$ | | −5230.5929 | −5280.282 | −5189.8208 |
| $LR\,stat$ | | | 89.12 *** | 83.3667 *** |
| The estimated index function parameters. | | | | |
| $X_1$ | 0.8 | 0.5561 *** <br> [0.4553;0.6568] | 0.7271 *** <br> [0.5877;0.8665] | 0.7618 *** <br> [0.6191;0.9046] |
| $X_2$ | −2 | −1.5742 *** <br> [−1.6786;−1.4698] | −2.0745 *** <br> [−2.2619;−1.8871] | −2.1414 *** <br> [−2.3319;−1.951] |
| $intercept$ | 1.5 | 1.2223 *** <br> [1.1361;1.3086] | 1.5999 *** <br> [1.4578;1.7419] | 1.638 *** <br> [1.4917;1.7842] |
| The variance parameters. | | | | |
| $Z_{\nu_{it}}$ | 0.6 | | 0.6417 *** <br> [0.5062;0.7772] | 0.603 *** <br> [0.4731;0.7329] |
| $\sigma_\mu$ | 0.45 | 0.3549 *** <br> [0.3144;0.4007] | | 0.44 *** <br> [0.3883;0.4987] |

95% level confident interval in brackets; ***: Significant at the 1% level. **: Significant at the 5% level.

*Appendix E.3. Application and Comparison for Data Generated with Both Individual Effects and Idiosyncratic Heteroskedastic*

**Table A6.** Estimated index function and variance parameters.

| Variables | DGP | Homoskedastic Panel Probit | Heteroskedastic Pooled Probit | Heteroskedastic Panel Probit |
|---|---|---|---|---|
| With $(N, T) = (500, 5)$ | | | | |
| $LogL$ | | $-1378.8704$ | $-1407.272$ | $-1371.3114$ |
| $LR\,stat$ | | | 11.00 *** | 15.1319 *** |
| The estimated index function parameters. | | | | |
| $X_1$ | 0.8 | 0.7061 *** [0.4946;0.9175] | 0.8069 *** [0.5124;1.1015] | 0.8886 *** [0.6023;1.1749] |
| $X_2$ | $-2$ | $-1.4263$ *** [$-1.6449;-1.2077$] | $-1.6477$ *** [$-2.0588;-1.2365$] | $-1.7918$ *** [$-2.1511;-1.4324$] |
| $intercept$ | 1.5 | 1.0207 *** [0.8499;1.1915] | 1.1883 *** [0.8931;1.4836] | 1.2692 *** [1.002;1.5365] |
| The variance parameters. | | | | |
| $Z_{\mu_i}$ | 0.7 | | 0.1062 [$-0.1847;0.3971$] | 0.6385 ** [0.0271;1.2499] |
| $\lambda_0$ | $-0.8$ | $-1.1807$ *** [$-1.5411;-0.8204$] | | $-0.767$ *** [$-1.2244;-0.3096$] |
| $Z_{v_{it}}$ | 0.6 | | 0.4681 *** [0.1734;0.7628] | 0.448 *** [0.155;0.7411] |
| With $(N, T) = (500, 20)$ | | | | |
| $LogL$ | | $-5245.4829$ | $-5447.011$ | $-5200.2317$ |
| $LR\,stat$ | | | 63.51 *** | 92.7282 *** |
| The estimated index function parameters. | | | | |
| $X_1$ | 0.8 | 0.5571 *** [0.4553;0.659] | 0.7218 *** [0.5743;0.8692] | 0.7407 *** [0.6011;0.8804] |
| $X_2$ | $-2$ | $-1.5148$ *** [$-1.62;-1.4096$] | $-1.9041$ *** [$-2.117;-1.6912$] | $-1.9696$ *** [$-2.1449;-1.7943$] |
| $intercept$ | 1.5 | 1.1794 *** [1.0879;1.271] | 1.468 *** [1.3057;1.6302] | 1.5026 *** [1.3669;1.6383] |
| The variance parameters. | | | | |
| $Z_{\mu_i}$ | 0.7 | | 0.1148 * [$-0.0189;0.2485$] | 0.7565 *** [0.4228;1.0902] |
| $\lambda_0$ | $-0.8$ | $-1.4512$ *** [$-1.6468;-1.2556$] | | $-0.9136$ *** [$-1.1258;-0.7013$] |
| $Z_{v_{it}}$ | 0.6 | | 0.5498 *** [0.4099;0.6896] | 0.5255 *** [0.4108;0.6603] |

95% level confident interval in brackets; ***: Significant at the 1% level. **: Significant at the 5% level. *: Significant at the 10% level.

## Appendix F. Estimates for Different Numbers of Quadrature Points

**Table A7.** Changes in Parameters and in log-likelihood with respect to the number of quadrature point $Q$.

| Variables | DGP | Q = 6 | Q = 8 | Q = 10 | Q = 12 | Q = 14 | Q = 16 | Q = 18 | Q = 20 |
|---|---|---|---|---|---|---|---|---|---|
| $LogL$ | | −5233.2859 | −5210.0882 | −5200.2317 | −5195.4211 | −5192.3082 | −5190.1575 | −5189.2627 | −5188.4647 |
| $Wald\ stat$ | | 66.3381 | 81.5596 | 94.3733 | 103.304 | 109.889 | 115.235 | 116.835 | 117.604 |
| $LR\ stat$ | | 67.0431 | 80.8018 | 92.7282 | 100.563 | 106.433 | 110.666 | 112.443 | 114.037 |
| $X_1$ | 0.8 | 0.6809 *** [0.5508;0.811] | 0.7143 *** [0.579;0.8495] | 0.7407 *** [0.6011;0.8804] | 0.7598 *** [0.6167;0.9028] | 0.7751 *** [0.6292;0.9211] | 0.7889 *** [0.6403;0.9375] | 0.7962 *** [0.646;0.9464] | 0.8035 *** [0.6517;0.9553] |
| $X_2$ | −2 | −1.8256 *** [−1.9899;−1.6612] | −1.9058 *** [−2.0755;−1.7361] | −1.9696 *** [−2.1449;−1.7943] | −2.0171 *** [−2.1978;−1.8365] | −2.0568 *** [−2.2428;−1.8709] | −2.0906 *** [−2.2814;−1.8999] | −2.1087 *** [−2.3029;−1.9145] | −2.1265 *** [−2.3244;−1.9287] |
| Intercept | 1.5 | 1.3551 *** [1.2299;1.4803] | 1.4401 *** [1.3098;1.5703] | 1.5026 *** [1.3669;1.6383] | 1.5479 *** [1.4067;1.6892] | 1.5852 *** [1.4384;1.7319] | 1.6162 *** [1.4642;1.7681] | 1.6332 *** [1.4774;1.7889] | 1.6497 *** [1.4901;1.8094] |
| $Z_{\mu_i}$ | 0.7 | 0.8259 *** [0.4615;1.1903] | 0.7739 *** [0.4341;1.1137] | 0.7565 *** [0.4228;1.0902] | 0.7366 *** [0.4036;1.0696] | 0.7436 *** [0.4094;1.0778] | 0.7582 *** [0.4241;1.0923] | 0.7598 *** [0.425;1.0947] | 0.7697 *** [0.4347;1.1046] |
| $\lambda_0$ | −0.8 | −1.0917 *** [−1.3352;−0.8482] | −0.984 *** [−1.2023;−0.7656] | −0.9136 *** [−1.1258;−0.7013] | −0.8603 *** [−1.0733;−0.6474] | −0.831 *** [−1.0463;−0.6157] | −0.8104 *** [−1.0266;−0.5941] | −0.7967 *** [−1.015;−0.5784] | −0.7887 *** [−1.0077;−0.5697] |
| $Z_{v_{it}}$ | 0.6 | 0.4013 *** [0.2782;0.5244] | 0.4761 *** [0.3529;0.5993] | 0.5355 *** [0.4108;0.6603] | 0.5783 *** [0.4512;0.7053] | 0.6142 *** [0.4842;0.7442] | 0.6435 *** [0.5109;0.7761] | 0.6593 *** [0.5243;0.7943] | 0.6747 *** [0.5372;0.8122] |
| $\Delta\ in\ LogL$ | | | 0.0044 | 0.0019 | 0.0009 | 0.0006 | 0.0004 | 0.0002 | 0.0002 |
| $\Delta\ in\ Wald\ stat$ | | | 0.2259 | 0.155 | 0.0936 | 0.0632 | 0.0482 | 0.0138 | 0.0065 |
| $\Delta\ in\ LR\ stat$ | | | 0.2022 | 0.1458 | 0.0836 | 0.0578 | 0.0394 | 0.0159 | 0.0141 |
| $\Delta\ in\ param.$ | | 0.162 | 0.1022 | 0.0631 | 0.0335 | 0.0341 | 0.0465 | 0.0533 | 0.0599 |
| $Time\ (sec)$ | | 41 | 53 | 62 | 79 | 96 | 106 | 130 | 133 |

95% level confident interval in brackets; ***: Significant at the 1% level. $\Delta$ denotes the relative difference defined as $\frac{|x-y|}{1+|y|}$. It is calculated to assess the variation in the log-likelihood ($LogL$), $LR\ stat$ and parameters when the number of quadrature points $Q$ increases. $\Delta\ in\ parameters$ is calculated as the maximum relative difference between parameters for two different $Q$.

## References

Baltagi, Badi H. 1988. An alternative heteroscedastic error component model, problem 88.2.2. *Econometric Theory* 4: 349–50. [CrossRef]

Baltagi, Badi H. 2008. *Econometric Analysis of Panel Data*, 4th ed. Hoboken: John Wiley & Sons.

Baltagi, Badi H., Georges Bresson, and Alain Pirotte. 2006. Joint lm test for homoskedasticity in a one-way error component model. *Journal of Econometrics* 134: 401–17. [CrossRef]

Bland, James R., and Amanda C. Cook. 2018. Random effects probit and logit: understanding predictions and marginal effects. *Applied Economics Letters* 26: 116–23. [CrossRef]

Davidson, Russell, and James G. MacKinnon. 1984. Convenient specification tests for logit and probit models. *Journal of Econometrics* 25: 241–62. [CrossRef]

Gould, William, Jeffrey Pitblado, and William Sribney. 2010. *Maximum Likelihood Estimation With Stata*, 4th ed. College Station: Stata Press.

Greene, William H. 2012. *Econometric Analysis*, 7th ed. Upper Saddle Rive: Prentice Hall.

Greene, William H. 2018. *Econometric Analysis*, 8th ed. New York: Pearson.

Johnston, John, and John DiNardo. 2001. *Econometric Methods*, 4th ed. New York: The McGraw-Hill Companies.

Lancaster, Tony. 2000. The incidental parameter problem since 1948. *Journal of Econometrics* 95: 391–413. [CrossRef]

Lechner, Michael. 1995. Some specification tests for probit models estimated on panel data. *Journal of Business and Economic Statistics* 13: 475–88. [CrossRef]

Liu, Qing, and Donald A. Pierce. 1994. A note on gauss-hermite quadrature. *Biometrika* 83: 624–29. [CrossRef]

Mazodier, Pascal, and Alain Trognon. 1978. Heteroskedasticity and stratification in error components models. *Annales de l'INSEE* 30: 451–82.

Montes-Rojas, Gabriel, and Walter Sosa-Escudero. 2011. Robust tests for heteroskedasticity in the one-way error components model. *Journal of Econometrics* 160: 300–10. [CrossRef]

Moussa, Richard, and Eric Delattre. 2018. On the estimation of causality in a bivariate dynamic probit model on panel data with stata software. a technical review. *Theoretical Economics Letters* 8: 1257–78. [CrossRef]

Naylor, Jennifer C., and Adrian F. M. Smith. 1982. Applications of a method for the efficient computation of posterior distributions. *Applied Statistics* 31: 214–25. [CrossRef]

Randolph, William C. 1988. A transformation for heteroscedastic error components regression models. *Economics Letters* 27: 349–54. [CrossRef]

Verbon, Harrie. 1980. Testing for heteroscedasticity in a model of seemingly unrelated regression equations with variance components. *Economics Letters* 5: 149–53. [CrossRef]

Wansbeek, Tom. 1989. An alternative heteroscedastic error components model, solution 88.1.1. *Econometric Theory* 5: 326. [CrossRef]

Wooldridge, Jeffrey M. 2001. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: The MIT Press.