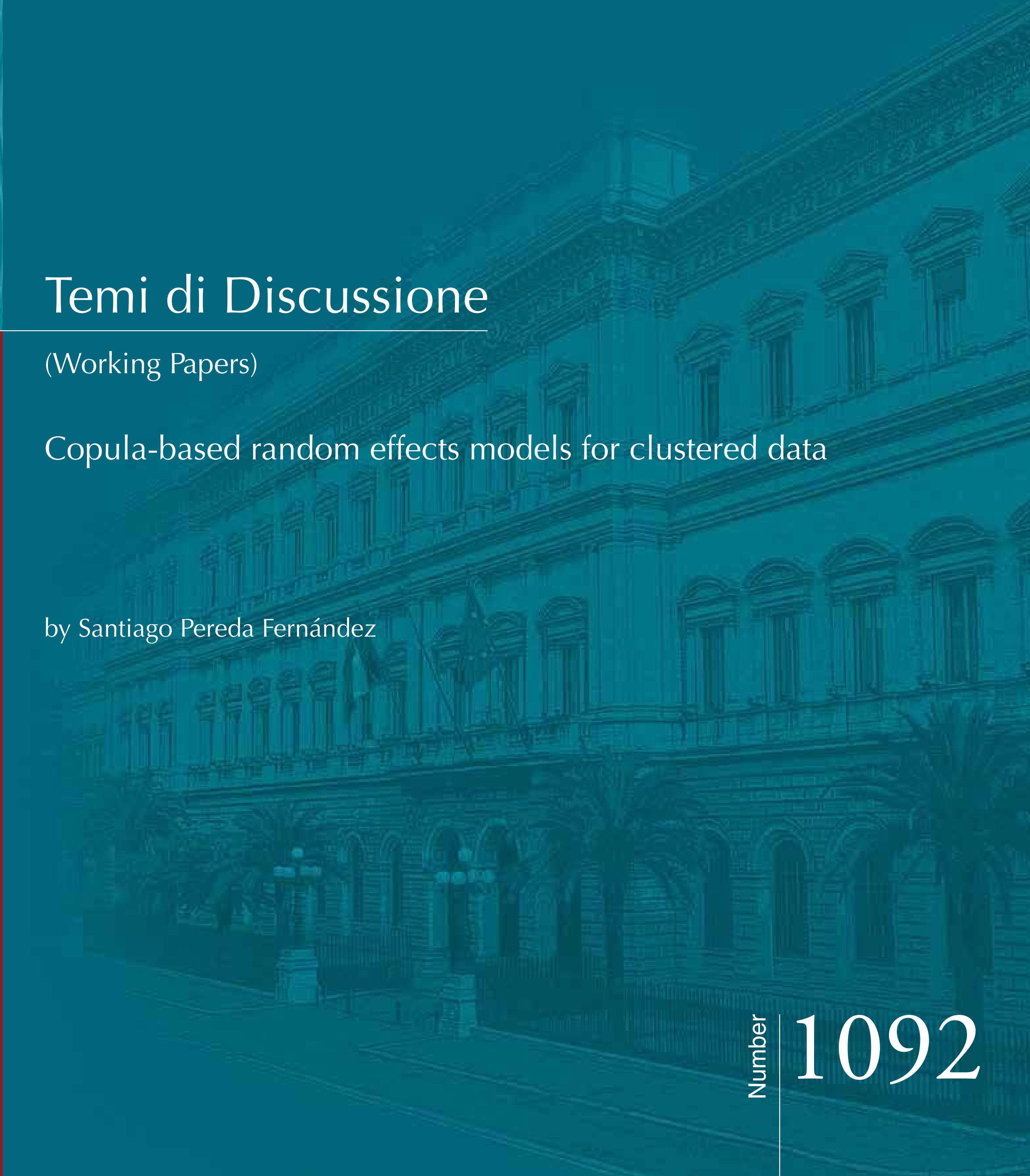
BANCA D'ITALIA
EUROSISTEMA

Temi di Discussione

(Working Papers)

Copula-based random effects models for clustered data

by Santiago Pereda Fernández



Number 1092

December 2016

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working papers)

Copula-based random effects models for clustered data

by Santiago Pereda Fernández

Number 1092 - December 2016

The purpose of the Temi di discussione *series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.*

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.

ISSN 1594-7939 (print)
ISSN 2281-3950 (online)

Printed by the Printing and Publishing Division of the Bank of Italy

COPULA-BASED RANDOM EFFECTS MODELS FOR CLUSTERED DATA

by Santiago Pereda Fernández*

Abstract

Sorting and spillovers can create correlation in individual outcomes. In this situation, standard discrete choice estimators cannot consistently estimate the probability of joint and conditional events, and alternative estimators can yield incoherent statistical models or intractable estimators. I propose a random effects estimator that models the dependence among the unobserved heterogeneity of individuals in the same cluster using a parametric copula. This estimator makes it possible to compute joint and conditional probabilities of the outcome variable, and is statistically coherent. I describe its properties, establishing its efficiency relative to standard random effects estimators, and propose a specification test for the copula. The likelihood function for each cluster is an integral whose dimension equals the size of the cluster, which may require high-dimensional numerical integration. To overcome the curse of dimensionality from which methods like Monte Carlo integration suffer, I propose an algorithm that works for Archimedean copulas. I illustrate this approach by analysing labour supply in married couples.

JEL Classification: C23, C25, J22.
Keywords: copula, high-dimensional integration, nonlinear panel data.

Contents

* Bank of Italy, Structural Economic Analysis Directorate.

1 Introduction[*]

There are several methods to accommodate individual heterogeneity in binary choice panel data models. When these unobserved individual effects are mutually independent, these methods can consistently estimate all the relevant parameters of a model.[1] However, if they display some degree of correlation, these estimators will fail to consistently estimate the probability that a joint or a conditional event occurs, as the joint distribution depends on the correlation structure of the unobservables. This is the case in many real world situations, as agents interact with each other, or they are influenced by other factors that are not observed by the econometrician. For example, students in the same classroom tend to interact with each other and they learn from the same teacher, leading to within classroom correlated test scores (Hanushek, 1971), or sorting in marriages (Bruze, 2011; Charles et al., 2013) may lead to both partners having a similar propensity to work.

There are two main contributions in this paper: first, I present the Copula-Based Random Effects estimator (CBRE) for clustered data. I show how to use this estimator to consistently estimate the probability of joint and conditional events, as well as average partial effects. Moreover, I adapt Vuong (1989) test to this framework, which results in a specification test that can be used to choose the copula with the best fit, and show the asymptotic efficiency of the CBRE estimator relative to standard random effects estimators. Second, I propose an algorithm to numerically approximate high-dimensional integrals with Archimedean copulas.

I consider a setup in which individual outcomes are correlated if they belong to the same cluster, but they are independent across clusters. Simultaneous equations models with limited dependent variables face several challenges, which could result in statistically

[*]Banca d'Italia, Via Nazionale 91, 00184 Roma, Italy. I would like to thank Lorenzo Burlon, Domenico Depalo, Simone Emiliozzi, Iván Fernández-Val, Giuseppe Ilardi, Marco Savegnago, Laura Sigalotti, Paolo Zacchia, and seminar participants at Banca d'Italia, the 2015 EWM of the Econometric Society, and the 3rd Annual Conference of the IAAE. All remaining errors are my own. This paper is based on data from Eurostat, EU Statistics on Income and Living Conditions. The views presented in this paper do not necessarily reflect those of the Banca d'Italia, and the responsibility for all conclusions drawn from the data lies entirely with the author. I can be reached via email at santiago.pereda@.bancaditalia.it.

[1]See Arellano and Bonhomme (2011) for a survey on existing approaches to deal with unobserved heterogeneity with fixed T.

incoherent models in which the sum of all probabilities does not equal one (Maddala and Lee, 1976; Heckman, 1978; Schmidt, 1981). Game theoretic models guarantee the existence of an equilibrium, but neither its unicity (De Paula, 2013), nor an easily implementable estimator are guaranteed (particularly if the number of players is large), and if the model is misspecified, then the estimator may be inconsistent (Brock and Durlauf, 2001). On the other hand, the CBRE estimator is statistically coherent, easy to implement, and allows to compute the conditional or joint probabilities that can arise in social interactions contexts.[2]

This estimator extends standard binary choice random effects estimators: the likelihood function takes the form of an integral of the probability of the observed data, conditional on the unobserved heterogeneity, over the joint distribution of the latter. Rather than directly specifying the joint multivariate distribution for the unobserved heterogeneity, I separately model their marginal distribution and their copula, which together characterize the joint distribution. This isolates the only difference with respect to regular random effects estimators, and provides a flexible way of modeling the correlation, as it allows the combination of a copula with different marginals, and the other way around.

Because of the within cluster correlation, the integrals are multidimensional, where the dimension equals the number of individuals in each cluster. Estimators of this kind face implementation challenges because of the curse of dimensionality. Thus, much of the empirical analysis has been constrained to use quadrature methods for the integration when the latent variable has a low dimension (up to 4), or simulation-based methods (Hajivassiliou et al., 1996), which are usually restricted to the multivariate normal distribution. I propose an algorithm that can be used to numerically approximate high-dimensional integrals which works for Archimedean copulas, and can be also extended to elliptical copulas.[3] I compare the performance of the algorithm to that of Monte Carlo integration, and show that this algorithm overcomes the curse of dimensionality for the type of approximations considered.

[2]It is left for the econometrician to determine whether the estimated probabilities have a causal interpretation in each empirical study.

[3]Belloni and Alessie (2013) used the GHK simulator to approximate integrals of dimension 11, which, as far as I know, is the highest-dimensional integral that has been approximated in applied work. On the other hand, the algorithm I propose was used in Pereda-Fernández (2015) to approximate integrals of dimension up to 35.

This paper builds on the literature of panel data discrete choice models (Chamberlain, 1980, 1984) and other nonlinear panel data models. The early focus was on identification of the slope parameter without imposing distributional assumptions on the unobserved heterogeneity, such as Chamberlain (1980), or Manski (1987), who proposed the Maximum Score Estimator (Manski, 1975) to consistently estimate the slope parameter. Lee (1999) proposed an alternative estimator that, unlike the Maximum Score Estimator, achieves the parametric convergence rate. More recently other authors have allowed the slopes to be heterogeneous (Altonji and Matzkin, 2005; Browning et al., 2007), including the possibility of a non-monotonic random coefficients model (Fox et al., 2012; Gautier and Kitamura, 2013). Many of these identification results require either large support of at least one of the covariates, or the idiosyncratic error term to be logistically distributed (Chamberlain, 2010). In contrast with these approaches, I attempt to provide a random effects estimator whose identification hinges on parametric assumptions, but which does not require any of the previous two conditions to hold.

There are two main alternatives to modeling individual heterogeneity with random effects: a fixed effects approach that eliminates the individual heterogeneity from the objective function, and estimating the individual effects as any other parameter of the model. In a linear panel data setup, differencing the dependent variable out eliminates the need to control for any correlation in the unobserved heterogeneity, but this is generally not possible in a nonlinear setup.[4] Also, because of the incidental parameter problem, using individual dummies to control for the individual heterogeneity would result in inconsistent estimates. There exist bias-corrected estimators (e.g. Hahn and Newey 2004; Fernández-Val 2009), although they often rely on a relatively large number of periods and they assume the unobserved heterogeneity to be independent across individuals.

To illustrate the applicability of the estimator, I use it to estimate the correlation in the individual propensity to work of married couples for several European countries using the

[4]Conditional fixed effects logit (Chamberlain, 1980) does not depend on the unobserved heterogeneity, although it cannot be used to estimate average partial effects or the probability of a joint event *unconditionally*. The same problem applies to any other deconvolution that differences out the distribution of the unobserved heterogeneity, as in Bonhomme (2012).

EU-SILC dataset, finding that there is at most a modest degree of correlation. Using these estimates, I compute some counterfactual probabilities, finding that ignoring this correlation leads to biases in the estimation of the probability that at least one member of the couple is employed in each period, or the probability of the wife being employed conditional on the employment status of the husband.

The rest of the paper is organized as follows: in section 2 I describe the econometric framework and present the estimator, along with the specification tests. In section 3 I discuss its asymptotic properties. Section 4 describes the algorithm used for the approximation of the multidimensional integral. The results of a Monte Carlo simulation are shown in section 5, whereas section 6 shows the results of the empirical application. Section 7 concludes. All proofs are shown in appendix A.

2 Framework and Estimation

To motivate and illustrate the econometric problem, consider the labor supply of married couples over time.[5] There is evidence that the probability of the women being employed is higher if the husband is also employed. A researcher trying to model this as a simultaneous equations model in which, in each period, the dependent variable of the partner has an impact on the probability of being employed, would run into a statistically incoherent model unless only one of the two can impact the partner's behavior. Formally, if $y_{it} = \mathbf{1}\left(\alpha_i y_{jt} + \eta_i + x_{it}'\beta_0 + \varepsilon_{it} \geq 0\right)$, for $i, j = 1, 2$, $i \neq j$, the model is statistically incoherent if $\alpha_i \alpha_j \neq 0$. This condition essentially eliminates the simultaneity of the model, and an analogous situation arises when the number of simultaneous equations (cluster size) increases.

To overcome this problem, one could try to model the labor supply of couples as a non-cooperative game. There exist several games to model it, each of which relies on different assumptions and leads to different results.[6] Also, because of the multiplicity of equilibria,

[5] It is also possible to consider quasi-panels. This was the case considered in Pereda-Fernández (2015), in which the different questions of an exam constituted the time dimension, and students' individual effects were correlated within classrooms.

[6] See e.g. Kaya (2014) for a study of the labor supply in couples in which the author compares the

these models open the door for an equilibrium selection mechanism (e.g. Pareto optimality) which may cause the model to be misspecified, or one could be agnostic about it at the cost of having partial identification of the parameters (De Paula, 2013). Further, games with 2 players and 2 possible actions can be easy to model, but as the number of players increases, the model becomes less tractable. Hence, if one wanted to model a game in which students either pass or fail an exam, and these events are simultaneously determined for all students in the same classroom, the number of equilibria could be too large.

The approach in this paper is to generalize a random effects model to allow the individual heterogeneity of those individuals in the same cluster (the two partners of a married couple, or all the students in a classroom) to be correlated within clusters. From a statistical standpoint, this model is coherent, and consequently it can be used to compute joint and conditional probabilities. For example, it can be used to estimate the probability of a woman being employed conditional on the husband being employed, and compare it to the probability when the husband is unemployed, as well as its evolution over time. Similarly, it can be used to estimate the probability that at least one of them is employed in every period. Formally, consider the following nonlinear panel data setup:[7]

$$\begin{aligned} y_{ict} &= \mathbf{1}\left(y^*_{ict} \geq 0\right) \\ y^*_{ict} &= \eta_{ic} + x'_{ict}\beta_0 + \varepsilon_{ict} \end{aligned} \tag{1}$$

where the econometrician observes the dependent variable y_{ict} and the covariates x_{ict} for agent $i = 1, ..., N_c$ in cluster $c = 1, ..., C$ at time $t = 1, ..., T$. The main departure in this framework from the usual one is that, for each cluster c, the individual effects of the N_c members of the cluster are correlated with each other, though they are independent of the individual effects of the members from other clusters.

I model this dependence by separately considering the marginal distribution of each

estimates for several game specifications.

[7] I present a binary choice model, but any other limited dependent variable model, such as censored data, can exhibit group dependence in their individual unobservables. It is also possible to model this dependence using copulas and adapt the techniques presented in this paper for the estimation in those setups.

individual effect, η_{ic}, and the underlying correlation among them using a copula.[8] Copulas are multivariate cdfs whose arguments are the ranks of the individual effects, *i.e.* $u_{ic} = F_\eta(\eta_{ic};\sigma_0)$. Hence, they are invariant to the marginal distribution of the effects. To complete the model, assume that the distribution of ε_{ict}, F_ε, satisfies $\mathbb{P}(y_{ict}|x_{ict},\eta_{ic}) = 1 - F_\varepsilon(-(x'_{ict}\beta_0 + \eta_{ic}))$, $\eta_{ic} \sim F_\eta(\sigma_0)$, and the joint distribution of the individual effects in cluster c is given by $u_c \equiv (u_{1c},...,u_{N_c c})' \sim C(u_c;\rho_0)$.[9] σ_0 and ρ_0 respectively denote the parameters of the marginal distribution and the copula of the individual effects. The latter determines the amount of correlation of these effects and typically nests the independence copula. Denote the vector of parameters by $\theta \equiv (\beta',\sigma',\rho')'$, $z_{ict} \equiv (y_{ict}, x'_{ict})'$, and define the vectors of stacked individual variables by the ic subscript, and the vectors of stacked group-individual variables by the c subscript. The log-likelihood function is given by

$$\mathcal{L}(\theta) = \sum_{c=1}^{C} \log\left(\int_{[0,1]^{N_c}} \prod_{i=1}^{N_c} P_{ic}(z_{ic},\eta_{ic};\beta)\, dC(u_c;\rho)\right) \tag{2}$$

where $P_{ic}(z_{ic},\eta_c;\beta) \equiv \prod_{t=1}^{T}\left[1 - F_\varepsilon(-(\eta_{ic} + x'_{ict}\beta))\right]^{y_{ict}} F_\varepsilon(-(\eta_{ic} + x'_{ict}\beta))^{1-y_{ict}}$ and $\eta_{ic} = F_\eta^{-1}(u_{ic};\sigma)$. The identification of θ is based on the parametric assumptions of the model. As pointed out by Arellano (2003), identification in a binary choice panel setup is fragile, and it usually hinges on assumptions that are not satisfied in certain applications, such as having at least a regressor with positive Lebesgue density over the whole real line. Chernozhukov et al. (2013) showed that when the distribution of the regressors is discrete, the distribution of the individual effects is not identified. This result can be extended to the lack of identification of their copula, which is shown in lemma 1 in appendix A.

The CBRE estimator is given by $\hat{\theta} = \arg\max_{\theta\in\Theta}\mathcal{L}(\theta)$. Note that this estimator requires the integration of a product over a potentially large dimensional space. In the following subsections, I show how to estimate joint and conditional events, and average partial effects. Moreover, to address the problem of the choice of the copula, I propose a specification test to

[8] Sklar (1959) showed that any continuous multivariate cdf can be written in terms of a copula whose arguments are the marginal distributions, *i.e.* $\mathbb{P}(X_1 \le x_1,...,X_d \le x_d) = C(F_1(x_1),...,F_d(x_d))$.

[9] It can be extended to allow for the existence of correlated random effects by letting the marginal distribution of η_{ic} depend on $x_{ic} \equiv (x_{ic1},...,x_{icT})'$. Consequently, the *conditional* ranks would be given by $u_{ic} = F_\eta(\eta_{ic}|x_{ic};\sigma_0)$. Similarly, one could allow the copula to depend on x_c.

choose between any two parametric copulas, and another one to check whether the estimated copula is significantly different from the independence copula.[10]

2.1 Estimation of Joint and Conditional Events

Let $\mathcal{S}$ denote all the permutations of $y_c \equiv (y_{1c1}, ..., y_{1cT}, ..., y_{N_c cT})$ that satisfy a condition $\mathcal{C}$. In the labor supply example, the set $\mathcal{S} = \{y_c : y_{1ct} + y_{2ct} \geq 1 \ \forall t\}$ denotes all the possible situations in which at least one of the two partners is employed in every period, where $y_{ict} = 1$ if individual i is employed at time t. The probability of such events is given by

$$\mathbb{P}\left(y_c \in \mathcal{S}|x_c\right) = \sum_{d\in\mathcal{S}} \mathbb{P}\left(y_c = d|x_c\right) = \sum_{d\in\mathcal{S}} \int_{[0,1]^{N_c}} \prod_{i=1}^{N_c} \prod_{t=1}^{T} \mathbb{P}\left(d|x_{ict}, \eta_{ic}\right) dC\left(u_c; \rho_0\right) \tag{3}$$

where $\mathbb{P}\left(d_{ict}|x_{ict}, \eta_{ic}\right) = \left[1 - F_\varepsilon\left(-\left(\eta_{ic} + x'_{ict}\beta_0\right)\right)\right]^{d_{ict}} F_\varepsilon\left(-\left(\eta_{ic} + x'_{ict}\beta_0\right)\right)^{1-d_{ict}}$. To estimate the probability that the outcome satisfies $\mathcal{C}$, replace θ_0 by $\hat{\theta}$ and approximate the integral as shown in section 4. Also, note that it is straightforward to compute the probability of conditional events once the joint and marginal events are known: if one wants to estimate the probability of an event A given B, estimate their joint probability and the marginal of the event B, and divide the joint by the marginal. Continuing with the labor supply example, one could estimate the probability of a woman being employed conditional on her husband being (un)employed.

2.2 Estimation of Average Partial Effects

Frequently, the econometrician is after the estimation of the average partial effect rather than the regressor coefficients. The average partial effect is defined as the marginal effect that increasing a regressor x_{ictj} would have on the probability of the dependent variable being equal to one, averaged over the whole population. Mathematically,

$$APE\left(x_{ictj}\right) \equiv \int_{\mathbb{R}} \frac{\partial}{\partial x_{ixtj}} \mathbb{P}\left(y_{ict} = 1|x_{ict}, \eta_{ic}\right) dF_\eta\left(\eta_{ic}; \sigma_0\right) \tag{4}$$

[10]It is worth noticing that the individual effects are unobserved, and therefore it is not possible to carry out a visual analysis to help the researcher choose the most appropriate copula.

Since it just depends on the marginal distribution of η_{ic}, there is no need to know the copula to identify them, and it can be computed the standard way using the sample analogue.[11]

2.3 Specification Tests

Testing a parametric copula against another

Consider two different parametric copulas, $C_1(u_c;\rho)$ and $C_2(u_c;\xi)$, where both ρ and ξ belong to the interior of their respective parameter spaces. A researcher who has no theoretical basis to choose one over the other may want to choose whichever has the best fit to the data. This is the strictly non-nested model considered by Vuong (1989), and neither of the two copulas are necessarily the true one. Denote their respective likelihoods by $\ell_{1,c}(z_c;\theta_1)$ and $\ell_{2,c}(z_c;\theta_2)$, where $\theta_1 \equiv (\mu',\rho')'$ and $\theta_2 \equiv (\mu',\xi')'$, where $\mu \equiv \left(\beta',\sigma_\eta'\right)'$ is the vector with the marginal parameters. The null hypothesis in this case is that both copulas are equivalent, *i.e.*

$$H_0 : \mathbb{E}\left[\log\frac{\ell_{1,c}(z_c;\theta_1)}{\ell_{2,c}(z_c;\theta_2)}\right] = 0$$

against the alternatives that C_1 is better than C_2,

$$H_1 : \mathbb{E}\left[\log\frac{\ell_{1,c}(z_c;\theta_1)}{\ell_{2,c}(z_c;\theta_2)}\right] > 0$$

or that C_2 is better than C_1,

$$H_2 : \mathbb{E}\left[\log\frac{\ell_{1,c}(z_c;\theta_1)}{\ell_{2,c}(z_c;\theta_2)}\right] < 0$$

The test statistic takes the form of a likelihood ratio, whose asymptotic distribution under the null is given by (theorem 5.1 in Vuong (1989))

$$\frac{1}{\sqrt{C}}\frac{LR\left(\hat{\theta}_1,\hat{\theta}_2\right)}{\hat{\omega}} \xrightarrow{d} \mathcal{N}(0,1)$$

[11] It is worth remembering however, that the APE depend on the parametric assumptions. There is a vast literature that focuses on the identification and estimation of APE in this and other related frameworks. See, for instance, Graham and Powell (2012), Chernozhukov et al. (2013), or Fernández-Val and Lee (2013).

where

$$LR\left(\hat{\theta}_1, \hat{\theta}_2\right) \equiv \sum_{c=1}^{C} \log \frac{\ell_{1,c}\left(z_c; \hat{\theta}_1\right)}{\ell_{2,c}\left(z_c; \hat{\theta}_2\right)}$$

$$\hat{\omega}^2 \equiv \frac{1}{C} \sum_{c=1}^{C} \left[\log \frac{\ell_{1,c}\left(z_c; \hat{\theta}_1\right)}{\ell_{2,c}\left(z_c; \hat{\theta}_2\right)}\right]^2 - \left[\frac{1}{C} \sum_{c=1}^{C} \log \frac{\ell_{1,c}\left(z_c; \hat{\theta}_1\right)}{\ell_{2,c}\left(z_c; \hat{\theta}_2\right)}\right]^2$$

Notice that this test could also be used to test the marginal distribution of the individual effects (e.g. normal versus Laplace distribution), the distribution of the idiosyncratic term (e.g. probit versus logit), or a combination of them.

Testing for independence of the copula

For most parametric copulas, the independence case is a particular value of the parameters of the copula, so testing for independence amounts to testing whether those parameters are statistically equal to ρ^{ind}. If ρ^{ind} is in the interior of the parameter space (e.g. for the bivariate Gaussian copula, in which the null hypothesis is $H_0 : \rho = 0$, where $\rho \in [-1, 1]$), it is easy to test the null hypothesis using standard tests, such as a t-test.

A more complicated situation arises if ρ^{ind} lies on the boundary of the parameter space.[12] Self and Liang (1987) showed that in this case, the maximum likelihood estimator is still consistent, but not asymptotically normal. For expositional clarity, I focus on the case in which ρ is univariate. Let $Z \sim \mathcal{N}(0, \sigma_Z^2)$. The limiting distribution of $\sqrt{C}\left(\hat{\rho} - \rho^{ind}\right)$ is given by $Z\mathbf{1}(Z > 0)$ where σ_Z^2 is the element of the inverse of the information matrix that corresponds to ρ. In words, the asymptotic distribution is the 50:50 mixture of a degenerate distribution at 0 and a χ_1^2. Then, one would not accept the null hypothesis of independence if $\hat{\rho}$ is greater than the 95th percentile of this distribution.

[12]For some copulas, the parameter space is bounded, with the value that makes the copula independent lying at the boundary. For example, a $Clayton\,(0)$, a $Gumbel\,(1)$, or a $Frank\,(0)$ are actually the independence copula.

3 Properties

Let $P_c(z_c, \eta_c; \beta) \equiv \prod_{i=1}^{N_c} P_{ic}(z_{ic}, \eta_{ic}; \beta)$, $\ell_c(z_c; \theta) \equiv \int_{[0,1]^{N_c}} \prod_{i=1}^{N_c} P_{ic}(z_{ic}, \eta_{ic}; \beta)\, dC(u_c; \rho)$, and $\ell_{ic}(z_{ic}; \mu) \equiv \int_{\mathbb{R}} P_{ic}(z_{ic}, \eta_{ic}; \beta)\, dF_\eta(\eta_{ic}; \sigma)$. In order to derive the asymptotic properties of the estimator, let the following assumptions hold:

Assumption 1. *η_c is iid for all $c = 1, ..., C$, x_{ic} are iid $\forall c = 1, ..., C, i = 1, ..., N_c$, and ε_{ict} are iid for all $c = 1, ..., C, i = 1, ..., N_c, t = 1, ..., T$.*

Assumption 2. *$\theta \neq \theta_0 \Rightarrow \ell_c(z_c; \theta) \neq \ell_c(z_c; \theta_0)$.*

Assumption 3. *$\theta \in int\Theta$, where Θ is compact.*

Assumption 4. *$\ell_{ic}(z_{ic}; \mu)$ is continuous for all $\theta \in \Theta$.*

Assumption 5. $\mathbb{E}\left[\sup_{\theta \in \Theta} |\log(\ell_{ic}(z_{ic}; \mu))|\right] < \infty$.

Assumption 6. *$\ell_{ic}(z_{ic}; \mu)$ is twice continuously differentiable with respect to θ; $\ell_{ic}(z_{ic}; \mu) > 0$ in a neighborhood $\mathcal{N}$ of θ_0.*

Assumption 7. $\int \sup_{\theta \in \mathcal{N}} \|\nabla_\mu \ell_{ic}(z_{ic}; \mu)\|\, dz_{ic} < \infty$, $\int \sup_{\theta \in \mathcal{N}} \|\nabla_{\mu\mu} \ell_{ic}(z_{ic}; \mu)\|\, dz_{ic} < \infty$.

Assumption 8. $\Sigma_\theta^{CBRE} \equiv \left[\nabla_\theta \log(\ell_c(z_c; \theta))\, \nabla_\theta \log(\ell_c(z_c; \theta))'\right]$ *exists and is nonsingular.*

Assumption 9. $\mathbb{E}\left[\sup_{\theta \in \mathcal{N}} \|\nabla_{\theta\theta} \log(\ell_c(z_c; \theta))\|\right] < \infty$.

Assumption 10. *The copula has pdf $c(u_c; \rho)$ which is twice continuously differentiable in ρ, and is bounded by $0 < \underline{c} < c(u_c; \rho) < \overline{c} < \infty$ $\forall \theta \in \Theta$. Moreover, $\forall \theta$ in a neighborhood $\mathcal{N}$ of θ_0, the first and second derivatives are bounded in absolute value by $\overline{c}_1$ and $\overline{c}_2$, respectively.*

Assumption 11. *Cluster size is either predetermined, or it is drawn from a distribution with bounded support, independently of all other variables: $N_c \sim F_N(n)$ $n \in \left\{1, ..., \overline{N}\right\}$, for some $\overline{N} \in \mathbb{N}$.*

Assumptions 1 to 9 mimic the assumptions in theorems 2.5 and 3.3 in Newey and McFadden (1994). With some small modifications, these assumptions work for standard binary choice random effects models[13]. In other words, they allow us to extend any binary choice random effects estimator to have the cluster dependence described in this paper. Notice that it would be possible to further relax some of these assumptions, such as allowing within group cross-sectional dependence in the covariates, but relaxing some of them could result in non-standard properties: if assumption 3 is relaxed and the true value of the parameter lies at the boundary of the parameter space, the asymptotic distribution will not be normal.

Assumption 10 imposes smoothness restrictions on the copula, as well as some bounds on its distribution function. An implication of this assumption is that it rules out the perfect correlation case, in which the copula has no proper pdf. However, the independence case is covered by this assumption, since in that case the pdf equals one everywhere. Assumption 11 limits cluster size to $\overline{N}$, ruling out the possibility that the size of a group grows to infinity as the sample size grows. This assumption is required to bound the likelihood function, and it should be satisfied in most applications. Regarding its independence with respect to all other variables, it could be relaxed at the cost of complicating the analysis[14].

The following proposition establishes the asymptotic distribution of the CBRE estimator.

Proposition 1. *Under assumptions 1 to 11, the CBRE estimator $\hat{\theta}$ is a consistent estimator for θ_0 and its asymptotic distribution is given by $\sqrt{C}\left(\hat{\theta}-\theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_\theta^{CBRE}\right)$.*

Estimation of the asymptotic variance is standard and it also requires a multi-dimensional numerical integration. See appendix B for more details and the exact form of the score used to find the maximum.

3.1 Comparison with Standard Random Effects Methods

If one were not interested in the estimation of ρ, an alternative estimator to the CBRE estimator would be the regular RE estimator, $\tilde{\mu}$, which ignores the within cluster correlation

[13]They should be rewritten to reflect the fact that the marginal distributions do not depend on ρ. Hence, one should appropriately redefine the parameter space, the hessian, etc.

[14]For example, if the distribution of class size depended on θ, the jacobian and the hessian would become harder to implement.

of the individual effects and is the maximizer of the following pseudo-likelihood function:[15]

$$\tilde{\mathcal{L}}(\mu) = \sum_{c=1}^{C} \sum_{i=1}^{N_c} \log\left(\ell_{ic}\left(z_{ic};\mu\right)\right) = \sum_{c=1}^{C} \log\left(\int_{[0,1]^{N_c}} \prod_{i=1}^{N_c} P_{ic}\left(z_{ic},\eta_{ic};\beta\right) \prod_{i=1}^{N_c} du_{ic}\right) \tag{5}$$

Notice that if we denote by ρ^{ind} the value of ρ that makes the copula independent, then $\tilde{\mathcal{L}}(\mu) = \mathcal{L}\left(\mu,\rho^{ind}\right)$, and hence equation 5 is a particular case of equation 2. Therefore $\mathcal{L}(\hat{\mu},\hat{\rho}) \geq \mathcal{L}\left(\tilde{\mu},\rho^{ind}\right)$, so in general each estimator yields a different estimate, and the CBRE estimator is in general more efficient than the RE estimator.[16] To see this, consider the expected scores of both estimators:[17]

$$\mathbb{E}\left[\sum_{i=1}^{N_c} \nabla_\mu \log\left(\ell_{ic}\left(z_{ic};\mu_0\right)\right)\right] \equiv \mathbb{E}\left[\frac{\int_{[0,1]^{N_c}} \nabla_\mu P_c\left(z_c,\eta_c;\beta_0\right) \prod_{i=1}^{N_c} du_{ic}}{\int_{[0,1]^{N_c}} P_c\left(z_c,\eta_c;\beta_0\right) \prod_{i=1}^{N_c} du_{ic}}\right] = 0$$

$$\mathbb{E}\left[\nabla_\mu \log\left(\ell_c\left(z_c;\theta_0\right)\right)\right] \equiv \mathbb{E}\left[\frac{\int_{[0,1]^{N_c}} \nabla_\mu P_c\left(z_c,\eta_c;\beta_0\right) dC\left(u_c;\rho_0\right)}{\int_{[0,1]^{N_c}} P_c\left(z_c,\eta_c;\beta_0\right) dC\left(u_c;\rho_0\right)}\right] = 0$$

$$\mathbb{E}\left[\nabla_\rho \log\left(\ell_c\left(z_c;\theta_0\right)\right)\right] \equiv \mathbb{E}\left[\frac{\int_{[0,1]^{N_c}} P_c\left(z_c,\eta_c;\beta_0\right) \nabla_\rho c\left(u_c;\rho_0\right) \prod_{i=1}^{N_c} du_{ic}}{\int_{[0,1]^{N_c}} P_c\left(z_c,\eta_c;\beta_0\right) dC\left(u_c;\rho_0\right)}\right] = 0$$

Under the assumption that the likelihood is correctly specified, the asymptotic variance of the CBRE estimator is not larger than the asymptotic variance of the RE estimator. The following proposition establishes under which condition both estimators are asymptotically efficient:

Proposition 2. *The RE estimator is as efficient as the CBRE estimator of* μ *if the asymptotic variance of the latter equals* $-\mathbb{E}\left[\sum_{i=1}^{N_c} \nabla_{\mu\mu} \log\left(\ell_{ic}\left(z_{ic},\mu_0\right)\right)\right]^{-1}$.

Consequently, if the copula is fully known, *i.e* if it does not depend on ρ or if ρ is known, then RE has the same asymptotic variance as CBRE if and only if the actual copula is the independence copula, as stated in the next corrolary:

[15] The most common estimators are *RE logit* and *RE probit*. See Wooldridge (2010) for further details.

[16] This is a conceptually similar problem to the one considered by Prokhorov and Schmidt (2009), but unlike their case, the copula here models a latent variable, complicating the analysis, as the likelihood cannot be decomposed into the sum of the logarithms of the marginals and the copula.

[17] Notice that the expectations are at the cluster level, to reflect the dependence across individual effects. Therefore, the asymptotic distribution of the RE estimator is not the usual one, which assumes no correlation.

Corollary 1. *If ρ_0 is known, then the asymptotic variance of the RE and CBRE estimators coincide if and only if $C(u_c;\rho_0) = \prod_{i=1}^{N_c} u_{ic}$.*

In any case, it is necessary to take into account the within cluster correlation to compute the standard errors of the maximum likelihood estimates, which the likelihood of the CBRE estimator does by default, unlike the RE estimator. For the latter, a correction is required to reflect the actual variability of the estimator.

4 Implementation Algorithm

As in any discrete choice random effects model, implementing the estimator requires the numerical approximation of an integral. Both the Jacobian and the Hessian of the likelihood function depend on the copula of the individual effects, which implies that the dimension of the integrals equals the size of the clusters. Simulation methods like Monte Carlo tend to perform slowly in this kind of setup, and have been outperformed by some recent advances in high dimensional numerical integration methods, although they are still subject to the curse of dimensionality.[18] I propose an algorithm to numerically approximate a class of integrals when the copula is Archimedean.[19] A variation of this algorithm can also be used with Elliptical copulas, as shown in appendix D.

By Corollary 2.2 in Marshall and Olkin (1988), an Archimedean copula is given by

$$C(u) = \int_{\mathbb{R}_+^N} \exp\left(-\sum_{i=1}^{N} \theta_i \phi_i^{-1}(u_i)\right) dG(\theta) \tag{6}$$

where G is the cdf of θ, and ϕ_i is the Laplace transform of the marginal distributions of G. Some of the most common copulas have θ unidimensional and $\phi_i = \phi \forall i$. Consider the following integral:

$$\mathcal{I} = \int_{[0,1]^N} \prod_{i=1}^{N} \ell_i(u_i)\, dC(u) = \int_{\mathbb{R}_+} \prod_{i=1}^{N} \left[\int_0^1 \ell_i(u_i)\, dF^\theta(u_i)\right] dG(\theta) \tag{7}$$

[18]See, for instance, Heiss and Winschel (2008) or Skrainka and Judd (2011).

[19]A copula is Archimedean if $C(u_1, ..., u_d; \rho) = \phi\left(\sum_{i=1}^{d} \phi^{-1}(u_i;\rho);\rho\right)$, where ϕ is the so called generator function. See appendix C for more details.

where $F^{\theta}(u_i) = \exp\left(-\theta\phi^{-1}(u_i)\right)$. The originally N-dimensional integral can be expressed as the integral of the product of N independent integrals, reducing the dimensionality from N to 2. Hence, an algorithm to approximate the integral is given by

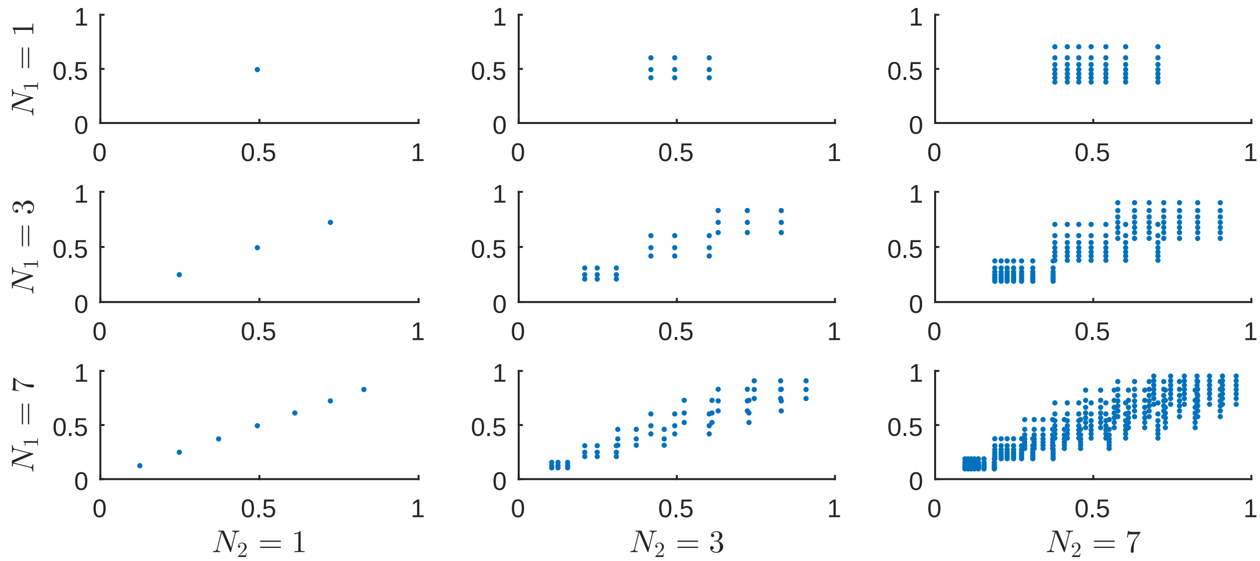
1. Compute a grid of values of θ, given by $\theta_j = G^{-1}\left(\frac{j}{N_1+1}\right)$, $\forall j = 1, ..., N_1$.
2. Compute a grid of values of u $\forall j$, given by $u_{jh} = \phi\left(-\frac{1}{\theta_j}\log\left(\frac{h}{N_2+1}\right)\right)$, $\forall h = 1, ..., N_2$.
3. Approximate the integral by $\hat{\mathcal{I}} = \frac{1}{N_1}\sum_{j=1}^{N_1}\prod_{i=1}^{N}\left[\frac{1}{N_2}\sum_{h=1}^{N_2}\ell_i(u_{jh})\right]$.

To understand how the algorithm works, consider the integration of a function g with respect to a distribution F: $\int g(x)\, dF(x)$. One possibility is to split the unit interval into $Q+1$ intervals of equal probability using Q evenly spaced quantiles: $0 < \frac{1}{Q+1} < ... < \frac{Q}{Q+1} < 1$. Then the integral is approximated by evaluating the function g at these quantiles and taking the average across quantiles: $\frac{1}{Q+1}\sum_{q=1}^{Q} g(x_q)$, where $x_q = F^{-1}\left(\frac{q}{Q+1}\right)$ is the qth quantile of the function F.[20]

The algorithm uses this approximation twice, and figure 1 shows how the selection of the points used for integration is done in practice. Suppose that we want to do an integration using a bivariate Clayton copula. For a fixed value of θ, we have N_2 different values for each u_i as shown in the upper graphs. These points split the unit interval into N_2+1 intervals that have the same probability of occurring, conditional on θ. Thus, it is possible to approximate the inner integral of each dimension j as $\frac{1}{N_2+1}\sum_{h=1}^{N_2}\ell_i(u_{jh})$. The symmetry of the copula means that the points u_{jh} are indeed the same for each dimension, so there is no need to compute a different number of points of support for each dimension. Then, to approximate the outer integral, one would do the same for the N_1 values of θ_j and calculate the average. Graphically, the number of squares increases, as shown in figure 1 as we move from the upper to the lower graphs. And as $N_1, N_2 \to \infty$, the unit square is covered by more points and $\hat{\mathcal{I}} \to \mathcal{I}$. For higher dimensions, the intuition remains the same, and for each value of θ_j there is a hypercube composed of N_2^d points.

[20] This is not the only possibility, and wherever it is applicable, one could use a quadrature rule to approximate the integral.

Figure 1: Algorithm grid to approximate the integral



To show the performance of the algorithm in terms of speed and precision, I numerically approximate the integral $\mathcal{I} \equiv \int_{[0,1]^d} \prod_{j=1}^{d} \sqrt{u_j} dC(u_1, ..., u_d; \rho)$ with a *Clayton* (4) copula and dimension $d = \{2, 3\}$. This integral has no closed form solution. For the algorithm used in this paper I set $N_1 = N_2 = \{9, 19, 49, 99\}$. The number of draws in the Monte Carlo equals the number of points evaluated by the algorithm, *i.e.* $N_1 N_2^d$ draws, and the number of repetitions was 100. The sampling algorithm for the Monte Carlo is the one proposed by Marshall and Olkin (1988). Table 1 shows the mean value of both methods, the standard deviation of the Monte Carlo across repetitions (the algorithm proposed in this paper yields always the same result.), and the average time spent per repetition.

Even when the dimensionality of the integral is low, the algorithm proposed in this paper is several orders of magnitude faster than the traditional Monte Carlo, and the difference increases with the dimension of the integral. When the dimension is 2, for a given number of points, their performance is quite similar, and the approximation is within two standard deviations of the Monte Carlo. Even though the new algorithm consistently reports a number inferior to the mean across repetitions of the Monte Carlo, increasing the number of points at which the integral is evaluated is not as costly as for the Monte Carlo, resulting in a more accurate approximation for a given amount of computational time. When $d = 3$, the

Table 1: Implementation algorithm & Monte Carlo comparison

$N_1 = N_2$		9	19	49	99
			$d = 2$		
$\hat{\mathcal{I}}$	Algorithm	0.4933	0.4934	0.4936	0.4937
	Monte Carlo	0.4942	0.4944	0.4940	0.4940
S.D.	Algorithm	-	-	-	-
	Monte Carlo	0.0112	0.0033	0.0008	0.0003
Time	Algorithm	0.0001	0.0003	0.0011	0.0032
	Monte Carlo	0.0027	0.0036	0.0569	0.4776
			$d = 3$		
$\hat{\mathcal{I}}$	Algorithm	0.3786	0.3827	0.3854	0.3863
	Monte Carlo	0.3868	0.3873	0.3873	0.3873
S.D.	Algorithm	-	-	-	-
	Monte Carlo	0.0039	0.0007	0.0001	0.0000
Time	Algorithm	0.0001	0.0003	0.0011	0.0035
	Monte Carlo	0.0252	0.0828	3.3281	58.8875

Notes: $\hat{\mathcal{I}}$ is the approximated value of the integral for the proposed algorithm, and the mean value across repetitions for the Monte Carlo simulations.

algorithm I propose loses some accuracy with respect to the Monte Carlo, but the time gains are so large that by increasing the number of integration points the proposed algorithm outperforms the Monte Carlo. Table 2 shows the performance of the algorithm in high dimensions: its accuracy decreases with the dimensionality of the problem, as reflected in the changes in the approximation when we increase the number of points used to evaluate the integral. However, for given N_1 and N_2, computational time remains unchanged despite the increase of the dimensionality.

5 Monte Carlo

To show the finite sample performance of the estimator, I run a Monte Carlo with the following data generating process: $y_{ict} = \mathbf{1}\left(\eta_{ic} + \xi_t + x'_{it}\beta + \varepsilon_{ict} > 0\right)$, where ε_{ict} is logistically distributed, $\eta_c \sim \mathcal{N}\left(0, \sigma_0^2\right)$, $u_c \sim Clayton\left(\rho_0\right)$, $x_{ict} \sim U\left(0,1\right)$, $\xi = (-1.5, -1, -0.5, 0)$, $\beta_0 = 1$, $\sigma_0 = 1$, and $\rho_0 = 0.5$ for $i = 1, ..., 5$, $c = 1, ..., C$ for $C = 200, 400, 1000$, and $t = 1, 2, 3, 4$. The total number of repetitions is 1000. The main results are shown in

Table 2: Performance in high dimensions

$N_1 = N_2$	9	19	49	99	199
			$d = 2$		
$\hat{\mathcal{I}}$	0.4933	0.4934	0.4936	0.4937	0.4938
Time	0.0001	0.0003	0.0011	0.0032	0.0108
			$d = 3$		
$\hat{\mathcal{I}}$	0.3786	0.3827	0.3854	0.3863	0.3868
Time	0.0001	0.0003	0.0011	0.0035	0.0107
			$d = 5$		
$\hat{\mathcal{I}}$	0.2452	0.2534	0.2584	0.2602	0.2610
Time	0.0001	0.0003	0.0011	0.0035	0.0109
			$d = 10$		
$\hat{\mathcal{I}}$	0.1076	0.1176	0.1238	0.1260	0.1271
Time	0.0001	0.0003	0.0011	0.0033	0.0109
			$d = 50$		
$\hat{\mathcal{I}}$	0.0017	0.0033	0.0049	0.0057	0.0062
Time	0.0001	0.0003	0.0011	0.0033	0.0108

Notes: $\hat{\mathcal{I}}$ is the approximated value of the integral for the proposed algorithm.

table 3. Both the CBRE and the RE estimators consistently estimate the time fixed effects and β. The standard deviation of the individual effects is consistently estimated by the RE estimator, but the CBRE estimator of this parameter is slightly upward biased. This bias is reduced as the number of points used for the approximation increases, and it is of the same magnitude if the sample size varies. The choice of N_1 and N_2 also results in a substantial bias of the correlation parameter when the two numbers are different: if N_1 is larger than N_2, then the parameter is downward biased, whereas if N_1 is smaller than N_2, the bias is positive. Therefore, when choosing N_1 and N_2 one should be aware that if they are different, the correlation parameter will be biased, and if they are too small, also the standard deviation of the individual effects will be biased. In terms of efficiency, both estimators have a similar performance, with the CBRE estimator having slightly smaller standard errors, but the difference was of the order of the fourth digit.[21] Finally, the likelihood function for the RE estimator is always smaller than those of the CBRE estimator.

[21]Results available upon request.

Table 3: Monte Carlo Results

N	1000					2000					5000					
	CBRE				RE	CBRE				RE	CBRE				RE	θ_0
N_1	10	10	50	50		10	10	50	50		10	10	50	50		
N_2	10	50	10	50		10	50	10	50		10	50	10	50		
ξ_1	-1.51	-1.53	-1.49	-1.51	-1.51	-1.50	-1.52	-1.49	-1.50	-1.50	-1.51	-1.53	-1.49	-1.50	-1.50	-1.5
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	
ξ_2	-1.01	-1.03	-1.00	-1.01	-1.01	-1.01	-1.03	-0.99	-1.01	-1.00	-1.01	-1.02	-0.99	-1.00	-1.00	-1
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	
ξ_3	-0.51	-0.53	-0.49	-0.51	-0.50	-0.51	-0.52	-0.49	-0.50	-0.50	-0.51	-0.52	-0.49	-0.50	-0.50	-0.5
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	
ξ_4	0.00	-0.02	0.02	0.00	0.00	0.00	-0.02	0.02	0.00	0.00	-0.01	-0.03	0.01	0.00	0.00	0
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	
β	1.01	1.01	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	
σ	1.24	1.12	1.19	1.06	1.00	1.24	1.12	1.18	1.06	1.00	1.24	1.12	1.18	1.06	1.00	1
	(0.09)	(0.08)	(0.08)	(0.07)	(0.07)	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	
ρ	0.52	0.69	0.39	0.51	-	0.53	0.69	0.39	0.52	-	0.52	0.69	0.39	0.51	-	0.5
	(0.15)	(0.19)	(0.11)	(0.15)		(0.11)	(0.14)	(0.08)	(0.11)		(0.06)	(0.08)	(0.05)	(0.06)		
$\mathcal{L}$	-2541	-2541	-2541	-2540	-2556	-5087	-5086	-5086	-5086	-5117	-12723	-12723	-12722	-12722	-12797	

Notes: Mean estimates of the parameters across repetitions, standard deviations across repetitions in parentheses, $\mathcal{L}$ denotes the maximized value of the likelihood function, and θ_0 the true value of the parameters. The RE estimates were calculated by approximating the integral with a Gauss-Hermite quadrature with 20 points.

Finally, table 4 shows the results of the tests of independence (T1 and T3), and between the Clayton and Gumbel copulas (T2 and T4), both when the true copula is a Clayton(2) (T1 and T2), and an independent copula (T3 and T4), which nests both the Clayton and Gumbel copulas. When the true copula is a Clayton(2), the first test always rejects the null hypothesis of independence. The second test accepts the alternative hypothesis of the Clayton copula being a better fit than the Gumbel more often than not, and with a higher probability as the sample size increases. It also accepts the null hypothesis that both copulas provide an equally good fit, but it never accepts the hypothesis that the Gumbel is the best fit.[22] In the second experiment however, the null hypothesis of independence is almost always accepted by the first test, whereas the second test shows that both copulas provide a statistically equal fit.

Table 4: Monte Carlo Tests

N	1000				2000				5000			
Test	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
H_0 accepted	0	34.8	99.6	97.6	0	13.8	99.6	99.2	0	0	96.8	100
H_1 accepted	100	65.2	0.4	2.4	100	86.2	0.4	0.8	100	100	3.2	0
H_2 accepted		0		0		0		0		0		0

Notes: The numbers represent the percentage each hypothesis was accepted in the 500 repetitions of each experiment; 5% test size; $N_1 = N_2 = 50$.

6 Empirical Application

To illustrate the estimator, I study the existence of labor supply in married couples. I use the 2012 wave of the EU-SILC (European Union Statistics on Income and Living Conditions) dataset, which follows 279,115 individuals for the 2009-2012 period.[23] I keep the sub-population of married couples, in which both individuals were aged 21-65 during the whole period, leaving us with 28,246 individuals.

[22]Even when the true copula was the Clayton(2), the estimates of the slope parameters, *i.e.* the β and ξ_j for $j = 1, ..., 4$, were centered around their true values. Results available upon request.

[23]Other papers that have used the EU-SILC dataset to study labor supply are Bredtmann et al. (2014), Kalíšková (2015), and Schlenker (2015), though their approaches are different from the discrete choice model presented in this paper.

For these couples I run a CBRE logit regression in which the dependent variable takes value one if they worked during the year and zero otherwise, and the covariates included are gender, age, level of education, and total household non-labor income, on top of yearly dummies. I assume that the random effects are normally distributed and that the copula is a Clayton. The results are shown in table 6.[24] The estimates are qualitatively the same for all countries: married females work with a smaller probability than their husbands, the older workers become, the smaller the probability that they work, and increasing the level of education is correlated with an increase in the probability of working (with the exception of Greece, where workers with primary or no education have a larger probability of working than those with secondary education). On the other hand, the coefficient of non-labor income is not significant in all countries. Regarding the distribution of the individual effects, its standard deviation is always significantly different from zero and substantially large, implying that the probability of working at a period greatly varies across individuals, but the correlation inside the couple is always small: Denmark, the country for which the correlation is the highest has a coefficient of 0.55, which in terms of the linear correlation would be approximately 0.33.[25]

Using these estimates, I compute the probability that at least one member of the couple was employed in every period, to which I refer as a *working household*, and then I change the estimated copula by the independence copula, obtaining the counterfactual probability when there is neither sorting nor spillovers in the individual heterogeneity, presenting these estimates in table 7. Consistently with the large differences in the labor market across countries, the probability of observing a working household has a lot of variation: countries with a low unemployment rate, such as Denmark, the Netherlands, or Norway, have a high probability, whereas countries with high unemployment rate, such as Greece or Spain, have a low probability. In the counterfactual scenario, this probability would increase in most countries, reducing the proportion of non-working households. However, this increase would

[24] For completeness, the results for the standard RE logit estimator are shown in appendix E.

[25] A *Clayton* (0.65) has a Kendall τ statistic of 0.22, which is the value attained by a Gaussian copula with a linear correlation of 0.33.

Table 5: CBRE logit estimates

	AT	BE	BG	CZ	DK	EL	ES	FI	FR	HU	IT	NL	NO	PL	PT	UK
FE	-2.06	-3.67	-1.11	-1.75	-0.97	-5.03	-4.08	-0.94	-1.96	-1.07	-4.23	-5.15	-2.17	-2.87	-3.01	-1.06
	(0.49)	(0.81)	(0.48)	(0.36)	(0.35)	(0.58)	(0.30)	(0.39)	(0.29)	(0.31)	(0.30)	(0.67)	(0.34)	(0.38)	(0.52)	(0.50)
C5	-0.55	-0.23	-1.41	0.29	-1.30	0.78	0.09	-1.47	-0.40	0.69	0.94	-0.52	-0.70	0.58	-0.03	1.00
	(0.70)	(0.68)	(1.21)	(0.79)	(0.68)	(0.59)	(0.32)	(0.51)	(0.42)	(0.46)	(0.39)	(1.21)	(0.50)	(0.45)	(0.91)	(1.00)
C5*FE	-2.93	-1.43	-1.61	-7.99	0.66	-1.69	-1.02	-3.93	-3.20	-5.56	-2.69	-2.33	-0.79	-3.55	-0.99	-3.98
	(0.83)	(1.04)	(1.77)	(0.88)	(0.78)	(0.74)	(0.43)	(0.64)	(0.49)	(0.47)	(0.46)	(1.25)	(0.58)	(0.53)	(1.05)	(1.05)
AGE	-0.19	-0.34	-0.17	-0.22	-0.08	-0.18	-0.11	-0.14	-0.31	-0.17	-0.17	-0.31	-0.08	-0.26	-0.23	-0.10
	(0.03)	(0.05)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)	(0.03)
SE	0.92	1.33	1.66	2.60	0.31	-0.15	2.05	1.90	1.21	1.57	2.53	2.08	2.01	2.12	1.63	-0.03
	(0.41)	(0.44)	(0.59)	(0.61)	(0.43)	(0.30)	(0.29)	(0.50)	(0.25)	(0.29)	(0.28)	(0.62)	(0.39)	(0.66)	(0.47)	(0.38)
TE	1.68	2.42	4.28	4.06	1.89	3.61	3.93	2.70	3.86	3.26	4.67	4.06	3.24	5.60	4.21	0.76
	(0.61)	(0.48)	(0.75)	(0.74)	(0.49)	(0.54)	(0.31)	(0.56)	(0.37)	(0.42)	(0.44)	(0.61)	(0.42)	(0.77)	(0.75)	(0.41)
IN	0.00	-0.01	-0.23	0.05	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.03	0.00	0.11	0.03	-0.03
	(0.01)	(0.01)	(0.42)	(0.05)	(0.00)	(0.04)	(0.03)	(0.01)	(0.00)	(0.10)	(0.02)	(0.01)	(0.01)	(0.01)	(0.08)	(0.03)
$\hat{\sigma}$	4.72	7.43	5.17	4.58	3.47	5.73	4.75	4.00	6.39	4.95	4.77	6.59	4.08	6.09	5.57	4.56
	(0.36)	(0.80)	(0.38)	(0.25)	(0.29)	(0.41)	(0.21)	(0.29)	(0.27)	(0.26)	(0.20)	(0.49)	(0.26)	(0.32)	(0.40)	(0.40)
$\hat{\rho}$	0.39	0.47	0.22	0.28	0.55	0.07	0.17	0.25	0.29	0.39	0.12	0.10	0.19	0.15	0.24	0.19
	(0.19)	(0.25)	(0.18)	(0.15)	(0.26)	(0.16)	(0.11)	(0.16)	(0.09)	(0.14)	(0.11)	(0.12)	(0.14)	(0.12)	(0.17)	(0.19)
N	764	604	780	1404	696	930	1906	882	3370	1422	2186	1360	1364	1848	814	640
T	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

Notes: Standard errors in parentheses for all coefficients except for $\hat{\rho}$, for which I report the critical values at the 95% level. FE, C5, AGE, SE, TE, and IN respectively denote female, number of children smaller than 5 years old, age, secondary education, tertiary education, and non-labor income (expressed in thousands of euros); N is the sample size of each country; T is the total number of periods. The countries whose estimates are shown in this table are Austria, Belgium, Bulgaria, Czech Republic, Denmark, Greece, Spain, Finland, Hungary, Italy, Netherlands, Norway, Poland, Portugal, and the United Kingdom. The estimates for the remaining countries (Cyprus, Estonia, Iceland, Lithuania, Luxemburg, Latvia, Malta, and Slovenia) are available upon request.

be concentrated in those countries with an already high probability: the largest one would be in Denmark, followed by the United Kingdom and Austria. On the other hand, countries with a relatively low probability of having working households (Greece, Spain, Italy) would only have a slightly higher probability in the counterfactual scenario.

Then, I redo the same analysis computing the probability that a wife is employed conditional on whether the husband employed or unemployed.[26] The probability of a woman being employed is higher in every country whenever the husband is employed, and this difference ranges from 5% in Greece to 25% in Belgium. There is a positive correlation between this difference and the unconditional probability of the woman being employed. However, when the husband is unemployed, the probability of the woman being employed is positively correlated to the probability of observing a working household in every period, with countries like Belgium, Greece, Italy, and Spain having the lowest probability, and Denmark, the Netherlands, and Norway having the highest probability. If no such correlation in the unobserved propensity to work existed, then the probability of the wife being employed would lie in between the two probabilities, but closer to the conditional probability when the husband is employed (and the copula is not independent) than when he is unemployed.

Some remarks are in order. These results cannot be extrapolated to the whole population, as the characteristics of married couples in working age, both observable and unobservable, differ from those of singles. Moreover, the coefficient ρ does not have a causal interpretation in this example: marrying someone with a higher propensity to work does not imply that the own propensity is changed in any direction, nor marrying people at random would lead to the counterfactual scenario if there are such spillovers inside the marriage. The goal of this exercise is to isolate the contribution of the correlation in the unobserved propensity to work inside couples to the probability of having working households.

[26] Notice that for the case in which the copula is independent, the conditional probability is the same regardless of the working status of the husband, since conditional on the covariates, the joint probability is the product of the marginals.

Table 6: Counterfactuals

	AT	BE	BG	CZ	DK	EL	ES	FI	FR	HU	IT	NL	NO	PL	PT	UK
$\underline{P}_\rho$	80.4	71.0	77.8	86.1	89.8	65.2	68.9	82.9	76.0	66.7	67.8	90.6	92.2	67.5	69.0	86.3
	(2.6)	(3.9)	(3.0)	(1.6)	(2.0)	(4.4)	(2.6)	(2.5)	(1.5)	(2.7)	(2.7)	(1.5)	(1.3)	(2.7)	(3.7)	(2.7)
$\underline{P}_I$	83.5	72.6	79.8	88.3	93.8	65.2	69.2	85.2	77.7	67.6	67.9	91.6	94.2	67.9	69.9	88.7
	(3.5)	(5.1)	(3.6)	(2.1)	(2.9)	(4.6)	(2.9)	(3.0)	(1.8)	(3.4)	(2.9)	(1.6)	(1.6)	(3.0)	(4.3)	(3.3)
DIF	3.1	1.6	2.0	2.1	3.9	0.1	0.3	2.3	1.7	0.9	0.1	1.0	2.0	0.5	0.9	2.4
	(1.4)	(1.4)	(0.8)	(0.7)	(1.5)	(0.2)	(0.3)	(0.8)	(0.4)	(0.8)	(0.2)	(0.2)	(0.4)	(0.3)	(0.7)	(0.8)
$CP1_\rho$	78.1	73.7	82.1	77.8	91.8	57.2	58.9	79.5	77.9	70.4	59.4	79.8	85.7	66.0	69.4	80.9
	(2.6)	(3.6)	(2.2)	(2.2)	(1.3)	(4.1)	(2.8)	(2.5)	(1.3)	(2.5)	(2.6)	(1.9)	(1.5)	(2.6)	(3.5)	(2.5)
$CP0_\rho$	56.6	48.4	71.5	62.1	72.8	52.5	49.1	67.5	61.7	52.1	51.7	71.3	74.4	57.3	56.4	69.9
	(9.9)	(13.1)	(6.9)	(8.4)	(7.0)	(10.9)	(7.0)	(6.5)	(4.5)	(6.5)	(7.2)	(8.2)	(6.1)	(6.0)	(8.8)	(9.0)
CP_I	75.2	69.3	80.1	76.3	90.3	56.4	57.3	77.7	75.2	66.1	58.3	79.2	84.8	64.2	66.8	79.5
	(3.0)	(4.4)	(2.5)	(2.3)	(1.6)	(4.2)	(3.0)	(2.8)	(1.5)	(2.9)	(2.7)	(2.0)	(1.6)	(2.7)	(3.8)	(2.7)
N	764	604	780	1404	696	930	1906	882	3370	1422	2186	1360	1364	1848	814	640
T	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

Notes: Standard errors of the estimated probabilities were computed using the delta method. $\underline{P}_\rho$ and $\underline{P}_I$ respectively denote the probability (in %) that at least one member of the couple was employed in every period when the parameter of the copula is the estimated one and when the copula is independent, and DIF denotes the difference between the two of them; $CP1_\rho$, $CP0_\rho$, and CP_I respectively denote the probability that a wife is employed conditional on her husband being employed when the copula is the estimated one, conditional on her husband being unemployed when the copula is the estimated one, and with the independence copula; N is the sample size of each country; T is the total number of periods. The countries whose estimates are shown in this table are Austria, Belgium, Bulgaria, Czech Republic, Denmark, Greece, Spain, Finland, Hungary, Italy, Netherlands, Norway, Poland, Portugal, and the United Kingdom. The estimates for the remaining countries (Cyprus, Estonia, Iceland, Lithuania, Luxemburg, Latvia, Malta, and Slovenia) are available upon request.

7 Conclusion

In this paper I present the CBRE estimator for clustered data, a random effects estimator for binary choice panel data in which the unobserved heterogeneity of individuals in the same cluster is correlated. The correlation of the unobserved heterogeneity is modeled using a copula. This estimator allows to easily compute joint and conditional events, which are inherently difficult to model in a simultaneous equations framework with limited dependent variables. It is more efficient than standard random effects estimators, and unlike the latter, it can be used to consistently estimate the probability of joint and conditional events. I also propose two types of tests of hypothesis, one in which the null hypothesis is that the value of the correlation parameter lies on the boundary of the parameter space (usually for the independence copula), and another one along the lines of Vuong (1989) that discriminates between two non-nested models, both of which could be potentially misspecified.

The computation of the estimator requires the numerical approximation of potentially high-dimensional integrals. To overcome this issue, I propose an algorithm that approximates such integrals for Archimedean copulas. This algorithm does not suffer from the curse of dimensionality unlike traditional simulation methods, such as Monte Carlo integration.

I illustrate the use of the estimator with a Monte Carlo simulation, and an empirical application of labor supply of married couples. The results show evidence of some degree of correlation in the unobserved propensity to be employed between the two members of the couple. Then I use these estimates to compute the probability that the wife is employed, conditional on the employment status of the husband, and compare it to the counterfactual probability when the individual effects are independent. The findings suggest that the probability of the woman being employed is substantially larger when the husband is also employed.

References

Altonji, J. G. and R. L. Matzkin (2005). Cross section and panel data estimators for nonseparable models with endogenous regressors. *Econometrica 73*(4), 1053–1102.

Arellano, M. (2003). *Modelling optimal instrumental variables for dynamic panel data models.* CEMFI.

Arellano, M. and S. Bonhomme (2011). Nonlinear panel data analysis. *Annual Review of Economics 3*(1), 395–424.

Belloni, M. and R. Alessie (2013). Retirement choices in italy: what an option value model tells us. *Oxford Bulletin of Economics and Statistics 75*(4), 499–527.

Bonhomme, S. (2012). Functional differencing. *Econometrica 80*(4), 1337–1385.

Bredtmann, J., S. Otten, and C. Rulff (2014). Husband's unemployment and wife's labor supply–the added worker effect across europe. Technical report, Ruhr Economic Paper.

Brock, W. A. and S. N. Durlauf (2001). Interactions-based models. *Handbook of Econometrics 5*, 3297–3380.

Browning, M., J. Carro, et al. (2007). Heterogeneity and microeconometrics modeling. *ECONOMETRIC SOCIETY MONOGRAPHS 43*, 47.

Bruze, G. (2011). Marriage choices of movie stars: Does spouse's education matter? *Journal of Human Capital 5*(1), 1–28.

Cambanis, S., S. Huang, and G. Simons (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis 11*(3), 368–385.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies 47*(1), 225–238.

Chamberlain, G. (1984). Panel data. *Handbook of econometrics 2*, 1247–1318.

Chamberlain, G. (2010). Binary response models for panel data: Identification and information. *Econometrica 78*(1), 159–168.

Charles, K. K., E. Hurst, and A. Killewald (2013). Marital sorting and parental wealth. *Demography 50*(1), 51–70.

Chernozhukov, V., I. Fernández-Val, J. Hahn, and W. Newey (2013). Average and quantile effects in nonseparable panel models. *Econometrica 81*(2), 535–580.

De Paula, A. (2013). Econometric analysis of games with multiple equilibria. *Annu. Rev. Econ. 5*(1), 107–131.

Embrechts, P., F. Lindskog, and A. McNeil (2001). Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*.

Fernández-Val, I. (2009). Fixed effects estimation of structural parameters and marginal effects in panel probit models. *Journal of Econometrics 150*(1), 71–85.

Fernández-Val, I. and J. Lee (2013). Panel data models with nonadditive unobserved heterogeneity: Estimation and inference. *Quantitative Economics 4*(3), 453–481.

Fox, J. T., K. il Kim, S. P. Ryan, and P. Bajari (2012). The random coefficients logit model is identified. *Journal of Econometrics 166*(2), 204–212.

Gautier, E. and Y. Kitamura (2013). Nonparametric estimation in random coefficients binary choice models. *Econometrica 81*(2), 581–607.

Graham, B. S. and J. L. Powell (2012). Identification and estimation of average partial effects in "irregular" correlated random coefficient panel data models. *Econometrica 80*(5), 2105–2152.

Hahn, J. and W. Newey (2004). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica 72*(4), 1295–1319.

Hajivassiliou, V., D. McFadden, and P. Ruud (1996). Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results. *Journal of econometrics 72*(1), 85–134.

Hanushek, E. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *The American Economic Review 61*(2), 280–288.

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica 46*(4), 931–959.

Heiss, F. and V. Winschel (2008). Likelihood approximation by numerical integration on sparse grids. *Journal of Econometrics 144*(1), 62–80.

Kalíšková, K. (2015). Tax and transfer policies and the female labor supply in the eu. Technical report, IZA Discussion Paper.

Kaya, E. (2014). Heterogeneous couples, household interactions and labor supply elasticities of married women. Technical report, Cardiff University, Cardiff Business School, Economics Section.

Lee, M.-J. (1999). A root-n consistent semiparametric estimator for related-effect binary response panel data. *Econometrica 67*(2), 427–433.

Maddala, G. and L.-F. Lee (1976). Recursive models with qualitative endogenous variables. In *Annals of Economic and Social Measurement, Volume 5, number 4*, pp. 525–545. NBER.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics 3*(3), 205–228.

Manski, C. F. (1987). Semiparametric analysis of random effects linear models from binary panel data. *Econometrica: Journal of the Econometric Society 55*(2), 357–362.

Marshall, A. W. and I. Olkin (1988). Families of multivariate distributions. *Journal of the American Statistical Association 83*(403), 834–841.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics 4*, 2111–2245.

Pereda-Fernández, S. (2015). Teachers and cheaters. just an anagram? Unpublished working paper.

Prokhorov, A. and P. Schmidt (2009). Likelihood-based estimation in a panel setting: robustness, redundancy and validity of copulas. *Journal of Econometrics 153*(1), 93–104.

Schlenker, E. (2015). The labour supply of women in stem. *IZA Journal of European Labor Studies 4*(1), 1–17.

Schmidt, P. (1981). Constraints on the parameters in simultaneous tobit and probit models. *Structural analysis of discrete data with econometric applications 1*, 423–434.

Self, S. G. and K.-Y. Liang (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association 82*(398), 605–610.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Istitut de Statistique de l'Universitè de Paris 8*, 229–231.

Skrainka, B. S. and K. L. Judd (2011). High performance quadrature rules: How numerical integration affects a popular model of product differentiation. Technical report, SSRN.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society 57*(2), 307–333.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Appendix

A Mathematical proofs

A.1 Proof to proposition 1

The proposition is shown by checking that assumptions 1 to 11 satisfy the assumptions in theorems 2.5 and 3.3 in Newey and McFadden (1994). Rather than considering the data *iid* at the individual level, I do it at the cluster level. Begin with the consistency result.

By assumptions 10 and 11,

$$\ell_c(z_c;\theta) = \int_{[0,1]}^{N_c} \prod_{i=1}^{N_c} P_{ic}(z_{ic},\eta_{ic};\beta)\, c(u_c;\rho)\, du_{ic} < \bar{c} \max_{i=1,\ldots,N_c} \ell_{ic}(z_{ic};\theta)^{\overline{N}} < \infty$$

So $\ell_c(z_{ci};\theta)$ is well defined and finite. By assumption 4, for any sequence $\theta_n : \theta_n \to \theta$, $\prod_{i=1}^{N_c} P_{ic}(z_{ic},\eta_{ic,n};\beta_n)\, c(u_c;\rho_n) \to \prod_{i=1}^{N_c} P_{ic}(z_{ic},\eta_{ic};\beta)\, c(u_c;\rho)$ for almost every u_c. Therefore, by the dominated convergence theorem, $\ell_c(z_c;\theta_n) \to \ell_c(z_c;\theta)$, so $\ell_c(z_c;\theta)$ is continuous with respect to θ.

$$\log(\ell_c(z_c;\theta)) = \log\left(\int_{[0,1]}^{N_c} \prod_{i=1}^{N_c} P_{ic}(z_{ic},\eta_{ic};\beta)\, c(u_c;\rho)\, du_{ic}\right) < \log(\bar{c}) + \sum_{i=1}^{N_c} \log(\ell_{ic}(z_{ic};\mu))$$

where the inequality follows from assumption 10. By a similar argument, $\log(\ell_c(z_c;\theta)) > \log(\underline{c}) + \sum_{i=1}^{N_c} \log(\ell_{ic}(z_{ic};\mu))$. Hence, by assumptions 5 and 11,

$$\mathbb{E}\left[\sup_{\theta\in\Theta} |\log(\ell_c(z_c;\theta))|\right] < \max\{|\log(\underline{c})|, |\log(\bar{c})|\} + \overline{N}\mathbb{E}\left[\max_{i=1,\ldots,N_c} \sup_{\theta\in\Theta} |\log(\ell_{ic}(z_{ic};\mu))|\right] < \infty$$

These two results, together with assumptions 1 to 3, verify the conditions in theorem 2.5 in Newey and McFadden (1994) and hence $\hat{\theta} \xrightarrow{P} \theta_0$. Note that it would be further possible to relax assumption 3 to allow θ_0 to be on the boundary of Θ and still get consistency.

By assumptions 7, 10, and 11, $\|\nabla_\mu \ell_c(z_c;\theta)\| < \bar{c}\overline{N} \max_{i=1,\ldots,N_c} \prod_{j\neq i} \ell_{jc}(z_{jc};\mu)\, \|(\nabla_\mu \ell_{ic}(z_{ic};\mu))\|$ for all $\theta \in \mathcal{N}$. Hence,

$$\int \sup_{\theta\in\mathcal{N}} \|\nabla_\mu \ell_c(z_c;\theta)\|\, dz_c < \bar{c}\overline{N} \max_{i=1,\ldots,N_c} \prod_{j\neq i} \int \sup_{\theta\in\mathcal{N}} \ell_{jc}(z_{jc};\mu)\, dz_{jc} \int \sup_{\theta\in\mathcal{N}} \|\nabla_\mu \ell_{ic}(z_{ic};\mu)\|\, dz_{ic} < \infty$$

By assumptions 10 and 11, $\|\nabla_\rho \ell_c(z_c;\theta)\| < \|\bar{c}_1\| \prod_{i=1}^{N_c} \ell_{ic}(z_{ic};\mu)$ for all $\theta \in \mathcal{N}$. Hence,

$$\int \sup_{\theta\in\mathcal{N}} \|\nabla_\rho \ell_c(z_c;\theta)\| \, dz_c < \|\bar{c}_1\| \prod_{i=1}^{N_c} \int \sup_{\theta\in\mathcal{N}} |\ell_{ic}(z_{ic};\mu)| \, dz_{ic} < \infty$$

By a parallel argument, the second derivatives can be bounded. Consequently, it follows that $\int \sup_{\theta\in\mathcal{N}} \|\nabla_\theta \ell_c(z_c;\theta)\| \, dz_c < \infty$ and $\int \sup_{\theta\in\mathcal{N}} \|\nabla_{\theta\theta} \ell_c(z_c;\theta)\| \, dz_c < \infty$. This, together with assumptions 1, 3, 6, and 8, and 10, the conditions in theorem 3.3 in Newey and McFadden (1994) are verified and $\sqrt{C}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_\theta^{CBRE}\right)$.

A.2 Proof to proposition 2

The asymptotic variance of the CBRE estimator $\hat{\theta}$ is given by

$$\Sigma_\theta^{CBRE} = -\mathbb{E}\begin{bmatrix} \nabla_{\mu\mu} \log(\ell_c(z_c;\theta_0)) & \nabla_{\mu\rho} \log(\ell_c(z_c;\theta_0)) \\ \nabla_{\mu\rho} \log(\ell_c(z_c;\theta_0))' & \nabla_{\rho\rho} \log(\ell_c(z_c;\theta_0)) \end{bmatrix}$$

Using the inverse block matrix formula, the variance of $\hat{\theta}$ is given by

$$\begin{aligned}\Sigma_\mu^{CBRE} &= -\left(\mathbb{E}\left[\nabla_{\mu\mu} \log(\ell_c(z_c;\theta_0))\right]\right. \\ &\quad \left. - \mathbb{E}\left[\nabla_{\mu\rho} \log(\ell_c(z_c;\theta_0))\right] \mathbb{E}\left[\nabla_{\rho\rho} \log(\ell_c(z_c;\theta_0))\right]^{-1} \mathbb{E}\left[\nabla_{\mu\rho} \log(\ell_c(z_c;\theta_0))\right]'\right)^{-1}\end{aligned}$$

Similarly, for the RE estimator, its asymptotic variance is given by

$$\Sigma_\mu^{RE} = -\mathbb{E}\left[\sum_{i=1}^{N_c} \nabla_{\mu\mu} \log(\ell_{ic}(z_{ic};\mu_0))\right]^{-1}$$

A.3 Proof to corollary 1

$$
\begin{aligned}
&-\mathbb{E}\left[\nabla_{\mu\mu}\sum_{i=1}^{N_c}\log\left(\ell_{ic}\left(z_{ic};\mu_0\right)\right)\right]^{-1} = -\mathbb{E}\left[\nabla_{\mu\mu}\log\left(\ell_c\left(z_c;\mu_0\right)\right)\right]^{-1}\\
&\Leftrightarrow \mathbb{E}\left[\nabla_{\mu\mu}\sum_{i=1}^{N_c}\log\left(\ell_{ic}\left(z_{ic};\mu_0\right)\right)-\nabla_{\mu\mu}\log\left(\ell_c\left(z_c;\mu_0\right)\right)\right]=0\\
&\Leftrightarrow \mathbb{E}\left[\int\left(\nabla_{\mu\mu}P_c\left(z_c,\eta_c;\beta_0\right)-\nabla_{\mu}P_c\left(z_c,\eta_c;\beta_0\right)\nabla_{\mu}P_c\left(z_c,\eta_c;\beta_0\right)'\right)\cdot\right.\\
&\left.\left(\frac{1}{\int P_c\left(z_c,\eta_c;\beta_0\right)\prod_{j=1}^{N_c}du_{jc}}-\frac{c\left(u_c;\rho_0\right)}{\int P_c\left(z_c,\eta_c;\beta_0\right)c\left(u_c;\rho_0\right)\prod_{j=1}^{N_c}du_{jc}}\right)\prod_{i=1}^{N_c}du_{ic}\right]=0\\
&\Leftrightarrow c\left(u_c;\rho_0\right)=1\\
&\Leftrightarrow C\left(u_c;\rho_0\right)=\prod_{i=1}^{N_c}u_{ic}
\end{aligned}
$$

A.4 Extra lemma

Lemma 1. *Assume that the distribution of X_{ict} is discrete with finite support, and let $\mathbb{P}\left(Y_c|X_c\right)=\int_{\mathcal{Y}}P_c\left(\eta_c;\beta\right)dC\left(u_c;\rho\right)$, where $P_c\left(\eta_c;\beta\right)$ is defined as in the main text and is a measurable function of η_c for each $\beta\in B$, and $\mathcal{Y}$ denotes the support of η_c. Then, for each β, every marginal distribution $F_\eta\left(\eta_{ic};\sigma\right)$ on the support of η_{ic}, and every copula $C\left(u_c;\rho\right)$ on $[0,1]^N$, there exists a discrete distribution $F_\eta^{k,N,T}$ with no more than 2^{NT} support points such that $\int_{\mathcal{Y}}P_c\left(z_c,\eta_c;\beta\right)dC\left(u_c;\rho\right)=\int_{\mathcal{Y}}P_c\left(z_c,\eta_c;\beta\right)dF_\eta^{k,N,T}\left(\eta_c\right)$.*

Proof. The definition of the copula implies the existence of a multivariate cdf F_η such that $C\left(u_c;\rho\right)=F_\eta\left(\eta_c;\sigma,\rho\right)$. For each $k=1,...,K$ of the possible values that the vector $\left(X_{11},...,X_{NT}\right)$ can take, there are $J=2^{NT}$ distinct values that the vector $\left(Y_{11},...,Y_{NT}\right)$ can take. Apply lemma 7 in Chernozhukov et al. (2013) to $\int_{\mathcal{Y}}P_c\left(z_c,\eta_c;\beta\right)dF_\eta\left(\eta_c;\sigma,\rho\right)$ to obtain the desired result. □

B Score and Hessian

Let F_{ict} and f_{ict} be shorthand for $F_{\varepsilon}\left(-\left(\eta_{ic}+x_{ict}^{\prime}\beta\right)\right)$ and $f_{\varepsilon}\left(-\left(\eta_{ic}+x_{ict}^{\prime}\beta\right)\right)$, denote the quantile function of η_{ic} by $Q_{\eta}\left(u\right)\equiv F_{\eta}^{-1}\left(u\right)$ and by $q_{\eta}\left(u;\sigma\right)$ its derivative with respect to σ. Then, the score is given by [27]

$$\frac{\partial\mathcal{L}\left(\theta\right)}{\partial\beta}=\sum_{c=1}^{N_c}\frac{\int_{[0,1]^{N_c}}P_c\left(z_c,\eta_c;\beta\right)\sum_{i=1}^{N_c}\sum_{t=1}^{T}\frac{f_{ict}}{F_{ict}(1-F_{ict})}\left(y_{ict}-\left(1-F_{ict}\right)\right)x_{ict}dC\left(u_c;\rho\right)}{\int_{[0,1]^{N_c}}P_c\left(z_c,\eta_c;\beta\right)dC\left(u_c;\rho\right)} \tag{8}$$

$$\frac{\partial\mathcal{L}\left(\theta\right)}{\partial\sigma}=\sum_{c=1}^{N_c}\frac{\int_{[0,1]^{N_c}}P_c\left(z_c,\eta_c;\beta\right)\sum_{i=1}^{N_c}\sum_{t=1}^{T}\frac{f_{ict}}{F_{ict}(1-F_{ict})}\left(y_{ict}-\left(1-F_{ict}\right)\right)q_{\eta}\left(u_{ict};\sigma\right)dC\left(u_c;\rho\right)}{\int_{[0,1]^{N_c}}P_c\left(z_c,\eta_c;\beta\right)dC\left(u_c;\rho\right)} \tag{9}$$

$$\frac{\partial\mathcal{L}\left(\theta\right)}{\partial\rho}=\sum_{c=1}^{N_c}\frac{\int_{[0,1]^{N_c}}P_c\left(z_c,\eta_c;\beta\right)\frac{\partial^{N_c+1}c(u_c;\rho)}{\prod_{j=1}^{N_c}\partial u_{jc}\partial\rho}\prod_{i=1}^{N_c}du_{ic}}{\int_{[0,1]^{N_c}}P_c\left(z_c,\eta_c;\beta\right)dC\left(u_c;\rho\right)} \tag{10}$$

It is immediate to approximate equations 8 and 9 using the proposed algorithm presented in this paper. Regarding equation 10, it is more convenient to numerically evaluate the derivative, *i.e.* $\frac{\partial\mathcal{L}(\theta)}{\partial\rho}\approx\frac{\mathcal{L}(\mu,\rho+\varepsilon)-\mathcal{L}(\mu,\rho)}{\epsilon}$. This approximation works reasonably well for values of the parameters such that the copula is not independent or close to independent, but is numerically unstable when the copula is approximately independent. Consequently, it is convenient to bound the parameter space used in the estimation. For example, for the Clayton copula $\rho\geq 0.0001$, and for the Gumbel copula $\rho>1.01$. Finally, the Hessian is easily computed by

$$\hat{H}\left(\hat{\theta},\right)=\frac{1}{C}\sum_{c=1}^{C}\frac{\partial\log\left(\hat{\ell}_c\left(z_c;\hat{\theta}\right)\right)}{\partial\theta^{\prime}}\frac{\partial\log\left(\hat{\ell}_c\left(z_c;\hat{\theta}\right)\right)}{\partial\left(\theta^{\prime}\right)}$$

[27] If η_{ic} belongs to a scale family of distributions, *i.e.* if $\eta_{ic}=\sigma\tilde{\eta}_{ic}$, where $\tilde{\eta}_{ic}\sim F_{\eta}\left(1\right)$, then $Q_{\eta}\left(u_{ic}\right)=\sigma\tilde{\eta}_{ic}$, and thus $q_{\eta}\left(u_{ic};\sigma\right)$ simplifies to $q_{\eta}\left(u_{ic};1\right)=\tilde{\eta}_{ic}$.

C Archimedean Copulas

The following table shows for the generator function and the sampler for several of the most popular Archimedean copulas. *Geo* denotes a geometric distribution with pmf $\theta(1-\theta)^k$ for $k \in \mathbb{N}$, Γ denotes a gamma distribution with pdf $x^{\frac{1}{\theta}-1}\frac{\exp(-x)}{\Gamma\left(\frac{1}{\theta}\right)}$ for $x \in \mathbb{R}_{++}$, *Log* a logarithmic distribution with pmf $\frac{\theta^k}{-k\log(1-\theta)}$ for $k \in \mathbb{N}$, *St* denotes a Stable distribution with characteristic function $\exp\left[it\mathbf{1}(\theta=1) - \left|\cos\left(\frac{\pi}{2\theta}\right)t^{\frac{1}{\theta}}\right|(1 - isgn(t)\Phi)\right]$ for $\Phi = \tan\left(\frac{\pi}{2\theta}\right)$ if $\theta \neq 1$, and $\Phi = -\frac{2}{\pi}\log|t|$ if $\theta = 1$, and *Sib* denotes a Sibuya distribution with pmf $\binom{\frac{1}{\theta}}{k}(-1)^{k-1}$ for $k \in \mathbb{N}$. Figure 2 shows the pdf for some of these copulas, and it compares them to the grid of points used by the algorithm to approximate the likelihood integrals.

Table 7: Generator and Sampler of Archimedean Copulas

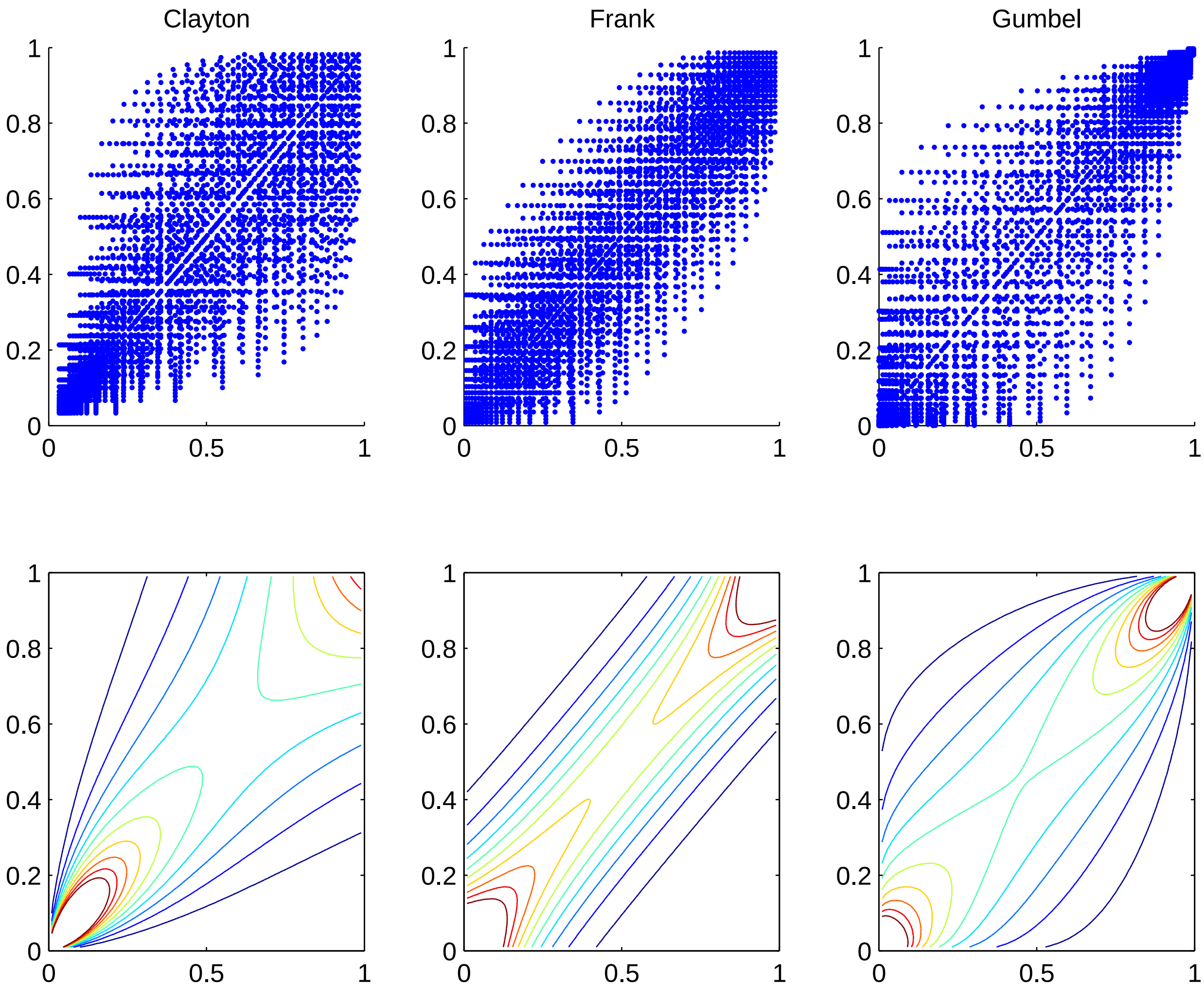
Copula	Θ	$\phi(t)$	$G(\theta)$
Ali-Mikhail-Haq	$[0,1]$	$\frac{1-\theta}{\exp(t)-\theta}$	$Geo(1-\theta)$
Clayton	$(0,\infty)$	$(1+t)^{\frac{1}{\theta}}$	$\Gamma\left(\frac{1}{\theta},1\right)$
Frank	$(0,\infty)$	$-\frac{1}{\theta}\log\left(1-\left(1-e^{-\theta}\right)e^{-t}\right)$	$Log(1-\exp(-\theta))$
Gumbel	$[1,\infty]$	$\exp\left(-t^{\frac{1}{\theta}}\right)$	$St\left(\frac{1}{\theta},1,\cos^{\theta}\left(\frac{\pi}{2\theta}\right),\mathbf{1}(\theta=1);1\right)$
Joe	$[1,\infty]$	$1-(1-\exp(-t))^{\frac{1}{\theta}}$	$Sib\left(\frac{1}{\theta}\right)$

D Elliptical copulas

The algorithm presented in section 4 is designed for Archimedean copulas. Elliptical copulas constitute another major parametric family, including two of the most widely used copulas, the Gaussian and the t.[28] In this section I show that, even though it is not possible to apply the proposed algorithm for these two copulas, it is possible to use Heiss and Winschel (2008) approximation using sparse grids to compute the likelihood which, despite being increasingly slower to compute as the dimension of the integral increases, does not suffer so heavily from the curse of dimensionality like other methods.

[28] See Cambanis et al. (1981) or Embrechts et al. (2001) for the definition of elliptical distributions and copulas, and thorough discussions of their properties.

Figure 2: Copula contour and approximation



Let R denote the linear correlation matrix of a d-variate normal distribution and Φ_R its cdf. The Gaussian copula with correlation R is given by

$$C(u;R) = \Phi_R\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right)$$

If R is positive definite, then it is possible to obtain the Cholesky decomposition, denoted by A. As shown by Embrechts et al. (2001), it is possible to express the Gaussian copula in terms of d independent normal distributions and the coefficients of A, where the (i,j) element is denoted by a_{ij}. Hence, it is possible to rewrite the integral that is required to evaluate the likelihood as

$$\begin{aligned}\mathcal{I} &= \int_{[0,1]^d} \prod_{i=1}^{d} \ell_i(u_i)\, dC(u;R) &(11)\\ &= \int_{[0,1]^d} \prod_{i=1}^{d} \ell_i\left(\Phi\left(\sum_{j=1}^{i} a_{ji}\Phi^{-1}(v_j)\right)\right) \prod_{j=1}^{d} dv_j &(12)\end{aligned}$$

Thus, it is not possible to reduce the dimensionality of this integral as it was done when the copula was Archimedean. Notice however that the likelihood is decomposed into d independent random variables, meaning that one can use Heiss and Winschel (2008) approximation, and the correlation structure is general, which is a richer structure than the one implied by standard Archimedean copulas.

A similar reformulation of the integral for the t copula is possible: denote by $t_{\nu,R}$ the cdf of the d-variate t distribution with ν degrees of freedom and correlation matrix R, then the t copula is given by[29]

$$C(u;\nu,R) = t_{\nu,R}\left(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d)\right)$$

Again, if R is positive definite, and following Embrechts et al. (2001), the copula can be written in terms of d independent normal variables and a χ^2 with ν degrees of freedom,

[29] Generalization of the previous algorithm to other elliptical copulas is conceptually straightforward, as it only requires changing the t distribution by another appropriately chosen continuous distribution on $\mathbb{R}^+$. See Cambanis et al. (1981) for details.

whose cdf I denote by F_ν. The integral $\mathcal{I}$ is then given by

$$\mathcal{I} = \int_{[0,1]^d} \prod_{i=1}^{d} \ell_i (u_i)\, dC(u; \nu, R) \tag{13}$$

$$= \int_{[0,1]^{d+1}} \prod_{i=1}^{d} \ell_i \left(t_\nu \left(\frac{\sqrt{\nu}}{\sqrt{F_\nu^{-1}(w)}} \sum_{j=1}^{i} a_{ji} \Phi^{-1}(v_j) \right) \right) \prod_{j=1}^{d} dv_j dw \tag{14}$$

With respect to the Gaussian copula, the only remarkable difference is the inclusion of the χ^2, which results in an increase of the dimension of the integral from d to $d+1$, but it is still possible to use Heiss and Winschel (2008) approximation.

If one were willing to adopt a symmetric correlation among the elements of the copula, $i.e$ if all the off-diagonal elements of R were equal to ρ, then it would be possible to obtain a reduction of the dimensionality of the integral similar to the one attained for the Archimedean copulas. To see this, notice that by the properties of the normal distribution, it is possible to obtain a d-variate normal distribution with covariance function $R = (1-\rho) I_d + \rho \iota_d \iota_d'$, where ι_d is a vector of ones, if each element is the sum of two independent random normals, one specific to each dimension, and one common to all, with weights $\sqrt{1-\rho}$ and $\sqrt{\rho}$. Hence, when the copula is Gaussian, the integral $\mathcal{I}$ can be rewritten as

$$\mathcal{I} = \int_0^1 \prod_{i=1}^{d} \left[\int_0^1 \ell_i \left(\Phi \left(\sqrt{\rho} \Phi^{-1}(z) + \sqrt{1-\rho} \Phi^{-1}(v_i) \right) \right) dv_i \right] dz$$

For the t copula a similar decomposition is feasible, but the dimensionality of the resulting integral is 3, because of the χ^2 distribution:

$$\mathcal{I} = \int_{[0,1]^2} \prod_{i=1}^{d} \left[\int_0^1 \ell_i \left(t_\nu \left(\frac{\sqrt{\nu}}{\sqrt{F_\nu^{-1}(w)}} \left(\sqrt{\rho} \Phi^{-1}(z) + \sqrt{1-\rho} \Phi^{-1}(v_i) \right) \right) \right) dv_i \right] dz dw$$

E Extra Results

Table 8: RE logit estimates

	AT	BE	BG	CZ	DK	EL	ES	FI	FR	HU	IT	NL	NO	PL	PT	UK
FE	-2.24	-1.09	-0.75	-1.75	-0.93	-4.48	-3.81	-1.06	-2.01	-0.99	-3.85	-5.08	-2.21	-1.97	-2.70	-1.11
	(0.43)	(0.39)	(0.27)	(0.33)	(0.39)	(0.37)	(0.23)	(0.35)	(0.19)	(0.25)	(0.22)	(0.44)	(0.33)	(0.20)	(0.30)	(0.41)
C5	-1.01	0.41	-1.20	0.19	-1.62	0.54	-0.15	-1.59	-1.03	0.08	0.59	0.44	-0.72	-0.49	-0.75	1.82
	(0.67)	(0.58)	(0.47)	(0.76)	(0.70)	(0.41)	(0.27)	(0.51)	(0.31)	(0.36)	(0.31)	(1.16)	(0.48)	(0.26)	(0.63)	(0.93)
C5*FE	-2.72	-2.07	-1.45	-7.30	0.85	-0.94	-0.68	-3.74	-2.24	-4.03	-2.16	-3.60	-0.82	-2.34	0.30	-4.49
	(0.77)	(0.66)	(0.46)	(0.81)	(0.91)	(0.44)	(0.34)	(0.65)	(0.36)	(0.43)	(0.34)	(1.18)	(0.57)	(0.28)	(0.76)	(1.01)
AGE	-0.21	-0.18	-0.14	-0.20	-0.09	-0.18	-0.12	-0.15	-0.31	-0.15	-0.13	-0.34	-0.08	-0.23	-0.19	-0.08
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)
SE	0.91	2.76	2.10	2.82	0.58	-0.79	1.58	1.97	1.68	1.67	2.48	2.39	2.00	3.01	1.35	0.26
	(0.37)	(0.31)	(0.37)	(0.47)	(0.43)	(0.22)	(0.23)	(0.46)	(0.18)	(0.24)	(0.19)	(0.37)	(0.36)	(0.30)	(0.36)	(0.34)
TE	1.66	4.07	4.51	4.33	2.36	3.10	3.57	2.79	3.90	3.25	4.52	4.20	3.39	5.54	3.98	0.88
	(0.51)	(0.35)	(0.48)	(0.59)	(0.50)	(0.37)	(0.24)	(0.50)	(0.23)	(0.33)	(0.33)	(0.43)	(0.38)	(0.40)	(0.60)	(0.37)
IN	0.00	-0.02	0.02	0.05	0.00	0.03	0.01	0.01	0.00	-0.02	0.02	0.02	0.00	-0.01	-0.08	-0.02
	(0.01)	(0.01)	(0.35)	(0.05)	(0.00)	(0.02)	(0.02)	(0.01)	(0.00)	(0.09)	(0.02)	(0.01)	(0.00)	(0.03)	(0.08)	(0.03)
$\hat{\sigma}$	4.21	6.38	4.35	3.87	3.25	5.43	4.21	3.86	5.51	4.35	4.32	5.97	3.73	5.28	4.63	4.04
	(0.41)	(0.56)	(0.35)	(0.27)	(0.39)	(0.42)	(0.23)	(0.35)	(0.24)	(0.27)	(0.23)	(0.49)	(0.32)	(0.28)	(0.36)	(0.44)
N	764	604	780	1404	696	930	1906	882	3370	1422	2186	1360	1364	1848	814	640
T	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

Notes: Standard errors in parentheses. FE, C5, AGE, SE, TE, and IN respectively denote female, number of children smaller than 5 years old, age, secondary education, tertiary education, and non-labor income (expressed in thousands of euros); N is the sample size of each country; T is the total number of periods. The countries whose estimates are shown in this table are Austria, Belgium, Bulgaria, Czech Republic, Denmark, Greece, Spain, Finland, Hungary, Italy, Netherlands, Norway, Poland, Portugal, and the United Kingdom. The estimates for the remaining countries (Cyprus, Estonia, Iceland, Lithuania, Luxemburg, Latvia, Malta, and Slovenia) are available upon request.

RECENTLY PUBLISHED "TEMI" (*)

N. 1068 – *The labor market channel of macroeconomic uncertainty*, by Elisa Guglielminetti (June 2016).

N. 1069 – *Individual trust: does quality of public services matter?*, by Silvia Camussi and Anna Laura Mancini (June 2016).

N. 1070 – *Some reflections on the social welfare bases of the measurement of global income inequality,* by Andrea Brandolini and Francesca Carta (July 2016).

N. 1071 – *Boulevard of broken dreams. The end of the EU funding (1997: Abruzzi, Italy),* by Guglielmo Barone, Francesco David and Guido de Blasio (July 2016).

N. 1072 – *Bank quality, judicial efficiency and borrower runs: loan repayment delays in Italy,* by Fabio Schiantarelli, Massimiliano Stacchini and Philip Strahan (July 2016).

N. 1073 – *Search costs and the severity of adverse selection,* by Francesco Palazzo (July 2016).

N. 1074 – *Macroeconomic effectiveness of non-standard monetary policy and early exit. A model-based evaluation*, by Lorenzo Burlon, Andrea Gerali, Alessandro Notarpietro and Massimiliano Pisani (July 2016).

N. 1075 – *Quantifying the productivity effects of global sourcing,* by Sara Formai and Filippo Vergara Caffarelli (July 2016).

N. 1076 – *Intergovernmental transfers and expenditure arrears,* by Paolo Chiades, Luciano Greco, Vanni Mengotto, Luigi Moretti and Paola Valbonesi (July 2016).

N. 1077 – *A "reverse Robin Hood"? The distributional implications of non-standard monetary policy for Italian households*, by Marco Casiraghi, Eugenio Gaiotti, Lisa Rodano and Alessandro Secchi (July 2016).

N. 1078 – *Global macroeconomic effects of exiting from unconventional monetary policy,* by Pietro Cova, Patrizio Pagano and Massimiliano Pisani (September 2016).

N. 1079 – *Parents, schools and human capital differences across countries*, by Marta De Philippis and Federico Rossi (September 2016).

N. 1080 – *Self-fulfilling deflations*, by Roberto Piazza, (September 2016).

N. 1081 – *Dealing with student heterogeneity: curriculum implementation strategies and student achievement*, by Rosario Maria Ballatore and Paolo Sestito, (September 2016).

N. 1082 – *Price dispersion and consumer inattention: evidence from the market of bank accounts*, by Nicola Branzoli, (September 2016).

N. 1083 – *BTP futures and cash relationships: a high frequency data analysis*, by Onofrio Panzarino, Francesco Potente and Alfonso Puorro, (September 2016).

N. 1084 – *Women at work: the impact of welfare and fiscal policies in a dynamic labor supply model*, by Maria Rosaria Marino, Marzia Romanelli and Martino Tasso, (September 2016).

N. 1085 – *Foreign ownership and performance: evidence from a panel of Italian firms*, by Chiara Bentivogli and Litterio Mirenda (October 2016).

N. 1086 – *Should I stay or should I go? Firms' mobility across banks in the aftermath of financial turmoil*, by Davide Arnaudo, Giacinto Micucci, Massimiliano Rigon and Paola Rossi (October 2016).

N. 1087 – *Housing and credit markets in Italy in times of crisis*, by Michele Loberto and Francesco Zollino (October 2016).

N. 1088 – *Search peer monitoring via loss mutualization*, by Francesco Palazzo (October 2016).

N. 1089 – *Non-standard monetary policy, asset prices and macroprudential policy in a monetary union*, by Lorenzo Burlon, Andrea Gerali, Alessandro Notarpietro and Massimiliano Pisani (October 2016).

N. 1090 – *Does credit scoring improve the selection of borrowers and credit quality?*, by Giorgio Albareto, Roberto Felici and Enrico Sette (October 2016).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

"TEMI" LATER PUBLISHED ELSEWHERE

2014

G. M. TOMAT, *Revisiting poverty and welfare dominance*, Economia pubblica, v. 44, 2, 125-149, **TD No. 651 (December 2007).**

M. TABOGA, *The riskiness of corporate bonds*, Journal of Money, Credit and Banking, v.46, 4, pp. 693-713, **TD No. 730 (October 2009).**

G. MICUCCI and P. ROSSI, *Il ruolo delle tecnologie di prestito nella ristrutturazione dei debiti delle imprese in crisi*, in A. Zazzaro (a cura di), Le banche e il credito alle imprese durante la crisi, Bologna, Il Mulino, **TD No. 763 (June 2010).**

F. D'AMURI, *Gli effetti della legge 133/2008 sulle assenze per malattia nel settore pubblico,* Rivista di politica economica, v. 105, 1, pp. 301-321, **TD No. 787 (January 2011).**

R. BRONZINI and E. IACHINI, *Are incentives for R&D effective? Evidence from a regression discontinuity approach,* American Economic Journal : Economic Policy, v. 6, 4, pp. 100-134, **TD No. 791 (February 2011).**

P. ANGELINI, S. NERI and F. PANETTA, *The interaction between capital requirements and monetary policy*, Journal of Money, Credit and Banking, v. 46, 6, pp. 1073-1112, **TD No. 801 (March 2011).**

M. BRAGA, M. PACCAGNELLA and M. PELLIZZARI, *Evaluating students' evaluations of professors,* Economics of Education Review, v. 41, pp. 71-88, **TD No. 825 (October 2011).**

M. FRANCESE and R. MARZIA, *Is there Room for containing healthcare costs? An analysis of regional spending differentials in Italy,* The European Journal of Health Economics, v. 15, 2, pp. 117-132, **TD No. 828 (October 2011).**

L. GAMBACORTA and P. E. MISTRULLI, *Bank heterogeneity and interest rate setting: what lessons have we learned since Lehman Brothers?,* Journal of Money, Credit and Banking, v. 46, 4, pp. 753-778, **TD No. 829 (October 2011).**

M. PERICOLI, *Real term structure and inflation compensation in the euro area*, International Journal of Central Banking, v. 10, 1, pp. 1-42, **TD No. 841 (January 2012).**

E. GENNARI and G. MESSINA, *How sticky are local expenditures in Italy? Assessing the relevance of the flypaper effect through municipal data,* International Tax and Public Finance, v. 21, 2, pp. 324-344, **TD No. 844 (January 2012).**

V. DI GACINTO, M. GOMELLINI, G. MICUCCI and M. PAGNINI, *Mapping local productivity advantages in Italy: industrial districts, cities or both?*, Journal of Economic Geography, v. 14, pp. 365–394, **TD No. 850 (January 2012).**

A. ACCETTURO, F. MANARESI, S. MOCETTI and E. OLIVIERI, *Don't Stand so close to me: the urban impact of immigration,* Regional Science and Urban Economics, v. 45, pp. 45-56, **TD No. 866 (April 2012).**

M. PORQUEDDU and F. VENDITTI, *Do food commodity prices have asymmetric effects on euro area inflation,* Studies in Nonlinear Dynamics and Econometrics, v. 18, 4, pp. 419-443, **TD No. 878 (September 2012).**

S. FEDERICO, *Industry dynamics and competition from low-wage countries: evidence on Italy*, Oxford Bulletin of Economics and Statistics, v. 76, 3, pp. 389-410, **TD No. 879 (September 2012).**

F. D'AMURI and G. PERI, *Immigration, jobs and employment protection: evidence from Europe before and during the Great Recession,* Journal of the European Economic Association, v. 12, 2, pp. 432-464, **TD No. 886 (October 2012).**

M. TABOGA, *What is a prime bank? A euribor-OIS spread perspective,* International Finance, v. 17, 1, pp. 51-75, **TD No. 895 (January 2013).**

G. CANNONE and D. FANTINO, *Evaluating the efficacy of european regional funds for R&D,* Rassegna italiana di valutazione, v. 58, pp. 165-196, **TD No. 902 (February 2013).**

L. GAMBACORTA and F. M. SIGNORETTI, *Should monetary policy lean against the wind? An analysis based on a DSGE model with banking,* Journal of Economic Dynamics and Control, v. 43, pp. 146-74, **TD No. 921 (July 2013).**

M. BARIGOZZI, CONTI A.M. and M. LUCIANI, *Do euro area countries respond asymmetrically to the common monetary policy?,* Oxford Bulletin of Economics and Statistics, v. 76, 5, pp. 693-714, **TD No. 923 (July 2013).**

U. ALBERTAZZI and M. BOTTERO, *Foreign bank lending: evidence from the global financial crisis,* Journal of International Economics, v. 92, 1, pp. 22-35, **TD No. 926 (July 2013).**

R. De Bonis and A. Silvestrini, *The Italian financial cycle: 1861-2011,* Cliometrica, v.8, 3, pp. 301-334, **TD No. 936 (October 2013).**

G. Barone and S. Mocetti, *Natural disasters, growth and institutions: a tale of two earthquakes,* Journal of Urban Economics, v. 84, pp. 52-66, **TD No. 949 (January 2014).**

D. Pianeselli and A. Zaghini, *The cost of firms' debt financing and the global financial crisis,* Finance Research Letters, v. 11, 2, pp. 74-83, **TD No. 950 (February 2014).**

J. Li and G. Zinna, *On bank credit risk: sytemic or bank-specific? Evidence from the US and UK,* Journal of Financial and Quantitative Analysis, v. 49, 5/6, pp. 1403-1442, **TD No. 951 (February 2015).**

A. Zaghini, *Bank bonds: size, systemic relevance and the sovereign,* International Finance, v. 17, 2, pp. 161-183, **TD No. 966 (July 2014).**

G. Sbrana and A. Silvestrini, *Random switching exponential smoothing and inventory forecasting,* International Journal of Production Economics, v. 156, 1, pp. 283-294, **TD No. 971 (October 2014).**

M. Silvia, *Does issuing equity help R&D activity? Evidence from unlisted Italian high-tech manufacturing firms,* Economics of Innovation and New Technology, v. 23, 8, pp. 825-854, **TD No. 978 (October 2014).**

2015

G. de Blasio, D. Fantino and G. Pellegrini, *Evaluating the impact of innovation incentives: evidence from an unexpected shortage of funds,* Industrial and Corporate Change, , v. 24, 6, pp. 1285-1314, **TD No. 792 (February 2011).**

M. Bugamelli, S. Fabiani and E. Sette, *The age of the dragon: the effect of imports from China on firm-level prices*, Journal of Money, Credit and Banking, v. 47, 6, pp. 1091-1118, **TD No. 737 (January 2010).**

R. Bronzini, *The effects of extensive and intensive margins of FDI on domestic employment: microeconomic evidence from Italy*, B.E. Journal of Economic Analysis & Policy, v. 15, 4, pp. 2079-2109, **TD No. 769 (July 2010).**

U. Albertazzi, G. Eramo, L. Gambacorta and C. Salleo, *Asymmetric information in securitization: an empirical assessment*, Journal of Monetary Economics, v. 71, pp. 33-49, **TD No. 796 (February 2011).**

A. Di Cesare, A. P. Stork and C. de Vries, *Risk measures for autocorrelated hedge fund returns,* Journal of Financial Econometrics, v. 13, 4, pp. 868-895, **TD No. 831 (October 2011).**

G. Bulligan, M. Marcellino and F. Venditti, *Forecasting economic activity with targeted predictors,* International Journal of Forecasting, v. 31, 1, pp. 188-206, **TD No. 847 (February 2012).**

A. Ciarlone, *House price cycles in emerging economies,* Studies in Economics and Finance, v. 32, 1, **TD No. 863 (May 2012).**

D. Fantino, A. Mori and D. Scalise, *Collaboration between firms and universities in Italy: the role of a firm's proximity to top-rated departments,* Rivista Italiana degli economisti, v. 1, 2, pp. 219-251, **TD No. 884 (October 2012).**

A. Bardozzetti and D. Dottori, *Collective Action Clauses: how do they Affect Sovereign Bond Yields?,* Journal of International Economics , v 92, 2, pp. 286-303, **TD No. 897 (January 2013).**

D. Depalo, R. Giordano and E. Papapetrou, *Public-private wage differentials in euro area countries: evidence from quantile decomposition analysis,* Empirical Economics, v. 49, 3, pp. 985-1115, **TD No. 907 (April 2013).**

G. Barone and G. Narciso, *Organized crime and business subsidies: Where does the money go?,* Journal of Urban Economics, v. 86, pp. 98-110, **TD No. 916 (June 2013).**

P. Alessandri and B. Nelson, *Simple banking: profitability and the yield curve,* Journal of Money, Credit and Banking, v. 47, 1, pp. 143-175, **TD No. 945 (January 2014).**

M. Taneli and B. Ohl, *Information acquisition and learning from prices over the business cycle,* Journal of Economic Theory, 158 B, pp. 585–633, **TD No. 946 (January 2014).**

R. Aaberge and A. Brandolini, *Multidimensional poverty and inequality,* in A. B. Atkinson and F. Bourguignon (eds.), Handbook of Income Distribution, Volume 2A, Amsterdam, Elsevier, **TD No. 976 (October 2014).**

V. CUCINIELLO and F. M. SIGNORETTI, *Large banks,loan rate markup and monetary policy,* International Journal of Central Banking, v. 11, 3, pp. 141-177, **TD No. 987 (November 2014).**

M. FRATZSCHER, D. RIMEC, L. SARNOB and G. ZINNA, *The scapegoat theory of exchange rates: the first tests,* Journal of Monetary Economics, v. 70, 1, pp. 1-21, **TD No. 991 (November 2014).**

A. NOTARPIETRO and S. SIVIERO, *Optimal monetary policy rules and house prices: the role of financial frictions,* Journal of Money, Credit and Banking, v. 47, S1, pp. 383-410, **TD No. 993 (November 2014).**

R. ANTONIETTI, R. BRONZINI and G. CAINELLI, *Inward greenfield FDI and innovation,* Economia e Politica Industriale, v. 42, 1, pp. 93-116, **TD No. 1006 (March 2015).**

T. CESARONI, *Procyclicality of credit rating systems: how to manage it,* Journal of Economics and Business, v. 82. pp. 62-83, **TD No. 1034 (October 2015).**

M. RIGGI and F. VENDITTI, *The time varying effect of oil price shocks on euro-area exports,* Journal of Economic Dynamics and Control, v. 59, pp. 75-94, **TD No. 1035 (October 2015).**

2016

E. BONACCORSI DI PATTI and E. SETTE, *Did the securitization market freeze affect bank lending during the financial crisis? Evidence from a credit register,* Journal of Financial Intermediation , v. 25, 1, pp. 54-76, **TD No. 848 (February 2012).**

M. MARCELLINO, M. PORQUEDDU and F. VENDITTI, *Short-Term GDP forecasting with a mixed frequency dynamic factor model with stochastic volatility,* Journal of Business & Economic Statistics , v. 34, 1, pp. 118-127, **TD No. 896 (January 2013).**

M. ANDINI and G. DE BLASIO, *Local development that money cannot buy: Italy's Contratti di Programma,* Journal of Economic Geography, v. 16, 2, pp. 365-393, **TD No. 915 (June 2013).**

F BRIPI, *The role of regulation on entry: evidence from the Italian provinces,* World Bank Economic Review, v. 30, 2, pp. 383-411, **TD No. 932 (September 2013).**

L. ESPOSITO, A. NOBILI and T. ROPELE, *The management of interest rate risk during the crisis: evidence from Italian banks,* Journal of Banking & Finance, v. 59, pp. 486-504, **TD No. 933 (September 2013).**

F. BUSETTI and M. CAIVANO, *The trend–cycle decomposition of output and the Phillips Curve: bayesian estimates for Italy and the Euro Area,* Empirical Economics, V. 50, 4, pp. 1565-1587, **TD No. 941 (November 2013).**

M. CAIVANO and A. HARVEY, *Time-series models with an EGB2 conditional distribution,* Journal of Time Series Analysis, v. 35, 6, pp. 558-571, **TD No. 947 (January 2014).**

G. ALBANESE, G. DE BLASIO and P. SESTITO, *My parents taught me. evidence on the family transmission of values,* Journal of Population Economics, v. 29, 2, pp. 571-592, **TD No. 955 (March 2014).**

R. BRONZINI and P. PISELLI, *The impact of R&D subsidies on firm innovation,* Research Policy, v. 45, 2, pp. 442-457, **TD No. 960 (April 2014).**

L. BURLON and M. VILALTA-BUFI, *A new look at technical progress and early retirement,* IZA Journal of Labor Policy, v. 5, **TD No. 963 (June 2014).**

A. BRANDOLINI and E. VIVIANO, *Behind and beyond the (headcount) employment rate,* Journal of the Royal Statistical Society: Series A, v. 179, 3, pp. 657-681, **TD No. 965 (July 2015).**

A. BELTRATTI, B. BORTOLOTTI and M. CACCAVAIO, *Stock market efficiency in China: evidence from the split-share reform,* Quarterly Review of Economics and Finance, v. 60, pp. 125-137, **TD No. 969 (October 2014).**

A. CIARLONE and V. MICELI, *Escaping financial crises? Macro evidence from sovereign wealth funds' investment behaviour,* Emerging Markets Review, v. 27, 2, pp. 169-196, **TD No. 972 (October 2014).**

D. DOTTORI and M. MANNA, *Strategy and tactics in public debt management,* Journal of Policy Modeling, v. 38, 1, pp. 1-25, **TD No. 1005 (March 2015).**

F. CORNELI and E. TARANTINO, *Sovereign debt and reserves with liquidity and productivity crises,* Journal of International Money and Finance, v. 65, pp. 166-194, **TD No. 1012 (June 2015).**

G. RODANO, N. SERRANO-VELARDE and E. TARANTINO, *Bankruptcy law and bank financing,* Journal of Financial Economics, v. 120, 2, pp. 363-382, **TD No. 1013 (June 2015).**

S. BOLATTO and M. SBRACIA, *Deconstructing the gains from trade: selection of industries vs reallocation of workers,* Review of International Economics, v. 24, 2, pp. 344-363, **TD No. 1037 (November 2015).**

A. CALZA and A. ZAGHINI, *Shoe-leather costs in the euro area and the foreign demand for euro banknotes,* International Journal of Central Banking, v. 12, 1, pp. 231-246, **TD No. 1039 (December 2015).**

E. CIANI, *Retirement, Pension eligibility and home production,* Labour Economics, v. 38, pp. 106-120, **TD No. 1056 (March 2016).**

L. D'AURIZIO and D. DEPALO, *An evaluation of the policies on repayment of government's trade debt in Italy,* Italian Economic Journal, v. 2, 2, pp. 167-196, **TD No. 1061 (April 2016).**

FORTHCOMING

S. MOCETTI, M. PAGNINI and E. SETTE, *Information technology and banking organization,* Journal of Financial Services Research, **TD No. 752 (March 2010).**

G. MICUCCI and P. ROSSI, *Debt restructuring and the role of banks' organizational structure and lending technologies,* Journal of Financial Services Research, **TD No. 763 (June 2010).**

M. RIGGI, *Capital destruction, jobless recoveries, and the discipline device role of unemployment,* Macroeconomic Dynamics, **TD No. 871 July 2012).**

S. FEDERICO and E. TOSTI, *Exporters and importers of services: firm-level evidence on Italy,* The World Economy, **TD No. 877 (September 2012).**

P. BOLTON, X. FREIXAS, L. GAMBACORTA and P. E. MISTRULLI, *Relationship and transaction lending in a crisis*, Review of Financial Studies, **TD No. 917 (July 2013).**

G. DE BLASIO and S. POY, *The impact of local minimum wages on employment: evidence from Italy in the 1950s,* Regional Science and Urban Economics, **TD No. 953 (March 2014).**

A. L. MANCINI, C. MONFARDINI and S. PASQUA, *Is a good example the best sermon? Children's imitation of parental reading,* Review of Economics of the Household, **TD No. 958 (April 2014).**

L. BURLON, *Public expenditure distribution, voting, and growth,* Journal of Public Economic Theory, **TD No. 961 (April 2014).**

G. ZINNA, *Price pressures on UK real rates: an empirical investigation,* Review of Finance, **TD No. 968 (July 2014).**

U. ALBERTAZZI, M. BOTTERO and G. SENE, *Information externalities in the credit market and the spell of credit rationing,* Journal of Financial Intermediation, **TD No. 980 (November 2014).**

A. BORIN and M. MANCINI, *Foreign direct investment and firm performance: an empirical analysis of Italian firms,* Review of World Economics, **TD No. 1011 (June 2015).**

R. BRONZINI and A. D'IGNAZIO, *Bank internationalisation and firm exports: evidence from matched firm-bank data,* Review of International Economics, **TD No. 1055 (March 2016).**