Size Metrics and Dynamics of Firm Expansion in the European Pharmaceutical Industry*

Franco Mariuzzo†
School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ, England.

Xiaoheng Zhang‡

May 7, 2013

Abstract

We pick the most discrete measure of firm size, the cumulative number of unitary expansions as pillar metric, and through competing dynamic panel Probit methods that control for size and deal with unobserved heterogeneity and initial conditions, investigate how pharmaceutical multinationals capture opportunities (expand/grow) during the European Single Market Programme era, 1993-2004. We extend the empirical Growth-of-Firms literature by linking discrete and continuous metrics of size via a copula approach, and simulate the remaining firm size metrics from the pillar metric and the forecasted pillar metric. We validate our approach by comparing the simulated dynamics with those obtained from the equation-by-equation estimation. We find that the copula is a superior approach to simulate size metrics distributions. Most importantly, it is able to provide us with a multivariate function whose dynamics offer a more comprehensive snapshot of firm growth.
KEYWORDS: Copula, dynamic Probit models with unobserved heterogeneity, firm-growth, pharmaceutical industry, Single Market Programme.
JEL classifiers: C10, C11, C23, L11, L65.

*This paper was presented at the Irish Economic Association conference, Westport, Ireland (2008), the Small Open Economies in a Globalized World conference, Waterloo, Canada (2008), and the European Economic Association, Barcelona, Spain (2009). We thank Wiji Arulampalam, Silvi Berger, Steve Davies, Mike Harrison, Ruud Koning, Bruce Lyons and Paul Walsh for their precious comments. We are grateful to referees for a number of helpful suggestions that have greatly improved our article.
†Corresponding author: Tel. +44 (0)1603 59 7530. Fax. +44 (0)1603 45 6259. E-mail: F.Mariuzzo@uea.ac.uk.
‡Email: zhangx@tcd.ie. A grant from the Institute for International Integration Studies, Trinity College Dublin, is gratefully acknowledged.

1 Introduction

The literature on firm growth has repeatedly relied on any readily available measure/metric of size to empirically test, either within an industry or across industries, the validity of Gibrat's law of proportionate effect, i.e., the independence between firm growth and firm size, which leads to a size probability distribution that commonly is lognormal. While for some research papers the choice of the measure is led by the research question, for others it is at discretion of the researcher. The problem of working with one measure of firm size is that in principle, there may be systematic differences in the dynamics of alternative metrics of size, with a implicit risk that the resulting analysis would suffer from being tailored to the chosen metric. This was a concern well recognized in the literature, and addressed in a non satisfactory way by Kumar (1985) with the comparison of different (quasi-continuous) metrics of size to examine firm growth theory for quoted UK companies during the period 1960-1976. According to his results, Gibrat's law is satisfied for some metrics, but not for others. This article extends his work and centers the research question on how one deals with alternative metrics of firm size, some discrete and others continuous. Unlike Kumar, we do not analyze firm growth equation-by-equation using Least Squares, but strive to build up a more robust procedure. We begin by studying how alternative measures of firm size relate to one another in a joint multivariate nonnormal probability distribution, where nonnormality is induced by combining different (quasi-discrete and quasi-continuous) distributions. Given nonnormality, we rely on a copula approach: a methodology that only requires that one knows the marginals of the variates to fit the multivariate distribution, as will be seen. We fully exploit such approach in a three-step procedure. The first step is devoted to studying scale-free population measures of association between the firm size random variables, so as to understand how similarly or differently the alternative metrics might behave.[1] If we were to find evidence of metrics being highly positively associated, we could interpret this result as indication that one does not have to worry much about devising a selected metric. The researcher could, in this case, stick to his most preferred (or a handier) measure of firm size and pursue the analysis on the growth dynamics of that measure; we proceed to do this here, studying the discrete expansion dynamics of the European pharmaceutical multinationals during the Single Market Programme era, 1993-2004 using our chosen pillar metric measuring new subsidiaries as indicators of expansions (opportunities). This is an ideal measure during the EU enlargement, for during this time period acquisition growth is the main engine for growth, with minimal effect of internal growth. The copula is then deployed to bridge from the chosen metric of firm size to the remaining metrics. We condition the best fitted copula first to the observed, and then to the forecasted, values of the cumulated number of opportunities captured, and simulate the remaining size metrics upon that. This methodology allows us to study firm growth dynamics on one size metric dimension and produce the firm growth dynamics on any other size metric dimension. We evaluate the accuracy of the copula procedure by confronting its simulations, with those forecasted from equation-by-equation size metrics estimation of linear and nonlinear dynamic panel models. The results suggest that the copula technique is preferred to recover the probability distributions of alternative size metrics.

In a last step the copula procedure is exploited to produce a more global measure of firm size. This new metric is the fitted multivariate cumulative density function of the various observable metrics of firm size. Though normalized, this global measure is very informative on how each individual metric contributes to firm growth. We are interested in decoding its behavior over time.

We have deliberately chosen the European pharmaceutical industry during the Single Market Programme era (1993-2004) to apply this extended version of firm growth theory for two reasons. Firstly, it is a top industry for the European economy in terms of manufacturing value added and R&D investment; in fact, it is the fifth largest industry in the European Union in terms of manufacturing value added, amounting to 3.5 per cent, and it accounts for about 17 per cent of total EU business R&D expenditures, according to EFPIA (2005). R&D activity is key variable in this industry, and the stochastic nature of innovation makes this industry ideal to test the stochastic firm growth theory. Secondly, the industry was heavily regulated at the national level and had to cope with non-tariff barriers (e.g. frontier controls, national differences in regulations, public procurements favorable to domestic producers) before the Single Market Programme era, as largely described by Cecchini et al. (1988). In such situation the enlargement of the European market gives the firms in this industry significant scope to expand, increase production efficiency and enhance R&D capacity through mergers and acquisitions, relocations or external

[1] A scale-free measure of dependence is a measure unchanged under positive monotonic transformations of the random variable.

collaborations. In turn, this translates into more business opportunities, expansions, available to utilize in our empirical analysis of acquisition growth.[2]

Gibrat's (1931) "law of proportionate effect" (English translation Gibrat (1957)) states that the expected *rate* of growth of a firm's size is, each time period, independent of its current magnitude. Such an early challenging theory, capable of explaining important regularities, has motivated an extensive literature in the joint field of Statistics and Industrial Economics, with clear focus on the understanding of *skewness* in firm size probability distributions. This Growth-of-Firms literature, also termed stochastic literature on firm size, sees skewness as the result of the relationship between a firm's size and its absolute growth (hereafter simply called growth), i.e., if larger firms tend to grow faster/slower than their smaller rivals then the industry size distribution tends to exhibit more/less positive skewness, this producing longer/shorter right tails. By modifying the boundary conditions and underlying assumptions, the literature has derived and discussed several probability distributions capable of explaining size, such as: Exponential, Fisher's log series, Geometric, Log-normal, Negative Binomial, Pareto (Zipf), Poisson, and Yule. Albeit merely statistical, this literature has received considerable attention in Industrial Economics, due to its contribution in explaining firm market concentration; one of the three components of the structure-conduct-performance paradigm that started with Mason (1939, 1948) and continued with Bain (1951, 1956), and a long list of followers.[3] It obscured a previous literature that used the feature of the cost function to provide arguments in favor of industry regularities. The link between the two literatures is the core of Simon and Bonini (1958).

The lack of economic theory behind this stochastic literature on firm size created room for a new stream of literature, originating in the 1970s, which employed either optimization theory, or game theory, to explain firm growth. In this latter approach skewness takes on a more deterministic nature, given by the observable role of measurable economic factors.[4]

Sutton (1997a, 1998) bridges the gap between these two streams of literature by proposing a game theoretical "independent submarkets" theory. His theory of firm size distribution relies on a simplified discrete metric that counts the number of opportunities a firm has captured; expressed in his terminology as the number of "independent submarkets" the firm has entered. This simple metric, suggested originally in Ijiri and Simon (1967) and Simon and Ijiri (1977), was termed by Sutton the number of "independent isolated islands" a firm expands to. Here, size is determined by the cumulated sum of (homogeneous) unitary expansions. In his book, Sutton (1998) sketches how to extend the theory of unitary expansions to the more realistic case of unequal (heterogeneous) discrete opportunities (pp. 258-9, 290-1), which previously Simon and Ijiri (1977) called "random increments". In short, Sutton's main finding is that a lower bound in a Lorenz curve distribution (i.e., a minimum degree of inequality in firm size) is the best that one can do to unravel the industry distribution of firm size: any distance between the actual data and the lower bound can be justified by, among other things, the degree of heterogeneity in the arrival of opportunities.[5]

Relevant to our pharmaceutical dataset, Sutton defines the submarkets of the pharmaceutical industry as chemically related groups within each therapeutical category of drugs. He documents and discusses the history of 50 top best-sellers and concludes that a lottery model of random selection is a first approximation of the R&D process in this industry, because the scope of one new drug leading to another drug in the same chemical group is limited. Hence, he rationalizes skewness to be the result of the limiting firm size distribution of the industry.

Our article builds up from Gibrat's law and promotes the theory of maturation, whose original concept was spread in the literature by Rostow (1959). This latter theory entails a more general way to describe skewness in the firm size distribution. It does not preclude Gibrat's law to be valid, but defers its validity to a set of restrictions to

[2]Studies on the pharmaceutical industry which are relevant to this paper can be found in Howells (1992), Matraves (1999), Bottazzi et al. (2001), Kotzian (2004), Bottazzi and Secchi (2005, 2006b), Buldyrev et al. (2007), Cefis et al. (2007) and Ornaghi (2009). In addition, studies on the industrial organization of the European Union prior to the implementation of the Single European Market can be found in Davies and Lyons (1996).

[3]A non-exhaustive list of followers includes: Simon (1955), Hart and Prais (1956), Adelman (1958), Hart (1962), Hymer and Pashigian (1962), Mansfield (1962), Ijiri and Simon (1964, 1967), Samuels (1965), Steindl (1965, 1968), Wedervang (1965), Quandt (1966), Singh et al. (1968), Davies and Lyons (1982), Simon and Ijiri (1977), Luttmer (2006, 2007, 2011), Buldyrev et al. (2007), Cefis et al. (2007). See de Wit (2005) for an overview of the stochastic models of firm size.

[4]Here a non-exhaustive list of papers includes: Hjalmarsson (1974), Lucas Jr (1978), Jovanovic (1982), Selten (1983), Hall (1987), Klepper and Graddy (1990), Cohen and Klepper (1992, 1996), Ericson and Pakes (1995), Klepper (1996), Klepper and Simons (2000), Cabral and Mata (2003), Klette and Kortum (2004), Klepper and Thompson (2006).

[5]Empirical support for this theory is available in Bottazzi et al. (2001), de Juan (2003), Amisano and Giorgetti (2005), Bottazzi and Secchi (2005, 2006b), Buzzacchi and Valletti (2006).

be met.

We stress the fact that while Simon and Ijiri (1977) in the section called "random increments" manage to extend the Growth-of-Firms theory to cover alternative discrete measures of size, their generalization to continuous metrics of size, is not explained, but is limited to the statement:

> Whether sales, assets, number of employees, value added, or profits, are used as the size measure, the observed distributions always belong to the class of highly skewed distributions $\cdots$,

the same statement is also available in Simon and Bonini (1958). Such limitation is made more transparent in Sutton (1997b):

> Size can be measured in a number of ways $\cdots$. Though we might in principle expect systematic differences between the several measures, such differences have not been a focus of interest in the literature.

This article fills in the generalization gap by empirically investigating the association between alternative discrete and continuous measures of firm size. Specifically, it is the relation between size measured as number of cumulated unitary opportunities and alternative heterogeneous metrics of size that motivates our work. Given that multinationals are the unit of investigation of this paper, we shall call firms "multinationals" and employ the abbreviation MNE. We obtained our data from a commercial database called *Amadeus*. It is worth mentioning that our data provide no reliable information on firm exit, meaning that exit dynamics cannot be studied.[6] We had access to three versions of Amadeus (v1997, v2000 and v2005) and technically a firm (subsidiary) observed in the first database (v1997) can be tracked along the time line up to 2005 using the unique *id* assigned to it. However, there are competing reasons for a firm being observed in v1997, but not in v2005. It could be that the firm (subsidiary) went out of the business or it was taken over, and lost its business entity (its unique *id*) becoming a department of the acquirer, or that Amadeus failed to track this firm down the time line. Given this impossibility to identify pure firm exit we prefer not to include such observations in our study.

Our paper considers an MNE catching a new opportunity (expanding), whenever it enlarges its number of subsidiaries in the market. In line with the Growth-of-Firms literature we treat arrival of opportunities as stochastic. However, aligned to its competing stream of literature, we examine the dynamics of a firm's success in capturing opportunities in a more structural way.

The rest of the article is organized as follows: in Section 2 we review a simple conceptual model; in Section 3 we outline the copula procedure. Section 4 presents the dynamic panel Probit approach with unobserved heterogeneity and compare alternative econometric techniques. Section 5 describes our data. Our results are discussed in Section 6. Section 7 concludes.

2 The Conceptual Model

We generalize Steindl's (1965) formulation of Gibrat's (1931) law by allowing MNE size to be measured by any metric $h \in \mathcal{H}$. Momentarily we do not model the correlation between different size metrics, but model them as independent, that is, equation-by-equation. We denote the size of the MNE i at time t with s_{hit} and the *iid* random variable of proportionate rate of growth between $t-1$ and t, with $\epsilon_{hit} \sim \left(\mu_h, \sigma_h^2\right)$. Absolute growth (onward, simply *growth*) is expressed as

$$s_{hit} - s_{hi,t-1} = s_{hi,t-1}\epsilon_{hit}. \tag{1}$$

Under the assumption of discrete time periods, short enough to make the variance of ϵ_{hit} small, log transformations of a time expansion of the relation in Eq. (1) lead to the random walk on a logarithmic scale approximation

$$logs_{hit} \simeq logs_{hi0} + \epsilon_{hi1} + \epsilon_{hi2} + \cdots + \epsilon_{hit}. \tag{2}$$

In the limit as $t \to \infty$, $logs_{hi0}$ can be omitted from Eq. (2), being small compared to $logs_{hit}$. If we make use of the Lindeberg-Levy central limit theorem we get that $logs_{hit}$ has a limiting normal distribution, with mean $\mu_h t$ and variance $\sigma_h^2 t$. Hence, s_{hit} is shaped as a skewed limiting lognormal distribution, which holds true if the underlying

[6] For an investigation of the sample selection bias generated by nonrandom firm exit, refer to Hall (1987), Evans (1987a,b), Dunne et al. (1988, 1989) and Geroski and Machin (1991).

assumptions are met.

We build upon the above relation to structure our econometric analysis. We replace the the right hand side of Eq. (1) with a composite function $f(\circ)$ of contemporaneous random profit, which maps the cumulative density function of the random component of profit (denoted with G) into growth. This modification allows us to relate growth and profit, bearing in mind that key determinant of profit in the pharmaceutical industry is, along with them market expansion itself, successful R&D. It is the randomness of the R&D success that makes the pharmaceutical industry suitable to study firm growth within the Growth-of-Firms literature. We express the monotonic function of profit as a function of the original one-period lagged size, observable covariates and a new random error term. We incorporate all observable effects into the unique vector $\mathbf{w}_h$. This modification, brings in a set of unknown parameters, that we denote with $\boldsymbol{\theta}_h$. Growth is now formulated as the function:

$$s_{hit} - s_{hi,t-1} = f_h \left[G_h \left(s_{hi,t-1}, \mathbf{w}_{hit}; \boldsymbol{\theta}_h \right) \right]. \tag{3a}$$

For notational convenience, we label the growth of the MNE i in period t, $s_{hit} - s_{hi,t-1}$, with y_{hit}.

Among the elements of the set of all metrics $\mathcal{H}$ we have a simple discrete measure of size: the number of cumulated opportunities (which we interchangeably call expansions) that an MNE has captured. We label this metric with *op* and update Eq. (3a) to be

$$y_{(op)it} = f_{(op)} \left[G_{(op)} \left(s_{(op)i,t-1}, \mathbf{w}_{(op)it}; \boldsymbol{\theta}_{(op)} \right) \right] = \mathbb{I} \left[G \left(\pi_{it} \geq 0 \right) \right], \tag{3b}$$

where $= \mathbb{I}$ is the indicator function and π is the profit function.[7] In this particular case, growth is a dichotomous variable that takes on integer values $y_{(op)it} \in \{0, 1\}$, with a value of 1 indicating an expansion, and $f_{(op)}(\circ)$ being the function that maps from the profit CDF to the binary growth, in this case $f_{(op)}(\circ) = \mathbb{I}$. To ease the notation, we omit the subscript $_{(op)}$ in the remainder of the article, but it should be understood that when no explicit notation on size metrics is used we are referring to size measured by the cumulative number of opportunities, also called pillar metric for reasons that shall become clear shortly.

With an obvious interest in the dynamics of expansion, and with the awareness of data limitation, we choose the variables entering $\mathbf{w}_{it}$ to be the lagged dependent variable $y_{i,t-1}$, along with a set of time dummies and observable MNE-specific characteristics. The lagged dependent variable is added to capture any form of observable persistence in the MNE expansion process. Its inclusion is in line with Ijiri and Simon's (1967) model of autocorrelated growth, whose simulations are the core of Ijiri and Simon (1964).[8]

Our decision to have lagged number of cumulated opportunities, $s_{i,t-1}$, directly influencing the growth is motivated by Gibrat's law, as re-interpreted in Simon and Ijiri (1977) and Sutton (1998). It is a result of the right hand side of Eq. (1). Each active MNE in the industry has a certain probability of capturing an opportunity, i.e., a certain probability of expanding. The number of opportunities captured may have direct influence on each MNE's propensity to expand, and perhaps it could conform to Gibrat's law: the probability that an MNE captures a new opportunity is proportionate to its size, which is another way to say that the expected growth *rate* is independent of firm size. As an MNE becomes larger it has more chances to catch new opportunities, and the continuation of this process is reflected in the skewness of the industry size distribution. Our econometric model shares the same spirit of the theoretical model of firm growth proposed in Bottazzi and Secchi (2006a) - in the context of finite opportunities, a firm's probability of capturing a new opportunity is positively related to the existing opportunities it has captured. However, because of the nonlinear way that lagged size enters the f composite function in Eq. (3b), Gibrat's relation between firm growth and size is no longer obvious, and it will be analyzed by stratifying groups of MNE by size.

We believe that studying the growth expansion process of the European pharmaceutical industry during the Single Market Programme era, does give satisfactory insight on the resulting industry size concentration. Though our analysis is at the MNE-level, evidence of a positive and significant coefficient, attached to an expansion that occurred at time $t-1$, would indicate the sustainability of this industry. In other words, it would suggest that

[7] We have arbitrarily set the profit cutoff at zero. Alternatively, we could have set it at $\underline{\pi}$, or at $\underline{\pi}_t$ if it varied over time. The implication of not observing the cutoff point is that in our econometrics we shall not be able to identify either the constant, if $\underline{\pi} \neq 0$, or the time dummies, if $\underline{\pi}_t \neq 0$.

[8] The serial correlation assumption states that the probability of growth of an existing MNE is proportional to the weighted sum of past increments of size, and the weight is decreasing the further the occurrence of each increment from the current period. Such a carry-over effect can be triggered by successful innovation in production or marketing processes.

the industry is able to generate profits that are large enough to support its expensive innovative activities and secure profits into the future. In this way, past expansions exert a behavioral effect on current expansion and this effect is termed "true state dependency" by Heckman (1981b). However, as is often the case, the examination of dynamics can be blurred by what Heckman (1981b) has termed "spurious state dependency". This occurs if unobservable multinational-level effects are serially correlated over time or are correlated with initial expansion, and these correlations have not been properly controlled for. In those cases, the lagged expansion will incorrectly capture this unobserved effect, behaving as if it is a driving force behind the current expansion, even if there is no state dependence at all. To tackle this problem we adopt a dynamic panel random effects model that deals with unobserved heterogeneity and the initial conditions problem, and estimate the $\boldsymbol{\theta}$ parameters. We will discuss the mechanics of this model in the next section.

Behind our structural model there is an underlying economic motivation. In every period of time, MNEs are given the chance to expand their size by capturing one of the opportunities available in the market, e.g. successful R&D. We motivate the "race" for opportunities in the following way. Depending on their profit realization, our companies succeed, or not, in retaining one of the available opportunities. However, though opportunities are homogeneous units, they also spread effects into alternative heterogeneous metrics of MNE size. Understanding how sum of opportunities and stocks of alternative measures of size relate to one another is our first ambition. Our data allow for six alternative metrics of MNE size. Two are discrete: i) the aforementioned cumulative number of opportunities captured, $s_{(op)}$; ii) the cumulative number of subsidiaries established or acquired, $s_{(su)}$; and four are continuous: iii) employment, $s_{(em)}$; iv) operational revenues (turnover), $s_{(or)}$; v) sales, $s_{(sa)}$; and vi) total assets, $s_{(ta)}$. In Table 1 we document a more exhaustive, though still incomplete, list of univariate measures of MNE size adopted in the literature, such as: capital, inputs, output, plants and equipment, and profit.

In Eq. (3b) we have suggested an econometric relation to study MNE growth, when MNE size is measured by number of cumulated opportunities. What happens if MNE size is measured by any other metric? Here we dispose of two alternative methods to investigate such dynamics and understand the skewness of MNE size probability distribution. The first approach estimates the MNE growth dynamics equation-by-equation, and for each size metric. One downside of this method is that, because of unobserved heterogeneity, the estimator which is consistent to study the growth dynamics for the discrete metrics is no longer consistent for the continuous metrics. We will need to use one type of estimator to fit the growth dynamics of continuous metrics, and another estimator to fit the growth dynamics of discrete metrics, though we will do our best to make the estimates as comparable as possible. Another obvious limitation is that when we estimate the growth dynamics of one size metric, we have very little to say on what would had been the growth dynamics if we were to use another metric, unless we would be willing to control for the correlation between the error terms of the various measures of growth and estimate a system of equations, accounting for the complexity of combining linear and nonlinear dynamic estimators. The second approach avoids this univariate analysis and exploits the possible cross-sectional association between the various MNE size metrics. While we research both methods, we give more room to the latter, as it also has the additional advantage of generating a global measure of firm size. In this second method we are interested in the joint cumulative density function of the various metrics of size, $G(\mathbf{s})$, with $\mathbf{s} \equiv [s_1, s_2, \cdots, s_H]$. Because we are mixing together discrete and continuous measures of size, it would be inappropriate to assume, as is often the case in the empirical literature, $G(\mathbf{s})$ to be cumulative multivariate normal, so the functional form of $G(\mathbf{s})$ has to be fitted. We exploit the information on the marginal distribution of each metric of size, $G_h(s_h)$, and employ the copula approach to gain knowledge of the underlying multivariate CDF. Crucial points in favor of this copula approach are: i) It only exploits information on the marginal distributions of the set of variables; ii) It allows us to investigate the growth dynamics of one metric and simulate the values of the remaining metrics based upon that variable; iii) The joint CDF is an important measure of (normalized) size, which we write as an element of the superset $\mathcal{H}^* \supset \mathcal{H}$. Hence we can now study the general dynamics of relative size expansion.

We outline the copula approach in the next section.

3 The copula

Most of the content of this section and of the extension developed in Appendix B is drawn from Nelsen (2006).
The copula is a statistical function that was first introduced by Hoeffding (1940), see collection Hoeffding (1994), but is commonly associated to Sklar (1973). It is a function that maps from marginal distributions to a multivariate joint

distribution. As such, it can be used to recover multivariate joint distributions from information on marginal distributions, and is particularly suitable to assembling non-normal data or marginals coming from different parametric families. Prior to outlining the copula, we introduce some notation. Let $(S_1, S_2, \cdots, S_H)$ be H random variables, with distribution functions $G_1(s_1) = P(S_1 \leq s_1), G_2(s_2) = P(S_2 \leq s_2), \cdots, G_H(s_H) = P(S_H \leq s_H)$, respectively, and joint distribution function $G(s_1, s_2, \cdots, s_H) = P(S_1 \leq s_1,\ S_2 \leq s_2, \cdots, S_H \leq s_H)$. The marginals map a vector of H random variables $(S_1, S_2, \cdots, S_H)$ to the point $[G_1, G_2, \cdots, G_H]$ in the H^{th} unit space, and to such point corresponds a real value in the unit interval $[0, 1]$. The function that maps from each point in the H^{th} unit space to a joint distribution function in the unit interval $[0, 1]$ is called copula.

A theorem of interest for us is:

Theorem 3.1 (Sklar's theorem) *Let G be an H-dimensional joint distribution function with distribution functions $G_1, G_2, \cdots, G_H$ and denote with $\boldsymbol{\varrho}$ the vector of parameters that measures the dependence between the marginals. Then, there exists an H-copula C such that $\forall\ (s_1, s_2, \cdots, s_H) \in \bar{\Re}^H$,*

$$G(s_1, s_2, \cdots, s_H) = C[G_1(s_1), G_2(s_2), \cdots, G_H(s_H); \boldsymbol{\varrho}]. \tag{4}$$

If $G_1, G_2, \cdots, G_H$ are all continuous, then C is unique; otherwise C is uniquely determined on $[RanG_1 \times RanG_2 \times \cdots \times RanG_H]$, where Ran denotes the range of the distribution function. Conversely, if C is an H-copula and $G_1, G_2, \cdots, G_H$ are distribution functions, then the function G defined in Eq. (4) is an H-dimensional joint distribution function with marginal distribution functions $G_1, G_2, \cdots, G_H$.

Sklar's (1973) theorem states that an $H-$dimensional copula is the function C that maps from the unit space $\Im^H$ to the unit interval $\Im$ and that satisfies the conditions:

1. $C(1, \cdots, 1, u_h, 1, \cdots, 1) = u_h \ \ \forall h \leq H \ \text{ and } \ u_h \in [0, 1]$;

2. $C(u_1, \cdots, u_H) = 0 \ \text{ if } \ u_h = 0 \ \text{ for any } \ h \leq H$;

3. C is H-increasing.

Hence, an H-dimensional copula is a multivariate distribution function that has all H one-dimensional marginals over the uniform distribution, $U(0, 1)$.

An important corollary of the theorem 3.1 sets out a method to construct copulas directly from joint distribution functions:

Corollary 3.2 (From joint distribution functions to copulas) *Let $G, C, G_1, G_2, \cdots, G_H$ be as in theorem 3.1, and let $G_1^{-1}, G_2^{-1}, \cdots, G_H^{-1}$, be inverse (quantile) functions of $G_1, G_2, \cdots, G_H$, respectively. Then for any point $(u_1, u_2, \cdots, u_H) \in \Im^H$,*

$$C(u_1, u_2, \cdots, u_H; \boldsymbol{\varrho}) = G\left[G_1^{-1}(u_1), G_2^{-1}(u_2), \cdots, G_H^{-1}(u_H)\right], \tag{5}$$

with $G_1^{-1}(u_1) = s_1, G_2^{-1}(u_2) = s_2, \cdots, G_H^{-1}(u_H) = s_H$. Corollary 3.2 is particularly useful for simulations.

Consult Appendix B for further notation and for a description of the parametric copula functions that are available for the multivariate case: $H > 2$.

The next section outlines the econometric methodologies we make use of.

4 Econometrics

4.1 Firm size measured by cumulative number of expansions

We specify the latent profit function for our dynamic model as the random linear dynamic equation

$$\begin{aligned} \pi_{it} &= \gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \tilde{\mathbf{z}}_{1i}\tilde{\boldsymbol{\lambda}}_1 + \tilde{u}_{it} \quad i = 1, \cdots, M; \quad t = \tau_i, \cdots, T, \\ \tilde{u}_{it} &= c_i + \varepsilon_{it} \end{aligned} \tag{6}$$

where $y_{i,t-1}$ is the lagged binary variable defined in Section 2. The observable explanatory covariates are decomposed into three components: the time-varying strictly exogenous variables, $\mathbf{x}_{1it}$, the time-varying endogenous variables, $\mathbf{x}_{2it}$, and the time invariant exogenous variables $\tilde{\mathbf{z}}_{1i}$. Similarly, the composite error term $\tilde{u}_{it}$ is decomposed into the unobserved MNE-level heterogeneity, c_i, and the time-varying function of the idiosyncratic error term, ε_{it}, which we assume to be identically distributed and independent of unobserved heterogeneity, and the observable covariates. The use of the $\tilde{}$ symbol shall be clarified in the remaining of the section. The exogeneity and endogeneity of the time-varying observable terms are the result of the three conditions $E\left(c_i|\mathbf{x}_{1it}\right) = E\left(c_i|\tilde{\mathbf{z}}_{1i}\right) = 0$ and $E\left(c_i|\mathbf{x}_{2it}\right) \neq 0$. We indicate with τ_i the period in which MNE i appears in the sample for the first time. So, for the balanced panel (the incumbents) we have $\tau_i = 1$ (and initial conditions at time $\tau_{i-1} = 0$) and for the new entrants $\tau_i > 1$. Our data share a pattern common in firm-level data: the number of MNEs M is large relative to the number of periods T, so asymptotics rely on $M \to \infty$.

The presence of a large cross-section in a nonlinear panel model rules out the possibility of modeling the c_i as parameters. In fact, because of the "incidental parameters" problem a fixed effects analysis would produce inconsistent estimates of the parameters, introduced by (Heckman (1981a,b). So, the rest of the section will treat the unobserved heterogeneity as a random variable.

What we observe in our data is not the latent random profit function shown in Eq. (6), but rather the binary outcome of an MNE expansion, whose relation with profit has been given previously in Eq. (3b). We assume the idiosyncratic error to be distributed as $NID(0, \sigma_\varepsilon^2)$ and given that y is a binary variable, we standardize the idiosyncratic error as $NID(0, 1)$. The implication is that all parameters, as well as the function of the unobserved MNE-level heterogeneity, will be re-scaled by σ_ε. Hence, the conditional probability that an MNE i expands in period t is

$$P\left(y_{it} = 1|y_{i,t-1}, \cdots, y_{i,\tau_{i-1}}, \mathbf{x}_{it}, \tilde{\mathbf{z}}_{1i}, c_i; \boldsymbol{\theta}\right) = \Phi\left(\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \tilde{\mathbf{z}}_{1i}\tilde{\boldsymbol{\lambda}}_1 + c_i\right), \tag{7}$$

where Φ is the standard normal cumulative density function, $\mathbf{x}_{it} \equiv [\mathbf{x}_{1it}, \mathbf{x}_{2it}]$ and $\boldsymbol{\theta}$ are all the parameters to be estimated. The joint conditional density for $(y_{i,\tau_i}, y_{i,\tau_i+1}, \cdots, y_{iT})$ results in the following dynamic unobserved effects Probit model,

$$P\left(y_{i,\tau_i}, y_{i,\tau_i+1}, \cdots, y_{iT}|y_{i,\tau_{i-1}}, \mathbf{x}_i, \tilde{\mathbf{z}}_{1i}, c_i; \boldsymbol{\theta}\right) = \prod_{t=\tau_i}^{T} \Phi\left[(\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \tilde{\mathbf{z}}_{1i}\tilde{\boldsymbol{\lambda}}_1 + c_i)(2y_{it} - 1)\right], \tag{8}$$

with $\mathbf{x}_i \equiv [\mathbf{x}'_{i1}, \mathbf{x}'_{i2}, \cdots, \mathbf{x}'_{iT}]'$.[9]

Because of data limitation our observable regressors are confined to include a minimal number of variables. The exogenous time-varying vector, $\mathbf{x}_{1it}$ is made only of time dummies: $\mathbf{x}_{1it} = \mathbf{x}_{1t}$. The endogenous time-varying vector is reduced to be the scalar variable lagged size: $\mathbf{x}_{2it} = s_{i,t-1}$. The time-invariant vector $\tilde{\mathbf{z}}_{1i}$ includes MNEs headquarter area dummies. The lagged size variable $s_{i,t-1}$ influences growth indirectly through determinants of extra profit, such as: scope economies, economies of scale in terms of production and R&D, better access to the demand or superior market power that can be gained as an MNE evolves.

The presence of unobserved heterogeneity makes the log-likelihood function of the above density not suitable to estimate the $\boldsymbol{\theta}$ parameters consistently, unless one has a way to integrate out the unobserved heterogeneity. In order to do so we need, first, to ensure that we account for any possible correlation between the unobserved heterogeneity and the regressors - given that not all our regressors, and surely not the lagged dependent variable, are orthogonal to the unobserved heterogeneity. Second, for the *balanced* sub-sample we must have a way to cope with the initial conditions problem, i.e., an existing relation between the initial observations of the dependent variable y_{i0} and the unobserved heterogeneity. The stochastic process that has determined an expansion in the initial period (period 0 in our notation), had been ongoing prior to that date and, as such, we cannot take it as exogenous. The initial conditions problem is particularly severe for small T.

Both of the above issues can be tackled. We make use of the Mundlak (1978) approach and account for the

[9]Eq. (8) combines the conditional probability rule with the assumption of sequential exogeneity. In case of three random variables $\{y_3, y_2, y_1\}$ the conditional probability rule writes the joint distribution can be expressed $P\left(y_3, y_2, y_1\right)$ as $P\left(y_3|y_2, y_1\right) P\left(y_2|y_1\right) P\left(y_1\right)$, and the sequential exogeneity assumption provides the simplification $P\left(y_3|y_2\right) P\left(y_2|y_1\right) P\left(y_1\right)$.

possible correlation between the unobserved heterogeneity and the subset of sequentially exogenous regressors. Hence we write the unobserved heterogeneity as:

$$c_i = \phi_0 + \overline{\mathbf{x}}_{2i}\phi_1 + \alpha_i, \tag{9}$$

and assume $E(\alpha_i|\mathbf{x}_{1it}) = E(\alpha_i|\mathbf{x}_{2it}) = 0$. We add to the time-invariant vector $\tilde{\mathbf{z}}_{1i}$ the Mundlak (1978) correction terms: the constant and $\overline{\mathbf{x}}_{2i}$ (which happens to be limited to $\overline{\mathbf{x}}_{2i} = \bar{s}_i$). We denote the new vector with $\mathbf{z}_{1i}$, postulate $E(\alpha_i|\mathbf{z}_{1i}) = 0$, and take note that the ˜ symbol is no longer part of the notation.[10] Hence, Eq. (9) relates the MNE unobserved profit heterogeneity to a linear function of the long run size. The coefficient ϕ_1 is expected to be positive, so that large MNEs are expected to generate a level of profits larger than that of small MNEs. The profit equation in (6) is upgraded to:

$$\begin{aligned} \pi_{it} &= \gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \mathbf{z}_{1i}\boldsymbol{\lambda}_1 + u_{it} \\ u_{it} &= \alpha_i + \varepsilon_{it}, \end{aligned} \tag{6a}$$

and obviously the probabilities in (7) and (8) have to be updated, accordingly.

Turning to the initial conditions problem, the panel data econometric literature has developed alternative ways to deal with it. Heckman (1981a,b) suggests recovering a full conditional density for $(y_{i0}, y_{i1}, \cdots, y_{iT}|\mathbf{x}_i, \mathbf{z}_{1i})$, by extending the original density from the initial period and integrating out the unobserved heterogeneity. To fulfill his idea he first specifies a parametric density for y_{i0} given $(\mathbf{x}_i, \mathbf{z}_{1i}, \alpha_i)$, thus giving the density to the initial period, and then a parametric density to integrate out the unobserved heterogeneity. The lack of an existing software code to estimate Heckman's full model has given rise to alternative estimation methods, such as: Orme (2001) and Wooldridge (2005). The former introduces an easy to implement two-step procedure, which is suitable to cases of low correlation between the initial conditions and the unobserved heterogeneity. The latter proposes parameterizing a conditional density for the unobserved heterogeneity only, so as to integrate out the unobserved heterogeneity, leaving the density for $(y_{i1}, y_{i2}, \cdots, y_{iT})$ conditional on $(y_{i0}, \mathbf{x}_i, \mathbf{z}_{1i})$.

As mentioned earlier, the initial conditions problem concerns only the MNEs that are in the sample the entire time period, the balanced panel, not those that enter at some time $\tau_i > 1$, the new entrants. In Appendix A we internalize this distinction in our modeling of probabilities for the two groups, and detail the necessary notation. Each of the three solutions can be estimated in `Stata` either directly, or via an add-on program called `gllamm`.

4.2 Other measures of firm size: $h \in (\mathcal{H} \setminus op)$

This section extends the relation between latent MNE profit and MNE growth to any other size metric. The latent profit equation expressed in (6a) becomes the metric-specific linear equation

$$\pi_{it} = \gamma_h y_{hi,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_{h1} + \mathbf{x}_{h2it}\boldsymbol{\beta}_{h2} + \mathbf{z}_{h1i}\boldsymbol{\lambda}_{h1} + u_{hit}. \tag{6b}$$

We must now carefully discriminate between discrete and continuous size metrics, for the former will lead to nonlinear growth dynamic equations and the latter to linear growth dynamic equations. We postulate that a mapping from the profit CDF to the growth variable exists for each of the size metrics.

If the size metric is continuous we expect the composite function not to be surjective, but bijective. We can directly estimate the MNE growth linear relation

$$y_{hit} = \gamma_h y_{hi,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_{h1} + \mathbf{x}_{h2it}\boldsymbol{\beta}_{h2} + \mathbf{z}_{h1i}\boldsymbol{\lambda}_{h1} + u_{hit}, \tag{6c}$$

being aware that the coefficients and error term have to be suitably rescaled to be comparable to those in Eq. (6b). On the other hand, if the size metric is discrete the equation of interest is, as mentioned in Section 2, $y_{hit} = f_h[G_h(\pi_{it})]$.

[10]The term $\bar{s}_i$ is computed as

$$\bar{s}_i \equiv \frac{1}{T - \tau_i + 1}\sum_{t=\tau_i}^{T} s_{i,t-1}. \tag{10}$$

.

The main econometric consequence of this particular demarcation between continuous and discrete size metrics is that we shall employ linear panel data techniques to estimate growth dynamics for continuous size metrics, and nonlinear econometric methodologies to estimate growth dynamics for discrete size measures. While both types of equations (linear and nonlinear) will require the Mundlak correction, only the nonlinear dynamic equations will necessitate the revision to correct for the initial conditions problem. Consequently, the discrete metric, cumulated number of subsidiaries, will involve a procedure that is much like that presented in the previous section, the only difference being that this time a different parametric distribution for the error term is assumed.

5 Data and Descriptive Statistics

We gathered information on pharmaceutical MNEs from a commercial database called *Amadeus*, which is published by Bureau van Dijk (BvD) Electronic Publishing. The publisher collects firms' account data from official or commercial sources of individual European countries and processes the data in order to achieve maximum comparability across countries. The database contains balance sheets, profits and losses tables and other information, among which is business activity, date of incorporation, location, ownership, etc. This database is by far the most comprehensive source of financial information on European firms.[11] A firm is recorded by *Amadeus* as an EU pharmaceutical MNE, if at some time t it has pharmaceutical-related subsidiaries (either production, research or marketing driven) in more than one country, and if at least one of these countries belongs to the EU-14 (the EU-15 minus Luxembourg): Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden, and the UK.[12] We examine the pharmaceutical business. We calculate the various measures of size for each MNE by summing the corresponding metrics of its pharmaceutical-related subsidiaries. Some of the large pharmaceutical firms listed in KPMG's report of the top 100 pharmaceutical companies have large non-pharmaceutical business; examples are: Bayer and Akzo Nobel. However, by summing the measures of each pharmaceutical-related subsidiary we are able to concentrate on the pharmaceutical part of the business. We treat acquired and new own-established subsidiaries in the same way, as we believe that both greenfield investments and acquisitions reflect MNEs' willingness and action to expand. Expansions through both platforms are influenced by industry-level trends, such as relocation of business activities or adjustments to the Single Market Programme.

The starting year of the sample period is chosen to be 1992 ($t = 0$ in our notation), which is just one year prior to the implementation of the European Single Market Programme, occurring on the first day of 1993. The last year in the period is 2004 (T in our notation). As mentioned in the introduction section, information on firm exit is not reliable in *Amadeus*. Also, mergers complicate the identification of the MNEs in the sample, as the *Amadeus* database does not report the MNEs that have been acquired, and only records the acquirer. This makes it impossible to construct a perfect panel of expansion history. To get around this difficulty, we adopt Bottazzi and Secchi's (2005; 2006b) procedure, which treats the merged enterprises as single entities throughout the period of investigation.

The maximum number of MNEs in the dataset is 268. For these 268 MNEs we count 624 new subsidiaries established or acquired and 974 existing subsidiaries, for a total of 1 598 subsidiaries. A more detailed description of the construction of the MNE panel dataset and the issue of missing values is described in Appendix C.

Given that our paper extends the study of firm growth to comparing alternative metrics of size, in the rest of this section we discuss the behavior of each measure that our database allows for. We directly use the information on subsidiaries to generate an initial discrete measure of MNE size: the number of cumulated subsidiaries acquired over time. We indirectly utilize the data on subsidiaries to construct a second discrete variable of MNE size: the number of cumulated opportunities, *alias* expansions, captured over time. So, the discrete size metrics we make use of, are:

1. $s_{(op)it}$ is the stock of business opportunities captured by the MNE i up to period t. It is our pillar measure of size, that we use to investigate the dynamics of MNE expansion, upon which we simulate the dynamics of

[11] At the time when we begun preparing the sample of the pharmaceutical industry, year 2004, *Amadeus* covered approximately 11 million public and private firms in 41 European countries.

[12] Following the EU statistical classification of economic activities, NACE Revision 1.1, pharmaceutical-related subsidiaries are classified as: 2 441 (manufacture of basic pharmaceutical products), 2 442 (manufacture of pharmaceutical preparations), 5 146 (wholesale of pharmaceutical goods), 5 231 (dispensing chemists), 5 232 (retail sale of medical and orthopaedic goods) and 7 310 (research and experimental development on natural sciences and engineering).

the other metrics. We have chosen this as pillar metric because it is by its nature the most discrete metric of size available in our data.

2. $s_{(su)it}$ is the total number of pharmaceutical-related subsidiaries that the MNE i has established or acquired up to year t.

Furthermore, in our dataset we have access to the following continuous size metrics:

3. $s_{(em)it}$ is MNE i's number of employees in period t.
4. $s_{(or)it}$ is MNE i's operational revenues (turnover) in period t.
5. $s_{(sa)it}$ is MNE i's sales in period t.
6. $s_{(ta)it}$ is MNE i's total assets in period t.

Table 2 discloses valuable information on each of these metrics. The first row displays several statistics related to the expansion process, $y_{(op)}$. We observe that the average growth of opportunities captured by the cross-section of MNEs, i.e., the share of MNEs that have expanded in a given period, ranges between 12 and 20 per cent most of the time, before shrinking to lower values after 2001. The second row gives the statistics for the stock of opportunities, $s_{(op)}$. We note that the French company Sanofi Aventis (S.A.) holds the largest number of cumulated opportunities, a number that grows over time from 26 to 33. The MNEs' average number of cumulated opportunities captured in the European pharmaceutical market increases about two-thirds of a unit during the period of investigation, a trend that comes along with a rise in the dispersion. Skewness does not follow the same trend, as it slightly contracts. In the next section we will investigate this particular trend. A similar set of statistics is offered for the number of subsidiaries. The number of subsidiaries is an important measure of MNEs size during the period of EU enlargement: strategic entry in the various European countries to access the market, along with the introduction of greenfield parks and tax credits, boosted the number of subsidiaries. However, one has to be aware of major distinction between internal and external/acquisition growth. A company can either expand the existing subsidiaries (by hiring additional employees, increasing the turnover of the same subsidiaries, etc), internal growth, or and the case of interest for us during the EU enlargement, where growth is mainly external and rather than occurring within the exiting subsidiaries, it occurs increase the portfolio of subsidiaries, external/acquisition growth. Our analysis focuses on the latter, and the existence of a strong association between the discrete measures of size and the continuous will provide reliability to our choice. We discuss the number of subsidiaries. We notice that the largest growth occurs in year 2000 when an MNE establishes or acquires 17 subsidiaries. Year 2000 is certainly a boom year, as it displays the largest increase in the market, with 78 new subsidiaries attained in total. Average growth in number of subsidiaries shares an inverse U shape with average growth in number of cumulated opportunities. What differs here is the dispersion, marked by a positive trend with a rather oscillating behavior. If we convey our attention to the stock of subsidiaries, $s_{(su)}$, we witness the dominance of the largest company Sanofi Aventis, which steadily grows over time, reaching a count of 61 subsidiaries after 2000. MNEs hold, on average, 4.75 subsidiaries in 1992 and that number grows to 5.96 in 2004. The inequality in subsidiary ownership increases over time, as indicated by the upward movement of standard deviation and to a certain extent skewness.

Table 2 additionally provides statistics for our continuous metrics of size. Here, the quality of the *Amadeus* dataset limits the time span of availability to the sub-period 1995-2004, as too many missing values were present prior to 1995. Missing values were also present post 1995, but to a much lower extent. Take the stock of employment for example, it starts with 76 observations (out of 228) missing in 1995 and drops to 10 observations (out of 268) missing in year 2004. Operational revenues and total assets lead to a similar pattern of missing observations. The remaining metric, sales, displays a heavier burden of missing data *and because of that* we will exclude it from our analysis. All statistics for operational revenues, sales and total assets, aside from skewness, are expressed in millions of dollars. While Sanofi Aventis is the leading MNE in terms of employment, operational revenues and sales, British company GlaxoSmithKline holds the largest total assets throughout the entire period and the highest employment in 1995. The average values of operational revenues, sales and total assets, as well as their dispersions, grow over time and are almost doubled by 2004. On the other hand, the growth of average employment is almost unnoticeable.

We provide further evidence on the behavior of the five size metrics by inspecting their distributions. We plot the empirical cumulative density functions (with each size measure normalized in the unitary interval) for the periods 1995 and 2004 in Figure 1. The skewness towards small firms is striking. By comparing the two time periods we notice a mild change in skewness for the discrete metrics, confirming the figures displayed in Table 2. We have highly emphasized the role of skewness, as this statistic has been at the core of study of the Growth-of-Firms literature.

Finally, Table 2 also unveils information on the break-down of the number of MNEs in our sample by geographical location of the headquarters (EU, US, or Other) and displays valuable statistics on the number of new entrants. We note that the number of new entrants reaches a peak of 9 MNEs in 1998 and decreases thereafter, vanishing by the end of the period in 2004. Relative to the full sample of MNEs, new entrants play a very modest role, never going beyond 3.6 per cent. The fact that this particular industry displays "weak" dynamics of entry is convenient here, because it allows one to give less weight to the aforementioned concern that Gibrat's law might be rendered invalid in light of excessive entry and exit, recalling that exit is not accounted for in our dataset.[13] Limited entry will also make evaluation of the theory of maturation more sound. The last line of the table reports the proportion of opportunities captured by new entrants. We observe an oscillating trend that goes from 0 to 21 per cent.

The cross-analysis of the descriptive statistics in Table 2 and the visual inspection of Figure 1, convey a general message. With the exception of skewness, all measures of size display the same sign in their trends. Skewness remains rather stable for the number of cumulated opportunities, heavily drops for the total assets and rises for the other metrics.

If we were to draw conclusions from the statistics that we have just discussed, we would state that all variables, to a certain extent, co-move but the degree of co-movement requires further investigation. In order to produce a more reliable study of the cross-relation of the variables, we appeal to the copula approach. If we can show with this method that all the metrics of size closely concord, then it would be enough to analyze the growth dynamics of any metric, and in that case, why not choose the simplest (or the most easily accessible) one as the pillar? The answer to this question is given in the next section.

6 Results

6.1 Step-One

6.1.1 Size Metrics Association

As already mentioned in the previous sections, the way to link a mixture of discrete and continuous variables or, generally speaking, variables that come from different distributions, is to utilize a multivariate distribution that has marginals uniform over the interval [0,1], i.e., to employ a copula. Prior to discussing the results obtained from applying such a technique to our data, we warn the reader to be cautious in running this methodology with discrete data, for as we detail in Appendix B, Sklar's theorem may no longer provide uniqueness. Genest and Neslehova (2007) enumerate a complete list of things that can go wrong when the copula links discrete distributions. Their criticism is not a severe concern here, for our discrete variables rely on a large number of integers with positive support, as confirmed in the empirical distributions plotted in Figure 1, up to the point that one could label our discrete measures of firm size as quasi-discrete metrics.

In order to estimate the copula, we first fit the parametric distributions of the alternative (discrete and continuous) measures of firm size. Table 3 reports, for each time period, the estimated parameters of distributions with positive support. We investigate four discrete distributions - Exponential, Geometric, Negative-Binomial and Poisson - and four continuous distributions - Lognormal, Exponential, Gamma and Weibull. Most of the distributions depend on one parameter, except for the Negative-Binomial, Lognormal and Gamma, which rely on two.[14] Because our ultimate goal is to fit the copula, we fit the distributions of the various metrics of firm size limited to the subperiod 1995-2004, for the data on continuous variables present too many missing values prior to 1995. In addition, because our discrete metrics have a rich domain, for each of them we also fit the most natural continuous distribution (the lognormal) and insure against treating as discrete, variables that should be treated as continuous. The implication is that for the two discrete size metrics in our sample we will fit five distributions.

The first panel of Table 3 investigates size measured by number of cumulated opportunities. It is the Geometric

[13] We still test its validity in the Results section.

[14] We have utilized a library available in R to fit the distributions. Ricci (2005) gives details on the procedures adopted.

distribution that gives the best fit, as suggested by the p-value of the Pearson χ^2 goodness-of-fit statistic. The same parametric distribution also provides the best approximation to size measured by the number of cumulated subsidiaries, though in this instance there is, in some cases, evidence/strong-evidence against the null hypothesis that the observed and expected frequency come from the same probability distribution. Yet, the geometric distribution is the one that provides the best fit among the selected distributions. As for the three continuous metrics, we select the Lognormal distribution as that giving the best fit, but remark that for operational revenues and total assets there are periods where there is evidence and strong evidence against the null hypothesis. The last column of Table 4 is for the pooled sample. Not surprisingly, due to the large number of observations there is strong evidence against the null hypothesis for all metrics but employment. The rule that we apply to the pooled panel, is to retain as best parametric distribution the one selected during the period-by-period analysis (the Geometric distribution) for the discrete metrics and the Lognormal for the continuous metrics.

With parametric distributions on hand for each time period we have all the information necessary to fit a parametric copula. We utilize the package `copula` written in `R` and fit, by maximum likelihood the Elliptical (Normal) and Archimedean (Clayton, Frank and Gumbel) popular classes.[15] Table 4 reports the estimated parameters, which are: the marginal parameters, and either the concordance coefficients for the Elliptical copula or the copula parameter for the Archimedean copulas. The arguments of the likelihood maximization are: the concordance coefficients, the copula parameter and the new marginal parameters. We compare the log-likelihood values and subsequently choose as best fitting copula the Elliptical Normal copula.

Relevant to the first aim of this article is the evidence of strong positive association between the five variates, as confirmed by the high positive values of the Spearman's ϱ estimated parameters. The variable total assets displays the lowest concordance with the number of cumulated opportunities, but that lowest level raises steadily over time, up to reaching a fair degree of positive association in the latest periods. We interpret the results as evidence of the relevance of external growth in the industry expansion process.

6.1.2 Conditionally Simulated Size Metrics

We simulate the size vector $\tilde{\mathbf{s}}_{-[h]}$ from the conditional distribution $G(\cdot|s_h)$, where s_h denotes the observed values for the selected size metric h (*op* in our case) and the tilde vector $\tilde{\mathbf{s}}_{-[h]}$ indicates the simulated values for the remaining size metrics. The simulated values are compared with the corresponding observed values and for each metric, and time period, we evaluate the reliability of the copula. In order to set the accuracy of the employed copula we simulate, for each time period and metric of size, 1000 observations.[16] Each period, of the simulated 1000 vectors of size metrics, we retain the M_t ones that have the element simulated number of cumulated opportunities closest to the true observed values, that is, we condition the copula simulations on the observed values of size measured by the number of cumulated opportunities.[17] We repeat this procedure $bs = 1000$ times. Table 5 displays two statistics on the goodness-of-fit of the copula, inclusive of a 95 per cent empirical confidence interval. In its first column we have the average (over the bs repetitions) simulated coefficient of determination, which tells us that conditioning on the number of cumulated opportunities, the copula is, on average, able to explain over 70 per cent of variation in the number of subsidiaries variation over time, between 19 and 33 per cent of employment and operational revenues, and a range between 4 and 19 per cent of total assets. The coefficients of determination are highly concentrated, as revealed by the 95 per cent empirical confidence intervals. Most importantly, to consolidate the estimated copula Table 5 displays the average p-value of the Pearson χ^2 test for the equality of simulated distributions and the empirical distributions observed in our data. With a probability value of 5 per cent we do not reject the null hypothesis, on average. The coefficient of determination and the Pearson χ^2 test suggest together that we have a good methodology to move from the most discrete metric to the others.

At this point, with a preferred metric on hand, and a consolidated mapping of metric to metrics at our disposal, we proceed to investigating the expansion dynamics of the European pharmaceutical MNEs.

[15]Consult Yan (2007) and Kojadinovic and Yan (2010) for a detailed exposition of the empirical technique.

[16]We employed the `mvdc` function, which is part of the `copula` package in `R`, to obtain our simulations. The `mvdc` function requires as inputs the parametric distributions for all variables (size metrics) and their estimated parameters.

[17]In case of multiplicity of simulated observations close to a certain value of the observed metric, we randomly pick one of them.

6.2 Step-Two

6.2.1 Forecasted Pillar Size Metric

Given the evidence of a positive relation between the various metrics of size highlighted in previous sections, and given that size has been key element in the firm size theory proposed first by Simon and Ijiri (1977), and reviewed by Sutton (1997b) later, we are happy to choose the most discrete metric of firm size, which is the number of cumulated expansions/opportunities, as pillar metric upon which to study the firm growth dynamics, but first we provide a brief recap. In Section 2 we have related the profit function to a functional form of any metric of lagged size and observable and unobservable shifters. We have proposed a linear relation in Eq. (6). In the econometrics section 4.1, we have set out the conditions required to estimate the parameters of a dynamic version of the binary relation expressed in Eq. (3b). We have warned about the issue of possible correlation between the unobserved heterogeneity and a subset of regressors that include the lagged dependent variable and the endogenous variables. We have suggested the Mundlak correction, as a solution. Now, we evaluate the benefit of introducing this correction term, in either a pooled dynamic Probit model or a random effects dynamic Probit model. The estimation results are documented in Table 6. The introduction of the correction term, on the one hand invigorates the impact of state dependence, on the other hand it produces a significant mean reversion effect on growth, and brings in rich trajectories of growth that depend, all else equal, on how far away the MNE is from its exogenous long-run size. Another interesting result is the reduced weight that unobserved heterogeneity carries, after the inclusion of the Mundlak correction term in the random effects estimator; as confirmed by the lower values of the estimated coefficient of intertemporal correlation, $\hat{\rho}$. A final important finding is the lack of predictive power of the specifications without the Mundlak adjustment. In particular, it strikes us to see how unsuccessful the estimates without the correction term are in predicting the positive outcomes of expansion. Upon the introduction of the correction term percentages are brought up to almost 17 per cent, perhaps this is not so surprising, given that the binary dependent variable displays zeros 84 per cent of the time. Now, we further analyze the results inclusive of the Mundlak correction. We begin our discussion with the pooled Probit estimation. The point estimate of the coefficient associated with the lagged dependent variable $y_{i,t-1}$ is positive and statistically significant at the 1 per cent level, providing evidence of persistence. If an MNE has expanded in the previous period it faces, *ceteris paribus*, a higher probability of expanding in the current period. Such an effect can be either reinforced, reduced or wiped out, by the combined estimated effects of average size $\bar{s}_i$ (a variable that ranges in the interval of $[0.5, 23.6]$) and lagged size $s_{i,t-1}$ (an integer variable that ranges in the interval of $[0, 24]$). The coefficients of $\bar{s}_i$ and $s_{i,t-1}$ are close to one another, but have an opposite sign. This implies that if a firm is at its early stage of growth, with size below its average long-run level, the overall effect on size growth is large at the beginning. As size reaches its average level, the positive boost on growth reduces, and the effect fades after size overtakes that average value; in other words, we find support for Rostow's theory of maturation. The area dummies are non-significant, suggesting that the physical location of the MNE headquarters gives no comparative advantage on growth. This is an effect that we believe is triggered by the strong deregulation that took place during the Single Market Programme era.

Our second estimate is the random effects estimator. The point estimate of the parameter associated with $y_{i,t-1}$ is smaller than that of the aforementioned pooled Probit model, though the probability distributions of the two estimators are not significantly different from one another at the 10 per cent level. The non-significant discrepancy between the two estimates is easily understood if one accounts for the different normalization of the error term, which imposes $\sigma_u^2 = 1$ in the pooled Probit and $\sigma_\varepsilon^2 = 1$ in the random effects estimation. The random effects coefficients need to be multiplied by a factor of $\sqrt{1-\rho}$ to be comparable to those of the pooled Probit (see Arulampalam (1999)), with ρ denoting the correlation between the composite error in two time periods.[18] However, in our context ρ is estimated to be very low at $\widehat{\rho} = 0.071$, suggesting that one has a good enough approximation if the effect of the normalization is disregarded; this makes the two sets of results directly comparable.[19] Although not too high, we note that the estimated coefficient $\hat{\rho}$ is significant at the 5 per cent level, which confirms our suspicion that some intertemporal correlation exists between the composite errors u_{it} and u_{is}, with $s \neq t$.

To deepen our understanding of the effect of state dependence, we calculate the average partial effect (APE) from the counterfactual outcome probabilities that rely, for all i and t, on the two extreme states: complete expansion,

[18] $\rho \equiv Corr\left(u_{it}, u_{is}\right) = \frac{\sigma_\alpha^2}{1+\sigma_\alpha^2}$, for $s, t = 1, 2, ..., T; s \neq t$. As ρ is bounded between 0 and 1, $\sqrt{1-\rho}$ is smaller than 1.

[19] Because $\sqrt{1-\hat{\rho}}$ is ≈ 1.

$y_{i,t-1}=1$, denoted below with subscript (1), and absence of expansion, $y_{i,t-1}=0$, denoted with subscript (0). In the random effects model we compute the sample average growth in the two states, as follows:

$$\begin{aligned}\widehat{y}_{(1)} &= \frac{1}{M}\sum_{i=1}^{M}\frac{1}{T-\tau_i+1}\sum_{t=\tau_i}^{T}\int\Phi\left[\left(\widehat{\gamma}+\mathbf{x}_{1it}\widehat{\boldsymbol{\beta}}_1+\mathbf{x}_{2it(1)}\widehat{\boldsymbol{\beta}}_2+\mathbf{z}_{1i(1)}\widehat{\boldsymbol{\lambda}}_1+\widehat{\sigma}_\alpha\alpha\right)\sqrt{1-\widehat{\rho}}\right]\phi(\alpha)d\alpha \\ \widehat{y}_{(0)} &= \frac{1}{M}\sum_{i=1}^{M}\frac{1}{T-\tau_i+1}\sum_{t=\tau_i}^{T}\int\Phi\left[\left(\mathbf{x}_{1it}\widehat{\boldsymbol{\beta}}_1+\mathbf{x}_{2i(0)}\widehat{\boldsymbol{\beta}}_2+\mathbf{z}_{1i(0)}\widehat{\boldsymbol{\lambda}}_1+\widehat{\sigma}_\alpha\alpha\right)\sqrt{1-\widehat{\rho}}\right]\phi(\alpha)d\alpha \end{aligned} \quad (11)$$

where the term $\sqrt{1-\widehat{\rho}}$ is used to make the random effects APE comparable to the pooled Probit.[20] The subscript (1) given to the time-varying endogenous variable (which we recall to be $s_{i,t-1}$) defines the relation $\mathbf{x}_{2it(1)}\equiv s_{i0}+t-1$, and because of the Mundlak correction element the same subscript is extended to the time-invariant variables. Similarly, the subscript (0) given to the time-varying endogenous variable defines $\mathbf{x}_{2i(0)}\equiv s_{i0}$, with obvious implications for the Mundlak correction element of the time-invariant vector. The APE is computed as the difference between $\widehat{y}_{(1)}$ and $\widehat{y}_{(0)}$. The results of the APE are reported in percentages at the bottom of Table 6. The APE of the pooled Probit model is larger than that of the random effects model by 1.3 percentage points, suggesting a minimal "spurious" state dependence.

We have earlier highlighted the problem of possible correlation between the initial conditions and the unobserved heterogeneity. In Appendix A we present and review solutions to this initial conditions problem, and set out several estimation techniques that deal with it directly in `Stata`. In the same appendix, we show that all of the solutions, but the "Wooldridge" one, require an initial conditions equation. We specify such an equation to depend on the constant, area dummies of the headquarters, the pre-sample information on age and the long-run average number of opportunities. We employ the pre-sample MNE age as recorded in 1991 (t=-1 in our notation), $age_{i,91}$, which is believed to be exogenous once lagged size, lagged expansion and average size have been accounted for. We do not model entry but treat it as exogenous, that is, as the result of a stochastic process believed to be independent of the conditional stochastic process that has determined the outcome of growth. There is a chunk of literature that disagrees with this view and thus models the full joint distribution of age and size. Theoretical foundations are given in Evans and Jovanovic (1989) and Cooley and Quadrini (2001). A non exhaustive list of empirical studies that have tested this joint relation lists: Evans (1987a,b), Hall (1987), Dunne et al. (1989), Farinas and Moreno (2000), Yasuda (2005), Oliveira and Fortunato (2008) and Huynh and Petrunia (2010).

An exhaustive comparison of alternative techniques to estimate a dynamic Probit model with unobserved heterogeneity is offered in Table 7. All estimation methods use the Mundlak correction term. We distinguish between equicorrelation and free-correlation. Such demarcation is determined by the way that we treat the unobserved heterogeneity. We have free-correlation if the unobserved heterogeneity is multiplied by time-varying parameters, $\delta_t\alpha_i$, and have equicorrelation if we fix $\delta_t=1$ for all $t=1,2,\cdots,T$ and $\delta_t=\delta_0$ for $t=0$. Among the competing estimation methodologies that we consider only Orme (2001) accommodates for both models. We start our discussion of results with Heckman's solution. Two alternative econometric estimation techniques are considered: Arulampalam and Stewart (2009) and Stewart (2006).[21] We estimate the simplified initial conditions equation, whose notation is deferred to Eq. (15a) of Appendix A. Controlling for average size in the initial conditions equation makes the *additional* exogenous variable, age, non-significant. The coefficient of intertemporal equicorrelation, ρ, is still low and significant at 5 per cent using Arulampalam and Stewart's (2009) procedure, but is nearly doubled and highly significant using Stewart's (2006) procedure. The discrepancy is induced by the different integration methodologies adopted by the competing methods, leading to different figures of the intertemporal correlation explained by the unobserved heterogeneity: 6 versus 14 per cent. Wooldridge (2005) offers a second approach to solving the initial conditions problem. Again, we restrict our interest to the case of equicorrelation. The results are documented in column three. We now have a new coefficient that is associated with MNE initial expansion. It has a positive parameter significant at 10 per cent, hinting that initial expansion produces a mild long-run effect on growth. The last attempt to solve the initial conditions problem is offered by Orme's (2001) two-step estimator, which is suitable to environments with weak correlation. His methodology brings in a new variable: the inverse Mill's ratio, *imr*.

[20] Thus, this is a term that at low values of $\hat{\rho}$ we suggest can be disregarded.

[21] Due to a failure of convergence of the estimation algorithm, we do not report or discuss the free correlation cases related to these two estimation techniques.

We start by explaining the equicorrelated estimation that is reported in column four. The inverse Mill's ratio is significant at 1 per cent, and the remaining parameters are very much in line with the estimates indicated by the previous techniques, so we omit discussing them again. Our last estimation, again using Orme's technique, is dedicated to the case of free correlation, which only imposes the restriction $\delta_T = 1$ and leaving all remaining δ_t parameters unconstrained. Table 7, column five, reports the estimates of the unconstrained coefficients δ_t, but none of them comes in significant, so the rest of the paper will concentrate only on equicorrelation.

We select as the best estimation technique, the one that predicts correctly the highest percentage of expansions. This is Arulampalam and Stewart's (2009) methodology, with an accuracy over 20 per cent. The results that are going to be discussed in the rest of the section will rely entirely on this estimation.

In addition to the analysis of state dependence discussed above, the effect of MNE size on expansion merits closer examination. We are interested in answering the following questions:

1. Does initial size matter in explaining MNEs' growth?

2. Do MNEs grow along different paths according to their initial sizes?

3. Does Gibrat's law hold for our richer functional form of size?

In order to respond to these questions we assign the MNEs to five size strata based on their initial size, $s_{i,91}$.[22] For each stratum, the average expansion is estimated in each period t as:

$$\widehat{y}_{rt} = \frac{1}{M_{rt}} \sum_{i \in \mathcal{M}_{rt}} \int \Phi \left[\left(\widehat{\gamma} + \mathbf{x}_{1it}\widehat{\boldsymbol{\beta}}_1 + \mathbf{x}_{2it}\widehat{\boldsymbol{\beta}}_2 + \mathbf{z}_{1i}\widehat{\boldsymbol{\lambda}}_1 + \widehat{\sigma}_\alpha \alpha \right) \sqrt{1-\widehat{\rho}} \right] \phi(\alpha) d\alpha \quad (12)$$

$$r = \{NE, 1, 2, 3, 4\}; \quad t = \{\tau_i, \tau_i + 1, \cdots, T\}. \quad (13)$$

where $\mathcal{M}_{rt}$ is the set of MNEs in the rth stratum that are active in period t and M_{rt} is its number of elements.

Figure 2 compares the predicted and observed average expansion (and its cumulated version) of the growth and proportional growth over time, and by strata for the period 1992-2004. Sub-figure 2(a) plots the average expected expansion, i.e., the average probability of expansion by strata. The figure sheds light on the first question. If we momentarily exclude the new entrants (NEs) from the analysis, we observe at the beginning of the period a clear pattern of positive monotonicity between initial size and growth. For example, by looking at stratum 4, which includes the largest MNEs in year 1992, we note that in the mid nineties its MNEs had more than double the average probability of success, relative to the MNEs belonging to the second largest stratum, stratum 3. This monotonicity vanishes over time, up to the point that a common converging path arises for all strata after year 2000. By the end of the period, year 2004, the average probabilities of expansion are clustered at a level of approximately 10 per cent. During the period the largest MNEs have lost a good bit of their positive momentum. Until now, we have purposely excluded from the discussion the role of new entrants, as these are known to have an atypical pattern that can undermine Gibrat's law, as demonstrated in Simon and Ijiri (1977).[23] From Sub-figure 2(a) we observe that the stratum that includes the new entrants carries, on average, a higher probability of success than stratum 1 and 2, and in certain years even stratum 3. More importantly, the NEs stratum always exhibits a higher probability of expansion than stratum 1, indicating that a portion of new entrants moves away from the lowest part of the distribution over time, a trend that is confirmed in Ijiri and Simon (1964). The fact that all strata converge to a similar value of growth by the end of the time period, can certainly be explained by the Mundlak correction factor which, being significant, introduces a mean reversion effect in the growth dynamics.

Sub-figure 2(c) is the cumulated version of Sub-figure 2(a). The particular trajectories of expected cumulated expansions that are graphed in the figure can be used to answer the second question posed above. Here the strong role of initial size is neatly visible. By the end of the period of investigation, year 2004, we find that an MNE from the largest stratum (stratum 4) is, on average, expected to gain about 3.5 new opportunities, versus the 1.75 opportunities of an MNE that belongs to the second largest stratum (stratum 3). A similar comparison could be

[22]Because of the discrete nature of $s_{i,91}$, each size stratum does not have an equal number of MNEs in it. The first stratum includes the MNEs that have captured only one opportunity up to year 1992 (27.2 per cent of the sample); the second stratum those that have $s_{i,91} \in \{2, 3\}$ (20 per cent of the sample); the third stratum have $s_{i,91} \in \{4, 5, 6\}$ (12.8 per cent of the sample); and the fourth stratum with $s_{i,91} \geq 7$ (14.7 per cent of the sample). A final stratum contains the subsample of new entrants (25.3 per cent of the sample).

[23]Even though new entry does not play a major role in this industry as highlighted in Table 2.

extended to the remaining strata. Strata with different initial sizes seem to converge to different steady state levels, where the convergence to a steady state here is suggested by the flattening of the paths toward the end of the period. Abstracting from time effects that illustrate yearly changes common to all MNEs, the strict concavity in the paths, followed by convergence to a steady state (more marked in the largest stratum (stratum 4)), could be signalling underlying diseconomies of scale, which might be induced by rising managerial costs, by transaction costs in large enterprises or, most plausibly, by the maturation of the European Single Market Programme. Alternatively, the convergence in the steady state may signal the exhaustion of external growth push and the beginning of internal growth.

We compare the predicted average growth plotted in Sub-figures 2(a),(c) with the average observed growth graphed in Sub-figures 2(b),(d). With the exception of NE (that show predicted values larger than observed values), the predicted values capture the trends of the observed values well.

In order to answer the third question, i.e., to have a proper say on Gibrat's law, we look at growth defined in proportional terms. Sub-figures 2(e),(f) graph this information for us. We analyze Sub-figure 2(e). Not only is Gibrat's law violated for the stratum of new entrants, a result that confirms the empirical findings of Lotti et al. (2003) and Calvo (2006), but its violation is also extended to the stratum of small MNEs that are part of the balanced aspect of the panel. This is a breach that might happen to be worsened if we had data on MNE exit.

We have largely investigated the dynamics of MNE growth under the number of cumulated opportunities metric. We wish now to study the dynamics of the remaining metrics. We begin with the equation-by-equation analysis.

6.2.2 Equation-By-Equation Forecasted Size Metrics

One way to proceed when various metrics of firm size are available is to estimate the growth dynamics of each individual metric, equation-by-equation. This is the procedure pursued by Kumar (1985). We enrich Kumar's approach by replacing his biased Least Squares estimator with more appropriate estimators that cope with dynamic panels with unobserved heterogeneity. We distinguish between discrete and continuous metrics because the former requires linear estimators, the latter nonlinear estimators. In our data the only additional (discrete) size metric that requires a nonlinear estimator is the cumulative number of subsidiaries captured. We model the growth dynamics of this variable as a random coefficient dynamic Poisson model, and deal with exactly the same set of issues that we have previously laid out for the dynamic Probit procedure: the initial conditions problem and the correlation between the unobserved heterogeneity and the explanatory variables. As for the remaining (continuous) metrics, their growth dynamics can be studied with a dynamic linear panel approach that accounts for the unobserved heterogeneity. As we wish to identify both time-varying and time-invariant effects, we choose as estimation procedure an extension of the Hausman and Taylor (1981) static panel model. We employ a two-stage technique. In the first stage we utilize the Arellano and Bond (1991) `xtabond` Stata module and estimate the time-varying variables, inclusive of the lagged dependent variable. In the second stage we estimate the time-invariant effects and correct the standard errors using the procedure suggested in Kripfganz and Schwarz (2012).

The equation-by-equation estimates of MNE size growth dynamics are presented in Table 8. The first column refers to the discrete cumulated number of subsidiaries. At the time of this writing, the only dynamic Poisson random effects estimator implemented in `Stata` is the one that relies on the Wooldridge methodology, and therefore it is the estimation technique that we use. Nevertheless, we compare the estimates in column one with those displayed in the Arulampalam and Stewart (2009) methodology of Table 7, that is, we confront the growth dynamics of cumulative number of subsidiaries and the cumulative number of opportunities, and notice that the persistent effect of lagged growth, lagged size and the long run average size all have the same sign and are statistically significant. Given that the error terms in the latent profit equations come from different probability distributions, we can only compare the ratio of average partial effects between two continuous variables. The ratio between the partial effect of lagged size and that of average size is estimated to be -0.94 for cumulated number of expansions, and -0.86 for cumulated number of subsidiaries. We extend the comparison to the continuous metrics and find the ratios to be -0.82, -0.69, -1.29, for employment, operational revenues, total assets, respectively. The estimated ratios have the same sign and are close in magnitude. Only total assets has a ratio below minus one, demonstrating another deviation from all other metrics. We read this result as signal that the MNEs size metrics display similar growth dynamics; a result that reinforces the findings of the copula approach discussed earlier.

The equation-by-equation estimated growth dynamics are then used to forecast the size dynamics, which we will use in the next section to evaluate the quality of the estimates and compare those with the simulations of size derived from the copula.

6.2.3 Conditionally Simulated Forecasted Size Metrics

What the equation-by-equation method has not allowed for is how to move from one size metric to another. One question that motivates this research is centered on the existence of a relation between expansions and other measures of growth, and therefore between the cumulated expansions and other size metrics. In Section 6.1.1 we have studied the population association between the various size metrics and found strong evidence of a common moving behavior. Now, we wish to exploit such degree of association and simulate the forecasted values of the remaining size metrics from the fitted copula conditional on the forecasted pillar metric - a method that saves us from estimating linear and nonlinear dynamic panel models again and again in the way that is executed in the previous section. We proceed as follows. We repeat the exercise of drawing 1000 copula simulations for all periods of time and for each metric of size, but this time we condition the simulations to $\hat{s}_{(op)}$. We simulate the new vector $\tilde{\tilde{\mathbf{s}}}_{-[h]}$ from the conditional distribution $G(\cdot|\hat{s}_h)$. We not only evaluate the quality of the new simulations but, foremost, derive the firm size dynamics of each individual size metric by capitalizing on their association with the pillar metric.

Similarly to Table 5, Table 9 displays the values of the coefficients of determination and those of the test for the equality of distributions. Columns $2-5$ of Table 9 can be compared with columns 1-4 of Table 5, and from their comparison we learn that the copula simulations conditional on the forecasted pillar metric are as reliable as those that revolve on the observed pillar metric. Because of the availability of $\hat{s}_{(op)}$, Table 9 has an additional first column with important information on the distance between the observed and the forecasted pillar metric.[24] The remaining columns $6-10$ are illuminating on the goodness-of-fit of the size metrics forecasted from the equation-by-equation growth estimations, inclusive as we can see of the copula metric (last column), of which we will say more in the next section.

Describing Table 9 in more detail, on its first column shows a high explanatory power of the coefficient of determination for the forecasted pillar metric, with an almost unnoticeable downtrend over time. Table 9 also offers useful comparison between the reliability of the copula approach and the equation-by-equation alternative. The comparison between the coefficients of determination of the two competing techniques highlights a better fit of the equation-by-equation forecasts for the metrics number of subsidiaries, employment and operational revenues. On the one hand, the equation-by-equation forecasts explain better the total variation of the corresponding observed metrics, with the exclusion of the problematic total assets. On the other hand, the equation-by-equation technique does worse in fitting the probability distributions of the corresponding observed metrics, with the exception of cumulated number of subsidiaries for part of the period. And when the decision is between choosing a method that performs better in explaining the total variation of an observed metric, the equation-by-equation estimation, versus one that achieves a better fit of the underlying probability distribution function, the copula method, we have no hesitation to prefer the latter to the former, for the size probability distribution has played a key role in the Growth-of-Firms literature, which we bridge to here.

Having a methodology that is able to estimate reliable dynamics of firm size growth within an industry, independently of the measure of size adopted is extremely important, for it gives any policy maker the magic wand to forecast size distribution in the future, and therefore to be put in the condition to anticipate the market structure that will occur in the market: the European pharmaceutical market here.

6.3 Step-Three

6.3.1 The Copula Metric

The last use of the copula methodology is to generate a comprehensive (normalized to one) continuous metric of firm size, which we label $s_{(co)}$. This additional firm size metric is nothing but the the multivariate CDF itself. Making use of the notation developed in Appendix 3, the MNE growth is given by the bounded difference $\hat{C}(\mathbf{u}_t)-\hat{C}(\mathbf{u}_{t-1})$, where $\hat{C}$ is the fitted copula and $\mathbf{u}$ is a point in the $H-$dimensional unit space. The added value of our new size metric is that it intrinsically carries the 'DNA' of each competing metric. With this new metric at disposal, the set of all metrics is now extended to the superset $\mathcal{H}^* \supset \mathcal{H}$.

[24]We utilize the estimates of our chosen dynamic Probit random effects and forecast the size-measure number of cumulated opportunities, $\hat{s}_{(op)}$. We construct the variable by adding to its initial size $\hat{s}_{(op)i,\tau_i-1}$ (set to 0 for new entrants) the forecasted cumulated unitary expansions.

We have largely talked over how to relate MNE growth metrics to a latent random profit function, via a set of correspondences that map from the CDF of the random profit function to the outcome of each size metric; the copula metric will require another of these correspondences.

Because of the continuity of the copula metric, the equation-by-equation MNE growth dynamics will be expressed as a dynamic linear panel model. We present our estimated version in the last column of Table 8. The ratio between the partial effect of lagged size and that of average size is estimated to be -0.98, which is not too far-off that of its input metrics.

7 Conclusions

Prior to selecting a pillar measure (metric) upon which to study the dynamics of firm growth, this paper undertakes a diligent study of the association between different measures of size. We rely on a copula approach and a dataset on pharmaceutical multinationals that allows us to construct six measures of size, two discrete and four continuous. Due to an excessive number of missing values we have discharged one of the continuous metrics, sales, from the analysis.

The results from the copula estimation provide evidence of concordance in the behavior of the alternative variates. We have some difficulty in explaining total assets, because this variable displays an anomalous behavior. We base our claim not only on the high values of the estimated concordance coefficients, but also on the ability of the copula conditioned on one metric, to simulate the remaining metrics - as confirmed by the statistics documented in Table 5. Evidence of strong association in the variates is the "no impediment to" that we needed to freely select the metric that we believed to be the most straightforward measure to employ: the number of cumulated opportunities.

With this metric at hand, we proceed with an analysis which studies the dynamics of the European pharmaceutical industry within a period of enlargement: the European Single Market Programme era, 1993-2004. Under appropriate assumptions outlined in the paper, we estimate a dynamic panel random effects Probit model of expansion. In our estimations we did our best to account, on the one hand for the correlation between the unobserved heterogeneity and the regressors, and on the other hand for the correlation between the unobserved heterogeneity and the initial conditions. We present results from alternative estimation techniques. We pick the one that predicts, correctly, the highest number of cumulated opportunities (expansions).

Understanding the evolution of the European pharmaceutical industry during the Single Market Programme period, is key to learning about the resulting configuration of market structure. Disentangling the MNE size dynamics and testing the theory of firm growth are the essential requirements. Relative to the former, one of the main findings of the paper is that there is a considerable positive relation between past and current expansion, i.e., we find evidence of strong state dependence. In a counterfactual, where we compare the two extreme scenarios of continuity of expansion and lack of expansion, we quantify the average partial effect exerted by the state dependence to be almost 30 per cent. We read this percentage as a sign of the healthy growth that the European pharmaceutical industry carries through the period. In addition, we uncover that size has a significant mean reversion effect on MNE growth. This effect is confirmed in our analysis of growth by strata. Initial size matters at the beginning of the period of investigation, but over time due to the mean reversion impact, such an initial advantage vanishes. This confirms the theory of maturation, which states that firms face a period of rapid growth, followed by a slow down, or even a stop in growth. We add onto the analysis the role of proportional growth by strata, and find out that new entrants have a higher probability of expansion than the subset of incumbents with low initial size. This finding suggests that Gibrat's law is invalid for firms at the lower end of the size distribution. Ultimately, we find that concentration measures based on the two least heterogeneous size metrics, number of cumulated opportunities and number of cumulated subsidiaries do not satisfy Sutton's lower bound, while the most heterogenous metrics lead to a concentration curve which is well above the lower bound. Surprisingly the global measure of firm size that we have computed does not satisfy Sutton's lower bound.

From the comparison between the copula simulations and the alternative equation-by-equation forecasts we have not gained sufficient support to claim that one method is univocally more accurate than the other. One method is preferable to fit the probability distribution functions of the size variables, the other to compute the total variation of the size variables and to determine the inequality measure, basically to provide information on the moments of the distributions. It is worth adding that the choice to condition the copula simulation on the most homogeneous metric also has implications on the accuracy of the methodology. We could have chosen a more heterogeneous metric as pillar, but that was beyond the message of this article. Our interest was on how a primary measure of firm size as the cumulated number of expansion is able to explain more heterogeneous metrics of size.

While this article extends the way that firm growth is measured and proposes a new way of thinking about firm growth, it has the limitation of having its results curtailed to the European pharmaceutical industry during a period of expansion.

At last, it is worth mentioning that we have limited the copula analysis to its static case. We leave to future research a dynamic investigation of alternative metrics of size.

Table 1: Size metrics used in previous studies

Size Metrics	Papers
Assets	Hart and Prais (1956), Adelman (1958), Simon and Bonini (1958), Ijiri and Simon (1964), Samuels (1965), Scherer (1965), Quandt (1966), Samuels and Smyth (1968), Singh et al. (1968), Radice (1971), Smyth et al. (1975), Dunne and Hughes (1994), Amaral et al. (1997).
Capital employed	Samuels and Chesher (1972), Chesher (1979).
Capital invested	Hall and Weiss (1967), Smyth et al. (1975).
Employment	Simon and Bonini (1958), Mansfield (1962), Saving (1965), Scherer (1965), Smyth et al. (1975), Evans (1987b), Hall (1987), Pavitt et al. (1987), Dunne et al. (1989), Dunne and Hughes (1994), Variyam and Kraybill (1994), Amaral et al. (1997), Axtell (2001), Blonigen and Tomlin (2001), Cabral and Mata (2003), Lotti et al. (2003), Calvo (2006), Huber and Pfaffermayr (2010).
Input	Amaral et al. (1997).
Market valuation	Samuels and Smyth (1968), Singh et al. (1968), Radice (1971), Smyth et al. (1975).
Output	Mansfield (1962).
Plants and equipments	Amaral et al. (1997).
Profit	Hart (1962).
Property	Amaral et al. (1997).
Sales	Hall and Weiss (1967), Samuels and Smyth (1968), Singh et al. (1968), Radice (1971), Smyth et al. (1975), Hart and Oulton (1996), Amaral et al. (1997), Geroski et al. (1997), Sutton (1997b), Bottazzi et al. (2001), Higson et al. (2002), Bottazzi and Secchi (2005), Bottazzi and Secchi (2006a), Bottazzi and Secchi (2006b), Buldyrev et al. (2007), Cefis et al. (2007), Amisano and Giorgetti (2005).
Turnover	Freeman (1986), Dunne et al. (1988).

Table 2: Size metrics yearly statistics

Metric	statistics	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	Pooled
$y_{(op)}$	max	na	1	1	1	1	1	1	1	1	1	1	1	1	1
	mean		0.12	0.20	0.17	0.18	0.15	0.17	0.15	0.18	0.16	0.10	0.08	0.03	0.14
	s.d.		0.33	0.40	0.37	0.39	0.35	0.38	0.35	0.39	0.36	0.30	0.27	0.17	0.35
	Σ_i		27	44	38	43	35	43	37	47	41	26	21	8	410
$s_{(op)}$	max	26	26	27	28	29	29	30	31	32	33	33	33	33	33
	mean	4.00	4.06	4.11	4.22	4.26	4.37	4.39	4.43	4.53	4.63	4.64	4.69	4.72	4.41
	s.d.	4.45	4.50	4.56	4.63	4.70	4.78	4.84	4.89	4.97	5.05	5.07	5.12	5.14	4.84
	skewness	2.31	2.29	2.31	2.27	2.25	2.20	2.25	2.25	2.28	2.28	2.25	2.24	2.24	2.27
	$\arg\max_i$	S.A.†	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.
$y_{(su)}$	max	na	2	3	3	4	3	6	7	17	5	2	11	2	17
	mean		0.13	0.22	0.21	0.24	0.19	0.21	0.19	0.31	0.21	0.11	0.11	0.03	0.18
	s.d.		0.37	0.53	0.54	0.65	0.55	0.64	0.64	1.36	0.58	0.40	0.72	0.20	0.66
	Σ_i		27	47	47	54	45	49	47	78	53	29	30	9	515
$s_{(su)}$	max	47	47	49	51	54	54	56	58	59	61	61	61	61	61
	mean	4.75	4.83	4.90	5.06	5.16	5.32	5.37	5.46	5.69	5.84	5.85	5.93	5.96	5.43
	s.d.	5.94	6.03	6.15	6.33	6.58	6.75	6.98	7.12	7.61	7.76	7.80	8.02	8.06	7.11
	skewness	3.10	3.09	3.18	3.14	3.24	3.13	3.29	3.28	3.28	3.29	3.25	3.29	3.30	3.32
	$\arg\max_i$	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.
$y_{(em)}$	max	na	na	na	na	5.86	8.14	3.35	12.00	12.92	6.64	3.19	11.91	1.12	12.92
	mean					-0.01	0.11	0.16	0.12	0.30	0.21	0.08	0.07	-0.11	0.10
	s.d.					1.01	0.86	0.43	1.02	1.25	0.79	0.68	0.95	0.70	0.89
	Σ_i					-1.48	20.71	31.70	24.71	68.72	50.93	19.83	18.18	-28.37	204.93
$s_{(em)}$	max	na	na	na	18.19	21.94	26.82	30.17	31.14	40.78	45.31	45.58	44.16	45.28	45.58
	mean				1.26	1.21	1.27	1.38	1.37	1.61	1.76	1.81	1.84	1.76	1.55
	s.d.				2.86	2.80	3.06	3.31	3.44	4.32	4.75	4.66	4.76	4.60	4.02
	skewness				3.66	4.34	4.85	4.96	5.01	5.72	5.96	5.50	5.16	5.54	5.71
	$\arg\max_i$				Glaxo	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.
	N. missing				76	69	61	55	34	24	16	11	6	10	362

† S.A.=Sanofi Aventis.

Cont.

Table 2: Size metrics yearly statistics (Cont.)

Metric	statistics	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	Pooled
$y_{(or)}$	max					0.73	2.56	3.61	5.98	4.73	2.48	5.93	8.93	0.45	8.93
	mean	na	na	na	na	0.03	0.02	0.09	0.06	0.07	0.06	0.18	0.19	-0.08	0.07
	s.d.					0.13	0.27	0.40	0.55	0.45	0.30	0.61	0.76	0.38	0.48
	Σ_i					5.97	4.05	20.01	12.77	15.34	14.64	45.60	49.14	-21.14	146.38
$s_{(or)}$	max				7.39	7.93	10.49	12.43	12.96	16.92	19.18	25.11	34.04	33.49	34.04
	mean				0.41	0.43	0.43	0.51	0.53	0.57	0.61	0.77	0.94	0.88	0.62
	s.d.				1.03	1.07	1.10	1.37	1.47	1.68	1.86	2.38	3.04	2.91	1.99
	skewness	na	na	na	4.27	4.35	5.13	4.99	4.71	5.66	5.97	6.25	6.79	7.22	7.91
	$\underset{i}{\text{argmax}}$				S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.
	N. missing				70	64	55	50	35	23	14	8	4	10	333
$y_{(sa)}$	max					0.61	2.63	2.84	3.54	3.77	1.97	5.59	7.76	0.71	7.76
	mean	na	na	na	na	0.02	0.05	0.06	0.01	0.03	0.05	0.15	0.17	-0.08	0.05
	s.d.					0.11	0.25	0.26	0.28	0.33	0.23	0.55	0.67	0.38	0.40
	Σ_i					3.67	8.76	10.80	1.66	6.47	10.04	30.83	36.61	-17.40	91.44
$s_{(sa)}$	max				6.25	6.73	9.36	10.84	11.32	15.09	16.44	22.04	29.80	28.70	29.80
	mean				0.35	0.37	0.39	0.44	0.43	0.45	0.48	0.61	0.77	0.71	0.51
	s.d.				0.82	0.85	0.97	1.14	1.16	1.34	1.47	1.91	2.53	2.40	1.62
	skewness	na	na	na	4.23	4.35	5.51	5.69	5.79	7.36	7.18	7.53	7.89	8.32	9.36
	$\underset{i}{\text{argmax}}$				S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.	S.A.
	N. missing				100	96	82	78	72	64	56	50	49	54	701
$y_{(ta)i}$	max					1.20	2.30	1.33	11.05	16.40	7.33	10.87	10.79	4.14	16.40
	mean	na	na	na	na	0.02	0.01	0.03	0.11	0.13	0.08	0.19	0.26	-0.18	0.07
	s.d.					0.17	0.64	0.20	0.95	1.11	0.52	0.93	1.13	1.47	0.92
	Σ_i					4.10	2.08	7.13	24.78	30.53	18.89	47.72	68.28	-45.62	157.90
$s_{(ta)i}$	max				32.10	30.34	21.83	21.22	23.16	39.56	39.70	36.06	46.85	31.79	46.85
	mean				0.51	0.52	0.50	0.52	0.60	0.70	0.76	0.94	1.19	1.03	0.75
	s.d.				2.44	2.32	1.89	1.85	2.15	3.02	3.14	3.40	4.35	3.55	2.97
	skewness	na	na	na	11.02	10.54	8.07	7.66	6.78	9.62	8.78	6.64	6.82	5.98	8.37
	$\underset{i}{\text{argmax}}$				Glaxo	Glaxo	Glaxo	Glaxo	Glaxo	Glaxo	Glaxo	Glaxo	Glaxo	Glaxo	Glaxo
	N. missing				64	56	47	41	30	19	13	9	6	12	297
MNEs	N. EU	142	143	149	150	154	155	160	162	163	166	169	170	170	-
	N. US	38	38	39	39	41	41	41	44	47	47	48	49	49	-
	N. Other	34	36	37	39	41	42	46	47	48	48	49	49	49	-
	N. Total	214	217	225	228	236	238	247	253	258	261	266	268	268	-
	N. NE	na	3	8	3	8	2	9	6	5	3	5	2	0	-
	$\frac{\text{N. NE}}{\sum_i y_{(op)i}}$	na	0.11	0.18	0.08	0.19	0.06	0.21	0.16	0.11	0.07	0.19	0.10	0.00	-

Table 3: Yearly fit of distributions with positive support

Distribution		1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	Pooled
		DISCRETE METRICS										
		$\mathbf{s_{(op)}}$										
	rate	0.237	0.235	0.229	0.228	0.226	0.221	0.216	0.216	0.213	0.212	0.222
Exponential		(0.016)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.004)
	p − val χ^2	[0.002]	[0.008]	[0.008]	[0.005]	[0.006]	[0.009]	[0.006]	[0.003]	[0.004]	[0.005]	[0.000]
	prob.	0.191	0.190	0.186	0.186	0.184	0.181	0.178	0.177	0.176	0.175	0.182
Geometric†		(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.003)
	p − val χ^2	[0.278]	[0.396]	[0.459]	[0.351]	[0.192]	[0.386]	[0.429]	[0.472]	[0.570]	[0.546]	[0.000]
	size	1.591	1.543	1.545	1.527	1.513	1.528	1.534	1.518	1.502	1.509	1.527
		(0.193)	(0.182)	(0.181)	(0.175)	(0.170)	(0.170)	(0.169)	(0.165)	(0.162)	(0.163)	(0.054)
Negative-Binomial	μ	4.224	4.263	4.374	4.389	4.431	4.527	4.632	4.643	4.687	4.717	4.497
		(0.260)	(0.261)	(0.265)	(0.262)	(0.262)	(0.264)	(0.267)	(0.266)	(0.268)	(0.269)	(0.084)
	p − val χ^2	[0.001]	[0.004]	[0.004]	[0.002]	[0.001]	[0.003]	[0.002]	[0.002]	[0.003]	[0.005]	[0.000]
	λ	4.224	4.263	4.374	4.389	4.431	4.527	4.632	4.643	4.687	4.716	4.497
Poisson		(0.136)	(0.134)	(0.136)	(0.133)	(0.132)	(0.132)	(0.133)	(0.132)	(0.132)	(0.133)	(0.042)
	p − val χ^2	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	mean	0.990	0.988	1.015	1.015	1.022	1.048	1.073	1.071	1.076	1.085	1.040
		(0.061)	(0.061)	(0.061)	(0.059)	(0.059)	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)	(0.019)
Lognormal	**s.d.**	0.917	0.931	0.934	0.935	0.939	0.937	0.939	0.944	0.949	0.948	0.938
		(0.043)	(0.043)	(0.043)	(0.042)	(0.042)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.013)
	p − val χ^2	[0.003]	[0.013]	[0.015]	[0.010]	[0.012]	[0.014]	[0.007]	[0.002]	[0.002]	[0.005]	[0.000]
		$\mathbf{s_{(su)}}$										
	rate	0.198	0.194	0.188	0.186	0.183	0.176	0.171	0.171	0.169	0.168	0.179
Exponential		(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.004)
	p − val χ^2	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	prob.	0.165	0.162	0.158	0.157	0.155	0.150	0.146	0.146	0.144	0.144	0.152
Geometric†		(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.003)
	p − val χ^2	[0.063]	[0.127]	[0.094]	[0.053]	[0.026]	[0.010]	[0.044]	[0.065]	[0.038]	[0.049]	[0.000]
	size	1.253	1.212	1.204	1.185	1.169	1.132	1.130	1.113	1.095	1.099	1.150
		(0.137)	(0.128)	(0.126)	(0.121)	(0.117)	(0.110)	(0.109)	(0.106)	(0.104)	(0.104)	(0.036)
Negative-Binomial	μ	5.057	5.161	5.323	5.372	5.462	5.686	5.835	5.855	5.929	5.962	5.581
		(0.334)	(0.339)	(0.348)	(0.347)	(0.350)	(0.364)	(0.371)	(0.371)	(0.377)	(0.378)	(0.114)
	p − val χ^2	[0.001]	[0.004]	[0.003]	[0.002]	[0.001]	[0.000]	[0.003]	[0.006]	[0.004]	[0.005]	[0.000]
	λ	5.057	5.161	5.324	5.372	5.462	5.686	5.835	5.853	5.929	5.963	5.581
Poisson		(0.149)	(0.148)	(0.150)	(0.148)	(0.147)	(0.148)	(0.150)	(0.148)	(0.149)	(0.149)	(0.047)
	p − val χ^2	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	mean	1.087	1.093	1.223	1.226	1.137	1.164	1.188	1.183	1.188	1.196	1.150
		(0.066)	(0.066)	(0.066)	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)	(0.064)	(0.020)
Lognormal	**s.d.**	0.996	1.009	1.011	1.011	1.017	1.024	1.032	1.042	1.048	1.047	1.025
		(0.047)	(0.046)	(0.046)	(0.045)	(0.457)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.014)
	p − val χ^2	[0.004]	[0.008]	[0.001]	[0.000]	[0.001]	[0.000]	[0.003]	[0.007]	[0.006]	[0.005]	[0.000]

Standard errors in round bracket.
† Selected distribution based on the Pearson χ^2 test.

Cont.

Table 3: Yearly fit of distributions with positive support (Cont.)

Distribution		**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**Pooled**
		QUASI-CONTINUOUS METRICS										
		$\mathbf{s_{(em)}}$										
	mean	-1.481	-1.466	-1.462	-1.275	-1.317	-1.165	-1.043	-0.994	-1.002	-0.996	-1.199
		(0.141)	(0.138)	(0.138)	(0.129)	(0.123)	(0.119)	(0.114)	(0.113)	(0.112)	(0.110)	(0.039)
Lognormal†	**s.d.**	1.958	1.954	1.979	1.888	1.881	1.862	1.810	1.806	1.816	1.773	1.877
		(0.100)	(0.098)	(0.097)	(0.091)	(0.087)	(0.084)	(0.081)	(0.080)	(0.079)	(0.078)	(0.028)
	$\mathbf{p-val\ \chi^2}$	[0.886]	[0.319]	[0.296]	[0.609]	[0.863]	[0.162]	[0.268]	[0.299]	[0.268]	[0.125]	[0.157]
	rate	0.791	0.824	0.790	0.723	0.728	0.622	0.568	0.553	0.542	0.568	0.644
Exponential		(0.057)	(0.058)	(0.055)	(0.050)	(0.048)	(0.040)	(0.036)	(0.035)	(0.034)	(0.035)	(0.013)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	shape	0.525	0.535	0.529	0.548	0.543	0.544	0.549	0.552	0.547	0.558	0.542
		(0.028)	(0.028)	(0.027)	(0.027)	(0.026)	(0.025)	(0.025)	(0.025)	(0.024)	(0.025)	(0.008)
Weibull	**scale**	0.609	0.611	0.619	0.717	0.688	0.793	0.881	0.922	0.921	0.907	0.772
		(0.089)	(0.085)	(0.086)	(0.095)	(0.088)	(0.099)	(0.107)	(0.111)	(0.110)	(0.107)	(0.031)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	shape	0.386	0.397	0.389	0.410	0.402	0.401	0.408	0.413	0.407	0.418	0.401
		(0.032)	(0.032)	(0.031)	(0.032)	(0.030)	(0.029)	(0.029)	(0.030)	(0.029)	(0.030)	(0.010)
Gamma	**rate**	0.305	0.327	0.308	0.296	0.293	0.250	0.231	0.229	0.220	0.238	0.258
		(0.043)	(0.045)	(0.042)	(0.039)	(0.037)	(0.031)	(0.028)	(0.028)	(0.026)	(0.028)	(0.010)
	$\mathbf{p-val\ \chi^2}$	[0.116]	[0.019]	[0.007]	[0.036]	[0.012]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
		$\mathbf{s_{(or)}}$										
	mean	-2.890	-2.774	-2.756	-2.585	-2.631	-2.703	-2.666	-2.428	-2.195	-2.249	-2.570
		(0.154)	(0.146)	(0.142)	(0.137)	(0.135)	(0.141)	(0.136)	(0.135)	(0.130)	(0.129)	(0.044)
Lognormal†	**s.d.**	2.171	2.085	2.070	2.022	2.061	2.204	2.193	2.169	2.119	2.076	2.131
		(0.109)	(0.103)	(0.100)	(0.097)	(0.095)	(0.100)	(0.097)	(0.095)	(0.092)	(0.091)	(0.031)
	$\mathbf{p-val\ \chi^2}$	[0.070]	[0.008]	[0.041]	[0.010]	[0.027]	[0.148]	[0.007]	[0.072]	[0.140]	[0.042]	[0.000]
	rate	2.461	2.346	2.343	1.963	1.876	1.754	1.645	1.299	1.058	1.131	1.602
Exponential		(0.175)	(0.164)	(0.161)	(0.133)	(0.123)	(0.112)	(0.103)	(0.081)	(0.065)	(0.070)	(0.033)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	shape	0.482	0.493	0.496	0.497	0.485	0.468	0.465	0.465	0.470	0.472	0.474
		(0.025)	(0.025)	(0.025)	(0.024)	(0.023)	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.007)
Weibull	**scale**	0.163	0.177	0.179	0.209	0.203	0.200	0.206	0.261	0.323	0.303	0.222
		(0.026)	(0.027)	(0.026)	(0.030)	(0.029)	(0.029)	(0.030)	(0.037)	(0.045)	(0.042)	(0.010)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	shape	0.340	0.350	0.353	0.352	0.338	0.319	0.315	0.316	0.319	0.321	0.324
		(0.027)	(0.028)	(0.027)	(0.027)	(0.025)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.008)
Gamma	**rate**	0.836	0.821	0.826	0.690	0.634	0.559	0.519	0.410	0.338	0.363	0.520
		(0.122)	(0.117)	(0.115)	(0.095)	(0.085)	(0.075)	(0.068)	(0.054)	(0.044)	(0.047)	(0.022)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
		$\mathbf{s_{(ta)}}$										
	mean	-3.047	-3.015	-2.959	-2.858	-2.841	-2.805	-2.669	-2.398	-2.161	-2.199	-2.673
		(0.150)	(0.150)	(0.148)	(0.145)	(0.142)	(0.141)	(0.133)	(0.127)	(0.126)	(0.124)	(0.044)
Lognormal†	**s.d.**	2.138	2.188	2.200	2.189	2.189	2.223	2.116	2.038	2.042	1.986	2.152
		(0.106)	(0.106)	(0.105)	(0.103)	(0.100)	(0.100)	(0.094)	(0.090)	(0.089)	(0.088)	(0.031)
	$\mathbf{p-val\ \chi^2}$	[0.292]	[0.198]	[0.676]	[0.257]	[0.113]	[0.030]	[0.001]	[0.016]	[0.074]	[0.066]	[0.000]
	rate	1.942	1.935	1.980	1.911	1.656	1.426	1.311	1.069	0.843	0.966	1.342
Exponential		(0.136)	(0.133)	(0.133)	(0.127)	(0.107)	(0.090)	(0.082)	(0.067)	(0.052)	(0.060)	(0.027)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	shape	0.451	0.453	0.457	0.464	0.451	0.443	0.448	0.453	0.453	0.463	0.449
		(0.022)	(0.022)	(0.022)	(0.022)	(0.021)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.006)
Weibull	**scale**	0.142	0.147	0.156	0.170	0.175	0.185	0.205	0.264	0.334	0.314	0.206
		(0.023)	(0.024)	(0.024)	(0.026)	(0.027)	(0.028)	(0.030)	(0.039)	(0.048)	(0.045)	(0.010)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	shape	0.291	0.294	0.303	0.310	0.296	0.284	0.289	0.296	0.296	0.308	0.291
		(0.023)	(0.022)	(0.023)	(0.023)	(0.021)	(0.020)	(0.020)	(0.021)	(0.020)	(0.022)	(0.007)
Gamma	**rate**	0.565	0.569	0.599	0.593	0.490	0.405	0.379	0.317	0.250	0.297	0.391
		(0.086)	(0.084)	(0.086)	(0.083)	(0.068)	(0.056)	(0.051)	(0.042)	(0.033)	(0.039)	(0.017)
	$\mathbf{p-val\ \chi^2}$	[0.000]	[0.000]	[0.007]	[0.003]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Table 4: Yearly fit of copulas

copula		Marginal parameter								Concordance coefficient					copula parameter	Log likelihood
		$\mathbf{prob}_{(op)}$	$\mathbf{prob}_{(su)}$	$\mathbf{mean}_{(em)}$	$\mathbf{sd}_{(em)}$	$\mathbf{mean}_{(or)}$	$\mathbf{sd}_{(or)}$	$\mathbf{mean}_{(ta)}$	$\mathbf{sd}_{(ta)}$		$\mathbf{s}_{(su)}$	$\mathbf{s}_{(em)}$	$\mathbf{s}_{(or)}$	$\mathbf{s}_{(ta)}$	ϱ	**ll**
									1995							
										$\mathbf{s}_{(op)}$	0.932	0.601	0.602	0.223		
										$\mathbf{s}_{(su)}$		0.701	0.712	0.389		
Elliptical	**Normal**†	0.185	0.163	-1.848	2.051	-3.167	2.190	-3.009	2.130	$\mathbf{s}_{(em)}$			0.960	0.702		-1,753
										$\mathbf{s}_{(or)}$				0.700		
										$\mathbf{s}_{(op)}$	(0.003)	(0.019)	(0.019)	(0.030)		
										$\mathbf{s}_{(su)}$		(0.016)	(0.015)	(0.027)		
		(0.005)	(0.005)	(0.062)	(0.043)	(0.066)	(0.046)	(0.067)	(0.047)	$\mathbf{s}_{(em)}$			(0.002)	(0.016)		
										$\mathbf{s}_{(or)}$				(0.016)		
Archimedean	**Clayton**	0.179	0.163	-1.958	2.302	-3.232	2.564	-3.361	2.480						2.656	-3,280
		(0.004)	(0.004)	(0.054)	(0.037)	(0.060)	(0.041)	(0.058)	(0.040)						(0.069)	
	Frank	0.174	0.159	-1.656	2.030	-2.999	2.180	-3.180	2.164						2.170	-4,233
		(0.005)	(0.004)	(0.061)	(0.046)	(0.065)	(0.049)	(0.065)	(0.049)						(0.104)	
	Gumbel	0.178	0.155	-1.772	2.088	-3.179	2.255	-3.325	2.300						2.079	-2,693
		(0.005)	(0.004)	(0.059)	(0.037)	(0.064)	(0.040)	(0.065)	(0.041)						(0.045)	
									1996							
										$\mathbf{s}_{(op)}$	0.927	0.655	0.616	0.318		
										$\mathbf{s}_{(su)}$		0.754	0.719	0.475		
Elliptical	**Normal**†	0.178	0.155	-1.800	2.167	-3.058	2.252	-3.094	2.192	$\mathbf{s}_{(em)}$			0.951	0.733		-2,170
										$\mathbf{s}_{(or)}$				0.785		
										$\mathbf{s}_{(op)}$	(0.004)	(0.017)	(0.019)	(0.029)		
										$\mathbf{s}_{(su)}$		(0.013)	(0.015)	(0.025)		
		(0.005)	(0.004)	(0.065)	(0.045)	(0.068)	(0.048)	(0.069)	(0.049)	$\mathbf{s}_{(em)}$			(0.003)	(0.015)		
										$\mathbf{s}_{(or)}$				(0.012)		
Archimedean	**Clayton**	0.194	0.170	-2.033	2.419	-3.369	2.571	-3.530	2.586						2.879	-2,531
		(0.005)	(0.004)	(0.057)	(0.038)	(0.060)	(0.041)	(0.061)	(0.041)						(0.072)	
	Frank	0.175	0.146	-1.642	2.047	-2.829	2.155	-2.953	2.139						2.180	-4,774
		(0.005)	(0.004)	(0.061)	(0.046)	(0.064)	(0.048)	(0.064)	(0.048)						(0.104)	
	Gumbel	0.177	0.153	-1.812	2.120	-3.137	2.293	-3.275	2.283						2.060	-2,842
		(0.005)	(0.004)	(0.060)	(0.038)	(0.065)	(0.041)	(0.065)	(0.041)						(0.045)	
									1997							
										$\mathbf{s}_{(op)}$	0.932	0.716	0.709	0.504		
										$\mathbf{s}_{(su)}$		0.806	0.804	0.642		
Elliptical	**Normal**†	0.181	0.154	-1.802	2.121	-3.089	2.222	-2.984	2.099	$\mathbf{s}_{(em)}$			0.965	0.840		-1,862
										$\mathbf{s}_{(or)}$				0.819		
										$\mathbf{s}_{(op)}$	(0.003)	(0.014)	(0.014)	(0.023)		
										$\mathbf{s}_{(su)}$		(0.010)	(0.010)	(0.018)		
		(0.005)	(0.004)	(0.063)	(0.042)	(0.066)	(0.044)	(0.064)	(0.045)	$\mathbf{s}_{(em)}$			(0.002)	(0.009)		
										$\mathbf{s}_{(or)}$				(0.010)		
Archimedean	**Clayton**	0.181	0.155	-1.849	2.418	-3.165	2.496	-3.185	2.443						2.633	-3,481
		(0.004)	(0.004)	(0.057)	(0.040)	(0.059)	(0.041)	(0.057)	(0.040)						(0.067)	
	Frank	0.185	0.147	-1.711	1.981	-2.838	2.075	-3.071	2.101						2.069	-4,430
		(0.005)	(0.004)	(0.059)	(0.046)	(0.062)	(0.047)	(0.063)	(0.047)						(0.102)	
	Gumbel	0.172	0.149	-1.713	2.064	-3.021	2.150	-3.088	2.192						2.047	-3,150
		(0.005)	(0.004)	(0.059)	(0.038)	(0.061)	(0.039)	(0.062)	(0.040)						(0.044)	

Standard errors in round bracket.

Cont.

Table 4: Yearly fit of copulas (Cont.)

copula		Marginal parameter								Concordance coefficient					copula parameter	Log likelihood
		$\mathbf{prob}_{(op)}$	$\mathbf{prob}_{(su)}$	$\mathbf{mean}_{(em)}$	$\mathbf{sd}_{(em)}$	$\mathbf{mean}_{(or)}$	$\mathbf{sd}_{(or)}$	$\mathbf{mean}_{(ta)}$	$\mathbf{sd}_{(ta)}$		$\mathbf{s}_{(su)}$	$\mathbf{s}_{(em)}$	$\mathbf{s}_{(or)}$	$\mathbf{s}_{(ta)}$	ϱ	ll
		1998														
Elliptical	**Normal**†	0.183	0.158	-1.600	1.936	-2.943	2.063	-2.938	2.076	$\mathbf{s}_{(op)}$	0.913	0.682	0.633	0.451		-2,341
										$\mathbf{s}_{(su)}$		0.796	0.749	0.613		
										$\mathbf{s}_{(em)}$			0.932	0.843		
										$\mathbf{s}_{(or)}$				0.785		
		(0.005)	(0.005)	(0.058)	(0.038)	(0.062)	(0.043)	(0.064)	(0.045)	$\mathbf{s}_{(op)}$	(0.004)	(0.016)	(0.018)	(0.025)		
										$\mathbf{s}_{(su)}$		(0.011)	(0.013)	(0.020)		
										$\mathbf{s}_{(em)}$			(0.004)	(0.009)		
										$\mathbf{s}_{(or)}$				(0.012)		
Archimedean	**Clayton**	0.185 (0.005)	0.159 (0.004)	-1.684 (0.052)	2.196 (0.035)	-3.007 (0.057)	2.421 (0.039)	-3.121 (0.057)	2.436 (0.038)						2.770 (0.070)	-3,505
	Frank	0.171 (0.005)	0.143 (0.004)	-1.394 (0.056)	1.872 (0.042)	-2.766 (0.062)	2.060 (0.047)	-2.869 (0.064)	2.115 (0.048)						2.041 (0.104)	-5,083
	Gumbel	0.179 (0.005)	0.148 (0.004)	-1.570 (0.054)	1.913 (0.039)	-2.882 (0.060)	2.122 (0.038)	-2.999 (0.060)	2.107 (0.038)						2.041 (0.044)	-3,312
		1999														
Elliptical	**Normal**†	0.181	0.152	-1.465	1.897	-2.817	1.997	-2.897	2.096	$\mathbf{s}_{(op)}$	0.920	0.657	0.662	0.478		-2,752
										$\mathbf{s}_{(su)}$		0.794	0.765	0.616		
										$\mathbf{s}_{(em)}$			0.916	0.827		
										$\mathbf{s}_{(or)}$				0.790		
		(0.005)	(0.004)	(0.056)	(0.037)	(0.060)	(0.041)	(0.065)	(0.046)	$\mathbf{s}_{(op)}$	(0.004)	(0.016)	(0.017)	(0.024)		
										$\mathbf{s}_{(su)}$		(0.011)	(0.013)	(0.020)		
										$\mathbf{s}_{(em)}$			(0.005)	(0.010)		
										$\mathbf{s}_{(or)}$				(0.012)		
Archimedean	**Clayton**	0.187 (0.005)	0.159 (0.004)	-1.687 (0.051)	2.154 (0.036)	-3.066 (0.054)	2.290 (0.038)	-3.245 (0.058)	2.439 (0.041)						2.658 (0.068)	-3,328
	Frank	0.174 (0.005)	0.155 (0.004)	-1.486 (0.056)	1.861 (0.042)	-2.790 (0.061)	2.036 (0.047)	-2.905 (0.062)	2.070 (0.047)						2.218 (0.108)	-4,751
	Gumbel	0.185 (0.005)	0.154 (0.004)	-1.548 (0.054)	1.895 (0.034)	-2.967 (0.058)	2.046 (0.037)	-3.100 (0.061)	2.142 (0.038)						2.025 (0.044)	-3,094
		2000														
Elliptical	**Normal**†	0.189	0.159	-1.485	1.924	-3.074	2.269	-2.918	2.126	$\mathbf{s}_{(op)}$	0.897	0.669	0.681	0.384		-2,397
										$\mathbf{s}_{(su)}$		0.831	0.798	0.597		
										$\mathbf{s}_{(em)}$			0.943	0.800		
										$\mathbf{s}_{(or)}$				0.749		
		(0.005)	(0.005)	(0.057)	(0.035)	(0.068)	(0.045)	(0.065)	(0.045)	$\mathbf{s}_{(op)}$	(0.005)	(0.016)	(0.016)	(0.026)		
										$\mathbf{s}_{(su)}$		(0.009)	(0.011)	(0.020)		
										$\mathbf{s}_{(em)}$			(0.003)	(0.011)		
										$\mathbf{s}_{(or)}$				(0.014)		
Archimedean	**Clayton**	0.187 (0.005)	0.161 (0.004)	-1.555 (0.050)	2.135 (0.034)	-3.195 (0.059)	2.496 (0.040)	-3.180 (0.057)	2.417 (0.040)						2.576 (0.066)	-3,510
	Frank	0.175 (0.005)	0.147 (0.004)	-1.141 (0.057)	1.906 (0.043)	-2.844 (0.068)	2.263 (0.051)	-2.744 (0.066)	2.193 (0.050)						2.190 (0.106)	-5,377
	Gumbel	0.179 (0.005)	0.151 (0.004)	-1.425 (0.055)	1.936 (0.036)	-3.009 (0.066)	2.333 (0.044)	-3.033 (0.062)	2.197 (0.040)						2.022 (0.044)	-3,557

Cont.

Table 4: Yearly fit of copulas (Cont.)

copula		Marginal parameter								Concordance coefficient					copula parameter	Log likelihood
		$\mathbf{prob}_{(op)}$	$\mathbf{prob}_{(su)}$	$\mathbf{mean}_{(em)}$	$\mathbf{sd}_{(em)}$	$\mathbf{mean}_{(or)}$	$\mathbf{sd}_{(or)}$	$\mathbf{mean}_{(ta)}$	$\mathbf{sd}_{(ta)}$		$\mathbf{s}_{(su)}$	$\mathbf{s}_{(em)}$	$\mathbf{s}_{(or)}$	$\mathbf{s}_{(ta)}$	ϱ	**ll**
									2001							
Elliptical	**Normal**†	0.170	0.141	-1.292	1.890	-3.002	2.308	-2.883	2.130	$\mathbf{s}_{(op)}$	0.906	0.682	0.691	0.501		-2,492
										$\mathbf{s}_{(su)}$		0.841	0.812	0.684		
										$\mathbf{s}_{(em)}$			0.961	0.878		
										$\mathbf{s}_{(or)}$				0.844		
		(0.005)	(0.004)	(0.055)	(0.034)	(0.065)	(0.045)	(0.064)	(0.045)	$\mathbf{s}_{(op)}$	(0.005)	(0.015)	(0.015)	(0.022)		
										$\mathbf{s}_{(su)}$		(0.008)	(0.010)	(0.016)		
										$\mathbf{s}_{(em)}$			(0.002)	(0.007)		
										$\mathbf{s}_{(or)}$				(0.009)		
Archimedean	**Clayton**	0.189 (0.005)	0.156 (0.004)	-1.516 (0.052)	2.204 (0.036)	-3.222 (0.062)	2.655 (0.044)	-3.156 (0.058)	2.477 (0.041)						2.732 (0.071)	-3,773
	Frank	0.172 (0.005)	0.149 (0.004)	-1.192 (0.055)	1.861 (0.042)	-2.765 (0.068)	2.276 (0.052)	-2.804 (0.060)	2.032 (0.046)						2.177 (0.105)	-5,298
	Gumbel	0.170 (0.005)	0.143 (0.004)	-1.263 (0.056)	1.978 (0.036)	-2.936 (0.067)	2.372 (0.042)	-2.908 (0.063)	2.213 (0.039)						2.121 (0.046)	-3,874
									2002							
Elliptical	**Normal**†	0.178	0.146	-1.230	1.940	-2.741	2.339	-2.620	2.088	$\mathbf{s}_{(op)}$	0.901	0.698	0.676	0.531		-3,025
										$\mathbf{s}_{(su)}$		0.845	0.815	0.705		
										$\mathbf{s}_{(em)}$			0.961	0.854		
										$\mathbf{s}_{(or)}$				0.895		
		(0.005)	(0.004)	(0.057)	(0.035)	(0.069)	(0.044)	(0.063)	(0.043)	$\mathbf{s}_{(op)}$	(0.005)	(0.014)	(0.015)	(0.021)		
										$\mathbf{s}_{(su)}$		(0.008)	(0.010)	(0.015)		
										$\mathbf{s}_{(em)}$			(0.002)	(0.008)		
										$\mathbf{s}_{(or)}$				(0.006)		
Archimedean	**Clayton**	0.176 (0.004)	0.149 (0.004)	-1.255 (0.050)	2.117 (0.033)	-2.745 (0.062)	2.608 (0.042)	-2.750 (0.056)	2.354 (0.037)						2.587 (0.066)	-4,736
	Frank	0.181 (0.005)	0.153 (0.004)	-1.113 (0.055)	1.849 (0.042)	-2.612 (0.066)	2.237 (0.050)	-2.550 (0.063)	2.134 (0.048)						2.281 (0.105)	-5,744
	Gumbel	0.175 (0.005)	0.145 (0.004)	-1.225 (0.054)	1.907 (0.034)	-2.670 (0.065)	2.276 (0.041)	-2.672 (0.061)	2.147 (0.038)						2.042 (0.044)	-4,279
									2003							
Elliptical	**Normal**†	0.171	0.141	-1.169	1.932	-2.390	2.264	-2.270	2.075	$\mathbf{s}_{(op)}$	0.898	0.687	0.646	0.505		-3,633
										$\mathbf{s}_{(su)}$		0.855	0.802	0.701		
										$\mathbf{s}_{(em)}$			0.956	0.866		
										$\mathbf{s}_{(or)}$				0.892		
		(0.005)	(0.004)	(0.056)	(0.033)	(0.067)	(0.043)	(0.062)	(0.042)	$\mathbf{s}_{(op)}$	(0.005)	(0.014)	(0.016)	(0.021)		
										$\mathbf{s}_{(su)}$		(0.007)	(0.010)	(0.015)		
										$\mathbf{s}_{(em)}$			(0.002)	(0.007)		
										$\mathbf{s}_{(or)}$				(0.006)		
Archimedean	**Clayton**	0.183 (0.004)	0.157 (0.004)	-1.387 (0.048)	2.047 (0.034)	-2.691 (0.057)	2.419 (0.040)	-2.686 (0.055)	2.343 (0.039)						2.611 (0.067)	-4,475
	Frank	0.170 (0.005)	0.138 (0.004)	-1.034 (0.054)	1.822 (0.041)	-2.398 (0.066)	2.211 (0.049)	-2.236 (0.062)	2.074 (0.047)						2.291 (0.106)	-6,402
	Gumbel	0.164 (0.004)	0.135 (0.004)	-1.062 (0.054)	1.944 (0.034)	-2.293 (0.065)	2.293 (0.041)	-2.279 (0.063)	2.204 (0.039)						2.165 (0.047)	-5,182

Cont.

Table 4: Yearly fit of copulas (Cont.)

copula		Marginal parameter								Concordance coefficient					copula parameter	Log likelihood
		$\mathbf{prob}_{(\mathbf{op})}$	$\mathbf{prob}_{(su)}$	$\mathbf{mean}_{(em)}$	$\mathbf{sd}_{(em)}$	$\mathbf{mean}_{(or)}$	$\mathbf{sd}_{(or)}$	$\mathbf{mean}_{(ta)}$	$\mathbf{sd}_{(ta)}$		$\mathbf{s}_{(su)}$	$\mathbf{s}_{(em)}$	$\mathbf{s}_{(or)}$	$\mathbf{s}_{(ta)}$	ϱ	**ll**
									2004							
Elliptical	**Normal**†	0.187	0.159	-1.426	1.873	-2.734	2.233	-2.656	2.125	$\mathbf{s}_{(op)}$	0.896	0.679	0.662	0.616		-2,510
										$\mathbf{s}_{(su)}$		0.830	0.792	0.757		
										$\mathbf{s}_{(em)}$			0.962	0.922		
										$\mathbf{s}_{(or)}$				0.929		
		(0.005)	(0.005)	(0.055)	(0.035)	(0.067)	(0.045)	(0.064)	(0.044)	$\mathbf{s}_{(op)}$	(0.005)	(0.015)	(0.016)	(0.018)		
										$\mathbf{s}_{(su)}$		(0.009)	(0.011)	(0.013)		
										$\mathbf{s}_{(em)}$			(0.002)	(0.004)		
										$\mathbf{s}_{(or)}$				(0.004)		
Archimedean	**Clayton**	0.177	0.144	-1.256	2.092	-2.599	2.494	-2.595	2.397						2.528	-5,147
		(0.004)	(0.004)	(0.050)	(0.034)	(0.060)	(0.041)	(0.057)	(0.040)						(0.067)	
	Frank	0.164	0.137	-0.969	1.686	-2.222	2.121	-2.274	1.992						2.033	-6,413
		(0.005)	(0.004)	(0.050)	(0.038)	(0.063)	(0.047)	(0.059)	(0.044)						(0.100)	
	Gumbel	0.169	0.140	-1.125	1.828	-2.467	2.157	-2.421	2.080						2.033	-4,792
		(0.005)	(0.004)	(0.052)	(0.033)	(0.062)	(0.039)	(0.059)	(0.038)						(0.044)	
									All							
Elliptical	**Normal**†	0.171	0.148	-1.456	1.960	-2.861	2.177	-2.846	2.085	$\mathbf{s}_{(op)}$	0.920	0.701	0.660	0.487		-2,694
										$\mathbf{s}_{(su)}$		0.822	0.758	0.636		
										$\mathbf{s}_{(em)}$			0.933	0.835		
										$\mathbf{s}_{(or)}$				0.841		
		(0.005)	(0.004)	(0.058)	(0.037)	(0.066)	(0.045)	(0.064)	(0.045)	$\mathbf{s}_{(op)}$	(0.004)	(0.014)	(0.017)	(0.023)		
										$\mathbf{s}_{(su)}$		(0.009)	(0.013)	(0.019)		
										$\mathbf{s}_{(em)}$			(0.004)	(0.010)		
										$\mathbf{s}_{(or)}$				(0.009)		
Archimedean	**Clayton**	0.183	0.160	-1.620	2.217	-3.025	2.517	-3.095	2.482						2.682	-3,793
		(0.004)	(0.004)	(0.052)	(0.036)	(0.059)	(0.041)	(0.059)	(0.040)						(0.069)	
	Frank	0.165	0.148	-1.223	1.869	-2.649	2.131	-2.667	2.096						2.020	-5,529
		(0.005)	(0.004)	(0.056)	(0.042)	(0.064)	(0.048)	(0.063)	(0.047)						(0.100)	
	Gumbel	0.173	0.141	-1.444	1.974	-2.778	2.212	-2.844	2.150						2.073	-3,812
		(0.005)	(0.004)	(0.056)	(0.034)	(0.063)	(0.039)	(0.061)	(0.038)						(0.045)	

Table 5: Conditional (on $s_{(op)}$) simulations from Elliptical Normal copula: Coefficient of determination and Pearson's χ^2 test of equality of distributions

Year	Copula Conditional Simulation							
	1	2	3	4	1	2	3	4
Year	$\tilde{s}_{(su)}$	$\tilde{s}_{(em)}$	$\tilde{s}_{(or)}$	$\tilde{s}_{(ta)}$	$\tilde{s}_{(su)}$	$\tilde{s}_{(em)}$	$\tilde{s}_{(or)}$	$\tilde{s}_{(ta)}$
	R^2				p − value χ^2			
1995	0.825 [0.798,0.843]	0.216 [0.189,0.276]	0.186 [0.172,0.210]	0.035 [0.033,0.041]	0.493 [0.023,0.979]	0.299 [0.001,0.939]	0.199 [0.000,0.872]	0.172 [0.000,0.810]
1996	0.814 [0.787,0.836]	0.273 [0.231,0.359]	0.198 [0.180,0.236]	0.049 [0.045,0.055]	0.440 [0.004,0.980]	0.203 [0.000,0.879]	0.119 [0.000,0.682]	0.266 [0.000,0.952]
1997	0.803 [0.773,0.826]	0.301 [0.257,0.394]	0.242 [0.222,0.283]	0.096 [0.086,0.110]	0.458 [0.011,0.978]	0.140 [0.000,0.756]	0.177 [0.000,0.830]	0.387 [0.002,0.963]
1998	0.801 [0.770,0.824]	0.294 [0.257,0.374]	0.212 [0.192,0.240]	0.109 [0.100,0.123]	0.440 [0.011,0.976]	0.231 [0.000,0.881]	0.113 [0.000,0.696]	0.264 [0.000,0.925]
1999	0.788 [0.758,0.813]	0.270 [0.236,0.336]	0.227 [0.209,0.251]	0.114 [0.106,0.130]	0.333 [0.006,0.916]	0.364 [0.001,0.947]	0.220 [0.000,0.920]	0.240 [0.000,0.881]
2000	0.738 [0.703,0.764]	0.294 [0.245,0.375]	0.236 [0.204,0.295]	0.085 [0.077,0.100]	0.204 [0.001,0.784]	0.366 [0.001,0.957]	0.097 [0.000,0.641]	0.197 [0.000,0.853]
2001	0.737 [0.704,0.763]	0.271 [0.237,0.323]	0.234 [0.210,0.274]	0.101 [0.094,0.114]	0.211 [0.001,0.789]	0.226 [0.000,0.880]	0.067 [0.000,0.516]	0.059 [0.000,0.448]
2002	0.738 [0.704,0.767]	0.315 [0.270,0.403]	0.246 [0.211,0.320]	0.137 [0.123,0.161]	0.233 [0.001,0.834]	0.162 [0.000,0.778]	0.145 [0.000,0.800]	0.081 [0.000,0.637]
2003	0.724 [0.690,0.751]	0.325 [0.280,0.408]	0.242 [0.209,0.316]	0.136 [0.122,0.162]	0.224 [0.001,0.807]	0.304 [0.001,0.913]	0.206 [0.000,0.868]	0.167 [0.000,0.785]
2004	0.742 [0.710,0.765]	0.295 [0.253,0.365]	0.237 [0.199,0.311]	0.189 [0.165,0.228]	0.306 [0.002,0.915]	0.158 [0.000,0.753]	0.074 [0.000,0.576]	0.154 [0.000,0.797]

In square brackets 95 per cent empirical confidence interval.

Table 6: MNEs expansion Pooled and Random Effects dynamic panel Probit estimation - with and without Mundlak correction

Variable	**Pooled**[†] no Mundlak correction	**Random Effects**[††] no Mundlak correction	**Pooled**[†] Mundlak correction	**Random Effects**[††] Mundlak correction
$y_{i,t-1}$ [γ]	0.395*** (0.074)	0.090 (0.091)	0.713*** (0.086)	0.597*** (0.099)
$s_{(op)i,t-1}$ [β_{21}]	0.035*** (0.007)	0.028*** (0.009)	-0.805*** (0.103)	-0.799*** (0.055)
$Cons$ [λ_{10}]	-2.067*** (0.185)	-2.203*** (0.203)	-1.613*** (0.183)	-1.643*** (0.176)
eu_i [λ_{11}]	-0.124 (0.090)	-0.154 (0.120)	-0.125 (0.087)	-0.132 (0.097)
us_i [λ_{12}]	0.017 (0.127)	0.024 (0.150)	0.003 (0.124)	-0.005 (0.121)
$\bar{s}_{(op)i}$ [λ_{13}]			0.853*** (0.101)	0.851*** (0.056)
Year dummies	Yes	Yes	Yes	Yes
ρ		0.208*** (0.041)		0.071** (0.032)
Statistics:				
APE of $y_{i,t-1}$	15.2	5.0	36.8	35.5
Percentage of positive predicted outcomes	1.9	0.0	16.8	14.9
Log-likelihood	-1 111.1	-1 088.5	-957.7	-954.4
BIC	2 350.1	2 312.8	2 051.3	2 052.8
N. Obs.	2 965	2 965	2 965	2 965

[†] Clustered standard errors in brackets.
[††] Robust standard errors in brackets, computed using Fisher's information matrix.

Table 7: Random coefficients dynamic Probit model: Alternative estimates MNEs expansion

Variable[†]	**Arulampalam & Stewart EC**[†††]	**Stewart EC**[†††]	**Wooldridge EC**[†††]	**Orme EC**[†††]	**Orme FC**[†††]
Main equation					
$y_{i,t-1}$ $[\gamma]$	0.622***	0.580***	0.566***	0.564***	0.559***
	(0.112)	(0.111)	(0.100)	(0.099)	(0.100)
$s_{(op)i,t-1}$ $[\beta_{21}]$	-0.830***	-0.680***	-0.788***	-0.798***	-0.797***
	(0.107)	(0.060)	(0.056)	(0.055)	(0.056)
$Const$ $[\lambda_{10}]$	-1.482***	-1.787***	-1.717***	-1.246***	-1.178***
	(0.155)	(0.223)	(0.183)	(0.170)	(0.232)
eu_i $[\lambda_{11}]$	-0.142	-0.160	-0.120	-0.121	-0.124
	(0.090)	(0.132)	(0.099)	(0.098)	(0.100)
us_i $[\lambda_{12}]$	-0.007	-0.002	-0.003	-0.009	-0.008
	(0.125)	(0.160)	(0.123)	(0.120)	(0.123)
$\bar{s}_{(op)i}$ $[\lambda_{13}]$	0.881***	0.746***	0.842***	0.847***	0.846***
	(0.104)	(0.061)	(0.056)	(0.056)	(0.057)
$y_{i,91}$ $[\lambda_{14}]$			0.158*		
			(0.085)		
Initial conditions equation (1992)					
$Const$ $[\lambda_{00}]$	-1.964***	-1.912***		-1.912***	-1.912***
	(0.377)	(0.363)		(0.363)	(0.363)
eu_i $[\lambda_{01}]$	0.302	0.295		0.295	0.295
	(0.385)	(0.356)		(0.356)	(0.356)
us_i $[\lambda_{02}]$	0.236	0.227		0.227	0.227
	(0.452)	(0.422)		(0.422)	(0.422)
$\bar{s}_{(op)i}$ $[\lambda_{03}]$	0.092***	0.092***		0.092***	0.092***
	(0.020)	(0.022)		(0.022)	(0.022)
$age_{i,91}$ $[\eta_1]$	0.009	0.009		0.009	0.009
	(0.037)	(0.037)		(0.037)	(0.037)
$imr\ \alpha_i$ $[\delta_0]$				-0.211***	-0.277*
				(0.053)	(0.164)
Year dummies	yes	yes	yes	yes	yes
ρ	0.064**	0.142***	0.080**	0.071**	0.082**
	(0.037)	(0.044)	(0.033)	(0.031)	(0.033)
α_i $[\delta_0]$	0.156				
	(0.917)				
α_i $[\delta_1]$					-0.049
					(0.302)
α_i $[\delta_2]$					-0.298
					(0.253)
α_i $[\delta_3]$					0.047
					(0.261)
α_i $[\delta_4]$					-0.174
					(0.234)
α_i $[\delta_5]$					0.221
					(0.242)
α_i $[\delta_6]$					-0.138
					(0.217)
α_i $[\delta_7]$					0.213
					(0.218)
α_i $[\delta_8]$					0.135
					(0.208)
α_i $[\delta_9]$					0.216
					(0.207)
α_i $[\delta_{10}]$					0.153
					(0.213)
α_i $[\delta_{11}]$					0.064
					(0.214)
Statistics:					
APE of $y_{i,t-1}$	28.8	38.9	35.2	34.8	34.3
Percentage of positive predicted outcomes	20.7	12.7	14.6	16.3	15.6
Log-likelihood	-1 032.9	-838.6	-952.6	-953.1	-947.3
BIC	2 259.3	1 870.8	2 057.2	2 058.1	2 134.4
N. Obs. main equation	2 971	2 971	2 965	2 971	2 971
N. Obs. initial equation	208	208	na	208	208

[†] The parameters in square brackets rely on the notation developed in Appendix A.
[††] Clustered standard errors in brackets.
[†††] Robust standard errors in brackets, computed using Fisher's information matrix.

Table 8: Equation-by-equation estimates MNEs size dynamics

Variable	$y_{(su)}$	$y_{(em)}$	$y_{(or)}$	$y_{(ta)}$	$y_{(co)}$
$y_{i,t-1}$ [γ]	0.261**	-0.145	-0.167	0.150	0.330***
	(0.114)	(0.122)	(0.120)	(0.140)	(0.027)
$s_{i,t-1}$ [β_{21}]	-0.274**	-0.485***	-0.123	-0.525***	-1.793***
	(0.113)	(0.085)	(0.106)	(0.115)	(0.050)
$Cons$ [λ_{10}]	-3.023***	-0.465***	-0.024	-0.319***	0.020
	(0.323)	(0.067)	(0.025)	(0.077)	(0.050)
eu_i [λ_{11}]	-0.092	0.016	-0.040	-0.036	-0.036
	(0.240)	(0.074)	(0.030)	(0.092)	(0.043)
us_i [λ_{12}]	0.345	-0.019	-0.054	0.206	-0.009
	(0.269)	(0.094)	(0.045)	0.177	(0.056)
$\bar{s}_i$ [λ_{13}]	0.317***	0.590***	0.179***	0.406***	1.860***
	(0.088)	(0.025)	0.038	(0.117)	(0.196)
$y_{(su)i,91}$ [λ_{14}]	0.180	-0.049	0.116	0.530*	-0.022
	(0.447)	(0.218)	(0.072)	(0.304)	(0.038)
Year dummies [β_1]	Yes	Yes	Yes	Yes	Yes
N. Obs.	2 197†	1 472	1 472	1 472	1 472
Gmm value	0.001				
Sargan test (P-value)		0.000	0.000	0.000	0.000
Serial corr. Lag 1 (P-value)		0.008	0.003	0.069	0.000
Serial corr. Lag 2 (P-value)		0.183	0.306	0.849	0.454

Robust standard errors in parentheses. † For this variable data were available from 1991.

Table 9: Conditional (on $\hat{s}_{(op)}$) simulations from Elliptical Normal copula and equation-by-equation forecasted values: Coefficient of determination and Pearson's χ^2 test of equality of distributions

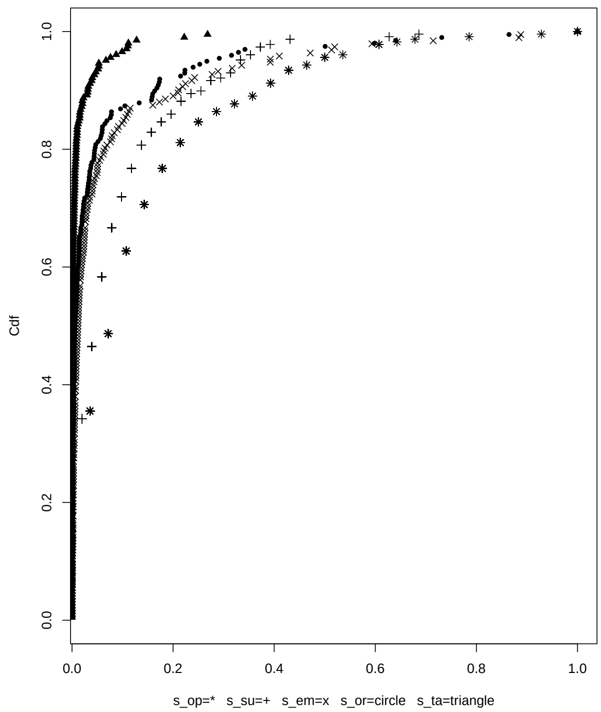
Year	**Eq.-by-Eq. Forecast**	**Conditional Simulation**				**Eq.-by-Eq. Forecast**			
	1	2	3	4	5	6	7	8	9
	$\hat{\mathbf{s}}_{(op)}$	$\tilde{\tilde{\mathbf{s}}}_{(su)}$	$\tilde{\tilde{\mathbf{s}}}_{(em)}$	$\tilde{\tilde{\mathbf{s}}}_{(or)}$	$\tilde{\tilde{\mathbf{s}}}_{(ta)}$	$\hat{\mathbf{s}}_{(su)}$	$\hat{\mathbf{s}}_{(em)}$	$\hat{\mathbf{s}}_{(or)}$	$\hat{\mathbf{s}}_{(ta)}$
				$\mathbf{R^2}$					
1995	1.000 [0.998,1.000]	0.826 [0.801,0.844]	0.215 [0.188,0.289]	0.186 [0.172,0.220]	0.035 [0.033,0.042]	0.989	na	na	na
1996	0.994 [0.993,0.995]	0.814 [0.787,0.835]	0.284 [0.240,0.373]	0.206 [0.189,0.244]	0.050 [0.046,0.059]	0.968	na	na	na
1997	0.991 [0.989,0.991]	0.796 [0.768,0.819]	0.306 [0.261,0.409]	0.246 [0.223,0.297]	0.093 [0.084,0.108]	0.945	0.888	0.573	0.010
1998	0.991 [0.988,0.992]	0.802 [0.772,0.824]	0.305 [0.265,0.389]	0.220 [0.200,0.247]	0.113 [0.103,0.129]	0.928	0.899	0.488	0.014
1999	0.990 [0.988,0.991]	0.786 [0.752,0.812]	0.283 [0.249,0.345]	0.234 [0.214,0.259]	0.117 [0.109,0.131]	0.912	0.858	0.524	0.000
2000	0.990 [0.984,0.991]	0.719 [0.687,0.742]	0.287 [0.239,0.358]	0.232 [0.201,0.284]	0.080 [0.072,0.094]	0.907	0.873	0.648	0.001
2001	0.988 [0.985,0.989]	0.711 [0.680,0.735]	0.262 [0.228,0.319]	0.227 [0.203,0.272]	0.092 [0.085,0.105]	0.900	0.875	0.671	0.001
2002	0.985 [0.981,0.986]	0.699 [0.667,0.726]	0.299 [0.257,0.372]	0.235 [0.200,0.310]	0.126 [0.114,0.146]	0.905	0.849	0.593	0.005
2003	0.983 [0.981,0.983]	0.680 [0.647,0.706]	0.307 [0.265,0.382]	0.229 [0.196,0.296]	0.124 [0.111,0.152]	0.913	0.786	0.408	0.004
2004	0.979 [0.973,0.980]	0.685 [0.652,0.708]	0.272 [0.232,0.343]	0.219 [0.182,0.282]	0.174 [0.151,0.209]	0.915	0.721	0.146	0.000
				$\mathbf{p-value\ \chi^2}$					
1995	1.000 [1.000,1.000]	0.484 [0.020,0.980]	0.318 [0.000,0.960]	0.195 [0.000,0.822]	0.173 [0.000,0.784]	1.000	na	na	na
1996	0.978 [0.978,0.978]	0.493 [0.007,0.984]	0.201 [0.000,0.861]	0.124 [0.000,0.756]	0.248 [0.000,0.915]	0.993	na	na	na
1997	0.812 [0.812,0.812]	0.567 [0.026,0.989]	0.124 [0.000,0.734]	0.156 [0.000,0.834]	0.386 [0.001,0.948]	0.954	0.000	0.000	0.000
1998	0.948 [0.945,0.945]	0.499 [0.018,0.985]	0.216 [0.000,0.871]	0.110 [0.000,0.709]	0.271 [0.000,0.902]	0.827	0.000	0.000	0.000
1999	0.741 [0.741,0.741]	0.439 [0.012,0.959]	0.350 [0.001,0.940]	0.211 [0.000,0.861]	0.217 [0.000,0.862]	0.478	0.000	0.000	0.000
2000	0.957 [0.957,0.957]	0.274 [0.002,0.908]	0.378 [0.002,0.956]	0.096 [0.000,0.619]	0.204 [0.000,0.869]	0.627	0.000	0.000	0.000
2001	0.999 [0.999,0.999]	0.218 [0.001,0.811]	0.210 [0.000,0.857]	0.058 [0.000,0.447]	0.061 [0.000,0.546]	0.151	0.000	0.000	0.000
2002	0.996 [0.930,0.930]	0.202 [0.000,0.825]	0.157 [0.000,0.781]	0.154 [0.000,0.826]	0.083 [0.000,0.655]	0.121	0.000	0.000	0.000
2003	0.930 [0.915,0.915]	0.154 [0.000,0.761]	0.317 [0.000,0.922]	0.213 [0.000,0.879]	0.164 [0.000,0.846]	0.096	0.000	0.000	0.000
2004	0.915 [0.000,0.000]	0.182 [0.000,0.852]	0.203 [0.000,0.850]	0.091 [0.000,0.600]	0.175 [0.000,0.803]	0.062	0.000	0.000	0.000

In square brackets 95 per cent empirical confidence interval.

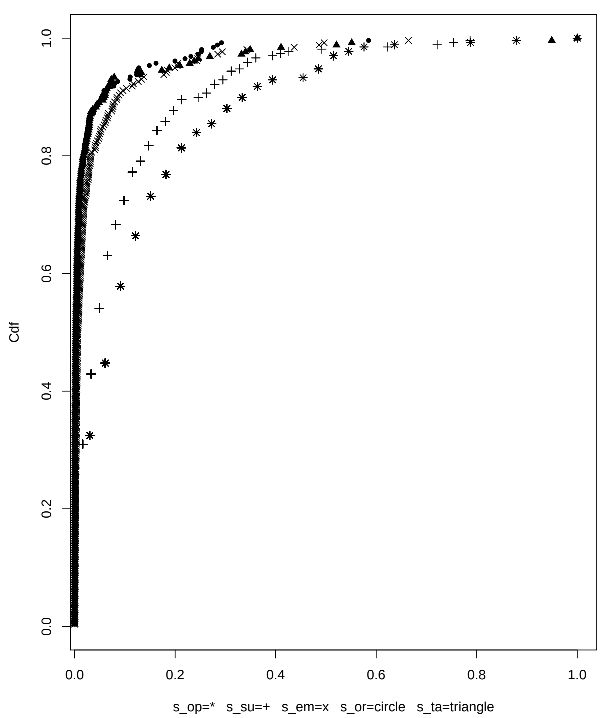
Table 10: Yearly percentage of missing values for continuous size metrics (prior to data imputation)

Year	$s_{(em)}$	$s_{(or)}$	$s_{(sa)}$	$s_{(ta)}$
	Prior to data imputation			
1992	99.0	98.7	98.8	98.6
1993	95.5	94.4	95.2	93.0
1994	78.4	73.4	80.3	69.9
1995	40.5	35.0	50.8	29.1
1996	38.8	33.3	50.5	28.3
1997	37.2	31.1	46.5	26.2
1998	33.8	27.3	45.3	25.5
1999	29.4	22.2	44.5	21.2
2000	24.8	19.1	44.7	18.2
2001	23.2	17.5	43.6	16.4
2002	19.1	12.2	38.9	11.5
2003	21.7	16.5	42.2	16.2
2004	60.1	62.3	75.6	64.4
	After data imputation			
1992	99.0	98.7	98.8	98.6
1993	95.5	94.4	95.2	93.0
1994	78.4	73.4	80.3	69.9
1995	36.1	30.0	46.6	24.7
1996	34.3	28.7	46.2	23.6
1997	32.4	26.2	41.3	20.8
1998	26.4	21.6	38.8	18.8
1999	22.4	16.0	37.5	15.3
2000	17.3	11.7	37.0	11.3
2001	15.4	9.3	35.5	9.0
2002	12.0	5.8	32.4	6.7
2003	10.9	5.0	32.0	6.2
2004	14.7	9.2	35.5	9.3

Figure 1: Empirical cumulative density function for normalized size measures



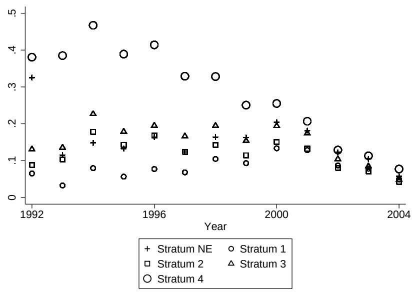
(a) Year 1995



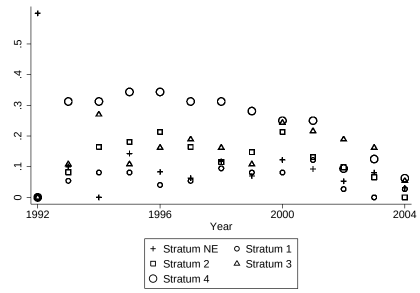
(b) Year 2004

Figure 2: Absolute and proportional growth of predicted and observed number of cumulated opportunities

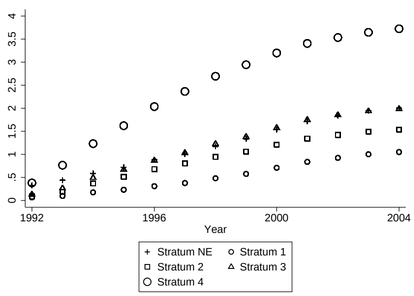
Absolute growth



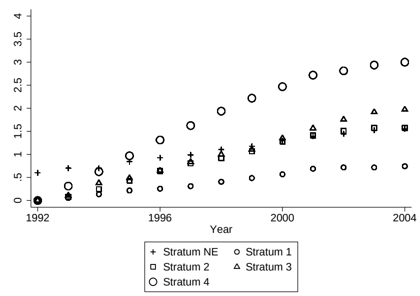
(a) $\overline{\hat{y}}_t$: Avg. (predicted)



(b) $\overline{y}_t$: Avg. (observed)

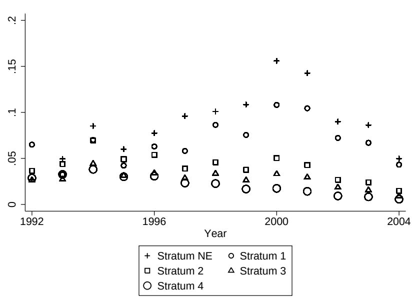


(c) $\sum_t \overline{\hat{y}}_t$: Cum. avg. (predicted)

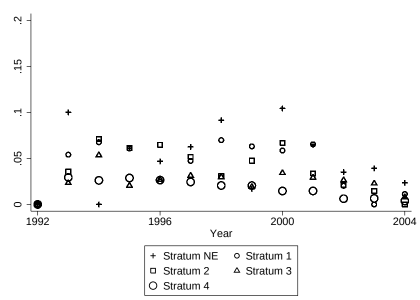


(d) $\sum_t \overline{y}_t$: Cum. avg. (observed)

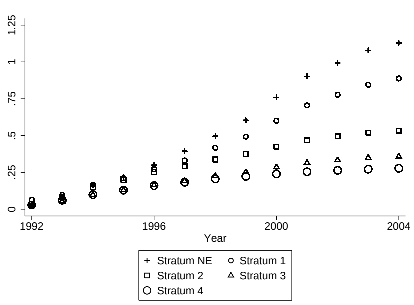
Proportional growth



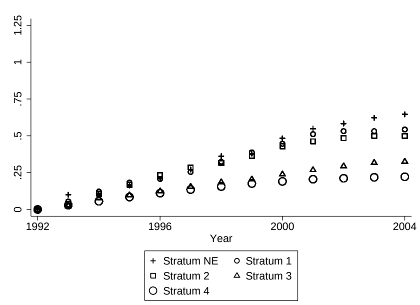
(e) $\overline{\left(\frac{\hat{y}}{s_{-1}}\right)_t}$: Avg. (predicted %)



(f) $\overline{\left(\frac{y}{s_{-1}}\right)_t}$: Avg. (observed %)



(g) $\sum_t \overline{\left(\frac{\hat{y}}{s_{-1}}\right)_t}$: Cum. avg. (predicted %)



(h) $\sum_t \overline{\left(\frac{y}{s_{-1}}\right)_t}$: Cum. avg. (observed %)

A Solutions to the Initial Conditions Problem

A.1 Heckman's Procedure

Heckman (1981a,b) suggests a parametric approach to account for the initial conditions and meantime integrate the unobserved heterogeneity out of the joint density function that we have formulated in Eq. (8). Heckman's idea yields to the following conditional joint density function:

$$P\left(y_{i0}, y_{i1}, \cdots, y_{iT} | y_{i,t-1}, \mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i; \boldsymbol{\theta}\right) = \tag{14a}$$
$$\int G\left(y_{i0} | \mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i, \tilde{\alpha}\right) \prod_{t=1}^{T} \Phi\left[\left(\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \mathbf{z}_{1i}\boldsymbol{\lambda}_1 + \tilde{\alpha}\right)\left(2y_{it} - 1\right)\right] h\left(\tilde{\alpha} | \mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i\right) d\tilde{\alpha},$$

where $G(y_{i0}|\mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i, \tilde{\alpha})$ is the conditional cumulative density function of the initial value (condition) of the dependent variable, $h(\tilde{\alpha}|\mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i)$ is the conditional probability density function of the unobserved heterogeneity, and $\mathbf{v}_i$ is a new vector of time-invariant variables that will be defined shortly. For the sub-sample of MNEs that enter in the industry at time $\tau_i > 1$, i.e., the new entrants, the above equation slightly modifies to:

$$P\left(y_{i\tau_i}, \cdots, y_{iT} | y_{i,\tau_i - 1} = 0, y_{i,t-1}, \mathbf{x}_i, \mathbf{z}_{1i}; \boldsymbol{\theta}\right) = \tag{14b}$$
$$\int \prod_{t=\tau_i}^{T} \Phi\left[\left(\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \mathbf{z}_{1i}\boldsymbol{\lambda}_1 + \tilde{\alpha}\right)\left(2y_{it} - 1\right)\right] h\left(\tilde{\alpha} | \mathbf{x}_i\right) d\tilde{\alpha}.$$

The remainder of the section concentrates on the balanced panel, as it is this one that suffers the initial conditions problem. We work with the additional assumption $\tilde{\alpha}| \left(\mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i\right) \sim NID(0, \sigma^2_{\tilde{\alpha}})$.[25] We write down the underlying profit function for the initial period as,

$$\pi_{i0} = \mathbf{v}_i\boldsymbol{\eta} + \mathbf{z}_{1i}\boldsymbol{\lambda_0} + u_{i0}, \tag{15a}$$

where $\mathbf{v}_i$, as suggested in Heckman (1981a,b), includes any exogenous pre-sample regressor. The initial period composite error, u_{i0} is given by the sum of $\delta_0\tilde{\alpha}_i + \varepsilon_{i0}$, with $\varepsilon_0 \sim N(0,1)$, and orthogonal to $\tilde{\alpha}$, $\mathbf{x}_i$, $\mathbf{z}_{1i}$, and $\mathbf{v}_i$.

The initial conditions problem arises because of the correlation between the random variables u_0 and $\tilde{\alpha}$, conditional on $\mathbf{x}_i$, $\mathbf{z}_{1i}$, and $\mathbf{v}_i$. In order to familiarize with the issue, we rewrite the initial profit function of Eq. (15a) as

$$\tilde{\pi}_{i0} = \mathbf{v}_i\tilde{\boldsymbol{\eta}} + \mathbf{z}_{1i}\tilde{\boldsymbol{\lambda}}_0 + \tilde{\delta}_0\tilde{\alpha}_i + \tilde{\varepsilon}_{i0}, \tag{15b}$$

which is nothing but Eq. (15a) with the error decomposition substituted in, and with both sides of the equation multiplied by $\sigma_{u_0|\tilde{\alpha}}$, a parameter to be defined below. In order to proceed with the parametric analysis, we specify the joint distribution for $(\tilde{\alpha}, u_0|\mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i)$ to be the bivariate normal

$$\begin{pmatrix} \tilde{\alpha} \\ u_0 \end{pmatrix} |\mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i \Bigg) \sim NID\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\tilde{\alpha}} & \rho_0\sigma_{\tilde{\alpha}}\sigma_{u_0} \\ \rho_0\sigma_{\tilde{\alpha}}\sigma_{u_0} & \sigma^2_{u_0} \end{pmatrix}\right], \tag{16}$$

which implies that the conditional distribution for u_0 given $\tilde{\alpha}$ is $\sim NID[\rho_0\frac{\sigma_{u_0}}{\sigma_{\tilde{\alpha}}}, \sigma^2_{u_0}(1-\rho^2_0)]$, where ρ_0 denotes the correlation coefficient between $\tilde{\alpha}$ and u_0. Because of the assumption of NID of $(\tilde{\alpha}, u_0)$, $\tilde{\varepsilon}_0$ is itself $\sim NID\left[0, \sigma^2_{u_0}\left(1-\rho^2_0\right)\right]$. Thus, we are allowed to re-formulate the initial latent profit function as

$$\tilde{\pi}_{i0} = \mathbf{v}_i\tilde{\boldsymbol{\eta}} + \mathbf{z}_{1i}\tilde{\boldsymbol{\lambda}}_0 + \rho_0\frac{\sigma_{u_0}}{\sigma_{\tilde{\alpha}}}\tilde{\alpha}_i + \left(\sigma_{u_0}\sqrt{1-\rho^2_0}\right)\varepsilon_{i0}. \tag{17}$$

If we divide both sides of the above equation by $\sigma_{u_0|\tilde{\alpha}} \equiv \sigma_{u_0}\sqrt{1-\rho^2_0}$, we have back the latent profit function (15a), $\pi_{i0} = \mathbf{v}_i\boldsymbol{\eta} + \mathbf{z}_{1i}\boldsymbol{\lambda_0} + \delta_0\tilde{\alpha}_i + \varepsilon_{i0}$, with $\pi_{i0} = \frac{\tilde{\pi}_{i0}}{\sigma_{u_0}\sqrt{1-\rho^2_0}}$, $\boldsymbol{\eta} = \frac{\tilde{\boldsymbol{\eta}}}{\sigma_{u_0}\sqrt{1-\rho^2_0}}$, $\boldsymbol{\lambda_0} = \frac{\tilde{\boldsymbol{\lambda}}_0}{\sigma_{u_0}\sqrt{1-\rho^2_0}}$, $\delta_0 = \frac{1}{\sigma_{\tilde{\alpha}}} \cdot \frac{\rho_0}{\sqrt{1-\rho^2_0}}$.

[25] So that $(\delta_t\tilde{\alpha}|\mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i) \sim NID(0, \delta^2_t\sigma^2_{\tilde{\alpha}})$ has a variance that is allowed to vary over time.

We now have all the components to parameterize the functions $G(\cdot)$ and $h(\cdot)$ displayed in Eq. (14b) with the CDF $\Phi(\mathbf{v}_i\boldsymbol{\eta} + \mathbf{z}_{1i}\boldsymbol{\lambda_0} + \delta_0\sigma_{\tilde{\alpha}}\alpha)$ and pdf $\phi(\alpha)$, respectively. The conditional joint density is formulated in full as:

$$P\left(y_{i0}, y_{i1}, \cdots, y_{iT}|y_{i,t-1}, \mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i, \boldsymbol{\theta^F}\right) = \tag{18a}$$
$$\int_{-\infty}^{\infty} \Phi\left[(\mathbf{v}_i\boldsymbol{\eta} + \mathbf{z}_{1i}\boldsymbol{\lambda_0} + \delta_0\sigma_{\tilde{\alpha}}\alpha)(2y_{i0}-1)\right]\prod_{t=1}^{T}\Phi\left[(\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \mathbf{z}_{1i}\boldsymbol{\lambda}_1 + \delta_t\sigma_{\tilde{\alpha}}\alpha)(2y_{it}-1)\right]\phi(\alpha)d\alpha,$$

with $\alpha = \frac{\tilde{\alpha}}{\sigma_{\tilde{\alpha}}}$ and δ_T set to one to allow for the identification of $\sigma_{\tilde{\alpha}}$, and $\boldsymbol{\theta}^F$ indicating the parameters to be estimated in the free correlation scenario.

It is very common in empirical papers that use dynamic Probit models, to impose equicorrelation in the composite error. Such an assumption sets the δ_t parameters (for $t = 1, \cdots, T$) to one, and consequently the density in Eq. (18a) simplifies to:

$$P\left(y_{i0}, y_{i1}, \cdots, y_{iT}|y_{i,t-1}, \mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i, \boldsymbol{\theta^E}\right) = \tag{18b}$$
$$\int_{-\infty}^{\infty} \Phi\left[(\mathbf{v}_i\boldsymbol{\eta} + \mathbf{z}_{1i}\boldsymbol{\lambda_0} + \delta_0\sigma_{\tilde{\alpha}}\alpha)(2y_{i0}-1)\right]\prod_{t=1}^{T}\Phi\left[(\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \mathbf{z}_{1i}\boldsymbol{\lambda}_1 + \sigma_{\tilde{\alpha}}\alpha)(2y_{it}-1)\right]\phi(\alpha)d\alpha,$$

with $\boldsymbol{\theta}^E$ denoting the set of parameters to be estimated in the equicorrelation case.

The next two sections outline two alternative estimation procedures to estimate Heckman's solution to the initial conditions problem within the `Stata` environment.

A.1.1 Arulampalam and Stewart Standard Random Effects

Arulampalam and Stewart (2009) propose an approach to estimate Heckman's estimator, that can easily be implemented in softwares that deal with heteroscedasticity, such as the add-on program for `Stata`, `gllamm`. For the free correlated case, the authors suggest employing $T+1$ time dummies, d_{it}^{τ}, to identify the free correlation parameters. Each time dummy is set to 1 if $\tau = t$ and 0 otherwise. The conditional probability function of an expansion at time t writes as follows:

$$P\left(y_{it} = 1|y_{i,t-1}, \mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i, \alpha_i; \boldsymbol{\theta^{AS,F}}\right) = \tag{19a}$$
$$\Phi\left[(\mathbf{v}_i\boldsymbol{\eta} + \mathbf{z}_{1i}\boldsymbol{\lambda_0} + \delta_0\tilde{\alpha}_i)\, d_{it}^0 + (\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \mathbf{z}_{1i}\boldsymbol{\lambda}_1)(1 - d_{it}^0) + \tilde{\alpha}_i\sum_{\tau=1}^{T}\delta_\tau d_{it}^{\tau}\right],$$

with $t = 1, \cdots, T$ and, as mentioned earlier, δ_T set to one.

To the cost of repeating ourselves, the alternative case of equicorrelation sets the $\delta_\tau$parameters to one, for $\tau = 1, \cdots, T$, and subsequently the conditional probability of an expansion simplifies to

$$P\left(y_{it} = 1|y_{i,t-1}, \mathbf{x}_i, \mathbf{z}_{1i}, \mathbf{v}_i, \alpha_i; \boldsymbol{\theta^{AS,E}}\right) = \tag{19b}$$
$$\Phi\left[(\mathbf{v}_i\boldsymbol{\eta} + \mathbf{z}_{1i}, \boldsymbol{\lambda_0} + \delta_0\tilde{\alpha}_i)\, d_{it}^0 + (\gamma y_{i,t-1} + \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \mathbf{z}_{1i}\boldsymbol{\lambda}_1 + \tilde{\alpha}_i)(1 - d_{it}^0)\right].$$

A.1.2 Stewart's Method

Stewart has written a `Stata` code `redprob` that can estimate both cases of free correlation, Eq. (18a), and equicorrelation, Eq. (18b), using Gaussian-Hermite quadrature. Details on the use of his program are given in Stewart (2006).

A.2 Orme's Two-Step Procedure

Orme (2001) proposes an easy-to-use two-step estimator that is suitable for cases of low correlation between the initial conditions and unobserved heterogeneity (weak correlation). Under the assumption that y_{it} and y_{i0} are independent of one another, one can integrate the unobserved heterogeneity out of the joint density in the following way

$$P\left(y_{i1}, \cdots, y_{iT}|y_{i0}; \boldsymbol{\theta}\right) = \frac{\int_{-\infty}^{\infty} F\left(y_{i1}, \cdots, y_{iT}|\alpha\right) G\left(y_{i0}|\alpha\right) h\left(\alpha\right) d\alpha}{\int_{-\infty}^{\infty} G\left(y_{i0}|\alpha\right) h\left(\alpha\right) d\alpha}. \tag{20}$$

If we consider y_{i0} as exogenous, the above conditional joint probability simplifies to

$$P\left(y_{i1},\cdots,y_{iT}|y_{i0};\boldsymbol{\theta}\right)=\int_{-\infty}^{\infty}F\left(y_{i1},\cdots,y_{iT}|\alpha\right)h\left(\alpha\right)d\alpha, \tag{21}$$

where we have exploited the condition $G\left(y_{i0}|\alpha\right)_{|\rho_0=0}=G\left(y_{i0}\right)$, with ρ_0 already defined as correlation coefficient between $\tilde{\alpha}$ and u_0. By virtue of local approximation, Orme (2001) extends the logic to low values of ρ_0, and approximates the conditional joint density of Eq. (20) with

$$P^a\left(y_{i1},\cdots,y_{iT}|y_{i0};\boldsymbol{\theta}^O\right)=\int_{-\infty}^{\infty}F^a\left(y_{i1},\cdots,y_{iT}|\alpha\right)h\left(\alpha\right)d\alpha, \tag{22}$$

where the superscript a denotes an approximation of the function.

Similarly to Heckman (1981a,b), Orme (2001) assumes bivariate normality for the unobserved heterogeneity and the initial composite error but, differently from Heckman, he conditions the unobserved heterogeneity to the initial composite error, so to have

$$\tilde{\alpha}_i=\rho_0\frac{\sigma_{\tilde{\alpha}}}{\sigma_{u_0}}u_{i0}+\sqrt{1-\rho_0^2}\sigma_{\tilde{\alpha}}\varepsilon_{i0}, \tag{23}$$

with ε_0 assumed to be conditionally independent of u_0, as well as $\mathbf{x}_i$, $\mathbf{z}_{1i}$, $\mathbf{v}_i$ and distributed as $N(0,1)$.

The estimation of the approximated conditional density formulated in Eq. (22) requires a two-step procedure:

Step 1 Normalize $\sigma^2_{u_0}=1$, so that $G\left(y_{i0}\right)$ can be estimated as an ordinary Probit. Having $E\left(\varepsilon_{i0}|y_{i0}\right)=0$, by construction, the conditional expectation for the initial composite error yields

$$E\left(u_{i0}|y_{i0}\right)=\frac{(2y_{i0}-1)\phi\left(\mathbf{v}_i\boldsymbol{\eta}+\mathbf{z}_{1i}\boldsymbol{\lambda_0}\right)}{\Phi\left[(2y_{i0}-1)\left(\mathbf{v}_i\boldsymbol{\eta}+\mathbf{z}_{1i}\boldsymbol{\lambda_0}\right)\right]}. \tag{24}$$

Step 2 Augment the specification in Eq. (6) in order to have:

$$\pi_{it}=\gamma y_{i,t-1}+\mathbf{x}_{1it}\boldsymbol{\beta}_1+\mathbf{x}_{2it}\boldsymbol{\beta}_2+\mathbf{z}_{1i}\boldsymbol{\lambda_1}+\rho_0\sigma_{\tilde{\alpha}}u_{i0}+\sqrt{1-\rho_0^2}\sigma_{\tilde{\alpha}}\varepsilon_{i0}+\varepsilon_{it},\quad t=1,2,\cdots,T, \tag{25}$$

where ε is $NID(0,1)$. Next, derive the following conditional joint density:

$$P^a\left(y_{i1},\cdots,y_{iT}|y_{i,t-1},\mathbf{x}_i,\mathbf{z}_{1i},\mathbf{v}_i,y_{i0};\boldsymbol{\theta^{O,E}}\right)= \prod_{t=1}^{T}\Phi\left[\left(\gamma y_{i,t-1}+\mathbf{x}_{1it}\boldsymbol{\beta}_1+\mathbf{x}_{2it}\boldsymbol{\beta}_2+\mathbf{z}_{1i}\boldsymbol{\lambda}_1+\delta_0\hat{u}_{i0}\right)\left(2y_{it}-1\right)\right]. \tag{26}$$

where $\hat{u}_{i0}$ is the Probit residual from the first stage, and the parameter δ_0 is the product $\rho_0\sigma_{\tilde{\alpha}}$.

Arulampalam and Stewart (2009), recognize that Orme's two-step procedure can be generalized to allow for free correlation. It is enough to interact the initial composite error with time dummies d_t, yielding the following modified version of Eq. (25):

$$\pi_{it}=\gamma y_{i,t-1}+\mathbf{x}_{1it}\boldsymbol{\beta}_1+\mathbf{x}_{2it}\boldsymbol{\beta}_2+\mathbf{z}_{1i}\boldsymbol{\lambda_1}+d_t\rho_0\sigma_{\tilde{\alpha}}u_{i0}+\sqrt{1-\rho_0^2}\sigma_{\tilde{\alpha}}\varepsilon_{i0}+\varepsilon_{it}, \tag{27}$$

which has augmented conditional joint density

$$P_i^a\left(y_{i1},\cdots,y_{iT}|,y_{iT}|y_{i,t-1},\mathbf{x}_i,\mathbf{z}_{1i},\mathbf{v}_i,y_{i0};\boldsymbol{\theta^{O,F}}\right)= \prod_{t=1}^{T}\Phi\left[\left(\gamma y_{i,t-1}+\mathbf{x}_{1it}\boldsymbol{\beta}_1+\mathbf{x}_{2it}\boldsymbol{\beta}_2+\mathbf{z}_{1i}\boldsymbol{\lambda}_1+\delta_t\hat{u}_{i0}\right)\left(2y_{it}-1\right)\right], \tag{28}$$

where $\delta_t\equiv d_t\rho_0\sigma_{\tilde{\alpha}}$.

A.3 Wooldridge's Method

Wooldridge (2005) proposes conditioning the distribution of the unobserved heterogeneity on the initial value of the dependent variable and full history of the (exogenous) time-varying covariates. Wooldridge's idea produces the following conditional joint density function

$$
\begin{aligned}
&P\left(y_{i1},\cdots,y_{iT}|y_{i0},y_{i,t-1},\mathbf{x}_i,\mathbf{z}_{1i},\boldsymbol{\theta}^{\boldsymbol{W}}\right)= \\
&\int_{-\infty}^{\infty}\left[\prod_{t=1}^{T}f_t\left(y_{it}=1|y_{i,t-1},\mathbf{x}_i,\mathbf{z}_{1i},c\right)\right]h\left(c|y_{i0},\mathbf{z}_{1i}\right)dc.
\end{aligned}
\tag{29}
$$

Using Mundlak's approach, we parameterize the general functions in Eq. (29) as

$$
\begin{aligned}
h\left(c|y_{i0},\mathbf{z}_{1i}\right) &\sim N\left(\mu_0 y_{i0}+\mathbf{z}_{1i}\boldsymbol{\lambda_0},\sigma_\alpha^2\right)\\
f_t\left(y_{it}=1|y_{i,t-1},\mathbf{x}_i,\mathbf{z}_{1i},c_i\right) &= \Phi\left[\left(\gamma y_{i,t-1}+\mathbf{x}_{1it}\boldsymbol{\beta}_1+\mathbf{x}_{2it}\boldsymbol{\beta}_2+\mathbf{z}_{1i}\boldsymbol{\lambda}_1+c_i\right)\left(2y_{it}-1\right)\right].
\end{aligned}
$$

One limitation of Wooldridge's methodology is that one cannot identify the time-invariant variables that are correlated with the unobserved heterogeneity. Nevertheless they should all be included in the exogenous variables, as they bring about explanatory power to the estimation process. The major advantage of Wooldridge's methodology is that it can be easily estimated in `Stata` using the command `xtprobit`.

The densities to be estimated, in case of equicorrelation are

$$
\begin{aligned}
&P\left(y_{i1},\cdots,y_{iT}|y_{i0},y_{i,t-1},\mathbf{x}_i,\mathbf{z}_{1i},\boldsymbol{\theta}^{\boldsymbol{W,E}}\right)= \\
&\int_{-\infty}^{\infty}\prod_{t=1}^{T}\Phi\left\{\left[\gamma y_{i,t-1}+\mathbf{x}_{1it}\boldsymbol{\beta}_1+\mathbf{x}_{2it}\boldsymbol{\beta}_2+\mathbf{z}_{1i}\left(1-d_{it}^0\right)\boldsymbol{\lambda}_1+\mathbf{z}_{1i}d_{it}^0\boldsymbol{\lambda_0}+\mu_0 y_{i0}+\sigma_{\tilde{\alpha}}\alpha\right]\left(2y_{it}-1\right)\right\}\phi(\alpha)d\alpha,
\end{aligned}
\tag{30a}
$$

and in case of free correlation

$$
\begin{aligned}
&P\left(y_{i1},\cdots,y_{iT}|y_{i0},y_{i,t-1},\mathbf{x}_i,\mathbf{z}_{1i},\boldsymbol{\theta}^{\boldsymbol{W,F}}\right)= \\
&\int_{-\infty}^{\infty}\prod_{t=1}^{T}\Phi\left[\left(\gamma y_{i,t-1}+\mathbf{x}_{1it}\boldsymbol{\beta}_1+\mathbf{x}_{2it}\boldsymbol{\beta}_2+\mathbf{z}_{1i}\left(1-d_{it}^0\right)\boldsymbol{\lambda}_1+\mathbf{z}_{1i}d_{it}^0\boldsymbol{\lambda_0}+\mu_0 y_{i0}+\delta_t\sigma_{\tilde{\alpha}}\alpha\right)\left(2y_{it}-1\right)\right]\phi(\alpha)d\alpha.
\end{aligned}
\tag{30b}
$$

B The copula

We enrich our notation further. Let $\bar{\Re}$ denote the extended real line $[-\infty,\infty]$ and $\Im$ the unit interval $[0,1]$. We define an H-dimensional joint distribution function G with domain $\bar{\Re}^H$ a function that has the following properties:

1. $G(\infty,\cdots,\infty,s_h,\infty,\cdots,\infty)=G_h(s_h)$ for any $h\leq H$;

2. $G(\infty,\cdots,\infty,\cdots,\infty)=1$;

3. $G(s_1,\cdots,s_H)=0$ if $s_h=-\infty$ for any $h\leq H$, and in this case we say that G is *grounded*;

4. G is H-increasing.

In order to explain what is an H-increasing function we need first to define how to compute the volume of a H-box.

Definition B.1 *Let $\mathcal{S}_1,\mathcal{S}_2,\cdots,\mathcal{S}_H$ be nonempty subsets of $\overline{R}^H$, and let G be an H-place real function such that the domain $DomG=\mathcal{S}_1\times\mathcal{S}_2\times\cdots\times\mathcal{S}_H$. Let $\mathbf{B}=[\mathbf{a},\mathbf{b}]$ be an H-box all of whose vertices $\mathbf{c}$ are in $DomG$. Then the H- volume of $\mathbf{B}$ is given by*

$$
V(\mathbf{B})=\sum sgn(\mathbf{c})\,G(\mathbf{c}),
\tag{31}
$$

where the sum is taken over all vertices $\mathbf{c}$ *of* $\mathbf{B}$*, and* $sgn(\mathbf{c})$ *is given by*

$$sgn(\mathbf{c}) = \begin{cases} 1, & \text{if } c_h = a_h \text{ for an even sum of } h's, \\ -1, & \text{if } c_h = a_h \text{ for an odd sum of } h's. \end{cases} \quad (32)$$

Definition B.2 *An H-place real function G is H-increasing if $V(\mathbf{B}) \geq 0$ for all H-boxes whose vertices lie in $DomG$.*

Because the marginal distribution functions may not be continuous, prior to defining a copula it is worth defining a subcopula.

Definition B.3 *An H-dimensional subcopula is a function C' with the following properties:*

1. *$DomC' = \mathcal{U}_1 \times \mathcal{U}_2 \times \cdots \times \mathcal{U}_H$, where each $\mathcal{U}_h$, $h \leq H$, is a subset of $\Im$ containing 0 and 1;*
2. *C' is grounded and H-increasing;*
3. *C' has margins C_h which satisfy*

$$C'_h(u) = u \text{ for all } u \text{ in } \mathcal{U}_h. \quad (33)$$

Note that for every $\mathbf{u} \in DomC'$, $0 \leq C'(\mathbf{u}) \leq 1$, so that $RanC'$ is also a subset of $\Im$.

Definition B.4 *An H-dimensional copula is an H-dimensional subcopula whose domain is $\Im^H$.*

In the rest of the section we introduce the parametric copula functions that are available for the multivariate case of $H > 2$. We separate them in two classes: Elliptical and Archimedean.

B.1 Elliptical copula

If we differentiate Eq. (5) we get the density of the Elliptical copula

$$c(u_1, u_2, \cdots, u_H; \boldsymbol{\varrho}) = \frac{g\left[G_1^{-1}(u_1), G_2^{-1}(u_2), \cdots, G_H^{-1}(u_H)\right]}{\prod_{h=1}^{H} g_h\left[G_h^{-1}(u_h)\right]}, \quad (34)$$

where g is the joint probability density function of the Elliptical distribution, and g_h the marginal density functions. If we assume the marginal distribution functions to be standard normals, $G_h = \Phi$, we have the H-dimensional copula standard normal (gaussian), whose density is given by

$$c(u_1, u_2, \cdots, u_H; \boldsymbol{\Gamma}) = |\boldsymbol{\Gamma}|^{-1/2} \exp\left\{-\frac{1}{2}\left[\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_H)\right]' \left(\mathbf{I}_H - \boldsymbol{\Gamma}^{-1}\right)\left[\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_H)\right]\right\}, \quad (35)$$

where $\mathbf{I}_H$ denotes the identity matrix and $\boldsymbol{\Gamma}$ the dispersion matrix.

B.2 Archimedean copula

Theorem B.5 (Nelsen (2006)'s Archimedean copula) *Let φ be a continuous strictly decreasing function from $\Im$ to $[0,\infty]$ such that $\varphi(0) = \infty$ and $\varphi(1) = 0$, and let φ^{-1} denote the inverse of φ. If C^H is the function from $\Im^H$ to $\Im$ given by*

$$C^H(u_1, u_2, \cdots, u_H; \varrho) = \varphi^{-1}\left[\varphi(u_1) + \varphi(u_2) + \cdots + \varphi(u_H)\right], \quad (36)$$

then C^H is a H-copula $\forall h \geq 2$, if and only if φ^{-1} is completely monotonic on $[0,\infty)$.

By parameterizing the generation function φ and its inverse φ^{-1} we obtain some well-known Archimedean copulas:

1. **Clayton copula**: If we let the generator function be $t = \varphi(u; \varrho) = u^{-\varrho} - 1$ for $\varrho > 0$, then its inverse is $\varphi^{-1}(t; \varrho) = (1 + t)^{-1/\varrho}$, and the resulting multivariate copula is

$$C^H(u_1, u_2, \cdots, u_H; \varrho) = \left(u_1^{-\varrho} + u_2^{-\varrho} + \cdots + u_H^{-\varrho} - H + 1\right)^{-1/\varrho}. \tag{37}$$

2. **Frank copula**: If we let the generator function be $t = \varphi(u; \varrho) = -\ln \frac{(e^{-\varrho u} - 1)}{(e^{-\varrho} - 1)}$ for $\varrho > 0$, then its inverse is $\varphi^{-1}(t; \varrho) = -\frac{1}{\varrho} \ln\left[1 - (1 - e^{-\varrho}) e^{-t}\right]$, and the resulting multivariate copula is

$$C^H(u_1, u_2, \cdots, u_H; \varrho) = -\frac{1}{\varrho} \ln \left[1 + \frac{(e^{-\varrho u_1} - 1)(e^{-\varrho u_2} - 1) \cdots (e^{-\varrho u_H} - 1)}{(e^{-\varrho} - 1)^{H-1}}\right]. \tag{38}$$

3. **Gumbel copula**: If we let the generator function be $t = \varphi(u; \varrho) = (-\ln u)^{\varrho}$ for $\varrho \geq 1$, then its inverse is $\varphi^{-1}(t; \varrho) = \exp\left(-t^{1/\varrho}\right)$, and the resulting multivariate copula is

$$C^H(u_1, u_2, \cdots, u_H; \varrho) = \exp\left\{-\left[(-\ln u_1)^{\varrho} + (-\ln u_2)^{\varrho} + \cdots + (-\ln u_H)^{\varrho}\right]^{1/\varrho}\right\}. \tag{39}$$

C Construction of the dataset of pharmaceutical MNEs

We utilize the 2005 version of the *Amadeus* database, which covers the time period 1992-2004 and extract data on pharmaceutical related subsidiaries (firms) located in one of the EU-14 countries (EU-15 without Luxembourg). We complement this version with previous 1997 version, which covers the time period 1988-1997. Since we are interested in the European Single Market Programme we focus our attention on the sub-period 1992-2004. We spot a total of 1 809 subsidiaries; of those, about a third overlaps in the two database versions. Furthermore, by investigating additional sources, such as: a built-in ownership database in *Amadeus*, pharmaceutical enterprises website and other formats, we manage to link each subsidiary to a corresponding pharmaceutical multinational enterprise (MNE), narrowing down the total number of units to 278 MNEs.

While the majority of these selected subsidiaries have a unitary unconsolidated account, a minority exhibits a consolidated account, that is, an account that sums over the subsidiary of interest and all its pharmaceutical-related subsidiaries. Luckily we know the total number of subsidiaries that are linked to each of these consolidated firms and recognize that they are a mixture of the aforementioned unconsolidated subsidiaries and a residual number of subsidiaries whose account information is not available in *Amadeus*.[26] Because our interest stands in recovering alternative metrics of size at the MNE-level, we face the issue of possible double-counting, i.e., the situation where an unconsolidated subsidiary is both directly and indirectly (via a consolidated subsidiary) associated to an MNE. We overcome this side effect in the following way. First, with the help of the ownership database available in *Amadeus*, we deduct from the size metrics of the consolidated subsidiaries the amount attributable to the identified European unconsolidated subsidiaries. Next, we average the "residual" size metrics over the remaining number of subsidiaries plus the consolidated subsidiary itself. The reason for this last step is that most of these left over subsidiaries that make the consolidated subsidiary are non European and therefore should not be included in our sample.

The 2005 version of the database registers the year of incorporation of subsidiaries in MNEs. However due to mergers and acquisitions, some firms exhibit a year of acquisition subsequent to the year of incorporation. These subsidiaries should only be counted as part of the MNE after the acquisition takes place. We investigate acquisition cases using the European Union's antitrust database, a database on mergers and acquisitions provided by Kiel Institute of World Economics and other online information sources and find the year of acquisition for 161 subsidiaries.[27]

The size metrics we manage to retain from the database are the discrete metrics - number of cumulated opportunities and number of cumulated subsidiaries- and the four continuous metrics - employment, operational revenues

[26]Though *Amadeus* only collects information on firms located in Europe, it lists the *id* number and country of residence of firms located outside Europe, if those were parents or subsidiaries of a European firm.

[27]European Union's antitrust cases database can be found at `http://ec.europa.eu/comm/competition/antitrust/cases/`. The mergers and acquisition database is called "DOME" (Database on Mergers in Europe). DOME is available at `http://www.ifw-kiel.de/academy/data-bases/dome_e/database-on-mergers-in-europe-dome/?searchterm=dome`.

(turnover), sales and total assets. The discrete metric number of cumulated subsidiaries each MNE owns in a certain year, accounts for the subsidiaries established or acquired up to that year. And anytime we observe an MNE having established or acquired one or more new subsidiaries in a certain year, we regard that MNE as if it has taken up (expanded by) one opportunity in that same year.[28] The continuous monetary size metrics, turnover, total assets and sales are deflated with national GDP price index obtained from EUROSTAT. The base year is 1995. All monetary values are retained in US dollars. We work under the maintained assumption that the volatility of the exchange rates between the EU-14 national currencies is not a severe concern during the time period under investigation, at least not during the sub-period of actual investigation, which begins in year 1995, as explained below.

Unfortunately the continuous size metrics of interest present a heavy number of missing values, as confirmed by Table 10. Take turnover in its first panel for example, the variable shows that among all subsidiaries active in 1992, 98.7 per cent of them are without information on turnover. The percentage of missing values for turnover reduces significantly in 1995 and reaches its lowest level in 2002, before slightly upturning thereafter. The percentages of missing values for the other variables follow a similar trend. The intensive proportion of missing values in early years motivates our choice of 1995 as starting year for our analysis of continuous size metrics. Whenever possible, we replace missing values in the 2005 version of dataset with non-missing values in the earlier 1997 version. We deepen our understanding of the severity of the issue and for each subsidiaries we generate the average percentage of missing values over the four continuous size metrics. About 50 per cent of subsidiaries display an average percentage of missing values below or equal 20 per cent. The percentage of missing values raises to 66 per cent if we consider the 90^{th} percentile of firms. We choose the value of this latter percentile as cut-off point and drop all subsidiaries with average percentages above it. This step brings the sample size of subsidiaries down to 1 598 and MNEs down to 268. We pursue data imputation for intermediary missing values, that is, each year we replace subsidiaries' missing values with their mean values over the previous three years. Since the panel of interest starts in year 1995, the first year the imputation kicks in is 1998. We present the new percentages in the second panel of Table 10. Compared to those in the first panel we observe a substantial gain in the latest years from the process of data imputation and removal of noisy firms.

References

Adelman I. 1958. A stochastic analysis of the size distribution of firms. *Journal of the American Statistical Association* **53**: 893–904.

Amaral L, Buldyrev S, Havlin S, Leschhorn H, Maass P, Salinger M, Stanley H, Stanley M. 1997. Scaling behavior in economics: empirical results for company growth. *Journal De Physique France* **7**: 621–633.

Amisano G, Giorgetti M. 2005. Entry in pharmaceutical submarkets: a Bayesian panel probit approach. *Working Papers ubs0408, University of Brescia, Department of Economics* .

Arellano M, Bond S. 1991. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies* **58**: 277.

Arulampalam W. 1999. A note on estimated effects in random effect probit models. *Oxford Bulletin of Economics and Statistics* **61**: 597–602.

Arulampalam W, Stewart M. 2009. Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators. *Oxford Bulletin of Economics and Statistics* **71**: 659–681.

Axtell R. 2001. Zipf distribution of US firm sizes. *Science* **293**: 1818–1820.

Bain J. 1951. Relation of profit rate to industry concentration: American manufacturing, 1936–1940. *The Quarterly Journal of Economics* **65**: 293–324.

[28]The opportunities taken over before year 1992 are not observable. We calculate it as the product of $s_{(su)}$ in 1992 and the ratio between the number of cumulated opportunities taken over during the period (1992-2004) and the number of cumulated subsidiaries established or acquired during the same period.

Bain J. 1956. *Barriers to New Competition.* Harvard University Press: Cambridge, MA.

Blonigen AB, Tomlin K. 2001. Size and growth of Japanese plants in the United States. *International Journal of Industrial Organization* **19**: 931–952.

Bottazzi G, Dosi G, Lippi M, Pammolli F, Riccaboni M. 2001. Innovation and corporate growth in the evolution of the drug industry. *International Journal of Industrial Organization* **19**: 1161–1187.

Bottazzi G, Secchi A. 2005. Growth and diversification patterns of the worldwide pharmaceutical industry. *Review of Industrial Organization* **26**: 195–216.

Bottazzi G, Secchi A. 2006a. Explaining the distribution of firm growth rates. *The RAND Journal of Economics* **37**: 235–256.

Bottazzi G, Secchi A. 2006b. Gibrat's Law and diversification. *Industrial and Corporate Change* **15**: 847–875.

Buldyrev S, Growiec J, Pammolli F, Riccaboni M, Stanley H. 2007. The growth of business firms: facts and theory. *Journal of the European Economic Association* **5**: 574–584.

Buzzacchi L, Valletti T. 2006. Firm size distribution: testing the "independent submarkets model" in the Italian motor insurance industry. *International Journal of Industrial Organization* **24**: 809–834.

Cabral L, Mata J. 2003. On the evolution of the firm size distribution: facts and theory. *The American Economic Review* **93**: 1075–1090.

Calvo J. 2006. Testing Gibrats Law for small, young and innovating firms. *Small Business Economics* **26**: 117–123.

Cecchini P, Catinat M, Jacquemin A. 1988. *The European challenge, 1992 : the benefits of a single market.* Wildwood House: Aldershot.

Cefis E, Ciccarelli M, Orsenigo L. 2007. From Gibrats legacy to Gibrats fallacy: a Bayesian approach to study the growth of firms. *Structural Change and Economic Dynamics* **18**: 348–369.

Chesher A. 1979. Testing the law of proportionate effect. *The Journal of Industrial Economics* **27**: 403–411.

Cohen W, Klepper S. 1992. The anatomy of industry R&D intensity distributions. *The American Economic Review* **82**: 773–799.

Cohen W, Klepper S. 1996. A reprise of size and R&D. *The Economic Journal* **106**: 925–951.

Cooley F, Quadrini V. 2001. Financial markets and firm dynamics. *The American Economic Review* **91**: 1286–1310.

Davies S, Lyons B. 1982. Seller concentration: The technological explanation and demand uncertainty. *The Economic Journal* **92**: 903–919.

Davies S, Lyons B. 1996. *Industrial organization in the European Union: Structure, strategy, and the competitive mechanism.* Oxford University Press, USA.

de Juan R. 2003. The independent submarkets model: an application to the Spanish retail banking market. *International Journal of Industrial Organization* **21**: 1461–1487.

de Wit G. 2005. Firm size distributions: An overview of steady-state distributions resulting from firm dynamics models. *International Journal of Industrial Organization* **23**: 423–450.

Dunne P, Hughes A. 1994. Age, size, growth and survival: UK companies in the 1980s. *The Journal of Industrial Economics* **42**: 115–140.

Dunne T, Roberts M, Samuelson L. 1988. Patterns of firm entry and exit in US manufacturing industries. *The RAND Journal of Economics* **19**: 495–515.

Dunne T, Roberts M, Samuelson L. 1989. The growth and failure of US manufacturing plants. *The Quarterly Journal of Economics* **104**: 671–698.

EFPIA. 2005. Pharmaceutical industry in figures. The European Federation of Pharmaceutical Industries and Associations (EFPIA).

Ericson R, Pakes A. 1995. Markov-perfect industry dynamics: a framework for empirical work. *The Review of Economic Studies* **62**: 53–82.

Evans D. 1987a. Tests of alternative theories of firm growth. *The Journal of Political Economy* **95**: 657–674.

Evans D. 1987b. The relationship between firm growth, size, and age: estimates for 100 manufacturing industries. *The Journal of Industrial Economics* **35**: 567–581.

Evans D, Jovanovic B. 1989. An estimated model of entrepreneurial choice under liquidity constraints. *The Journal of Political Economy* **97**: 808–827.

Farinas C, Moreno L. 2000. Firms' growth, size and age: a nonparametric approach. *Review of Industrial Organization* **17**: 249–265.

Freeman J. 1986. A piecewise regression analysis for firm-size distributions. *The Statistician* **35**: 547–559.

Genest C, Neslehova J. 2007. A primer on Copulas for count data. *Astin Bulletin* **37**: 475–515.

Geroski P, Machin S. 1991. *Market Dynamics and Entry*. Blackwell: Oxford.

Geroski P, Machin S, Walters C. 1997. Corporate growth and profitability. *The Journal of Industrial Economics* **45**: 171–189.

Gibrat R. 1931. *Les Inégalités Économiques*. Librairie du Recueil Sirey: Paris.

Gibrat R. 1957. On economic inequality. *International Economic Papers* **7**: 53–70.

Hall B. 1987. The relationship between firm size and firm growth in the US manufacturing sector. *The Journal of Industrial Economics* **35**: 583–606.

Hall M, Weiss L. 1967. Firm size and profitability. *The Review of Economics and Statistics* **49**: 319–331.

Hart P. 1962. The size and growth of firms. *Economica* **29**: 29–39.

Hart P, Oulton N. 1996. Growth and size of firms. *The Economic Journal* **106**: 1242–1252.

Hart P, Prais S. 1956. The analysis of business concentration: a statistical approach. *Journal of the Royal Statistical Society. Series A (General)* **119**: 150–191.

Hausman J, Taylor W. 1981. Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric Society* : 1377–1398.

Heckman J. 1981a. Heterogeneity and state dependence. *Studies in Labor Markets* **31**: 91–140.

Heckman J. 1981b. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In Manski C, McFadden D (eds.) *Structural Analysis of Discrete Data and Econometric Applications*. MIT Press: Cambridge, MA, 179–195.

Higson C, Holly S, Kattuman P. 2002. The cross sectional dynamics of the US business cycle: 1950-1999. *Journal of Economic Dynamics and Control* : 1539–1555.

Hjalmarsson L. 1974. The size distribution of establishments and firms derived from an optimal process of capacity expansion. *European Economic Review* **5**: 123–140.

Hoeffding W. 1940. Masstabinvariante Korrelationstheorie. *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin* **5**: 179–233.

Hoeffding W. 1994. Scale-invariant correlation theory. In Fisher N, Sen P (eds.) *The Collected Works of Wassily Hoeffding*. Springer-Verlag: New York, 57–107.

Howells J. 1992. Pharmaceuticals and europe 1992: the dynamics of industrial change. *Environment and Planning A* **24**: 33–48.

Huber P, Pfaffermayr M. 2010. Testing for conditional convergence in variance and skewness: The firm size distribution revisited. *Oxford Bulletin of Economics and Statistics* **72**: 648–668.

Huynh P, Petrunia J. 2010. Age effects, leverage and firm growth. *Journal of Economic Dynamics & Control* **34**: 1003–1013.

Hymer S, Pashigian P. 1962. Firm size and rate of growth. *The Journal of Political Economy* **70**: 556–569.

Ijiri Y, Simon H. 1964. Business firm growth and size. *The American Economic Review* **54**: 77–89.

Ijiri Y, Simon H. 1967. A model of business firm growth. *Econometrica: Journal of the Econometric Society* **35**: 348–355.

Jovanovic B. 1982. Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society* **50**: 649–670.

Klepper S. 1996. Entry, exit, growth, and innovation over the product life cycle. *The American Economic Review* **86**: 562–583.

Klepper S, Graddy E. 1990. The evolution of new industries and the determinants of market structure. *The RAND Journal of Economics* **21**: 27–44.

Klepper S, Simons K. 2000. The making of an oligopoly: firm survival and technological change in the evolution of the US tire industry. *The Journal of Political Economy* **108**: 728–760.

Klepper S, Thompson P. 2006. Submarkets and the evolution of market structure. *The RAND Journal of Economics* **37**: 861–886.

Klette T, Kortum S. 2004. Innovating firms and aggregate innovation. *The Journal of Political Economy* **112**: 986–1018.

Kojadinovic I, Yan J. 2010. Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software* **34**: 1–20.

Kotzian P. 2004. Pharmaceutical r&d in the setting of incomplete european integration. *International Journal of the Economics of Business* **11**: 175–195.

Kripfganz S, Schwarz C. 2012. Estimation of linear dynami panel data models with time-invariant regressors. Manuscript.

Kumar M. 1985. Growth, acquisition activity and firm size: evidence fomr the united kingdom. *The Journal of Industrial Economics* **33**: 327–338.

Lotti F, Santarelli E, Vivarelli M. 2003. Does Gibrat's Law hold among young, small firms? *The Journal of Evolutionary Economics* **13**: 213–235.

Lucas Jr R. 1978. On the size distribution of business firms. *The Bell Journal of Economics* **9**: 508–523.

Luttmer E. 2006. Consumer search and firm growth .

Luttmer E. 2007. Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics* **122**: 1103–1144.

Luttmer E. 2011. On the mechanics of firm growth. *The Review of Economic Studies* **78**: 1042–1068.

Mansfield E. 1962. Entry, Gibrat's Law, innovation, and the growth of firms. *The American Economic Review* **52**: 1023–1051.

Mason E. 1939. Price and production policies of large-scale enterprise. *The American Economic Review* **29**: 61–74.

Mason E. 1948. The current status of the monopoly problem in the United States. *Harvard Law Review* **62**: 1265–1285.

Matraves C. 1999. Market structure, r&d and advertising in the pharmaceutical industry. *The Journal of Industrial Economics* **47**: 169–194.

Mundlak Y. 1978. On the pooling of time series and cross section data. *Econometrica* **46**: 69–85.

Nelsen R. 2006. *An Introduction to Copulas*. Springer: New York.

Oliveira B, Fortunato A. 2008. The dynamics of the growth of firms: evidence from the services sector. *Empirica* **35**: 293–312.

Orme C. 2001. Two-step inference in dynamic non-linear panel data models. *Unpublished paper, University of Manchester* .
URL `http://personalpages.manchester.ac.uk/staff/chris.orme/documents/Research\\\%20Papers/initcondlast.pdf`

Ornaghi C. 2009. Mergers and innovation in big pharma! *International Journal of Industrial Organization* **27**: 70–79.

Pavitt K, Robson M, Townsend J. 1987. The size distribution of innovating firms in the UK: 1945-1983. *The Journal of Industrial Economics* **35**: 297–316.

Quandt R. 1966. On the size distribution of firms. *The American Economic Review* **56**: 416–432.

Radice H. 1971. Control type, profitability and growth in large firms: an empirical study. *The Economic Journal* **81**: 547–562.

Ricci V. 2005. Fitting distributions with R. *R project website* .
URL `http://cran.r-project.org/doc/contrib/Ricci-distributions-en.pdf`

Rostow W. 1959. The stages of economic growth. *Economic History Review* : 1–16.

Samuels J. 1965. Size and the growth of firms. *The Review of Economic Studies* **32**: 105–112.

Samuels J, Chesher A. 1972. Company growth, survival and size. In Cowling K (ed.) *Market Structure and Corporate Behaviour*. Gray Mills: London.

Samuels J, Smyth D. 1968. Profits, variability of profits and firm size. *Economica* **35**: 127–139.

Saving T. 1965. The four-parameter lognormal, diseconomies of scale and the size distribution of manufacturing establishments. *International Economic Review* **6**: 105–114.

Scherer F. 1965. Firm size, market structure, opportunity, and the output of patented inventions. *The American Economic Review* **55**: 1097–1125.

Selten R. 1983. A model of oligopolistic size structure and profitability. *European Economic Review* **22**: 33–57.

Simon H. 1955. A behavioral model of rational choice. *The Quarterly Journal of Economics* **69**: 99–118.

Simon H, Bonini C. 1958. The size distribution of business firms. *The American Economic Review* **48**: 607–617.

Simon H, Ijiri Y. 1977. *Skew Distributions and the Sizes of Business Firms*. North Holland: Amsterdam.

Singh A, Whittington G, Burley H. 1968. *Growth, Profitability and Valuation: a Study of United Kingdom Quoted Companies*. Cambridge University Press: Cambridge.

Sklar A. 1973. Random variables: joint distribution functions and copulas. *Kybernetika* **9**: 449–460.

Smyth D, Boyes W, Peseau D. 1975. The measurement of firm size: theory and evidence for the United States and the United Kingdom. *The Review of Economics and Statistics* **57**: 111–114.

Steindl J. 1965. *Random Process and the Growth of Firms: a study of the Pareto Law*. Hafner: New York.

Steindl J. 1968. Size distributions in economics. In *International Encyclopedia of the Social Sciences*, volume 14. New York: MacMillan and Free Press, 294–300.

Stewart M. 2006. -redprob- A Stata program for the Heckman estimator of the random effects dynamic probit model. *mimeo, University of Warwick* .

Sutton J. 1997a. Game-theoretic models of market structure. In Kreps D, Wallis K (eds.) *Advances in Economics and Econometrics: Theory and Applications. Seventh World Congress*, 26. Econometric Society, Cambridge University Press: Cambridge, 66–86.

Sutton J. 1997b. Gibrat's legacy. *Journal of Economic Literature* **35**: 40–59.

Sutton J. 1998. *Technology and Market Structure: Theory and History*.

Variyam J, Kraybill D. 1994. Managerial inputs and the growth of rural small firms. *American Journal of Agricultural Economics* **76**: 568–575.

Wedervang F. 1965. *Development of a Population of Industrial Firms*. Universitetsforlaget: Blindern, Oslo.

Wooldridge J. 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Jouranal of Applied Econometrics* **20**: 39–54.

Yan J. 2007. Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software* **21**. URL `http://ideas.repec.org/a/jss/jstsof/21i04.html`

Yasuda T. 2005. Firm growth, size, age and behavior in Japanese manufacturing. *Small Business Economics* **24**: 1–15.