The Dynamics of Multinationals Expansion in the European Pharmaceutical Industry (DRAFT)

Franco Mariuzzo*
Department of Economics and Econometrics
University of Groningen, The Netherlands.

Xiaoheng Zhang†
Institute for International Integration Studies (IIIS)
Trinity College Dublin, Ireland.

June 11, 2008

Abstract

This paper avails of the multinational-level *Amadeus* dataset to study the dynamics of the European pharmaceutical industry in the Single Market Programme era: 1990-2004. Our unit of investigations are multinationals that every period decide whether to expand or not, by capturing at least one of the arisen opportunities. By using a dynamic panel probit approach with unobserved heterogeneity we show that the industry presents a positive persistent pattern over time. Comparisons of traditional and Bayesian econometrics provide robustness to this result. Besides that, we exploit a parametric copula technique to relate the history of multinationals binary decisions to their resulting size (measured either as number of subsidiaries acquired or as operational revenues). An important by-product of this procedure is that we can use both the primitives of our dynamic probit estimations and the simulations derived from the copula procedure to predict the probability each firm expands each subsequent period. We employ our predictions both to choose the best dynamic probit approach and to check how this fits into Sutton's lower bound measure of market concentration.

KEYWORDS: Bayesian and Classic Econometrics, Dynamic Entry, Multinational Growth, Pharmaceutical Industry, Single Market Programme.

JEL classifiers: C01, C11, C23, L10, L65.

---

*Corresponding author: +31-50-363-3789. Email: f.mariuzzo@rug.nl. Funding from Marie Curie Excellence Grant MEXT-CT-2006-042471 is gratefully aknowledged.

†A grant from DEPFA BANK PLC is gratefully aknowledged.

1 Introduction

Gibrat (1931)'s "*law of proportionate effect*" has motivated a growing literature in industrial economics to study and explain skewness in firms' size distributions.[1] This growth-of-firms literature (also known as stochastic literature on firms size) sees skewness as the resulting of a relation between firm's size and its rate of growth: if larger firms tend to grow faster (or slower) than their smaller rivals, the size distribution of the firms tend to exhibit more (or less) skewness. The popularity of this literature stands in its strong link with the structure-conduct- performance literature that started with Mason (1939, 1949) and Bain (1951, 1956). Sutton (1998) abandons this pure stochastic view and makes use of a game theoretical approach.[2] His main finding is that a lower bound in a Lorenz curve distribution (a minimum degree of firms' size inequalities) is the best one can do in explaining firms' size distribution.[3] Any distance of actual data from his lower bound would be explained, among other things, by the degree of heterogeneity in the arrival of opportunities.

Our paper studies growth-of-firms under a different perspective. It models firms choice in a discrete way. Every period firms optimally decide whether to expand their size by capturing one of the heterogeneous opportunities that have arisen. Opportunities are described as the opening up of new subsidiaries or acquisition of existing ones, and since every period firms can capture more than one subsidiary, a certain degree of heterogeneity in opportunities arise. Firms base their optimal decisions on underlying profits that depend, among other things, on demand shocks, scope economies, entry barriers, and the underlying level of competition.

We have a twofold goal. Firstly we want to test whether firms' decisions are persistent over time. To this aim we make use of dynamic panel probit models with unobserved heterogeneity, and to give robustness to our results we compare Classical and Bayesian approaches. Secondly, we relate firms' expansion to their underlying size. We do that indirectly, investigating the relation between

[1] Hart and Prais (1956), Adelman (1958), Simon and Bonini (1958), Hart (1962), Hymer and Pashigian (1962), Mansfield (1962), Steindl (1965, 1968), Ijri and Simon (1964, 1977), Quandt (1966), Hjalmarsson (1974), Lucas (1978), Hall (1987), Sutton (1997), Bottazzi et Al. (2001), Bottazzi and Secchi (2003, 2006), Buldyrev et al. (2007), Cefis et al. (2007).

[2] An alternative literature beholds firms' growth in a more structural way, by assuming heterogeneous firms as rational strategic maximizer agents [Selten (1983), Pakes and McGuire (1994), Ericson and Pakes (1995)].

[3] The lorez curve represents a convenient way to represent the size distribution of firms.

number of opportunities captured over time and size of these opportunities. This is a way to understand whether firms that have captured more opportunities grow faster and is also a way to answer the question: Do mnes that capture the first opportunities available in the market grow larger than those that enter later? In order to capture this idea, we use a parametric copula approach which allows us to derive the joint distribution (size, number of opportunities captured) from the empirical marginals of the two variates. Then we simulate firms' size by conditioning it to the number of opportunities captured up that period. An important by-product of this procedure is that we can use both the primitives of our dynamic probit estimations and the simulations to predict the probability each firms has to expand each subsequent period. One can therefore start predictions in period 1 and extend them up to period T. We use predictions as a way to indicate the best dynamic probit approach and to check the structure opportunities-firm size-market concentration.

The European pharmaceutical industry during the EU enlargement, 1990-2004, offers us a valid firm-level dataset. We obtain our data from a commercial database called *Amadeus* which has information on multinationals. Given the particular data, onwards a firm will be labeled a multinational, or mne. Our dataset has information on opportunities captured by each mne and on their respective size. We use two measures of size:

1. Number of subsidiaries acquired up to time t;
2. Operational revenues in period t.[4]

We treat the dataset as a balanced panel over the full period, assuming entrants are sleepers until they do not capture their first opportunity. It is important to mention that our dataset does not provide information on mnes exit, for exitors are not reported in the data. This means exit dynamics cannot be studied.

The rest of the paper is organized as follows: in Section 2 we present a simple conceptual model; in Section 3 we outline the dynamic panel probit approach with unobserved heterogeneity and compare Classical and Bayesian econometric

[4]Due to lack of data, when size is measured in terms of operational revenues the period is confined to 1995-2004.

techniques. Section 4 describes our data. Our results are discussed in Section 5. Section 6 concludes.

2 The Model

Multinationals expansion can be modelled in a simple way as the discrete choice to expand or not, based on an underlying latent profit relation,

$$y^*_{it} = \psi_{i,t-1} + \varepsilon_{it}, \tag{1}$$

where the latent profit of multinational i in period t, y^*_{it}, is explained by a idiosyncratic part, ε_{it}, and a function of pre-determined variables, $\psi_{i,t-1}$. Multinational i decide to expand in period t, $y_{it} = 1$, if $y^*_{it} \geq 0$; and not to expand, $y_{it} = 0$, otherwise. The actual expansion process occurs by the opening up of new subsidiaries or acquisition of existing ones. The number of mne with non-negative profits in period t matches exactly the number of opportunities that have arisen in that period. We know that opportunities can be heterogeneous, thus we can expect their difference in size to some extent reflect mnes differences in profits.

With our interest directed towards the dynamics of expansion, we model $\psi_{i,t-1}$ other than inclusive of mne specific effects, as a function of lagged expansion, $y_{i,t-1}$, observable multinational-specific characteristics such as a vector of time-invariant characteristics, O_i and a lagged measure of size $S^h_{i,t-1}$ (where $h = n$ denotes number of subsidiaries, $h = s$ denotes operational revenues).[5] The resulting relation can be expressed as,

$$\psi_{it} = f(\alpha_i, y_{i,t-1}, S^h_{i,t-1}, O_i), \tag{2}$$

Lagged expansion is aimed at capturing any form of persistence in the expansion process. This should help our understanding of the European pharmaceutical industry dynamics and the resulting industry concentration. That is to say, the evidence of a positive state dependency of expansion at time t on expansion at time $t-1$ indicates the sustainability of this industry, for example, this industry is able to generate profits to support its expensive innovative activities through

[5] O_i contains the nationality of MNEs, for example if they are EU or US based, as well as a variant indicating if these MNEs are leading pharmaceutical manufacturers in the world. O_i's presence in (2) is merely functional, in order to meet the technical requirements of instrumental variables and comparability of the econometric models.

a persistent expansion behavior. In this way, past expansions exert a behavior effect to current expansion and this effect is termed as "*true state dependency*" in Heckman (1981). However, as is often the case, the examination of dynamics can be blurred by what Heckman (1981) has termed "*spurious state dependency*". This occurs if unobservable multinational-level effects, α_i is serially correlated over time but hasn't been properly controlled for by researchers. In this case, the past expansion decision will wrongly capture this unobserved effect and behave as if it is a driving force of the current expansion decision, even if there is no state dependence at all. To tackle this problem we adopt a dynamic panel random effects econometric model. How we model random effects in a dynamic model using the Classic and Bayesian econometrics approaches is going to be the core of the next section.

Our choice of lagged number of subsidiaries, $S^n_{i,t-1}$, and operational revenues as measures of size, $S^s_{i,t-1}$, is motivated by Sutton (1998)'s theory, where he sees new investment opportunities as many equal-sized "*islands*". Each active multinational in an industry has a certain probability of capturing that opportunity, and subsequently to expand. Size of each multinational is related to the number of opportunity "*islands*" it has taken before. Sutton's concept implies that the number of opportunities captured and size of a multinational have a direct influence on each multinational's propensity to expand. Visually, as a multinational becomes larger it has more chances to catch a new coming opportunity. The direct effect of number and size on expansion is also corresponding to an important empirical tradition in the IO literature, Gibrat's Law (see Sutton, 1997), which states that "the multinational's proportionate rate of growth is given by a random variable whose mean is independent of the multinational's current size". Gibrat's Law implies that the larger the multinational is, it is more likely to catch a new opportunity. However, consensus does not exist among economists with regard to the validity of Gibrat's Law (see Thompson and Klepper, 2003, Evans, 1987 and Dunne et al., 1988).[6] Bottazzi et al. (2001) and Bottazzi and Secchi (2006) provided evidences against the validity of Gibrat's law in the

[6] Theoretical work done in Thompson and Klepper (2003) shows that "mean multinational growth is strictly decreasing in multinational size conditional on multinational age (Proposition 7)." Empirically, Evans (1987) using employment to proxy multinational size and Dunne et al.(1988) using output as proxy provided evidence that for American large manufacturing multinationals, the proportional rate of growth of a multinational (or plant) conditional on survival is decreasing in size.

pharmaceutical industry.[7]

Besides the direct effects that the lagged observables, $S^h_{i,t-1}$, have on the probability of expansion, they can also influence expansion in an indirect way through scope economies, economies of scale in terms of production and R&D, or superior market power as gained as multinationals grow up. If indirect effects of scope economies, economies of scale and market power are growing faster, they probably impose a "*behavior effect*" on future expansion, termed as such by Heckman (1981), which can be effectively captured by the lagged expansion variable, $y_{i,t-1}$. Therefore, controlling for $S^h_{i,t-1}$ in the latent profit function is necessary if one wants to model the state dependence explicitly.

The next section outlines the econometric methodologies we used.

3 Econometric Methods

3.1 A Dynamic Probit Classical Approach

We employ a dynamic extension of probit model to address state dependence of multinationals' expansion decisions. Following Wooldridge (2002), we use a generic notation to model the probability multinational i expands in the European pharmaceutical market (via opening up/acquiring a subsidiary) at year t,

$$\begin{aligned} P\left(y_{it}=1|\alpha_i, y_{i,t-1}, \mathbf{x}_{i,t-1}, \boldsymbol{\lambda}\right) &= \Phi\left(\alpha_i+\rho y_{i,t-1}+\mathbf{x}'_{i,t-1}\boldsymbol{\beta}\right) \\ \boldsymbol{\lambda} &\equiv \left(\rho, \boldsymbol{\beta}'\right)'; \alpha_i \sim N\left(0, \sigma^2_\alpha\right). \end{aligned} \tag{3}$$

The above relation states that: The probability of expansion at time t is a function of the lagged dependent variable, $y_{i,t-1}$, a vector of all lagged and multinational-specific independent variables, $\mathbf{x}'_{i,t-1}$, and a time-invariant multinational-specific effect, observables, O_i, and unobservables (to us), α_i. This later is aimed at capturing all unobservable factors that could influence multinationals' expansion propensity.

[7] In Bottazzi and Secchi(2006), they find empirically for 198 world top multinationals in the pharmaceutical industry from 1987 to 1997, the variance of the growth rate to be negatively associated with size, is contrast to Gibrat's Law. But they further find that this dependence can be explained by the diversification of the MNEs in sub-markets: "bigger multinationals operate in more sub-markets and the variance of their growth is consequently reduced." Bottazzi et al. (2001) used detailed sales data for 150 top pharmaceutical companies in seven major Western markets(US, UK, France, Germany, Spain, Italy and Canada) for the long time span from 1987 to 1997 to test Gibrat's Law. In the aggregated level they find the statistical evidence to challenge Gibrat's Law because growth rates depend on size.

The discrete choice of mnes' expansion modeled in (3) can be expressed linearly to the latent profits, as seen in (1). Anytime a multinational realizes a positive y_{it}^{*}, it will have generated enough resources to capture one of the arisen opportunities, $y_{it} = 1$. Furthermore, although not explicitly modeled, we account for the fact that higher positive profits y_{it}^{*} give mnes the ability to capture larger opportunities. Hence the model of expansion can be written as,

$$y_{it}^{*} = \alpha_i + \rho y_{i,t-1} + \mathbf{x}_{i,t-1}^{'}\boldsymbol{\beta} + u_{it}, \tag{4}$$

where u_{it} is an *iid* term distributed as a standard normal.

However, this model faces some difficulties to hold, as pointed out in Wooldridge (2002). Assumption of strict exogeneity between all regressors in the right hand side (RHS) of Equation (4) and the unit-specific effect α_i is required to insure the consistent estimation of the parameters $\boldsymbol{\lambda}$ as in a traditional static version of the probit model. But this assumption rules out the possibility of introducing dynamic mechanism in the RHS of (3) or (4) because lagged y_{it} are inevitably correlated with time-invariant unit-specific effect α_i and inclusion of the lagged expansion variable raises an issue called "initial condition problem". Besides this difficulty, to consistently estimate the partial effects for independent variables, the α_i has to be independent of $\mathbf{x}_{i,t-1}$ and $y_{i,t-1}$, but this assumption is often violated. For example in our case, a multinational's reputation is influenced by its expansion decisions at current time and past years and consequently size is correlated with the unit-specific effect α_i.

In the next subsubsections we outline two alternative methods to handle the initial condition problem.

3.1.1 Wooldridge's Method

Wooldridge (2002) proposes to condition the random effect, α_i, explicitly on the initial value of the dependent variable, y_{i0}, and to the almost full history of the independent variables, $\mathbf{x}_i^{T-1} = \left(\mathbf{x}_{i1}^{'}, \mathbf{x}_{i2}^{'}, ..., \mathbf{x}_{i,T-1}^{'}\right)^{'}$ as follows,

$$\begin{aligned} h\left(\alpha_i | y_{i0}, \mathbf{x}_i^{T-1}, \boldsymbol{\gamma}\right) &= \gamma_{[0]} + \gamma_{[1]} y_{i0} + \left(\mathbf{x}_i^{T-1}\right)^{'} \boldsymbol{\gamma}_{[2]}, \\ \boldsymbol{\gamma}_2 &\equiv \left(\gamma_{[2]11}, ..., \gamma_{[2]k1}, ..., \gamma_{[2]1,T-1}, ..., \gamma_{[2]k,T-1}\right)^{'}. \end{aligned} \tag{5}$$

However, for each multinational i in our balanced panel, $\mathbf{x}_i^{T-1}$ is a $k\left(T-1\right) \times 1$ vector of independent variables and this configuration causes inefficiency in estimation. Mundlak (1978) suggested replacing the full time series of variables with their time-mean to reduce the number of excessive regressors. So that,

$$\overline{\mathbf{x}}_{i[-1]} \equiv \frac{1}{T-1}\sum_{t=2}^{T}\mathbf{x}_{i,t-1}.$$

Such that (5) can be rewritten as,

$$h\left(\alpha_i|y_{i0},\mathbf{x}_i^{T-1},\boldsymbol{\gamma}\right)=\gamma_{[0]}+\gamma_{[1]}y_{i0}+\left(\overline{\mathbf{x}}_{i[-1]}\right)^{\prime}\overline{\boldsymbol{\gamma}}_{[2]},\quad \overline{\boldsymbol{\gamma}}_{[2]}=\left(\overline{\gamma}_{[2]1},...,\overline{\gamma}_{[2]k}\right)^{\prime}.$$

Once α_i is conditioned, the task of finding a joint density function of y_{it}, for $t=1,2,...,T$, is transformed into the task of finding joint density function of y_{it}, for $t=2,...,T$, conditional on the initial value y_{i0} and the regressors $\overline{\mathbf{x}}_{i[-1]}$. The conditional joint distribution of $(y_{i2},y_{i3},...,y_{iT})$ can be expressed as an integral with the random effect being integrated out,

$$f\left(y_{i2},y_{i3},...,y_{iT}|y_{i0},y_{i,t-1},\mathbf{x}_i^{T-1},\boldsymbol{\gamma},\boldsymbol{\lambda}\right)=\int_{-\infty}^{\infty}\left[\begin{array}{c} f\left(y_{i2},y_{i3},...,y_{iT}|\alpha_i,y_{i,t-1},\mathbf{x}_{i,t-1},\boldsymbol{\lambda}\right) \\ h\left(\alpha_i|y_{i0},\mathbf{x}_i^{T-1},\boldsymbol{\gamma}\right)d\alpha_i \end{array}\right] \tag{6}$$

The advantages of the above approach are: (i) It is easy to find the conditional density of α_i given y_{i0} and $\mathbf{x}_i^{T-1}$, and it is also computationally easy to estimate maximum likelihood estimators using regular statistical software; (ii) The average partial effect for every regressors can be computed; (iii) It is easily extended to the Bayesian case as we will see later.[8]

3.1.2 Heckman's Method

Heckman (1981) proposes an alternative approach to handle the initial condition problem. He suggests to find the joint distribution for time series of binary outcomes $(y_{i2},y_{i3},...,y_{iT})$ by specifying two conditional distributions $\Phi_1\left(y_{i0}|\mathbf{x}_i,\alpha_i,\boldsymbol{\eta}\right)$ for y_{i0} and $f_2\left(\alpha_i|\mathbf{x}_i,\boldsymbol{\theta}\right)$ for α_i. Substituting into (6) the two conditional distributions and integrating the product with respect to α_i, we recover the joint distribution of $(y_{i2},y_{i3},...,y_{iT})$ conditional only on $\mathbf{x}_{i,t-1}$,

$$f\left(y_{i2},y_{i3},...,y_{iT}|\mathbf{x}_{i,t-1},\boldsymbol{\lambda}\right)=\int_{-\infty}^{\infty}\left\{\begin{array}{c}\Phi_1\left(\cdot\right)^{y_{i1}}\left[1-\Phi_1\left(\cdot\right)\right]^{(1-y_{i1})} \\ \prod_{t=2}^{T}\Phi\left(\cdot\right)^{y_{it}}\left[1-\Phi\left(\cdot\right)\right]^{(1-y_{it})}\end{array}\right\}f_2\left(\alpha_i|\mathbf{x}_i,\boldsymbol{\theta}\right)d\alpha_i. \tag{7}$$

[8] For example, -xtprobit- in Stata.

Likelihood for the units across whole time span to be maximized is

$$L = \prod_{i=1}^{N} f\left(y_{i2}, y_{i3}, ..., y_{iT} | \mathbf{x}_{i,t-1}, \boldsymbol{\lambda}\right).$$

To apply this approach we use a Stata program “redprob”, which is provided in Stewart (2006) and described in Stewart (2005). In Stewart’s program, conditional distribution $\Phi_1(\cdot)$ is approximated by a reduced-form linear function of initial latent profit

$$y_{i0}^{*} = \acute{\mathbf{z}}_{i0}\boldsymbol{\eta}_1 + \eta_2\alpha_i,$$

where $\mathbf{z}_{i0}$ is a vector containing exogenous instruments and $\mathbf{x}_{i0}$. The parameter η_2 measures correlation between the initial condition and the unit-specific effect. This program evaluates the integral (7) by Gaussian-Hermite quadrature assuming that $\alpha_i \sim N\left(0, \sigma_\alpha^2\right)$.

3.2 A Dynamic Probit Bayesian Approach

Amisano and Giorgetti (2005) use a Bayesian approach to study expansion into submarkets and diversification in the pharmaceutical industry across seven developed countries. In this section we propose our modification to their methodology.

The Bayesian approach has not become popular in the industrial-organization research, but its flexibility and computational easiness encourage us to use it. Particularly, the random-effect models can easily be accommodated with the Bayesian approach; not having to deal with solving prohibitive high-dimension integrals involved, as in traditional econometric models. Unlike traditional econometrics, the Bayesian approach does not require strong parametric assumptions over the sampling or strong asymptotic theories (Lancaster, 2004). In addition the Bayesian methodology uses either previously gained or available information on the parameters of model as priors in the estimation, saving therefore time in the estimation. This is drastically contrasted to the traditional econometrics, which requires researchers to be blind to the data and outcomes.

Bayesian modelling and computation of the discrete responses model have been studied in various papers (see Chib, 1992, Albert and Chib, 1993 and Chib and Greenberg, 1996). In the probit model in (3), what we are interested are the parameters $\boldsymbol{\lambda} \equiv \left(\rho, \boldsymbol{\beta}'\right)'$, but we have as well to specify the unit-specific α_i to accommodate for random effects. Therefore, expressed in the Bayesian way, the joint posterior distribution of the parameters of interests (and those of no interest) conditional on the data $\mathbf{z}_{it} = \left(y_{it}, y_{i,t-1}, \mathbf{x}_{i,t-1}^{\prime}\right)$ (a general expression for both dependent and independent variables) is,

$$p\left(\boldsymbol{\lambda},\alpha_i|\mathbf{z}_{it}\right) \propto p\left(\mathbf{z}_{it}|\boldsymbol{\lambda},\alpha_i\right) p\left(\boldsymbol{\lambda},\alpha_i\right), \tag{8}$$

where $p\left(\mathbf{z}_{it}|\boldsymbol{\lambda},\alpha_i\right)$ is the likelihood of $\mathbf{z}_{it}$ and $p\left(\boldsymbol{\lambda},\alpha_i\right)$ is the prior distribution of random parameters. We also assume that α_i is independent of $\boldsymbol{\lambda}$. This equation reads like "the conditional joint posterior distribution of $\boldsymbol{\lambda}$ and α_i is proportional to, $\propto$, the product of likelihood of data $\mathbf{z}_{it}$ conditional on $\boldsymbol{\lambda}$ and α_i and their prior". Intuitively, if the posterior joint distribution of $\boldsymbol{\lambda}$ and α_i conditional on data $\mathbf{z}_{it}$ is known, the task of estimating parameters of interests is nothing more than simulating the joint posterior distribution of $\boldsymbol{\lambda}$ and α_i. However, as pointed out in Chib and Greenberg (1996), $p\left(\mathbf{z}_{it}|\boldsymbol{\lambda},\alpha_i\right)$ is intractable because its discrete nature does not allow normalization and hence $p\left(\boldsymbol{\lambda},\alpha_i|\mathbf{z}_{it}\right)$ cannot be easily simulated. Conditioning the latent variable y^*, as well as $\boldsymbol{\lambda}$ and α_i on data $\mathbf{z}$, Expression (8) becomes,

$$\begin{aligned} p\left(\boldsymbol{\lambda},\alpha_i|\mathbf{z}_{it}\right) &\propto p\left(\mathbf{z}_{it}|\boldsymbol{\lambda},\alpha_i\right) p\left(\boldsymbol{\lambda},\alpha_i\right) \\ &= p\left(z_{it}|y_{it}^*, \lambda, \alpha_i\right) p\left(y_{it}^*|\lambda, \alpha_i\right) p\left(\lambda, \alpha_i\right). \end{aligned}$$

Although (posterior) distribution $p\left(y_{it}^*, \lambda, \alpha_i|z_{it}\right)$ is complicated, the conditional distributions $p\left(y_{it}^*|\lambda, \alpha_i, z_{it}\right)$ and $p\left(\lambda, \alpha_i|y_{it}^*, z_{it}\right)$ are easy to simulate (Albert and Chib, 1993). So that,

$$\begin{aligned} p\left(y_{it}^*|\lambda, \alpha_i, z_{it}\right) &= p\left(y_{it}^*|\rho, \boldsymbol{\beta}, \alpha_i, y_{it}, y_{i,t-1}, \mathbf{x}_{i,t-1}\right) \propto \\ &\propto \phi\left(y_{it}^*\right) \prod_{n=1}^{N} \prod_{t=2}^{T} \left[\begin{array}{c} I\left(y_{it}=1\right) I\left(\rho y_{i,t-1} + \mathbf{x}_{i,t-1}^{\prime}\boldsymbol{\beta} + \alpha_i > 0\right) + \\ +I\left(y_{it}=0\right) I\left(\rho y_{i,t-1} + \mathbf{x}_{i,t-1}^{\prime}\boldsymbol{\beta} + \alpha_i \leq 0\right) \end{array} \right], \end{aligned}$$

and,

$$\begin{aligned} p\left(\lambda, \alpha_i|y_{it}^*, z_{it}\right) &= p\left(\rho, \boldsymbol{\beta}, \alpha_i|y_{it}^*, y_{it}, y_{i,t-1}, \mathbf{x}_{i,t-1}\right) \propto \\ &\propto \phi\left(\rho\right) \phi\left(\boldsymbol{\beta}\right) \phi\left(\alpha_i\right) \prod_{n=1}^{N} \prod_{t=2}^{T} \left\{ \begin{array}{c} \Phi\left(\rho y_{i,t-1} + \mathbf{x}_{i,t-1}^{\prime}\boldsymbol{\beta} + \alpha_i\right)^{y_{it}} \\ \left[1 - \Phi\left(\rho y_{i,t-1} + \mathbf{x}_{i,t-1}^{\prime}\boldsymbol{\beta} + \alpha_i\right)\right]^{(1-y_{it})} \end{array} \right\}, \end{aligned}$$

where $I\left(\cdot\right)$ is an indicator function for $y_{it} = 1$ and $y_{it} = 0$ and $\phi\left(y_{it}^*\right)$, $\phi\left(\rho\right)$, $\phi\left(\boldsymbol{\beta}\right)$ and $\phi\left(\alpha_i\right)$ are all prior density functions which reflect the subjective beliefs of the researcher.

The Gibbs sampling for the $\boldsymbol{\lambda} \equiv \left(\rho, \boldsymbol{\beta}'\right)'$ and α_i parameters under this specification is based on the following steps,

1. Choose the initial values of ρ, $\boldsymbol{\beta}$ and α_i;

2. Sample y_{it}^* for $i = 1, 2, ..., N$ from a normal distribution $(y^*|\rho, \boldsymbol{\beta}, \alpha_i, y_{it}, y_{i,t-1}, \mathbf{x}_{i,t-1})$ that is truncated on the left if $y_{it}^* > 0$, $y_{it} = 1$, and on the right if $y_{it}^* \leq 0$, $y_{it} = 0$;

3. Sample ρ, $\boldsymbol{\beta}$, α_i, in the normal distribution $(\rho, \boldsymbol{\beta}, \alpha_i|y_{it}^*, y_{it}, y_{i,t-1}, \mathbf{x}_{i,t-1})$, where the means $\rho, \boldsymbol{\beta}$, are equalized to $(\mathbf{X}\mathbf{X})^{-1}\,\mathbf{X}\acute{\mathbf{y}}^*$ and variance $(\mathbf{X}\mathbf{X})^{-1}$, where $\mathbf{X} \equiv (\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_N)$ and $\mathbf{X}_i$ is a $(k \times T)$ matrix of all explanatory variables $\left(y_{i,t-1}, \acute{\mathbf{x}}_{i,t-1}\right)^{\prime}$.

4. Repeat step 1 to 3 ns times. The number ns should be large enough to allow the posterior distribution of the parameters of interests converge to a stationary state.

We use the Bayesian statistic software "WinBUGS" to implement our Bayesian analysis. We apply the Bayesian analysis to the dynamic Probit model based on Wooldridge's approach. In this way we make our Bayesian results comparable to those presented in the first of the Classical approaches presented.

4 Data and Variables

We extract information and data of pharmaceutical Multinationals and their subsidiaries from a commercial database called "*Amadeus*". This database is so far the most comprehensive one collecting accounts data and other relevant information of mnes located in Europe. A brief description of the "*Amadeus*" database and details about how we identify mnes and their subsidiaries is set in Appendix 1.

We look for pharmaceutical multinationals whose pharmaceutical-related production subsidiaries, research subsidiaries or marketing subsidiaries were established or acquired in the EU-15 countries in the time span 1990-2004. We only consider those multinational parents who are either incumbent in this industry before 1990 or new entrants, coming into it after 1990. These inclusion criteria strictly comply with Sutton's theory of submarket and firm's size distribution, which only applies to incumbents and new entrants, but irrelevant to companies who exit. One additional issue of identifying mne parents is mergers, which took place frequently in the period of interest in this industry.[9] Mergers complicate the identification of an mne's expansion history because companies who merged together later will appear as separate businesses in the earlier year and this make it is impossible to construct a balanced panel of expansion history. To get around this difficulty, we adopt Bottazzi and Secchi (2006)'s procedure of treating merged companies as a single entity as if it has come into being at the beginning of period under investigation. This "super company" procedure will bias the inter-temporal size distribution of firms, but will stress the effects of mne's size on its expansion probability. Finally, we are left with a total of 265 pharmaceutical multinationals, among them 31 mnes are deemed as new entrants. This sample represents the vast majority of the Multinational players in the European pharmaceutical industry who were active before and after the implementation of the Single Market Programme.[10] The Single Market Programme is treated as a single geographic market. Our sample can also be regarded as a good approximation of the major European pharmaceutical producers because during the collection of information for the European companies it is hard to find any major pharmaceutical producer that did not expand its business beyond national borders.

Of these 265 pharmaceutical multinationals, we collected information on their subsidiaries, amounting to 1757, existing and active until 2004. Detailed information are provided in tables A1.1 and A1.2 in the Appendix. Specifically, in the time period 1990-2004 we observe 827 new subsidiaries established or acquired. These subsidiaries are in one or more of three categories of business activities: i) production; ii) wholesaler or dispensing and iii) research and development.[11] The sources of these subsidiaries are either green-field investment by multinational parent themselves or acquisition through mergers or takeover.

[9] Two eminent examples of mergers are British company AstraZeneca and Swiss company Novartis. British company Zeneca merged with Swedish company Astra in 1999 to create AstraZeneca (UK) and two Swiss companies Ciba-Geigy and Sandoz Laboratories merged in 1996 to create Novartis (Switzerland).

[10] According to "The Top 100 Companies" in KPMG (2005/2006), 78 out of 100 world largest pharmaceutical companies are in the sample, 22 companies do not have subsidiaries in the EU (according to the "*Amadeus*" data).

[11] In the Amadeus, business activities are classified by NACE Rev. 1.1 Industry Code. We are primarily looking for the subsidiaries with code 2441(manufacture of basic pharmaceutical products), 2442 (manufacture of pharmaceutical preparations), 5146 (wholesale of pharmaceutical goods), 5231 (dispensing chemists), 5232 (retail sale of medical and orthopaedic goods) and 7310 (research and experimental development on natural sciences and engineering). We also check some subsidiaries that do not have above codes if their businesses are actually dealing with pharmaceutical products using information on pharmaceutical MNEs' official websites.

Multinational's expansion in each year is defined as a binary variable equal to 1 if a multinational has at least one pharmaceutical subsidiary established or acquired in that year and zero, otherwise. Entries of new-entrant mnes are treated as zero in the year before these new entrants came into the market and for the years after they came into the market, entries are counted in usual way that we deal with incumbent mnes. This treatment virtually assumes new mnes were already in the market but they were not active, and it helps to obtain a balanced panel. Eventually, of the 265 multinationals over the 15-year period, we count 564 expansion observations.

Our idea of treating acquired and new own-established subsidiaries indifferently is justified by the fact that either green-field investment or acquisition reflects each multinational's willingness and movements of expansion. They are all influenced by industry-level trends such as relocation of business activities or adjustments to the Single Market, and consequently they are components of the industrial dynamics. Figure 3 shows us the cumulative density functions of number of subsidiaries and size over the period 1996-2003. As the figure shows, almost eighty percent of the distribution of number of subsidiaries is made of very small mnes (with only one or two subsidiaries). We expect number of subsidiaries and size to be related and our conjecture is confirmed by Figures 4 and 5 where it appears that for the right side of the distribution there is a strong positive relation between the two random variables.

The explanatory variables we use in our estimations are described below:

1. $y_{i,t-1}$ is one-year lagged dependent variable, which equals 1 if some subsidiaries were established or acquired in year $t-1$.

2. $S^n_{i,t-1}$ the total number of subsidiaries multinational i has established or acquired, regardless of their activities (production, wholesales or research), in the previous period; and its quadratic term, $S^{n2}_{i,t-1}$. Although we argue that acquiring a subsidiary may be motivated by different considerations other than those motivating green-field investments, we consider the effect of additional acquired subsidiary on future expansion should be the same as an additional own-established subsidiary.

   The above-discussed positive effect of number of subsidiaries $S^n_{i,t-1}$ may be counteracted by external and internal factors. Externally, the availability of new business opportunities may be scarce, because the European market for the pharmaceutical products is well-established for many decades. Some of famous players in this industry have already been in Europe for more than a hundred years. Therefore, a multinational having many subsidiaries may merely indicate well-exploited opportunities. Internally, as the number of subsidiaries increases, managerial costs or transaction costs in a multinational also rise, which will raise the long-run average cost of the multinational. The diseconomies of scale are discussed in Williamson (1975) for the transaction costs due to the bureaucratic failure, atmospheric consequences, information distortion and incentive limits, and are tested in Canback (2004). These possible negative connections motivate us to model the $S^n_{i,t-1}$ variable in a non-linear relationship with the probability of expansion. Is there an optimum number of subsidiaries a multinational, allowing for its characteristics, should have? This translates into another question, isn't that beyond that optimal number of subsidiary the probability a multinational will open a new subsidiary will decrease? Our way to answer these questions is by including a quadratic term of $S^n_{i,t-1}$ and studying its relationship with the probability of expansion.

3. $S^s_{i,t-1}$ is multinational i's one-period lagged size measured by operational revenues.[12] This is calculated by summing up the operational revenues of one mne's pharmaceutical-related subsidiaries for each year. Unfortunately Amadeus failed to report operational revenues for some subsidiaries in earlier years, and we report in Table 1 below how many mnes in the sample have missing values of operational revenues in each year. As a safe measure we consider that reliable aggregated size variable only starts from 1995 and hence we use data from 1995 in our estimations.

   Table1. Number of Missing Values (out of 265 mnes) in Size in each year.[13]

[12] Following a usual practice in the Bayesian analysis, $S^s_{i,t-1}$ is normalized to follow a $N(0,1)$ distribution to guarantee fast convergence to the stable posterior distribution in the Gibbs Sampling.

[13] Note: since size is lagged by one year, size in 2004 is not used for estimation of entry from 1990 to 2004.

Year	1990	1991	1992	1993	1994	1995	1996
Number of Missing	234	235	239	173	103	44	42
1997	1998	1999	2000	2001	2002	2003	2004
34	33	22	17	13	10	7	note

4. Besides these explanatory variables, we use three dummy variables (the instruments in Heckman's approach). They are TOP_i which is equal to one for those multinationals among the top 100 pharmaceutical companies in the world rank in 2004 (KPMG, 2005/2006) and EU_i or US_i which are equal one for those multinationals having headquarters located either in the EU or US. These instruments are also included in Wooldridge's approach (both in traditional economics or the Bayesian analysis) as explanatory variables to make results from different models comparable.

Along with the above variables we use initial condition of the explanatory variables, $y_{i,90}$, $S^h_{i,90}$, $S^{h2}_{i,90}$, in addition to the time mean variables, $\overline{S^h_i}$, and $\overline{S^{h2}_i}$. We report in Appendix 1 some descriptive statistics of expansion, changes in size and the number of subsidiaries and explanatory variables and two empirical densities of the number of subsidiaries. The next section presents our main results.

5 Results

5.1 Size proxied by Number of Subsidiaries

Tables 2 and 3 report our estimates in a reduced form (excluding variables related to operational revenues) for a 15-year sample. Estimates of a Pooled Probit, a Heckman approach, a Wooldridge approach and a Bayesian approach are reported side by side. Comparing the pooled probit with the other random effects approaches we need to remember that the pooled probit model treats observations as belonging to simple cross-section, in this way the temporal correlation between unobservable unit-specific effects is ignored and the terms α_i and true error term μ_{it} are treated as a composite error term,

$$c_{it} = \alpha_i + \mu_{it}.$$

The implication is that the pooled probit uses as normalization of the error term $\sigma_c^2 = 1$, whereas the other random effects approaches adopt $\sigma_\mu^2 = 1$. Therefore, the random effects coefficients need to be multiplied by a factor $\sqrt{1-\iota} = \sigma_\mu / \sigma_c$ to make them comparable to the pooled probit model ones.[14]and this preparation will reduce the estimates of random-effect models even further.

In terms of the lagged dependent variable $y_{i,t-1}$, the pooled probit model significantly over-estimates the level of state dependence when compared to the other random effect models. The Bayesian approach is the one that offers the most conservative but significant estimate of lagged expansion. This is evidence that the Bayesian approach outperforms classic econometric in terms of controlling for random effect. This is because in the Bayesian analysis what you get are the exact estimates from the simulation in the posterior distributions of the parameters of interests whereas in the classic econometrics, the maximum-likelihood estimates (MLEs) are calculated by mathematical approximation Lancaster (2004).

To evaluate the effect of state dependence, we calculate the average partial effect (APE) by constructing counter-factual outcome probabilities and calculating those in the states $y_{i,t-1} = 0$ or $y_{i,t-1} = 1$; with associated explanatory variables $\mathbf{x}_{i,t-1} = \left(S_{i,t-1}^n\right)$ or $\mathbf{x}_{i,t-1} = \left(S_{i,t-1}^n, S_{i,t-1}^{n2}\right)$ and their sample mean values $\overline{\mathbf{x}}_{(-1)i,\cdot}$ (Wooldridge, 2006 and Stewart, 2006). In this way we compute the following probabilities,

$$\begin{aligned} \widehat{p}_{(1)} &= \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=2}^{T} \Phi \left[\left(\widehat{\rho} + \mathbf{x}'_{i,t-1} \widehat{\boldsymbol{\beta}} + \overline{\mathbf{x}}'_{i[-1]} \widehat{\overline{\boldsymbol{\gamma}}}_{[2]} \right) \sqrt{1-\widehat{\iota}} \right] \\ \widehat{p}_{(0)} &= \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=2}^{T} \Phi \left[\left(\mathbf{x}'_{i,t-1} \widehat{\boldsymbol{\beta}} + \overline{\mathbf{x}}'_{i[-1]} \widehat{\overline{\boldsymbol{\gamma}}}_{[2]} \right) \sqrt{1-\widehat{\iota}} \right]. \end{aligned}$$

APEs, $\widehat{p}_{(1)} - \widehat{p}_{(0)}$, are reported at the bottom of Table 2 and Table 3 and they are all at very similar level, but the APE in the pooled probit model are larger than those in the other three approaches.

Turn to the estimates of the lagged number of subsidiaries $S_{i,t-1}^n$ and $S_{i,t-1}^{n2}$, the four models all estimate a statistically significant negative effect of $S_{i,t-1}^n$(Tables 2-3). In Table 3, we find that adding quadratic form $S_{i,t-1}^{n2}$ leads to an increase in the effect of lagged expansion for all four models. The effect of $S_{i,t-1}^n$ also

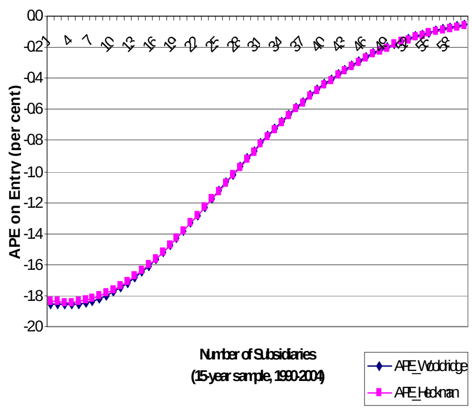
[14] $\iota = Corr\left(u_{it}, u_{is}\right) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\mu^2}$, for $t, s = 2, 3, ..., T; s \neq t$, which is the equicorrelation between the error in two time periods (Stewart, 2006).

changes drastically (having a stronger negative impact on expansion). The two coefficients of $S_{i,t-1}^{n2}$ and $\overline{S_i^{n2}}$ are both statistically significant but have only weak effects on expansion.

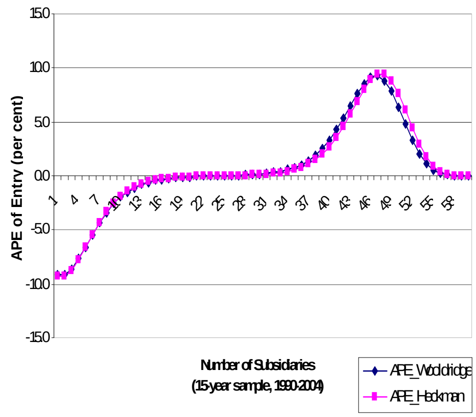
We calculate the APEs on expansion that derived from a change in the numbers of subsidiaries for models without and with $S_{i,t-1}^{n2}$. To include the time-mean number of subsidiaries in an straightforward way, we assume a hypothetical multinational that enters the European market at year $t = 1$ with only one subsidiary and expands at the speed of one additional subsidiary each year, such that the time-mean number of subsidiaries at year $t = 1$ is $\frac{T+1}{2}$ and the time-mean of quadratic term of number for this mne equals $\frac{T(2T+1)}{6}$. Figure 1 shows the APEs of $S_{i,t-1}^{n}$ for models using Wooldridge and Heckman's approach. $S_{i,t-1}^{n}$ is seen to have a negative effect on the probability of expansion; however this effect gets weaker and closer to zero as the number of subsidiaries increases up to 60 (notice that 54 is the maximum number of pharmaceutical subsidiaries observed in our sample). This finding seems to be in contrast with Sutton's "opportunity island" idea. Figure 2 shows the APEs of the number of subsidiaries when non-linear relation is captured by including a quadratic term of $S_{i,t-1}$. At the early stage of expansion, one additional subsidiary has a negative effect on the probability of expansion but this effect turns out to diminish as the number of subsidiaries increases. After a threshold of 22 the APE becomes positive. When the number approaches 40, the APE increases at faster pace and it reaches the highest level at 48, then declines. This finding enriches our understanding of the effect of the number of subsidiaries because in the model without quadratic term, we can only find a constantly negative effect of it.

The variant effect of APE of subsidiaries in models with quadratic terms can be understood by a dynamic trade-off between Sutton's "opportunity island" theory and negative effects such as managerial costs and exploited opportunities.

Starting from multinationals with very few subsidiaries, the priority of small multinationals is surviving instead of opening up new subsidiaries, which is an expensive regularity for new entrants due to managerial costs and the risk in the market. As a consequence, negative APE when the number is small merely indicates unfavorable situation small entrants are in. After the first stage of struggling, surviving multinationals gradually overcome their size disadvantages to those incumbents and begin to expand; hence the negative effect of number



(a) Figure 1: Average Partial Effect of number of subsidiaries (model without squared terms)



(b) Figure 2: Average Partial Effect of number of subsidiaries (model inclusive of squared terms)

becomes weaker as small multinationals grow up. Empirically we find that, in the mid-stage of expansion, APE virtually becomes very weak and stable for a long term, say, approximately in the interval of $number = (22, 35)$ for a 15-year sample, Finally, as the number approaches 40, APE boosts up. This evidence suggests that after a long term of stable growth, the multinationals become large when economies of scale and scope overweigh diseconomies and eventually allow the multinational to gain superior position in the market, which makes it more capable of capturing more opportunities.

Initial value of expansion ($y_{i,90}$, $S^n_{i,90}$, $S^{n2}_{i,90}$ and instruments (TOP_i, EU_i, US_i) are all time invariant so their effects cannot be identified (as pointed in Wooldridge, 2002, p. 488). But statistically insignificant coefficients of initial expansion $y_{i,90}$ in Wooldridge's approach and the Bayesian approach still can provide information about the dependence between the unit-specific effect α_i and initial expansion. And these evidences are reinforced by the insignificant value of η_2 in the Heckman's approach in Table 2 and 3. Therefore, we can conclude safely that there is no evidence of correlation between unit-specific effect α_i and initial condition of expansion. Moreover, ι, the coefficient of inter-temporal correlation between u_{it} is reported in Heckman and Wooldridge's approach and none is statistically significant.

5.2 Size Proxied by Operational Revenues

For the short sample period we replace the number of subsidiaries $S^n_{i,t-1}$ with the operational revenues $S^s_{i,t-1}$. This variable introduces more heterogeneity of size into the expansion model. Results are reported Table 4. First of all, we find that for three correct models (Heckman, Wooldridge and Bayesian), the coefficients of $y_{i,t-1}$ are much smaller than those in 15-year sample and they are merely statistically significant at 10 per cent level (expect Bayesian model). Again Pooled-probit model overestimates the state dependence.

Turning to the effect of size, lagged size shows negative but insignificant effects on expansion in three correct models. While the coefficients of its time mean are significantly positive in two out of three models. The negative effect of operational revenues is consistent with negative effect of the subsidiaries seen in the long sample. Again, the effect of initial value $S^s_{i,90}$ cannot be identified and interpreted.

The coefficients of inter-temporal correlation between errors, ι, in Heckman and Wooldridge's approach are positive and statistically significant at 1 per cent level. This justifies our choice of random-effect models to correct this issue.

5.3 Predictions

5.3.1 Size Proxied by Number of Subsidiaries

We make use of the primitives of the dynamic probit estimations and perform a forecasting experiment on mnes' growth in terms of number of subsidiaries. It is as if the estimated primitives would be the long run decision parameters that explain mnes expansion. Our prediction exercise starts in period 2, 1992 where we observe the initial conditions (period 0 is 1990) and the values of the variables in period 1. We use the primitives of the model (the estimated parameters) to predict the expected number of subsidiaries captured by each mne in period 2. We do not account for individual heterogeneity, given the parameters from the model are all non-significant (see Tables 2-3). We then extend the prediction to period 3, 1993, but this time we make use of the predicted number of subsidiaries resulting from period 2. We repeat our predictions up to period T (2003), of course every time updating the predicted number of subsidiaries. We then compute the R^2 (Table 5) between the predicted number of subsidiaries (derived from our model) and the true number of subsidiaries owned by each mne. We eventually plot the concentration Lorenz curves and Sutton's lower bounds for the entire period (Figure 6a).

5.3.2 Size Proxied by Operational Revenues

Figure 3 compares the cumulative density functions of size, expressed in terms of operational revenues, and number of opportunities. What is noticeable from the figure is that the smooth increase in number of opportunities is not followed by an increase in the size of the mnes. Only in the last 10%-20% of the distributions number of opportunities and size track each other. This interesting pattern is more readable in the next figure, Figure 4, which shows the relation between operational revenues and number of opportunities [both normalized to the space $(0, 1)$]. A 45 degree line would exhibit a pattern of opportunity-homogeneity, with the implication that the more opportunities a mne captures, the larger that mne would be. Figure 4 exhibits also a rather unexpected interesting pattern: a large heterogeneity in opportunities seems to be peculiar to those mnes that have captured a small number of opportunities. There is a coexistence of smaller mnes, in terms of number of opportunities, that have captured large opportunities, and smaller mnes that have captured small opportunities. In order to capture the particular relation between number of opportunities and size, we make use a Gumbel copula approach which has the idiosyncrasy of capturing positive correlations in the highest part of the distributions. The copula approach is outlined in Appendix 3.

We estimate the correlation coefficients (see Table 7) to recover the joint distribution (size, number of opportunities) and subsequently we simulate size conditioning to the observable number of opportunities, as shown in Figure 5. The final step is to integrate these simulations along with the primitives of the dynamic probit estimations (Table 4) and predict operational revenues for the period 1997-2003 and use our predictions to evaluate the goodness of the various modeled use based the R^2 (Table 6).

Also in this case we plot the concentration Lorenz curves and Sutton's lower bound (Figure 6b).

6 Conclusions

Understanding the dynamics of the European pharmaceutical industry is crucial for industrial analysts and policy makers to evaluate the impact of the SMP on this industry and to predict its development. We study this industry's expansion dynamic in terms of mnes' expansion through the channel of establishing new subsidiaries or acquiring subsidiaries from other multinationals. We find that state dependence of the current expansion and past expansion in this industry is modest in the period of the SMP that last year's expansion increases the probability of expansion in the current year by around 6 to 17 per cent. Although no comparable researches are available, we argue that this is an indication that the pharmaceutical industry in the European Union is at an equilibrium state.

Measured by the number of subsidiaries and multinational's operational revenues in the last year, multinational heterogeneity is also found to be crucial in the expansion process. In the model without quadratic terms of the number of subsidiaries, subsidiaries show a negative impact on the probability of expansion. But when quadratic term of the number is included, effects of the number become varying from negative to positive as number increases.

In the model where size is measured by operational revenues, we find negative effect of size, which is in line with the negative effect of subsidiaries.

Sophisticated econometric methods are used in this study to prevent unobserved Random effects of multinationals from contaminating the estimation of state dependence. We use two traditional approaches and a Bayesian approach on the dynamic-Probit model to control the initial-condition problem and Endogeneity caused by Random effects. The results are robust using three approaches. Novel Bayesian approach was rarely applied in the literature of industrial organization before and its performance in this study is promising.

References

[1] Albert, James H. and Chib, Siddhartha "Bayesian Analysis of Binary and Polychotomous Response Data." Journal of the American Statistical Association, 1993, 88, pp.669–679.

[2] Amisano, Gianni and Giorgetti, Maria L. "The Dynamics of Multinationals' Entry and Diversification: A Bayesian Panel Probit Approach." Diparti-

mento di Scienze Economiche, Università di Brescia, Working Paper, May 2005.

[3] Barbosa, Natália "What Drives New Multinationals into an Industry? An Integrative Model of Entry." Working Paper Series no. 23, 2003, Nucleo de Investigacao em Microeconomia Aplicada, Universidade do Minho.

[4] Berry, S. T. "Estimation of a Model of Entry in the Airline Industry." Econometrica, July 1992, 60(4), pp. 889-918.

[5] Bottazzi, G.; Dosi, G.; Lippi, M.; Pammolli, P. and Riccaboni, M. "Innovation and Corporate Growth in the Evolution of the Drug Industry, International Journal of Industrial Organization, 2001, 19, pp. 1161–1187.

[6] Bottazzi, G. and Secchi, S. "Gibrat's Law and Diversification." Industrial and Corporate Change, 2006, 15, pp. 847–875.

[7] Brooks, Stephen P. and Gelman, Andrew. "General Methods for Monitoring Convergence of Iterative Simulations." Journal of Computational and Graphical Statistics, 1998, 7(4), pp.434-455.

[8] Canback, Staffan "Diseconomies of Scale in Large Corporations: Theory and Empirical Analysis." Henley Management College, Working Paper, February 2004.

[9] Cassiman, Bruno; Colombo, Massimo; Garrone, Paolo and Veugelers, Reinhilde "The Impact of M&A on the R&D Process. An Empirical Analysis of the Role of Technological and Market Relatedness." IESE Business Scholl, University of Navarra, Working Paper No. 500, April 2003.

[10] Caves, Richard E.; Whinston, Michael D.; Hurwitz, Mark A.; Pakes, Ariel and Temin, Peter "Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry." Brookings Papers on Economic Activity. Microeconomics, 1991, pp. 1-66.

[11] Chib, S. 1992, "Bayes Inference in the Tobit Censored Regression Model." Journal of Econometrics, 51, pp. 79-99.

[12] Chib, Siddhartha, and Greenberg, Edward. "Understanding the Metropolis-Hastings Algorithm." The American Statistician, 1995, 49, pp.327–336.

[13] Cowles, Mary K. and Carlin, Bradley P. "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review." Journal of the American Statistical Association, 1996, pp. 883-904.

[14] Danzon, Patricia M.; Nicholson, Sean and Pereira, Nuno Sousa "Productivity in Pharmaceutical-Biotechnology R&D: The Role of Experience and Alliances." NBER Working Paper No.9615, April 2003.

[15] Dunne, Timothy; Roberts, Mark J. and Samuelson, Larry "Patterns of Multinational Entry and Exit in U.S. Manufacturing Industries." Rand Journal of Economics, Winter 1988, 19(4), p 495-515.

[16] Evans, David S. "The Relationship between Multinational Growth, Size, and Age: Estimates for 100 Manufacturing Industries," Journal of Industrial Economics, June 1987, 35(4), pp. 567-81.

[17] Gelfand, Alan E. and Smith, Adrian F. M. "Sampling-Based Approaches to Calculating Marginal Densities." Journal of the American Statistical Association, 1990, 85, pp.398-409.

[18] Geman, Stuart and Geman, Donald. 1984. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images." IEEE Transactions of Pattern Analysis and Machine Intelligence, 1984, 6, pp.721-741.

[19] Gelman, Andrew and Rubin, Donald B. "Inference from Iterative Simulation Using Multiple Sequences (with discussion)." Statistical Science, 1992, 7(4), pp. 457-472.

[20] Geweke, John "Evaluating the Accuracy of Sampling Based Approaches to the Calculation of Posterior Moments." In Bernardo, Berger, Dawid and Smith, eds., Bayesian Statistics 4, pp. 169-194, Oxford: Oxford University Press, 1992.

[21] Hastings, W. K. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." Biometrika, 1979, 57, pp.97-109.

[22] Heckman J. J. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process", in C.F. Manski and D. McFadden (eds.) Structural Analysis

of Discrete Data with Econometric Applications, Cambridge: MIT Press, 1981.

[23] Kass, Robert E. and Wasserman, Larry "The Selection of Prior Distributions by Formal Rules." Journal of the American Statistical Association, 1996, 91, pp.1343-1370.

[24] Keane, Michael P and Wolpin, Kenneth I. "The Career Decisions of Young Men." Journal of Political Economy, June 1997, 105(3), pp. 473-522.

[25] Klepper, Steven and Thompson, Peter "Submarkets and the Evolution of Market Structure." Department of Economics, Florida International University, Working Papers No. 0303, 2003.

[26] Kyle, Margaret K. "Pharmaceutical Price Controls and Entry Strategies." The Review of Economics and Statistics, 2007, 89(1), pp. 88-99.

[27] Lancaster, Tony "An Introduction to Modern Bayesian Economics." Blackwell Publishing Ltd., 2004.

[28] Metropolis, Nicholas; Rosenbluth, Arianna W.; Rosenbluth, Marshall N.; Teller, Augusta H. and Teller, Edward. "Equation of State Calculations by Fast Computing Machines." The Journal of Chemical Physics, 1953, 21, pp.1087-1092.

[29] Morton, Scott M. "Entry Decisions in the Generic Pharmaceutical Industry." RAND Journal of Economics, Autumn 1999, 30(3), pp. 421-440.

[30] Mundlak Y. "On the Pooling of Time Series and Cross Section Data", Econometrica, 1978, 46, pp. 69-85.

[31] Ornaghi, Carmine "Mergers and Innovation: The Case of the Pharmaceutical Industry." Economics Division, School of Social Sciences, University of Southampton, Discussion Papers in Economics and Econometrics No. 0605, 2006.

[32] Raftery, Adrian E. and Lewis, Steven M. "How Many Iterations in the Gibbs Sampler?" In Bernardo, Berger, Dawid and Smith, eds., Bayesian Statistics 4, pp. 763-773, Oxford: Oxford University Press, 1992.

[33] SCRIP 100 "An analysis of the Pharmaceutical Industry's Performance and Prospects." Winter 2005/2006, KPMG.

[34] SCRIP Reports "Pharmaceutical Mergers & Acquisitions: A Critical Analysis." Surrey: PJB Publications Ltd., 2003.

[35] Stewart, Mark B. "The Inter-related Dynamics of Unemployment and Low-wage Employment." Department of Economics, University of Warwick, The Warwick Economics Research Paper Series (TWERPS), No. 741, 2005.

[36] Stewart, Mark B. "-redprob- A Stata Program for the Heckman Estimator of the Random Effects Dynamic Probit Model." University of Warwick, mimeo, 2006.

[37] Sutton, John "Gibrat's Legacy." Journal of economic Literature, 1997, 35, pp.40-59.

[38] Sutton, John "Technology and Market Structure: Theory and History." Cambridge, MA: The MIT Press, 1998.

[39] Tanner, Martin A. and Wing Hung Wong "The Calculation of Posterior Distributions by Data Augmentation (with Discussion)." Journal of the American Statistical Association, 1987, 82, pp. 528-50.

[40] Williamson, O. E. "Markets and Hierarchies: Analysis and Antitrust Implications." New York: Free Press, 1975.

[41] Wooldridge, Jeffery M. "Econometric Analysis of Cross-Section and Panel Data." Cambridge: The MIT Press, 2002.

Table 2. Dynamic Random-Effect Probit Model without quadratic terms Sample period: 1990 - 2004.[15]

Variable	Pooled Probit	Heckman	Wooldridge	Bayesian
Lagged Entry ($y_{i,t-1}$)	0.414*** (0.075)	0.361*** (0.084)	0.372*** (0.084)	0.307*** (0.080)
Lagged Number of Subsidiaries ($S^n_{i,t-1}$)	-0.137*** (0.018)	-0.135*** (0.018)	-0.136*** (0.018)	-0.133*** (0.018)
Time mean Number of Subsidiaries ($\overline{S^n_i}$)	0.180*** (0.017)	0.178*** (0.018)	0.177*** (0.018)	0.177*** (0.018)
Initial Entry ($y_{i,90}$)	-	-	-0.050 (0.087)	-0.053 (0.099)
Initial Number of Subsidiaries ($S^n_{i,90}$)	0.028 (0.018)	0.035* (0.021)	-	-
TOP	0.446* (0.247)	0.420 (0.260)	0.088 (0.081)	0.091 (0.092)
EU	0.284 (0.288)	0.287 (0.301)	-0.057 (0.085)	-0.056 (0.095)
US	0.394 (0.322)	0.377 (0.337)	0.038 (0.098)	0.041 (0.111)
Constant	-1.752*** (0.116)	-1.767*** (0.118)	-1.742*** (0.136)	-1.780*** (0.138)
ι	-	0.026 (0.022)	0.021 (0.022)	-
η_2	-	2.009 (2.558)	-	-
Average Partial Effect of Lagged Entry (%)	14.3	12.7	12.9	10.8
Log-likelihood	-1229.5228	-1336.0645	-1226.6468	-

[15] * significant at 10 per cent level, ** significant at 5 per cent level, *** significant at 1 per cent level; standard error in parentheses.

Table 3. Dynamic Random-Effect Probit Model with quadratic terms Sample period: 1990 - 2004.[16]

Variable	Pooled Probit	Heckman	Wooldridge	Bayesian
Lagged Entry ($y_{i,t-1}$)	0.552*** (0.078)	0.550*** (0.079)	0.547*** (0.079)	0.484*** (0.082)
Lagged Number of Subsidiaries ($S^n_{i,t-1}$)	-0.568*** (0.043)	-0.568*** (0.043)	-0.567*** (0.043)	-0.568*** (0.042)
Lagged Square Number of Subsidiaries ($S^{n2}_{i,t-1}$)	0.008*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.008*** (0.001)
Time Mean Number of Subsidiaries ($\overline{S^n_i}$)	0.630*** (0.043)	0.629*** (0.043)	0.635*** (0.044)	0.640*** (0.044)
Time Mean Square Number of Subsidiaries ($\overline{S^{n2}_i}$)	-0.008*** (0.001)	-0.008*** (0.001)	-0.008*** (0.001)	-0.009*** (0.001)
Initial Entry ($y_{i,90}$)	-	-	-0.121 (0.087)	-0.123 (0.100)
Initial Number of Subsidiaries ($S^n_{i,90}$)	0.157*** (0.054)	0.414 (2.091)	-	-
Initial Square Number of Subsidiaries ($S^{n2}_{i,90}$)	-0.007** (0.003)	-0.017 (0.087)	-	-
TOP	0.380 (0.249)	0.973 (4.966)	0.048 (0.080)	0.050 (0.091)
EU	0.169 (0.298)	0.440 (2.338)	-0.086 (0.084)	-0.093 (0.095)
US	0.368 (0.329)	0.943 (4.775)	0.005 (0.096)	0.001 (0.109)
Constant	-2.345*** (0.140)	-2.345*** (0.140)	-2.308*** (0.154)	-2.355*** (0.159)
ι	-	0.000 (0.001)	0.000 (0.000)	-
η_2	-	183.959 (1332.585)	-	-
Average Partial Effect of Lagged Entry (%)	16.2	16.2	15.9	14.6
Log-Likelihood	-1157.1961	-1261.3211	-1154.4514	-

[16] * significant at 10 per cent level, ** significant at 5 per cent level, *** significant at 1 per cent level; standard error in parentheses.

Table 4. Dynamic Random-Effect Probit Model without quadratic terms Sample period : 1995 - 2004.[17]

Variable	Pooled Probit	Heckman	Wooldridge	Bayesian
Lagged Entry ($y_{i,t-1}$)	0.453*** (0.089)	0.192* (0.112)	0.194* (0.111)	0.150 (0.109)
Lagged Size ($S^s_{i,t-1}$)	-0.106 (0.131)	-0.095 (0.138)	-0.099 (0.141)	-0.109 (0.143)
Time mean Size ($\overline{S^s_i}$)	0.275** (0.134)	0.297** (0.144)	0.241 (0.147)	0.257* (0.152)
Initial Entry ($y_{i,95}$)	-	-	0.397*** (0.123)	0.415*** (0.131)
Initial Size ($S^s_{i,95}$)	-0.005 (0.088)	0.006 (0.090)	-	-
TOP	0.803*** (0.232)	0.776*** (0.235)	0.168 (0.114)	0.171 (0.124)
EU	0.730** (0.302)	0.741** (0.304)	-0.172 (0.123)	-0.176 (0.132)
US	0.404 (0.343)	0.401 (0.344)	0.012 (0.143)	0.012 (0.153)
Constant	-1.144*** (0.098)	-1.185*** (0.107)	-1.175*** (0.149)	-1.194*** (0.158)
ι	-	0.158*** (0.045)	0.121*** (0.043)	-
η_2	-	0.284 (0.386)	-	-
Average Partial Effect of Lagged Entry (%)	17.3	8.2	7.2	5.9
Log-likelihood	-788.9012	-892.8201	-769.8126	-

[17]* significant at 10 per cent level, ** significant at 5 per cent level, *** significant at 1 per cent level; standard error in parentheses.

Table 5: R^2 of the various specifications with size proxied by number of subsidiaries.

Year	No squared terms				Inclusive of squared terms			
	Pooled Probit	***Heckman***	***Wooldridge***	***Bayesian***	***Pooled Probit***	***Heckman***	***Wooldridge***	***Bayesian***
1992	0.94	0.94	0.97	0.96	0.96	0.93	0.97	0.93
1993	0.91	0.91	0.97	0.94	0.95	0.90	0.97	0.90
1994	0.85	0.84	0.94	0.91	0.89	0.83	0.94	0.84
1995	0.77	0.77	0.88	0.85	0.82	0.77	0.88	0.76
1996	0.68	0.68	0.82	0.78	0.74	0.68	0.82	0.69
1997	0.64	0.61	0.77	0.74	0.68	0.63	0.78	0.64
1998	0.58	0.56	0.73	0.71	0.63	0.57	0.74	0.59
1999	0.55	0.53	0.72	0.70	0.59	0.51	0.73	0.57
2000	0.45	0.45	0.62	0.62	0.50	0.50	0.64	0.49
2001	0.42	0.42	0.61	0.61	0.47	0.47	0.62	0.48
2002	0.41	0.40	0.60	0.61	0.45	0.44	0.62	0.47
2003	0.37	0.37	0.56	0.57	0.42	0.41	0.58	0.43
2004	0.35	0.35	0.53	0.54	0.40	0.40	0.55	0.40

Table 6: R^2 of the the various specifications when size is proxied by operational revenues.

Year	No squared terms				Inclusive of squared terms			
	Pooled Probit	***Heckman***	***Wooldridge***	***Bayesian***	***Pooled Probit***	***Heckman***	***Wooldridge***	***Bayesian***
1997	0.88	0.88	0.89		nd	nd	nd	nd
1998	0.57	0.55	0.59		nd	nd	nd	nd
1999	0.67	0.65	0.62		nd	nd	nd	nd
2000	0.39	0.40	0.45		nd	nd	nd	nd
2001	0.46	0.46	0.52		nd	nd	nd	nd
2002	0.55	0.48	0.67		nd	nd	nd	nd
2003	0.37	0.34	0.48		nd	nd	nd	nd

Table 7: Estimated correlation coefficient from Gumbel copula for size proxied by operational revenues

Year	Correlation coefficient
1997	6.15
1998	6.15
1999	6.32
2000	6.50
2001	6.47
2002	5.97
2003	5.90

A Appendix

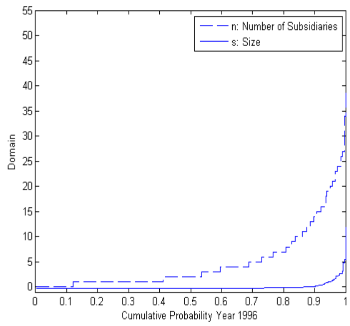
A.1 The Amadeus Database and Identification of Subsidiaries

Compiled by the Bureau Van Dijk, the Amadeus collects both public and private firm accounts for 38 European countries. It is able to provide researchers with comprehensive information on increasing numbers of firms from 1992. This information covers the balance sheet, the profit and loss account, various financial ratios, the ownership data, the industry classification code, address details and the year of incorporation. Therefore, it allows one to trace a firm's birth and evolution over time.

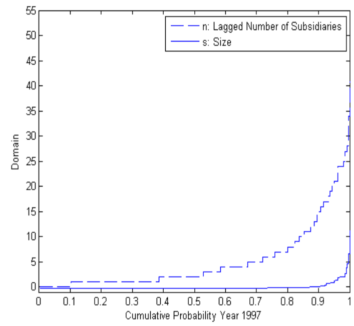
In the build-in ownership database in the Amadeus, each firm is linked to its shareholders and the value of each shareholder's share in that firm is available. We mainly rely on this ownership database to identify a subsidiary firm's parent. Sometimes, a firm may have more than one mne shareholder, and its mne shareholders may be inter-linked. Usually the Amadeus marks one of the mne shareholders as the ultimate owner of the subsidiary firm. In the event that it does not, we define, from among all mne shareholders a subsidiary has, the mne shareholder that has the largest share (directly, or indirectly through other subsidiaries) as its ultimate owner. By doing so, each subsidiary is linked to only one mne shareholder as its mne parent, and all ultimate mne parents defined by this way are independent of each other.

By defining an ultimate owner as having the largest share in a firm, we avoid the complication of joint ventures. According to the ownership database in the Amadeus, only one joint-venture case where two parents have exactly 50 per cent shares each in a subsidiary is found, which is Bracco Spa, an Italian company owned equally by E.MERCK (Germany) and Brafin Finanziaria Spa (Italy). We somewhat arbitrarily treat Bracco Spa as the subsidiary of E.MERCK because E.MERCK is a leading European pharmaceutical multinational.

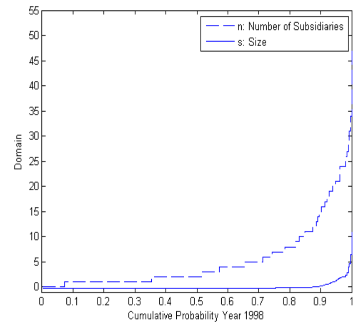
Figure 3: Cumulative density functions of number of subsidiaries and operational revenues



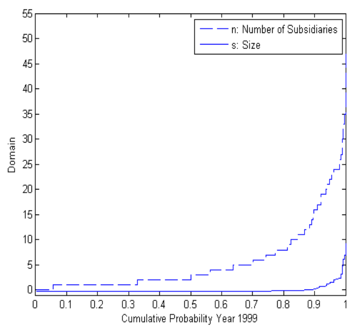
(a) Year1996



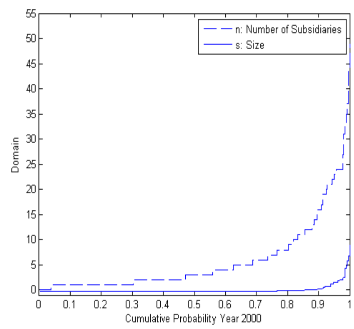
(b) Year1997



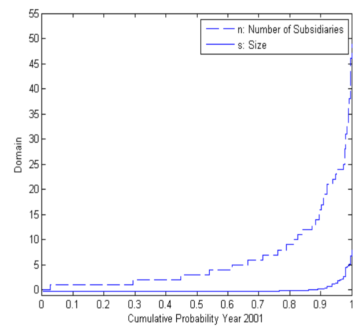
(c) Year1998



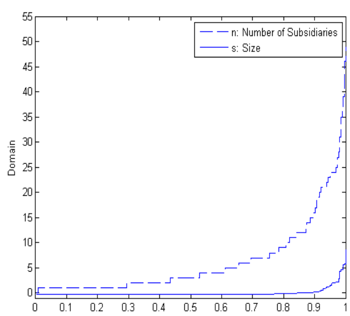
(d) Year1999



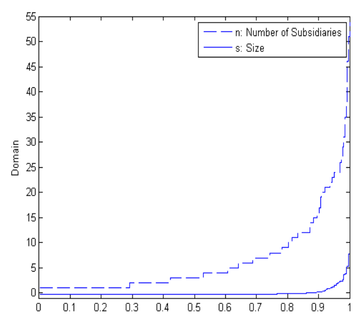
(e) Year2000



(f) Year2001

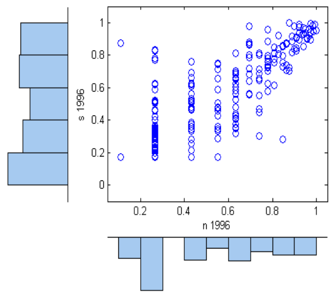


(g) Year2002

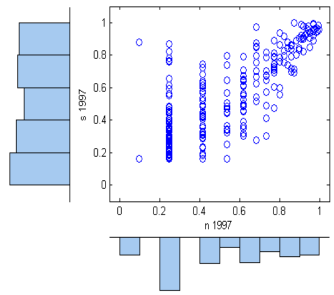


(h) Year2003

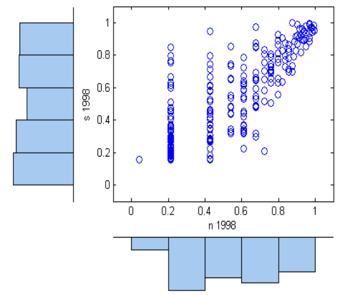
Figure 4: Relation number of subsidiaries and operational revenues



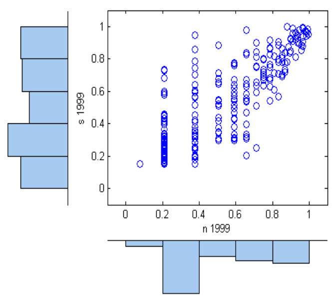
(a) Year1996



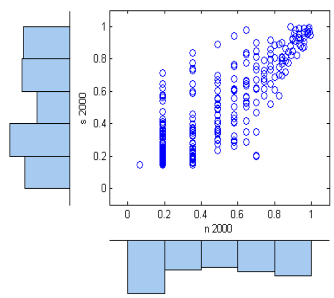
(b) Year1997



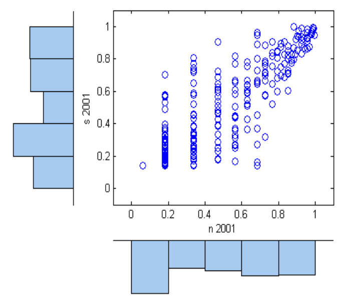
(c) Year1998



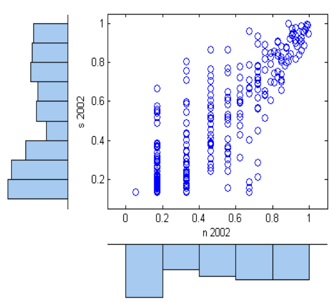
(d) Year1999



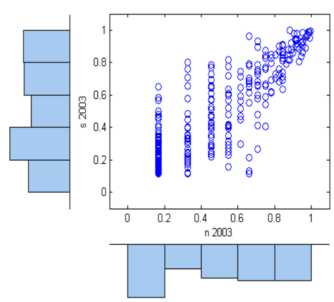
(e) Year2000



(f) Year2001

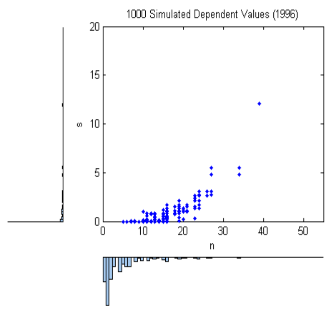


(g) Year2002

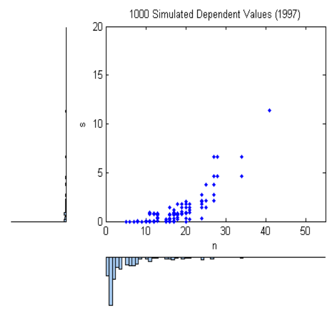


(h) Year2003

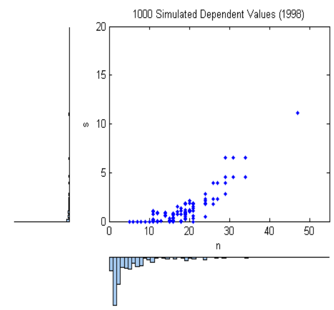
Figure 5: Simulations number of subsidiaries and operational revenues



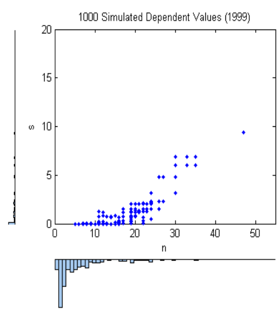
(a) Year1996



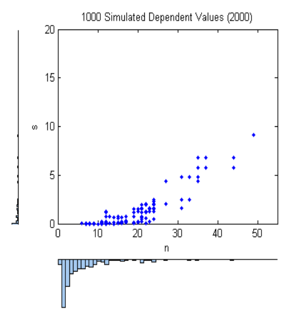
(b) Year1997



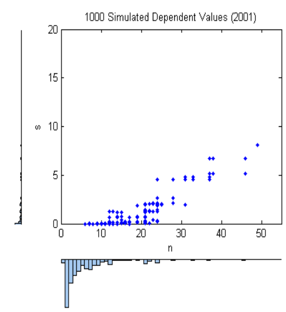
(c) Year1998



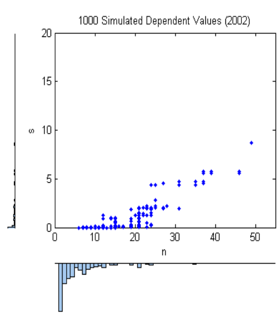
(d) Year1999



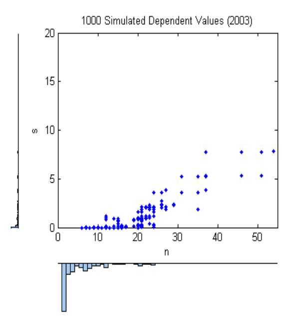
(e) Year2000



(f) Year2001

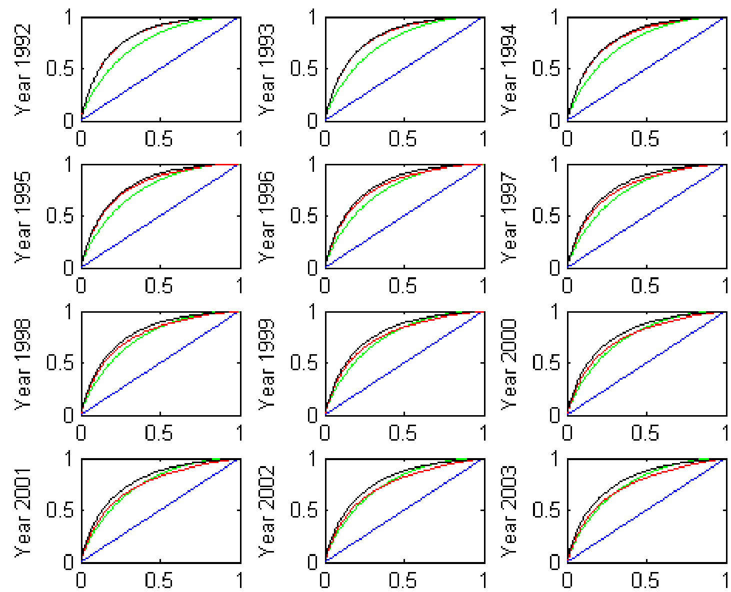


(g) Year2002

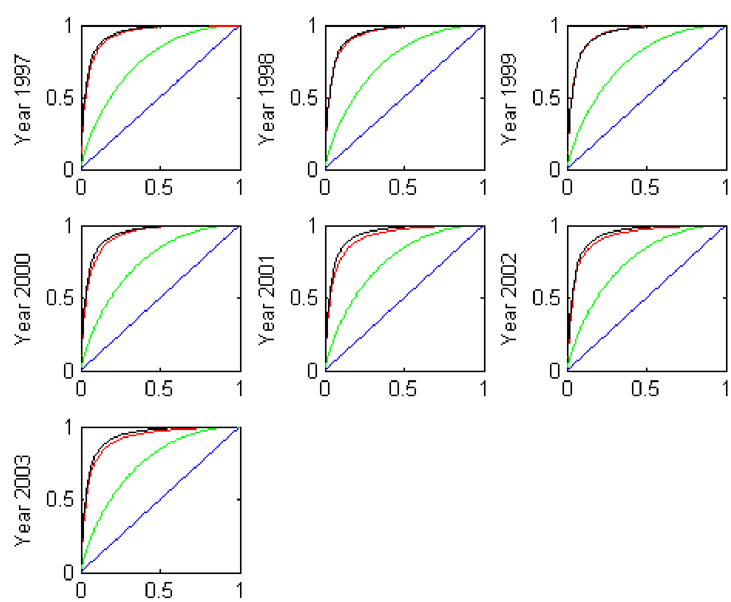


(h) Year2003

Figure 6: Lorenz concentration curves*



(a) Number of subsidiaries



(b) Operational revenues

*The blue line is the symmetric case; the green line is Sutton's lower bound
the red line is the predicted number of subsidiaries; the black line is the actual number of subsidiaries.

Table A1.1: Statistics of the subsidiaries established or acquired and entry by mnes in the period between 1990 – 2004.[18], [19], [20]

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
New subsidiaries established or acquired by each MNE*	na	8 (0.23) [0.73] {60}	3 (0.16) [0.45] {42}	1 (0.11) [0.31] {29}	3 (0.23) [0.53] {61}	4 (0.19) [0.54] {51}	5 (0.27) [0.71] {71}	4 (0.22) [0.62] {59}	6 (0.24) [0.67] {63}	4 (0.20) [0.52] {53}	17 (0.34) [1.42] {90}	7 (0.25) [0.72] {65}	4 (0.15) [0.50] {40}	15 (0.17) [0.99] {44}	15 (0.14) [1.01] {38}
Changes in size** (million USD)	na	na	na	na	na	na	3,805.1 (91.4) [381.1] {18,455.7}	19,352.6 (166.6) [1,374.6] {35,151.3}	3,601.7 (80.4) [322.2] {17,439.7}	5,420.0 (6.2) [584.3] {1,403.5}	3,969.7 (44.4) [300.1] {10,651.5}	10,568.6 (110.3) [788.1] {27,358.2}	7,172.8 (211.4) [757.1] {53,895.0}	15,192.2 (293.3) [1,205.1] {74,510.1}	na
1 new subsidiaries entry %	na	15.5	12.9	10.9	18.5	14.3	16.9	15.1	16.2	15.9	17.8	15.9	10.9	9.5	6.4
1 new subsidiary entry %	na	12.1	10.2	10.9	14.3	10.9	10.9	10.2	11.7	12.8	13.2	10.6	8.3	7.2	4.5
2 new subsidiaries entry %	na	2.3	2.3	0.0	3.8	2.3	3.4	3.4	3.0	2.3	2.3	3.8	1.5	1.9	1.1
> 2 new subsidiaries entry %	na	1.1	0.4	0.0	0.4	1.1	2.6	1.5	1.5	0.8	2.3	1.5	1.1	0.4	0.8
Entry	na	60	42	29	61	51	71	59	63	53	90	65	40	44	38

[18] Note: Geographic coverage: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden, and UK (excl. Luxembourg).

[19] * The first value is the maximum number of new subsidiaries established or acquired by any MNE, the mean of this item is in (), then standard deviation in [] and sum of all new subsidiaries in {}.

[20] ** Size is measured by operational revenue. The first value is the maximum change in size of any MNE, the mean of this item is in (), then standard deviation in [] and sum of all changes in size {}. We only report these statistics from 1995 as too many missing data for the previous years. The number of MNEs with missing size data is reported in Table 1 in text.

Table A1.2: Statistics of Explanatory Variables.[21], [22], [23]

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Number of Subsidiaries*	3.74 (5.35) [32] Novartis	3.97 (5.57) [32] Novartis	4.12 (5.70) [32] Novartis	4.23 (5.77) [33] Novartis	4.46 (5.92) [34] Novartis	4.66 (6.16) [34] Novartis	4.92 (6.48) [39] Novartis	5.15 (6.71) [41] Novartis	5.38 (6.95) [47] Novartis	5.58 (7.10) [47] Novartis	5.92 (7.67) [49] Novartis	6.17 (7.85) [49] Novartis	6.32 (7.94) [49] Novartis	6.49 (8.31) [54] Pfizer	6.63 (8.58) [55] Pfizer
Operational Revenue** (million USD)	-	-	-	-	-	772.8 (2,527.8) [27,991] Bayer	836.2 (2,779.8) [30,436] Bayer	960.2 (3,260.4) [34,620] Bayer	1,017.4 (3,316.6) [34,737] Bayer	971.0 (3,071.2) [28,155] Bayer	978.9 (3,123.5) [28,235] Bayer	1,065.5 (3,320.1) [27,211] Bayer	1,237.9 (3,883.3) [34,384] Bayer	1,510.2 (4,827.3) [38,788] Bayer	-
Top MNEs	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73
EU MNEs	169	169	169	169	169	169	169	169	169	169	169	169	169	169	169
US MNEs	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47
OTHER MNEs	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49
Total number of MNEs	265	265	265	265	265	265	265	265	265	265	265	265	265	265	265

Table A1.3. Transition probabilities

Entry at year t / Entry at year t-1	0	1	Sum
0	88.24	11.76	100.00
1	73.41	26.59	100.00

[21] Note: Geographic coverage: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden, and UK (excl. Luxembourg).

[22] * The first value is the mean of the number of subsidiaries, its standard deviation in (), its maximal value in [] and the MNE with maximal subsidiaries follows.

[23] ** The first value is the operational revenues, its standard deviation in (), its maximal value in [] and the MNE with maximal operational revenues follows.

Figure A1.1: Empirical Density of the Number of Subsidiaries[24]

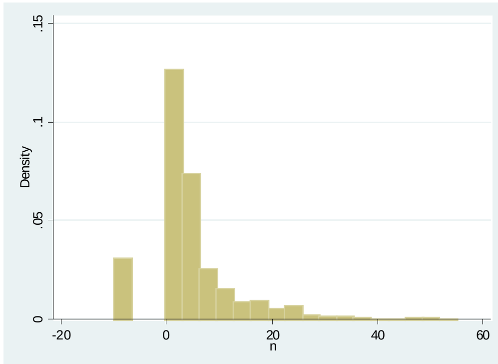
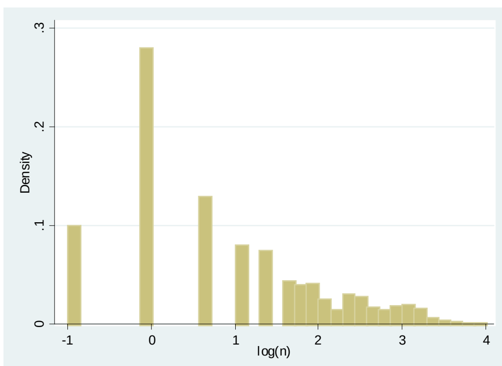


Figure A1.2: Empirical Density of the Number of Subsidiaries (in log form)[25]



A.2 A study of the variance – size relation comparable to Bottazzi and Secchi (2006)

Our data allow us to conduct a study on the relation between firm's growth rate and its size. Previously, Bottazzi and Secchi (BS hereafter) (2003, 2006) studies this relation using a data of 198 world top pharmaceutical mnes and they find evidence that size is negatively correlated with standard deviation of growth rate. We replicate their approach in a strict way to make our results comparable to theirs.

[24]Note : 15 year's data of 265 MNEs is pooled together, after checking the stability of this empirical density for each year. Bin # = 20. n is set to -10 for MNEs with zero subsidiaries to indicate these sleeping MNEs.

[25]Note : 15 year's data of 265 MNEs is pooled together, after checking the stability of this empirical density for each year. Bin # = 20. log(n) is set to -1 for MNEs with zero subsidiaries to indicate these sleeping MNEs.

Let firm size (operational revenues) be S_{it}, then define,

$$s_{it} \equiv \log(S_{it}) - \frac{1}{N}\sum_{i=1}^{N}\log(S_{it})$$

as normalized size for year t. N is the number of mnes in each year with reported value of operational revenue. This normalization is necessary to eliminate the nominal trend of currency in the growth of size. The empirical density and its kernel estimates of normalized size is shown in Figure A2.

Figure A2.1 Empirical density of size and its kernel estimate[26]

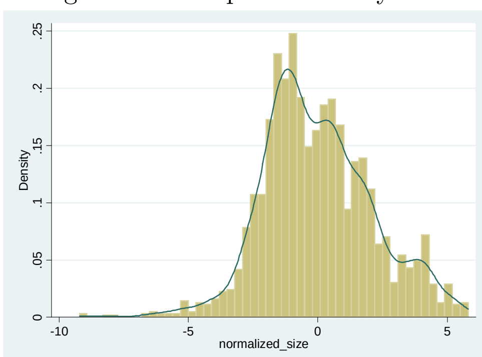
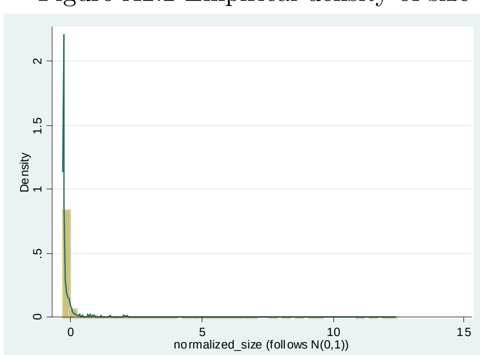


Figure A2.2 Empirical density of size and its kernel estimate[27]



[26] Note: Sizes are available from 1995 to 2003. 9 year's data of maximum 265 MNEs is pooled together, after checking the stability of this empirical density for each year. Bin # = 50. Gaussian kernel density is used.

[27] Note: Sizes are available from 1995 to 2003. 9 year's data of maximum 265 MNEs is pooled together, after checking the stability of this empirical density for each year. Sizes are normalized to follow N(0,1). Bin # = 50. Gaussian kernel density is used.

This empirical density and kernel estimate are similar to what have been found in BS (2003, 2006). With one exception that their empirical density shows two modes near 0 and 4 but ours shows three modes.

We further define log growth rate as

$$g_{it} = s_{i,t+1} - s_{it}.$$

If substitute $s_{i,t+1}$ and s_{it} into this formula, we will have

$$g_{it} = \log\left(S_{i,t+1}\right) - \log\left(S_{it}\right) - \left[\frac{1}{N}\sum_{i=1}^{N}\log\left(S_{i,t+1}\right) - \frac{1}{N}\sum_{i=1}^{N}\log\left(S_{it}\right)\right]$$

and the last term on right hand side in the bracket is an adjustment term for inflation.

We split the pooled sizes, s_{it}, into 51 bins and calculate the mean, standard deviation, skewness and kurtosis of firms in each bin and then fit the following linear model of growth rate and size

$$f\left(g|s\right) = \alpha + \beta s,$$

where g are in turn four moments of growth rate of firms in each bin and s is average log size, s_i, in the bin (note subscript is dropped because sizes are pooled). In each of 51 bins there are 36 mnes and in BS (2003, 2006) they have 50 bins and 36 mnes in each bin).

The results of above linear regression are reported in Table A4. We also report Bottazzi and Secchi's results in the same table. BS (2003) also run the regressions for each year. We did the same thing and find the results are quite similar to the results reported below.

Table A2.1[28]

[28] The differences between BS (2006) and our results are that

1) In our case, mean growth rate and mean size have a statistically significant linear relation, but the reported coefficient is much weaker than that reported in BS (2006) (beta = -0.11(0.08));

2) In our case, the coefficient of size in the model "g s.d. – size" is statistically significant, but is much weaker than that found in BS (2006) (beta = -0.21(0.02)).

Relation	β	Std error	Obs.	β	Std error	Obs.
	Our results			Bottazzi and Secchi (2006)		
g_{it} Mean – Size	-0.036***	0.009	51	-0.11	0.08	50
g_{it} s.d.– Size	-0.062***	0.012	51	-0.21***	0.02	50
g_{it} Skewness – Size	-0.036	0.106	51	-0.12	0.09	50
g_{it} Kurtosis - Size	0.131	0.349	51	-0.03	0.11	50

The differences between BS (2006) and our results are that:

1. In our case, mean growth rate and mean size have a statistically significant linear relation, but the reported coefficient is much weaker than that reported in BS (2006) (beta = -0.11(0.08));

2. In our case, the coefficient of size in the model "g s.d. − size" is statistically significant, but is much weaker than that found in BS (2006) (beta = -0.21(0.02)).

A.3 The Copula Approach

This is a technique used to model joint parametric distributions from marginal distributions of nonnormal data. We use the empirical marginal distributions We can define a copula in the following way. Let F be a bivariate cumulative distribution function with random variables R_1 and R_2 and one-dimensional margins F_1 and F_2, then there exists a bivariate distribution function, copula, C. such that,

$$F(r_1, r_2) = C(F_1(r_1), F_2(r_2); \theta),$$

where is a parameter that measures the dependence between the marginals. Let $U(0,1)$ be the uniform distribution and u_1 and u_2 uniform distributed variates. We can then write $r_1 = F_1^{-1}(u_1) \sim F_1$ and $r_2 = F_2^{-1}(u_2) \sim F_2$. Hence,

$$F(r_1, r_2) = F\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right) = C(u_1, u_2; \theta).$$

This means that if $r \sim F$, and F is countinuous then $(F_1(r_1), F_2(r_2)) \sim C$, and if $U \sim C$, then $\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right) \sim F$.

For the properties and a more detailed description of the copula we refer to Trivedi and Zimmer (2007).

Given our data present strong correlation at high values (see Figures XXX the kernels) we decide to make use of the Gumbel copula which takes the form,

$$C(u_1, u_2; \theta) = \exp\left[-\left(\widetilde{u}_1^{\theta} + \widetilde{u}_2^{\theta}\right)^{1/\theta}\right],$$

where $\widetilde{u}_j = -\log(u_j)$, $j = 1, 2$ with $\theta \in [1, \infty)$.[29]

[29] A value of $\theta = 1$ indicates independence between F_1 and F_2.