



Learning from foreign investment by rival firms: Theory and evidence

Carlo Altomonte^{a,*}, Enrico Pennings^b

^a *Bocconi University, Milan, Italy*

^b *Erasmus University and Tinbergen Institute, Rotterdam, Netherlands*

Received 16 May 2006; received in revised form 10 October 2007; accepted 11 December 2007
Available online 23 December 2007

Abstract

We offer an alternative explanation for follow-the-leader behavior in foreign investment decisions based on Bayesian learning by rival firms. We test the implications of the model through a panel count data sample of MNEs that have invested in Central and Eastern Europe over the period 1990–1997. Interacting the measure of rivals' investment in country-industry pairs with uncertainty, we are able to identify the channel of Bayesian learning about revenue postulated by the model as the only one consistently generating the detected follow-the-leader behavior of foreign investments. The empirical findings are robust with respect to different model specifications.

© 2007 Elsevier B.V. All rights reserved.

JEL classification: C25; D81; F21; L10

Keywords: Discrete choice panel data; Uncertainty; FDI; Bayesian learning

1. Introduction

In the literature on foreign direct investment (FDI) it is well established the idea that foreign entry by a firm may trigger a rival reaction, leading to a follow-the-leader (FTL) behavior in foreign investment decisions: firms (the followers) invest abroad as a reaction to the set up of a foreign affiliate by a first-mover competitor (the

leader). A possible rationale of such a behavior has been originally discussed in the business literature by Knickerbocker (1973) and Flowers (1976), and it is known as, 'oligopolistic reaction'¹: the intuition is that firms, uncertain of production costs in the country to which they currently export, run the risk of being underpriced by a rival that switches from exporting to establishing a manufacturing subsidiary in the host country. By imitating the behavior of the lead investor,

* Corresponding author. Faculty of Economics H13-31, Erasmus University Rotterdam, Burg. Oudlaan 50, 3062PA Rotterdam, Netherlands. Tel.: +31 10 4082166; fax: +31 10 4089149.

E-mail address: pennings@few.eur.nl (E. Pennings).

¹ The concept has been labeled, 'oligopolistic reaction' to foreign investment since the effect was found to be particularly relevant in a market characterized by some degree of industry concentration.

the follower firms can instead match the production cost of the rival firm abroad and thus avoid being underpriced.

This paper tests for the presence of learning to rationalize the follow-the-leader behavior observed in the patterns of foreign direct investment in the market. To the best of our knowledge, this is the first paper proposing a learning mechanism to rationalize the observed FTL behavior of multinational enterprises (MNEs) and testing for it in the data.² In our model, learning can be either about costs (as in [Vettas, 2000](#)) or about revenues. Moreover, in line with a recent literature on firm heterogeneity, the assumption we make on the prior distribution of the variables of concern implies a Pareto distribution for the observations through which learning takes place.

In general, follow-the-leader behavior in FDI decisions is supported by broad empirical evidence. Controlling for variables relevant for the decision to undertake FDI (e.g. the market size of the host country and the distance from the investor's home to its host country), [Yu and Ito \(1988\)](#) consider FTL behavior in two industries, the US tire and textiles. By finding follow-the-leader FDI only in the tire industry, they conclude that firms only react oligopolistically in moderately concentrated industries such as the tire one and not in more competitively structured industries as textiles. More in general, by examining all Japanese investment into the U.S., [Hennart and Park \(1994\)](#) find evidence that FDI by a Japanese enterprise group in the U.S. is more likely if other Japanese rivals have already invested in the U.S.; [Ito and Rose \(2002\)](#) show that, in the same tire industry, firms like Continental and Bridgestone imitate FDI decisions by leading firms like Goodyear and Michelin, with follow-the-leader behavior measured as the impact of the total number of

foreign firms (regardless of when they entered) on the probability of investment by another foreign firm in a given year.

Though these results provide compelling evidence for the phenomenon and have a straight-forward economic interpretation, they however fail to identify the theoretical channels through which the reaction of rival multinational enterprises can arise. Moreover, all the previously quoted studies are based on the study of only one or two industries, while a broader analysis encompassing the relation between industry-specific characteristics and rivals' reaction is lacking.

In our paper, we link the finding of our theoretical model with the recent flow of FDI to Central and Eastern European Countries (CEECs). The sample is chosen since it provides an interesting 'natural experiment': first, the existence of a learning effect seems plausible after the fall of the iron curtain, as many firms considered investing in Eastern Europe because of the expected lower marginal cost and/or possible new market opportunities in the region. Second, the fall of the Berlin wall in 1989 enables us to monitor over time the number of foreign investments taking place in CEECs and the follow-up behavior by rivals, thus controlling for initial conditions. In particular, it is possible to exclude the effect of learning from domestic firms, since these companies were either non-existing or subject to a heavy restructuring process in the early years of transition. Our sample therefore consists of the yearly number of European Union's foreign investors over the period 1990–1997, over a large set of industries and the most important CEECs.³ By identifying the order of entry from the very first investor to late investors, and using a panel negative binomial regression model relating foreign investment in a given year, industry and country to changes in the total number of investors operating in the same industry and country in the previous year,⁴ we are able to explicitly test for a foreign firm's reaction to other firms' entry. In addition we test

² Many other studies (e.g. [Grossman et al., 1977](#); [Cukierman, 1980](#); [Vettas, 1998](#)) have considered the role of (acquiring) information before making an investment in a Bayesian framework. In particular [Cukierman \(1980\)](#) analyzes the effects of uncertainty on the timing of investment of a risk-neutral firm. These studies, however, examine endogenous information arrival, whereas our approach takes the arrival of information as exogenous. A more closely related theoretical model is given in [Hoff \(1997\)](#) who examines the impact of pioneering firms on the entry decision of risk averse potential followers. Nevertheless, in our approach potential entrants all exist in the first period, whereas [Hoff \(1997\)](#) assumes a two-period model where there is a new generation of investors in the second period. Moreover, no empirical evidence is provided of the latter findings. [Chang \(1995\)](#) is the first to test empirically the learning channel as a possible explanation for the sequential entry of MNEs, in the absence however of a structural theoretical framework.

³ We exploit the PECODB database, developed by ISLA-Bocconi, Milan within the EURECO Research Training Network program funded by the EC (www.eureco.org). The database is a firm-specific collection of 4200 FDI operations in the CEECs in the period 1990–2001. In terms of validation, the database records virtually all the first mover investors in the region, and it is able to account for almost 70% of the region's total FDI inward stock up to 1997.

⁴ The negative binomial distribution assumption for the number of investors in a year is the most flexible, leaving the Poisson distribution as a special case. In general, previous applications of econometric count models for panel data are rare. An exception is the relationship between firms' investment in R&D and the number of patent applications, starting with the seminal paper by [Hausmann et al. \(1984\)](#).

for information spillovers from foreign investors in the previous year in the same industry but in other countries.

An additional advantage of our approach is related to the comparison of alternative channels put forward by the literature for explaining the FTL behavior. Head, Mayer and Ries (2002) formalize the original Knickerbocker (1973) rationale within a Cournot-type model where there is cost uncertainty and a certain minimal degree of risk aversion by firms. Though their model is elegant, the derivation of their main result pays however tribute to strong restrictions on the parameters' space and on the underlying assumptions of the model (e.g. equal slopes of the demand curves in both the home and the host country). In addition, if there is not enough uncertainty and/or firms are not sufficiently risk averse, FDI decisions turn out to be, in their framework, strategic substitutes rather than strategic complements.⁵

Leahy and Pavelin (2003) provide a simpler theoretical explanation for follow-the-leader behavior in FDI decisions. In their model domestic rivals may be motivated to imitate the leader's FDI when this facilitates collusive behavior in the foreign market. However, since FTL foreign investment only hinges on the possibility to collude, neither uncertainty nor risk aversion play a role in driving their main result; moreover, their framework is also exclusively based on the existence of an oligopolistic market structure.

An alternative explanation for follow-the-leader FDI can be inferred within the theories of economic geography. FTL foreign investment is in fact consistent with a pattern of FDI that is spatially agglomerated: if the trade-off between competition effects and agglomeration forces is solved in favor of the latter, it becomes profitable to follow abroad the leader investor.⁶ Thus, the agglomeration channel for FTL foreign investment predicts that the latter is more likely the higher the number of early investors, in contrast with the findings of the previously quoted papers, where an oligopolistic market structure is crucial for the generation of a follow-the-leader behavior. In particular, Crozet et al. (2004) perform an analysis of FDI in French regions, showing that the location of new entrants is positively and significantly influenced by the proximity of other MNEs of

the same nationality.⁷ Analogously, Buch et al. (2005) in their study of German FDI determinants find positive agglomeration effects working through the number of other German firms that are active in a given host country.

Finally, a follow-the-leader pattern in the undertaking of foreign direct investment might be exogenously generated by firm or industry-specific characteristics: firms (thus with no leader-follower relation) would commonly observe a signal, unobserved by the econometrician, that reduces their uncertainty, but some of them might be more efficient in reacting to this signal and exploit first-mover advantages, with the 'losers' bunching behind them in terms of investment timing. In a similar way, time-to build heterogeneity might generate an exogenous sequential pattern in FDI inflows.

Our framework is able to encompass the main implications of all these alternative FTL models. Therefore, the exercise allows us to precisely identify the channel which, among the possible alternative explanations, is more consistent with the detected pattern of sequential investment by MNEs. To this extent, our results show that, alongside more traditional determinants of FDI, follow-the-leader behavior driven by our proposed channel of Bayesian learning by rival firms plays a significant role in driving MNEs' decisions to invest abroad. More specifically, the results indicate that firms learn about revenue rather than cost. This result is robust with respect to different model specifications which control for both industry and country heterogeneity.

The paper is structured as follows. Section 2 presents a simple model of rival MNEs' reaction through Bayesian learning, whose implications are tested through the econometric approach presented in Section 3 against possible alternative channels driving FTL foreign investment. The results are discussed in Section 4, while Section 5 extends the empirical approach by considering learning from investments in other CEECs than the host country. Finally, Section 6 concludes.

2. FDI and Bayesian learning

Following related studies on herd behavior and informational cascades (e.g. Banerjee, 1992 and Bikhchandani et al., 1998), we assume that the expected payoff from adopting or rejecting an action, i.e. investing abroad in this case, depends on observable signals,

⁵ Firms actions are strategic complements (substitutes) when an increase in the action of one firm raises (lowers) the marginal benefit of an increase in the action for another firm.

⁶ Head et al. (1995) are among the first to put forward the idea that proximity to other firms may play a role in the spatial distribution of multinational activities. Head and Mayer (2004) provide a recent empirical application of economic geography concepts to the location patterns of FDI.

⁷ Crozet et al. (2004) indeed suggest that learning might explain the difference in the coefficients of French rather than foreign rivals as a determinant of new FDI location: new foreign entrants seem to "learn" more from the proximity with French firms, although no specific test of this hypothesis is carried out.

e.g. the cost of production that a rival pays in the host country. The undertaking of a FDI by a rival firm resolves some uncertainty surrounding the profitability of the investing firm in the host country. This information then becomes common knowledge to all other potential investors in the same country/industry, and thus can generate, *ceteris paribus*, a FTL foreign investment.⁸ The analysis concentrates on investments in the entire sector, aggregated in such a way that (many) firms are producing different goods. As the level of analysis is not very disaggregated, we exclude possible strategic effects through price or capacity by first movers.

2.1. Theory

To formalize Bayesian learning, suppose that the (conditional) distribution of a variable x that is relevant for the investment decision (e.g. cost or revenues) in a country i and industry j at time t is exponential with parameter $\lambda_{ijt} > 0$, where λ_{ijt} is the expected value of the variable in the given country/industry/year. The probability density function can then be written as

$$h(x_{ijt}|\lambda_{ijt}) = \lambda_{ijt}^{-1} e^{-x_{ijt}/\lambda_{ijt}} \quad (1)$$

For a risk-neutral firm, the decision to invest will be made based on the expected value of x_{ijt} , which is uncertain. Now, let the prior distribution for λ_{ijt} be inverse gamma with parameters $\alpha > 0$ and $\beta > 0$. The distribution can then be written as

$$g(\lambda_{ijt}) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda_{ijt}^{-\alpha-1} e^{-\beta/\lambda_{ijt}} \quad (2)$$

where $\Gamma(\alpha)$ is the gamma function. Both the mean and the variance of the inverse gamma distribution exist for $\alpha > 1$ and $\alpha > 2$, respectively, and are known to be

$$E[\lambda_{ijt}] = \mu = \frac{\beta}{\alpha - 1}; \quad \text{VAR}[\lambda_{ijt}] = \frac{\mu^2}{\alpha - 2} \quad (3)$$

We can now prove the following:

Lemma 1. *Let x_{ijt} be an observation from an exponential distribution where the mean λ_{ijt} is inverse gamma*

⁸ Given the scope of the paper, we neglect here related issues such as the rival's decision on sharing information of its costs. Shapiro (1986) provides the condition under which full revelation on cost information is optimal from the firm's perspective. We also abstract from strategic considerations on the timing of the first movers, since wait and see strategies of MNEs in a real option framework have already been analyzed by Pennings and Altomonte (2006) in the same context of the CEECs.

with parameters $\alpha > 0$ and $\beta > 0$. Then, the unconditional distribution of x_{ijt} , denoted by $f(x_{ijt})$ is Pareto distributed.

Proof. The unconditional distribution is $f(x_{ijt}) = \int_0^\infty h(x_{ijt}|\lambda_{ijt})g(\lambda_{ijt})d\lambda_{ijt}$. Using Eqs. (1) and (2) we have that $f(x_{ijt}) = \frac{\alpha\beta^\alpha}{(x_{ijt}+\beta)^{1+\alpha}}$ which is the Pareto distribution. \square

Lemma 2. *The sample mean of x_{ijt} is greater than the prior mean.*

Proof. The sample mean is the mean of the Pareto distribution, which equals $\alpha\mu$. The prior mean is the mean of the inverse gamma distribution, which equals μ under the condition $\alpha > 1$ (see Eq. (3)). It then follows that $\alpha\mu > \mu$. \square

Note that the reason for the difference between the prior mean of the distribution and the mean of the observations goes back to the skewness of the exponential distribution. As the expected value of the exponential distribution is uncertain, and the distribution is skewed with a non-negative support, higher uncertainty about the mean increases the right tail of the distribution and therefore the mean of the sample observations.⁹ Analogously, it is relatively straightforward to prove the following:

Lemma 3. *The coefficient of variation of x_{ijt} is decreasing in α .*

Proof. If observations are Pareto-distributed, the variance equals $\frac{\mu^2\alpha}{\alpha-2}$. The coefficient of variation, defined as the standard deviation over the mean, then equals $\frac{1}{\sqrt{\alpha^2+2\alpha}}$, which is decreasing in α . \square

Endowed with these results, let c_{ijt} be the total cost of production and R_{ijt} the revenue for a specific country and industry at a specific year. Furthermore, suppose that for each firm f there is an idiosyncratic shock to profitability π^f . We will distinguish between two cases: (i) c_{ijt} is uncertain while R_{ijt} is certain and (ii) c_{ijt} is certain while R_{ijt} is uncertain. Denoting μ_c and μ_R as the prior mean $E[\lambda_{ijt}] = \mu$ in the case of cost and revenue uncertainty, respectively, the sequence of the investment decision is as follows. If there is cost uncertainty, a risk-neutral firm for which $\pi^f + R_{ijt} > \mu_c$ at the arrival of the investment opportunity will invest first; with revenue uncertainty, instead, firms will start to invest when $\pi^f + \mu_R > c_{ijt}$. In other words, there is a critical

⁹ The assumption of a Pareto-type distribution of firms is being increasingly used by the literature on international investment. Helpman, Melitz and Yeaple (2004) provide some empirical tests of its validity.

level of π^f above which firms become first-mover investors, conditional on their prior expectations on costs (μ_c) or revenues (μ_R). After the initial investment (s), firms learn about revenue in the case of revenue uncertainty or about cost in the case of cost uncertainty. This information then becomes common knowledge to all other potential investors in the same country/industry, and thus can generate a change in their expected profitability, causing follow-on investment.¹⁰

Potential followers thus observe either the cost or the revenue of investment and update their prior distribution for λ_{ijt} . Define λ_{ijt+1}^c and λ_{ijt+1}^R respectively, as the *posterior* mean of their cost and revenue distribution. When uncertainty is about cost (revenue), followers would then invest when $\pi^f + R_{ijt+1} > \lambda_{ijt+1}^c$ (or $\pi^f + \lambda_{ijt+1}^R > c_{ijt+1}$ under uncertain revenues). In the case of cost updating, given cost observations c_1, \dots, c_n the posterior distribution for λ_{ijt+1}^c is also inverse gamma¹¹ with parameters $\alpha+n$ and $\beta+\sum_{i=1}^n c_i$. Hence, the posterior mean can be written as $\lambda_{ijt+1}^c = (\beta+n\bar{c})/(\alpha-1+n)$, where \bar{c} stands for the mean of the cost observations up to time t . Recalling Eq. (3) we then obtain $\lambda_{ijt+1}^c = [\mu_c(\alpha-1) + n\bar{c}] / (\alpha-1+n)$. With Bayesian updating, an increase in the observations n then implies a greater weight given to the mean of the data (\bar{c}), and a lower weight to the prior mean (μ_c). A similar reasoning applies when uncertainty is about revenue, letting \bar{R} denote the average revenue of the revenue observations in that case.

For our analysis, we are interested in the probability of investment given that n foreign firms have invested before. More in particular, we want to examine the impact of the number of previous investments, and of uncertainty about the cost or revenue, on the conditional probability to invest. We can now prove the following:

Proposition 1. *Both the posterior mean of cost λ_{ijt+1}^c and of revenue λ_{ijt+1}^R , on average, increase in the number of firms n that have invested.*

Proof. From Lemma 2, we have that $E[\bar{c}] > \mu_c$ and $E[\bar{R}] > \mu_R$. Hence, $E\left[\frac{\partial \lambda_{ijt+1}^c}{\partial n}\right] = E\left[\frac{(\alpha-1)(\bar{c}-\mu_c)}{(\alpha-1+n)^2}\right] > 0$ and $E\left[\frac{\partial \lambda_{ijt+1}^R}{\partial n}\right] = E\left[\frac{(\alpha-1)(\bar{R}-\mu_R)}{(\alpha-1+n)^2}\right] > 0$.

Proposition 2. *Both the posterior mean of cost λ_{ijt+1}^c and of revenue λ_{ijt+1}^R , on average, increase in the coefficient of variation of the distribution.*

¹⁰ Clearly, when revenues and costs are time dependent, firms can also invest at a later date for reasons other than Bayesian learning.

¹¹ The family of gamma distributions serves as a conjugate prior in this case, meaning that the posterior distribution has the same distribution as the prior.

Proof. Given Lemma 3, it suffices to calculate $E\left[\frac{\partial \lambda_{ijt+1}^c}{\partial \alpha}\right] = E\left[-\frac{n(\bar{c}-\mu_c)}{(\alpha-1+n)^2}\right] < 0$ and $E\left[\frac{\partial \lambda_{ijt+1}^R}{\partial \alpha}\right] = E\left[-\frac{n(\bar{R}-\mu_R)}{(\alpha-1+n)^2}\right] < 0$.

Proposition 1 shows that cost uncertainty and revenue uncertainty have different impacts on investors that update their prior means in a Bayesian way. When uncertainty is about cost, the probability of investment (i.e. the probability that $\pi^f + R_{ijt+1} > \lambda_{ijt+1}^c$) decreases for a Bayesian investor, while it increases for a Bayesian investor that faces revenue uncertainty (since in this case she will invest if $\pi^f + \lambda_{ijt+1}^R > c_{ijt+1}$). Hence, uncertainty about revenue can lead to follow-the-leader investment, as discussed in the introduction. For later investors, the information content from investments decreases and there is less reaction to previous investments. For very high n , the posterior approximately equals the average of the number of observations, which will hardly be affected by an additional entry.

Proposition 2 allows us to distinguish the different explanations of follow-the-leader investment. It shows in fact that firms learn more about cost or revenues with increasing uncertainty. So, the probability of follow-the-leader investment becomes greater if uncertainty increases and there is learning about revenue. In that case, a Bayesian investor puts more weight on the observation revenue of the first investor and less on the prior mean. If learning is about cost, however, there will be a negative reaction to early investment, where the size of the change in probability increases in uncertainty.

In order to identify the Bayesian learning channel as the one relevant for FTL foreign investment, we will model in Section 2.2 the variables driving the unconditional probability of MNEs' entry from the very first investor on. Section 2.3 controls for the variables driving the conditional probability, leading to the FTL behaviour, while Section 2.4 comes up with an identification strategy that allows us to validate our proposed learning channel and exclude alternative explanations for FTL foreign investment.

2.2. Empirical model design: unconditional probability

We suppose that both the (expected) revenue R_{ijt} and the (expected) cost c_{ijt} depend on a vector of exogenous variables (X_{ijt}), country and industry fixed-effects and time dummies. As a change in the probability of investment in the sector-country pair leads to a change in the expected number of investors, we use the number y_{ijt} of foreign investments undertaken in sector i of a particular CEEC j at time t , as derived from the previously discussed PECODB dataset. We consider 48 NACE2 and 3 sectors in both manufacturing and

services, and 7 countries in Central and Eastern Europe (Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovakia, Slovenia) over the period 1990–1997. The first year is the one in which the investment opportunities were created by the start of the transition process, with these opportunities eventually exploited by first-mover multinationals.

In particular, the number of first-mover investors amounts on average to 7.9% of all investments in our sample, while the percentage of investors rises to 10.5% in the year after the first investment.¹² The latter finding thus provides some preliminary evidence of follow-the-leader FDI within the industries of the CEECs considered. Although the method and speed of liberalization differed across the CEECs, our study can pick-up follow-the-leader behavior as far as we pick-up the correct order of investment, starting from the very first. As foreign investments were virtually prohibited in the CEECs before 1990, with very few exceptions in particular in Hungary, this condition seems to hold in our sample.

Given our model design, on the right-hand side of the estimating equation we include standard covariates for revenues and costs used in the literature on FDI determinants: $SIZE_{jt}$ (log of country j 's market size in year t proxied by its population), $GDPPC_{jt}$ (log of gross domestic product per capita at time t), $DIST_j$ (log of kilometeric distance between the capital city of the host country and an average EU location, as a proxy for transport costs), $RELWAGE_{jt}$ (the yearly average monthly gross wage of each country divided by the yearly average of the countries considered, as a proxy for relative labor costs).¹³ In order to control for industry-specific effects on revenues and costs, we also include two industry-related variables, namely the average size of the industry ($INDSIZE_{it}$), proxied by the share of each industry considered in each country's gross value-added in year t , and dummies for industries with high sunk costs ($HIGH_i$) and moderate sunk costs (MED_i), respectively. The dummies are constructed with a reference to Davies and Lyons (1996) who classified

industries based on their NACE-codes as advertising and/or R&D intensive. The dummy for high sunk cost industries takes a value of 1 if the industry is both advertising and R&D intensive, while the dummy for moderate sunk costs is 1 if the industry is either advertising or R&D intensive (see our working paper for a more detailed description of all variables). The countries included in the present analysis are Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovak Republic, Slovenia. A list of the considered industries is reported in our working paper.

2.3. Empirical model design: conditional probability

According to our theoretical model, the variables driving the Bayesian learning channel are related to the uncertainty with which costs and revenues are observed and to the number n of early-mover, information-revealing firms that a potential investor observes in a period. In order to condition our FDI probability upon the presence of early-mover investors (the parameter n), a series of categorical dummies is constructed indicating the amount of investment in the previous years. The categorical dummies are L_{ijt} , M_{ijt} , H_{ijt} , and $V H_{ijt}$. They indicate the number of early investors observed by a firm in industry i of country j at time t : more specifically, the dummies take on a value of 1 if the firm observes, respectively, the first and/or second early investment (L_{ijt}), the third, fourth or fifth early investment (M_{ijt}), the sixth until tenth early investment (H_{ijt}) or the eleventh or later ($V H_{ijt}$) early investment in the same country and industry pair.¹⁴

Clearly, when all dummies take value zero, we will be considering the first mover in each country and industry pair, and thus we will be modelling the unconditional probability of undertaking a FDI. On the contrary, when there is entry by rival investors in the prior years in the analysed country/industry pair, we will be de facto measuring FTL foreign investment under its various possible channels.¹⁵ Moreover, all investment dummies implicitly capture the competition effect of rivals on revenues (neglected in the theoretical model),

¹² Detailed data on FDI dynamics as well as a Statistical Annex can be found in the working paper version of the article available on the authors' web pages.

¹³ Our FDI covariates are virtually identical to those used, for example, in Buch et al. (2005) in their study of German FDI determinants. In their paper, the authors provide a very exhaustive synthesis of the most recent theories on foreign investment, including agglomeration effects, as well as a discussion of the standard proxies used to measure FDI determinants. In particular, they also use distance as a proxy for transport costs in order to discriminate between horizontal and vertical FDI.

¹⁴ The latter assumption is then relaxed in Section 5 of the paper, where we consider investments also from other countries.

¹⁵ In the construction of the dummies, we have opted for a continuous learning hypothesis, i.e. we have taken into account previous investments in any year before t , rather than only in the year $t-1$. In any case, looking at the follow-the-leader pattern, there are very few country/industry pairs where the investment remained constant in the year immediately after the entry of the first mover. As a result, either methodology yields two sets of dummies highly correlated, without significant differences on the final results.

which predicts a negative relation between rival entry in an industry and profitability of entry.

We have proxied uncertainty within a given country and industry pair ij through a variable ($INDUNC_{ij}$) measuring the average coefficient of variation of EBIT (earnings before interest and taxes) of a sample of firms currently operating in the countries and industries under consideration.¹⁶ The source is the AMADEUS dataset, provided by Bureau van Dijck, a consulting firm operating in Brussels, and containing balance sheet data of a sample of roughly 5,000,000 companies operating in Europe. Of the almost 180,000 companies recorded in the seven countries considered in our sample, we have restricted our analysis to the 32,083 firms for which data are available for at least four consecutive years in order to have a meaningful estimate of each firm's EBIT standard deviation. Hence, on average the EBIT coefficient of variation is calculated with 95 firms per observation.

2.4. Empirical model design: identification

First of all, if only the traditional FDI determinants appear significant in our estimates, then all the different channels insofar discussed are not relevant as drivers of FTL foreign investments. If however the investment dummies driving the conditional FDI probability appear to matter for the analysis, then their sign and significance would allow to discriminate among different alternative channels explaining FTL foreign investment:

- a. If the follow-the-leader behavior exclusively depends on a collusive argument à la [Leahy and Pavelin \(2003\)](#), we should observe a significance of the L_{ijt} and, eventually, the M_{ijt} variables, i.e. the dummies measuring the impact of the first/second or third/fifth mover on later investments, but no other dummies, measuring a higher number of early movers, should be significant.
- b. If instead FTL is driven by oligopolistic reaction, [Head, Mayer and Ries \(2002\)](#) show that there is a negative relation between the incentive to invest and previous FDI by a rival investor in the absence of uncertainty. This means that in a regression that includes as explanatory variables the L_{ijt} variable and its interaction with uncertainty, the first variable

should be negative and significant, irrespectively from the interaction term; again, the significance should fade away for other dummies measuring a higher number of early movers. As FDI decisions are complements for sufficiently high uncertainty, the interaction term of L_{ijt} with uncertainty should instead be positive and significant.

- c. On the other side of the extreme, $V H_{ijt}$ shows the relevance of an increase in an already large number of firms in the industry. As such, a significant coefficient of this dummy only, together with the traditional FDI determinants, would reveal that the agglomeration channel drives FTL foreign investment in the considered country/industry pair, as postulated, e.g., by [Crozet et al. \(2004\)](#) or [Buch et al. \(2005\)](#).
- d. If the relevant channel for explaining FTL foreign investment is the one of Bayesian learning, then the interaction between L_{ijt} and $INDUNC_{ij}$ should stand out as the crucial variable. Without uncertainty, there is no learning and the prior mean equals the posterior mean.¹⁷ As the interaction variable is equal to zero when uncertainty equals zero, the coefficient for L_{ijt} alone should not be significantly different from zero. If instead investors learn by Bayesian updating, but the interaction is not included in the model, the coefficient for L_{ijt} should be positive when investors learn about revenue, and negative when they learn about cost (Proposition 1). If the interaction is included, the model predicts (Proposition 2) that the interaction term ($L_{ijt} * INDUNC_{ij}$) is positive and significant when there is learning about revenue, and negative and significant when there is learning about cost. At the same time, as argued before, the coefficient for L_{ijt} should be insignificant, irrespectively of the type of learning (whether about cost or revenue). Moreover, the significance of the interaction term should fade away as higher order dummies (M_{ijt} , H_{ijt} , and $V H_{ijt}$) are considered, since changes in the posterior mean affecting the propensity to invest are *ceteris paribus* highest when the very first observations of production cost are made.¹⁸ This can be easily shown by taking the second order derivative

¹⁶ Note that the coefficient of variation, defined as the standard deviation over the mean, is dimensionless. The EBIT, being an indicator of profitability, incorporates in principle uncertainty (and thus potential learning) on both costs and revenues. We will discuss the robustness of our results to different proxies of uncertainty.

¹⁷ This can be seen by substituting a very high value of α in the equation for the posterior mean.

¹⁸ Incidentally, the progressive loss of significance of the interaction between uncertainty measured through the EBIT and the higher order dummies (M_{ijt} , H_{ijt} , and $V H_{ijt}$) also contributes to rule out the criticism of a potential endogeneity of the $INDUNC$ variable: if the volatility of earnings increases with the number of investors, then the interaction term should in principle remain always significant.

of the probability of investment with respect to the number of investments.

Finally, it can be the case, as already discussed, that exogenous firm or industry-specific characteristics matter in driving the sequential cascade of investments, thus leading to a simultaneity problem and spurious correlation that we would incorrectly interpret as FTL behavior. Clearly there might be a few cases where such an event can take place. However, in order to acknowledge that commonly received signals, interacting with firm or industry-specific characteristics, entirely drive the significance of our results, we should assume that these effects operate systematically across the 336 country/industry heterogeneous pairs considered over time in our model design, and that they cannot be picked-up by industry, country or time-fixed effects, which is unlikely. Moreover, such an explanation can hardly be compatible with the identification strategy proposed for the Bayesian learning channel (d): if common unobserved shocks drive investors, in fact, it is true that the interaction term between our L_{ijt} dummy and uncertainty might become significant, since this information would be more powerful in situations characterized by higher uncertainty, but there is no a priori explanation, at least according to this channel, should we detect a loss of significance of L_{ijt} when the interaction term is included, or a progressive insignificance of the interaction term with higher order dummies.

Thus, the Bayesian learning alternative would stand as the actual channel through which follow-the-leader FDI takes place.

3. Econometric approach

In order to avoid a simultaneity bias, we have lagged all the covariates (not only the investment dummies) one year.¹⁹ Note that by using categorical dummies for modelling previous investments, rather than lagged FDI numbers, we do not introduce serial correlation in the error term, a bias which would have required a dynamic discrete choice panel data model design, a class of models whose properties have however been assessed only for some specific distributions.²⁰

¹⁹ The economic rationale for lagging the covariates is also related to the evidence of the so-called “time to build” period elapsing between the actual timing of investment and the decision to invest. The one-year lag also reduces the problem of spurious correlation from commonly observed signals, as discussed in the previous footnote.

²⁰ See Honoré and Kyriazidou (2000) for a reference to this class of models.

Since the dependent variable is a count variable, the most basic assumption on its distribution is that it is Poisson distributed, hence with a density function which equals

$$f(y_{ijt}|\lambda_{ijt}) = \frac{\exp(-\lambda_{ijt})\lambda_{ijt}^{y_{ijt}}}{y_{ijt}!} \quad (4)$$

Parameter λ_{ijt} represents the mean and depends on covariates by the function $\ln(\lambda_{ijt})=X'_{ijt}\zeta$, where ζ is a parameter vector.

However, a key assumption of the Poisson distribution is that the variance is equal to the mean. Such an assumption is likely to be violated when dealing with our sample, since it is well known that count data typically show overdispersion (i.e. variance greater than the mean) when there is either unobserved heterogeneity and/or, ‘positive contagion’ (one event increases the likelihood of another), two features which are likely to arise given the economic nature of our data. In the case of overdispersion, the Poisson estimates are inefficient, with standard errors biased downwards.

As a result, in line with Hausman et al. (1984) we generalize the Poisson model by introducing an individual unobserved effect in the conditional mean of the Poisson distribution. For mathematical convenience a gamma distribution with parameters ϑ_{ijt} and $1/\gamma$ is assumed for the conditional mean, with ϑ_{ijt} now a function of covariates so that $\ln(\vartheta_{ijt}) = X'_{ijt}\zeta$. The resulting distribution of the dependent variable is a (panel) negative binomial (NB1 according to the specification in Cameron and Trivedi, 1998), the density of which equals

$$f(y_{ijt}|\vartheta_{ijt}, \gamma) = \frac{\Gamma(\vartheta_{ijt} + y_{ijt})}{\Gamma(y_{ijt} + 1)\Gamma(\vartheta_{ijt})} \left(\frac{\gamma}{1 + \gamma}\right)^{y_{ijt}} \left(\frac{1}{1 + \gamma}\right)^{\vartheta_{ijt}} \quad (5)$$

where $\Gamma(\cdot)$ is a standard gamma distribution and $\gamma > 0$. The main advantage of the negative binomial model over a standard Poisson model is that the former allows for a different mean and variance. More specifically, in Eq. (5) the ratio of variance and mean can be calculated as $1 + \gamma$. So the parameter γ can be interpreted as a dispersion parameter. The negative binomial distribution thus becomes a Poisson distribution as $\gamma \downarrow 0$.

Since we are dealing with an industry-, country- and time-specific dimension in count data, where observed heterogeneity or positive contagion are not unlikely, it matters how both the panel nature of the conditional mean and the overdispersion parameter γ are modelled. Hence, in order to prove the robustness of our results, we have provided several different model specifications.

As a benchmark for the econometric analysis, it is convenient to start from the standard Poisson model reported in Eq. (4), thus ignoring the panel dimension in the data (Model 1). Next, a gamma distribution with parameters ϑ_{ijt} and $1/\gamma$ is assumed for each conditional mean λ_{ijt} , leading to the standard negative binomial model of Eq. (5), with overdispersion held constant across all industries and countries pairs (Model 2). Model 3 and 4 use the same density function of Model 2, but tackle the three-dimensional nature of the conditional mean considering industry — and both industry- and country-fixed effects in ϑ_{ijt} , respectively.

As a next step, we explicitly deal with the panel dimension of our data, in which we model the industry mean number of investments in the countries under consideration. More specifically, a gamma distribution is assumed for each industry mean in a given year (λ_{it})²¹ in Model 5. An industry-specific overdispersion parameter (γ_i) is considered in Model 6. Note that in the latter model, since the mean of the $\Gamma(\vartheta_{ijt}, 1/\gamma_i)$ distribution equals ϑ_{ijt}/γ_i , in this case the industry-specific overdispersion parameter also acts as an industry-fixed effect in the mean, along the same lines of Model 3, but this time taking into account the panel nature of the data. Finally, as a further robustness check, we estimate the same negative binomial panel model allowing the overdispersion parameter to vary randomly across groups. More specifically, in Model 7 it is assumed that a monotone transformation of the overdispersion parameter, $(1/1 + \gamma_i)$, is drawn from a beta distribution with parameters ν_1 and ν_2 , allowing for a more parsimonious way to account for heterogeneity in the overdispersion parameter.²²

The following scheme summarizes the different model specifications employed in the analysis.

- Model 1 $y_{ijt} \sim \text{Poisson}(\lambda_{ijt})$
 Model 2 $\lambda_{ijt} \sim \Gamma(\vartheta_{ijt}, 1/\gamma)$
 Model 3 $\lambda_{ijt} \sim \Gamma(\vartheta_{ijt}, 1/\gamma)$ with i -fixed effects in ϑ_{ijt}
 Model 4 $\lambda_{ijt} \sim \Gamma(\vartheta_{ijt}, 1/\gamma)$ with i - and j -fixed effects in ϑ_{ijt}
 Model 5 $\lambda_{it} \sim \Gamma(\vartheta_{ijt}, 1/\gamma)$
 Model 6 $\lambda_{it} \sim \Gamma(\vartheta_{ijt}, 1/\gamma_i)$
 Model 7 $\lambda_{it} \sim \Gamma(\vartheta_{ijt}, 1/\gamma_i)$ with $(1/1 + \alpha_i) \sim \text{Beta}(\nu_1, \nu_2)$

²¹ In other words, the industry mean is assumed to be constant across all countries. This assumption is relaxed in Section 5 of the paper.

²² In this last case the mean of the beta distribution is known to be $\nu_1/(\nu_1 + \nu_2)$. The assumption of a beta distribution leads to a tractable joint probability distribution (see Hausman et al., 1984), so maximum likelihood estimation of the parameters is straightforward.

4. Results

Table 1 shows the results for the pooled specifications of the estimation (i.e. Models 1 to 4 in the previous scheme). The first set of control variables have the expected sign and are overall significant. More specifically, when considering the unconditional regression excluding the investor dummies and their interactions (Model 1), traditional FDI determinants measuring population size and GDP per capita are positive. These results show that horizontal (market-seeking) investment explains a significant portion of the total number of incoming investment. Nevertheless, the relative wage variable is negative and significant, indicating that vertical (efficiency-seeking) investment, where firms outsource activities to the CEECs, is also important.²³ As a result, the sign of distance, which is related with opposite signs for vertical and horizontal FDI, does not appear to be significant in the unconditional regression. The industry size variable is positive and significant, showing that investment is more likely in sectors that are relatively large. Without taking into account the Bayesian learning channel, uncertainty is negative and significant, in line with the finding in the literature on uncertainty and FDI.²⁴ The industry dummies measuring sunk costs are negatively signed and significant and, as expected, industries characterized by high sunk costs seem to deter more FDI than medium ones. The results are robust across the different model specifications: only when the regression takes into account both sector and country fixed effects, some of these variables lose their significance (e.g. Model 4 in Table 1).

Looking at the investment dummy variables across the various models presented in Table 1, there is strong evidence of follow-the-leader behavior through Bayesian learning about revenue, since the previously discussed conditions on the parameter for L_{ijt} and its interaction with uncertainty are met for the different model specifications. Leaving out the interaction between uncertainty and total early investment in the previous year (all models labeled ‘a’ in Table 1), we find a positive and significant effect of the latter variable, and hence evidence that firms react to rivals’ entry especially when the number of previous entrants is low. However, when we include the interaction with uncertainty (all models labeled ‘b’ in Table 1), the dummies measuring early investment tend to be less significant, in particular

²³ Many authors (e.g. Buch et al., 2005) report evidence that both market-seeking and efficiency-seeking strategies have been pursued by MNEs investing in the CEECs.

²⁴ See Brunetti and Weder (1998) for an empirical analysis.

Table 1
Baseline (pooled) models

	Poisson			Negative binomial					
	(1)	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Const	−33.4**(-8.79)	.68(.17)	−.73(-.18)	6.89(1.31)	5.25(1.00)	−14.9**(-2.78)	−16.5(-3.09)	−155.9*(-1.93)	−127.2(-1.57)
Country size	1.03**(29.2)	.54**(13.7)	.52**(13.2)	.50**(9.59)	.49**(9.36)	.84**(15.0)	.83**(14.8)	17.6*(1.90)	14.1(1.52)
GDP p.c.	2.88**(12.3)	.43*(1.78)	.53**(2.20)	.01(.05)	.14(.45)	1.54**(4.60)	1.67**(5.00)	−.75(-.89)	−.29(-.34)
Relative wages	−1.81**(-10.5)	−.64**(-3.68)	−.71**(-4.05)	−.36*(-1.72)	−.45**(-2.11)	−1.05**(-4.77)	−1.14**(-5.15)	−.20(-.30)	−.47(-.70)
Distance	.39(1.28)	−1.27**(-4.05)	−1.16**(-3.67)	−1.67**(-3.85)	−1.54**(-3.54)	−.64(-1.52)	−.51(-1.22)	−	−
Industry size	.16**(13.1)	.08**(6.96)	.07**(6.46)	.06**(3.92)	.06**(3.68)	.06(1.41)	.06(1.48)	.06(1.37)	.06(1.38)
INDUNC	−.06**(-2.64)	−.04*(-1.74)	−.03(-.66)	−.06**(-1.96)	−.13*(-1.76)	−.01(-.43)	−.11(-1.45)	−.03(-.89)	−.12(-1.46)
Medium sunk costs	−.20**(-4.33)	−.07(-1.64)	−.05(-1.19)	−.08(-1.21)	−.06(-.97)	−	−	−	−
High sunk costs	−.31**(-4.60)	−.14**(-2.07)	−.15**(-2.27)	−.11(-1.26)	−.12(-1.37)	−	−	−	−
<i>Number of early investors</i>									
L (0–2)	−	1.05**(13.6)	.87**(4.89)	.94**(9.32)	.60**(2.61)	.46**(4.40)	−.03(-.12)	.40**(3.81)	−.05(-.22)
M (3–5)	−	.58**(8.15)	.68**(4.24)	.57**(5.49)	.71**(3.14)	.21**(2.08)	.39*(1.73)	.19*(1.86)	.39*(1.73)
H (6–10)	−	.23**(3.15)	.13(.73)	.27**(2.52)	.01(.06)	.01(.09)	−.13(-.53)	−.05(-.55)	−.22(-.90)
VH (>10)	−	.51**(6.39)	1.29**(6.53)	.47**(3.99)	1.25**(4.55)	.08(.64)	.85**(3.06)	.07(.63)	.85**(3.08)
VH*Econ Scale	−	.25**(3.26)	.25**(3.18)	.23*(1.95)	.24**(2.02)	.05(.39)	.03(.23)	.06(.47)	.04(.33)
L*INDUNC	−	−	.08(1.16)	−	.16(1.60)	−	.24**(2.30)	−	.22**(2.14)
M*INDUNC	−	−	−.04(-.69)	−	−.07(-.73)	−	−.09(-.95)	−	−.10(-1.05)
H*INDUNC	−	−	.04(.60)	−	.11(1.11)	−	.06(.59)	−	.07(.72)
VH*INDUNC	−	−	−.36**(-4.29)	−	−.36**(-3.14)	−	−.35**(-3.09)	−	−.36**(-3.11)
Time dummies	341.4**	261.2**	271.9**	102.7**	108.9**	116.1**	123.2**	47.1**	49.4**
Industry dummies	−	−	−	−	−	288.1**	290.1**	306.3**	306.4**
Country dummies	−	−	−	−	−	−	−	36.7**	34.4**
Overdispersion γ	−	−	−	1.39**(11.58)	1.37**(12.45)	.95**(10.55)	.93**(10.33)	.89**(9.88)	.87**(9.67)
Log likelihood	−3169.5	−2727.6	−2713.3	−2393.2	−2386.8	−2240.4	−2232.5	−2222.9	−2215.4
N. of obs.	2208	2208	2208	2208	2208	2208	2208	2208	2208
Specification test ^a	−	−	−	668.8**	652.9**	336.1**	326.2**	300.2**	293.8**

Note: T-statistics in parentheses. For time, industry and country dummies the joint test of significance is reported.

** significant at the 5 percent level or more; * significant at the 10 percent level.

^a χ^2 test statistic of LR specification test. Model 1a is the benchmark.

when industry- and country-fixed effects are included in the estimation (Models 3b and 4b in Table 1). At the same time, the interaction variable in these latter cases tends to become positive and significant. As these empirical findings are not consistent with other explanations for FTL behavior, rivals' reaction can only be attributed to the channel of Bayesian learning about revenue that we have put forward. Moreover, our finding of follow-the-leader behavior is robust with respect to industry- and country-fixed effects (Models 3 and 4).

The results also show that the effects of agglomeration on rivals' entry outweigh on average the competition effect for those country/industry pairs in which the threshold of ten previous investments is crossed (the VH_{ijt} dummy takes value 1). The result is consistent with our model: we recall in fact that follow-the-leader FDI induced by Bayesian learning is not in contrast with the agglomeration channel, since, when the number of firms operating in the market is above a critical threshold, the Bayesian learning channel will be less and less influencing rivals' reaction, with FTL foreign investment driven by alternative determinants, i.e. the agglomeration effects detected here. In particular, since it is well-known that economies of scale are one of the sources of agglomeration benefits, as a robustness check the VH_{ijt} variable is interacted with a dummy variable ES that takes the value of 1 in industries where economies of scale are important (Pavitt, 1984). The interaction has, as expected, a positive sign, thus providing some further evidence of agglomeration effects, although it is significant only when country- and industry-fixed effects are not considered, as it can be seen from comparing Models 1 and 2 with Models 3 and 4 in Table 1. Finally, the interaction between uncertainty and VH_{ijt} is negative, a result not in contrast with the combined predictions of Propositions 1 and 2 and the conditions they identify on the total number of firms.²⁵

In terms of model specification, the benchmark Model 1 in Table 1 relies on some restrictive assumptions. In particular, as discussed earlier, the Poisson distribution (Model 1a and 1b) is very restrictive in the sense that it imposes the mean to be equal to the variance. Models 2 to 4 show the results for several specifications of the negative binomial model, which generalises the Poisson distribution allowing for over-

dispersion.²⁶ The estimates of the overdispersion parameter reported in the last row of Table 1 show that the hypothesis of no overdispersion is clearly rejected. Under these more flexible model specifications, our main findings remain however valid, illustrating their robustness.

As to give an idea of the quantitative impact of the Bayesian learning, it is possible to estimate the expected number of entrants using one of the specifications where the Bayesian channel is explicitly measured, by fixing the value of all explanatory variables at the average in a reference sector and reference year (i.e. setting at zero industry and time dummies). Using Model 3b of Table 1, for example, one has that without learning from foreign investors (i.e. when all investor dummies are equal to zero), the average country receives on average 0.57 foreign investments in the reference sector/year. With the first mover dummy L set at 1, the expected number of foreign investments in the reference sector increases to 0.92. These numbers are calculated for average levels of uncertainty. For very low levels of uncertainty, there would be no increase in the expected number of foreign investors, while high levels of uncertainty would raise the expected number of investors to a doubling of the expected number of investors.

As a further step, in Table 2 we have explicitly modelled the panel nature of our data, to better control for exogenous characteristics that might spuriously induce the FTL behavior. The first columns (Model 5) show the results for the regression with a constant overdispersion parameter across groups. As discussed in the previous section, the main difference with Model 2 is that the latter considers random effects for each observation ijt while Model 5 estimates random effects for industries it only. The previous result on the FTL behavior remains industry-fixed effects in the dispersion parameter. As discussed earlier, Model 3 (industry-fixed effects in ϑ) and Model 6 (industry-fixed effects in ϑ) are very close in structure. Since the empirical results are roughly equal, it appears unimportant whether fixed effects are only in the mean (Model 3) or also in the overdispersion parameter (Model 6). Finally, Model 7 reports the specification where dispersion in each sector is randomly drawn from a beta distribution. Both Models 6 and 7 perform significantly better than the panel specification with a

²⁵ An explanation might be linked to the possibility that agglomeration effects are higher in industries where firms tend to be more similar. A higher uncertainty might hence be related to a lower degree of similarity between firms, and hence induce a negative sign in the interaction effect.

²⁶ The specification tests reported in Table 1 are LR specification tests, all nested starting from Model 1. Essentially, in Model 2 the restriction that the overdispersion parameter is zero is rejected. In Model 3 and 4, together with the restriction on the overdispersion parameter, the hypothesis of joint industry (Model 3) and country/industry (Model 4) fixed effects equal to zero is also rejected.

Table 2
Negative binomial panel models

	(5a)	(5b)	(6a)	(6b)	(7a)	(7b)
Const	−19.8**(-4.74)	−21.1**(-5.03)	−13.4**(-2.45)	−14.9**(-2.71)	−10.1*(-1.85)	−11.5**(-2.12)
Country size	.83**(.19.1)	.81**(.18.7)	.82**(.14.4)	.81**(.14.3)	.77**(.13.5)	.76**(.13.4)
GDP p.c.	1.86**(.7.00)	1.96**(.7.37)	1.51**(.4.44)	1.63**(.4.80)	1.26**(.3.71)	1.38**(.4.09)
Relative wages	−1.29**(-6.96)	−1.36**(-7.29)	−1.07**(-4.79)	−1.15**(-5.14)	−.94**(-4.25)	−1.03**(-4.61)
Distance	−.23(-.73)	−.12(-.39)	−.73*(-1.69)	−.61(-1.41)	−.88**(-2.03)	−.75*(-1.75)
Industry size	.05**(.2.06)	.06**(.2.10)	.01(.13)	.01(.06)	.03(.94)	.02(.84)
INDUNC	−.01(-.06)	−.04(-.73)	−.02(-.72)	−.12(-1.51)	−.03(-.80)	−.13(-1.61)
Medium sunk costs	−.16(-.67)	−.14(-.60)	.16(.75)	.13(.61)	−.03(-.20)	−.04(-.28)
High sunk costs	−.12(-.39)	−.15(-.48)	−.41(-1.57)	−.36(-1.36)	−.23(-1.12)	−.21(-1.01)
<i>Number of early investors</i>						
<i>L</i> (0–2)	.62**(.7.51)	.26(1.42)	.49**(.4.74)	.06(.26)	.59**(.5.58)	.14(.59)
<i>M</i> (3–5)	.22**(.2.98)	.36**(.2.12)	.19*(1.87)	.37(1.63)	.26*(2.52)	.44*(1.95)
<i>H</i> (6–10)	.01(.07)	−.04(.27)	.02(.22)	−.20(-.82)	.06(.63)	−.16(-.64)
VH (>10)	.11(1.19)	.85**(.4.05)	.06(.53)	.85**(.2.94)	.12(.98)	.90**(.3.17)
VH*Econ Scale	.03(.32)	.01(.07)	.02(.16)	.01(.09)	.04(.28)	.03(.23)
<i>L</i> *INDUNC	−	.17**(.2.15)	−	.21**(.2.01)	−	.22**(.2.10)
<i>M</i> *INDUNC	−	−.06(-.97)	−	−.09(-.93)	−	−.09(-.95)
<i>H</i> *INDUNC	−	.02(.28)	−	.10(.98)	−	.10(.98)
VH*INDUNC	−	−.34**(-3.91)	−	−.36**(-3.01)	−	−.36**(-3.06)
Time dummies	237.0**	245.5**	111.9**	117.7**	110.2**	116.4**
Overdispersion γ	.59**(.4.92)	.58**(.4.83)	−	−	−	−
ν_1	−	−	−	−	4.38**(.4.13)	4.46**(.4.13)
ν_2	−	−	−	−	2.76**(.4.00)	2.76**(.4.00)
Industry-specific γ_i	No	No	Yes	Yes	No	No
Log likelihood	−2506.9	−2494.0	−2088.3	−2081.7	−2312.8	−2305.9
<i>N.</i> of obs.	2208	2208	2208	2208	2208	2208
Goodness of fit test ^a	441.3**	438.5**	1278.5**	1263.2**	829.5**	814.7**

Note: T-statistics in parentheses. For time dummies the joint test of significance is reported.

** significant at the 5 percent level or more; * significant at the 10 percent level.

^a χ^2 test statistic of LR specification test. Model 1 is the benchmark.

constant dispersion reported in Model 5. Still, the main results with respect to follow-the-leader behavior and the control variables are hardly affected by these more sophisticated specifications.

Finally, as a robustness check, we have recalculated Table 2 using the average coefficient of variation of value added of our firms, limiting it to the manufacturing sector, since value-added in the services industry lacks a clear economic meaning and it is often measured very poorly. The limitation to the manufacturing sector has however the advantage to check whether our results depend on the inclusion or not of services firm, which might behave differently with respect to the other covariates. The results, reported in the working paper version of the article, show that the FTL behavior remains valid also under this specification.²⁷

²⁷ Similar results have also been obtained proxying uncertainty through the coefficient of variation of turnover (sales), although the standard errors of our estimates are slightly higher using the latter proxy.

5. An extension: country-specific heterogeneity

In Table 2 we have reported the results for model specifications in which the panel group structure of the data was considered only across industries i , ignoring possible biases arising from country heterogeneity. A possible additional set of dummies convenient for our analysis considers the fact that information about production costs may also be revealed by firms that invest in the same industry i , but in a country $s \neq j$. Investment dummies CL_{ijt} , CM_{ijt} , CH_{ijt} , CVH_{ijt} therefore indicate, respectively, that in the previous year the first and/or second investment (CL_{ijt}), the third, fourth or fifth investment (CM_{ijt}), the sixth until tenth investment (CH_{ijt}) or the eleventh or later (CVH_{ijt}) investment took place in the same industry i but in another country $s \neq j$ of the seven CEECs considered at time t .²⁸

²⁸ As Yu and Ito (1988) suggest for further research, rival MNEs' reaction may be cross-border through the effect of global competition on firm's activities.

Table 3
Negative binomial panel models controlling for country heterogeneity

	(5a)	(5b)	(6a)	(6b)	(7a)	(7b)
Const	−19.5**(-4.64)	−20.6**(-4.88)	−12.3**(-2.23)	−13.9**(-2.51)	−10.0*(-1.83)	−11.6**(-2.12)
Country size	.83**(19.0)	.82**(18.8)	.81**(14.2)	.81**(14.2)	.76**(13.3)	.76**(13.4)
GDP p.c.	1.84**(6.91)	1.97**(7.35)	1.46**(4.25)	1.63**(4.71)	1.23**(3.63)	1.40**(4.10)
Relative wages	−1.29**(-6.94)	−1.37**(-7.28)	−1.05**(-4.70)	−1.15**(-5.10)	−.93**(-4.18)	−1.03**(-4.58)
Distance	−.27(-.83)	−.14(-.45)	−.81*(-1.87)	−.64(-1.47)	−.88**(-2.05)	−.72*(-1.66)
Industry size	.06**(2.03)	.05**(1.98)	.01(.24)	.01(.10)	.03(1.07)	.03(.92)
INDUNC	.01(.08)	−.28**(-1.96)	−.03(-.73)	−.42**(-2.35)	−.03(-.83)	−.40**(-2.28)
Medium sunk costs	−.20(-.77)	−.18(-.71)	.13(.59)	.09(.44)	−.04(-.25)	−.05(-.33)
High sunk costs	−.19(-.55)	−.22(-.64)	−.54(-2.03)**	−.49*(-1.84)	−.22(-1.02)	−.21(-.99)
<i>Number of early investors</i>						
<i>L</i> (0–2)	.61**(7.33)	.30(1.59)	.49**(4.66)	.10(.40)	.59**(5.55)	.18(.73)
<i>M</i> (3–5)	.19**(2.55)	.31*(1.82)	.18*(1.74)	.29(1.30)	.26*(2.48)	.39*(1.74)
<i>H</i> (6–10)	.03(.42)	.02(.01)	.04(.44)	−.18(-.72)	.09(.84)	−.14(-.56)
<i>VH</i> (>10)	.09(1.01)	.81**(3.86)	.03(.25)	.80**(2.76)	.14(1.05)	.90**(3.17)
<i>VH</i> *Econ Scale	.01(.07)	−.01(.13)	.04(.26)	.02(.17)	.03(.17)	.01(.06)
<i>L</i> *INDUNC	−	.15*(1.83)	−	.19*(1.70)	−	.19*(1.77)
<i>M</i> *INDUNC	−	−.06(-.88)	−	−.06(-.64)	−	−.07(-.75)
<i>H</i> *INDUNC	−	.01(.01)	−	.09(.91)	−	.09(.92)
<i>VH</i> *INDUNC	−	−.33**(-3.83)	−	−.36**(-3.00)	−	−.36**(-3.05)
<i>CL</i>	.07(.49)	−.42(-1.16)	−.09(-.54)	−.89**(-2.02)	.15(.78)	−.54(-1.24)
<i>CM</i>	.15(1.36)	.10(.38)	.15(1.02)	.16(.48)	.26*(1.77)	.24(.70)
<i>CH</i>	−.23**(-2.44)	−.09(-.41)	−.28**(-2.26)	.05(.17)	−.14(-1.13)	.13(.46)
<i>CVH</i>	−.22**(-2.14)	−.37*(-1.89)	−.24*(-1.87)	−.39(-1.54)	−.06(-.46)	−.16(-.66)
<i>CVH</i> *ES	.16(1.51)	.15(1.43)	.08(.62)	.09(.69)	.04(.35)	.05(.44)
<i>CL</i> *INDUNC	−	.24(1.47)	−	.40***(1.96)	−	.35*(1.74)
<i>CM</i> *INDUNC	−	.02(.18)	−	−.01(-.10)	−	.01(.03)
<i>CH</i> *INDUNC	−	−.07(-.74)	−	−.16(-1.32)	−	−.13(-1.12)
<i>CVH</i> *INDUNC	−	.09(1.12)	−	.08(.77)	−	.05(.57)
Time dummies	214.6**	225.7**	101.6**	109.2**	99.5**	107.4**
Overdispersion γ	.66***(4.71)	.67***(4.79)	−	−	−	−
ν_1	−	−	−	−	4.73***(3.61)	4.76***(3.60)
ν_2	−	−	−	−	3.02***(3.21)	2.99***(3.25)
Industry-specific γ_i	No	No	Yes	Yes	No	No
Log likelihood	−2499.2	−2484.7	−2082.5	−2073.3	−2309.5	−2300.4
<i>N.</i> of obs.	2208	2208	2208	2208	2208	2208
Goodness of fit test ^a	456.7**	457.1**	1290.1**	1279.9**	836.1**	825.8**

Note: T-statistics in parentheses. For time dummies the joint test of significance is reported.

** significant at the 5 percent level or more; * significant at the 10 percent level.

^a χ^2 test statistic of LR specification test. Model 1 is the benchmark.

Table 3 presents the results of the same panel estimations reported in Table 2 enriched with the dummies modelling country-specific heterogeneity. Our main findings for both the control variables and the follow-the-leader behavior remain valid. In fact, when controlling for information spillovers from investment flowing into other countries, the previously discussed conditions on the parameter for L_{ijt} and its interaction with uncertainty are also satisfied in these model specifications, albeit with a smaller degree of significance. Once the channel of Bayesian learning from other countries is duly considered in explaining follow-up investments (Models ‘b’ in Table 3), uncertainty per se in a given country j ($INDUNC_{ij}$) also turns out to be a

significant, negative determinant of FDI. The reason is that uncertainty has both a positive (through Bayesian learning) and a negative effect on investment.²⁹ Without controlling for the Bayesian learning channel (the interaction term $L_{ijt} * INDUNC_{ij}$), the two effects cancel out. Once we control for the latter interaction, instead, a negative significant sign appears in the estimates of uncertainty alone.

²⁹ In general three channels through which uncertainty can negatively affect investment can be identified: the option theory of irreversible investment, financing constraints and risk-aversion. For a short discussion of this literature see Ghosal and Loungani (2000).

As far as the cross-country investment dummies are concerned, we have some evidence of FDI from rival firms acting as strategic substitutes rather than complements. In two of the model specifications where the Bayesian learning channel is not considered (Models 5a' and 6a'), it can be seen that a higher number of previous investments in other countries $s \neq j$ (dummies *CH* and *CV H*) affects negatively and significantly new FDI undertakings in country j , thus suggesting a tendency toward industry concentration/geographical specialization by MNEs. Finally, the positive and significant interaction between *CL* and *INDUNC* in Models 6 and 7 of Table 3 provides some evidence of learning from rivals that established a first or second investment in a given industry in a country different from the host one. Rivals' reaction to investments in other CEECs however is not a robust finding.

6. Conclusion

Paying tribute to the original intuition by Knickerbocker (1973), we have been able to derive a general theoretical model of rival MNEs' reaction based on Bayesian learning from first mover investors, encompassing the main implications of alternative models developed by the literature, and testing the resulting propositions on the actual behavior of rival MNEs. We find evidence for our theoretical propositions, showing that, alongside more traditional determinants of FDI, follow-the-leader behavior driven by Bayesian learning by rival firms plays a significant role in driving MNEs' decisions to invest abroad. More specifically, the results indicate that firms learn about revenue rather than cost. This result is robust with respect to different model specifications which control for both industry and country heterogeneity.

Two future lines of research are evident to us. First of all, the long studied issue of FDI seen as strategic substitutes or complements might be worth another closer look. In our paper, FDI are strategic substitutes only after a certain threshold in the number of rivals is reached, and only with respect to FDI undertaken in countries different than the one in which the considered investment is taking place. When previous investments in the same country are considered, instead, our study suggests that agglomeration effects and Bayesian learning make FDI decisions strategic complements. As a result, the effects of FDI substitution or complementarity seem to be a function of the geographical space in which rivals are considered.

Second, it is obvious that the use of categorical dummies for modelling previous investments suffers from some potential shortcomings, threshold effects being the

most evident ones. Therefore, the results of this paper should be validated as soon as the new econometric techniques on dynamic discrete panel data models move from the frontier of theoretical research to more routinely methodological tools.

Acknowledgements

We wish to thank the Editor (Ali Hortacsu), an anonymous referee and seminar participants at the 5th annual meeting of the European Trade Study Group in Madrid, Erasmus University in Rotterdam and the 31st EARIE meeting in Berlin for helpful comments and suggestions. The usual disclaimer applies. This paper has been written as part of a CEPR research network on 'The Economic Geography of Europe: Measurement, Testing and Policy Simulations', funded by the EC under the Research Training Network program (Contract No: HPRN-CT-2000-00069).

References

- Banerjee, A.V., 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107, 797–817.
- Bikhchandani, S., Hirshleifer, D., Welch, I., 1998. Learning from the behavior of others: conformity, fads, and informational cascades. *Journal of Economic Perspectives* 12, 151–170.
- Brunetti, A., Weder, B., 1998. Investment and institutional uncertainty: a comparative study of different uncertainty measures. *Weltwirtschaftliches Archiv*, vol. 134, pp. 513–533.
- Buch, C., Kleinert, J., Lipponer, A., Toubal, F., 2005. Determinants and effects of foreign direct investment: evidence from German firm-level data. *Economic Policy* 20, 52–110.
- Cameron, A.C., Trivedi, P.K., 1998. *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, UK.
- Chang, S.J., 1995. International expansion strategy of Japanese firms: capability building through sequential entry. *Academy of Management Journal* 38, 383–407.
- Crozet, M., Mayer, T., Mucchielli, J.-L., 2004. How do firms agglomerate? A study of FDI in France. *Regional Science and Urban Economics* 34, 27–54.
- Cukierman, A., 1980. The effects of uncertainty on investment under risk neutrality with endogenous information. *Journal of Political Economy* 88, 462–480.
- Davies, S., Lyons, B., 1996. *Industrial Organization in the European Union*. Clarendon Press, London.
- Flowers, E.B., 1976. Oligopolistic reactions in eastern and Canadian direct investment in the United States. *Journal of International Business Studies* 43–55.
- Ghosal, V., Loungani, P., 2000. The differential impact of uncertainty on investment in small and large businesses. *Review of Economics and Statistics* 82, 338–343.
- Grossman, S., Kihlstrom, R.E., Mirman, L.J., 1977. A Bayesian approach to the production of information and learning by doing. *Review of Economic Studies* 44, 533–557.
- Hausman, J., Hall, B., Griliches, Z., 1984. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica* 52, 909–938.

- Head, K., Mayer, T., 2004. Market potential and the location of Japanese investment in the European Union. *Review of Economics and Statistics* 86, 959–972.
- Head, K., Ries, J., Swenson, D., 1995. Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States. *Journal of International Economics* 38, 223–247.
- Head, K., Mayer, T., Ries, J., 2002. Revisiting oligopolistic reaction: are decisions on foreign direct investment strategic complements? *Journal of Economics and Management Strategy* 11, 453–472.
- Helpman, E., Melitz, M., Yeaple, S.R., 2004. Export versus FDI with heterogeneous firms. *American Economic Review* 94, 300–316.
- Hennart, J.-F., Park, Y.-R., 1994. Location, governance, and strategic determinants of Japanese manufacturing investments in the United States. *Strategic Management Journal* 15, 419–436.
- Hoff, K., 1997. Bayesian learning in a infant industry model. *Journal of International Economics* 43, 409–436.
- Honoré, B.E., Kyriazidou, E., 2000. Panel data discrete choice models with lagged dependent variables. *Econometrica* 68, 839–874.
- Ito, K., Rose, E., 2002. Foreign direct investment location strategies in the tire industry. *Journal of International Business Studies* 33, 593–602.
- Knickerbocker, F.T., 1973. *Oligopolistic Reaction and Multinational Enterprise*. Harvard University Press, Cambridge, MA.
- Leahy, D., Pavelin, S., 2003. Follow-my-leader FDI and tacit collusion. *International Journal of Industrial Organization* 21, 439–453.
- Pavitt, K., 1984. Sectoral patterns of technological change: towards a taxonomy and a theory. *Research Policy* 13, 343–373.
- Pennings, E., Altomonte, C., 2006. The hazard rate of foreign direct investment: a structural estimation of a real option model. *Oxford Bulletin of Economics and Statistics* 68, 569–593.
- Shapiro, C., 1986. Exchange of cost information in oligopoly. *Review of Economic Studies* 53, 433–446.
- Vettas, N., 1998. Demand and supply in new markets: diffusion with bilateral learning. *RAND Journal of Economics* 29, 215–233.
- Vettas, N., 2000. Investment dynamics in markets with endogenous demand. *Journal of Industrial Economics* 48, 189–203.
- Yu, C.M., Ito, K., 1988. Oligopolistic reaction and foreign direct investment: the case of the U.S. tire and textiles industries. *Journal of International Business Studies* 449–460.