

Firm heterogeneity in modelling foreign direct investment location decisions

Simona Rasciute, Eric Pentecost and Ben Ferrett¹

Division of Economics
School of Business and Economics
Loughborough University
Loughborough
LE11 3TU
UK

¹ The authors are extremely grateful to Professor William Greene for detailed discussions of the mixed logit estimation technique used in this paper.

Firm heterogeneity in modelling foreign direct investment location decisions

Abstract

This paper investigates the probability of the FDI location decisions of multinational enterprises using a mixed logit panel data model, which is the most flexible discrete choice model. We employ a three-level dataset, which includes over 1100 foreign direct investment (FDI) location decisions into 13 alternative Central and Eastern European countries (CEECs) over an eleven-year period. Our empirical results on the effect of host-country, industry and firm characteristics on the probability of undertaking FDI in a particular location are significant, and consistent with the predictions of our theoretical model.

JEL classification Nos: C25, F23

Keywords: Mixed logit model, firm heterogeneity, random parameters, foreign direct investment, multi-level data, Halton draws

1. Introduction

This paper investigates, both theoretically and empirically, the effect of investing firm heterogeneity on firms' location choices for their foreign direct investment (FDI) projects. In a large sample of foreign direct investments into the Central and Eastern European Countries (CEECs) over the past decade, we find strong support for our theoretical model. We show how the characteristics of the investing firm, the host country and the industry concerned are all statistically important for the investment location decision.

The traditional empirical literature on FDI has been primarily concerned with the determinants of aggregate FDI flows, using either macroeconomic time series or panel data sets and methods. Important contributions have examined the influence of variables such as market size and relative factor prices (Goldsbrough 1979; Culem 1988; Carstensen & Toubai 2004; and Clausing & Dorobantu 2005). A related strand of literature has examined the positive effects of FDI on economic growth (Barrell & Pain 1997). A relatively recent development in the empirical FDI literature is a microeconomic focus on the determinants of the investment location choices of multinational enterprises (MNEs). This empirical work is usually motivated by the so-called 'New Economic Geography' framework (Krugman, 1991), which explains the geographical agglomeration of industries and workers. Empirically, the determinants of the choice of FDI location have been modelled using either aggregate data (Barrell & Pain, 1999) or firm-level data sets and a discrete choice empirical methodology (Head et al., 1999; Guimaraes et al., 2000 and Kim et al., 2003).

This literature, however, is subject to two principal limitations. First, it has applied basic discrete choice models, namely the Multinomial logit (MNL) and the Nested logit (NL) models, to investigate where firms choose to locate their capital². The former is subject to restrictive assumptions regarding the substitution patterns across investment location alternatives and in the absence of investing firm heterogeneity, the latter only partially relaxes the independence from irrelevant alternatives assumption in order to accommodate some substitution across alternatives³.

A second limitation is that the existing literature on the location of FDI usually uses country- and industry-specific data but does not incorporate investing firm characteristics. However, greater estimation efficiency can be achieved by using a multi-level data set, where country, industry and firm-level factors simultaneously determine the firm-level investment location decision. Perhaps the first theoretical rationale for incorporating investing-firm heterogeneity was provided by Nocke & Yeaple (2007), who show that a firm's decision to select FDI rather than exports depends critically on its productivity level. Our theoretical model in the next section builds on this approach to show how firm, industry, and host-country-specific variables drive the FDI location decision.

² The exception to this is Basile et al. (2008), who apply the Mixed logit model in the context of EU cohesion policy, however, the interpretation of results and policy implications are based on the estimated coefficients. Neither the sign nor the magnitude of the coefficients is very informative and further estimation of elasticities and marginal effects is needed.

³ The independence from irrelevant alternatives (IIA) assumption in discrete choice models differs somewhat from the IIA assumption in standard microeconomics: here, it means that introducing an extra choice affects all probabilities equi-proportionately. This can be understood as requiring that all potential FDI locations are equally close substitutes – a very strong assumption.

This paper, therefore, extends the recent literature by making two principal contributions. First, it explicitly allows for the effects of investing firm heterogeneity on the investment location decision by using the Mixed logit (ML) panel data model to investigate investment location choices by MNEs. Second, it makes use of a novel, multi-level data set, allowing firm, industry (or sector) and country effects to simultaneously determine firm-level FDI location decisions. This data set has been constructed partly from the Zephyr data base of the Bureau van Dijk, which reports some source-firm characteristics. The sample used in this paper covers 1,108 independent brownfield investment location choices of 693 firms in the EU(15), Norway, Switzerland, Russia, Japan and the USA into 13 Central and Eastern European Countries (CEECs) – the 12 recent EU member states excluding Cyprus and Malta, but including Croatia, Russia and Ukraine – over an eleven year period from 1997 to 2007. By applying the ML model to multi-level data, this paper seeks to investigate how important the investing firm’s characteristics are in the choice of investment location and if alternative locations tend to attract different types of FDI, i.e. is the effect of location factors such as market size, transport cost and labour costs, among others, universal across different investors like it has been mainly found in the existing literature or it varies depending on investing firms characteristics? The results show that firms investing in different sectors and firms of different characteristics benefit from country-level factors to different degrees, which offers strong support for our theoretical model.

The rest of the paper is set out as follows. The next section presents the theoretical model. The section after explains the ML model. The forth section

discusses the dataset and the construction of the variables and the fifth section presents the econometric results. Finally, the last section concludes.

2. Theoretical model

In this section we develop a simple but novel theoretical model to motivate the empirical analysis that follows. We examine a representative firm's FDI location decision. We simplify the analysis by taking the economic environment facing the firm as given – so, for example, we do not incorporate the reactions of rival firms into the analysis. This simplification is extremely common (e.g. it is a feature of all monopolistically competitive models). Moreover, in the classic models of merger in oligopoly, the reactions of rival firms to a merger generate results that are both intuitively unappealing and highly sensitive to specific assumptions on market conduct. Thus, for the purposes of our simple illustrative model, it is easiest to ignore the reactions of rivals.

We assume that the firm has decided to purchase a plant in a CEEC in order to serve the entire European product market and that it is now choosing where to invest. Thus, the firm is following a decision tree similar to that set out in Devereux and Griffith (1998), where the export versus FDI decision has already been settled in favour of FDI and the firm is now choosing the location of its FDI.

In any given FDI location, the profits of the multinational enterprise (MNE) are given by

$$\Pi = (1 - t)[(1 - b)P - K] \quad (1)$$

where $t, b \in [0,1]$ denote the profit tax rate and the level of risk respectively.⁴ More precisely, we are thinking of b as the probability that, following purchase, production in the plant will turn out to be impossible – perhaps because the MNE has been unable to obtain the relevant permits from government officials.

The plant’s per-period operating profits, which we will endogenise below, are denoted by P ; and per-period capital costs (i.e. the interest on the sum borrowed to purchase and refit the plant) are denoted by K . Along with t and b , we take K as exogenous – perhaps because, in the absence of acquisition, the target firm would not compete with the MNE.

We assume that operating profits P are given by

$$P = s\pi + (1 - s)(1 - d)\pi = \pi[1 - d(1 - s)] \quad (2)$$

where π is the level of profits (exclusive of shipping costs) per consumer; $s \in [0,1]$ is the share of European consumers that are located locally in the FDI host country; and $d \in [0,1]$ is an increasing measure of “distance”, which can be thought of as encompassing the effects of host-country EU membership and a shared border between the source and host countries.

The expression for P thus normalises the total number of European consumers to 1. It also simplifies the geography of Europe by assuming that, outside the host country, there exists an integrated European product market. Shipping costs are assumed to be *ad valorem* or “iceberg” in that transport consumes a given proportion of the product.

⁴ Because the interpretation of our theoretical model is straightforward, we omit firm, industry and country subscripts to avoid unnecessary clutter.

Our model is completed by a third equation that determines per-consumer profits, π :

$$\pi = \theta(U)[p - Wl] \quad (3)$$

For simplicity, we assume that each consumer buys 1 unit of the good that is produced following FDI. The product price p is therefore equal to the reservation price. W is the wage rate (we assume that labour is the only variable input), and l is the labour-output ratio, an indication of the labour intensity of production. Therefore, Wl is the unit production cost.

Finally, $\theta(U) \in [0,1]$ is the probability that the firm is able to recruit enough appropriately-skilled local workers to operate its purchased plant. We assume that this probability depends on the host country's unemployment rate, U , although we do not model this dependence explicitly. Indeed, the functional relationship between U and θ might be quite complex.

For example, in “traditional” industries, which tend to be relatively intensive in their use of unskilled labour, we might expect to observe $\theta'(U) > 0$ – that is, higher host-country unemployment makes recruiting workers easier. By contrast, in “non-traditional” industries, such as those based around science or services, skilled labour is important, and we might observe $\theta'(U) < 0$ because unemployment leads to a loss of skills and thus reduces the proportion of skilled workers in the total labour force. To capture these complex relationships, we could write

$$\theta = \theta(U, IND)$$

where $IND \in \{TRAD, NT\}$ is a dummy that indicates whether the industry concerned is traditional or non-traditional. Thus, the preceding discussion can be summed up as

$$\theta_U(U, TRAD) > 0 > \theta_U(U, NT)$$

Equations (1), (2) and (3) together constitute our theoretical model. It would be straightforward to embed our model of the MNE within a market context (e.g. with industry rivals and a downward-sloping demand curve), although we would not expect that to alter its key properties. Therefore, on grounds of simplicity, we prefer our representative-firm framework.

By repeated substitution, we can write the firm's profits from FDI in a given location in terms of exogenous variables:

$$\Pi = (1 - t)\{(1 - b) \cdot \theta(U) \cdot [p - Wl] \cdot [1 - d(1 - s)] - K\} \quad (4)$$

We can now use differentiation on (4) to analyse the determinants of FDI profitability, Π . We assume that the FDI project generates strictly positive pre-tax profits (i.e. $\{\cdot\} > 0$ in (4)), which in turn requires $p > Wl$.

We obtain the following results:

- An increase in a country's profit tax rate makes inward FDI less likely ($\partial\Pi/\partial t < 0$).
- An increase in a country's risk level makes inward FDI less likely ($\partial\Pi/\partial b < 0$).
- A larger host-country market makes inward FDI more likely ($\partial P/\partial s = \pi d > 0$) – especially by MNEs that produce high-value products

($\partial^2 P / \partial s \partial \pi = d > 0$). The latter result is consistent with the finding of Helpman *et al* (2004) that FDI is chosen by high-productivity firms.

- A greater distance from the core European product market makes inward FDI less likely ($\partial P / \partial d = -\pi(1 - s) < 0$) – especially by MNEs that produce high-value products ($\partial^2 P / \partial d \partial \pi = -(1 - s) < 0$). The latter result is due to the assumption of *ad valorem* trade costs, which imply that shipping costs more in total for high-value firms. If trade costs were instead specific, then we would expect this result to be reversed because firms with high margins are more able to shoulder specific trade costs.
- We hypothesized that the effect of higher host-country unemployment on inward FDI flows is likely to be ambiguous ($\partial \pi / \partial U \geq 0$) and dependent on the type of industry under consideration.
- Higher host-country wages discourage inward FDI ($\partial \pi / \partial W = -\theta l < 0$) – especially in “traditional” sectors that are relatively labour intensive ($\partial^2 \pi / \partial W \partial l = -\theta < 0$).

In the remainder of the paper, we investigate empirically these predictions from our theoretical model. We will concentrate on examining the factors that determine the profitability of investing in a particular location. We will assume that anything that raises profits makes FDI in that location “more likely.” Although our theoretical model is deterministic, it can easily be interpreted in such a probabilistic fashion (and thus related to our empirical methodology) by assuming that the firm’s profits from investment in any particular location contain an additive random element, which for simplicity we do not model explicitly. Thus, the

profit expressions may be thought of as giving the *expected* profits from FDI in a given location.

3. Econometric estimation

The mixed logit (ML)⁵ model is chosen because it is the most flexible discrete choice model that approximates any random profit model (McFadden & Train, 2000). In contrast to the MNL (multinomial logit) model, for example, the ML model allows for investing firm heterogeneity and unrestricted substitution patterns across investment location alternatives. In the most general form, the ML probabilities (P_{ic}) are the integrals of standard logit probabilities (L_{ic}) over a density of unobserved random parameters (Train, 2003), such that

$$P_{ic} = \int L_{ic}(\beta_i) f(\beta_i) d\beta_i \quad (5)$$

where $f(\beta_i)$ is the random parameter density function, which is specified to be continuous. In the mixed logit model, $\beta_i = \beta + w_i$ where w_i is uncorrelated with the attributes of the host country characteristics. This is essentially a random effects approach, but it is more general than the familiar "random effects" model as in the familiar case only the constant term is random. The panel data "aspect" used in this model is that w_i is time invariant. This aspect works the same way as a random effects model in the more familiar case.

⁵ In the literature also referred to as "random parameter logit", "mixed multinomial logit", "kernel logit" or "hybrid logit".

The mixed logit probability is a weighted average of the logit formula evaluated at different values of β_i , with the weights given by the density $f(\beta_i)$. The logit probability evaluated at parameters β_i is expressed as

$$L_{ic}(\beta_i) = \prod_{t=1}^T \left(\frac{e^{\beta'_i x_{ict}}}{\sum_g e^{\beta'_i x_{igt}}} \right) \quad (6)$$

where x_{ict} is a vector of observed investment location-specific variables (i , c , and t indicate individual firm, alternative location⁶ and time period respectively) and β_i are individual firm-specific parameters.

Since the integral in (5) cannot be calculated analytically, it has to be approximated through simulation by maximising the simulated log-likelihood function. A value of β is drawn from the distribution $f(\beta | \psi)$ for a given value of ψ , where ψ defines the parameters of the distribution (i.e. mean and scale parameter, and in the estimated model in the next section consists of 7 means and 7 scale parameters) and labelled β^r with the subscript $r=1$ referring to the first draw. The logit formula $L_{ci}(\beta^r)$ is calculated with this draw. Finally, the two steps are repeated many times and the results are averaged, which gives the simulated probability:

$$\tilde{P}_{ic} = \frac{1}{R} \sum_{r=1}^R L_{ic}(\beta^r) \quad (7)$$

where R is a number of draws. \tilde{P}_{ci} is an unbiased estimator of P_{ic} , whose variance decreases as R increases. This estimator is strictly positive so below, $\ln P_{ci}$ is defined. Furthermore, \tilde{P}_{ci} sums to one over all alternatives which is helpful when

⁶ Firms in the EU(15) individual countries, Norway, Switzerland, Russia, Japan and the USA choose to locate their investment in 13 Central and Eastern European Countries (CEECs) – the 12 recent EU member states excluding Cyprus and Malta, but including Croatia, Russia and Ukraine.

interpreting the results. The simulated probabilities are substituted into the log-likelihood function to give the simulated log likelihood:

$$SLL = \sum_{c=1}^C \sum_{i=1}^I \delta_{ic} \ln \check{P}_{ic} \quad (8)$$

where $\delta_{ic}=1$ if i chose c and zero otherwise. Note, however, that the estimator of $\ln \check{P}_{ic}$ is not unbiased because the expectation of $\ln P$ is not equal to the logarithm of the expected value of P , but it is consistent in R (Train, 2003). The maximum simulated likelihood estimator (MSLE) is the value of ψ that maximizes SLL .

4. The data set and variable specification

Table 1 gives a summary of variable definitions and sources. There are 1,108 firm-level data observations on brownfield FDI flows from firms of 20 market economies (the EU15 countries, USA, Japan, Russia, Norway and Switzerland) into firms in 13 transition economies (the 12 new EU member states (except for Malta and Cyprus) plus Croatia, Russia and Ukraine) from 1997 to 2007. Most of the empirical literature on FDI focuses on Greenfield investment, excluding other entry modes such as mergers and acquisitions (M&A), joint ventures and institutional buy-outs (Brownfield investment), which are very important for the CEECs (Head and Ries, 2008). For example, in 2005 the share of cross-border M&As in FDI was about 96 percent in Czech Republic, 84 percent in Estonia, 82 percent in Ukraine and 68 percent in Bulgaria (UNCTAD⁷ statistics)⁸. In order to test for the location

⁷ United Nations Conference on Trade and Development

⁸ It is difficult to estimate precisely what share of FDI flows is accounted for by cross-border mergers and acquisitions (M&A) because the values of cross-border M&As cannot be directly compared with FDI

determinants of essentially Brownfield investment, the MNEs in the sample enter a foreign market via M&A, joint ventures or institutional buy-outs.

Of all 13 host CEECs in the sample, Poland has been chosen the most times by MNEs as an FDI location (about 21% of all observations), followed by Russia with about 17% of foreign investment location choices. Slovenia and Latvia, on the other hand, have received the smallest share of foreign capital allocations (2% and 3% respectively). The two major source countries for investment in CEECs in the sample are Finland and the UK with shares of approximately 12% and 11% respectively. Japan and Ireland were at the other end of the scale as source countries with shares of about 1% each.

The host country's characteristics may have different effects on firms investing in different sectors. For example, our theoretical model showed that higher host-country wages discourage inward FDI most in "traditional" sectors that are relatively labour intensive. Our sample includes four types of industry or sector:

- Scale-intensive sectors include typical oligopolistic, large firm industries, with high capital intensity, extensive economies of scale and high technical and managerial complexity; for example, automobiles, aircrafts, chemicals, petrol and coal products, shipbuilding, industrial chemicals, drugs and medicines,

flows registered in the balance of payments. More specifically, first, the values of cross-border M&As include funds raised in local and international financial markets, while FDI data do not. Second, FDI data are reported on a net basis using balance-of-payment concept, while data on cross-border M&A purchases report only the total value of purchases abroad and do not subtract the amounts received from the sales of foreign affiliates. Finally, payments of cross-border M&As are not necessarily made in a single year, but can be phased over several years. As a result, calculating the value of cross-border as a percentage of FDI inflows in a given year may not be precise (UNCTAD, 2000).

petrol refineries, non-ferrous metals and railroad equipment (Midelfart-Knarvik *et al.*, 2000).

- Science-based sectors are characterised by innovative activities directly linked to high R&D expenditures; for example, fine chemicals, electronic components, telecommunications and aerospace.
- Traditional (supplier-dominated) sectors include such industries as textiles, clothing, furniture, leather and shoes, ceramics and the simplest metal products.
- Finally, banking, insurance and retail are examples of service sectors.

The country-specific determinants of FDI location decisions into the CEECs can be loosely divided into the traditional determinants and the transition-specific determinants. The transition-specific determinants are proxied by the risk associated with each host country; this is the variable b in our theoretical model. The Transparency International Corruption Perception Index (TICP) is used as a measure of the extent of corrupt practices in the host country. It ranks countries in terms of the degree to which corruption is perceived to exist among public officials and politicians and it varies from 1 (high corruption) to 10 (no corruption).

The traditional determinants of FDI are the market size of the host country, the cost of capital in the host country, distance from the source country and tax rates in the host country. As Table 1 shows, market size is simply the real GDP of the country and the cost of capital is measured as the real discount (interest) rate.

To capture the effect of host-country corporate income taxation, we use the effective corporate income tax rate, which is calculated by dividing profit taxes paid by corporations and other enterprises by a host country's GDP. Although other

studies (e.g. Bellak & Leibrecht, 2005; Carstensen & Toubal, 2004; Clausing & Dorobantu, 2005; Wei, 2000) tend to use the statutory corporate income tax rate, we consider the effective rate to be superior because the statutory rate is only one of the determinants of the total tax burden. Our approach allows comparisons of different national tax systems, taking into account such important aspects as untaxed reserves, tax enforcement and the treatment of losses.

Finally, three other country-specific factors are included in the empirical model: the national rate of unemployment and two dummy variables, one for European Union membership and another for a common border between the source and host countries. Both dummy variables are expected to have a positive effect on the probability of choosing certain CEECs as investment locations because common borders reduce distance, and EU membership implies free trade and the adoption of a Western business and legal environment.

Unemployment is included as an indicator of work-force availability. As we discussed in the theoretical section, we expect the relationship between host-country unemployment and inward FDI to be complex, and – plausibly – dependent on the industry concerned.

The characteristics of individual investing firms are also important to the FDI decision. The firm-level variables include the turnover of the investing firm as a proxy for its size and earnings before interest and tax as a proxy for its profitability.

5. Estimation and results

There are no rules that help choose appropriate distributions for random parameters in the specification of the ML model, apart from the requirement that the mean and standard deviation of the random parameter to be statistically significant. As a result, the analysis starts by treating each coefficient of the country-level variables separately as random by imposing various distributions, which gives an indication of which distribution should be used for random parameters when they are combined in the final specification⁹ (Table 2). The results of the final specification with the best model fit when random parameters with the most appropriate distributions are combined are presented in Table 3. A triangular distribution is imposed on the variables for market size in the host country, the wage variable, the distance variable and the unemployment variable¹⁰. A restricted uniform distribution is imposed on the two dummy variables: the dummy variable for common border and the dummy variable for EU membership¹¹. The means of the uniform distribution for the dummy variables are restricted to be equal to their variances; as a result, the sign of the estimated random parameters of the two dummy variables will be the same for all the investing firms. As the standard deviations of various distributions are not statistically significant for the cost of capital variable, it is treated as fixed and it is assumed that all information is captured by the mean. Initially 100 Halton

⁹ However, when all random parameters are combined in the final specification, the imposed distributions may not be the same as when each parameter is treated separately.

¹⁰ It is preferred to impose the same distribution for all continuous explanatory country-level variables. Uniform distribution is typically imposed on dummy variables.

¹¹ The country risk variable proved insignificant and correlated with the EU membership dummy and so was dropped from the remainder of the empirical analysis.

intelligent draws are used to estimate the model, while increasing the number to 1000 for the final model specification.

The ML is not only able to determine the existence of heterogeneity around the mean parameter, through the estimation of the standard deviation parameter, but it can also indicate the source of the heterogeneity through the interaction between the random parameters and other attributes (moderator variables). For example, observed heterogeneity in some country-level determinants of the investment location choice can be due to differences in industry and investing firms' characteristics. Table 3 shows that the following interaction terms are statistically significant: the interaction terms between the dummy variable for traditional sectors and unemployment rate in the host country; the interaction term between the dummy variable for traditional sectors and the wage rate in the host country; the interaction term between investing firms profitability and distance between investing and investment receiving countries; and finally, the interaction term between investing firm's size and host country's market size.^{12,13}

Despite the widespread use of interaction terms in the discrete choice methodology, the majority of applied researchers misinterpret the coefficients of interaction terms (Ai and Norton, 2003). Unlike in linear models, the interaction effect in non-linear models is a function of not only the coefficient for the

¹² The potential profit function is then given by

$$\pi_{icst} = \beta_0 + \beta_1 Wage_{sct} + \beta_2 D_{ct} + \beta_3 \tau_{cd} + \beta_4 u_{ct} + \beta_5 CBD_{cd} + \beta_6 EU_{ct} + \beta_7 K_{ct} + \beta_8 Profit_{it} \times \tau_{cd} + \beta_9 Size_{it} \times D_{ct} + \beta_{10} Tradit_{st} \times u_{ct} + \beta_{11} Tradit_{st} \times Wage_{sct} + \varepsilon_{icst}$$

where the notation is explained in Table 1.

¹³ The results of the Multinomial Logit (MNL) model estimation are presented in the last two columns of Table 3 for comparison. The estimated coefficients look stable across the two models. The LR test statistic of 292.0014 is greater than the Chi-square critical value of 12.5916, as a result, we are able to reject the hypothesis that the ML model does not statistically improve the MNL model.

interaction, but also the coefficients for each interacted variable and the values of all the variables in the model (Greene, 2008). Therefore, the sign of the interaction coefficient may not indicate the direction of the interaction effect, as the interaction effect may have different signs for different values of covariate. Furthermore, the interpretation of a separately included variable in the model changes if it is also a part of an interaction term (Jaccard, 2001). It does not represent a “main effect” but a conditional effect instead: that is, the effect of the variable when the values of the moderator variable (the other interacted variable) are zero. For example, the wage variable is not only included in the model separately, but also interacted with the dummy variable for traditional sectors as a result the wage variable represents the effect of wage costs on the probability of choosing a particular country as an investment location, when the foreign firm chooses to invest in non-traditional sectors¹⁴ (*Science-based, Service and Scale-intensive sectors*).

As neither the sign nor the magnitude of the interaction terms or separately included variables in the model are informative, the elasticities and marginal effects have to be estimated for continuous and dummy variables respectively as shown in Table 4.¹⁵ When elasticities are estimated for the interaction term where one of the interacted variables is an industry dummy, the sample is restricted to include investment in traditional and non-traditional sectors separately in order to

¹⁴ As the dummy variable for traditional sectors is equal to zero.

¹⁵ The elasticities for separately included variables in the model like market size and distance variables do not have much explanatory value, as they indicate the effect when moderator variables (investing firm’s size and profitability respectively) are equal to zero. The tax variable and the risk variable do not appear statistically significant.

reveal elasticities of the wage and unemployment variables in the two groups of sectors.

The results presented in Table 4 show:

- negative estimated elasticities for the wage variable in non-traditional sectors (except for Slovenia), indicating that the higher are wage costs in a host country (especially Slovakia and Croatia), the less likely that country is to be chosen as a location for foreign capital. The elasticities of the wage variable in the traditional sectors are negative and larger in absolute values, especially for Croatia and Lithuania, showing that firms in traditional sectors are more sensitive to higher wages than firms in non-traditional sectors. This is in line with our theoretical model. Firms investing in non-traditional sectors, for example, science-based industries, are less likely to be discouraged to pay higher wages (or even encouraged in case of Slovenia) for skilled workers, which reflects a skill premium.
- The unemployment variable is not only included in the model separately, but also interacted with the dummy variables for traditional sectors. Negative elasticities for the unemployment variable in non-traditional sectors indicate that the higher the unemployment rate in the host country, the less likely it is that the country will be chosen by foreign firms in non-traditional sectors (the effect is the strongest for Czech Republic and Poland). Negative, but smaller in absolute value (Czech Republic, Estonia, Romania and Hungary), or even positive (the rest of the countries) elasticities for the unemployment variable in traditional sectors indicate that higher host-country unemployment has a smaller negative or even a positive effect on the probability of investment for firms in traditional sectors. This is also in line with the discussion in the

theoretical section. Higher unemployment rate may be an indication of availability of unskilled labour, as unemployed workers loose their skills through time.

- Positive and statistically significant estimated marginal effects for the EU dummy variable, in Table 4, indicate that countries which became members of the EU by January 2007 are more likely to be chosen as FDI locations. The common border dummy also shows a positive marginal effect, as expected.
- Interpreting the positive and significant coefficient on the real interest rate in the host country is difficult, given that the real interest rate is conventionally regarded as the cost of capital. However, we speculate as follows. If internal sources (e.g. retained earnings) are key to financing corporate investment, as is common for large firms, and if the real interest rate is viewed as a proxy for the rate of profits in the host country, then our observed positive relationship can be rationalised.

When two continuous variables are interacted, the interpretation of the interaction terms is much more complicated, as the moderator variables (the size and profitability of the investing firms) take more than two values. Simulation¹⁶ is used to investigate the effect of the gradual increase in the moderator variable on the shares of foreign capital allocations across different countries and a number of foreign firms investing in those countries. The top half of Table 5 shows the effect of 1, 10, 50 and 100 % increase in the investing firm's *size* on the market shares of

¹⁶ The simulation capability allows testing how changes in certain variables impact upon the choice probabilities for each of the investment receiving countries. In this paper simulation is used as a tool to reveal the direction of the effect between two interacted continuous variables. If the increase (decrease) in moderator variables results an increase (decrease) in the choice probabilities of the interacted variable, we can conclude that the two interacted variables are positively related.

foreign capital allocations across the countries (first column) as well as the change in the number of investing firms (second column) in each country in the sample. The bottom half of Table 5 shows the effect of the equivalent increase in the investing firm's *profitability* on the two variables. The results show that a gradual increase in the firm profitability and size variables result in positive and increasing changes in foreign capital allocation shares and in the number of firms investing in a particular country. For example:

- The results for interaction term between investing firm's profitability and distance indicate that more profitable investing firms are less likely to be discouraged from investing in more remote countries as compared to less profitable firms. As discussed in the theoretical section, this result might arise because high-margin firms find it easier to shoulder the higher specific trade costs that are associated with greater distance.
- The results for the interaction between investing firm's size and host country's market size indicate that the larger the host country is, the more likely it is to be chosen by an investing firm to locate its capital. Moreover, this effect is stronger for larger investing firms. Both of these results are consistent with our theoretical model.

Statistically significant interaction terms between country-level variables and industry dummies together with the firm-level variables confirm the existence of heterogeneity revealed by statistically significant standard deviations of the parameters of certain country-level variables.

6. Conclusions

This paper applies the Mixed logit (ML) model, which is probably the most flexible discrete choice model, to investigate investment location choices by MNEs. Our empirical hypotheses are derived from a simple yet novel representative-firm model. We also make use of a multi-level data set – allowing firm, industry (or sector) and country effects to simultaneously determine firm-level FDI location decisions. The highly significant empirical results support the presence of source-firm heterogeneity in the investment location decisions, which is not only revealed by statistically significant interaction terms, but also by statistically significant standard deviations of the random parameters. These results show that country-level factors can have different effects on firms of different size and profitability and firms investing in different sectors.

For example, firms investing in traditional sectors are less likely to be discouraged by higher host-country unemployment but more likely to be discouraged by higher wage rates, as compared to MNEs in non-traditional sectors. The larger is the host country, the more likely it is to be chosen by foreign investors; and the effect is stronger for larger investing firms. On the other hand, more profitable firms are less likely to be discouraged from investing in more remote countries, as compared to less profitable firms.

Both our theoretical and empirical approaches to modelling the FDI decision allow for heterogeneity in firm, industry and host-country characteristics. We have shown that allowing for firm heterogeneity is important if robust estimates are to be found of the general determinants of the choice of investment location.

References

- Ai, C., Norton, E.C. 2003. Interaction terms in logit and probit models. *Economics Letters* 80, 123-129.
- Barrell, R., Pain, N., 1997. Foreign direct investment, technological change and economic growth within Europe. *Economic Journal* 107, 1770-86.
- Barrell, R., Pain, N., 1999. Domestic institutions, agglomerations and foreign direct investment in Europe. *European Economic Review* 43, 925-934.
- Basile R., Castellani, D., Zanfei, A., 2008. Location Choices of Multinational Firms in Europe: the Role of EU Cohesion Policy. *Journal of International Economics* 74, 328-340.
- Bellak, C., Leibrecht, M., 2005. Do low corporate income tax rates attract FDI? Evidence from Eight Central- and East European Countries. Research Paper Series, Globalisation, Productivity and Technology, No.2005/43. The University of Nottingham. Leverhulme Centre for Research on Globalisation and Economic Policy
- Carstensen, K., Toubal, F., 2004. Foreign direct investment in Central and Eastern European countries: a dynamic panel analysis. *Journal of Comparative Economics* 32, 3-22.
- Clausing, K.A., Dorobantu, C.L., 2005. Re-entering Europe: Does European Union candidacy boost foreign direct investment? *Economics of Transition* 13, 77-103.
- Culem, C.G., 1988. The locational determinants of direct investments among industrialised countries. *European Economic Review* 32, 885-904.

- Devereux, M.P., Griffith, R., 1998. Taxes and the location of production: evidence from a panel of US multinationals. *Journal of Public Economics* 68, 335-367.
- Goldsborough, D., 1979. The role of foreign direct investment in the external adjustment process. *IMF Staff Papers* 26, 725-754.
- Greene, W.H., 2008. *Econometric Analysis*, sixth edition, Pearson-Prentice Hall. New York.
- Guimaraes, P., Figueiredo, O., Woodward, D., 2000. Agglomeration and the location of foreign direct investment in Portugal. *Journal of Urban Economics* 47, 115-135.
- Head, C.K., Ries, J.C., 2008. FDI as an outcome of the market for corporate control: Theory and evidence. *Journal of International Economics* 72, 2-20.
- Head, C.K., Ries, J.C., Swenson, D.L., 1999. Attracting foreign manufacturing: Investment promotion and agglomeration. *Regional Science and Urban Economics* 29, 197-218.
- Helpman, E., Melitz, M.J., Yeaple, S.R., 2004. Exports versus FDI with Heterogeneous Firms. *American Economic Review* 94(1), 300-316.
- Jaccard, J., 2001. *Interaction effects in logistic regression*, London, New Delhi: Sage Publications.
- Kim, S.H., Pickton, T.S., Gerking, S., 2003. Foreign Direct Investment: Agglomeration Economies and Returns to Promotion Expenditures. *The Review of Regional Studies* 33(1), 61-72.
- Krugman, P., 1991. Increasing Returns and Economic-Geography. *Journal of Political Economy* 99, 483-499.

- McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15, 447-470.
- Midelfart-Knarvik, K.H., Overman, H.G., Redding, S.J., Venables, A.J., 2000. The Location of European Industry. *European Economy - Economic Papers* 142, Commission of the EC, Directorate-General for Economic and Financial Affairs (DG ECFIN).
- Nocke, V., Yeaple, S. 2007. Cross-border mergers and acquisitions vs. greenfield foreign direct investment: The role of firm heterogeneity. *Journal of International Economics*, 72(2), 336-365.
- Train, K., 2003. *Discrete Choice Methods with Simulation*, Cambridge University Press. Cambridge, Mass.
- Wei, S.J., 2000. How taxing is corruption on international investors? *Review of Economics and Statistics* 82, 1-11.

TABLE 1

List of Variables, Definitions and Sources

<i>Variable</i>	<i>Definition</i>	<i>Source</i>
$Choice_c$	a CEEC, in which firm n chooses to locate its investment over the period of time from 1997 to 2007 (it gets the value of 1 if the country received investment and 0 otherwise)	Bureau van Dijk Zephyr database
τ_{cd}	distance between the capital cities of the source country d and the host country c in kilometres	http://www.indo.com/distance/
D_c	Real GDP of the host country c of the year investment took place	IFS
b_c	Corruption perception index of the host country c of the year investment took place	Transparency International
u_c	unemployment rate of country c (percentage per annum) of the year investment took place	IFS
T_c	effective corporate income tax rate in country c of the year investment took place	Calculated using data from IFS
K_c	the real discount (interest) rate	IFS
CBD_{cd}	a dummy variable that takes a value 1 if both source country d and host country c share a border, and 0 otherwise	constructed
EU_c	dummy variable that takes value 1 if country c joined EU before January 2007, and 0 otherwise	constructed
$Scale_s$	dummy variable that takes a value 1 if industry s is a scale-scale industry, and 0 otherwise	constructed
$Science_s$	dummy variable that takes a value 1 if industry s is a science-based industry, and 0 otherwise	constructed
$Tradit_s$	dummy variable that takes a value 1 if industry s is a traditional industry, and 0 otherwise	constructed
$Service_s$	dummy variable that takes a value 1 if industry s is a service sector, and 0 otherwise	constructed
$Wage_{sc}$	hourly real wage rates in the industry s in the country c of the year investment took place	International Labour Organisation
$Size_i$	turnover of the investing firm i in Euros of the year investment took place	Bureau van Dijk Zephyr database
$Profit_i$	earnings before interest and taxes of the investing firm i in Euros of the year investment took place	Bureau van Dijk Zephyr database

TABLE 2
The Imposition of Various Distributions for Country-level Random Variables

	Distance					Risk				
	Coef	z-stats	Coef	z-stats	Log-likelihood	Coef	z-stats	Coef	z-stats	Log-likelihood
Normal			Log-likelihood is flat			0.3906	{5.354}	0.5161	{6.012}	-2507.072
Log-normal			Log-likelihood is flat					Log-likelihood is flat		
Triangular	-1.1821	{-11.007}	3.0309	{7.569}	-2501.44	0.3964	{5.394}	1.2748	{6.218}	-2506.846
	Unemployment					Interest				
	Coef	z-stats	Coef	z-stats	Log-likelihood	Coef	z-stats	Coef	z-stats	Log-likelihood
Normal	-10.3786	{-5.634}	18.3814	{8.110}	-2498.291	0.19	{1.122}	2.7461	{0.727}	-2514.21
Log-normal	1.4548	{1.2147}	1.2147	{4.719}	-2508.554	-0.3434	{-0.077}	0.9028	{0.178}	-2514.339
Triangular	-10.5425	{-5.674}	44.5571	{8.379}	-2498.102	1.2036	{1.151}	6.7393	{0.955}	-2514.191
	Tax					Wage				
	Coef	z-stats	Coef	z-stats	Log-likelihood	Coef	z-stats	Coef	z-stats	Log-likelihood
Normal	-0.0766	{-1.503}	35.1249	{3.971}	-2511.83	-0.105	{-1.990}	0.0392	{1.451}	-2514.294
Log-normal	0.0975	{0.051}	1.7071	{1.552}	-2513.612	-2.9966	{-3.077}	1.9047	{2.696}	-2510.452
Triangular	-0.8978	{-0.347}	87.1765	{4.013}	-2511.847	-0.1071	{-1.990}	0.1191	{1.550}	-2514.26
	GDP									
	Coef	z-stats	Coef	z-stats	Log-likelihood					
Normal	0.497	{7.635}	0.9543	{6.857}	-2485.322					
Log-normal			Log-likelihood is flat							
Triangular	0.4956	{7.497}	2.286	{6.988}	-2484.991					

t-statistics in parenthesis

The estimated mean and standard deviation parameters may not be comparable across explanatory variables when different distributions are imposed. For example, when the log-normal distribution is imposed, the median, mean and standard deviation of the coefficients themselves are given by $exp(b)$, $exp(b + s^2/2)$ and $exp(2b + s^2) [exp(s^2) - 1]$, respectively, where b is the estimated mean and s is the estimated standard deviation.

In addition restricted triangular, dome, erlang, weibull and exponential distribution have been used and the results are available upon request.

TABLE 3
The Results of the Mixed Logit Model Estimation

Variables	Distribution	The Mixed Logit model				Multinomial Logit model	
		Mean		Stand. Dev.			
Wage	Triangular	-.47091***	{-5.50}	1.64107***	{6.54}	-.19718***	{-4.71}
GDP	Triangular	.52431***	{6.93}	2.63609***	{8.93}	.49133***	{13.96}
Distance	Triangular	-1.82280***	{-10.80}	3.90893***	{6.47}	-.92397***	{-11.89}
Unempl	Triangular	-3.68843**	{-2.14}	48.0144***	{7.98}	-2.05055*	{-1.89}
Border	Restricted Uniform	.62247***	{3.46}	.62247***	{3.46}	.56302***	{5.18}
EU	Restricted Uniform	1.02096***	{5.40}	1.02096***	{5.40}	.57576***	{5.61}
Interest	-	7.50841***	{5.61}	-	-	0.9408	{0.95}
Prof_Dist	-	.18935***	{2.59}	-	-	.10592***	{2.69}
Size_GDP	-	.17762***	{3.59}	-	-	.10446***	{4.43}
Tr_Unemp	-	8.62223***	{3.18}	-	-	7.87587***	{4.69}
Tr_Wage	-	-.54867***	{-4.41}	-	-	-.42856***	{-5.39}
Log-likelihood		-2381.7243				-2527.725	
Chi-squared		920.47918				121.8558	
Pseudo R-sqr		0.1619				0.023536527	

t-statistics in parenthesis. The triangular distribution is imposed on the wage, market size, distance and unemployment variables, while the restricted uniform distribution is imposed on the dummy variables for common border and EU membership. The third and the fourth column present the estimated mean coefficients and their t-statistics respectively, while the fifth and the sixth columns present the estimated standard deviation parameters and their t-statistics respectively. For comparison reasons the results of the Multinomial Logit model are presented in the last two columns.

TABLE 4
Elasticities and Marginal Effects

Country	Wage (non-traditional)	Wage (traditional)	Unempl (non-traditional)	Unempl (traditional)	Border	EU	Interest
BG	-0.4262	-0.7024	-1.9547	9.4202	2.7006	8.3885	1.9444
CZ	-0.6131	-1.0561	-5.3513	-1.8953	1.4487	8.8180	2.7967
EE	-0.3824	-0.9455	-3.2253	-2.1385	0.2680	9.1136	0.5564
CR	-0.8099	-1.4806	0.8866	9.7087	0.0000	0.0000	1.4705
HU	-0.6602	-1.2896	-4.8826	-3.2887	1.4052	7.0228	2.6012
LT	-0.6726	-1.3382	-2.2621	0.0741	0.0000	4.8799	1.0054
LV	-0.6511	-1.1812	-2.3794	0.5118	0.2232	6.1675	0.3955
PL	-0.4704	-0.6636	4.9983	26.2240	0.8514	10.0009	3.1063
RO	-0.3813	-0.5462	-3.2006	-1.4245	0.0000	6.2603	0.4751
RU	-0.1324	-0.1860	-2.8028	2.3572	1.2924	0.0000	-1.0022
SI	0.1674	0.7328	-2.2684	0.1792	0.7493	4.6052	1.3605
SK	-0.7711	-1.3029	1.5027	15.8986	1.2243	7.0153	1.9972
UA	-0.2542	-0.4777	-1.6207	1.4936	0.5347	0.0000	3.2269

Elasticities are estimated for the continuous variables and marginal effects are estimated for the dummy variables. A *direct elasticity* measures the percentage change in the probability of choosing a particular alternative in the choice set with respect to a given percentage change in an attribute of that same alternative. *Direct marginal effects* represent the change in the choice probability for an alternative given a unit change in a variable related to that alternative.

TABLE 5
Simulation Results of Changes in Market shares and the Number of Firms

Size_GDP	1%		10%		50%		100%	
BG	0.001	0	0.005	0	0.022	0	0.045	1
CZ	0.004	0	0.038	0	0.202	2	0.435	4
EE	0.000	0	0.002	0	0.011	0	0.023	1
CR	0.000	0	0.003	0	0.017	0	0.035	1
HU	0.002	0	0.021	0	0.111	1	0.236	3
LT	0.001	0	0.004	0	0.021	1	0.042	0
LV	0.001	0	0.003	0	0.011	0	0.023	1
PL	0.018	0	0.185	2	1.007	11	2.148	24
RO	0.001	0	0.007	0	0.039	1	0.081	1
RU	0.025	0	0.254	2	1.230	3	2.373	26
SI	0.000	0	0.003	0	0.018	0	0.037	0
SK	0.001	0	0.008	0	0.039	0	0.079	1
UA	0.001	0	0.007	0	0.034	1	0.070	1
Prof_dist	1%		10%		50%		100%	
BG	0.008	0	0.089	1	0.973	11	1.788	20
CZ	0.007	0	0.09	1	0.899	10	1.640	18
EE	0.007	0	0.071	1	0.732	8	1.723	19
CR	0.003	0	0.034	1	0.538	6	1.21	14
HU	0.007	0	0.085	1	0.914	10	1.667	18
LT	0.006	0	0.063	0	0.747	8	1.639	18
LV	0.006	0	0.062	1	0.738	8	1.633	18
PL	0.022	0	0.261	3	1.330	15	2.164	24
RO	0.008	0	0.102	1	1.051	12	1.905	21
RU	0.032	0	0.300	3	1.267	14	2.103	23
SI	0.003	0	0.037	0	0.508	5	1.134	12
SK	0.005	0	0.059	0	0.727	8	1.476	16
UA	0.005	0	0.060	1	0.833	10	1.626	18

The first column shows the change in the market shares of the host countries regarding foreign capital allocations due to 1, 10, 50 and 100 % increase in the profitability and firm size variables. The second column shows the change in the number of firms investing in a particular country due to 1, 10, 50 and 100 % increase in the two variables.