Does it Take Time to Travel Distance?
Geography, Entry Timing and Knowledge Spillovers

Bruno Merlevede*
Ghent University and HUBrussel

August 2011
Preliminary Draft

Abstract

This paper decomposes traditional country-wide FDI spillovers in different components determined by geographical and entry-timing dimensions. We use a sufficiently long panel (1996-2005) of Romanian manufacturing firms that allows us to investigate whether spillover are limited to the regional level or whether it may simply take more time for spillovers to manifest themselves over longer distances.

JEL Classification: F2
Keywords: FDI, spillovers, dynamics, timing, regions, distance

*Corresponding author, Department of Economics, Ghent University, Tweeckerkenstraat 2, B-9000 Ghent, Belgium, Bruno.Merlevede@UGent.be
1 Introduction

Nowadays, countries actively and fiercely compete to attract foreign direct investment (see Harding and Javorcik, 2010). Policymakers are eager to do so because they expect to benefit in terms of faster economic growth in their country. Foreign direct investment brings additional capital and is expected to create new jobs. Moreover, a foreign firm will also bring more advanced technology to the host country to compete successfully with indigenous firms (Markusen, 1995). The belief that this transferred technology will be adopted by domestic firms is another reason for policymakers to attract foreign investors. These productivity spillovers from foreign to domestic firms have been investigated at least since Caves (1974). The current literature distinguishes between horizontal spillovers to firms in the same industries and vertical spillovers to firms in other industries linked to foreign firms through the supply chain. Notwithstanding thirty years of research, the literature surveys of Görg and Greenaway (2004) and Crespo and Fontoura (2007) on FDI spillovers conclude that there is no clear evidence of aggregate positive FDI spillovers. Following new theoretical insights that stress the importance of firm level heterogeneity in the study of firms’ participation in international markets (see e.g. Melitz, 2003 and Helpman et al., 2004), the literature has moved away from the idea that spillovers are unconditional and uniform. The focus has turned to the identification of characteristics that facilitate positive spillovers. These characteristics most often concern domestic firms’ characteristics such as absorptive capability (e.g. Merlevede and Schoors, 2007) or foreign firms’ characteristics such as ownership structure (Javorcik and Spatareanu, 2008). Several authors have also investigated whether spillovers may entail a regional dimension. Aitken et al. (1999) found no evidence for positive local horizontal spillovers, nor for country-wide spillovers in Venezuela. More recently, Altomonte and Colantone (2008) do find positive horizontal and backward spillovers in the top performing NUTS 2 regions in Romania, but no or negative spillovers in lagging regions. Using physical distance measures rather than administrative borders Halpern and Muraközy (2007) find that for horizontal spillovers distance matters: horizontal spillovers are only significantly positive for domestic firms close to foreign-owned firms.

This paper revisits the issue of local vs. country-wide spillovers and incorporates an important additional element in the analysis: foreign entry-timing.

Merlevede et al. (2010) show that adequately accounting for entry-timing effects reveals important new insights in the case of country-wide spillovers. FDI spillovers are typically treated as additional ‘inputs’ explaining domestic firms’ productivity in a production function estimation where the size and significance of the resulting coefficients are then taken as evidence of FDI spillovers. Spillover variables are typically based on foreign firms’ share
in total region or country-wide industry output (or employment). These spillover variables, however, lump together all foreign investment, new and old, in a single variable thus masking possible effects of foreign entry-timing. Indeed, a newly arriving foreign firm is highly unlikely to have the same impact on domestic firms’ productivity as a foreign firm that has been present in the host country for ten years. Contrary to the current spillover literature that only allows for differences between these firms to the extent that their share in industry output differs, Merlevede et al. (2010) allow for time variation of the intensity of the spillover effect, i.e. the estimated coefficient.

This paper combines the geographic and entry-timing dimensions in the analysis of regional spillovers. We use a sufficiently long panel (1996-2005) of Romanian manufacturing firms that allows us to investigate whether spillover are really limited to the regional level or whether it may simply take more time for spillovers to manifest themselves over longer distances. We do not only focus on statistical significance, i.e. the existence and direction of different spillover effects, but we also investigate the relative contributions of the different spillovers to productivity. Our data allow for a detailed regional disaggregation at NUTS 3 level. Furthermore the time-varying input-output tables for the calculation of supply chain spillovers use a Romanian industry code that maps into NACE 3-digit, which is more detailed than the typical NACE 2-digit tables (or comparable) used throughout the literature. Our final sample consists of 41 regions and 59 industries.

2 Literature on spillover effects

Horizontal spillovers run from a foreign firm to a host country firm in the same industry. Teece (1977) suggests two main channels for horizontal spillovers: technology imitation (the demonstration effect) and mobility of workers trained by foreign firms (see also Fosfuri et al., 2001, and Görg and Strobl, 2005). Marin and Bell (2006) find that training activities by foreign subsidiaries are related to stronger horizontal spillovers. Foreign entry may also fuel competition in the domestic market. Fiercer competition urges host country firms to either use existing technologies and resources more efficiently or adopt new technologies and organizational practices, which provides another important channel of horizontal spillovers (see Aitken and Harrison, 1999, and Glass and Saggi, 2002). None of these effects is necessarily positive. Labor market dynamics may entail negative spillovers such as a brain drain of local talent to foreign firms to the detriment of local firm productivity (Blalock and Gertler, 2004) or an overall increase in wages irrespective of productivity improvements caused by foreign firms paying higher wages (Aitken et al., 1996). Where foreign technology is easily copied, the foreign investor may choose to avoid leakage costs on state-of-the-art technology
by transferring technology that is only marginally superior to technology found in the host
country (see Glass and Saggi, 1998). Such policies obviously limit the scope for horizontal
spillovers via demonstration effects. The higher productivity of foreign affiliates may also
lead to lower prices or less demand for the products of domestic competitors. If domestic
firms fail to raise productivity in response to the increased competition, they will be pushed
up their average cost curves. Ultimately, domestic producers may not merely fall behind,
but fall by the wayside, driven out of business by the shock of foreign entry (see Aitken and
Harrison, 1999, on this market-stealing effect). These partial effects are hard to disentangle empirically and a general measure for horizontal spillovers will identify the net effect of all
these channels.

Figure 1 shows how backward spillovers run from the foreign firm to its upstream local
suppliers. Thus, even if foreign firms attempt to minimize their technology leakage to direct
competitors (horizontal effect), they may still want to assist their local suppliers in providing
inputs of sufficient quality in order to realize the full benefits of their investment. In other
words, they want the inputs from the host country to be lower cost yet similar in quality
to inputs in the home country. If the foreign firm decides to source locally, it may transfer
technology to more than one domestic supplier and encourage upstream technology diffusion
to circumvent a hold-up problem. Rodriguez-Clare (1996) shows that the backward linkage
effect is more likely to be favorable when the good produced by the foreign firm uses inter-
mediate goods intensively and when the home and host countries are similar in terms of the
variety of intermediate goods produced. Under reversed conditions, the backward linkage
effect could even damage the host country’s economy. Figure 1 also suggests how a forward spillover goes from the foreign firm to its downstream local buyer of inputs. The availability
of better inputs due to foreign investment enhances the productivity of firms that use these
inputs. However, there is also a danger that inputs produced locally by foreign firms are
more expensive and less adapted to local requirements. In this case there would be a negative
forward spillover.

The empirical framework to analyse spillover effects can be seen as an ‘augmented’ pro-
duction function, where spillover variables are added to other explanatory variables such as
labour, capital, and material inputs. The typical measure employed to identify horizontal or
within-industry spillovers is given in (1). For a (domestic) firm \(i\) in industry \(j\) at time \(t\) it is
of the following form:

\[
HR_{jt} = \sum_{f \in \mathcal{J}} F_{ft} Y_{ft} / \sum_{f \in \mathcal{J}} Y_{ft}
\]

(1)

where \(Y\) is either output or employment and \(F\) is an indicator of a firm \(f\)’s foreign ownership
status. $F$ could be the exact percentage of shares owned by foreign investors\(^1\) or it could be a dummy variable that takes the value 1 independent of the exact amount of shares woned by foreign investors. In line with the definition commonly applied by e.g. the OECD or the IMF, at least 10% of shares should be owned by a single foreign investor for a firm to be seen as foreign.\(^2\) $HR_{jt}$ in (1) is the share of output that is produced by foreign firms in industry $j$. This spillover variable is build up to industry level from firm-level data, implying that $HR_{jt}$ has the same value for all firms $f$ in industry $j$ at time $t$. This will be important for calculating appropriate standard errors on these variables (cf. infra).

The definition of the backward spillover variable, $BK_{jt}$, starts from the horizontal measure and combines it with information from input-output tables as in

$$BK_{jt} = \sum_{k \text{ if } k \neq j} \gamma_{jkt} * HR_{kt}$$

where $\gamma_{jkt}$ is the proportion of industry $j$’s output supplied to sourcing industry $k$ at time $t$. The $\gamma$s are calculated from (time-varying) IO-tables for intermediate consumption. Inputs sold within the firm’s industry are excluded ($k \neq j$) because this is captured by $HR_{jt}$. Since firms cannot easily or quickly switch industries to buy inputs, this approach avoids the problem of endogeneity by using the share of industry output sold to downstream domestic markets $k$ with some level of foreign presence $HR_{kt}$. Employing the share of firm output sold to foreign firms in different industries would cause endogeneity problems if the latter prefer to buy inputs from more productive domestic firms. In the same spirit, the forward spillover variable $FW_{jt}$ is defined as:

$$FW_{jt} = \sum_{l \text{ if } l \neq j} \delta_{jlt} * HR_{lt}$$

where the IO-tables reveal the proportion $\delta_{jlt}$ of industry $j$’s inputs purchased from upstream industries $l$. Inputs purchased within the industry ($l \neq j$) are again excluded, since this is already captured by $HR$.

The spillover variables $HR_{jt}$, $BK_{jt}$, and $FW_{jt}$ are then regressed on the productivity of (domestic) firm $i$ in industry $j$. The size, sign, and significance of the resulting coefficients are then taken as evidence on spillovers.

---

\(^1\)e.g. if 80% of shares is foreign-owned, $F$ equals 0.8.

\(^2\)A firm where only 5% of shares are owned by one or more foreigners is considered as a domestic firm.
3 A regional dynamic approach to spillovers

The typical measure to capture within-industry spillover effects $HR_{jt}$ in (1) is thus the share of output that is produced by foreign firms in industry $j$. For a given firm $i$ in industry $j$ in region $r$ at time $t$ we can now break down (1) into different geographical subcomponents as follows:

$$HR_{jt} = \frac{\sum_{f \in j} F_{ft}Y_{ft}}{\sum_{f \in j} Y_{ft}}$$

$$= \frac{\sum_{f \in j, f \in r} R_{frt}F_{ft}Y_{ft}}{\sum_{f \in j} Y_{ft}} + \frac{\sum_{f \in j, f \in r} NB_{frt}F_{ft}Y_{ft}}{\sum_{f \in j} Y_{ft}} + \frac{\sum_{f \in j, f \in r} (1 - R_{frt} - NB_{frt}) F_{ft}Y_{ft}}{\sum_{f \in j} Y_{ft}}$$

where $R_{frt}$ indicates whether firm $f$ is located in region $r$, and $NB_{frt}$ indicates whether firm $f$ is located in a contiguous region of $r$. Finally, $(1 - R_{frt} - NB_{frt})$ will equal 1 if if firm $f$ is located in a further-away non-neighbouring region, resulting in a rest-of-country spillover. Clearly, (5) could be further decomposed in a straightforward manner to account for second- or even higher-order neighbours. Introducing $HR_{jt}$ as a single variable in a regression thus involves the implicit assumption that the spillover intensity -as measured by the coefficient obtained on $HR_{jt}$, say $\alpha$- is the same within and across regions. In our empirical test we relax this assumption and allow the coefficients to differ between the different subcomponents in (5), i.e. obtain estimates $\alpha_R$, $\alpha_{NB}$, and $\alpha_{RoC}$.

How does this link to the existing research into regional spillovers? By introducing all three subcomponents of (5) in a regression analysis we differ from at least part of the literature on regional FDI spillovers that does not allow for cross-regional spillovers. Only the first term of the decomposition is included among the regressors, implicitly assuming that spillovers are confined to region boundaries and do not cross borders. This runs counter to Halpern and Murakozy (2008) who find that horizontal spillovers vary with distance, but do not disappear. It also runs counter to macro-spillover studies as e.g. Keller (2002) who finds that spillovers between countries are declining with distance.

There is a further difference with the regional spillover literature. To see this more clearly, rewrite (5) as follows:
\[
HR_{jt} = \frac{\sum_{f \in j} R_{ft} Y_{ft}}{\sum_{f \in j} R_{ft} Y_{ft}} * \frac{\sum_{f \in j} R_{ft} Y_{ft}}{\sum_{f \in j} Y_{ft}} + \frac{\sum_{f \in j} R_{ft} Y_{ft}}{\sum_{f \in j} Y_{ft}} * \frac{\sum_{f \in j} R_{ft} Y_{ft}}{\sum_{f \in j} Y_{ft}} + \frac{\sum_{f \in j} (1 - R_{ft} - NB_{ft}) Y_{ft}}{\sum_{f \in j} Y_{ft}} * \frac{\sum_{f \in j} (1 - R_{ft} - NB_{ft}) Y_{ft}}{\sum_{f \in j} Y_{ft}}
\]

Studies that investigate regional spillovers typically define their regional spillover variable as the first part of the first term. This spillover variable is constructed as output produced by foreign firms in industry \( j \) in region \( r \) as a share of total regional industry \( j \) output, rather than as a share of country-wide industry \( j \) output.

\[
\frac{\sum_{f \in j} R_{ft} Y_{ft}}{\sum_{f \in j} R_{ft} Y_{ft}}
\]

The definition of an appropriate measure relates to one’s idea about spillover potential (this is what the variables are intended to capture). Consider two neighbouring regions A and B. In region A 10 out of the total of 100 units are produced by foreign firms, while in region B 10,000 out of the total 100,000 units are produced by foreigners. If one believes that the spillover potential is the same in both regions, (7) is appropriate. On the other hand if one believes that there is an absolute scale aspect to spillovers, the first term of (5) is a more appropriate choice. In the former case, spillovers should be thought of as limited to the region level since it is difficult to carry this definition through to the cross-region level. Following the logic of relative within territorial unit presence, the spillover potential from neighbours should be measured as:

\[
\frac{\sum_{f \in j} NB_{ft} Y_{ft}}{\sum_{f \in j} NB_{ft} Y_{ft}}
\]

Definition (8) results in a value of 0.10 for both region A and B. However, it is quite counter-intuitive that the spillover potential from region A to B would equal the spillover potential from B to A. This is not the case when one uses the second subcomponent in (5). In our example, this yields a spillover potential from A to B of \( \frac{10}{100 \times 100} \) and a spillover potential from B to A of \( \frac{10,000}{100 \times 100} \). These values seem to be in line with the cross-region spillover potential one would expect. Note that the subcomponents in (5) will also show a relatively high correlation with simple count measures of foreign investment (e.g. Driffield (2000X) uses this type of measure). Regional definitions for \( BK \) and \( FW \) follow from (2) and (3) above. Since we
only have input-output tables at the country-level, we will have to assume that the technical coefficients are similar across regions and equal those derived from the country-level input output table.

3.1 Regional spillovers and the timing of foreign entry

Abstracting for the moment from the implicit geographical dimension in $HR_{jt}$ discussed in the previous section, (1) hides another important dimension that deserves a closer inspection. To see things more clearly, consider a different breakdown of (1):

$$HR_{jt} = \sum_{f \in j} \frac{\tilde{F}_f Y_{ft}}{\sum_{f \in j} Y_{ft}} + \sum_{f \in j} \frac{\tilde{F}^2_f Y_{ft}}{\sum_{f \in j} Y_{ft}} + ... + \sum_{f \in j} \frac{\tilde{F}^n_f Y_{ft}}{\sum_{f \in j} Y_{ft}}$$

(9)

here $\tilde{F}^x$ is a variable indicating foreign ownership status and entry timing. $\tilde{F}^x_{ft}$ will equal the percentage of shares owned by foreign investors in firm $f$ if at least 10% of shares were owned by a single foreign investor in year $t - x + 1$ and firm $f$ was not foreign owned in year $t - x$, i.e. the investment took place between $t - x + 1$ and $t - x$. So $\tilde{F}^x_{ft}$ is set to the percentage of shares owned by foreign investors if

$$\left( \sum_{v=0}^{x-1} F_{f,t-v} = x \right) \land \left( \sum_{w=x}^{\infty} F_{f,t-w} = 0 \right)$$

$HR_{jt}$ is thus broken down into $HR^1_{jt}$, $HR^2_{jt}$, ... along the lines of the time since foreign entry (note the difference with pure calendar time or taking lags of $HR_{jt}$). Time varying definitions for $BK^x_{jt}$, and $FW^x_{jt}$ follow from (2) and (3) above

$$BK^x_{jt} = \sum_{k \neq j \neq k} \gamma_{jk} * HR^x_{kt}$$

(10)

$$FW^x_{jt} = \sum_{l \neq j \neq l} \delta_{jl} * HR^x_{lt}$$

(11)

Various transmission channels discussed above imply an impact of foreign entry timing. The mobility of workers trained by foreign firms, nor technology imitation are likely to materialize in the very short run. Likewise, vertical spillovers driven by access to better inputs produced by foreign firms or by supplying inputs to multinational companies are not necessarily instantaneous nor permanent. There is also further circumstantial evidence on the importance of entry timing. For a long panel (1982-95) of firms in the Irish electronics
sector Görg and Ruane (2001) find indications that foreign firms start off with a relatively low extent of local linkages, but as they get accustomed, they proceed to develop more local input linkages. Giroud (2007) finds that local suppliers benefit significantly less form foreign presence in Vietnam than in Malaysia that opened up its economy to foreign entry much earlier. Ivarsson and Alvstam (2005) conclude that technology transfer to suppliers seems more efficient in older MNE plants based on their Volvo case study. Within multinationals technology is also not necessarily easily or rapidly transferred (see Urata and Kawai, 2000). This may give rise to specific time patterns in the transfer of technology to foreign affiliates and the ensuing spillovers. Merlevede et al. (2010) allow for an entry-timing pattern of spillover effects, i.e. every term in (9) has its own coefficient $\alpha_F(x)$. They find that newly arriving foreign entrants generate spillover effects that are very different from those generated by more mature foreign firms that have been present in the domestic economy for a longer period.

The decomposition along foreign entry timing is highly relevant for the study of the regional aspects of spillovers as well. It is not unlikely that the timing pattern for within-region spillovers is different from the timing pattern for cross-region spillovers. It may e.g. take more time for domestic firms to absorb spillovers from foreign firms in further-away regions. A failure to find cross-region spillovers on the basis of aggregate variables as in (5) or (??) could simply be due to the fact that entry timing has been neglected, rather than that these spillovers are truly absent. Further note that some papers limit the scope of spillovers to the boundaries of the region by the construction of the spillover variable without accounting for possible cross-region effects (e.g. Nicolini and Resmini, 2010, or Crespo et al., 2009). Figure 1 gives an overview of the dimensions that we will try to capture in our empirical framework.
4 Empirical approach and data

4.1 Empirical approach

As indicated above FDI spillovers are commonly analyzed in a production function framework. Firm level total factor productivity is obtained in a first step estimation and in a second step the FDI spillover variables together with some further controls are treated as additional ‘input’ explaining domestic firms’ productivity. The careful estimation of the production function is thus an important building block in the analysis. The basic problem in estimating productivity is that firms react to firm-specific productivity shocks that are not observed by the researcher. Griliches and Mairesse (1995) provide a detailed account of this problem and make the case that inputs should be treated as endogenous variables since they are chosen on the basis of the firm’s unobservable assessment of its productivity. OLS estimates of production functions therefore yield biased estimates of factor shares and biased estimates of productivity. The semi-parametric approaches by Olley and Pakes (1996) (OP) and a more recent modification of it by Levinsohn and Petrin (2003) (LP), and the dynamic panel data approach by Blundell and Bond (1998) (DPD) are alternative methodologies to overcome the endogeneity bias in estimating production functions. Both types of methodologies have been widely used in the recent literature on firm level heterogeneity for derivation of total factor productivity measures. More recently, Ackerberg et al. (2008) (ACF) argue that, while there are some solid and intuitive identification ideas in the paper by Levinsohn and Petrin (2003), their semi-parametric techniques suffer from collinearity problems casting doubt on the methodology. They suggest an alternative methodology that make use of the ideas in these papers, but do not suffer from these collinearity problems. We will use ACF tfp as our base case, but we will check the robustness of our results with respect to other tfp-measures.

We estimate domestic industry production functions separately for each Nace 2-digit manufacturing industry $j$ in the period 1996–2005, excluding firms that are foreign at some point in time from the estimation. Capital, labor, and material inputs elasticities are thus treated as industry-specific. A measure of total factor productivity $tfp_{ijt}$ for firm $i$ in industry $j$ in region $r$ at time $t$ is obtained as the difference between output and capital, labor, and material inputs, multiplied by their estimated coefficients:

$$tfp_{ijt} = Y_{ijt} - \beta_{lj} l_{ijt} - \beta_{kj} k_{ijt} - \beta_{mj} m_{ijt}$$ (12)

In the second step, we relate $tfp_{ijt}$ to a firm specific effect, a vector of spillover variables, $FDI_{jt}$, and firm and industry level controls. Note that (13) now pools firms from all indus-
tries together in one large panel, whereas (12) is estimated by industry. This specification follows the standard in the literature (e.g. Smarzynska Javorcik, 2004)

\[
\text{tfp}_{ijrt} = \alpha_i + \Psi_1 f (\text{FDI}_{jt-1}) + \Psi_2 Z_i(j)t + \xi_{ijt} \quad (13)
\]

The vector of spillover variables \((\text{FDI}_{jt-1})\) covers different horizontal and vertical spillover variables described above. More specifically, \(HR, BK,\) and \(FW\) are decomposed in function of geography and foreign entry timing as follows (for clarity industry and time subscripts are dropped). We consider three different regional dimensions: within-region spillovers, \(HR_{\text{reg}}\), first-order neighbours, \(HR_{\text{nb}}\), and the regions that make up the rest-of-country, \(HR_{\text{roc}}\). Considering the time span of our dataset (1996-2005, cf. infra) we opt to include \(HR_{X1}\) to \(HR_{X4}\) and create a variable \(HR_{X5^+}\) which aggregates all foreign firms that have been present for at least four full years on the domestic market, hence the summation from 1 to \(5^+\) in (14). Since we do not have information on foreign entry time prior to 1996, the time span of the dataset for the estimations reduces to 2001-2005 because of missing values.

\[
\Psi_1 f (\text{FDI}_{jt-1}) = \sum_{x=1}^{5^+} (\alpha_{\text{reg}}^x HR_{\text{reg}} + \alpha_{\text{nb}}^x HR_{\text{nb}} + \alpha_{\text{roc}}^x HR_{\text{roc}})
\]

\[
+ \sum_{x=1}^{5^+} (\beta_{\text{reg}}^x BK_{\text{reg}} + \beta_{\text{nb}}^x BK_{\text{nb}} + \beta_{\text{roc}}^x BK_{\text{roc}})
\]

\[
+ \sum_{x=1}^{5^+} (\delta_{\text{reg}}^x FW_{\text{reg}} + \delta_{\text{nb}}^x FW_{\text{nb}} + \delta_{\text{roc}}^x FW_{\text{roc}})
\]

\(Z_{i(j)t}\) is a vector of control variables. Specifically we control for competition within the industry, measured by the Herfindahl index, import competition in the industry, the share of intermediates supplied in total industry output, and firm age. Further we use the region-industry share of national industry activity and the region’s share of national manufacturing activity to control for region and region-industry agglomeration effects.

Specification (13) is first-differenced and then estimated by OLS, including industry \((\alpha_j)\), region \((\alpha_r)\), and time dummies \((\alpha_t)\). Because \(\text{FDI}_{jt}\) and some control variables are defined at the industry level, and estimations are performed at the firm level, standard errors need to be adjusted (Moulton, 1990). Standard errors are therefore clustered for all observations in the same region, industry and year (cf. Smarzynska Javorcik, 2004). This results in (15) as final specification to be estimated.

\[
\Delta \text{tfp}_{ijrt} = \Psi_1 \Delta f (\text{FDI}_{jt-1}) + \Psi_2 \Delta Z_{i(j)t} + \alpha_t + \alpha_j + \alpha_r + \xi_{ijrt} \quad (15)
\]
4.2 Data

We use firm-level data for a panel of Romanian manufacturing firms during 1996–2005. Because most foreign investment entered Romania after 1996, Romania makes a very good candidate to study the dynamic impact of recent foreign investment on domestic firm productivity. Our firm-level data are taken from the Amadeus database by Bureau Van Dijk. Amadeus is a pan-European database of financial information on public and private companies. Every month Bureau Van Dijk issues a new DVD with updated information. A single issue of the DVD contains only the latest information on ownership and firms that go out of business are dropped from the database fairly rapidly. Furthermore, because Bureau Van Dijk updates individual ownership links between legal entities rather than the full ownership structure of a given firm, the ownership information on a specific DVD-issue often consists of a number of ownership links with different dates, referring to the last verification of a specific link. To construct our dataset with entry, exit, and time-specific foreign entry in local Romanian firms, we therefore employed a series of different issues of the database. However, since ownership information is gathered at irregular intervals, we do not have ownership information for all firm-owner-year combinations.\(^3\) Given these specificities of Amadeus, we first created a dataset at the firm-owner-year-level with the available information from Amadeus. We then filled out missing firm-owner-year-entries under restriction that the full ownership structure cannot exceed 100%. In case of time gaps between entries for the same owner-firm combination but with a different share-size we assume that changes show up immediately in the database. We then fill out the gaps with the older information.\(^4\)

Data are deflated using industry price level data at Nace rev.1.1 2-digit level.\(^5\) These are taken from the Industrial Database for Eastern Europe from the Vienna Institute for International Economic Studies and from the Statistical Yearbook of the Romanian National Statistical Office (RNSO). Real output \(Y\) is measured as operating revenues deflated by producer price indices of the appropriate Nace industry; real material inputs \(M\), are deflated by a weighted intermediate input deflator where the industry-specific weighting scheme is drawn from the IO tables. Labor \(L\) is expressed as the number of employees. Real capital \(K\) is measured as fixed assets, deflated by the average of the deflators for the following five Nace

\(^3\)Identifying the same owner in different issues is not always straightforward since an ID is only listed in case the owner is a firm that is listed in Amadeus itself. For all other owners matching is done on the basis of the name. Differences in spacings, plurals, addition to the name of a company-type, the use of characters specific to Romanian versus standard Roman characters, ... in different issues are corrected for.

\(^4\)e.g. | Amadeus | immediate |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>40</td>
</tr>
<tr>
<td>2001</td>
<td>40</td>
</tr>
<tr>
<td>2002</td>
<td>50</td>
</tr>
</tbody>
</table>

\(^5\)Nomenclature générale des activités économiques dans les Communautés européennes.
Figure 2: Geographic distribution of domestic and foreign firms across NUTS3 regions, evolution 2000-05 (similar scaling on left and righthandside panels)
Figure 3: Geographic distribution of horizontal spillover potential measures based on different normalisations, evolution 2000-05 (similar scaling on left and righthandside panels)
industries: machinery and equipment (29); office machinery and computing (30); electrical machinery and apparatus (31); motor vehicles, trailers, and semi-trailers (34); and other transport equipment (35) (see Smarzynska Javorcik, 2004). IO tables for the period 1996–2005 were obtained from the RNSO. The tables are in national industry classification, but the RNSO provided a mapping into Nace rev. 1.1. The RNSO tables are fairly detailed and identify 59 manufacturing sectors. This provides us with richer detail in vertical relationships than the more common IO-tables at Nace 2-digit that only have 22 manufacturing sectors.

We restrict the dataset to firms with on average 5 employees over the sample period. The dataset is further trimmed for outliers by removing the top and bottom percentiles of the annual growth rates of real operating revenues, real capital, labour, and real material inputs. The share of foreign firms in the total number of sample firms steadily increased from 16% to 22% (10 to 15% if small firms are not excluded). The 2003 exit rate is high, but this pattern is confirmed by the pattern in the Romanian Trade Register (Trade Register data also include agriculture and services though). The stylized facts commonly found in the literature are confirmed in our dataset. Foreign firms are larger in terms of employment and capital, produce more output and are more productive. The latter holds across different estimation techniques. The productivity bonus of foreign over domestic firms ranges between 14% in case of the Olley-Pakes methodology (OP) and 36% in case of the Levinsohn-Petrin methodology (LP). There is a fairly high correlation between the tfp-measures resulting from different estimation techniques. For our empirical results we will rely on the tfp measure obtained by the methodology proposed by Ackerberg et al. (2008). Left aside the highly concentrated tobacco industry (Nace 16), on average (over industries) some 15% of industry output was produced by foreign firms in 1996. The share of foreign firms varies between 7% and 30%. In 2005 on average 39% of industry output was produced by foreign firms, while shares varied between 15% and 57% across industries. The correlation across years and spillovers is limited.

As a unit of geographical observation we focus on the NUTS 3 level. From the Amadeus database we know in which NUTS 3 region a firm is located. The NUTS (Nomenclature of territorial units for statistics) classification is a hierarchical system for dividing up the economic territory of the EU. NUTS 3 regions are defined as ‘small regions for specific diagnoses’ (population between 150,000 and 800,000). At this level Romania is divided in 42 territorial units, i.e. 41 counties and the capital Bucharest. Our data, however, do not allow us to discriminate between Bucharest and the surrounding county Ilfov. Therefore we have

---

6If the 'outlier' is the first or last observation for a specific firm and other datapoints are normal, the other firm-year data are kept. If not all observations for this firm are dropped from the dataset.

7We apply the procedure from Amiti and Konings (2007) to calculate investment from our data.

8Including or excluding the tobacco industry does not affect our results.
41 territorial units in our analysis. As a control we will also collapse the data to the NUTS 2 level that divides Romania in 8 regions. The use of region or location in the remainder of the paper always refers to the NUTS 3 classification, unless explicitly mentioned otherwise. To avoid problems with multi-plant firms we focus on unconsolidated data. The geographic distribution of across NUTS 3 regions is represented in Figures 2 and 3.

5 Results
to be completed

6 Conclusions
Table 2: Spillover patterns in above and below median productivity regions

Below median TFP regions

<table>
<thead>
<tr>
<th>MNE entered when?</th>
<th>t</th>
<th>t-1</th>
<th>t-2</th>
<th>t-3</th>
<th>t-4 or earlier</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HORIZONTAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MNE entered where?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>same region</td>
<td>0.287</td>
<td>-1.475</td>
<td>-11.982**</td>
<td>3.057</td>
<td>1.331</td>
</tr>
<tr>
<td>1st order neighbouring region</td>
<td>1.582</td>
<td>-3.898***</td>
<td>-3.619***</td>
<td>0.690</td>
<td>2.049***</td>
</tr>
<tr>
<td>rest-of-country</td>
<td>0.522</td>
<td>-1.732***</td>
<td>-1.301***</td>
<td>0.190</td>
<td>2.069***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MNE entered when?</th>
<th>t</th>
<th>t-1</th>
<th>t-2</th>
<th>t-3</th>
<th>t-4 or earlier</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BACKWARD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MNE entered where?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>same region</td>
<td>-0.264</td>
<td>81.087</td>
<td>76.331**</td>
<td>39.236</td>
<td>6.637</td>
</tr>
<tr>
<td>1st order neighbouring region</td>
<td>21.516*</td>
<td>14.896**</td>
<td>18.111***</td>
<td>2.970</td>
<td>-5.063**</td>
</tr>
<tr>
<td>rest-of-country</td>
<td>2.132</td>
<td>4.449***</td>
<td>3.731***</td>
<td>4.097***</td>
<td>0.100</td>
</tr>
</tbody>
</table>

Above median TFP regions

<table>
<thead>
<tr>
<th>MNE entered when?</th>
<th>t</th>
<th>t-1</th>
<th>t-2</th>
<th>t-3</th>
<th>t-4 or earlier</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HORIZONTAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MNE entered where?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>same region</td>
<td>-0.617</td>
<td>-2.194**</td>
<td>-1.002*</td>
<td>0.345</td>
<td>1.628***</td>
</tr>
<tr>
<td>1st order neighbouring region</td>
<td>0.820</td>
<td>-0.923</td>
<td>-2.128***</td>
<td>-0.209</td>
<td>1.737***</td>
</tr>
<tr>
<td>rest-of-country</td>
<td>0.071</td>
<td>-0.863*</td>
<td>-1.501***</td>
<td>0.733***</td>
<td>2.061***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MNE entered when?</th>
<th>t</th>
<th>t-1</th>
<th>t-2</th>
<th>t-3</th>
<th>t-4 or earlier</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BACKWARD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MNE entered where?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>same region</td>
<td>18.482**</td>
<td>11.719***</td>
<td>3.427***</td>
<td>3.939***</td>
<td>0.389</td>
</tr>
<tr>
<td>1st order neighbouring region</td>
<td>-2.015</td>
<td>2.276</td>
<td>0.190</td>
<td>3.320</td>
<td>-0.520</td>
</tr>
<tr>
<td>rest-of-country</td>
<td>3.350</td>
<td>10.601***</td>
<td>4.465***</td>
<td>4.422***</td>
<td>-0.603</td>
</tr>
</tbody>
</table>
References


Bureau van Dijk Electronic Publishing, Amadeus Database, Brussels, Belgium, various issues.


Harind, T., and Javorcik, B., 2010, "Roll out the Red Carpet and they Will Come: Investment Promotion, Information Asymmetries and FDI Inflows",


Ivarsson, I. and Alvstam, C., 2005, "Technology Transfer from TNCs to Local Suppliers in Developing Countries: A Study of AB Volvo’s Truck and Bus Plants in Brazil, China, India, and Mexico", *World Development, 33*(8), pp. 1325-1344


Romanian National Statistical Office (RNSO), *Statistical Yearbook*, Bucharest, Romania, 2005


