Työpapereita
Working Papers

243

THE RETURN TO THE TECHNOLOGICAL FRONTIER: THE CONDITIONAL EFFECTS OF PLANTS' R&D ON THEIR PRODUCTIVITY IN FINNISH MANUFACTURING

Eero Lehto
Petri Böckerman
Janne Huovari

PALKANSAAJIEN TUTKIMUSLAITOS LABOUR INSTITUTE FOR ECONOMIC RESEARCH

PALKANSAAJIEN TUTKIMUSLAITOS •TYÖPAPEREITA
LABOUR INSTITUTE FOR ECONOMIC RESEARCH • DISCUSSION PAPERS

243

The return to the technological frontier: The conditional effect of plants' R&D on their productivity in Finnish manufacturing*

Eero Lehto**
Petri Böckerman**
Janne Huovari***

* This paper is a part of the project "Regional productivity growth in Finland" (Tuottavuuden alueellinen kasvu Suomessa) financed by TEKES (National Technology Agency of Finland). The matched data that are used to produce the results can be accessed on site at the Research Laboratory of the Business Structures Unit of Statistics Finland. Paul A. Dillingham has kindly checked the English language. The usual disclaimer applies.

** Labour Institute for Economic Research. Address: Pitkänsillanranta 3A, 6. krs., FI-00530 Helsinki, FINLAND. Phone: +358-9-25357330. Fax: +358-9-25357332. E-mails: eero.lehto@labour.fi and petri.bockerman@labour.fi

*** Pellervo Economic Research Institute PTT. E-mail: janne.huovari@ptt.fi

Helsinki 2008

ISBN 978–952–209–057–7
ISSN 1795–1801

TIIVISTELMÄ

Tutkimuksessa tarkastellaan sitä, riippuuko tutkimus- ja kehitysmenojen vaikutus tuottavuuteen toimipaikan sijainnista toimialan tehokkuusrintaman suhteen. Aineistona käytetään teollisuuden toimipaikka-aineistoa vuosilta 1995-2005. Tutkimuksessa eritellään sekä toimipaikan oman että sen emoyrityksen tutkimus- ja kehitystoiminnan vaikutuksia. Tämän lisäksi tarkastellaan maantieteellisellä läheisyydellä painotetun muiden yritysten T&K -kannan vaikutuksia. Tulosten mukaan toimipaikan omalla ja sen emoyrityksen TK:lla on positiivinen vaikutus tuottavuuteen. Toimipaikan oman T&K:n vaikutus tuottavuuteen vähenee toimipaikan etäisyyden kasvaessa toimialan tehokkuusrintamasta. Muiden yritysten T&K-kannan vaikutus on myös keskimäärin positiivinen, mutta tämä vaikutus kasvaa toimipaikan etäisyyden kasvaessa toimialan tehokkuusrintamasta. Tämän lisäksi havaitaan se, että toimipaikoilla on taipumus lähentyä toimialan tehokkuusrintamaa riippumatta muiden yritysten T&K-kannan vaikutuksista.

ABSTRACT

This paper examines, through the use of plant-level data, whether R&D's productivity impact is contingent on the distance of a plant's productivity from the industry's technological frontier. R&D is specified as an accumulated stock from R&D investments. We analyse the productivity effect of a plant's own R&D as well as the productivity impact of the plant's parent firm's and other firms' proximity-weighted R&D stocks. The results show that a plant's own and a parent firm's R&D have a positive productivity impact and that the former impact decreases as the distance from the industry's technological frontier increases. Furthermore, the productivity effect of other firms' proximity-weighted R&D is, on average, positive, but this impact increases in the distance from the technological frontier. Another important finding is that all the plants tend to converge towards the industry's technological frontier despite the size of external R&D spillovers.

JEL Codes: D24, L00

Keywords: productivity, efficiency, technological frontier, spillovers, convergence

1. INTRODUCTION

This paper explores the productivity impact of R&D through the use of plant-level data. The effect from a plant's own R&D and that of other firms' R&D is considered. The other firms' R&D is seen, above all, as a source of the existing knowledge which firms by various schemes – in the form of spillovers (technological externalities) or through the market by means of pecuniary externalities[1] – adopt to improve their own productivity. We also examine the geographical proximity of knowledge spillovers and the other effects from other firms' knowledge capital that are intermediated through the market. In particular, this paper tests whether the productivity effects of a plant's own R&D and other firms' R&D are conditional on the plant's efficiency, more specifically on the distance of the plant's productivity from the industry's technological frontier. Furthermore, our empirical approach allows us to detect possible convergence towards the efficiency frontier.

A plant's R&D and its parent firm's own R&D are usually seen as describing efforts to create the firms' own knowledge. The firm's own R&D's function to strengthen a plant's absorption capacity – the possibility to which Cohen and Levinthal (1989) have referred – cannot, however, be forgotten.[2] Anyway, insofar as the firm's own R&D really resembles efforts to innovate, its positive impact on productivity is believed to be greatest in the neighbourhood of the industry's technological frontier according to the hypothesis put forward by Vandenbussche *et al.* (2006). Other firms' R&D stock, on the other hand, describes the potential to absorb from the other players in the market. Because this activity is based on imitation, its productivity impacts are assumed to be greatest far away from the industry's technological frontier.

Acemoglu *et al.* (2006) and Vandenbussche *et al.* (2006) test the above hypotheses by using country-level data. Acemoglu *et al.* (2006) show that R&D intensity increases as a country approaches the world technology frontier. Vandenbussche *et al.* (2006) explain total factor productivity by dividing the labour force into subgroups according to the educational level. In particular, Vandenbussche *et al.* (2006) discover that the productivity impact of the highly educated decreases as the distance from the technological frontier increases. For the less educated the finding is the opposite. In contrast, Girma (2005) analyses the productivity impacts of foreign direct investments (FDI). Girma (2005) uses firm-level data, and studies whether the productivity effects differ as a function of distance from the technological frontier. The nonlinearity of the impacts is examined through the use of threshold regression techniques. The results show that the productivity benefit from FDI increases with R&D-intensity until some threshold level beyond which it becomes less pronounced.

The theoretical literature on the subject has considered the implications of relative efficiency on the orientation of activities and on the use of resources. Being close to the industry's technological frontier, one cannot learn much from others, by definition. This implies that one should concentrate on innovation rather than on imitation (see Acemoglu *et al.,* 2006; Vandenbusschen *et al.,* 2006). Although the literature has stressed the relatively high requirements for the absorption of external knowledge (Cohen and Levinthal, 1989), one can follow Vandenbusschen *et al.* (2006) and take it as a starting point that the adoption of existing knowledge is almost always easier than the creation of new knowledge.

The sensitivity to circumstances and the tacitness of knowledge both imply that technological externalities are geographically restricted (see Breschin and Lissonin, 2001a, 2001b; Morgan, 2004). The geographical proximity of spillovers – which we also implicitly assume in the construction of the variables in our empirical analysis – is extensively considered in the empirical research (see e.g. Orlando, 2004; Lehto, 2007).

We contribute to the existing research by focusing on the impacts of R&D in the context that specifies a unit's position in relation to the industry's technological frontier at the plant level. How the distance from the frontier affects the productivity impacts of R&D is considered, in particular, and, moreover, whether the distance from the frontier affects the productivity impacts from the plant's own and external R&D differently. By using plant-level data we are able to separately control the impacts that arise from a unit's own actions and the impacts generated by the other plants' R&D through externalities. These effects are absent in the existing literature that relies on country- and industry-level data sets. We also introduce a theoretical analysis to the mechanism that produces the productivity impacts. Our hypotheses are related to the productivity impacts of R&D from various sources and the interaction of these effects and a plant's distance from the industry's technological frontier.

This paper evaluates the impact of R&D on both the total factor productivity and the labour productivity of the Finnish manufacturing plants over the period 1995-2005. Our study takes advantage of a large plant-level data in which firms' R&D is allocated to plants that actually carry R&D projects. This thoroughness of the data makes it possible to explain a plant's productivity by means of its own R&D and proximity-weighted R&D of a parent firm's other plants and other firms in the relevant market. The R&D variable that we use is the R&D stock. This allows us to take into account previous R&D investments that have been discovered to affect productivity by Rouvinen (2002).

The structure of the paper is as follows. Section 2 illustrates the forces that weaken and strengthen the productivity impacts. Section 3 describes the data. Section 4 introduces the estimated model and its variables. Section 5 reports our results. Section 6 considers the robustness of the results and the last section concludes.

2. THEORETICAL CONSIDERATIONS AND THE HYPOTHESES

This section examines the productivity mechanism in the setting in which a plant's relative efficiency at the starting point is allowed to vary and the unit costs related to the productivity project are specified to be a function of the distance from the industry's technological frontier. The productivity project either uses the existent knowledge, being imitative, or it creates new knowledge, being innovative. In Figure 1, the unit costs of an additional output generated by the productivity project and the unit income from that project vary along the vertical axis and a plant's productivity at the starting point varies along the horizontal axis. The vertical line F describes the industry's technological frontier and i gives the rank of a productivity project, which is to be applied within a given time period. Curve MC(i) (i = 1 or 2) represents the unit marginal costs the adoption of the existent knowledge to generate an additional unit of output, represented by the horizontal line M. The price of an output is normalized to be one. Curve DC(i) describes the additional unit costs of innovative activity when i additional units of output are produced. The fact that the curve MC(2) is above MC(1) shows that unit costs become higher when a plant tries to make several leaps in productivity within a given time period. The stickiness of information (see von Hippel, 1994), and the existence of frictions related to learning at the human and organizational level produce this phenomenon.

Figure 1 around here

MC(i) bends upwards because a low productivity unit has much more to learn from others than a high productivity unit. We believe that a plant can learn from other firms whose productivity is at a higher level and which are, also in other ways, within reach of its efforts to increase productivity. The latter requirement refers to the stickiness of technology transfers and its effect on the geographical limits of knowledge potential in the absorption of new technology. Breschi and Lissonin (2001a, 2001b) and Morgan (2004) argue that the tacit, complex and ambiguous nature of transferred information creates significant geographical limits. As a plant approaches the technological frontier, useful and available knowledge for productivity improvements becomes more scarce and MC(i) bends upwards. When a plant innovates, it moves the technological frontier outwards to the position F'. The possibilities for innovating plants to move the technological frontier improve when a plant approaches the technological frontier. The slow learning of humans and organizations explains why it is costly to make a big leap from backwardness to the industry's technological frontier. Therefore, CD(i) is downward-sloping. One could imagine that curve CD(2) would also be above curve CD(1).

The production of new knowledge is profitable in the range that is to the right from the point A. In the range that is to the left from the point C it pays to imitate. Accordingly, in the range between the points A and C it is profitable both to innovate and imitate. This behaviour corresponds to the analysis of Vandenbuschen *et al.* (2006) and their lemma 1. The intersection point A could lie on the right-hand side of the point C, representing a situation that corresponds to the development trap. Firms whose productivity is originally low would then never reach the technological frontier. This kind of reasoning, which puts an emphasis on technology, has become more common in the growth theory that has previously stressed, instead, the importance of incentives to invest in physical and human capital. For example, Feyrer (2003) stresses the central role of the adoption of technology and the creation of new technology for economic growth.[3]

In the empirical part of this study we test the impact of a plant's own and another firm's R&D on the plant's total factor and labour productivity. Furthermore, the productivity effect of R&D that is conducted in the parent firm's other plants is evaluated. The R&D variables of this study are R&D stocks, and all R&D outside the plant considered is weighted according to the geographical proximity. The other firms' R&D stock is hypothesised to be a source of the existing technological knowledge that can be utilised in a plant considered. So, in the framework of Figure 1 MC(i) describes the costs which the utilization of other firms' R&D stock creates. On the other hand, a plant's own R&D stock represents either the potential to absorb the existing knowledge or the total effort that is oriented to produce new knowledge as Cohen and Levinthal (1989) have proposed. The use of a plant's own R&D may be contingent on the distance from the industry's technological frontier. Therefore, the most advanced plants use their own R&D to create new technology and push the technological frontier outwards. According to this, the impact of a plant's own R&D may follow either the lines of curve MC(i) or the curve CD(i) in Figure 1. This feature can easily make the productivity impacts generated by a plant's own R&D a non-linear function of the distance from the industry's technological frontier.

The hypotheses to be empirically tested can be stated as follows:

(i) A plant tends to converge towards the industry's technological frontier,

(ii) A plant's own R&D is expected to have a positive influence on the plant's productivity,

(iii) The productivity impact of the plant's own R&D decreases as the distance from the industry's technological frontier increases, at least when the plant is not far from the technological frontier,

(iv) The impact of R&D in the parent firm's other plants is positive on the plant's productivity,

(v) Other firms' R&D contributes positively to the plant's productivity,

(vi) The productivity impact from other firms' R&D increases as the distance from the industry's technological frontier increases.

Without taking into account the productivity impacts generated by various types of R&D, all plants tend to converge towards the industry's technological frontier. This tendency was also discovered by Vandenbussche *et al.* (2006). Therefore, we propose in the hypothesis (i) that in Finland all firms have rather good possibilities, despite their location and their special field, to use the available information – that is not included in the firms' own R&D stocks – to strengthen their productivity. Further, the plant's own R&D is expected to improve the plant's productivity as hypothesis (ii) states. The fact that Finland is rather close to the global technological frontier in several manufacturing industries (see Scarpetta and Tressel, 2004) provides the motivation for the hypothesis (iii), which also presumes that the plant's own R&D is, on average, used for innovative activity. The hypothesis (iii) allows that a plant – whose productivity is low – uses its own R&D for imitation. For these plants an increase in the distance cannot be expected to reduce the productivity impact of the plant's own R&D. The hypothesis (vi) is based on the idea that external R&D is used to absorb the existing knowledge. In particular, the empirical findings by Vandenbussche *et al.* (2006) encourage us to expect convergence according to this hypothesis.

3. DATA

We use two main sources of data by Statistics Finland over the period 1995-2005. The first one is based on the Annual Industrial Statistics surveys that basically cover all manufacturing plants owned by firms that have no fewer than 20 persons. Output is measured by value added for the purpose of calculating labour and total factor productivity indicators. For the TFP indicator we use capital stock estimates, which are constructed from each plant's past investments through the use of the perpetual inventory method.

The second source of data consists of R&D surveys that incorporate information about R&D expenditures at the firm level and, in addition, the municipality level distribution of the firm level R&D. Using the plant and firm codes of the Annual Industrial Statistics surveys, we generate an algorithm that allocates firm-level R&D expenditures to its plants. The algorithm resembles the one in Lehto (2007). Most firms in the manufacturing sector consist only of one plant, which eases the process. In particular, in the case that the firm has only one plant in a municipality in which the firm has reported that it has pursued R&D activities, the firm's R&D is allocated to this plant. For other plants, we have utilised information about the geographical location of plants and information about the geographical location of R&D expenditures at the municipal level, as recorded in the R&D surveys. In addition, we have taken advantage of industry structure, employees' educational levels and the intended use of R&D expenditures.

We have interpolated the R&D expenditures for those units that are not included in the R&D surveys in all the years. Nominal R&D expenditures are converted to real R&D expenditures by using the average earnings index. The reason for this is that the labour costs of highly educated employees are an important part of overall R&D expenditures. We accumulate R&D stock from the real R&D expenditures by using the same method as Lehto and Lehtoranta (2004). In this calculation we assume the 15 per cent depreciation rate for R&D stock. R&D stock is arguably a better measure for the firm's stock of knowledge, because it is not nearly as volatile as R&D expenditures from year to year. R&D expenditures are almost exclusively allocated to the firm's production sites. Hence, R&D expenditures are not, in most cases, allocated to research laboratories that have entirely specialised in research and development. Despite the fact that the analysis is focused on the production sites, the R&D expenditures of all plants of the firm have been taken into account in the analysis.

4. SPECIFICATION OF THE VARIABLES AND MODELLING APPROACHES

4.1 Productivity

A logarithmic multilateral index for total factor productivity (tpf) – which assumes cost minimization – is calculated according to the principles introduced by Caves *et al.* (1982).[4] This index – in which a plant under consideration is compared with a hypothetical plant in the same (three-digit NACE) industry – is for a plant i in a year t

$$\mathrm{tfp}_{i,t} = \ln(\frac{Q_{i,t}/H_{i,t}}{\overline{Q}/\overline{H}}) - \frac{(S_{i,t}+\overline{S})}{2}\ln(\frac{K_{i,t}/H_{i,t}}{\overline{K}/\overline{H}}), \quad (1)$$

where

$Q_{i,t}$ = value added in real prices

$H_{i,t}$ = labour input in working hours

$K_{i,t}$ = fixed capital in real prices

$S_{i,t}$ = the share of the capital costs of the total costs.

The variables $\overline{Q}$, $\overline{H}$, $\overline{S}$ and $\overline{K}$ denote geometric means for the above variables at the three-digit NACE level. The capital rent $C_{i,t}$ for a plant i in year t is calculated by means of the user cost formula

$$C_{i,t} = P_{i,t} * (R_t + \delta_t - \pi_{i,t}),$$

where

$P_{i,t}$ = the price of capital calculated by Statistics Finland

R = the interest rate for a five-year bond

δ = 0.06 (the depreciation rate for manufacturing industries)[5]

$\pi_i = \log(p_{i,t} / p_{i,t-1})$

The capital costs $UC_{i,t}$ for plant i in year t are then obtained from $UC_{i,t} = C_{i,t} * K_{i,t}$, and so we obtain for $S_{i,t}$

$$S_{i,t} = \frac{UC_{i,t}}{W_{i,t} + UC_{i,t}},$$

where $W_{i,t}$ are the total labour costs.

For the logarithmic labour productivity $lp_{i,t}$ we use the formula

$$lp_{i,t} = \log(\frac{Q_{i,t}}{H_{i,t}}). \quad (2)$$

4.2. R&D variables

The other plants' R&D stock is a source of imitated knowledge. Based on the localised nature of knowledge spillovers, geographical proximity is taken into account in the construction of the variable for other plants' R&D stock. For every plant the variable for the other plants' R&D stock is calculated by using a modified version of the gravity model à la Harris (1954). The parent firm's other plants and other firms' plants are treated separately.

Basically the other plants' R&D stocks are weighted by the inverse of the geographic distance. However, the threshold distance of 10 kilometres is assumed. Without the threshold distance the relative weights would decrease very fast as the distance between plants increases and the R&D that is located in the same commuting area would obtain unrealistic small weight. With the threshold distance the weight coefficient for plant j's R&D stock for plant h is defined as $\frac{1}{d_{hj}+10}$, where d_{hj} is the distance between plants h and j. The measure of distance used is the road distance in kilometres[6] between the municipalities where plant h and plant j are located. For plants located in the same municipality, an internal distance of 7 kilometres is assumed.

The proximity-weighted R&D stock of other firms' plants for a plant h in a firm i is defined to be

$$\mathrm{RDM}_{hj} = \sum_{\substack{k=1 \\ k \neq i}}^{m} \sum_{j=1}^{n} \frac{1}{(10 + d_{hj})}(RDS_{jk}),$$

where

RDS_{jk} = plant j's own real R&D stock in a firm k.

Similarly, for a plant h in a firm i the external R&D stock in the parent firm i's other plants is also proximity-weighted and it is obtained from

$$\mathrm{RDE}_{hi} = \sum_{\substack{j=1 \\ j \neq h}}^{n} \frac{1}{(10 + d_{hj})}(RDS_{ji}),$$

where

RDS_{ji} = plant j's own real R&D stock in a parent firm i.

4.3 Determination of industry's technological frontier

Let $maxtfp_{k,t}$ be the maximum for the logarithmic total factor productivity index (1) in year t in industry k when industries are divided according to the three-digit NACE classification. Suppose that a plant h in a firm i belongs to an industry k. The dynamics of productivity is then supposed to be contingent on the distance of a plant's productivity from its technological frontier. For plant h in firm i the distance considered is

$$dist_{hi,t} = maxtfp_{k,t} - tfp_{hi,t}.$$

4.4 Specifications of the model

Let $TFP_{hi,t} = \exp(tfp_{hi,t})$ and assume that it is determined by

$$TFP_{ih,t} = TFP_{ih,t-1} * \alpha * RDS_{hi,t-1}^{(\beta_1+\beta_2 * dist_{hi,t-1})} * RDM_{hi,t-1}^{(\chi_1+\chi_2 * dist_{hi,t-1})} * \mathrm{RDE}_{\mathrm{hi,t\text{-}1}}^{\eta} * \exp(\sum_{k=1}^{m} \theta_k X_k), \quad (3)$$

where X_k represents other variables: the logarithm of a gross value for a plant's output (the scale variable), export dummy, industry-level dummies and year dummies. Taking logarithms of (3) we obtain for $dtfp_{hi,t}$ ($\equiv tfp_{hi,t} - tfp_{hi,t-1}$) the representation

$$dtfp_{hi,t} = \alpha + \beta_1 rds_{hi,t-1} + \beta_2 croso_{hi,t-1} + \gamma_1 rdm_{hi,t-1} + \gamma_2 cros_{hi,t-1} + \eta rde_{hi,t-1} + \theta_k \sum_{k=1}^{m} X_k , \quad (4)$$

where the small letters refer to the logarithmic values and the notation

$$croso_{hi,t-1} \equiv rds_{hi,t-1} * dist_{hi,t}$$

$$cros_{hi,t-1} \equiv rdm_{hi,t-1} * dist_{hi,t}$$

is used for the interaction terms. Owing to high correlation (0.998) between the distance variable $dist_{hi,t}$ and the interaction variable $cros_{hi,t-1}$ and implied multicollinearity, the variable $dist_{hi,t}$ has been omitted from the analysis. The convergence towards the industry's technological frontier can, however, be evaluated on the basis of coefficient γ_2 and the variation of the variables $rdm_{hi,t-1}$ and $dist_{hi,t}$ in $cros_{hi,t-1}$.

To test in detail how the productivity impact either from the plant's own R&D or from external R&D develops as a function of the distance from the industry's technological frontier we formulate the following variables:

$croso25_{hi,t-1} \equiv rds_{hi,t-1} * d25$

$croso50_{hi,t-1} \equiv rds_{hi,t-1} * d50$

$croso75_{hi,t-1} \equiv rds_{hi,t-1} * d75$,

where

d25 = 1, when dist < 25th percentile of dist. Otherwise d25 = 0.

d50 = 1, when dist ≥ 25th percentile and dist < 50th percentile. Otherwise d50 = 0.

d75 = 1, when dist ≥ 50th percentile and dist < 75th percentile. Otherwise d75 = 0.

Similarly we define

$cros25_{hi,t-1} \equiv rdm_{hi,t-1} * d25$

$cros25_{hi,t-1} \equiv rdm_{hi,t-1} * d50$

$cros25_{hi,t-1} \equiv rdm_{hi,t-1} * d75$.

The nonlinear transformation implemented first standardizes the distance variable to belong to the unit interval in each NACE three-digit industry and then divides the unit interval according to the percentiles defined. On the other hand, in the linear model the distance variable is a logarithmic transformation of the productivity index, whose range is allowed to vary from one industry to another. This may be an advantage as long as the industries are genuinely different but may cause harm in cases when the measurement errors are reflected in the variation over industries.

The non-linear model for total factor productivity can then be written in the form

$$dtfp_{hi,t} = \alpha + \beta_1 rds_{hi,t-1} + \beta_{225} croso25_{hi,t-1} + \beta_{250} croso50_{hi,t-1} + \beta_{275} croso75_{hi,t-1} + \gamma_1 rdm_{hi,t-1} + \gamma_{225} cros25_{hi,t-1} + \gamma_{250} cros50_{hi,t-1} + \gamma_{275} cros75_{hi,t-1} + \eta rde_{hi,t-1} + \theta_k \sum_{k=1}^{m} X_k \,. \quad (5)$$

In the equation (5), for example, the impact of a plant's own R&D in the distance which is below 25th percentile is indicated by $\beta_1 + \beta_{225}$. The coefficient β_1 alone shows how much a plant's own R&D affects productivity when the distance is above the 75th percentile. The interpretation for the other coefficients related to external R&D follows a similar pattern.

In testing the hypotheses one should pay attention to the net effects, too. When, for example, β_1 and β_2 have different signs, the linear model easily specifies for the distance variable $dist_{hi,t}$ a threshold value above which the combined effect $\beta_1 rds_{hi,t-1} + \beta_2 croso_{hi,t-1}$ has a positive productivity impact independently of the value of $rds_{hi,t-1}$. Furthermore, to test the presence of nonlinearities,[7] we estimate a specification (5) which makes it even easier to interpret whether the hypotheses (iii) and (vi) are fulfilled.

The verification of the hypothesis (i) is not straightforward. In the equation (4) it depends both on the signs of the coefficients β_2 and γ_2 and on the variation of the distance variable in the interaction variables $croso_{hi,t-1}$ and $cros_{hi,t-1}$. The other hypotheses introduced above give unambiguous expectations for the signs of the coefficients in equation (4).

- According to the hypothesis (ii), in (4) $\beta_1 > 0$ and the net effect $\beta_1 * rds + \beta_2 * croso$ is, on average, positive. In (5) we expect that the individual effects (β_1, $\beta_1 + \beta_{2j}$, j = 25, 50, 75) are, on average, positive.

- According to the hypothesis (iii), $\beta_2 < 0$ in (4). In (5) we expect that $\beta_{225} > 0$ so that $\beta_{225} > \beta_{250}$ or at least $\beta_{225} > \beta_{275}$.

- According to the hypothesis (iv), $\eta > 0$,

- According to the hypothesis (v), $\gamma_1 > 0$ or, at least, the net effect $\gamma_1 * rdm + \gamma_2 * cros$ is, on average, positive in (4). In (5) we expect that the individual effects (γ_1, $\gamma_1 + \gamma_{2j}$, j = 25, 50, 75) are, on average, positive.

- According to the hypothesis (vi), $\gamma_2 > 0$ in (4) and in (5) $\gamma_{225} < \gamma_{250} < \gamma_{275} < \gamma_2$.

The determination of labour productivity is also examined. Then the distance variable is

$$distl_{hi,t} = maxlp_{k,t} - lp_{hi,t}$$

and the corresponding interaction variables $crosol_{hi,t-1}$ and $crosl_{hi,t-1}$ are obtained from

$$crosol_{hi,t-1} \equiv rds_{hi,t-1} * distl_{hi,t}$$

$$crosl_{hi,t-1} \equiv rdm_{hi,t-1} * distl_{hi,t} .$$

The estimated linear equation is then

$$dlp_{hi,t} = \alpha + \beta_1 rds_{hi,t-1} + \beta_2 crosol_{hi,t-1} + \gamma_1 rdm_{hi,t-1} + \gamma_2 crosl_{hi,t-1} + \eta rde_{hi,t-1} + \theta_k \sum_{k=1}^{m} X_k \; . \quad (6)$$

For the labour productivity we also specify a nonlinear equation which is principally the same as the equation (5).

4.5. Estimation methods

All models are first estimated with OLS. To weaken the impact of the endogeneity of the explanatory variables on the results, endogenous variables are lagged with one year in the estimated equations. Furthermore, one should notice that R&D variables are stocks, which also lessens the seriousness of the possible endogeneity problem.

To tackle the possible endogeneity bias we also use the method of instrumental variables to estimate equations (4) and (6) and apply generalized two stage least squares (G2SLS) estimation with the random effects. The plant's own R&D ($rds_{hi,t-1}$), the interaction variables $croso_{hi,t-1}$ and $cros_{hi,t-1}$ and gross value for a plant's output (lagged with one year) are specified as endogenous variables. As instruments we use the original exogenous variables, which are in (4) and (6) the other firms' R&D ($rdm_{hi,t-1}$), R&D in the parent firm's other plants ($rde_{hi,t-1}$), export dummy,[8] year dummies and industry dummies. Furthermore, endogenous variables lagged with two years – which are the plant's own R&D ($rds_{hi,t-2}$), both the interaction variables (in equation (4) $cros_{hi,t-2}$ and $croso_{hi,t-2}$), the squared variables ($cros_{hi,t-2}*dist_{hi,t-2}$ and $croso_{hi,t-2}*dist_{hi,t-2}$ in equation (4)) and gross value for a plant's output – are used as instruments. The additional instruments, which are not included in (4) and (6), are the capital stock lagged with two years and two industry-structure variables which are the number of plants in other firms in the same three-digit industry and the number of the parent firm's other plants in the same three-digit industry.

In the non-linear model (5) the distance is converted into a dummy variable from the index which belongs to the interval [0, 1]. This makes it rather difficult to endogenize the interaction variables of the plant's own R&D. However, the use of dummies in the specification of the interaction variables decreases the possible correlation between the interaction variables and the error term. Because the total productivity effect of the plant's own R&D depends on its own R&D variable and the interaction variables, one cannot endogenize plant's own R&D variable alone. Therefore, in the 2GSLS estimation we have endogenized only the scale variable: the gross value for a plant's output (lagged with one year).

5. RESULTS

The estimation results for the change in total factor productivity and labour productivity from the linear models are reported in Table 1.[9] It is worth noting that total factor productivity may evolve differently than labour productivity when capital is used to replace labour or when the cost share of capital increases, as there are changes in relative prices. The standard deviation of the labour productivity level is smaller than the standard deviation of the total factor productivity level. This may partly reflect problems that are associated with the accurate assessment of capital input and user cost. Despite this, the main effects of interest are more or less the same for both total factor productivity and labour productivity.

Table 1 around here

The plant's own R&D – its direct and indirect effects – has a positive and statistically significant impact, on average, on total factor productivity in the specifications of Table 1. The same concerns the effect of the plant's own R&D on labour productivity in the OLS specification in Table 1. For example, in the OLS model for total factor productivity in Table 1 the net effect of the plant's own R&D – which also takes into account the impact of the interaction variable – is positive for almost all values of the distance variable ($dist_{hi,t}$). Only for the most inefficient plants, whose distance is above the 94th percentile, is the net effect in question negative.

The quantitative magnitude of the estimated direct effects of the plant's own R&D seems to be rather moderate at first sight. For example, the coefficient of the plant's own R&D is 0.008 (Table 1, Column 1). This means that as R&D increases by 1 per cent it increases the growth rate of total factor productivity by 0.008 percentage points. However, one has to bear in mind that R&D's share of the firm's total costs is, on average, small and therefore there is a large variation in the plants' R&D stocks. In particular, for over half of all plants the R&D stock is zero and for some other plants it is very large. The huge percentual increases in R&D expenditures and even in the R&D stock are not uncommon. For example, doubling the R&D stock increases the growth rate of total factor productivity by 0.8 percentage points. That is not a small change, as the mean of the TFP growth in data is 3.8%.

We also discover that the effect of the plant's own R&D on total factor productivity and labour productivity decreases as the plant deviates from the industry's technological frontier. This pattern holds in the OLS models. Furthermore, we find that the parent firm's proximity-weighted R&D stock in its other plants is positive and statistically significant in the linear models for total factor productivity and in the OLS model for labour productivity.

According to the estimation results, other firms' proximity-weighted R&D stock does not differ statistically from zero in the linear models of Table 1. However, the indirect effect of other firms' proximity-weighted R&D stock – being conditioned on the distance from the technological frontier – tends to be positive and statistically significant. This pattern is robust, because it prevails in all models in Table 1. We discover that other firms' R&D stock increases productivity when a plant is located far away from the industry's technological frontier. This confirms the hypothesis (vi).

The results for the nonlinear models are reported in Table 2. The coefficient for a plant's own R&D and for the respective interaction variables reveal that a plant's own R&D's impact on the total factor productivity and labour productivity is greatest when the plant is located close to the industry's technological frontier (the distance from it is below the 25th percentile). This effect seems to weaken when a plant is located far away from the technological frontier. In particular, this pattern seems to be clear for the total factor productivity. In the models for labour productivity (Table 2, Columns 2 and 4) the productivity impact concerned is also greatest for the most efficient units and smallest for the most inefficient plants (the coefficient of plant's own R&D variable alone). These results verify the hypothesis (iii).

Table 2 around here

Other firms' R&D affects productivity in accordance with the hypothesis (vi) in the nonlinear models of Table 2. The external R&D's impact on the total factor productivity and labour productivity is largest for the inefficient plants. The effect is diluted when a plant becomes more efficient and it is roughly zero for the most efficient plants.

Table 3 (Panel A) describes the effect of other firms' R&D on the growth rate of the total factor productivity using the coefficient of 0.0044 for the CROS variable (Table 1, Column 3). Given a plant's distance from the technological frontier, the increase in other firms' R&D stock from the minimum of the industry to the maximum of the industry increases the growth rate of total factor productivity by roughly 2 percentage points when the distance from the industry's technological frontier is high (i.e. a plant is located in the 80th percentile). The respective effect becomes smaller when the distance shortens. Given a plant's distance from the industry's technological frontier, the differences from the minimum value to the maximum value or from the minimum value to the average value in the other firms' R&D produce, in any case, a significant positive effect on the growth rate of the total factor productivity. Therefore, the spillover effects can be of a considerable size.

Table 3 around here

Table 3 (Panel B) illustrates the effect of a plant's distance from the industry's technological frontier on the growth rate of total factor productivity using the coefficient of 0.0044 for the CROS variable. The results reveal that given the other firms' R&D level, the distance from the industry's technological frontier has a substantial positive influence on productivity, independent of the level of other firms' R&D. This is the convergence effect in productivity. It shows that the inefficient plants in all regions tend to convergence towards the productivity level of the efficient plants. This confirms, through the use of plant-level data, the results presented in Vandenbussche *et al.* (2006). From Table 3 (Panel B) we also see that a big leap in the distance can create an approximately 0.5-1.5 percentage points larger productivity impact in the regions – where other firms' R&D is concentrated (80th percentile row) compared with the regions that lack other firms' R&D (20th percentile row).

We must also notice that the interaction variable CROSO generates an opposite effect on the convergence tendency discussed above, for example, in the OLS model in the first column of Table 1, where the coefficient of CROSO is negative. This effect dilutes the power of the convergence tendency but does not change its overall direction.

6. THE ROBUSTNESS OF THE RESULTS AND GEOGRAPHICAL PROXIMITY

The high correlation of 0.998 between the original distance variable ($dist_{hi,t}$) and the interaction variable ($CROS_{hi,t}$) leads to biased estimates and was thus the reason to omit the distance variable from the linear OLS regressions. For the same reason, the distance dummies d125, d150 and d175 were not included in the non-linear regression for the total factor productivity. The same applies, of course, to the estimation of labour productivity models. If the interaction variable for the external R&D were replaced by the pure distance variables ($dist_{hi,t}$ or dummies d125, d159 and d175), the distance from the industry's technological frontier would have a statistically significant positive impact on the change in productivity and the external R&D's productivity impact would also be positive and of the same size as its impact is, on average, in the models that are reported in Tables 1-2. If, on the other hand, the distance variables were inserted in the models reported in Tables 1-2 the interaction variable for the external R&D would become statistically insignificant. This leaves us somewhat uncertain about the actual conditionality of the external R&D's productivity impact on the distance from the industry's technological frontier.

We also experimented with replacing the original external R&D variable – which gives greater weight for R&D locating close – by the R&D stock which gives greater weight for R&D locating far away. The new variable is actually an R&D aggregate for all the other firms minus the original external R&D variable. After this replacement the productivity impact of the new external R&D variable turned out to be negative and the declining pattern of the impact as a function of the distance from the industry's technological frontier broke down. Therefore, for the most inefficient plants the impact was no longer the largest. This experiment demonstrates that geographically-determined weights in the external R&D make sense and that the original external R&D variable does not break the productivity pattern which the distance variable alone generates. These findings can be interpreted to indicate that the pattern mentioned in the hypothesis (vi) is just a part of the overall convergence tendency mentioned in the hypothesis (i), according to which the most inefficient plants tend to converge towards the industry's technological frontier, which indicates that they have more potential to absorb all kinds of existing knowledge than the other plants. The other firms' R&D stock that is located close by can then be interpreted as being only a part of this larger knowledge base to which R&D stock that is located far away does not belong.

The relevance of geographically-determined weights was also investigated by replacing the original R&D variable for the parent firm's other plants by the variable which no longer weights the other plant's R&D according to their geographical location. According to the results, the original coefficient for this variable, which is 0.0026 (Table 1, Column 1), turned out to be 0.0015, and 0.0020 (Table 1, Column 3) turned out to be 0.0011. Being geographically close also seems to be relevant within multi-plant firms.

7. CONCLUSIONS

This paper examines, through the use of plant-level data, whether R&D's productivity impact is contingent on the distance of a plant's productivity from the industry's technological frontier. R&D is specified as an accumulated stock from R&D investments. We analyse the productivity effect of a plant's own R&D as well as the productivity impact of the plant's parent firm's and other firms' proximity-weighted R&D stocks. The results show that a plant's own and a parent firm's R&D have a positive productivity impact and that the former impact decreases as the distance from the industry's technological frontier increases. Furthermore, the productivity effect of other firms' proximity-weighted R&D is, on average, positive, but this impact seems to increase in the distance from the technological frontier. Another important finding is that all the plants tend to converge towards the industry's technological frontier despite the size of external R&D spillovers.

ACKNOWLEDGEMENTS

This paper is a part of the project "Regional productivity growth in Finland" (Tuottavuuden alueellinen kasvu Suomessa) financed by TEKES (National Technology Agency of Finland).

The matched data that are used to produce the results can be accessed on site at the Research Laboratory of the Business Structures Unit of Statistics Finland. Paul A. Dillingham has kindly checked the English language. The usual disclaimer applies.

REFERENCES

Acemoglu, D., P. Aghion and F. Zilibotti (2006), 'Distance to frontier, selection and economic growth,' *Journal of European Economic Association,* **4**, 37-74.

Breschi, S. and F. Lissoni (2001a), 'Localised knowledge spillovers versus innovative milieux: knowledge "tacitness" reconsidered,' *Papers in Regional Science,* **80**, 255-273.

Breschi, S. and F. Lissoni (2001b), 'Knowledge spillovers and local innovative systems: a critical survey,' *Industrial and Corporate Change,* **10**, 975-1005.

Caves, D., L. Christensen and W. Diewert (1982), 'Multilateral comparisons of output, input and productivity using superlative index numbers,' *Economic Journal,* **92**, 73-86

Cohen, W. and D. Levinthal (1989), 'Innovation and learning: the two faces of R&D,' *The Economic Journal,* **99**, 569-596.

Feyrer, J. (2003), 'Convergence by parts,' Manuscript, Dartmouth College, (Available at http://www.dartmouth.edu/~jfeyrer/parts.pdf).

Girma, S. (2005), 'Absorptive capacity and productivity spillovers from FDI: A threshold regression analysis,' *Oxford Bulletin of Economics and Statistics,* **67**, 281-305.

Griffith, R., S. Redding and J. Van Reenen (2004), 'Mapping the two faces of R&D: Productivity growth in a panel of OECD industries,' *Review of Economics and Statistics,* **86**, 883-895.

Harris, C. (1954), 'The market as a factor in the localization of industry in the United States,' *Annals of the Association of American Geographers*, **64**, 315-348.

Ilmakunnas, P. and M. Maliranta (2004), 'Foreign medicine: a treatment effect analysis of the productivity effects of foreign ownership,' *Applied Economics Quarterly,* **50**, 41-59.

Lehto, E. (2007), 'Regional impact of research and development on productivity,' *Regional Studies,* **41**, 623-638.

Lehto, E. and O. Lehtoranta (2004), 'Becoming an acquirer and becoming acquired,' *Technological Change and Social Change,* **71**, 635-650.

Morgan, K. (2004), 'The exaggerated death of geography: learning, proximity and territorial innovation systems,' *Journal of Economic Geography,* **4**, 3-21.

Nadiri, M. and I. Prucha (1996), 'Estimation of the depreciation rate of physical and R&D capital in the U.S. total manufacturing sector,' *Economic Inquiry,* **34**, 43–56.

Orlando, M. (2004), 'Measuring spillovers from industrial R&D: on the importance of geographical and technological proximity,' *RAND Journal of Economics,* **35**, 777-786.

Ottaviano, G. and J-F Thisse (2001), 'On economic geography in economic theory: increasing returns and pecuniary externalities,' *Journal of Economic Geography,* **1**, 153-179.

Rouvinen, P. (2002), 'R&D-productivity dynamics: causality, lags and 'dry holes',' *Journal of Applied Economics,* **5**, 123-156.

Scarpetta, S. and T. Tressel (2004), 'Boosting productivity via innovation and adoption of new technologies: any role for labour market institutions?,' Working Paper No. 3273, World Bank.

Scitovsky, T. (1954), 'Two concepts of external economies,' *Journal of Political Economy*, **62**, 143-151.

Vandenbussche, J., P. Aghion and C. Meghir (2006), 'Growth, distance to frontier and composition of human capital,' *Journal of Economic Growth,* **11**, 97-127.

Von Hippel, E. (1994), 'Sticky information and the locus of problem solving: implications for innovation,' *Management Science,* **40**, 429-439.

Figure 1. An illustration of convergence to the technological frontier.

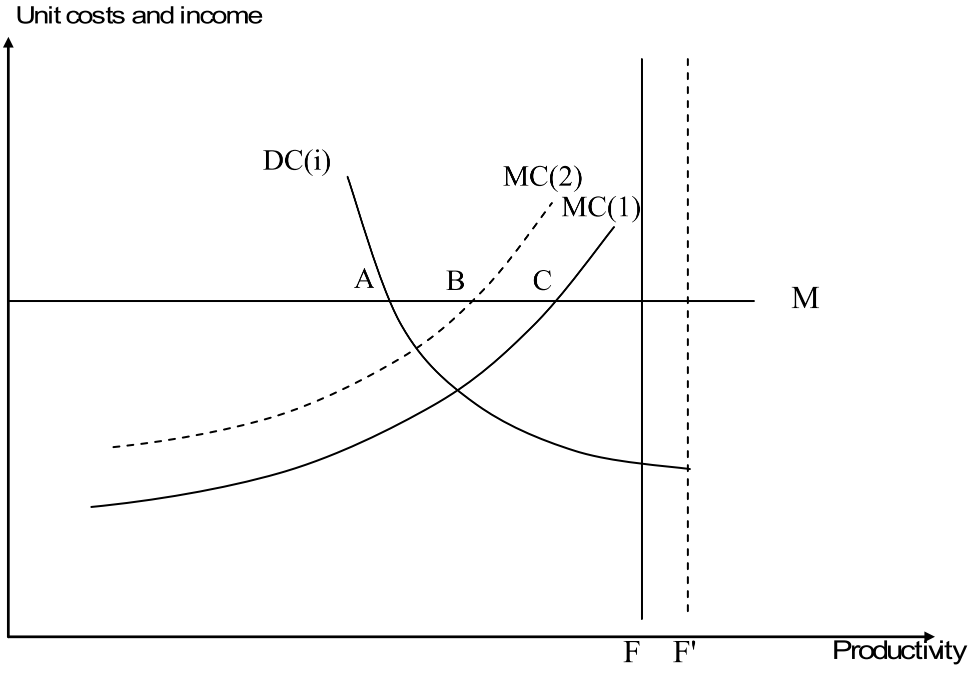

Table 1. The effect of R&D on the change in a plant's total factor productivity (Dtfp) and labour productivity (Dlp) in the linear model.

	OLS		G2SLS, Random effects	
	Total factor productivity	Labour productivity	Total factor productivity	Labour productivity
Own R&D$_{t-1}$	0.0076***	0.0061***	0.0033***	-0.0002
	(0.0009)	(0.0007)	(0.0012)	(0.0011)
Own R&D*distance$_{t-1}$	-0.0032***	-0.0044***	-0.0010	0.0008
	(0.0005)	(0.0005)	(0.0007)	(0.0009)
R&D in parent firm's other plants $_{t-1}$	0.0026***	0.0026***	0.0020**	0.0010
	(0.0008)	(0.0006)	(0.0009)	(0.0007)
Other firms' R&D$_{t-1}$	0.0040	-0.0061	-0.0017	-0.0018
	(0.0058)	(0.0038)	(0.0062)	(0.0049)
Other firms' R&D*distance$_{t-1}$	0.0077**	0.0111***	0.0044***	0.0024***
	(0.0003)	(0.0003)	(0.0004)	(0.0005)
Gross value of plant's output$_{t-1}$	-0.0206***	0.0045***	-0.0069*	0.0054*
	(0.0033)	(0.0022)	(0.0036)	(0.0029)
Export dummy	0.0746***	0.0490***	0.0467***	0.0141*
	(0.0094)	(0.0060)	(0.0101)	(0.0081)
Year dummies	Yes	Yes	Yes	Yes
Industry-level dummies	Yes	Yes	Yes	Yes
R-sq within			0.093	0.0811
R-sq between			0.035	0.0249
R-sq overall			0.077	0.0508
R-sq adjusted	0.0982	0.0843		
Number of observations	17 886	23 750	14 810	15 083

Notes: Standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%.

Table 2. The effect of R&D on the change in a plant's total factor productivity (Dtfp) and labour productivity (Dlp) in the non-linear model.

	OLS		G2SLS, Random effects	
	Total factor productivity	Labour productivity	Total factor productivity	Labour productivity
Own R&D$_{t-1}$	0.0006 (0.0011)	-0.0038 (0.0006)	-0.0002 (0.0009)	-0.0012* (0.0007)
Own R&D*distance$_{t-1}$ (with distance below 25th percentile)	0.0044*** (0.0016)	0.0034*** (0.0009)	0.0048*** (0.0010)	0.0040*** (0.0008)
Own R&D*distance$_{t-1}$ (with distance above 25th percentile and below 50th percentile)	0.0025* (0.0015)	0.0011 (0.0010)	0.0018 (0.0014)	0.0005 (0.0011)
Own R&D*distance$_{t-1}$ (with distance above 50th percentile and below 75th percentile)	0.0011 (0.0016)	0.0020** (0.0009)	0.0016 (0.0014)	0.0018* (0.0011)
R&D in parent firm's other plants $_{t-1}$	0.0041*** (0.0008)	0.0034*** (0.0006)	0.0034*** (0.0009)	0.0019*** (0.0007)
Other firms' R&D$_{t-1}$	0.0226*** (0.0059)	0.0104*** (0.0039)	0.0144** (0.0061)	0.0102** (0.0050)
Other firms' R&D*distance$_{t-1}$ (with distance below 25th percentile)	-0.0205*** (0.0008)	-0.0104*** (0.0005)	-0.0180*** (0.0006)	-0.0099*** (0.005)
Other firms' R&D*distance$_{t-1}$ (with distance above 25th percentile and below 50th percentile)	-0.0133*** (0.0008)	-0.0042*** (0005)	-0.0122*** (0.0008)	-0.0044*** (0.0006)
Other firms' R&D*distance$_{t-1}$ (with distance above 50th percentile and below 75th percentile)	-0.0091*** (0.0008)	-0.0023*** (0.0004)	-0.0087*** (0.0008)	-0.0019*** (0.0006)
Gross value of plant's output$_{t-1}$	-0.0065** (0.0033)	-0.0005 (0.0022)	0.0029 (0.0036)	0.0077** (0.0030)
Export dummy	0.0464*** (0.0094)	0.0224*** (0.0061)	0.0314*** (0.0099)	0.0104 (0.0081)
Year dummies	Yes	Yes	Yes	Yes
Industry-level dummies	Yes	Yes	Yes	Yes
R-sq within			0.1353	0.0574
R-sq between			0.0481	0.0195
R-sq overall			0.0943	0.0366
R-sq adjusted	0.0917	0.0359		
Number of observations	17 886	23 750	14 905	15 139

Notes: Standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%.

Table 3. Panel A shows the effect of other firms' R&D on the growth rate of total factor productivity, given a plant's distance from the industry's technological frontier and using the coefficient of 0.0044 for the CROS variable. Panel B shows the effect of a plant's distance from the technological frontier on the growth rate of total factor productivity, given the other firms' level of R&D and using the coefficient of 0.0044 for the CROS variable. The estimates are presented as averages of all plants, percentage points. (Percentiles, the minimum and maximum values and averages are calculated yearly from each three-digit NACE.).

Panel A:

Distance from the industry's technological frontier	Other firms' R&D, the difference between maximum and minimum	Other firms' R&D, the difference between average and minimum	Other firms' R&D, the difference between 80th percentile and 20th percentile
20th percentile	1.01	0.48	0.40
50th percentile	1.43	0.69	0.58
80th percentile	2.02	0.90	0.78

Panel B:

Other firms' R&D	Distance, the difference between maximum and minimum	Distance, the difference between average and minimum	Distance, the difference between 80th percentile and 20th percentile
20th percentile	23.96	10.03	6.36
50th percentile	24.67	10.32	6.54
80th percentile	25.38	10.61	6.73

Appendix 1

Descriptive Statistics

	Number of observations	Mean	Standard deviation	Minimum value	Maximum value
Dtfp	15 030	0.038	0.450	-4.236	3.869
Dlp	15 230	0.027	0.351	-9.822	9.547
Dist	15 186	1.420	1.162	0	20.908
Distl	15 692	1.095	0.695	0	9.360
Croso	15 186	8.961	13.662	0	153.75
Crosol	15 692	6.573	13.663	0	77.505
Square of croso	15 186	266.964	1045.257	0	23639.06
Square of crosol	15 692	116.631	244.958	0	6007.034
Rde	15 800	3.652	4.713	0	17.889
Rdm	15 800	17.317	0.692	14.812	18.961
Cros	15 186	24.524	19.915	0	388.377
Crosl	15 692	18.959	12.051	0	170.923
Square of cros	15 186	997.996	2679.483	0	150836.5
Square of crosl	15 692	504.655	685.333	0	29214.55

Appendix 2

OLS models for the instrumented variables in the models of Table 1 (after taking logarithm of R&D variables, gross value of a plant's output, fixed capital and the number of firms).

	Own R&D$_t$	Own R&D* distance$_t$	Other firms' R&D *distance$_t$	Gross value of plant's output$_t$
Own R&D$_{t-1}$	0.9688***	0.0948***	-0.2948***	0.0019***
	(0.0031)	(0.0191)	(0.0325)	(0.0007)
Own R&D*distance$_{t-1}$	0.0014	1.0699***	0.3126***	-0.0006
	(0.0022)	(0.0136)	(0.0232)	(0.0005)
Square of own R&D *distance$_{t-1}$	-2.3e-05	-0.0061***	-0.0053***	1.74e-06
	(2.1e-05)	(0.0001)	(0.0002)	(4.6e-06)
R&D in parent firm's other plants$_t$	0.0151***	0.0310*	-0.0160	0.0008
	(0.0028)	(0.0175)	(0.0299)	(0.0006)
Other firms' R&D$_t$	0.0165	-0.3766***	-0.3865**	-0.0016
	(0.0167)	(0.1036)	(0.1767)	(0.0038)
Other firms' R&D*distance$_{t-1}$	0.0007	-0.0438***	0.6431***	0.0013***
	(0.0010)	(0.0064)	(0.0109)	(0.0002)
Square of other firms' R&D*distance$_{t-1}$	-1.1e-06	9.9e-05***	-0.0007***	-1.3e-06
	(3.7e-06)	(2.3e-06)	(3.9e-05)	(8.2e-07)
Gross value of plant's output$_{t-1}$	0.0887***	0.1706*	-1.0321***	0.9690***
	(0.0139)	(0.0868)	(0.1482)	(0.0031)
Export dummy	0.1056***	-0.5265***	-2.0135**	0.0496***
	(0.0273)	(0.1694)	(0.9363)	(0.0061)
Fixed capital$_{t-1}$	0.0151	0.5613***	1.8881***	0.0081***
	(0.0105)	(0.0658)	(0.1123)	(0.0024)
Number of other firms' plants in own 3-digit industry	0.0129	0.7657***	1.9892***	-0.0030
	(0.0121)	(0.0755)	(0.1288)	(0.0027)
Number of own firm's other plants in own 3-digit industry	-0.0965***	0.2771**	1.6783***	-0.0174***
	(0.0211)	(0.1315)	(0.2243)	(0.0047)
Year dummies	Yes	Yes	Yes	Yes
Industry-level dummies	Yes	Yes	Yes	Yes
R-sq adjusted	0.9581	0.6513	0.5396	0.9546
Number of observations	17 798	17 566	17 566	17 793

Notes: Standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%.

[1] Scitovsky (1954), and Ottaviano and Thisse (2001) define externalities in this way.

[2] In particular, Griffith *et al.* (2004) provide evidence by using a panel of industries across twelve OECD countries that stresses the importance of R&D in increasing possibilities to technology transfers through the build-up of absorptive capacity.

[3] Later Howitt (2002), who considered the growth to be contingent on the distance from the technological frontier, obtained the result according to which the economies may settle down into three different stationary equilibriums. The economies that were originally not so advanced will never converge to the technological frontier.

[4] See also Ilmakunnas and Maliranta (2004).

[5] This is roughly the same as the estimated depreciation rate for the fixed capital in U.S. manufacturing (= 0.059) in Nadiri and Prucha (1996).

[6] The road distance data originates from the Finnish Road Administration. It is the distance between the economic centres of municipalities on main roads.

[7] Estimating quadratic model or using threshold regression techniques, Girma (2005) discovered nonlinear threshold effects. In the quadratic model the interaction between FDI in the region and absorptive capacity (the distance from the technological frontier) had a nonlinear U-shaped impact on the output.

[8] The export dummy does not usually change its value for a given firm over the period 1995-2005.

[9] Descriptive statistics for the variables are documented in the Appendix 1. The estimation results from the first stage regression are reported in the Appendix 2.