ETLA

ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B 00120 Helsinki Finland Tel. 358-9-609 900
Telefax 358-9-601 753 World Wide Web: http://www.etla.fi/

Keskusteluaiheita – Discussion papers

No. 1106

Esteban Fernández Vázquez* – Bart Los**

A MAXIMUM ENTROPY APPROACH TO THE IDENTIFICATION OF PRODUCTIVE TECHNOLOGY SPILLOVERS

* University of Oviedo (Department of Applied Economics); email: evazquez@uniovi.es

** University of Groningen (Faculty of Economics and Business & Groningen Growth and Development Centre); email: b.los@rug.nl

This paper is part of the project “Intangible Capital and Economic Growth” funded by TEKES, the Finnish Funding Agency for Technology and Innovation (project nr. 10169/25/06). Raportti on julkaistu myös Pellervon taloudellisen tutkimuslaitoksen työpaperina (nro. 101, ISBN 978-952-5594-72-0).

Acknowledgment:
The authors thank Edwin Stuivenwold for his able assistance in constructing the R&D data. Participants at the 16th International Input-Output Conference (Istanbul, July 2007), the Workshop on Intangible Capital and Economic Growth (Helsinki, June 2007) and a seminar at the University of Oviedo (September 2007) are gratefully acknowledged for their suggestions.

ISSN 0781-6847 08.11.2007

FERNÁNDEZ VÁZQUEZ, Esteban – **LOS,** Bart, **A MAXIMUM ENTROPY APPROACH TO THE IDENTIFICATION OF PRODUCTIVE TECHNOLOGY SPILLOVERS.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 2007, 21 p. (Keskusteluaiheita, Discussion papers, ISSN 0781-6847; No. 1106).

ABSTRACT: R&D activities by one industry often have positive effects on the productivity performance of other industries, as a consequence of technology spillovers. Econometric problems (such as multicollinearity), however, have prevented researchers from identifying the industries that have been responsible for the most important technology spillovers. This paper proposes an alternative estimation approach (Generalized Maximum Entropy econometrics), which can cope with datasets characterized by a high degree of multicollinearity. For a number of industries, rates of return to R&D expenditures by other industries are estimated on a bilateral basis. Furthermore, productivity effects of spillovers from the foreign counterparts of the industry are estimated. The analysis is done for eighteen industries in twelve OECD countries in the period 1976-1999.

1. Introduction

Knowledge has some characteristics of a public good. It is partly nonrival and partly non-excludable, which implies that it can give rise to externalities. In mainstream theory, externalities often call for public policy. If the externalities are mainly positive, governments should take care of additional supply of the public good. Since most theories stress the positive externalities of knowledge, its purposeful production (by means of R&D activities) should be stimulated. Since the extent and nature of R&D activity varies considerably across industries, policy effectiveness would be helped considerably if the industries that generate the most important externalities (or, 'spillovers') could be singled out. Despite the by now vast empirical literature on this topic, one cannot but observe that this identification objective has still not been attained. This paper proposes a less traditional approach, to come closer to the production of a matrix that indicates the productivity effects industries experience as a consequence of R&D activities done in each of the remaining industries. Productivity effects of foreign counterparts will also be estimated.

Due to econometric problems (i.e. multicollinearity), empirical research into productivity effects has so far relied on composite spillover variables. Constructing such variables involves the definition of a weighting scheme, to approximate the relevance of industry-specific contributions to the industry under consideration. Including relevance weights is important, since it is implausible to assume that the electronics industry will enjoy similar benefits from a euro spent on R&D in the computer industry to those from a euro spent on R&D in the furniture industry. Several weighting schemes have been proposed, however, based on different channels of technology flows. Studies like Los and Verspagen (2000) compared results for a couple of such composite variables to find out which type of spillovers would have the most prominent effects, but did not find very strong results. Keller (1997, 1998) went much further, by arguing that the theoretically underpinned composite variables do not perform any better than composite variables based on randomly chosen weights. Although Keller's results were not left uncriticized, a very inconvenient situation emerged: almost all studies (see Nadiri, 1993, and Mohnen, 1996, for early surveys) agree that technology spillovers have substantial positive effects on productivity, but it is impossible to assess which industries are best at "radiating" productive spillovers and whether the most important spillovers are of the rent spillover or the knowledge spillover kinds.

This paper attempts to shed new light on the discussion, by adopting a non-classical regression approach, which does not suffer from the problems that caused researchers using classical regression analysis to use composite spillover variables. Generalized Maximum Entropy (GME) econometrics can deal with multicollinearity in data (see the excellent introduction by Golan *et al.*, 1996). We applied our GME analysis on data for 12 developed countries, for the period 1976-1999. The data on industry-level value added growth and labor inputs were taken from the very recent EUKLEMS (2007) database. OECD's STAN-ANBERD dataset was used as the source for the industry-level R&D data. Our analysis cannot tell which spillover channels have had the strongest impact, but it gives indications about the main suppliers of technology spillovers for each of the 18 manufacturing industries for which we run the analysis.

The paper is organized as follows. Section 2 reviews the general setup of studies into the productivity effects of technology spillovers and discusses the current state of affairs. In Section 3, we give an introduction in the intuition behind GME estimation and present the equations we will estimate using GME techniques. Section 4 is devoted to a brief discussion of the data, after which the estimation results are presented in Section 5. Section 6 concludes.

2. A Brief Non-Chronological History of Spillover Effects Estimation

Since the early 1960s, many studies have tried to estimate the empirical importance of technology spillovers for productivity growth. Generally, these productivity studies start from a production function, most often an extended Cobb-Douglas specification. Not only the traditional production factors physical capital and labor are included, but also two kinds of R&D stocks: R&D investments by the unit (firm, industry, region or country) itself and R&D obtained through spillovers from other units (so-called indirect R&D). If we denote the former by *R* and the latter by *IR*, the production function looks like

$$Q_{jt} = A(IR)^{\eta}_{jt} K^{\alpha}_{jt} L^{\beta}_{jt} R^{\gamma}_{jt} \tag{1}$$

Q stands for value added, *A* is a constant, *K* indicates the stock of physical capital, *L* denotes employment, *t* is the time index and *j* is the unit index. The elasticities η, α, β and γ can be estimated, if sufficient observations on each of the variables are available. Alternatively, β can be measured as the labor share in total income (this approach is commonly known as 'growth accounting'). If constant returns to scale with respect to capital and labor are imposed, α equals 1-β. In this way, a measure for total factor productivity (TFP) growth is obtained, and this can be related to the changes in both R&D stocks.[1] Both approaches yield estimates for output elasticities with respect to indirect R&D, $(dQ/dIR)\cdot(IR/Q)$, or rates of return to indirect R&D, dQ/dIR. These are considered to be measures for the impact of spillovers. As explained by Van Meijl (1995), estimating a common rate of return is often less data-demanding than estimating a common elasticity. Under the (admittedly strong) assumption that R&D stocks are not subject to depreciation, rates of return can be estimated by linking total factor productivity growth to R&D intensities, defined as RE/Q and IRE/Q (*E* indicates expenditures).[2]

---

1 A third approach is to use the dual of the production function, i.e., the cost function. Changes in the costs per unit of output are regressed on changes in the prices and quantities of various inputs (see Bernstein and Nadiri, 1988).

2 This procedure is sometimes referred to as the 'Terleckyj transformation', after Terleckyj (1974).

In principle, the simplest way to estimate the influence of R&D efforts in other industries is the one applied by Bernstein and Nadiri (1988). They specify one indirect R&D variable for each of the (other) industries. For example, the decrease in unit costs in the U.S. chemical industry is related to the R&D expenditures of the industries that manufacture non-electrical machinery, electrical products, transportation equipment and scientific instruments. This approach lets the data speak for themselves to see which (other) industries influence the productivity of a particular industry. The method has one important drawback: most industry R&D budgets have risen during the last decades and are relatively high for the same set of countries, which causes huge multicollinearity problems. The method we propose below could be seen as a way of following up to the lead by Bernstein and Nadiri (1988), using an alternative regression technique.

Since classical regression analysis is not suitable to solve the problems encountered by Bernstein and Nadiri, many authors have proceeded along an alternative avenue of research. They continued in the way proposed much earlier already by Terleckyj (1974), using weights to construct aggregate indirect R&D investment variables (*IRE*):

$$IRE_j = \sum_i \omega_{ij} RE_i \qquad \forall i \neq j \tag{2}$$

In this expression, *i* and *j* denote the 'spillover producing' and 'spillover receiving' units, respectively. The weights ω_{ij} are the crucial elements distinguishing the different approaches to measuring spillovers. They indicate to what extent the R&D undertaken by *i* may be considered to be part of the indirect R&D expenditures of *j*. A number of weighting schemes have been proposed. We will describe them briefly (see Los and Verspagen, 2007, for much more detailed discussions).[3]

Unit Weights

In his firm level study emphasizing the effects of intraindustry spillovers, Bernstein (1989) circumvents the weighting problem by setting all weights equal to one. So did Los and Verspagen (2000) in their attempt to evaluate the empirical performance of four different interindustry spillover measures. The most important disadvantage of this method is that no account is taken of the theory of spillovers, which argues that due to differences in technological opportunities, appropriability of knowledge, differences in trade intensities among industries etc., the weights should in fact be very heterogeneous.

Weights Based on Transaction Input or Output Shares

Early attempts to include spillovers in productivity analysis at the industry level (Terleckyj, 1974) used trade statistics to construct industry weights ω_{ij}. Input-output tables are converted into

[3] See Griliches (1979, 1992) for classic contributions on channels through which innovations in one industry can affect the (sometimes misperceived) productivity performance of other industries. Van Pottelsberghe (1997) expresses views that are not in every sense in line with Los and Verspagen's (2000, 2007) opinions.

tables of output coefficients. Such coefficients indicate the share of industry *i*'s output delivered to industry *j*. Next, R&D weights are set equal to the output coefficients, except for the diagonal elements. Terleckyj also calculated similar output coefficients from capital flow matrices to account for interindustry investment flows. In this output shares approach, 'second-round' effects might also be important. This occurs when spillovers are transmitted to industries down the production chain, for example, when advances in semi-conductors spill over to the computer industry, and from there to the banking industry (see, e.g. Sakurai *et al.*, 1997).

Input-output tables are also used to compute spillover measures in which the ω_{ij}s are defined as the input coefficients a_{ij}. Wolff (1997), among others, used this measure in an interindustry context. In their highly influential international spillover study, Coe and Helpman (1995) construct a similar measure (using import weights). A disadvantage of these approaches is that only trade-related knowledge flows are taken into account. It is well-known that several other channels provide opportunities for technology spillovers.

Weights Based on Patent and Innovation Output Shares

Scherer (1982) pioneered another approach, because he felt that economic transactions often do not entail exchange of technology. A procedure based on true technological data should be used. First, he assigned a sample of patents granted in a certain period to an industry-of-origin, i.e., the producer of the technology described in the patent. Next, all patents were assigned to one or more industries-of-use, on the basis of information in the patent document.[4] Finally, output shares were computed in a way directly comparable to the way output coefficients are constructed for traditional input-output tables based on economic transactions.

Numbers of innovations could be used as an alternative for patent counts. Sterlacchini (1989) used a large innovation survey undertaken by Robson *et al.* (1988). In this survey, innovations were assigned to an industry-of-origin (or industry-of-manufacture) and an industry-of-use. Next, he used this 'innovations input-output table' to calculate innovation share weights ω_{ij}, denoting the share of innovations of industry *i* used by industry *j*. DeBresson *et al.* (1994) followed this lead. A disadvantage of both approaches is that the focus is on innovations traded between industries, usually embodied in product innovations. Knowledge flows not related to economic transactions are not considered. In this sense, the main disadvantage of input-output based weights is not addressed by these methods.

Weights Based on Patent Information Output Shares

Verspagen (1997a) derived different spillover measures from patent office documents. Using a concordance that maps patent classification codes onto manufacturing industry classes,

---

[4] Johnson and Evenson (1997) proposed a concordance that maps patent classification codes assigned by the Canadian Patent Office onto industry codes, which enabled them to construct their matrix without the need to examine every patent document individually.

Verspagen derived the industry most likely to have produced the knowledge described in the patent document, and the industries that have been most likely to benefit from this knowledge (not the patented product itself).[5] This yielded a 'patent information input-output table' similar in format to the ones described above. The ω_{ij}s were then, set equal to the output coefficients of this table.

Verspagen constructed a second type of patent information input-output tables using patent citations. The patent citation output share weights method has the disadvantage that it relates to a very specific channel of spillovers and implicitly assumes that each cited patent is equally relevant to the spillover receiver.

Weights Based on Technological Proximity

The first spillover measure explicitly focusing on non-traded knowledge spillovers was constructed by Jaffe (1986). He argued that knowledge generated by R&D investments flows into a 'spillover pool', which is accessible to all firms. Some firms or industries benefit more from firm *i*'s contribution to the pool than others, because not all knowledge is relevant to their R&D. To measure the part of the contribution of the *i*th firm that is relevant to firm *j*, Jaffe (1986) used a 'technological proximity' measure:

$$\omega_{ij} = \frac{\sum_{k=1}^{F} f_{ik} \cdot f_{jk}}{\sqrt{\left(\sum_{k=1}^{F} f_{ik}^2 \cdot \sum_{k=1}^{F} f_{jk}^2 \right)}} \, , \qquad (3)$$

Equation (3) gives the cosine of two vectors consisting of the shares of the *F* patent classes in the 'patent portfolio' of a firm. Goto and Suzuki (1989) chose a similar spillover measure in their productivity study at the industry level, but used Japanese information on the shares of product classes to which the R&D of an industry is devoted, instead of patent classes.[6]

A disadvantage of these methods is that symmetry is imposed, while it is very awkward to suppose that if industry *i* would generate knowledge useful for industry *j*, industry *i* will automatically benefit to the same extent from knowledge generated in *j*.

The discussion above shows that a number of approaches have been adopted to weight R&D expenditures to arrive at composite indirect R&D or spillover variables. The main result of most

---

5 Whereas a patent originating from the aircraft industry might have the airlines industry as its main beneficiary in terms of the use of the patented *product*, the main user of the *knowledge* documented in the patent might be the motor vehicles industry.

6 Comparable approaches can be found in Adams (1990), who used the shares of various categories of scientists in the research work force of an industry as determinants of its position 'in technological space', and in Los (2000), who proposed to compute weights analogously on the basis of columns of input-output tables.

studies is that technology spillovers do have a substantive impact on productivity growth, irrespective of the weighting scheme applied. As a matter of fact, Keller (1997) claimed that most sets of *randomly* generated weights yielded virtually identical rates of return and goodness of fit statistics. Later on, in a critique of the influential article by Coe and Helpman (1995), he also claimed to find such a result for the effects of international R&D spillovers (Keller, 1998). This result got a lot of attention. Although Keller's claims had to be modified somewhat because of the peculiar way in which he had constructed his random weights, the bottomline was a negative one: Unit weights as discussed above did not yield better or worse results than sets of weights constructed along ways grounded in theory. This more or less led to a standstill with regard to this kind of research. Case study research into sources of technology for specific industries and countries largely replaced systematic comparisons.

In our view, not much more can be gained from the composite spillover variable approach. We feel, however, that new developments in non-classical econometrics make it possible to deal with data characterized by strong violations of the requirements for sensible application of classical least squares approaches. Hence, we propose to return to the original Bernstein and Nadiri (1988) approach of specifying an equation with several industry-specific R&D variables in the right hand side of the equation. These equations will be riddled with multicollinearity problems. Since Generalized Maximum Entropy methods are capable of dealing with problems like these, we aim at estimating rates of return to R&D expenditures by individual industries, including by the industry considered. Furthermore, we will estimate the productivity effects of R&D expenditures by competing industries abroad.

3. The Maximum Entropy Approach

In this section, the basics of (Generalized) Maximum Entropy (ME) econometrics will be introduced. We will limit our discussion to methods used to obtain estimates for the type of linear regression models we use to assess the productivity effects of technology spillovers. More extensive introductions can be found in Kapur and Kesavan (1993) and Golan *et al.* (1996).

The essential property of the ME principle is that it chooses the 'most uncertain', 'most uniform' or 'least information-requiring' distribution for the estimate of a parameter that agrees with the data observed. This is fundamentally different from a more classical least squares approach, in which several assumptions on the distribution of the error term must be taken for granted. The main idea is that a random variable (such as an estimator) z can take on K values (z_1, …, z_K), with unknown probabilities $\mathbf{p} = (p_1, …, p_K)$. Following the formulation proposed by Shannon (1948), the entropy of this distribution $\mathbf{p}$ is:

$$H(\mathbf{p}) = -\sum_{k=1}^{K} p_k \ln p_k \qquad (4)$$

The entropy function H measures the 'uncertainty' of the outcomes of the event. This function reaches its maximum when $\mathbf{p}$ has a uniform distribution: $p_k = 1/K$ for all k. On the other extreme, this function takes a value of zero (no uncertainty) when the probability of one of the outcomes goes to one. If some information about the variable (for example, observations on the dependent and independent variables) is available, it can be used as constraints in a linear programming model aimed at maximizing (4). Each piece of information will lead to a Bayesian update of $\mathbf{p}$. In the linear regression framework, the estimator of a coefficient is found by computing the expected value of z, given $\mathbf{p}$. It is important to note that even for a situation with only one observation, the ME approach yields an estimate of the probabilities, since this observation will generally lead to a difference between the a priori uniformly distributed $\mathbf{p}$ and the posterior $\mathbf{p}$. Hence, in situations in which the number of observations is not large enough to apply classical econometrics, this approach can be used to obtain robust estimates of unknown parameters. Standard errors (required to judge the statistical significance of the point estimates) can be obtained as well, provided that the number of observations exceeds the number of parameters estimated.[7] In Appendix A (and the references therein), information can be found about the statistical properties of the GME estimators used in this paper.

The problem at hand is the estimation of a linear model where a variable y depends on R explanatory variables x_j:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{5}$$

in which $\mathbf{y}$ is the $(N \times 1)$ vector of observations for y, $\mathbf{X}$ is the $(N \times R)$ matrix of observations for the R explanatory variables, $\boldsymbol{\beta}$ is the $(R \times 1)$ vector of unknown parameters $\boldsymbol{\beta} = (\beta_1, ..., \beta_R)'$ to be estimated, and $\mathbf{e}$ is the $(N \times 1)$ vector with random disturbances. As mentioned, each β_r is assumed to be a discrete random variable in the GME approach. A priori beliefs about their $K \geq 2$ possible realizations are included in the estimation procedure by means of supporting vectors $\mathbf{b}_r = (b_{r1}, b_{r2}, \ldots, b_{rK})'$ with corresponding probabilities $\mathbf{p}_r = (p_{r1}, ..., p_{rK})'$, for $r = 1, \ldots, R$. The vectors $\mathbf{b}_r$ are based on the researcher's *a priori* beliefs about the likely values of the parameters. Now, vector $\boldsymbol{\beta}$ can be written as:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_R \end{bmatrix} = \mathbf{Bp} = \begin{bmatrix} \mathbf{b}_1' & \mathbf{0} & . & \mathbf{0} \\ \mathbf{0} & \mathbf{b}_2' & . & \mathbf{0} \\ . & . & . & . \\ \mathbf{0} & \mathbf{0} & . & \mathbf{b}_R' \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \cdots \\ \mathbf{p}_R \end{bmatrix} \tag{6}$$

7 More precisely, if the observations on the independent variables are contained in the matrix **X, X'X** should be of rank R or higher, if R is the number of unknown parameters.

Then, given the vectors $\mathbf{p}_r$ the initial estimate for each parameter is given by

$$\beta_r = \mathbf{b}'_r \mathbf{p}_r = \sum_{k=1}^{K} b_k p_{rk}; \quad r = 1,...,R \tag{7}$$

For the random term, a similar approach is followed. To express the lack of information about the actual values contained in **e**, we assume a distribution for each e_i, with a set of $Q \geq 2$ values $\mathbf{v}_i = \left(v_{i1},...,v_{iQ}\right)'$ with respective probabilities $w_i = \left(w_{i1}, w_{i2},..., w_{iQ}\right)'$.[8] Hence, we can write:

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \cdots \\ e_N \end{bmatrix} = \mathbf{Vw} = \begin{bmatrix} \mathbf{v}_1' & \mathbf{0} & . & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_2' & . & \mathbf{0} \\ . & . & . & . \\ \mathbf{0} & \mathbf{0} & . & \mathbf{v}_N' \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \cdots \\ \mathbf{w}_N \end{bmatrix} \tag{8}$$

and the value of the random term for an observation *i* equals

$$e_i = \mathbf{v}_i' \mathbf{w}_i = \sum_{q=1}^{Q} v_{iq} w_{iq}; \quad i = 1,..., N \tag{9}$$

Consequently, model (5) can be transformed into:

$$\mathbf{y} = \mathbf{XBp} + \mathbf{Vw} \tag{10}$$

Now, the estimation problem for the unknown vector of parameters $\boldsymbol{\beta}$ is reduced to the estimation of $R+N$ probability distributions of the support vectors, and the following constrained entropy maximization problem can be solved to obtain these estimates:

$$\underset{\mathbf{p},\mathbf{w}}{\text{Max}}\, H(\mathbf{p},\mathbf{w}) = -\sum_{r=1}^{R}\sum_{k=1}^{K} p_{rk} \ln p_{rk} - \sum_{i=1}^{N}\sum_{q=1}^{Q} w_{iq} \ln w_{iq} \tag{11a}$$

subject to:

$$\sum_{r=1}^{R}\sum_{k=1}^{K} x_{ri} b_k p_{rk} + \sum_{q=1}^{Q} v_{iq} w_{iq} = y_i; \quad i = 1,...,N \tag{11b}$$

---

8 The distribution for the errors is usually assumed symmetric and centered around 0. Therefore $v_{i1} = -v_{iQ}$. A usual procedure for giving values to this vector is following the so-called 3-sigma rule, which amounts to fixing the extreme bounds as ± 3 times the standard deviation of variable *y*. In our empirical analysis, we will assume identical a priori support vectors for each of the random disturbances.

$$\sum_{k=1}^{K} p_{rk} = 1; \quad r = 1,\dots,R \tag{11c}$$

$$\sum_{q=1}^{Q} w_{iq} = 1; \quad i = 1,\dots,N \tag{11d}$$

The restrictions in (11b) ensure that the posterior probability distributions of the estimates and the errors are compatible with the observations. The restrictions in (11c) and (11d) are just normalization constraints. The estimated value of β_r will be (cf. equation (7), but the vectors **p** now reflect a posteriori distributions):

$$\beta_r = \sum_{k=1}^{K} b_k p_{rk}; \quad r = 1,\dots,\mathrm{R} \tag{12}$$

For GME regressions, a pseudo-R^2 can be computed based on the concept of normalized entropy (see Golan *et al*., 2001). More specifically, the following formula has been applied:

$$pseudo\text{-}\mathrm{R}^2 = 1 - \left[\frac{H(p)}{\max H(p)} \right] \tag{13}$$

This expression compares the value of the entropy function obtained in the ME program with the maximum value Shannon's entropy could take, given the number of probabilities to estimate. A value equal to zero means that the entropy is maximum, which would mean that the information (the data sample) included as constraints in the ME program are not informative at all. The closer the value of this pseudo-R^2 to one, the more information the sample contains. Multiple types of pseudo-R^2s can be reported, the differences depending on the inclusion or exclusion of entropy related to the error term. In the results documented below, we report pseudo-R^2s related to the coefficients to be estimated only, following Golan *et al.* (1999).

4. Data Issues

We use the methodology introduced above to estimate equations resembling production function (1), for 18 industries. The industry classification is given in Appendix B. Our choice for this specific aggregation level is mainly driven by data availability in the EUKLEMS (2007) database, which is the most extensive set of data currently available. Despite the opportunities offered by this database, we are faced with some data restrictions. In order not to loose too much industry detail, we cannot include growth of capital intensities as a source of productivity growth. Although it does not fit standard mainstream production theory, one might argue that a lot of investment is induced by the emergence of improved or new capital goods. This would imply that parts of the returns to R&D carried out in capital goods industries are 'misallocated' to the investing industry if capital intensity is included as a separate determinant.

Further, since the number of countries for which the required data are available is relatively small, we decided to consider three subperiods, 1976-1983, 1984-1991 and 1992-1999. This leaves us with 36 observations per industry, since data for Denmark, Finland, France, Germany, Ireland, Italy, Japan, The Netherlands, Spain, Sweden, the UK and the US have been available. For each of the three time periods we added a dummy, which yields the set of regression equations

$$\left(\frac{\hat{Q}}{L}\right)_{ict} = \alpha_i + \sum_{j=1}^{18} \beta_{ij}^D \frac{RE_{jct}}{Q_{ict}} + \beta_i^F \frac{\sum_{d \neq c} RE_{idt}}{Q_{ict}} + \gamma_{i2} d_2 + \gamma_{i3} d_3 + \varepsilon_{ict} \quad i = 1, \ldots, 18;\ t = 1, 2, 3 \quad (14)$$

The abovementioned subperiods are indicated by *t*. The left hand side of the equation represents the annual average labor productivity growth for industry *i* in country *c*, as taken from the EUKLEMS (2007) database, variable LP_I (gross value added per hour worked, volume index).[9]

The second term of the right hand side of (14) contains eighteen R&D intensities and the corresponding rates of return (the β_{ij}^D coefficients). These refer to R&D expenditures by domestic industries, including the industry under consideration (*i*) itself. The third term captures important parts of the effect of foreign spillovers. In order to remain able to derive standard errors, the limited number of observations led us to the decision not to estimate effects of international interindustry spillovers, but to focus on the effects of international intraindustry spillovers (as opposed to, for example, Verspagen, 1997b). Neither did we include separate effects of R&D spillovers from individual countries, which would have been in the spirit of, among others, Coe and Helpman (1995). The effects of international intraindustry spillovers are captured by the rate of return β_i^F .

All R&D expenditures were taken from OECD's STAN-ANBERD database. In order to arrive at an industry-level classification compatible with the EUKLEMS productivity data some updating procedures comparable to EUKLEMS procedures had to be adopted, for instance in linking ISIC2 and ISIC3 industries to EUKLEMS industries.[10] The value added figures were taken from EUKLEMS (2007), variable VA. Both the R&D expenditures and the value added indicators are expressed in national currency and in current prices. To arrive at average

9 See Timmer *et al.* (2007) for an overview of the data and a description of construction procedures.

10 The R&D dataset is available from the authors on request.

observations for the subperiods, the annual R&D expenditures for the eight year-periods were added and so were the value added figures, after which the ratios of the sums were computed.[11]

The two dummy variables are included to take differences in technological opportunities and other period-specific differences into account. The reference period is the early period, 1976-1983. By adopting this specification, we assume that the rates of return to R&D remained equal over time. We acknowledge the restrictive nature of this approach, although mainstream economists would argue that profit-maximizing firms with rational expectations would lower their R&D expenditures in periods in which returns to a given level of R&D investments decline.

Alternatively, we could have opted for a specification in which we would have looked at just one, 24 year-period. To obtain a reasonable number of degrees of freedom, we should have assumed that industries within a few categories would have had identical rates of return to R&D. This approach was followed by Verspagen (1997a), who assigned industries to the categories "high-tech", "medium-tech" and "low-tech". We feel such an approach is more restrictive than ours, since it would imply that returns would be equal even though R&D activities in different industries are characterized by different degrees of uncertainty (and, therefore, risk).

We estimated equation (14) for 18 manufacturing industries. The industry classification can be found in the Appendix B.

5. Results

In order to estimate equation (14), we specified a maximum entropy problem shaped like equations (11). We took a common support vector with 3 elements (0, 0.5, 1) for all β_{ij}^{D} parameters, for all industries *i*. This implies that we assume that the range of feasible rates of return for own R&D efforts in industry *i* is *a priori* the same as the rate of return to R&D expenditures in other industry *j*. With this support vector we are impose our belief that only nonnegative rates of return of R&D are feasible in the medium- to long-run and averaged over firms in a country. We cannot think of a reason why R&D activity in one sector could affect the medium-run labor productivity performance in an industry negatively. Additionally, we set an upper bound to the rates of return of 100%. In view of the high rates of return to knowledge spillovers (118-147%) as reported by Scherer (1982), this might seem restrictive. If so, this would be indicated by the estimation results, because the additional information produced by the observations would push the estimates close to the 100%-bound set by the support vector.

A common support has been chosen too for the intercept α_i and the time dummies γ_{i2} and γ_{i3}. For these parameters, we took (-5, 0, 5) as the common support. The support vector for the rates of return to foreign intraindustry spillovers β_i^F was uniformly set as (-1, 0, 1). Finally, for

---

11 We could also have opted for a fully dynamic specification. This would have required the determination of a lag structure, which we consider an issue beyond the scope of this paper.

specifying the vector $\mathbf{v}_i$ of feasible values for the error term, a three-point vector centered around 0 has been taken in each regression, following the 3-sigma rule of variable y. This is common practice in most empirical studies that apply the ME approach (following Pukelsheim, 1994).

We first attempted to estimate (14) by means of traditional least squares techniques. We do not present the full set of results for reasons of space, but discussion of a few results suffices to conclude that the estimation problem at hand is not suitable to be tackled by OLS. Hardly any rate of return to own R&D is significant (3 out of 18, at 10%) and estimated rates of return to R&D done in other industries range from -2357% to +3857%. Many of these huge (in an absolute sense) estimates are not significant, however. Foreign intraindustry spillovers have significant effects in just 3 industries. The R^2s range from 0.42 to 0.78. Thus, the results suggest that R&D intensities are able to explain a substantial part of labor productivity growth rates indeed, but no reasonable interpretation can be given to estimates for single coefficients.

Tables 1 reports the results for the estimations of equation (14) obtained by GME along the lines set out above. The estimates for the rates of return of the own R&D intensity in each industry are emphasized. The final columns shows the pseudo-R^2 values.

The own R&D efforts are pointed out as a relevant variable for explaining variations in labor productivity for most industries. Except for industries 2 ("textiles, leather and footwear)", 3 ("wood"), 10 ("metal products") and 13 ("electrical machinery"), the estimates of the return of rate of the own R&D intensity are higher than zero. Furthermore, in almost all industries the own R&D intensity appears to be the highest rate of return estimated. In other words, although our neutral prior inserted into the GME program suggests that the R&D in one industry was expected to produce the same rate of return (50%, since 0.5 is the central value in the support vector) as R&D investments in any other industry, the additional information contained in the observations yields estimates that differ from our prior and are in line with intuition. One might expect that purposefully conducted R&D will have higher rates of return than R&D done by other industries, with different objectives in mind. We find only two cases in which the upper bound of our support vectors (reflecting a rate of return of 100%) might have been restrictive. The estimated rates of return in "fuels" (5) and "radio, TV and communication equipment" (14) amount to 94% and 86%, respectively. Sensitivity analysis with respect to the support vectors might yield evidence that the actual rates of return in these industries could have exceeded 100%. This is an issue will study more systematically below. For many other industries, the estimated returns are in plausible ranges.

Next, let us turn to the results for the effects of R&D spillovers from other industries. In many cases, the information contained in the observations drive the estimated rates of return down to a value close to 0. Apparently, the industries considered did no get relevant technology transmitted from the R&D-performing industry, or it did not manage to use it in a productivity-enhancing way. Two industries stand out in apparently generating hardly any positive productivity effects for other industries. These are “chemicals”(6) and “radio, TV and communication equipment” (14).

Six industries generated technology spillovers that raised labor productivity in at least ten industries. These were: “pulp, paper and printing” (4), “fuels” (5), “rubber and plastics”(7), “office machinery” (12), “instruments” (15) and “other manufacturing” (18). With the exception of (7) and (12), these industries might not be the ones one would have in mind as being very important in shaping manufacturing-wide productivity growth. It should be borne in mind, however, that rates of return do not say much about the contributions of these industries to productivity growth. The correct interpretation of the reported estimates is: if industry *i* would have spent the same amount of money on R&D activities as industry *j* (and with the same success rate), it would have enjoyed a rate of return on this investment equal to the value reported. In view of the fact that the amounts of R&D spent in an industry like “other manufacturing” (18) are rather small in comparison to other industries, the contribution to productivity growth on other industries is probably relatively modest.

It is also interesting to see which industries gained much from R&D activities in other industries. In this respect, we find considerable differences. Industries that appear not to have experienced a lot of positive spillovers (at least not originating from a diversity of industries) are “textiles and leather” (2), “wood” (3), “metal products” (10, “electrical machinery” (13) and “motor vehicles” (16). At the other end of the spectrum, we also identify industries that received productive spillovers from virtually all industries. Examples are “chemicals” (6), “machinery” (11) and “radio, TV and communication equipment” (14). It is important, however, to note that in particular for the last two industries, the value of the pseudo-R^2s is low. This implies that the positive estimates are mainly due to the uniform prior we used, which amounted to a rate of return of 50%. As can easily be checked, many of the estimated rates of return are close to this value indeed.

Table 1. GME estimations of rates of return to R&D and R&D spillovers

	α_i	β^D_{i1}	β^D_{i2}	β^D_{i3}	β^D_{i4}	β^D_{i5}	β^D_{i6}	β^D_{i7}	β^D_{i8}	β^D_{i9}	β^D_{i10}	β^D_{i11}	β^D_{i12}	β^D_{i13}	β^D_{i14}	β^D_{i15}	β^D_{i16}	β^D_{i17}	β^D_{i18}	β^F_i	γ_{i2}	γ_{i3}	Pseudo -R^2
s1	0.03**	0.38**	0.49**	0.47**	0.16**	0.34**	0.01	0.40**	0.45**	0.27**	0.20**	0.01	0.12**	0.06**	0.00	0.14**	0.00	0.00	0.55**	0.41**	-0.05**	-0.05**	0.32
s2	0.02**	0.01	0.40	0.35	0.00	0.31**	0.00	0.09	0.17	0.00	0.01	0.00	0.07**	0.00	0.00	0.01	0.00	0.00	0.33**	0.10**	-0.02**	-0.03**	0.57
s3	0.01**	0.01	0.19	0.30	0.38**	0.06	0.00	0.08	0.04	0.01	0.12	0.00	0.00	0.00	0.00	0.05**	0.00	0.02	0.09	0.18**	-0.03**	-0.04**	0.58
s4	0.01**	0.35**	0.37**	0.42**	0.28**	0.18**	0.00	0.04**	0.09**	0.01**	0.17**	0.00	0.00	0.00	0.00	0.00**	0.00	0.00**	0.18**	1.00**	-0.02**	0.07**	0.43
s5	-0.12**	0.01	0.15**	0.30**	0.09**	0.94**	0.00	0.05**	0.11**	0.34**	0.00	0.00	0.00	0.00	0.00	0.09**	0.00	0.22**	0.01	0.00**	0.00**	0.31**	0.61
s6	0.08**	0.22**	0.43**	0.44**	0.15**	0.32**	0.01**	0.37**	0.34**	0.18**	0.21**	0.01**	0.12**	0.02**	0.00**	0.09**	0.04**	0.01**	0.42**	0.01**	-0.06**	-0.04**	0.30
s7	-0.04**	0.09**	0.46**	0.41**	0.03**	0.37**	0.00	0.49**	0.40**	0.15**	0.16**	0.00	0.00**	0.01	0.00	0.31**	0.00**	0.11**	0.38**	0.04**	-0.02**	0.02**	0.41
s8	0.02**	0.04	0.38**	0.39**	0.01	0.07**	0.00	0.12**	0.23**	0.04	0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19**	0.13**	-0.02**	-0.02**	0.53
s9	-0.01	0.17	0.45	0.48	0.39**	0.53**	0.01	0.52*	0.43	0.52*	0.40	0.12*	0.10**	0.15**	0.00	0.46**	0.25**	0.15**	0.31	-0.02**	-0.14**	-0.21**	0.22
s10	-0.03**	0.26	0.42	0.47	0.31**	0.35	0.02	0.28	0.34	0.24	0.38	0.06	0.07**	0.16**	0.00	0.02	0.05	0.02	0.46*	0.13**	0.00**	-0.02**	0.33
s11	-0.01**	0.31**	0.49**	0.49**	0.40**	0.38**	0.04**	0.57**	0.50**	0.47**	0.44**	0.24**	0.10*	0.10**	0.02**	0.60**	0.21**	0.04**	0.46**	0.03**	-0.05**	-0.07**	0.20
s12	0.29**	0.27**	0.47	0.48	0.16**	0.52**	0.00	0.32**	0.41**	0.23**	0.33**	0.00	0.69**	0.22**	0.00	0.06**	0.00	0.17**	0.44**	-0.73**	0.06**	2.77**	0.37
s13	-0.09**	0.17	0.51	0.50	0.29	0.33	0.00	0.46	0.55	0.44	0.37	0.05	0.23**	0.07	0.00	0.11	0.06	0.04	0.49	0.01	0.04**	-0.01	0.30
s14	-0.34**	0.47**	0.46**	0.50**	0.65**	0.47**	0.12**	0.43**	0.45**	0.50**	0.44**	0.27**	0.30**	0.17**	0.86**	0.49**	0.12**	0.15**	0.46**	0.00**	-0.12**	0.63**	0.18
s15	-0.06**	0.17	0.41	0.37	0.07	0.24	0.00	0.21	0.22	0.08	0.26	0.00	0.02	0.00	0.00	0.24**	0.03	0.37**	0.41**	0.05**	-0.05**	-0.04**	0.39
s16	-0.02**	0.04	0.32	0.31	0.01	0.19	0.00	0.14	0.16	0.13	0.14	0.00	0.05**	0.00	0.00	0.00	0.25**	0.00	0.11	0.00	-0.02**	0.00	0.50
s17	-0.04**	0.04**	0.37**	0.46**	0.10**	0.70**	0.00	0.16**	0.24**	0.10**	0.19**	0.00	0.05**	0.00*	0.00	0.00	0.01**	0.06**	0.16**	0.04**	0.03**	0.00**	0.45
s18	-0.01**	0.02	0.30	0.25	0.08*	0.03	0.00	0.39**	0.22	0.03	0.02	0.00	0.00	0.00	0.00	0.05**	0.00	0.00	0.34**	0.05**	0.01**	0.00	0.55

* Estimates significantly different from 0 at 10%; ** Estimates significantly different from 0 at 5%;
Shaded cells on the main diagonal refer to productivity effects of "own" R&D.

Support vectors for all α_i, γ_{i2} and γ_{i3}: (-5, 0, 5); Support vectors for all β^D_{ij} : (0, 0.5, 1); Support vectors for all β^F_i : (-1, 0, 1).

Most of the estimated rates of return to foreign intraindustry spillovers are positive, such as in "food" (1), and to a lesser extent, "wood" (3), "non-metallic mineral products" (8) and "metal products" (10). As was mentioned already, we specified the support vector for the associated coefficient as (-1, 0, 1). We did so to consider that foreign R&D can lead to positive productivity effects through knowledge externalities, but also to business stealing effects. Such business stealing effects appear to dominate in the "computers" industry (12). One might speculate that the successful R&D projects in the US and Japan have largely eradicated high-productivity activities from many of the other countries in the sample, such as France and Italy.

For "pulp, paper, and printing" (4) we find a rate of return to foreign intraindustry spillovers equal to the upper boundary of the prior distribution. This is another case that asks for sensitivity analysis. Results for a different set of support vectors are documented in Appendix C. As can be concluded from the pseudo-R^2s, the wider support vectors and the higher expected rates of returns to domestic interindustry spillovers implied by the prior uniform distributions yield lower explanatory power. Nevertheless, most results do not change in a qualitative sense. The estimated rates of return are generally a bit higher, which should not come as a surprise: the expected value of the prior distributions is higher than in the baseline case reported in Table 1, after which exactly the same information (in the form of observations) is fed to the entropy maximizing program.

Some of the results we find are in line with intuition, others were less expected. Particularly, the result that our estimation framework singles out own R&D as yielding significant returns (while we treated these intensities symmetrically with indirect R&D intensities reflecting spillovers) makes us think that interesting and insightful results can be attained by further exploring GME estimation of technology spillover effects. Testing several specifications of the regression equation should prove a useful avenue for future research.

6. Conclusions

In this paper, we introduced a novel approach to the assessment of the impact of interindustry technology spillovers on labor productivity. Unlike the vast majority of empirical studies undertaken so far, we do not use classical least-squares estimation techniques, but rely on Generalized Maximum Entropy techniques. This toolbox of econometric methods is particularly geared towards situations in which data are ill-behaved. In studies linking productivity growth to sources of spillovers, multicollinearity is often a big problem, as a consequence of which it is impossible to estimate the effects of R&D done in individual industries. This paper is the first one to approach these problems using GME techniques.

GME estimation yields much more plausible results than ordinary least squares estimation. Our results show that with just a few exceptions, industries attain highly positive rates of return to their R&D investments. Moreover, some industries benefit from innovation in many other industries, whereas others mainly rely on own R&D activities. Generally, industries benefit from technology generated in similar firms abroad. In the computer industry, however, very successful R&D in a few countries appear to have had very negative effects in other countries. This business stealing effect dominated strongly.

The analysis in this paper can be extended in various ways. First, we do not employ the full potential of our dataset in terms of dynamic analyses. It should be possible to replicate studies like Los and Verspagen (2000), especially because GME can deal with non-stationary series of observations without having to incorporate cointegration formulations and the like. A second extension would not relate to the application, but more to a potential improvement in terms of the methodology. In this paper, our prior has been that all industries benefit to an identical extent from technology spillovers. Although the previous literature experienced difficulties in assessing the exact origins of productivity-enhancing spillovers, it made clear that industries differ in the extent to which spillovers play a role. Such information could be used to work with industry-specific priors, which might further improve the accuracy of the estimates. Finally, our paper does not shed light on the importance of the channels through which the most important technology spillovers flow. In our view, it should be possible to estimate a single productivity parameter for an indirect R&D variable for which the weights are estimated simultaneously. In a next step, these weights could be compared to the weights found by researchers who constructed their weights based on a specific idea of how technological spillovers emerge.

References

Adams, J.D. (1990), "Fundamental Stocks of Knowledge and Productivity Growth", *Journal of Political Economy*, vol. 98, pp. 673-702.

Bernstein, J.I. (1989), "The Structure of Canadian Inter-Industry R&D Spillovers, and the Rates of Return to R&D", *Journal of Industrial Economics*, vol. 37, pp. 315-328.

Bernstein, J.I. and M.I. Nadiri (1988), "Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries", *American Economic Review (Papers and Proceedings)*, vol. 78, pp. 429-434.

Coe, D.T. and E. Helpman (1995), "International R&D Spillovers", *European Economic Review*, vol. 39, pp. 859-887.

DeBresson, C., G. Sirilli, X. Hu and F. Kwan Luk (1994), "Structure and Location of Innovative Activity in the Italian Economy, 1981-85", *Economic Systems Research*, vol. 6, pp. 135-158.

EUKLEMS (2007), EUKLEMS Database, March 2007, http://www.euklems.com.

Fraser, I. (2000), "An Application of Maximum Entropy Estimation: The Demand for Meat in the United Kingdom", *Applied Economics*, vol. 32, pp. 45-59.

Golan, A., G. Judge and D. Miller (1996), *Maximum Entropy Econometrics: Robust Estimation with Limited Data* (Chichester UK, John Wiley).

Golan, A., E. Moretti and J.M. Perloff (1999), "An Information-Based Sample-Selection Estimation Model of Agricultural Workers' Choice between Piece-Rate and Hourly Work", *American Journal of Agricultural Economics*, vol. 81, pp. 735-741.

Golan, A., J.M. Perloff and E.Z. Shen (2001), "Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand", *Review of Economics and Statistics*, vol. 83, pp. 541-550.

Goto, A. and K. Suzuki (1989), "R&D Capital, Rate of Return on R&D Investment and Spillover of R&D in Japanese Manufacturing Industries", *Review of Economics and Statistics*, vol. 71, pp. 555-564.

Griliches, Z. (1979), "Issues in Assessing the Contribution of Research and Development to Productivity Growth", *The Bell Journal of Economics*, vol. 10, pp. 92-116.

Griliches, Z. (1992), "The Search for R&D Spillovers", *Scandinavian Journal of Economics*, vol. 94, pp. S29-S47.

Jaffe, A.B. (1986), "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value", *American Economic Review*, vol. 76, pp. 984-1001.

Johnson, D. and R.E. Evenson (1997), "Innovation and Invention in Canada", *Economic Systems Research*, vol. 9, pp. 177-192.

Kapur, J.N. and H.K. Kesavan (1993), *Entropy Optimization Principles with Applications* (New York: Academic Press).

Keller, W. (1997), "Technology Flows between Industries: Identification and Productivity Effects", *Economic Systems Research*, vol. 9, pp. 213-220.

Keller, W. (1998), "Are International R&D Spillovers Trade-Related? Analyzing Spillovers among Randomly Matched Trade Partners", *European Economic Review*, vol. 42, pp. 1469-1481.

Los, B. (2000), "The Empirical Performance of a New Interindustry Technology Spillover Measure", in: B. Nooteboom and P. Saviotti (eds.), *Technology and Knowledge: From the Firm to Innovation Systems*, Edward Elgar, Cheltenham UK, pp. 118-151.

Los, B. and B. Verspagen (2000), "R&D Spillovers and Productivity: Evidence from U.S. Manufacturing Microdata", *Empirical Economics*, vol. 25, pp. 127-148.

Los, B. and B. Verspagen (2007), "Technology Spillovers and Their Impact on Productivity", in: H. Hanusch and A. Pyka (eds.), *Elgar Companion to Neo-Schumpeterian Economics*, Edward Elgar, Cheltenham UK, pp. 574-593.

Mittelhammer R.C. and N.S. Cardell (1997), "On the Consistency and Asymptotic Normality of Data-constrained GME Estimators in the General Linear Model", mimeo, University of Washington.

Mohnen, P. (1996), "R&D Externalities and Productivity Growth", *STI Review*, vol. 18, pp. 39-66.

Nadiri, M.I. (1993), "Innovations and Technological Spillovers", NBER Working Paper 4423, Cambridge MA.

Pukelsheim, F. (1994), "The 3 Sigma Rule", *The American Statistician*, vol. 48, pp. 88-91.

Robson, M., J. Townsend and K. Pavitt (1988), "Sectoral Patterns of Production and Use of Innovations in the UK: 1945-83", *Research Policy*, vol. 17, pp. 1-14.

Sakurai, N., G. Papaconstantinou and E. Ioannidis (1997), "Impact of R&D and Technology Diffusion on Productivity Growth: Empirical Evidence for 10 OECD Countries", *Economic Systems Research*, vol. 9, pp. 81-109.

Scherer, F.M. (1982), "Inter-Industry Technology Flows and Productivity Measurement", *Review of Economics and Statistics*, vol. 64, pp. 627-634.

Shannon, J. (1948), "A Mathematical Theory of Communication", *Bell System Technical Bulletin Journal*, vol. 27, pp. 379-423.

Sterlacchini, A. (1989), "R&D, Innovations, and Total Factor Productivity Growth in British Manufacturing", *Applied Economics*, vol. 21, pp. 1549-1562.

Terleckyj, N.E. (1974), *Effects of R&D on the Productivity Growth of Industries: An Exploratory Study*, National Planning Association, Washington DC.

Timmer, M.P., M. O'Mahony and B. van Ark (2007), "The EU KLEMS Growth and Productivity Accounts: An Overview", University of Groningen & University of Birmingham.

Van Meijl, H. (1995), *Endogenous Technological Change: The Case of Information Technology*, Ph.D. Thesis, University of Limburg, Maastricht.

Van Pottelsberghe de la Potterie, B. (1997), "Issues in Assessing the Effect of Interindustry Spillovers", *Economic Systems Research*, vol. 9, pp. 331-356.

Verspagen, B. (1997a), "Measuring Inter-Sectoral Technology Spillovers: Estimates from the European and US Patent Office Databases", *Economic Systems Research*, vol. 9, pp. 47-65.

Verspagen, B. (1997b), "Estimating International Technology Spillovers Using technology Flow Matrices", *Weltwirtschaftliches Archiv*, vol. 133, pp. 226-248.

Wolff, E.N. (1997), "Spillovers, Linkages, and Technical Change", *Economic Systems Research*, vol. 9, pp. 9-23.

Appendix A: Statistical properties of the GME estimator

The large sample properties of the ME estimators are analyzed in Golan *et al.* (1996, chapter 6). ME estimators are shown to be consistent and asymptotically normal. Golan *et al.* also analyze small sample properties, using Monte Carlo simulation. They numerically compare the GME estimators to traditional least squares and maximum likelihood estimators. Their results show a good performance in terms of the accuracy of the estimates.

In order to do inference in the GME approach, the procedure suggested by Mittelhammer and Cardell (1997), Fraser (2000) and Golan *et al.* (2001) can be followed. Under assumptions on the behavior of model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ that guarantee the consistency and asymptotical normality of the estimator, the distribution of the estimates follows $\hat{\boldsymbol{\beta}} \rightarrow \mathbf{N}\left[\boldsymbol{\beta}, \frac{\sigma_\lambda^2}{\varkappa^2}(\mathbf{X}'\mathbf{X})^{-1}\right]$, where σ_λ^2 is the variance of the Lagrange multipliers of (13b) or (17b). It is possible to estimate σ_λ^2 consistently, as $\hat{\sigma}_\lambda^2 = \frac{\sum_{i=1}^{N}\hat{\lambda}_i^2}{N}$, where $\hat{\lambda}_i$ is the Lagrange-multiplier associated to observation *i*.

On the other hand, κ^2 is a scalar related to the variance of the error term. The parameter κ can be estimated consistently as $\hat{\kappa} = \frac{1}{N\sum_{i=1}^{N}\mathrm{Var}(\hat{e}_i)}$, in which $\mathrm{Var}(\hat{e}_i) = \left[\sum_{q=1}^{Q} v_q^2 \hat{w}_{iq}\right] - \hat{e}_i^2$, and $\hat{e}_i$ is the estimate of the error for each observation *i* defined by $\hat{e}_i = \sum_{q=1}^{Q} v_q \hat{w}_{iq}$. Hence, it is possible to estimate the variance of GME estimators and obtain the t-ratios as $\frac{\hat{\boldsymbol{\beta}}}{\sqrt{\mathrm{V\hat{a}r}(\hat{\boldsymbol{\beta}})}}$.

Appendix B: Industry Classification

1.	Food, beverages and tobacco	10.	Fabricated metal products
2.	Textiles, textile, leather and footwear	11.	Machinery, n.e.c.
3.	Wood and products of wood and cork	12.	Office, accounting and computing machinery
4.	Pulp, paper, paper products, printing and publishing	13.	Electrical machinery and apparatus, n.e.c.
5.	Coke, refined petroleum and nuclear fuel	14.	Radio, television and communication equipment
6.	Chemicals and chemical products	15.	Medical, precision and optical instruments
7.	Rubber and plastics	16.	Motor vehicles, trailers and semi-trailers
8.	Other non-metallic mineral products	17.	Other transport equipment
9.	Basic metals	18.	Manufacturing, n.e.c.

Appendix C: Sensitivity Analysis

Table C1. Sensitivity of the GME estimates

	α_i	β^D_{i1}	β^D_{i2}	β^D_{i3}	β^D_{i4}	β^D_{i5}	β^D_{i6}	β^D_{i7}	β^D_{i8}	β^D_{i9}	β^D_{i10}	β^D_{i11}	β^D_{i12}	β^D_{i13}	β^D_{i14}	β^D_{i15}	β^D_{i16}	β^D_{i17}	β^D_{i18}	β^F_i	γ_{i2}	γ_{i3}	Pseudo -R^2
s1	0.03**	0.48**	0.72**	0.67**	0.14**	0.42**	0.01	0.52**	0.62**	0.29**	0.20**	0.00	0.10**	0.04**	0.00	0.11**	0.00	0.00	0.84**	0.44**	-0.04**	-0.05**	0.38
s2	0.02**	0.01	0.55	0.48	0.00	0.36**	0.00	0.08	0.18	0.00	0.01	0.00	0.06**	0.00	0.00	0.00	0.00	0.00	0.38**	0.10**	-0.02**	-0.03**	0.61
s3	0.01**	0.00	0.21	0.38	0.45**	0.06	0.00	0.09	0.03	0.00	0.12	0.00	0.00	0.00	0.00	0.05**	0.00	0.01	0.09	0.19**	-0.03**	-0.04**	0.63
s4	-0.01**	0.46**	0.58**	0.69**	0.55**	0.42**	0.00	0.13**	0.26**	0.06**	0.35	0.00	0.00**	0.00**	0.00	0.01**	0.00	0.01**	0.36**	1.44**	-0.04**	0.07**	0.44
s5	-0.16**	0.01	0.25**	0.46**	0.08**	1.36**	0.00	0.10**	0.19**	0.43**	0.01	0.00	0.00	0.00	0.00	0.07**	0.00	0.29**	0.02*	-0.03**	-0.02**	0.25**	0.61
s6	0.07**	0.28**	0.63**	0.64**	0.17**	0.45**	0.00	0.51**	0.46**	0.23**	0.29**	0.00	0.13**	0.02*	0.00	0.09**	0.04**	0.01	0.58**	0.01**	-0.06**	-0.04**	0.37
s7	-0.05**	0.06**	0.63**	0.55**	0.01**	0.44	0.00	0.67**	0.47**	0.11**	0.15**	0.00	0.00	0.00**	0.00	0.32**	0.00	0.10**	0.45**	0.04**	-0.02**	0.02**	0.47
s8	0.02**	0.02	0.51**	0.51**	0.01	0.05	0.00	0.10	0.25*	0.03	0.05	0.00	0.00	0.00	0.00	0.00**	0.00	0.00	0.17**	0.12**	-0.02**	-0.02**	0.59
s9	-0.02**	0.13	0.62	0.70	0.50**	0.79**	0.00	0.74	0.56	0.74	0.53	0.08	0.06	0.11	0.00	0.44**	0.24**	0.11	0.36	-0.02**	-0.12**	-0.20**	0.29
s10	-0.03**	0.26	0.56	0.68	0.35**	0.44*	0.01	0.30	0.40	0.24	0.49*	0.03	0.05	0.14**	0.00	0.01	0.04	0.01	0.62**	0.12**	0.00	-0.01**	0.40
s11	-0.02**	0.35**	0.72**	0.72*	0.53**	0.51**	0.02	0.83**	0.71**	0.64**	0.61**	0.23**	0.08**	0.09**	0.01	0.66**	0.19**	0.03*	0.67**	0.03**	-0.04**	-0.06**	0.26
s12	0.12**	0.31**	0.68**	0.71**	0.19**	0.77**	0.00	0.41**	0.57**	0.26**	0.44**	0.00	1.05**	0.22**	0.00	0.05**	0.00	0.15**	0.62**	-0.76**	0.12**	2.76**	0.39
s13	-0.10**	0.14	0.75	0.74	0.32	0.42	0.00	0.62	0.79	0.52	0.46	0.02	0.22**	0.04	0.00	0.07	0.04	0.03	0.67	0.01	0.04**	0.00	0.37
s14	-0.41**	0.69**	0.63**	0.73**	1.14**	0.68**	0.03	0.52**	0.58**	0.74**	0.54**	0.14**	0.24**	0.07**	1.45**	0.66**	0.02**	0.05**	0.63**	-0.01**	-0.14**	0.56**	0.20
s15	-0.06**	0.15	0.56	0.48	0.04	0.25	0.00	0.22	0.22	0.06	0.28	0.00	0.01	0.00	0.00	0.23**	0.02	0.37**	0.51**	0.05**	-0.05**	-0.04**	0.47
s16	-0.01**	0.03	0.38	0.37	0.00	0.19	0.00	0.12	0.14	0.12	0.12	0.00	0.04**	0.00	0.00	0.00	0.24**	0.00	0.10	0.00	-0.02**	0.01	0.57
s17	-0.04**	0.02	0.49**	0.64**	0.08**	0.92**	0.00	0.16**	0.28**	0.08**	0.21**	0.00	0.03**	0.00	0.00	0.00	0.01*	0.04**	0.20**	0.05**	0.02**	0.00**	0.51
s18	-0.01**	0.01	0.35	0.27	0.07	0.02	0.00	0.49**	0.21	0.02	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.39**	0.05**	0.01**	-0.01**	0.60

* Estimates significantly different from 0 at 10%; ** Estimates significantly different from 0 at 5%;
Shaded cells on the main diagonal refer to productivity effects of "own" R&D;
The support vectors were fixed as (0, 0.75, 1.5) for the interindustry spillovers, in (-1.5,0,1.5) for the aggregated R&D in the same industry abroad and in (-7.5,0,7.5) for the intercept and the time dummies.

E L I N K E I N O E L Ä M Ä N   T U T K I M U S L A I T O S   (ETLA)

THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
LÖNNROTINKATU 4 B, FIN-00120 HELSINKI

Puh./Tel. (09) 609 900
Int. 358-9-609 900
http://www.etla.fi

Telefax (09) 601753
Int. 358-9-601 753

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

Julkaisut ovat saatavissa elektronisessa muodossa internet-osoitteessa:
http://www.etla.fi/finnish/research/publications/searchengine

No 1074 KARI E.O. ALHO – NUUTTI NIKULA, Productivity, Employment and Taxes – A Svar Analysis of the Trade-offs and Impacts. 22.02.2007. 19 p.

No 1075 RAINE HERMANS – MARTTI KULVIK, Simulaatio lääkekehitysalan kannattavuudesta ja riskeistä. 26.02.2007. 25 s.

No 1076 TERHI HAKALA – OLLI HALTIA – RAINE HERMANS – MARTTI KULVIK – HANNA NIKINMAA – ALBERT PORCAR-CASTELL – TIINA PURSULA, Biotechnology as a Competitive Edge for the Finnish Forest Cluster. 26.02.2007. 76 p.

No 1077 RITA ASPLUND, Finland: Decentralisation Tendencies within a Collective Wage Bargaining System. 02.03.2007. 31 p.

No 1078 PAVEL FILIPPOV – VLAD YURKOVSKY, Essay on Internationalisation Potential of Northwest Russian and Finnish Energy Clusters. 20.03.2007. 36 p.

No 1079 MARKKU KOTILAINEN, Determinants of Finnish-Russian Economic Relations. 22.03.2007. 39 p.

No 1080 JYRKI ALI-YRKKÖ – MIKA PAJARINEN – PETRI ROUVINEN – PEKKA YLÄ-ANTTILA, Family Businesses and Globalization in Finland. 03.04.2007. 35 p.

No 1081 JYRKI ALI-YRKKÖ, Ulkomaalaisomistuksen vaikutus yritysten kasvuun. 29.03.2007. 24 s.

No 1082 MIKKO KETOKIVI – JYRKI ALI-YRKKÖ, Determinants of Manufacturing-R&D Colocation. 30.03.2007. 28 p.

No 1083 VILLE KAITILA, Suomen ulkomaankaupan erikoistuminen – keiden kanssa kilpailemme? 05.04.2007. 25 s.

No 1084 MIKKO MÄKINEN, CEO Compensation, Firm Size and Firm Performance: Evidence from Finnish Panel Data. 30.03.2007. 38 p.

No 1085 MIKKO MÄKINEN, Do Stock Option Schemes Affect Firm Technical Inefficiency? Evidence from Finland. 30.03.2007. 26 p.

No 1086 CHRISTOPHER PALMBERG, Modes, Challenges and Outcomes of Nanotechnology Transfer – A Comparative Analysis of University and Company Researchers. 05.04.2007. 33 p.

No 1087 VILLE KAITILA, Free Trade between the EU and Russia: Sectoral Effects and Impact on Northwest Russia. 05.04.2007. 23 p.

No 1088 PAAVO SUNI, Oil Prices and The Russian Economy: Some Simulation Studies with NiGEM. 16.04.2007. 15 p.

No 1089 JUKKA LASSILA – NIKU MÄÄTTÄNEN – TARMO VALKONEN, Vapaaehtoinen eläkesäästäminen tulevaisuudessa. 16.04.2007. 38 s.

No 1090 VILLE KAITILA, Teollisuusmaiden suhteellinen etu ja sen panosintensiivisyys. 25.04.2007. 31 s.

No 1091 HELI KOSKI, Private-collective Software Business Models: Coordination and Commercialization via Licensing. 26.04.2007. 24 p.

No 1092 PEKKA ILMAKUNNAS – MIKA MALIRANTA, Aging, Labor Turnover and Firm Performance. 02.05.2007. 40 p.

No 1093 SAMI NAPARI, Gender Differences in Early-Career Wage Growth. 03.05.2007. 40 p.

No 1094 OLAVI RANTALA – PAAVO SUNI, Kasvihuonekaasupäästöt ja EU:n päästörajoituspolitiikan taloudelliset vaikutukset vuoteen 2012. 07.05.2007. 24 s.

No 1095 OLAVI RANTALA, Kasvihuonekaasupäästöjen ennakointi ja EU:n päästörajoituspolitiikan vaikutusten arviointi. 07.05.2007. 22 s.

No 1096 JANNE HUOVARI – JUKKA JALAVA, Kansainvälinen ja vertaileva näkökulma Suomen tuottavuuskehitykseen. 12.06.2007. 36 s.

No 1097 JARLE MØEN, Should Finland Introduce an R&D Tax Credit? Reflections Based on Experience with Norwegian R&D Policy. 12.06.2007. 14 p.

No 1098 RITA ASPLUND – OUSSAMA BEN-ABDELKARIM – ALI SKALLI, An Equity Perspective on Access to, Enrolment in and Finance of Tertiary Education. 09.08.2007. 48 p.

No 1099 TERTTU LUUKKONEN, Understanding the Strategies of Venture Capital investors in Helping their Portfolio Firms to Become International. 17.08.2007. 24 p.

No 1100 SARIANNA M. LUNDAN, The Home Country Effects of Internationalisation. 21.08.2007. 43 p.

No 1101 TUOMO NIKULAINEN, Identifying Nanotechnological Linkages in the Finnish Economy. An Explorative Study. 25.09.2007. 31 p.

No 1102 HELI KOSKI, Do Technology Diffusion Theories Explain the OSS Business Model Adoption Patterns? 29.10.2007. 26 p.

No 1103 JUKKA JALAVA – PIRKKO AULIN-AHMAVAARA – AKU ALANEN, Intangible Capital in the Finnish Business Sector, 1975-2005. 29.10.2007. 25 p.

No 1104 BÖCKERMAN, Petri – JOHANSSON, Edvard – HELAKORPI, Satu – UUTELA, Antti, Economic Inequality and Health: Looking Beyond Aggregate Indicators. 05.11.2007. 21 p.

No 1105 MIKA MALIRANTA – RITA ASPLUND, Training and Hiring Strategies to Improve Firm Performance. 08.11.2007. 45 p.

No 1106 ESTEBAN FERNÁNDEZ VÁZQUEZ – BART LOS, A Maximum Entropy Approach to the Identification of Productive Technology Spillovers. 08.11.2007. 21 p.