2001s-31

R&D and Patents: Which Way Does the Causality Run?

Hans van Ophem, Erik Brouwer,
Alfred Kleinknecht, Pierre Mohnen

Série Scientifique
Scientific Series

CIRANO
Centre interuniversitaire de recherche en analyse des organisations

Montréal
Avril 2001

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.

Les organisations-partenaires / The Partner Organizations

- École des Hautes Études Commerciales
- École Polytechnique
- Université Concordia
- Université de Montréal
- Université du Québec à Montréal
- Université Laval
- Université McGill
- MEQ
- MRST
- Alcan inc.
- AXA Canada
- Banque du Canada
- Banque Laurentienne du Canada
- Banque Nationale du Canada
- Banque Royale du Canada
- Bell Québec
- Bombardier
- Bourse de Montréal
- Développement des ressources humaines Canada (DRHC)
- Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec
- Hydro-Québec
- Imasco
- Industrie Canada
- Pratt & Whitney Canada Inc.
- Raymond Chabot Grant Thornton
- Ville de Montréal

© 2001 Hans van Ophem, Erik Brouwer, Alfred Kleinknecht et Pierre Mohnen. Tous droits réservés. All rights reserved.
Reproduction partielle permise avec citation du document source, incluant la notice ©.
Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.

Ce document est publié dans l'intention de rendre accessibles les résultats préliminaires de la recherche effectuée au CIRANO, afin de susciter des échanges et des suggestions. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs, et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.
This paper presents preliminary research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

ISSN 1198-8177

R&D and Patents: Which Way Does the Causality Run?*

Hans van Ophem[†], *Erik Brouwer*[‡], *Alfred Kleinknecht*[§], *Pierre Mohnen*[¶]

Résumé / Abstract

A partir de données transversales de 460 entreprises néerlandaises ayant répondu aux enquêtes innovation de 1988 et 1992, nous réexaminons le sens de la causalité entre la R-D et les brevets. Les deux équations de comportement ont été estimées simultanément en supposant une distribution bivariée conditionnelle entre ces deux variables, dont l'une est discrète et l'autre continue. Nous avons essayé différentes spécifications pour les données de comptage sur les brevets. Nous trouvons que la causalité à la Granger va des brevets à la R-D dans toutes les spécifications. Un brevet en plus augmente la R-D quatre ans plus tard de 7,5 %. La causalité dans l'autre sens disparaît dès que l'on s'écarte le moindrement d'une distribution Poisson des données de brevets.

From cross-sectional data of 460 firms that responded to both the 1988 and the 1992 Dutch innovation surveys we have reexamined the causality direction between R&D and patents, using data on contemporaneous and four-year lagged patent applications and R&D expenditures. The two equations have been estimated jointly assuming a bivariate conditional distribution between the two variables, one being discrete and the other one continuous. We have experimented with different specifications of the count data for patent applications. We find that patents Granger-cause R&D in all specifications. One additional patent increases R&D four years later by 7.5%. The reverse causality from R&D to patents vanishes as soon as we depart in one way or another from the simple Poisson specification of patent counts.

Mots Clés : Enquêtes innovation, brevets, R-D, données de comptage

Keywords: Innovation survey data, patents, R&D, count data

JEL: C35, D21

---

* Corresponding Author: Pierre Mohnen, CIRANO, 2020 University Street, 25th floor, Montréal, Qc, Canada H3A 2A5 Tel.: (514) 985-4028 Fax: (514) 985-4039 email: mohnenp@cirano.qc.ca

† University of Amsterdam and Tinbergen Institute

‡ University of Amsterdam

§ Technical University Delft

¶ Université du Québec à Montréal and CIRANO

1. Introduction.

Because of the lack of long time series on R&D and patents, few studies have examined the causal link in the Granger sense between these two innovation indicators. The existing evidence gathered by Griliches (1981), Pakes (1985), Hall et al. (1986) and Griliches et al. (1991) points to an almost instantaneous relationship, and, to the extent that any lag effects can be established, causality seems to run from R&D to patents. The Dutch innovation surveys which contain questions on patent applications allow us to re-examine this issue using information of four year lagged amounts of R&D expenditures and numbers of patent applications for a sample of 460 firms of the 1992 survey that also responded to the 1988 survey.

Section 2 discusses the model which relates R&D expenditures and patent count mutually. Section 3 introduces the data and the estimation results are presented in section 4. Section 5 concludes.

2. Model.

As discussed in the introduction, we want to investigate the relation between R&D expenditures and our measure of the rate of success of these expenditures, the number of patents acquired. Clearly, stating that R&D in year t ($R\&D_t$) depends on the patent count of year t (y_t) and that the patent count of year t depends on $R\&D_t$ gives rise to a model with an internal inconsistency which is impossible to be estimated. The problem is that y_t cannot cause $R\&D_t$ and at the same time be caused by $R\&D_t$. Also from economic theory one would not expect such a relation. Economic theory suggests that y_t and $R\&D_t$ depend on past R&D expenditures and past acquired patents respectively. In the panel data set used in this paper, we have 2 waves available with a time lag of 4 years (1988-1992).We will denote these past R&D expenditures by $R\&D_{t-1}$ and the past patent counts by y_{t-1}. It is not very likely that using these past observations specify the determination process of $R\&D_t$ and y_t correctly. One would expect many past values having an influence on this process (cf. Bound et al., 1884). However, as we do not have better information, we have to rely on our expectation that $R\&D_{t-1}$ and y_{t-1} are good proxies. We do not expect that R&D expenditures fluctuate wildly across time. Moreover, once a firm has established a high quality research department probably a relatively constant flow of patents will be generated.[1]

[1] This argument would allow us to use $R\&D_t$ and y_t as proxies for the past R&D expenditures and the patent count. Still a statisitical inconsistency would be present in the model (cf.

Assume that the distribution of the patent count given y_{t-1} is given by:

$$\Pr(y_{ti} = y_i | R\&D_{t-1i}) = f(y_i, R\&D_{t-1}; \lambda_i) \tag{1}$$

where λ_i is a vector of parameters of the distribution. One (or more) of the elements of λ_i can be made a function of the characteristics of the firm. If we choose the Poisson distribution λ_i only consists of the expectation of the count. If X_i represents a vector of characteristics of firm i, the usual specification for this expectation is:

$$\lambda_i = \exp(\beta' X_i + \alpha R\&D_{t-1i}) \tag{2}$$

For the R&D expenditures we specify the following relation:[2]

$$R\&D_{ti} = \gamma' Z_{ti} + \delta y_{t-1i} + \varepsilon_{ti} \tag{3}$$

where Z_{ti} is again a vector of explanatory variables of firm i and ε_{ti} an error term. Given the fact that both $R\&D_{t-1i}$ and y_{t-1i} are proxies of the past, we cannot expect that the distributions of y_{ti} and $R\&D_{ti}$ are independent. The relevant conditional loglikelihood looks something like:[3]

$$\begin{aligned} \log(L|y_{t-1i}, R\&D_{t-1i}) &= \sum_{i=1}^{N} \log(L_i|y_{t-1i}, R\&D_{t-1i}) \\ &= \sum_{i=1}^{N} \log(g(y_{ti}, R\&D_{ti}; \rho|y_{t-1i}, R\&D_{t-1i})) \end{aligned} \tag{4}$$

where N is the number of observations and $g(.,.;\rho|.)$ the conditional bivariate distribution of y_{ti} and $R\&D_{ti}$ where the correlation between the variables is given by ρ. Very few distribution functions combining a discrete and a continuous variable are available. However, by following the method of Van Ophem (2000) g can be specified for any combination of a discrete count and a continuous distribution.

We start by specifying the marginal distributions of the random variables under investigation. The conditional marginal distribution of the patent count is given by (1) and the conditional marginal distribution of the R&D expenditures is based on the assumption that the ε_{ti} has a conditional distribution $h(\varepsilon_{ti}|y_{t-1i})$.

Maddala, 1983, p.118).

[2]In fact we use the log of $R\&D_{ti}$ and the log of $R\&D_{t-1i}$ in the empirical analysis. For notational conciseness we will delete log in the section.

[3]Here we have abstracted from conditioning on the explanatory variables.

If we assume that the patent count takes a finite number of outcomes (say a maximum of Y) and that λ_i takes on a particular value, we can always determine numbers $\eta_0, \eta_1, .., \eta_{Y-1}$ such that:

$$\Pr(y_{ti} = y; \lambda_i | R\&D_{t-1i}) = \int_{\eta_{y-1}}^{\eta_y} \phi(u)du \qquad y = 1, .., Y-1 \qquad \text{(5a)}$$

where $\phi(.|.)$ is the density of the standard normal distribution function. For observations $y = 0$ and $y = Y$ we specify:

$$\begin{aligned} \Pr(y_{ti} &= 0; \lambda_i | R\&D_{t-1i}) = \int_{-\infty}^{\eta_0} \phi(u)du \\ \Pr(y_{ti} &= Y; \lambda_i | R\&D_{t-1i}) = \int_{\eta_{K-1}}^{+\infty} \phi(u)du \end{aligned} \qquad \text{(5b)}$$

(5a) and (5b), given some λ_i, define $\eta_0, \eta_1, .., \eta_{Y-1}$ uniquely. The relation between the probabilities $P(y_{ti} = y; \lambda_i | R\&D_{t-1i})$ and the η_y's can also be written as:

$$\Pr(y_{ti} \leq \kappa; \lambda_i | R\&D_{t-1i}) = \sum_{y=0}^{\kappa} P(y_{ti} = y; \lambda_i | R\&D_{t-1i}) = \Phi(\eta_\kappa) = \int_{-\infty}^{\eta_k} \phi(u)du \quad (6)$$

where we define $\eta_Y = +\infty$. Consequently:

$$\eta_\kappa = \Phi^{-1}\left(\sum_{y=0}^{\kappa} P(y_{ti} = y; \lambda_i | R\&D_{t-1i})\right) \qquad (7)$$

This relation defines η_y ($y = 0, .., Y-1$) uniquely for any value of θ, and clearly η_y is a function of λ_i: $\eta_y(\lambda_i)$. As a result we have related an arbitrary discrete distribution to the normal distribution. For count data the random variable y_{ti} is unbounded. However, this does not cause any problems in the estimation of the model. The loglikelihood function only contains the probability of actual, and therefore bounded, observations.

The density of $R\&D_{ti}$ can also be written in terms of the normal distribution whatever the distribution $h(.)$ (Lee, 1983):

$$R\&D^*_{ti} = \Phi^{-1}(H(R\&D_{ti} | y_{t-1i})) \qquad (8)$$

where $\Phi^{-1}(.)$ is the inverse of the standard cumulative normal distribution function and $H(R\&D_{ti}) = \int_{-\infty}^{R\&Dt_i} h(v)dv$. $R\&D^*_{ti}$ has a standard normal distribution. Following Lee (1983), if ε_j follows the distribution $F_j(\varepsilon_j)$, $u_j = \Phi^{-1}(F_j(\varepsilon_j))$ has a standard normal distribution.

A cumulative bivariate distribution having marginal distributions $F_1(\varepsilon_1)$ and $F_1(\varepsilon_1)$ and a correlation ρ_ε between ε_1 and ε_2 is given by:

$$H(\varepsilon_1, \varepsilon_2; \rho_\varepsilon) = B(u_1, u_2; \rho_u) = B(\Phi^{-1}(F_1(\varepsilon_1)), \Phi^{-1}(F_2(\varepsilon_2)); \rho_u) \tag{9}$$

where $B(.,.;\rho_u)$ is the cumulative bivariate normal distribution with zero means, unit variances and correlation ρ_u.[4] In our case we have a combination of a discrete and an actually observed continuous random variable, and therefore we need the first derivative of (9) with respect to the continuous variables (say, ε_2):

$$\frac{\partial}{\partial \varepsilon_2} H(\varepsilon_1, \varepsilon_2; \rho_\varepsilon) = \frac{f_2(\varepsilon_2)}{\phi(F_2(\varepsilon_2))} \frac{\partial B(u_1, u_2; \rho_u)}{\partial u_2} \tag{10}$$

The likelihood function of the observed pair $(y_{ti}, R\&D_{ti})$ equals:

$$\begin{aligned} L_i &= g(y_{ti} = y, R\&D_{ti} = r | R\&D_{t-1i}, y_{t-1i}) \qquad (2.1) \\ &= g(y_{ti} \leq y, R\&D_{ti} = r | R\&D_{t-1i}, y_{t-1i}) - g(y_{ti} \leq y-1, R\&D_{ti} = r | R\&D_{t-1i}, y_{t-1i}) \end{aligned}$$

Using (10) we have:

$$g(y_{ti} \leq y, R\&D_i = r | R\&D_{t-1i}, y_{t-1i}) = \frac{h(r)}{\phi(H(r))} \frac{\partial B\left(\Phi^{-1}\left(\sum_{k=0}^{y} P(y_{ti} = k)\right), \Phi^{-1}(H(r)); \rho_u\right)}{\partial \Phi^{-1}(H(r))} \tag{12}$$

This simplifies to (Maddala, 1983, p. 273):

$$g(y_{ti} \leq y, R\&D_i = r | R\&D_{t-1i}, y_{t-1i}) = h(r)\Phi\left(\frac{\Phi^{-1}\left(\sum_{k=0}^{y} P(y_{ti} = k)\right) - \rho_u \Phi^{-1}(H(r))}{\sqrt{1-\rho_u^2}}\right) \tag{13}$$

which can be substituted into (11) and then (4).

[4]Although ρ_ε and ρ_u are closely related, they are not the same. Numerical analysis shows that the signs and the order of magnitude are always the same (cf. Van Ophem, 1999).

3. Data.

We use data from the Dutch part of the *Community Innovation Survey (CIS)* which is available for the years 1988 and 1992. The population of interest in this survey are the firms with ten or more employees in all manufacturing and service sectors of the Dutch economy. The original sample size consists of about 4000 firms for both years. We restrict this sample to the firms conducting permanent R&D activities and of course, for which we have information for both years. The resulting sample contains 460 observations. Additional information about the CIS can be found in Brouwer (1997). The endogenous variables of the analysis are the patent count and the natural logarithm of R&D expenditures. The patent count relates to the patents submitted to the European Patent Office in München. We must content with using R&D expenditures since we have no information of past R&D to construct an R&D stock. Summary statistics are listed in Table 1. A salient detail is the large difference between the mean and variance of the patent counts. At first sight it is thus questionable to assume a Poisson distribution for the count, where the mean and the variance are the same. The patent count ranges from 0 to a maximum of 22, but the mean is below 1 patent per firm. There is thus a high frequency of zero counts.

-Insert Table 1-

The explanatory variables used in the estimations all relate to 1988. They are:

- firm size: the logarithm of the number of employees in the firm.

  Firm size is expected to have a positive effect on R&D. What is less clearcut is whether R&D increases more or less proportionately with firm size. There might be scale advantages, but there might also be a threshold effect (see Bound et al. (1984) for empirical evidence). Since the size effect is not our primary concern, we only introduce a linear term. The impact of size on patent count is also debatable. On the one hand, large firms might exploit their first-mover advantage rather than patenting to secure the appropriation of R&D benefits, but, on the other hand, large firms are better equipped to apply for patents and face future litigation battles.

- R&D collaboration: a dummy variable equal to 1 if the firm is engaged in an R&D collaboration with other firms.

R&D collaboration will probably have a positive impact on both the patent count and R&D. Collaborating firms will have a higher propensity to seek patent protection since they have to reveal information to their partners. We expect the effect of R&D collaboration on R&D to be positive. R&D collaboration allows firms to internalize their mutual R&D spillover and thereby to increase the returns to their R&D efforts (see d'Aspremont and Jacquemin (1988) for a theoretical discussion of this result).

- sectoral dummy variables equal to 1 if the firm's principal activity is in that sector and zero otherwise.

  The sectors are service, food and beverages, wood processing, chemicals, plastics and rubber, metals, machinery, electrical equipment, and transportation equipment. The reference group is all other sectors.

- average product life cycle in a sector in years.

  The average product life cycle should show up with a negative coefficient in both the patent count and R&D since a shorter life-cycle of products will increase the efforts of firms to renew their products more frequently.

4. Empirical results.

Tables 2 and 3 give the estimation results of the model described in section 2 under the assumptions of a Poisson distributed patent count (1) and a normal distribution for the error term of the $\log(R\&D_t)$ equation (3). Table 2 contains the estimates for the case that patent count and $\log(R\&D_t)$ are not correlated. This case can be estimated by a Poisson regression and ordinary least squares. R&D increases with size, patents not, whereas cooperative R&D increases the number of patents but not the amount of R&D. R&D and patents are not related to the average product life cycle but are cross-correlated over time. Table 3 contains the full maximum likelihood estimation for the specification that allows a non zero correlation between the error terms of the R&D and patent equations. Size now exerts a positive effect on both patent applications and R&D expenditures. The estimated marginal effect of size on the number of patent applications increases. The average product life cycle shows a negative coefficient on the amount of R&D. The correlation coefficient between the two error terms is, however, insignificantly different from zero. Causality between patents and R&D still seems to run both ways. The causality running from R&D to patents weakens when a contemporaneous correlation between both is introduced through the error term.

-Insert Tables 2 and 3-

If we look at the distribution of the count (cf. Table 7, second column), we observe a very large proportion of zero counts. The question is whether this is not a little out of line with the Poisson distribution estimated. To check this we have estimated a Without Zero Poisson model as proposed by Mullahy (1986). Broadly speaking, only the positive counts follow a Poisson distribution in this model and the probability of the zero count is estimated by a constant. The following distribution is estimated:

$$\begin{aligned} \Pr(y_{ti} &= 0) = \psi + (1-\psi)\phi(0) \\ \Pr(y_{ti} &= y) = (1-\psi)\phi(y) \text{ for } y > 0 \end{aligned} \tag{4.1}$$

where $\phi(y)$ is the chosen basic distribution function of the count. As we can see from table 4, treating the frequent occurence of zero patents differently from the other patent counts increases the contemporaneous correlation between R&D and patents but reduces their cross intertemporal effects. The coefficient of R&D on patents becomes insignificant. The estimates of the R&D equations are robust to this change of specification. The additional parameter ψ is significant.

-insert Table 4-

To investigate whether the assumption of a Poisson distributed count is too restrictive, we now turn to the Katz family of distributions (Katz, 1965 or Winkelmann, 1997, p.35). This type of distribution nests several other distributions for non-negative integers, while maintaining a parsimonious parameterization. It is defined by the recursive probabilities:

$$\frac{\Pr(y_{ti} = y | R\&D_i)}{\Pr(y_{ti} = y-1 | R\&D_i)} = \frac{\omega + \theta(y-1)}{y} \qquad y = 1, 2, ...; \omega > 0 \text{ and } y \leq \omega/\theta \text{ for } \theta < 0 \tag{15}$$

The Poisson ($\omega = \lambda$, $\theta = 0$), Negative Binomial, Geometric and Binomial distributions are special cases (Winkelmann, 1997, p. 36). To make the Poisson specification of section 3 a special case we assume that ω is individual specific

($\omega = \exp(X_i'\beta)$) and that θ is not. Because the implicit probabilities have to sum up to 1, the probability of a zero count can be simply derived:

$$\begin{aligned}
\Pr(y_{ti} = 0|R\&D_i) &= 1 - \sum_{k=1}^{\infty} \Pr(y_{ti} = k|R\&D_i) \qquad (16)\\
&= 1 - \sum_{k=1}^{\infty} \prod_{j=1}^{k} \frac{\omega + \theta(j-1)}{j} \Pr(y_{ti} = 0|R\&D_i)\\
&= \left(1 + \sum_{k=1}^{\infty} \prod_{j=1}^{k} \frac{\omega + \theta(j-1)}{j}\right)^{-1}
\end{aligned}$$

From this probability all other probabilities can be derived and the application of the method proposed in section 3 is straightforward. The estimation results can be found in Table 5 and Table 6 (respectively the with and without zero patent count specifications). The new parameter θ is significantly different from zero, proving formally our suspicion that the Poisson distribution is too restrictive. If we compare tables 3 and 5, we notice that the estimated correlation between the error terms increases. In the patent equation size and lagged R&D become insignificant. The estimates of the R&D equation remain pretty much the same. Combining the Katz distribution and the without zero patent specification increases the contemporaneous correlation coefficient of the error terms, reinforces the size elasticity in the patent equation, but does not resurrect any causality running from R&D to patents. The absence of Granger-causality from R&D to patents is broadly consistent with the predominantly contemporaneous R&D effect and the non-significance of lagged R&D effects on patents found in most other studies (see Cincera, 1997). Again, the R&D estimates remain largely unaffected.

-insert Tables 5 and 6-

To investigate the fit we employ the method discussed in Winkelmann (1997, p. 162). We predict the patent count distribution of the firms and after aggregating across all firms, we obtain a sampling distribution under the specified model. A Pearson χ^2-test is performed to check whether the sampling distribution fits the data. In doing this we need to limit the count and we opted for a maximum count of 8 (only 2 firms where granted more than 8 patents). The results of these calculations are listed in Table 7.

-insert Table 7-

From this table we can conclude that the best performing model is the one based on the Katz-system of distributions and a without zero patent count specification. Based on these estimates, we can assert that one additional patent applied for (assuming that on average an application gets granted) yields four years down the road a 7.5% increase in R&D expenditures. The interpretation could be that patents pave the way to a stream of development expenditures in order to bring the patented product to the market or to additional R&D aiming to develop complements to the patented product. The elasticity of patents with respect to R&D (from the contemporaneous correlation of the error terms) is on the order of 0.4, which is closer to the time-series estimates reported in the literature (see Griliches, 1990) than to the cross-section estimates that they are supposed to be comparable with.

5. Conclusion

From cross-sectional data on contemporaneous and four-year lagged patent applications and R&D expenditures we have reexamined the causality direction between R&D and patents. The two equations have been estimated jointly with a contemporaneous correlation working through the error terms. We have experimented with different specifications of the count data for patent applications: the Poisson distribution, the negative binomial distribution and the Poisson and negative binomial without zero patent specifications. The negative binomial without zero patent specification model performs best (as was also found by Licht and Zoz (1998) and Crépon and Duguet (1997)).

We find that patents Granger-cause R&D in all specifications. One additional patent increases R&D four years later by 7.5%. The reverse causality from R&D to patents vanishes as soon as we depart in one way or another from the simple Poisson specification of patent counts. Although our result should be confirmed by analyzing other datasets and by checking how sensitive our estimates are to other specifications (such as including non-linear size effects, modeling more explicitly the contemporaneous linkage between R&D and patents, and introducing innovative sales in the picture) we might have uncovered a different causality from the conventional one estimated by other authors.

References

[1] Acs, Z., and D.B. Audretsch, 1989, Patents as a Measure of Innovative Activity, *Kyklos*, vol. 42, p 171-180.

[2] D'Aspremont, C. and A. Jacquemin, 1988, Cooperative and noncooperative R&D in duopoly with spillovers, *American Economic Review*, 78(5), 1133-1137.

[3] Bound, J., C. Cummins, Z. Griliches, B.H. Hall and A. Jaffe, 1984, Who Does R&D and Who Patents?, chapter 2 of *R&D, Patents, and Productivity*, Z. Griliches (ed.),National Bureau of Economic Research, The University of Chicago Press, Chicago.

[4] Brouwer, E., 1997, *Into Innovation: Determinants and Indicators*, Ph.D.-thesis, University of Amsterdam.

[5] Cincera, M., 1997, Patents, R&D and Technological Spillovers at the Firm Level: Some Evidence from Econometric Count Models for Panel Data, *Journal of Applied Econometrics*, 12(3), 265-280.

[6] Crépon, B. and E. Duguet, 1997, Estimating the Innovation Function from Patent Numbers: GMM on Count Panel Data, *Journal of Applied Econometrics*, 12(3), 243-264.

[7] Dosi, G., 1988, Sources, Procedures, and Microeconomic Effects of Innovation, *Journal of Economic Literature*, vol. 26, pp. 1120-1171.

[8] Griliches, Z., 1981, Market Value, R&D and Patents, *Economics Letters*, vol. 7, pp. 183-187.

[9] Griliches, Z., (ed.), 1984, R&D, Patents, and Productivity, National Bureau of Economic Research, The University of Chicago Press, Chicago.

[10] Griliches, Z., 1990, Patent Statistics as Economic Indicators: a Survey, *Journal of Economic Literature*, vol. 28, pp. 1661-1707.

[11] Griliches, Z., 1995, R&D and Productivity: Econometric Results and Measurement Issues, chapter 3 of the *Handbook of the Economics of Innovation and Technological Change*, P. Stoneman (ed.), Basil Blackwell, Oxford.

[12] Griliches, Z., B. Hall and A. Pakes (1991), "R&D, patents and market value revisited: Is there a second (technological opportunity) factor?, *Economics of Innovation and New Technology*, 1(3), 183-201.

[13] Hall, B.H., J. Hausman and Z. Griliches, 1986, Patents and R and D: Is There a Lag?, *International Economic Review*, vol. 27, pp. 265-283.

[14] Hausman, J., B.H. Hall and Z. Griliches, 1984, Econometric Models for Count Data with an Application to the Patents-R&D Relationship, *Econometrica*, vol. 52, pp. 909-938.

[15] Katz, L., 1995, Unified Treatment of a Broad Class of Discrete Probability Distributions, in *Classical and Contagious Discrete Distributions*, G.P. Patil (ed.), Statistical Publishing Society, Pergamon Press, New York.

[16] Lee, L.F., 1983, Generalized Econometric Models with Selectivity, *Econometrica*, vol. 51, pp. 507-512.

[17] Licht, G. and K. Zoz, 2000, Patents and R&D, An Econometric Investigation Using Applications for German, European and US Patents by German Firms, in D. Encaoua, B. Hall, F. Laisney and J. Mairesse (eds.), *The Economics and Econometrics of Innovation*. Kluwer Academic Publishers, Boston.

[18] Maddala, G.S., 1983, *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge, Massachusetts.

[19] Montalvo, J.G., 1997, GMM Estimation of Count-Panel-Data Models With Fixed Effects and Predetermined Instruments, *Journal of Business and Economic Statistics,* vol, 15, pp. 82-89.

[20] Mullahy, J., 1986, Specification and Testing of Some Modified Count Data Models, *Journal of Econometrics*, vol. 33, pp. 341-365.

[21] Pakes, A. (1985), On patents, R&D, and the Stock Market Rates of Return, *Journal of Political Economy*, 93, 390-409.

[22] Patel, P. and K. Pavitt, 1995, Patterns of Technologival Activity: their Measurement and Interpretation, chapter 2 of the *Handbook of the Economics of Innovation and Technological Change*, P. Stoneman (ed.), Basil Blackwell, Oxford.

[23] Stoneman, P., (ed.), 1995, *Handbook of the Economics of Innovation and Technological Change*, Basil Blackwell, Oxford.

[24] Van Ophem, H., 1999, A General Method to Estimate Correlated, Discrete Random Variables, *Econometric Theory*, 15, 228-237.

[25] Van Ophem, H., 2000, Modeling Selectivity in Count-Data Models, *Journal of Business & Economic Statistics*, 18, 503-511.

[26] Wang, P., I.M. Cockburn and M.L. Puterman, 1998, Analysis of Patent Data - A Mixed-Poisson-Regression-Model Approach, *Journal of Business & Economic Statistics*, vol. 16, p. 27-41.

[27] Winkelmann, R., 1997, *Econometric Analysis of Count Data*, second edition, Springer-Verlag, Berlin.

Table 1: Sample characteristics of the variables.

variable	*mean*	*st. dev.*	*minimum*	*maximum*
patent count 1992	0.457	1.662	0	22
log(R&D) 1992	1.270	1.603	-2.794	8.939
firm size (log of employees)	4.998	1.214	1.386	11.082
R&D collaboration	0.452	0.498	0	1
service	0.211	0.408	0	1
food & beverages	0.091	0.288	0	1
wood processing	0.072	0.258	0	1
chemicals	0.080	0.272	0	1
plastics & rubber	0.035	0.183	0	1
metals	0.126	0.332	0	1
machinery	0.174	0.379	0	1
electric equipment	0.050	0.218	0	1
transportation equipment	0.065	0.247	0	1
average product life cycle	10.638	1.945	4.526	14.821
patent count 1988	0.663	2.200	0	33
log(R&D) 1988	1.315	1.530	-3.219	9.616

Table 2: ML-estimation of a Poisson distributed patent equation and OLS estimation of log (R&D) in 1992

variables	*patent-count*		*log(R&D)*	
constant	-3.087	(0.886)**	-2.248	(0.452)**
firm size (log of employees)	0.165	(0.089)	0.679	(0.049)**
R&D collaboration	0.369	(0.160)*	0.194	(0.117)
service	0.320	(0.513)	0.289	(0.221)
food & beverages	-0.124	(0.652)	0.251	(0.267)
wood processing	0.942	(0.579)	0.066	(0.284)
chemicals	2.085	(0.520)**	1.648	(0.282)**
plastics & rubber	2.078	(0.550)**	0.221	(0.361)
metals	1.336	(0.528)**	0.234	(0.251)
machinery	1.901	(0.483)**	0.663	(0.232)**
electric equipment	0.049	(0.586)	1.421	(0.317)**
transportation equipment	0.543	(0.577)	0.773	(0.294)**
average product life cycle	-0.055	(0.073)	-0.049	(0.035)
log($R\&D_{1988}$)	0.301	(0.067)**		
patent count 1988			0.101	(0.027)**

Mean log-likelihood (count part) = -0.877; number of observations = 460. R^2 (R&D-part) = 0.445, variance R&D-error = 1.468
Asymptotic standard errors in parentheses.* = significant at 5%; ** = significant at 1% (two-sided test).

TABLE 3: ML-estimation patents/R&D-model for 1992: Poisson specification of the count.

variables	*patent-count*		*log(R&D)*	
constant	-3.001	(0.936)**	-2.883	(0.561)**
firm size (log of employees)	0.157	(0.095)*	0.822	(0.062)**
R&D collaboration	0.317	(0.164)*	0.207	(0.143)
service	0.304	(0.519)	0.354	(0.275)
food & beverages	-0.250	(0.649)	0.423	(0.330)
wood processing	0.918	(0.580)*	0.101	(0.355)
chemicals	1.721	(0.523)**	2.068	(0.344)**
plastics & rubber	1.942	(0.543)**	0.493	(0.440)
metals	1.189	(0.525)*	0.422	(0.312)
machinery	1.679	(0.486)**	0.959	(0.287)**
electric equipment	0.386	(0.587)	1.620	(0.383)**
transportation equipment	0.609	(0.574)	0.977	(0.360)**
average product life cycle	-0.017	(0.072)	-0.092	(0.043)*
log($R\&D_{1988}$)	0.189	(0.075)*		
patent count 1988			0.095	(0.032)**
	model parameters			
correlation	0.082	(0.062)		
variance error term log(R&D)	2.042	(0.159)**		

Mean log-likelihood = -2.385; number of observations = 460.
Standard errors in parentheses.* = asymptotically significant at 10%; ** = significant at 1% (two-sided test)

TABLE 4: ML-estimation patents /R&D-model for 1992: Without zero specification specification for the count.				
variables	*patent-count*		*log(R&D)*	
constant	-2.689	(1.352)*	-2.841	(0.559)**
firm size (log of employees)	0.422	(0.121)**	0.824	(0.061)**
R&D collaboration	0.462	(0.186)*	0.240	(0.141)*
service	0.603	(0.628)	0.394	(0.275)
food & beverages	-0.146	(0.788)	0.440	(0.330)
wood processing	1.196	(0.792)	0.106	(0.354)
chemicals	1.727	(0.693)*	2.023	(0.341)**
plastics & rubber	1.320	(0.695)*	0.394	(0.435)
metals	1.225	(0.670)*	0.434	(0.311)
machinery	1.833	(0.603)**	0.939	(0.285)**
electric equipment	0.695	(0.644)	1.562	(0.379)**
transportation equipment	0.971	(0.732)	0.984	(0.358)**
average product life cycle	-0.071	(0.126)	-0.098	(0.043)*
log($R\&D_{1988}$)	0.053	(0.095)		
patent count 1988			0.078	(0.032)*
	model parameters			
correlation	0.322	(0.095)**		
variance error term log(R&D)	2.047	(0.159)**		
ψ	0.692	(0.039)**		
Mean log-likelihood = -2.213; number of observations = 460. Standard errors in parentheses.* = asymptotically significant at 10%; ** = significant at 1% (two-sided test)				

TABLE 5: ML-estimation patents/R&D-model for 1992: Katz-system specification of the count.				
variables	*patent-count*		*log(R&D)*	
constant	-4.430	(1.329)**	-2.891	(0.562)**
firm size (log of employees)	0.193	(0.151)	0.817	(0.061)**
R&D collaboration	0.085	(0.240)	0.203	(0.143)
service	-0.213	(0.721)	0.350	(0.275)
food & beverages	-0.189	(0.816)	0.429	(0.330)
wood processing	1.036	(0.718)	0.119	(0.356)
chemicals	1.602	(0.678)*	2.067	(0.344)**
plastics & rubber	1.643	(0.731)**	0.497	(0.440)
metals	0.924	(0.674)**	0.422	(0.313)
machinery	1.465	(0.620)*	0.958	(0.287)**
electric equipment	1.062	(0.751)	1.639	(0.383)**
transportation equipment	0.759	(0.725)	0.987	(0.361)**
average product life cycle	0.027	(0.099)	-0.089	(0.043)*
log(R&D$_{1988}$)	0.055	(0.126)		
patent count 1988			0.082	(0.032)*
	model parameters			
correlation	0.319	(0.118)**		
variance error term log(R&D)	2.050	(0.159)**		
θ	0.712	(0.052)**		

Mean log-likelihood = -2.229; number of observations = 460.
Standard errors in parentheses.* = asymptotically significant at 10%; ** = significant at 1% (two-sided test)

Table 6: ML-estimation patents/R&D-model for 1992: Without zero Katz-system specification of the count.

variables	*patent-count*		*log(R&D)*	
constant	-4.314	(1.409)**	-2.961	(0.563)**
firm size (log of employees)	0.500	(0.151)**	0.838	(0.061)**
R&D collaboration	0.452	(0.233)*	0.255	(0.142)**
service	-0.129	(0.738)	0.345	(0.276)
food & beverages	-0.328	(0.854)	0.414	(0.332)
wood processing	1.207	(0.811)	0.117	(0.358)
chemicals	1.838	(0.703)**	2.046	(0.342)**
plastics & rubber	1.204	(0.748)	0.374	(0.437)
metals	1.231	(0.697)*	0.444	(0.313)
machinery	1.982	(0.641)**	0.980	(0.287)**
electric equipment	1.071	(0.736)	1.602	(0.380)**
transportation equipment	1.111	(0.764)	1.009	(0.360)**
average product life cycle	-0.030	(0.108)	-0.094	(0.043)*
log(R&D$_{1988}$)	0.057	(0.115)		
patent count 1988			0.075	(0.032)*
	model parameters			
correlation	0.390	(0.102)**		
variance error term log(R&D)	2.059	(0.161)**		
θ	0.402	(0.103)**		
ψ	0.589	(0.061)**		

Mean log-likelihood = -2.202; number of observations = 460.
Standard errors in parentheses.* = asymptotically significant at 10%;
** = significant at 1% (two-sided test)

Table 7: Predicted patent-count .						
count (Y)	sample	Poisson $\rho = 0$	Poisson	Poisson WZ	Katz	Katz WZ
0	381	360	317	375	374	379
1	36	75	101	41	42	37
2	21	16	28	21	19	18
3	6	4	8	10	10	10
4	3	2	3	5	6	5
5	5	1	1	3	3	3
6	5	1	1	1	2	2
7	0	0	0	1	1	1
8	1	0	0	1	1	1
$\geq$9	2	0	0	3	2	4
Pearson-χ^2	-	66.5*	114.0*	15.9	10.0	10.6

* = H_o: predicted count fits sample count rejected at 1%.

Liste des publications au CIRANO*

Série Scientifique / *Scientific Series* (ISSN 1198-8177)

2001s-30 Contracting under Ex Post Moral Hazard and Non-Commitment / M. Martin Boyer

2001s-29 Project Financing when the Principal Cannot Commit / M. Martin Boyer

2001s-28 Complementarities in Innovation Policy / Pierre Mohnen et Lars-Hendrick Röller

2001s-27 Bankruptcy Cost, Financial Structure and Technological Flexibility Choices / Marcel Boyer, Armel Jacques et Michel Moreaux

2001s-26 Inflation as a Strategic Response / M. Martin Boyer et Pierre Thomas Léger

2001s-25 Simulation-Based Finite-Sample Tests for Heteroskedasticity and ARCH Effects / Jean-Marie Dufour, Lynda Khalaf, Jean-Thomas Bernard et Ian Genest

2001s-24 The Role of Organizational Commitment and Citizenship Behaviors in Understanding Relations between Human Resources Practices and Turnover Intentions of IT Personnel / Guy Paré, Michel Tremblay et Patrick Lalonde

2001s-23 A Resource-Based Analysis of Outsourcing: Evidence from Case Studies / Vital Roy et Benoit Aubert

2001s-22 Short and Long Memory in Equilibrium Interest Rate Dynamics / Jin-Chuan Duan et Kris Jacobs

2001s-21 Unemployment Insurance and Subsequent Job Duration: Job Matching vs Unobserved Heterogeneity / Christian Belzil

2001s-20 Estimating the Intergenerational Education Correlation from a Dynamic Programming Model / Christian Belzil et Jörgen Hansen

2001s-19 The Bootstrap of the Mean for Dependent Heterogeneous Arrays / Sílvia Gonçalves et Halbert White

2001s-18 Perspectives on IT Outsourcing Success: Covariance Structure Modelling of a Survey of Outsourcing in Australia / Anne C. Rouse, Brian Corbitt et Benoit A. Aubert

2001s-17 A Theory of Environmental Risk Disclosure / Bernard Sinclair-Desgagné et Estelle Gozlan

2001s-16 Marriage Market, Divorce Legislation and Household Labor Supply / Pierre-André Chiappori, Bernard Fortin et Guy Lacroix

2001s-15 Properties of Estimates of Daily GARCH Parameters Based on Intra-Day Observations / John W. Galbraith et Victoria Zinde-Walsh

2001s-14 A Ricardian Model of the Tragedy of the Commons / Pierre Lasserre et Antoine Soubeyran

2001s-13 Carbon Credits for Forests and Forest Products / Robert D. Cairns et Pierre Lasserre

---

* Consultez la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site Internet :

http://www.cirano.qc.ca/publication/documents.html