Knowledge spillovers in U.S. patents: A dynamic patent intensity model with secret common innovation factors.

Szabolcs Blazsek[a] and Alvaro Escribano[b],[*]

[a]*Department of Business Administration, Universidad de Navarra, Ed. Bibliotecas-Este, 31080, Pamplona, Spain*

[b]*Department of Economics, Universidad Carlos III de Madrid, Getafe, Spain*

Abstract

During the past two decades, innovations protected by patents have played a key role in business strategies. This fact enhanced studies of the determinants of patents and the impact of patents on innovation and competitive advantage. Sustaining competitive advantages is as important as creating them. Patents help sustaining competivite advantages by increasing the production cost of competitors, by signaling a better quality of products and by serving as barriers to entry. If patents are rewards for innovation, more R&D should be reflected in more patents applications but this is not the end of the story. There is empirical evidence showing that patents through time are becoming easier to get and more valuable to the firm due to increasing damage awards from infringers. These facts question the constant and static nature of the relationship between R&D and patents. Furthermore, innovation creates important knowledge spillovers due to its imperfect appropriability. Our paper investigates these dynamic effects using U.S. patent data from 1979 to 2000 with alternative model specifications for patent counts. We introduce a general dynamic count panel data model with dynamic observable and unobservable spillovers, which encompasses previous models, is able to control for the endogeneity of R&D and therefore can be consistently estimated by maximum likelihood. Apart from allowing for firm specific fixed and random effects, we introduce a common unobserved component, or secret stock of knowledge, that affects differently the propensity to patent of each firm across sectors due to their different absorptive capacity.

[*] Corresponding author. Telefonica Chair of Economics of Telecommunications. Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, 28903, Getafe (Madrid), Spain.
Telephone: +34 916249854 (A. Escribano).
E-mail addresses: sblazsek@unav.es (S. Blazsek), alvaroe@eco.uc3m.es (A. Escribano).

1. Introduction

This paper uses firm-level data in order to study the dynamic interaction among inventive activity, R&D expenses and observable and unobservable knowledge of several firms jointly. The objective of the paper is to identify the observable and latent dynamics of firms' patent applications activity and to measure intra-industry and inter-industry spillovers of knowledge using inter-firm patent citations information over the 1979-2000 period in the United States. When an innovation produced by a particular firm increases the productivity of other companies then we say that the firm generates an R&D spillover. The analysis of the spillover process is a complex issue because both the R&D activity and the spillover process are partly unobservable. Previous studies have used either 'input side' measures (R&D expenditure) or 'output side' proxies (number of patents per year) of innovative activity. Present paper takes an output side view and we use patent data to measure the inventive activity of firms. Griliches (1990) states that the main advantages of patent data are the followings: (1) By definition patents are closely related to inventive activity. (2) Patent documents are objective because they are produced by an independent patent office and their standards change slowly over time. (3) Patent data are widely available in several countries, over long periods of time, and cover almost every field of innovation.

According to Fung (2005) patents have two economic functions: (1) legal protection of knowledge capital and (2) disclosure of the specifications of innovations. The second function facilitates information-sharing agreements among firms (such as licensing and patent-sharing agreements) and imitation of past innovations. Legal protection against imitation can be limited in those industries where competitors can develop close substitutes that are not in conflict with previous patents. (A good example is the computer software industry.) This suggests that patents and patent citations may be good measures of R&D spillovers.

Technological improvement gives the innovators a competitive advantage. However, the non-rival nature of knowledge creates a business-stealing (competitive) effect by decreasing the cost of subsequent own innovations. A spillover of knowledge occurs when a new innovation created by a technological leader firm is adopted by another (follower) firm. In the economic literature many researchers have analyzed knowledge spillovers.[1] Scherer (1981) constructs an inter-industry technology flows matrix to measure knowledge spillovers between industries. Jaffe (1988) classifies firms into different technological clusters to identify the proximity of firms in the technology space. Lanjouw and Schankerman (1999) and Hall et al (2001) validates the use of patent statistics in economic research. They suggest that the intensity of forward citations (the number of citations received from subsequent patents) can be used to measure the significance of innovations, while backward citations (citations made to previous patents) can be used to capture R&D spillovers. Harhoff et al (1999) combine German and U.S. patent value survey and backward citation data. They find that patents reported to be relatively more valuable by the companies holding them are more

---

[1]See Nabseth and Ray (1974), Mansfield et al (1981), Scherer (1981), Rogers (1983), Pakes and Schankerman (1984), Jaffe (1986), Jaffe (1988), Cockburn and Griliches (1988), Jovanovic and MacDonald (1994), Jaffe et al (1998), Jaffe et al (2000), McGahan and Silverman (2003) and Fung (2005).

heavily cited in subsequent patents. Fung and Chow (2002) look at potential knowledge pools at the industrial level. Fung (2005) uses patent citations data to analyze the impact of knowledge spillovers on the convergence of productivity among firms. Knowledge spillovers are decomposed into: (1) intra-industry spillovers, (2) inter-industry spillovers and (3) within-firm knowledge flows. The main finding of this study is that each firm is converging to its own steady-state productivity growth rate, which is conditional on the firm's R&D efforts and the intensity of intra-industry spillovers it receives. Jaffe, Trajtenberg and Fogarty (2000) survey R&D managers in order to validate the use of patent citations to approximate the unobservable process of knowledge transfer.

Jovanovic and MacDonald (1994) point out that innovation and imitation tend to be substitutes. Though, the benefits generated by spillovers depend on the technological differences among firms and the absorptive capacity of the imitator firm. Escribano et al (2009) argue that those firms with higher levels of absorptive capacity can manage external knowledge flows more efficiently and stimulate innovation outputs. Firms may enhance their R&D absorptive capacity by investing in their R&D activity. Naturally, these factors create time lags in the adoption of technologies. For example, Nabseth and Ray (1974) and Rogers (1983) report that it may take a decade for some firms to adopt an innovation developed by others. Mansfield et al (1981) and Pakes and Schankerman (1984) also suggest that knowledge 'spills over' gradually to competitors.

Although Jaffe (1986) also reports that the R&D spillover process is inherently dynamic, most of these papers have applied cross-sectional econometric frameworks to study knowledge flows among firms or industries. The first R&D papers, which employed dynamic econometric models (Hausman et al, 1984 or Pakes, 1985) focused on the within-firm relationships among input and output side measures of R&D activity, and various measures of firm performance (market or accounting value based measures). Recently some papers have appeared in dynamic panel data econometrics, which introduced new models to capture dynamics in count data (see for example Blundell et al, 2002 and Wooldridge, 2005). Jaffe (1986) mentioned that the main reason for the lack of dynamic R&D spillover models in the literature is that 'data are not rich enough to support a dynamic specification with spillover effects' (Jaffe, 1986). We think that besides the availability of good quality data, part of the problem is the econometric specification used in previous papers. Therefore, this paper proposes a new dynamic panel data model of knowledge flows. In addition, we also contribute to the R&D literature by estimating the model to a new extended patent data set, which covers significantly longer time period (1979-2000: 22 years) than other recent papers in this research field (Peters, 2007; Cockburn and Wagner, 2007; Song and Jones, 2007).

The major econometric problem related to the modeling of R&D spillovers is that several factors, which influence the innovative activity of firms are unobservable for the researcher. Therefore, in this paper we assume that the inventive activity of a firm is the product of the next two components: (1) Observable R&D intensity component and (2) Latent R&D intensity component.

The observable R&D intensity component accounts for econometrically observable factors like own and other firms' R&D expenses or the number of citations made by patent documents to own or other firms' patents. This component is modeled as a multivariate dynamic time series model. Inclusion of the latent R&D component can be justified by Griliches (1990), who notes that some

innovations are unobservable either because they do not meet the standards set by the patent office or because they are held in secrecy.

Similarly to several past studies, which used patent data to measure the inventive activity of firms, present paper employs a Poisson-type patent count model. By patent counts we mean the number of successful patent applications assigned to firms during a given year.[2] Common characteristic of these models is that patent counts are treated as discrete-valued random variables and are analyzed by a *count data model*. In these models it is assumed that: (1) the arrival rate (or *conditional intensity*) of patents has some parametric functional form, and (2) the arrival rate is constant over a period of time. The consequence of the second assumption is that the statistical inference of the model can be done based on the number of patent applications during each period and the exact time of the innovation is irrelevant. Although in recent patent databases the application date of patents is available with daily precision, this information is a noisy measure of the time of innovations. Therefore, following Hausman et al (1984) most authors aggregate patent data over the year. Thus, the patent counts are assumed to follow a Poisson distribution. Fortunately, Wooldridge (2002) notes that the Poisson distribution has a very nice robustness property: whether or not the Poisson distribution holds, we still get consistent, asymptotically normal estimators of the parameters provided that the conditional mean function is correctly specified.

The main contribution of this paper is that we include more general dynamic components, observable spillovers and unobservable spillovers by including a dynamic latent factor into the classical model of Hausman et al (1984). The latent variables commonly drive the patent application intensity of all firms and the specification makes possible to separate observable and latent R&D spillovers. This paper also contributes to the most recent literature in that we concentrate on a well specified general dynamic model which may be a contribution to solving the endogeneity problem of R&D. This is important since most of the previous results from alternative GMM procedures and maximum likelihood (Cincer, 1997, Montalvo, 1997 and Blundell et al, 2002) obtained different numerical values for the coefficient of R&D but the sign was always positive. The specification of our model is estimated by maximum simulated likelihood (MSL) method.

A question that has generated long debates is how monopoly rights (patents, etc.) and competition affect innovation and productivity growth.

In the recent theoretical and empirical literature on the competition effects on innovation and productivity growth[3], there are two clear opposite views: Innovation under competition reduces *innovations rents*, relative to the monopoly rents, but innovation is also a mechanism to escape competition (*competitive advantage*) and in that sense increases *innovation rents*.

These two opposite views are also express as follows: First, a "*rent dissipation effect of competition*" which says that tough competition *discourage innovation* and productivity growth by reducing the expected rents from innovation. By reducing the monopoly rents, competition discourage firms from doing R&D activities which lower the innovation rate and the long run growth. The initial endogenous growth models of technical change of Romer (1990), Aghion and Howitt

---

[2]See Hausman et al (1984), Pakes (1985), Lanjouw et al (1998) or Trajtenberg (2002).

[3]Aghion and Griffith (2005) provide an interesting overview.

(1992), Grossman and Helpman (1991), predict that competition (or the imitation rate) has a negative effect on entry and innovation and therefore on productivity growth. That is, patent protection protects monopoly rents from innovation, enhancing further innovation and growth (Schumpeterian view). Second the "*escape competition effect*", followed by most competition authorities, says that competition is a necessary input for innovation both because it encourages new entry and because it *forces incumbent firms to innovate* and reduce costs to survive and therefore is productivity and growth enhancing.

Which of the two competition effects of dominates is an empirical question. For example, Aghion, Boom, Blundell, Griffith and Howit (2003) estimate an inverted-U relationship between innovation (citation-weighted patent count) and product market competition which is steeper for more neck-to-neck industries. In our empirical application, we also obtained an inverted-U relationship between R&D (after controlling for patent citations) and innovation (measured by patent application counts).

Finally, Crépon et al (1998) study the relationship between productivity, innovation and research at the firm level using a structural model. In particular they found that firm innovation output raises with research effort and other indicators through theirs effects on research and that firm productivity correlates positively with innovation output (patents).

Remaining part of the paper is structured as follows. First, we introduce the econometric model in Section 2. Then, Section 3 discusses the estimation method. Section 4 describes the patent and firm specific data. Section 5 summarizes our results. Finally, Section 6 concludes.

2. The model

2.1. Notation and definitions

We model the R&D activity of $i = 1, \ldots, N$ firms jointly over $t = 1, \ldots, T$ periods. Denote n_{it} the number of patent applications of the i-th firm in the t-th period. Let z_i denote firm size measured by the logarithm of the inflation adjusted book value. Let r_{it} denote log-R&D expenditure corresponding to the i-th firm in the t-th period. Furthermore, let $c_{it} = (c_{1it}, c_{2it}, c_{3it})'$ denote a 3×1 vector capturing observable R&D spillovers. The elements of c_{it} represent three components of the observable spillover of knowledge from three knowledge pools (see Fung, 2005): (1) *Own knowledge pool*: knowledge produced by the firm in the past, c_{1it}, (2) *Intra-industry knowledge pool*: knowledge produced by other firms in the same industry, c_{2it}, (3) *Inter-industry knowledge pool*: knowledge produced in other industries, c_{3it}.

In order to simplify the notation, we shall also use the $n_t = \{n_{it}\}, z = \{z_i\}, r_t = \{r_{it}\}, c_t = \{c_{it}\}$ for $i = 1, \ldots, N$ matrix notation corresponding to the above random variables. Thus, the observable variables form an $N \times T$ dimensional panel. Finally, let l_t^* denote the value of a latent variable in the t-th period, interpreted as an econometrically unobservable common factor (unobservable innovations) affecting the patenting behavior of all firms.

In this paper, similarly to Hausman, Hall and Griliches (HHG, 1984), we characterize the stochastic process $\{z, n_t, r_t, c_t, l_t^*\}$ by the conditional intensity. Nevertheless, there are at least two

important differences between our framework and HHG. First, we model the interaction among the patent application intensities by a multivariate specification of lagged R&D expenses and other variables, which measure patent citations. Second, we include a set of serially correlated latent variables into the model, which commonly drive the patent intensity of all firms. These latent variables are interpreted as the impact of unobserved common stock of knowledge (secret specific production processes, secret innovations, etc.) on the patent intensity of firms. The information set in the t-th period is generated as follows:

$$\mathcal{F}_t = \sigma[z, (n_1, r_1, c_1, l_1^*), \ldots, (n_{t-1}, r_{t-1}, c_{t-1}, l_{t-1}^*), (r_t, c_t, l_t^*)]. \tag{2.1}$$

Notice that the information set $\mathcal{F}_t$ includes the contemporaneous values of z, r_t, c_t and l_t^*. Similarly to HHG, the patent application intensity of firms is modeled by specifying the conditional hazard function of the point process formed by the patent arrival times. Define the $\mathcal{F}$-conditional hazard function corresponding to the i-th firm in the t-th period as follows (see Cox and Isham, 1980):

$$\lambda_{it}(\tau) = \lim_{\delta \to 0} \frac{\Pr\{n_{it}(\tau + \delta) - n_{it}(\tau) > 0 | \mathcal{F}_t\}}{\delta}, \tag{2.2}$$

where $\delta > 0$ and $n_{it}(\tau)$ is the number of patents of the i-th firm until time τ in the t-th period. In the remaining part of this paper the conditional hazard is assumed to be constant during each period, therefore, it can be indexed by t as follows: $\lambda_{it} = \lambda_{it}(\tau)$. The λ_{it} can be interpreted as the instantaneous probability that the i-th firm has a new patent at any point of time of the t-th period given all information available in the beginning of the t-th period. Thus, the conditional hazard, λ_{it} represents the patent application intensity of the i-th firm in the t-th period. Since the conditional hazard is assumed to be constant during each period, the statistical inference of the model can be done based on the number of patents occurred in each time interval.

2.2. Existing models

First, we shall give a brief overview of the patent applications 'basic Poisson models' of HHG that we extend in this paper. The basic specification of the conditional hazard function in HHG is

$$\lambda_{it} = \exp(\mu_0 + X_{it}^{'}\beta), \tag{2.3}$$

where μ_0 is a constant and

$$X_{it}^{'}\beta = \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \beta_5(L)r_{it}, \tag{2.4}$$

where D_i is a scientific sector dummy taking the value one for firms in the drug, computer, scientific instrument, chemical and electronic components industries ('high-tech' industries) and zero otherwise. Furthermore, $\beta_5(L) = \sum_{k=0}^{5} \beta_k L^k$ is the lag polinomial of five lags, which capture the contemporaneous and lagged impact of R&D expenses on patent counts. HHG estimates various specifications of this model by maximum likelihood as well as OLS methods and in both cases find "a somewhat disturbing U-shaped distributed lag, $\beta_5(L)$, which may well indicate a substantial truncation effect" after five lags and can create serial correlation given the persistence in R&D.

In order to solve the misspecification problem of the basic Poisson model, HHG proposes two alternative specifications for unobserved heterogeneity, which allow for latent firm specific *random effects* or *fixed effects*. Their proposed specification of the conditional hazard is the following:

$$\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \mu_i) = \exp(\mu_0 + X'_{it}\beta)\alpha_i, \tag{2.5}$$

where μ_i represents latent firm specific effects and α_i is defined by the second equality.

In case of *random effects* specification, HHG assumes that $\{\alpha_i\}$ are i.i.d. gamma distributed random variables with $E[\alpha_i] = 1$ and $Var[\alpha_i] = 1/\delta$. Analyzing the estimation results, HHG concludes that the inclusion of firm specific random effects somewhat "attenuates the U-shaped lag structure of R&D, but there is still a significant positive coefficient on the first lag".

When the *fixed effects* specification is estimated, the model takes the same functional form as in case of random effects besides that the $\{\alpha_i\}$ are treated as parameters of the model and not as latent i.i.d. variables, therefore, the μ_0 constant is not identified. The authors find that the coefficients of past R&D expenditure are all small and insignificant and the U-shape of the distributed lag no longer appears. In this specification, the firm specific effect, α_i represents both the accumulated stock of knowledge from past R&D in the firm and unobserved permanent differences across the firms, which affect their propensity to patent.

2.3. Extensions

We extend the HHG model in several directions. Most of these extensions are related to the fact that the HHG models intend to capture the lagged impact of R&D on own patent intensity and do not attempt to measure knowledge spillovers (observable and unobservable), which is the main goal of our paper. Essentially, we extend the HHG model in three ways: First, we introduce dynamic structures into the model of R&D activity. We concentrate on well specified dynamic models, which may be a contribution to solving the endogeneity problem of R&D. This is important since many previous results from alternative GMM procedures and maximum likelihood obtained different numerical values for the coefficient of R&D although the sign was always positive. Second, we use additional explanatory variables capturing observable knowledge spillovers such as patent citations, c_{it}. Moreover, we also include several industry dummy variables and the square of R&D expenditure into the HHG model. Third, we include a set of serially correlated latent variables into the model in order to account for the R&D spillover process not captured by observable variables.

The first extension that we consider is the 'dynamic Poisson model', which introduces serial correlation into the Poisson model of HHG. We use the dynamic panel count data model of Blundell et al (2002) and Wooldridge (2005):

$$\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \mu_i), \tag{2.6}$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1} + \Psi_{it} \tag{2.7}$$

and the dynamic term is given by:

$$\Psi_{it} = \beta_q(L)r_{it} + \kappa n_{it-1}, \tag{2.8}$$

7

where $\beta_q(L) = \sum_{k=0}^{q} \beta_k L^k$ measures the impact of lagged R&D expenses and κ measures the impact of lagged number of patent applications.

Next, we extend the dynamic Poisson model of Wooldridge (2005) by including square-R&D expenditure and 15 industry dummies into the model. Cockburn and Griliches (1988) state that patents are one of the mechanisms by which firms may appropriate returns from their R&D investments, and the effectiveness of patent protection varies significantly across industries. Therefore, we also include several industry dummies into the specification. The extended Poisson model is formulated as follows:

$$\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \mu_i), \tag{2.9}$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1} + \Psi_{it} + \bar{\beta} r_{it}^2 + \sum_{j=1}^{15} \rho_j I_{ij} \tag{2.10}$$

where ρ_j measures the impact of the 15 industry dummy variables, I_{ij}, $\bar{\beta}$ captures the non-linearities in R&D expenditure and the dynamic term is given by equation (2.8).

Furthermore, in our next extension of the models, we apply the 'dynamic Poisson model with observable R&D spillovers', which permits us to study the impact of past R&D and patent citations by the following formulation:

$$\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \mu_i), \tag{2.11}$$

$X'_{it}\beta$ is given by equation (2.10) and the dynamic term is given by:

$$\Psi_{it} = \beta_q(L)r_{it} + \kappa n_{it-1} + r_{it}[\omega_q(L)c_{2it}] + r_{it}[\phi_q(L)c_{3it}]. \tag{2.12}$$

where $\beta_q(L) = \sum_{k=0}^{q} \beta_k L^k$ measures the impact of lagged R&D expenses and κ measures the impact of lagged number of patent applications. Finally, $\omega_q(L) = \sum_{k=0}^{q} \omega_k L^k$ and $\phi_q(L) = \sum_{k=0}^{q} \phi_k L^k$ measure the product of lagged patent citations, c_{2it} and c_{3it} and contemporaneous R&D expenses, r_{it}. Remember that the c_{2it} denotes total patent citations received from the same industry, while c_{3it} is total patent citations received from other industries. Negative values of the corresponding parameters, ω and ϕ mean negative interaction between cited past knowledge and present research expenses. This way we can measure significant lagged observable R&D spillovers.

Finally, as previous literature reported that significant portion of knowledge spillovers cannot be observed directly, we include additional, time dependent latent variables into the model, which are possibly cross correlated along the lines of Pesaran (2006) and Holly et al (2006). Thus, the proposed structure captures the latent correlation structure of various firms' R&D activity. Our 'latent-factor dynamic Poisson model' takes the following form:

$$\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \mu_{it}) = \exp(\mu_0 + X'_{it}\beta)\alpha_{it}, \tag{2.13}$$

where α_{it} is defined in equation (2.13) and $X'_{it}\beta$ is given by equation (2.10). Moreover, from equation (2.12) it follows that the dynamic term is:

$$\Psi_{it} = \sum_{k=0}^{q} \beta_k r_{it-k} + \kappa n_{it-1} + \sum_{k=0}^{q} \omega_k c_{2it-k} r_{it} + \sum_{k=0}^{q} \phi_k c_{3it-k} r_{it} \tag{2.14}$$

The dynamic latent patent intensity component, μ_{it} measures the impact of unobserved knowledge flows on patent intensity. The latent patent intensity component is an extension of the stochastic conditional intensity (SCI) model of Bauwens and Hautsch (2006):

$$\alpha_{it} = \exp(\sigma_i l_t^*) \tag{2.15}$$

and

$$l_t^* = \mu l_{t-1}^* + \eta_t, \tag{2.16}$$

where $\sigma_i \in \mathbb{R}$ measures the firm specific impact of the latent R&D common factor, $\eta_t \sim \mathcal{N}(0,1)$ are i.i.d. innovation terms, which measure the general stock of unobserved knowledge and $|\mu| < 1$ captures the dynamics of the latent variables.[4] Since the sign of σ_i is not restricted, this unobserved common stock of knowledge is absorbed differently by each firm. The latent factor may have positive impact on the patent intensity of some firms and in the same time negative impact on the patent application intensity of other firms (we allow for cross section dependence in the panel data count model). This specification allows us to separate observable and latent knowledge spillovers and also to study the dynamics of the observable and unobservable R&D spillover process among firms due to the AR(1) specification of the latent process. More specifically, the latent component allows for correlation of the patent application intensity of various firms because:

$$cov(\sigma_i l_t^*, \sigma_j l_t^*) = \frac{\sigma_i \sigma_j}{1 - \mu^2} \tag{2.17}$$

Thus, following Pesaran (2006), the latent common factor is responsible for the unobservable knowledge cross section interaction between firms.

3. Inference

3.1. Log-likelihood functions

All models presented in the previous section are estimated by maximum likelihood or maximum simulated likelihood (MSL) method (see Appendix A).

Basic Poisson model. The log-likelihood of the basic Poisson model is given by (Gouriéroux, 1984, Chapter XI, pp.302):

$$\ln \mathcal{L} = \sum_{t=1}^{T} \sum_{i=1}^{N} n_{it} \ln \lambda_{it} - \ln(n_{it}!) - \lambda_{it}. \tag{3.1}$$

---

[4]We assume that the innovation terms in the latent AR(1) process of l_t^* are i.i.d. because we follow the specification of the latent intensity component and the statistical inference procedure proposed by Bauwens and Hautsch (2006). The statistical inference of the latent-factor Poisson model would be more complicated if $\{\eta_t\}$ formed a serially correlated time series.

Poisson model with random effects. The log-likelihood of the random effects model can be obtained by integrating the likelihood of n_{it} with respect to the assumed gamma distribution (Gouriéroux, 1984, Chapter XI, pp.304):

$$\int_0^\infty \frac{\exp(-\lambda_{it})\lambda_{it}^{n_{it}}}{n_{it}!} \tilde{f}(\alpha_i)d\alpha_i, \tag{3.2}$$

where the density of the gamma distribution with $E[\alpha_i] = 1$ and $Var[\alpha_i] = 1/\delta$ is given by

$$\tilde{f}(u) = \frac{\delta^\delta u^{\delta-1}\exp(-\delta u)}{\Gamma(\delta)}, \tag{3.3}$$

where $\Gamma(x)$ denotes the gamma function. Integrating by parts we obtain that the log-likelihood of the random effects model is the following:

$$
\begin{aligned}
\ln \mathcal{L} \ &= \sum_{t=1}^{T}\sum_{i=1}^{N} \ln\Gamma(\delta + n_{it}) - \ln\Gamma(\delta) - \ln\Gamma(n_{it}+1) + n_{it}\ln\left[\frac{\exp(\mu_0+X_{it}'\beta)}{\delta}\right] \\
&\quad - (n_{it}+\delta)\ln\left[1 + \frac{\exp(\mu_0+X_{it}'\beta)}{\delta}\right].
\end{aligned}
\tag{3.4}
$$

Poisson model with fixed effects. As we have seen in the previous section, the Poisson model with fixed effects has the following form:

$$\lambda_{it} = \exp(\mu_i + X_{it}'\beta). \tag{3.5}$$

Substituting this equation into (3.1) we obtain the log-likelihood of the fixed effects model:

$$\ln\mathcal{L} = \sum_{t=1}^{T}\sum_{i=1}^{N} n_{it}(\mu_i + X_{it}'\beta) - \ln(n_{it}!) - \exp(\mu_i + X_{it}'\beta). \tag{3.6}$$

Solving the first-order condition $\partial \ln\mathcal{L}/\partial\mu_i = 0$ for μ_i we get:

$$\exp(\mu_i) = \frac{\sum_{t=1}^{T} n_{it}}{\sum_{t=1}^{T}\exp(X_{it}'\beta)}. \tag{3.7}$$

Substituting this equation into (3.6) and introducing the notation $p_{it} = \exp(X_{it}'\beta)/\sum_{s=1}^{T}\exp(X_{is}'\beta)$ we get the log-likelihood of the Poisson model with fixed effects:

$$\ln\mathcal{L} = \sum_{t=1}^{T}\sum_{i=1}^{N} n_{it}\ln\left[p_{it}\sum_{s=1}^{T} n_{is}\right] - \ln(n_{it}!) - p_{it}\sum_{s=1}^{T} n_{is}. \tag{3.8}$$

HHG (1984) and Greene (2001) point out that this likelihood is conditional on the sum of the number of patents in the sample, $\sum_{s=1}^{T} n_{is}$. An alternative approach would be to estimate directly the μ_i parameters by (3.6). This would not require conditioning on $\sum_{s=1}^{T} n_{is}$, but for our data set this is computationally not feasible because the number of parameters would be very high.

Latent-factor Poisson model. The latent-factor patent count model is estimated by MSL method (see Gouriéroux and Monfort, 1991). First, denote the density of the latent factor l_t^* conditional on l_{t-1}^* as follows:

$$f_t^*(l_t^*|l_{t-1}^*) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(l_t^* - \mu l_{t-1}^*)^2}{2}\right). \tag{3.9}$$

Second, denote the conditional density of $n_{it}|\mathcal{F}_t$ as follows:

$$f_t(n_{it}|\mathcal{F}_t) = \frac{\exp(-\lambda_{it})\lambda_{it}^{n_{it}}}{n_{it}!} \tag{3.10}$$

Notice that the λ_{it} intensity is conditional on l_t^*. (See also the equations 2.1 and 2.2.) Therefore, if all latent variables $\{l_t^*\}$ were observable then the joint likelihood of a realization $\{n_{it}, l_t^*\}$ could be written as the product of $f_t(n_{it}|\mathcal{F}_t)$ and $f_t^*(l_t^*|l_{t-1}^*)$ as follows:

$$\prod_{t=1}^{T}\prod_{i=1}^{N} f_t(n_{it}|\mathcal{F}_t)f_t^*(l_t^*|l_{t-1}^*) = \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{\exp(-\lambda_{it})\lambda_{it}^{n_{it}}}{n_{it}!}f_t^*(l_t^*|l_{t-1}^*). \tag{3.11}$$

However, the $L^* = \{l_t^* : t = 1, \ldots, T\}$ are not observed therefore we integrate out all latent variables from the likelihood function with respect to the assumed normal distribution. Since the number of l_t^* is equal to the number of periods observed, the integrated likelihood function is the following T-dimensional integral:

$$\mathcal{L} = \int_{\mathbb{R}^T} \prod_{t=1}^{T}\prod_{i=1}^{N} \frac{\exp(-\lambda_{it})\lambda_{it}^{n_{it}}}{n_{it}!}f_t^*(l_t^*|l_{t-1}^*)dL^*. \tag{3.12}$$

The major difficulty related to the statistical inference of the model is the precise evaluation of the T-dimensional integral in $\mathcal{L}$ for given parameter values. This is performed numerically by Monte Carlo (MC) simulation method. However, it is well-known that the natural MC estimator of $\mathcal{L}$, which is based on trajectory draws from the assumed distribution of the latent variables (i.e. from the natural sampler) is downward biased in practice (see Richard and Zhang, 2007). In order to avoid this problem, we use the *efficient importance sampling* (EIS) method of Richard and Zhang (2007). The EIS technique is presented in details in Appendix A of this paper.

3.2. Consistency of the Poisson QMLE: Specification tests

We show the consistency of the latent-factor Poisson QMLE applied in two steps. First, we show that the regularity conditions of consistency are satisfied for our model. Second, we perform two alternative functional form tests of the exponential conditional mean.

First, we state the regularity conditions of Theorem 12.2 of Wooldridge (2002) in Appendix B. Conditions (a), (b) and (d) of Appendix B follow from the definitions of the random variables and the specifications of the count data models presented in Section 2. In addition, conditions (c), (e), (f), (g) and (h) stated in Appendix B follow from the properties of the log-likelihood functions presented in Section 3.

11

Second, we employ Wooldridge (1997, Section 3.5) and Wooldridge (2002, Section 19.2.5) to perform the conditional mean specification tests. For the latent-factor dynamic Poisson model, we estimate $E[n_{it}|\mathcal{F}_t] = \lambda_{it}(\theta) = \exp(\mu_0 + X'_{it}\beta + \sigma_i l^*_t)$. In this section, we test this model against two more complicated alternatives for each period $t = 1, \ldots, T$ and for each trajectory of the latent factor $r = 1, \ldots, R$. Write each alternative model as $E[n_{it}|\mathcal{F}_t] = \tilde{\lambda}_{it}(\theta, \delta)$, where δ is a $Q \times 1$ vector of additional parameters. We assume that for some known δ_0, and all θ, $\lambda_{it}(\theta) = \tilde{\lambda}_{it}(\theta, \delta_0)$. Thus, a test of null $H_0 : E(n_{it}|\mathcal{F}_t) = \lambda_{it}(\theta)$ is the same as testing $H_0 : \delta = \delta_0$. Let $\hat{\theta}$ be the Poisson QMLE of the parameters under the exponential regression null. Let $\hat{\lambda}_{it} = \lambda_{it}(\hat{\theta})$ and define $\hat{u}_{it} = n_{it} - \hat{\lambda}_{it}$.

In the following part of this section, the two alternative functional form tests are presented. Both functional form tests are robust in the sense that they do not require the *generalized linear models* (GLM) variance assumption:

$$Var[n_{it}|\mathcal{F}_t] = \sigma^2 E[n_{it}|\mathcal{F}_t] \tag{3.13}$$

where $\sigma^2 > 0$.

Functional form alternative 1: (Box-Cox like model). We test $H_0 : \delta = 0$ by nesting λ_{it} into the following Box-Cox like model:

$$\tilde{\lambda}_{it}(\theta, \delta) = \begin{cases} (1 + \delta \ln \lambda_{it})^{1/\delta}, & \delta \neq 0 \\ \lambda_{it}, & \delta = 0 \end{cases} \tag{3.14}$$

First, regress

$$\sqrt{\hat{\lambda}_{it}}(\ln \hat{\lambda}_{it})^2 \quad \text{on} \quad \sqrt{\hat{\lambda}_{it}}(1, X_{it}, l^*_t) \tag{3.15}$$

with $i = 1, \ldots, N$. Let the residuals of this regression be $\tilde{r}_{it}$ and let $\tilde{u}_{it} = \hat{u}_{it}/\sqrt{\hat{\lambda}_{it}}$. Then, compute $N - SSR$ from the regression

$$1 \quad \text{on} \quad \tilde{u}_{it}\tilde{r}_{it} \tag{3.16}$$

with $i = 1, \ldots, N$, where SSR denotes the *regression sum of squares* value of this regression. Use the $N - SSR$ statistic as an asymptotically χ^2_1 random variable.

Functional form alternative 2: (Non-linear RESET model). We test $H_0 : \delta_1 = 0, \delta_2 = 0$ by nesting λ_{it} in the following model that is analogous to Ramsey's (1969) RESET model for the linear regression model:

$$\tilde{\lambda}_{it}(\theta, \delta) = \exp[\ln \lambda_{it} + \delta_1(\ln \lambda_{it})^2 + \delta_2(\ln \lambda_{it})^3] \tag{3.17}$$

First, regress each of the last two terms on the right-hand side of the previous regression on $\sqrt{\hat{\lambda}_{it}}(1, X_{it}, l^*_t)$:

$$\sqrt{\hat{\lambda}_{it}}(\ln \hat{\lambda}_{it})^2 \quad \text{on} \quad \sqrt{\hat{\lambda}_{it}}(1, X_{it}, l^*_t) \tag{3.18}$$

12

and

$$\sqrt{\hat{\lambda}_{it}}(\ln \hat{\lambda}_{it})^3 \quad \text{on} \quad \sqrt{\hat{\lambda}_{it}}(1, X_{it}, l_t^*) \tag{3.19}$$

Call the residuals of these two regressions $\tilde{r}_{it1}$ and $\tilde{r}_{it2}$, respectively. Then, obtain $N - SSR$ from the regression

$$1 \quad \text{on} \quad \tilde{u}_{it}\tilde{r}_{it1}, \tilde{u}_{it}\tilde{r}_{it2} \tag{3.20}$$

where $\tilde{u}_{it} = \hat{u}_{it}/\sqrt{\hat{\lambda}_{it}}$. The $N - SSR$ statistic is distributed asymptotically as χ_2^2.

3.3. Asymptotic variance estimators

In this section, we refer to Wooldridge (2002, Sections 19.2.3 and 19.6.1). The asymptotic normality of the Poisson QMLE follows from Theorem 12.3 of Wooldridge (2002). The result is

$$\sqrt{N}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, A_0^{-1}B_0 A_0^{-1}) \tag{3.21}$$

where

$$A_0 = \sum_{t=1}^{T} E[\nabla_\theta \lambda_{it}(\theta_0)' \nabla_\theta \lambda_{it}(\theta_0)/\lambda_{it}(\theta_0)] \tag{3.22}$$

and

$$B_0 = E[s_i(\theta_0)s_i(\theta_0)']. \tag{3.23}$$

In this work, we impose the following Assumption 1 to derive the asymptotic variance estimators for individual parameters that can be used for inference in Poisson panel data models (see Wooldridge, 2002, Sections 19.2.3 and 19.6.1.):

Assumption 1 (Exponential mean function): The conditional mean of n_{it} is specified as the exponential of a linear function: $E[n_{it}|\mathcal{F}_t] = \lambda_{it}(\theta) = \exp(\mu_0 + X_{it}'\beta + \sigma_i l_t^*)$.

This assumption yields the fully robust variance matrix estimator that requires only the regularity conditions stated in Step 1 of the conditional mean specification tests section (see Wooldridge, 2002, Section 19). With this assumption, consistent estimators of A_0 and B_0 are

$$\hat{A} = N^{-1} \sum_{t=1}^{T} \sum_{i=1}^{N} \nabla_\theta \hat{\lambda}_{it}' \nabla_\theta \hat{\lambda}_{it}/\hat{\lambda}_{it} \tag{3.24}$$

and

$$\hat{B} = N^{-1} \sum_{i=1}^{N} s_i(\hat{\theta})s_i(\hat{\theta})' \tag{3.25}$$

and we use $\hat{A}^{-1}\hat{B}\hat{A}^{-1}/N$ for $A\hat{var}(\hat{\theta})$.

3.4. Endogeneity of R&D expenditure: A test for the latent-factor dynamic Poisson model

In the previous literature of R&D and patents, one of the most difficult issues to solve is that R&D expenses are endogenous in the patent count data models. In this section, we investigate whether our latent-factor dynamic Poisson model helps to solve the endogeneity problem reported previously. We employ the two-stage approach suggested by Wooldridge (1997, Section 6.1 and 2002, Section 19.5.1) designed for Poisson models with exponential mean function like our latent-factor specification. We employ the cross-sectional setup of Wooldridge (1997 and 2002) for the panel data framework of our latent-factor model.

Recall that the exponential mean function of the latent-factor dynamic Poisson model is $\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \sigma_i l^*_t)$. First, consider that the contemporaneous R&D expenses could be endogenous while the rest of the variables in X_{it} are exogenous. Denote the variables of contemporaneous R&D expenses by Y_{it}. Notice that in equations (2.10) and (2.14) we have five terms with contemporaneous R&D expenses, i.e. $Y_{it} = (tr_{it}, r^2_{it}, r_{it}, r_{it}c_{2it}, r_{it}c_{3it})$. Let Z_{it} denote the exogenous variables in X_{it}. Then, the mean function of the latent-factor Poisson model can be written as $\lambda_{it} = \exp(\mu_0 + Y'_{it}\gamma + Z'_{it}\delta + \sigma_i l^*_t)$.

In the first stage, we obtain the cross-sectional OLS estimates from the following equations for each $t = 1, \ldots, T$ conditional on the latent component, $\sigma_i l^*_t$:

$$
\begin{aligned}
tr_{it} &= \omega_1 + Z_{it}\Pi_1 + (\sigma_i l^*_t)\psi_1 + v_{1it} \\
r^2_{it} &= \omega_2 + Z_{it}\Pi_2 + (\sigma_i l^*_t)\psi_2 + v_{2it} \\
r_{it} &= \omega_3 + Z_{it}\Pi_3 + (\sigma_i l^*_t)\psi_3 + v_{3it} \\
r_{it}c_{2it} &= \omega_4 + Z_{it}\Pi_4 + (\sigma_i l^*_t)\psi_4 + v_{4it} \\
r_{it}c_{3it} &= \omega_5 + Z_{it}\Pi_5 + (\sigma_i l^*_t)\psi_5 + v_{5it}
\end{aligned}
\tag{3.26}
$$

with $i = 1, \ldots, N$, where v_{1it}, v_{2it}, v_{3it}, v_{4it} and v_{5it} are the error terms of each regression. Then, compute the values of the estimated OLS residuals as

$$
\begin{aligned}
\hat{v}_{1it} &= tr_{it} - \hat{\omega_1} - Z_{it}\hat{\Pi}_1 - (\sigma_i l^*_t)\hat{\psi}_1 \\
\hat{v}_{2it} &= r^2_{it} - \hat{\omega_2} - Z_{it}\hat{\Pi}_2 - (\sigma_i l^*_t)\hat{\psi}_2 \\
\hat{v}_{3it} &= r_{it} - \hat{\omega_3} - Z_{it}\hat{\Pi}_3 - (\sigma_i l^*_t)\hat{\psi}_3 \\
\hat{v}_{4it} &= r_{it}c_{2it} - \hat{\omega_4} - Z_{it}\hat{\Pi}_4 - (\sigma_i l^*_t)\hat{\psi}_4 \\
\hat{v}_{5it} &= r_{it}c_{3it} - \hat{\omega_5} - Z_{it}\hat{\Pi}_5 - (\sigma_i l^*_t)\hat{\psi}_5
\end{aligned}
\tag{3.27}
$$

with $i = 1, \ldots, N$.

In the second stage, include these residuals into the original formulation of the exponential mean function of the latent-factor model:

$$
\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \sigma_i l^*_t + \rho_1\hat{v}_{1it} + \rho_2\hat{v}_{2it} + \rho_3\hat{v}_{3it} + \rho_4\hat{v}_{4it} + \rho_5\hat{v}_{5it})
\tag{3.28}
$$

We test for exogeneity of contemporaneous R&D as follows: test the null hypothesis that Y_{it} is exogenous, $H_0 : \{\rho_i = 0 : i = 1, \ldots, 5\}$. We estimate the extended conditional mean specification by MSL-EIS technique and check if the $\hat{\rho}_i$ estimates are significant. If they are non-significant then we

conclude that there is no evidence against the hypothesis that the R&D expenditure variable is exogenous in our model.

4. Data

We use data from several sources. The U.S. utility patent data set for the January 1979 - June 2005 period was purchased from MicroPatents and for the 1963-1978 period was obtained from the NBER patent data files. The U.S. patent database includes the USPTO patent number, application date, publication date, USPTO patent number of cited patents, 3-digit U.S. technological class and assignee name (company name if the patent was assigned to a firm) for each patent. Company specific information was downloaded from the Standard & Poor's Compustat data files. For each firm we downloaded the book value for the sample midpoint year (following HHG, 1984), and R&D expenses for each year. Then, we created a match file and crossed the patent data set with the firm database via the 6-digit Compustat CUSIP codes. Firm-specific data was corrected for inflation using consumer price index data from the U.S. Department of Labor, Bureau of Labor Statistics. Our sample includes 559,729 U.S. utility patents with application dates in the 1979-2004 period (22 years) of 4,476 U.S. firms. (See Table 1 and 2.) In the data procedures, we closely followed the recommendations of Hall et al (2001). In the remaining part of this chapter, we describe some details of the database procedures and construction of additional exogenous variables.

4.1. Time of patents

The patent data set contains application date and issue (publication) date for each patent. As proposed by Hall et al (2001) we use the application date in order to determine the time of an innovation because inventors have incentive to apply for patent as soon as possible after completing the innovation.

4.2. Application-publication-lag

The U.S. patent database contains patents published until June 2005. This means that the data set excludes patents, which were submitted to the Patent Office before June 2005 but were not published before the end of our sample. It order to investigate the impact of the sample truncation, we analyze the distribution of the application-grant-lag (i.e., time elapsed between the publication date and the application date of a patent) in 1997, a year which is already not affected by the sample truncation bias. We find that 95.7 percent of patents are granted within 4.5 years after submission. (See Figure 4.) Therefore, we use a 4.5-year safety-lag and include data on patents with application dates until December 2000. (Hall et al, 2001 recommend an at least 3-year safety lag.) The total number of patent applications in the U.S. between 1979 and 2000 are presented on Figure 1. Moreover, we also present the patent propensity between 1979 and 2000 on Figure 2 in order to show that patent propensity measured by the number of patent applications divided by R&D expenditure has been increasing during the sample period. Finally, Figure 3 shows the positive relationship between mean R&D expenditure and mean number of patent applications between 1979 and 2000.

4.3. Quality of knowledge, citation-lag

We compute a measure of patent quality based on the number of citations received by each patent granted between January 1963 and June 2005. We measure the *quality of knowledge* represented by a patent by computing the number of citations the patent receives from future patents (see also Hall et al, 2001). Nevertheless, the number of citations a patent receives from future patents is subject to sample truncation bias because the sample excludes future patents, which may potentially cite the observed patents. (See Figure 5.) In order to solve the truncation problem related to citation-lag, we employ the fixed-effects approach of Hall et al (2001) that is we divide the number of citations received figures by the average number of citations in the corresponding year and technological category. The technological categories are defined as in Hall et al (2001) that is (1) chemical, (2) computers and communications, (3) drugs and medical, (4) electrical and electronics, (5) mechanical and (6) others. (See Figure 6 and 7.)

4.4. Industry classification

We use the modified standard industry classification (SIC) of Hall and Mairesse (1996) that is (1) paper and printing, (2) chemicals, (3) rubber and plastics, (4) wood and misc., (5) primary metals, (6) fabricated metals, (7) machinery, (8) electrical machinery, (9) autos, (10) aircrafts and other trans., (11) textiles and leather, (12) pharmaceuticals, (13) food, (14) computers and inst., (15) oil, (16) non-manufacturing. Moreover, we classify firms into high-tech and non-high-tech sectors following the same authors.

4.5. Observable knowledge spillovers

Observable knowledge flow occurs between two firms if a patent of a company cites a previous patent of another firm. Our data set contains all U.S. utility patent citations made by patents granted during the observation period. Using the patent citations information, for each patent, we compute the *quality of knowledge* received through (a) self-citations: the patent cites previous patents of the same firm, i.e. it builds on past knowledge produced in the same firm, (b) intra-industry spillovers: the patent cites previous patents of other firms in the same industry, i.e. knowledge spills over from the same sector and (c) inter-industry spillovers: the patent cites previous patents in different industries, i.e. knowledge spills over from other sectors.

Figure 9 shows that high-tech firms benefit more from intra-industry spillovers than non-hi-tech firms. Nevertheless, non-hi-tech firms benefit more from inter-industry spillovers. Not surprisingly, the total volume of knowledge flow is significantly higher for hi-tech firms than for non-high-tech firms. Figure 8 shows some interesting differences between industries. For example, intra-industry spillovers seem to be very important in the computer industry, while the aircrafts, car, rubber, metals, textil, food and non-manufacturing industries benefit more from knowledge produced in other industries. Self-citations seem to be more significant in the drugs, oil, paper, chemical and electrical machinery industries.

Clearly, the citation information is again subject to the citation-lag truncation bias because patents in the beginning of the observation period have less chance to cite previous and observed patents. After analyzing the citation-lag distribution (i.e. the distribution of time elapsed between citing and cited patent publication dates) we decided to use a 10-year safety-lag (from 1969) and only include patents in the sample from 1979.

5. Empirical results

5.1. Replication of the HHG models

In the first part of this section we replicate the maximum likelihood estimates of the Poisson models of HHG for the extended U.S. patent data set. The original data set of HHG included 8 years (1968-75) and 121 U.S. firms. We estimate the HHG Poisson models for a longer, 22-year time period (1979-2000) for 4,476 U.S. companies. Table 3 shows the results for the HHG Poisson model for our data set. We obtain similar results as HHG. We find a significant U-shape of the lagged R&D expenditure coefficients around 5 lags for the basic Poisson and random effects models, and, similarly to HHG, the U-shape is less significant when the fixed effects model is estimated. R&D expenditure in the current period has highly significant and positive parameter for each specification. In addition, firm size (log book value) and industry position (hi-tech dummy) have highly significant and positive coefficients as we observe in HHG's paper

5.2. Dynamic Poisson models

In Table 4, we present the estimation results of the dynamic Poisson models (basic Poisson model, random effects and fixed effects). We find significant positive coefficients for the dynamic term, $\kappa > 0$ and also for the initial value of patent counts, $\gamma_5 > 0$. Moreover, we observe significantly positive coefficient for current R&D expenses for each model, $\beta_0 > 0$. Finally, we evidence that the remaining trend, hi-tech dummy and firm size coefficients are very similar to the HHG results.

5.3. Extended dynamic Poisson models

Table 5 presents the estimation results of the extended dynamic Poisson models (basic Poisson model, random effects and fixed effects). The model is an extension of the dynamic Poisson model in the sense that it includes non-linearities in R&D expenditure, $\bar{\beta}$, 10 lags of past R&D expenditure, and several industry dummy variables, ρ_j as well. Results for the dynamic panel data coefficients are similar to the dynamic Poisson models. Moreover, contemporaneous and lagged R&D expenses are found to be significantly different from zero. The evolution of the contemporaneous values of R&D expenditure for the extended Poisson models is presented on Figure 10. Finally, industry dummy variables and the parameters measuring non-linearities in R&D expenditure are also significant.

5.4. Dynamic Poisson models with observable spillovers

Table 6 presents the estimation results for the three dynamic Poisson models with observable R&D spillovers (basic Poisson, random effects and fixed effects). The specifications are estimated for 10 lagged values of patent citations. Regarding the ω and ϕ parameters, which measure knowledge spillovers we observe interesting findings: For each model estimated in Table 6, we find significant inter-industry observable R&D spillovers for the first, second and third lag ($\omega_1, \omega_2, \omega_3 < 0$) and also for the 9-10th lags ($\omega_9, \omega_{10} < 0$). The most significant spillover process is found for the random effects specification where we find significantly negative parameters for the 9-10th lags. Moreover, for each model presented in Table 6 we evidence significant intra-industry observable R&D spillovers for the first and also for the 4-10th lags. Finally, we find that the deterministic trend, and firm size coefficients are similar to the HHG results.

5.5. Exogeneity and conditional mean specification tests

In this subsection, we perform several exogeneity and conditional mean specification tests on the latent-factor dynamic Poisson model of Table 8. Estimation results of individual parameters will be presented in the following Section 5.6.

We do the GLM robust conditional mean specification tests stated in Section 3.2 for the latent-factor Poisson model. We perform the test for each period $t = 1, \ldots, T$ and trajectory $r = 1, \ldots, R$ and we obtain SSR for the entire panel for each trajectory of the latent factor.[5] Then, we compute the average of the $N - SSR$ test statistic over R trajectories of the latent factor for each t and compute the corresponding p-value (see Table 11). The p-values of these tests indicate that we cannot reject to null hypothesis of exponential mean function against alternative functional forms of the latent-factor Poisson specification for all periods considered at the 5 percent level of significance.

We perform the exogeneity test presented in Section 3.4 to see if the R&D expenses variable was endogenous. In Tables 3, 5 and 8, we can see that the contemporaneous impact of R&D expenditure on patent intensity (the β_0 coefficient) for the latent-factor Poisson model is larger than the β_0 estimates of the HHG (1984) and the extended Poisson models. This may be a sign of biased parameters estimates of alternative models due to the endogeneity of R&D without conditioning in unobservable innovations. In order to test these exogeneity hypotheses, we also evaluated the endogeneity of R&D expenses in the HHG model with fixed effects (see Table 3) and the extended Poisson model with fixed effects (see Table 5), using the two-stage procedure described in Section 3.4 but applied to these models. We present the corresponding estimates of ρ_i in Tables 12A and 12B, respectively.[6] We evidence that some of the parameters estimates of ρ_i are significant at the 5 percent level of significance. Therefore, the null hypothesis that the R&D expenses variable is exogenous in these models is rejected. Finally, we performed the exogeneity test in the latent-factor Poisson model

---

[5]We exclude the first ten years of our sample (1979-1988) in the conditional mean specification test as our model is dynamic that includes ten lags of R&D expenses and patent citations data.

[6]Notice that for the HHG model and the extended Poisson model we only have two and three ρ_i coefficients in the endogeneity test, respectively. This is due to the fact that we have less terms that include r_{it} in these models.

of Table 8. In this test, we re-estimate the latent-factor Poisson model with five additional parameters that capture the impact of the residuals estimated by the OLS regressions (3.26). We present the estimates of these additional parameters ρ_i in Table 12C. We find that $\{\rho_i : i = 1, \ldots, 5\}$ are non-significant. Therefore, by using our dynamic latent-factor Poisson model, we fail to reject that R&D is exogenous and therefore we can interpret the estimated results with certain confidence. In the next subsection, we concentrate on the economic interpretation based on the parameter estimates presented in Table 7 and Table 8 since those are the only specifications that are able to pass our specification tests.

5.6. Latent-factor dynamic Poisson models

Our most general patent specifications are presented in Table 7 and Table 8, where we add a latent factor component into the extended dynamic Poisson model presented in Table 4. The results for the latent-factor Poisson model without R&D spillovers are presented in Table 7. Interestingly, we find the dynamic latent factor significant and highly persistent, with an AR(1) coefficient μ equal to 0.85, and with several of the corresponding industry parameters σ_j significantly different from zero. Thus, we can estimate the correlation caused by the latent factor across various industries. For example, we find evidence that the latent innovation is: a) negatively correlated in the (i) paper and printing, (ii) machinery, (iii) autos, (iv) computers and (v) oil industries since the corresponding $\sigma_j < 0$ and b) positively correlated in the (1) chemicals, (2) rubber and plastics, (3) wood and misc., (4) primary metals, (5) fabricated metals, (6) electrical machinery, (7) aircrafts, (8) textiles and leather, (9) pharmaceuticals, (10) food and (11) non-manufacturing industries since the corresponding parameters are $\sigma_j > 0$. Our more parsimonious dynamic representation maintains the basic properties of previous patent models while allowing for new features derived from the unobservable innovation components which also make R&D exogenous.

Finally, we estimated this latent-factor dynamic Poisson model controlling for observable R&D spillovers with interaction terms. The estimation results for this general model are included in Table 8. We also present the impact of lagged R&D expenditure on patent application intensity for alternative models in Figures 11-13. Moreover, the observable dynamic intra-industry and inter-industry R&D spillovers are presented in Figure 14. Notice that the inter-industry spillovers are always larger, in absolute value, than the intra-industry ones, having the largest effects at 6 and 10 year lags.

The main results for the latent-factor Poisson model with R&D spillovers are presented in Table 8. The dynamic latent innovation factor is significant and highly persistent, with an AR(1) coefficient μ equal to 0.88, and with significant effects on several industries. The elasticity of the *expected number of patents with respect to R&D* has some direct effects and some indirect effects from interaction terms. The direct contemporaneous effect of R&D on innovation output (measured by the expected number of patent applications) has an inverted U-shape. Those R&D elasticities are not independent of the quality of the innovation. For example, those indirect effects increase with the number of the current citations received and the number of citations received at the truncation lag (10 years ago). This is true for citations received from firms of the same industry (intra-industry) as

19

well as for citations from different industries (inter-industries). The elasticity with respect to lagged values of R&D is small in all cases, having the largest value equal to 0.06 at the truncation lag of 10 years. The percentage of expected patent applications *increases through time but at a decreasing rate* which depends on the level of R&D. The *size of the firm has a significant but small positive* elasticity equal to 0.01. Both the initial number of patent applications n_{i1} as well as the number of patent applications of the previous year n_{it-1} have small but significant positive effects. Finally, the firms that belong to the hi-tech sector, $D_i = 1$ have an elasticity that is 12% higher than the rest of the sectors.

In order to evaluate the main marginal contribution of the explanatory variables on the expected number of patent applications, we evaluate them at the corresponding sample averages. For that, we obtained the partial effects from the derivatives of λ_{it} with respect to each of the explanatory variables.[7] The estimated partial effects, evaluated at the sample mean of each explanatory variable, are presented in Table 9. From the first column of Table 9, we can see that the most important variable is R&D expenditure, r_{it} with partial effect of 4.33 followed by firm size, z_i with partial effect of 0.76. In addition, we also computed the partial effect of the latent factor, l_t^* on the patent application intensity. This effect depends on the specific j-industry considered, σ_j. Therefore, we present the partial effects of l_t^* for each industry $j = 1, \ldots, 16$ in Table 9. Finally, in the second and third columns of Table 9, we also present the partial effects from the two alternative specifications that do not include unobservable heterogeneity; the basic Poisson model (Table 3) and the extended Poisson model (Table 5). Notice that the total partial impacts of R&D expenditure are 3.93 and 3.00, respectively, which are lower (downward bias) than the corresponding partial effects of our the latent-factor dynamic Poisson specification of Table 8.

We present the overall average of the patent intensity estimates, $\overline{\lambda} = 4.9$ and the overall average of the number of patent applications, $\overline{n} = 5$, in Tables 10A and 10B, respectively. We also computed four other averages of $\hat{\lambda}_{it}$ and n_{it} for: (1) 1979-2000 for all firms, (2) 1979-2000 for each industry, (3) each year for all firms and (4) each year for each industry. In Tables 10A and 10B, we see that the patterns of the evolution of the estimated patent intensity and the average number of patents are very similar for the U.S. economy as a whole and for each industry during 1979-2000. However, we are able to identify three innovation intensity groups in terms of the expected number of patent applications. *Level 1*, with the expected number patent applications per year ranging from 1 to 4 which includes the following seven industries: (3) rubber and plastics, (4) wood and miscellaneous, (5) primary metals, (6) fabricated metals, (7) machinery, (11) textiles and leather and (16) non-manufacturing. *Level 2*, ranging the expected number patent applications per year from 5 to 10 with four industries: (1) paper and printing, (9) autos, (12) pharmaceuticals and (14) computers. These expected number of patent applications had their maximum values for those industries from the period 1995 to 2000. And finally *level 3*, which include the industries with the four highest

---

[7]Notice that estimating partial effects is not possible in alternative count data formulations that include unobservable heterogeneity. This is because the value of $\hat{\lambda}_{it}$ is not known when unobserved heterogeneity is included in the model, which is needed to estimate partial effects. However, in the latent-factor Poisson model it is possible to evaluate partial effects.

expected number patent applications per year: (2) chemicals, (8) electrical machinery, (10) aircrafts and other trans. and (15) oil. In those cases, the maximum values of the expected number of patent applications for those industries occurred from the period 1989 to 1992.

6. Conclusions

We have introduced new methods to control for firm-level R&D observed and unobserved spillovers in the U.S. economy over a long period of 22 years (1979-2000) merging patent data form MicroPatents and from the NBER data files. Previous R&D literature realized that knowledge spillovers are partly observable and partly latent but they were only able to control for observable spillovers. Hall et al (2001) suggests using patent citation data, which is fully available for a very long time period for all U.S. firms, to measure observable knowledge spillovers with the citations published in patent documents. By extending the latent-factor intensity approach of Bauwens and Hautsch (2006) to dynamic patent count data models, we are able to identify unobserved knowledge flows among several companies. Another important contribution of our approach is that we explicitly allow for cross dependence (co-movements) in the panel data model through the common unobserved stock of knowledge. Therefore, our dynamic count data model with unobservable innovation components is an extension of the canonical panel data count model of patent applications of HHG (1984) and of more recent contributions by Blundell et al (2002) and Wooldridge (2005).

Our modeling approach of patent counts is able to overcome the important endogeneity problem of R&D. This is important since previous results from alternative GMM procedures, and alternative maximum likelihood estimation procedures, obtained very different numerical values (unstable) for the coefficients of R&D and found also a very large instability in the other parameters depending on model specification issues, see for example Cincer (1997) and Montalvo (1997). Furthermore, Blundell et al (2002) found among alternative GMM procedures, different values and signs in the parameter of the lagged value of the number of patents. They mentioned that this instability was mainly due to the weak instruments problem, as both patents and R&D series are highly persistent.

In this paper, we overcome previous misspecification problems discussed in HHG (1984) and Blundell et al (2002) along the following lines: a) we consider a dynamic conditional parameterization of the patent application intensity of U.S. firms following Wooldridge (2005); b) we use a parametric model of observable knowledge spillovers by using intra-industry and inter-industry patent citation data; c) we introduce a dynamic unobserved stock of knowledge component, with persistent influences (AR(1) coefficient equal to 0.88) on the patent applications, allowing for a different impacts on each industry. In the preferred pre-sample mean model of Blundell et al (2002), the lagged value of the number of patents is equal to 0.84. In our model, this lagged value of the number of patents is reduced to 0.002 after controlling for unobserved innovations; d) we study the truncation R&D lag effect ('5 lags only' in previous papers) on the expected patent applications by allowing for a longer lag structure (10 lags). e) We found linear and a nonlinear terms in log(R&D) entering in the propensity to patent equation with an inverted U-shape relationship and with complementarities in spillovers based on patent citations. f) Both observable and latent knowledge spillovers are significant with significant positive observable R&D spillovers, both at the intra-industry and at the

21

inter-industry level and with differential effects by industry with the unobserved spillovers. This result is consistent with the complementarities obtained by Cassiman and Veugelers (2002) between legal protections of innovations (patents) and the strategic protections (secret or latent innovations). In a separate paper, we study the relationship between firm economic performance (market value of firms, etc.) and innovation output (patents) for leaders and followers that compete in the same market, after controlling for observable and latent R&D spillovers, but this is out of the scope of this paper; see Blazsek and Escribano (2009). Finally, g) regarding the patent intensity correlation measured by the latent process, we found evidence as in Pesaran (2006) and Holly et al (2006), of co-movements among unobserved knowledge components (unobserved innovation spillovers) in several industries.

Acknowledgments

Appendix A

The *efficient importance sampling* (EIS) method has been successfully applied for the evaluation of likelihood functions involving high-dimensional integrals for example in stochastic volatility models (Liesenfeld and Richard, 2003) and stochastic conditional intensity models (Bauwens and Hautsch, 2006) similar to our model.

Before explaining the details of the EIS technique, we note that this procedure is nested into a typical likelihood function maximization procedure. In order to maintain the stability of the maximization procedure, for every set of parameters we use the same set of i.i.d. $\mathcal{N}(0,1)$ so-called *common random numbers* (see Richard and Zhang, 2007) to estimate the value of the log-likelihood function.

Let us introduce some notation first. Denote the set of patent counts by $N_{ij} = \{n_{it} : t = 1, \ldots, j\}$ with $1 \leq j \leq T$, denote $N_t = \{N_{it} : i = 1, \ldots, N\}$, denote the set of latent variables by $L_j^* = \{l_t^* : t = 1, \ldots, j\}$ with $1 \leq j \leq T$ and let Ω denote the $5N \times T$ data matrix of exogenous variables. More precisely, the structure of Ω is the following:

$$\Omega = \left(\begin{array}{ccc} \Omega_1 & \cdots & \Omega_T \end{array} \right) = \left(\begin{array}{cccc} z & z & \cdots & z \\ r_1 & r_2 & \cdots & r_T \\ c_1 & c_2 & \cdots & c_T \end{array} \right) \tag{A1}$$

Then, let $Q_j = \{\Omega_t : t = 1, \ldots, j\}$ with with $1 \leq j \leq T$ and let

$$g(N_T, L_T^* | \Omega, \theta) = \prod_{t=1}^{T} g_t(n_t, l_t^* | N_{t-1}, L_{t-1}^*, Q_t, \theta_t), \tag{A2}$$

where

$$\ln g_t(n_t, l_t^* | N_{t-1}, L_{t-1}^*, Q_t, \theta_t) = \sum_{i=1}^{N} n_{it} \ln \lambda_{it} - \ln(n_{it}!) - \lambda_{it} + \ln f_t^*(l_t^* | l_{t-1}^*). \tag{A3}$$

The EIS methodology consists of the following elements. First, we introduce an *auxiliary sampler*, m, which is included into the likelihood function and then factorized into the product of T sequential auxiliary densities, $\{m_t : t = 1, \ldots, T\}$ as follows:

$$\mathcal{L}(N_T, \theta) = \int_{\mathbb{R}^T} \prod_{t=1}^{T} \frac{g_t(n_t, l_t^* | N_{t-1}, L_{t-1}^*, Q_t, \theta_t)}{m_t(l_t^* | L_{t-1}^*, \theta_t^*)} \times m_t(l_t^* | L_{t-1}^*, \theta_t^*) dL_T^*, \tag{A4}$$

where $\mathcal{L}(\Omega, \theta) = \mathcal{L}$ and θ^* is the vector of parameters of the auxiliary sampler defined as the *union* of all θ_t^*s, which denote the parameters of the i-th auxiliary sampler. Then, the *importance MC estimate* of $\mathcal{L}(N_T, \theta)$ for given θ^* is:

$$\hat{L}_R(N_T, \theta, \theta^*) = \frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T} \frac{g_t(n_t, l_{tr}^* | N_{t-1}, L_{t-1r}^*, Q_t, \theta_t)}{m_t(l_{tr}^* | L_{t-1r}^*, \theta_t^*)}, \tag{A5}$$

where $\{l^*_{tr} : t = 1, \ldots, T\}$ denotes the r-th trajectory of i.i.d. draws from $\{m_t : t = 1, \ldots, T\}$ and $r = 1, \ldots, R$. There are two questions arising related the above formulation: (A) *How to choose* m_t? and (B) *How to choose the* θ^*_t *parameters?*

(A) Richard and Zhang (2007) suggests to define the auxiliary sampler, m_t with its associated density kernel, k_t:

$$k_t(L^*_t, \theta^*_t) = m_t(l^*_t | L^*_{t-1}, \theta^*_t) \chi_t(L^*_{t-1}, \theta^*_t) \tag{A6}$$

where

$$\chi_t(L^*_{t-1}, \theta^*_t) = \int_{\mathbb{R}} k_t(L^*_t, \theta^*_t) dl^*_t \tag{A7}$$

denotes the t-th integrating constant associated to k_t. Richard and Zhang (2007) suggests to choose k_t as a normal density kernel. Following Bauwens and Hautsch (2006), we include the conditional density of the latent variable l^*_t into m_t. Thus, the t-th normal density kernel has the following form:

$$k_t(L^*_t, \theta^*_t) = \exp[\theta^*_{t1} l^*_t + \theta^*_{t2}(l^*_t)^2] \times \exp\left[-\frac{(l^*_t - \mu l^*_{t-1})^2}{2}\right], \tag{A8}$$

where $\theta^*_t = (\theta^*_{t1}, \theta^*_{t2})$ determine the conditional mean and variance of the t-th auxiliary sampler m_t. We can find relatively easily that the conditional mean, μ_t and conditional variance, π^2_t of the normal auxiliary sampler, m_t are given by the following expressions (Bauwens and Hautsch, 2006 have the same expressions for the conditional moments of the latent factor):

$$\mu_t = \pi^2_t(\theta^*_{t1} + \mu l^*_{t-1}) \tag{A9}$$

$$\pi^2_t = \frac{1}{1 - 2\theta^*_{t2}} \tag{A10}$$

Therefore, for given parameters of the auxiliary sampler a trajectory of $\{l^*_t\}$ can be generated from the following autoregressive process:

$$l^*_t = \pi^2_t \theta^*_{t1} + \pi^2_t \mu l^*_{t-1} + \pi_t \eta_t, \tag{A11}$$

where $\eta_t \sim \mathcal{N}(0, 1)$ i.i.d.. Moreover, from (A8) we may deduce that the t-th integrating constant is given by:

$$\chi_t(L^*_{t-1}, \theta^*_t) = \sqrt{2\pi\pi^2_t} \times \exp\left[-\frac{\mu^2(l^*_{t-1})^2}{2} + \frac{\mu^2_t}{2\pi^2_t}\right]. \tag{A12}$$

(B) The EIS methodology relies on the optimal choice of parameters of the auxiliary samplers in the sense that for given m the variance of $\hat{L}_R(N_T, \theta, \theta^*)$ is minimized, i.e.:

$$\theta^*(N_T, \theta) = \arg\min_{\theta^*} Var[\hat{L}_R(N_T, \theta, \theta^*)]. \tag{A13}$$

Loosely speaking, from equation (A5) one can see that this variance is 'small' if the auxiliary sampler m_t provides a 'good fit' to the g_t function. Expressing the auxiliary sampler by its associated density kernel and integrating constant from (A8), we may say that m_t provides 'good fit' to g_t if

$$\ln g_t(n_t, l_t^* | N_{t-1}, L_{t-1}^*, Q_t, \theta_t) + \ln \chi_t(L_{t-1}^*, \theta_t^*) \simeq \ln k_t(L_t^*, \theta_t^*). \tag{A14}$$

Richard and Zhang (2007) shows that if the auxiliary samplers are normal distributions then the MC variance minimization problem stated in equation (A13) can be reduced to a recursive sequence of T ordinary least squares (OLS) problems, each of the following form (see also Bauwens and Hautsch, 2006):

$$\ln g_t(n_t, l_t^* | N_{t-1}, L_{t-1}^*, Q_t, \theta_t) + \ln \chi_{t+1}(L_{tr}^*, \hat{\theta}_{t+1}^*) = \theta_{t0}^* + \theta_{t1}^* l_{tr}^* + \theta_{t2}^* (l_{tr}^*)^2 + u_{tr}, \tag{A15}$$

for $t = T, \ldots, 1$, $r = 1, \ldots, R$, $\chi_{T+1}(L_T^*, \hat{\theta}_{T+1}^*) = 1$ and $\hat{\theta}_{t+1}^*$ is the OLS estimate of θ_{t+1}^*. Thus, for each observation t, one has to compute the OLS estimate of the parameters of the auxiliary sampler, m_t. The regressions have a recursive structure because we use the $\hat{\theta}_{t+1}^*$ estimates in order to compute the integrating constant for the next, t-th OLS regression. (This is based on the permutation of the integrating constants in equation $(A5)$, see Richard and Zhang (2007) for more details.) Thus, the regressions are run backwards, i.e. from T to 1. The sample size of each regression is equal to the number of trajectories drawn, R. One of the advantages of the EIS algorithm is that these auxiliary regressions are typically run with relatively low sample sizes. In our case the number of trajectories of the latent variables is $R = 50$. In summary, the EIS technique consists of the following steps:

Step 1: Draw R trajectories $\{l_{tr}^*\}_{t=1}^T$ from the natural sampler, $\mathcal{N}(\mu l_{t-1r}^*, 1)$.

Step 2: for each t (from T to 1), estimate the regression in (A15).

Step 3: Given the OLS estimates of θ^* obtained in *Step 2*, draw R trajectories $\{l_{tr}^*\}_{t=1}^T$ from the auxiliary samplers, $\{m_t\}_{t=1}^T$.

Step 4: Compute the importance MC estimate of $\hat{L}_R$ according to (A5).

Richard and Zhang (2007) suggest to iterate Steps 2 and 3 several times to get a precise estimate of the value of the high-dimensional integral in (A4).

Appendix B

In order to show that the Poisson QMLE is weakly consistent given that the conditional mean function is correctly specified, we use Theorem 12.2 of Wooldridge (2002). This theorem states under which conditions the M-Estimators are consistent. This is useful to us as the Poisson QMLE is a special case of the M-Estimation methods (see Wooldridge, 2002, pp.341).

Theorem 12.2 requires the following conditions:

(a) Let w be a random vector of variables taking values in $\mathcal{W} \in \mathbb{R}^M$, where M is the number of variables.

(b) Let the Θ parameter space be a subset of $\mathbb{R}^p$, where p is the number of parameters.

(c) Let the log QML function $\ln \mathcal{L}(w, \theta) : \mathcal{W} \times \Theta \to \mathbb{R}$ be a real valued function.[8]

(d) The $\Theta \in \mathbb{R}^p$ parameter space is compact.

(e) For each $\theta \in \Theta$, $\ln \mathcal{L}(., \theta)$ is Borel measurable on $\mathcal{W}$.

(f) For each $w \in \mathcal{W}$, $\ln \mathcal{L}(., \theta)$ is continuous on Θ.

(g) $|\ln \mathcal{L}(w, \theta)| \leq b(w)$ for all $\theta \in \Theta$, where b is a non-negative function on $\mathcal{W}$ such that $E[b(w)] < \infty$.

(h) Identification assumption: the solution of the QMLE problem, θ_0 is unique: $E[\ln \mathcal{L}(w, \theta_0)] > E[\ln \mathcal{L}(w, \theta)]$ for all $\theta \in \Theta$, $\theta \neq \theta_0$.

---

[8]The $\ln \mathcal{L}(w, \theta)$ function corresponds to the $q(w, \theta)$ function in equation 12.8 of Wooldridge (2002) in the definition of the M-Estimator.

References

Aghion Ph., Boom R., Blundell R., Griffith R., P. Howit, 2003, Competition and Innovation: An Inverted U Relationship, NBER Working Paper 9269.

Aghion Ph., R. Griffith, 2005, Competition and Growth: Reconciling Theory and Evidence, Cambridge, MA: MIT Press.

Aghion Ph., P. Howitt, 1992, A Model of Growth Through Creative Destruction. Econometrica, 60, 323-351.

Bauwens, L., N. Hautsch, 2006, Stochastic Conditional Intensity Processes. Journal of Financial Econometrics, 4, 3, 450-493.

Blazsek, S., A. Escribano, 2009, Innovation, Interactions and Market Value of Firms: Differential Effects of Leaders and Followers. Mimeo, Universidad Carlos III de Madrid.

Blundell, R., R. Griffith, F. Windmeijer, 2002, Individual Effects and Dynamics in Count Data Models. Journal of Econometrics, 108, 113-131.

Cassiman, B., R. Veugelers, 2002, R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. The American Economic Review, 92, 4, 1169-1184.

Cincer 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, 265280.

Cockburn, I., Z. Griliches, 1988, The Estimation and Measurement of Spillover Effects of R&D Investment: Industry Effects and Appropriability Measures in the Stock Markets Valuation of R&D and Patents. The American Economic Review, 78,2, 419-423.

Cockburn, I., S. Wagner, 2007, Patents and the Survival of Internet-Related IPOS, paper presented at the EEA-ESEM Conference, August 2007, Budapest.

Cox, D. R., V. Isham, 1980, Point Processes, Chapman and Hall, London.

Crépon B., E. Guguet, J. Mairesse, 1998, Research, Innovation and Productivity: An Econometric Analysis at the Firm Level. NBER Working Paper Series, 6696, August.

Escribano A., A. Fosfuri, J. A. Tribó, 2009, Managing External Knowledge Flows: The Moderating Role of Absortive Capacity. Research Policy, 38, 96-105.

Fung, M. K., 2005, Are Knowledge Spillovers Driving the Convergence of Productivity among Firms?. Economica, 72, 287-305.

Fung, M. K., W. W. Chow, 2002, Measuring the Intensity of Knowledge Flow with Patent Statistics. Economics Letters, 74, 353-358.

Gouriéroux, C., 1984, Econometrie des Variables Qualitatives, Economica, Paris.

Gouriéroux, C., A. Monfort, 1991, Simulation Based Inference in Models with Heterogeneity. Annales d'Économie et de Statistique, 20/21, 69-107.

Greene, W., 2001, Estimating Econometric Models with Fixed Effects, Working Paper, Stern School of Business, New York University.

Griliches, Z., 1990, Patent Statistics as Economic Indicators: A Survey. Journal of Economic Literature, 28, 1661-1707.

Grossman G., E. Helpman, 1991, Innovation and Growth in a Global Economy, Cambridge, MA: MIT Press.

Hall, B., J. Mairesse, 1996, Estimating the Productivity of Research and Development in French and United States Manufacturing Firms, in B. van Ark, K. Wagner (eds.), International Productivity Differences, Measurement and Explanations, Amsterdam: Elsevier Science, 1996, NBER Working Paper No. 5501.

Hall, B., A. B. Jaffe, M. Trajtenberg, 2001, The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools, NBER Working Paper 8498.

Harhoff, D., F. Narin, F. M. Scherer, K. Vopel, 1999, Citation Frequency and the Value of Patented Inventions. The Review of Economics and Statistics, 81, 3, 511-515.

Hausman, J., B. Hall, Z. Griliches, 1984, Econometric Models for Count Data with an Application to the Patents-R&D Relationship. Econometrica, 52, 4, 909-938.

Holly S., M. H. Pesaran, T. Yamagata, 2006, A Spatio-Temporal Model of House Prices in the US, Working Paper, University of Cambridge.

Jaffe, A. B., 1986, Technological Opportunity and Spillovers of R&D: Evidence from Firms Patents, Profits, and Market Value. The American Economic Review, 76, 5, 984-1001.

Jaffe, A. B., 1988, Demand and Supply Influences in R&D Intensity and Productivity Growth. The Review of Economics and Statistics, 70, 3, 431-437.

Jaffe, A. B., M. S. Fogarty, B. A. Banks, 1998, Evidence from Patents and Patent Citations on the Impact of NASA and other Federal Labs on Commercial Innovation. The Journal of Industrial Economics, 46, 2, 183-205.

Jaffe, A. B., M. Trajtenberg, M. S. Fogarty, 2000, Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. The American Economic Review, 90, 2, 215-218.

Jovanovic, B., G. M. MacDonald, 1994, Competitive Diffusion. Journal of Political Economy, 102, 1, 24-52.

Lanjouw, J. O., A. Pakes, J. Putnam, 1998, How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data. The Journal of Industrial Economics, 46, 4, 405-432.

Lanjouw, J. O. and M. Schankerman., 1999, The Quality of Ideas: Measuring Innovation with Multiple Indicators, NBER Working Paper no. 7345.

Liesenfeld, R., J.-F. Richard., 2003, Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics. Journal of Empirical Finance, 10, 505-531.

Mansfield, E., M. Schwartz, S. Wagner., 1981, Imitation Costs and Patents: An empirical study. Economic Journal, 91, 907-918.

McGahan, A. M., B. S. Silverman., 2003, Innovation and Competition, Discussion Paper, University of Toronto.

Montalvo, J. G., 1997, GMM Estimation of Count-Panel-Data Models With Fixed Efects and Predetermined Instruments. Journal of Business and Economic Statistics, Vol. 15.

Nabseth, L., G. Ray, 1974, The Diffusion of new Industrial Processes: An International Study, Cambridge, Cambridge University Press.

Pakes, A., M. Schankerman, 1984, The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of return to Research Resources, in Griliches, Z. (ed.) R&D, Patents and Productivity, Chicago: University of Chicago Press (for NBER).

Pakes, A., 1985, On Patents, R&D, and the Stock Market Rate of return. Journal of Political Economy, 93, 2, 390-409.

Pesaran, M. H., 2006, Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. Econometrica, 74, 967-1012.

Peters, B., 2007, Nothing's Gonna Stop Innovators Now? An Empirical Investigation of the 'Success Breeds Success' Hypothesis, paper presented at the EEA-ESEM Conference, August 2007, Budapest.

Ramsey, J. B., 1969, Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. Journal of the Royal Statistical Society, Series B, 31, 350-371.

Richard, J. F., W. Zhang, 2007, Efficient High-Dimensional Importance Sampling. Journal of Econometrics, 141, 1385-1411.

Rogers, E. M., 1983, Diffusion of Innovations, 3rd edn. New York: Free Press.

Romer P., 1990, Endogenous Technical Change. Journal of Political Economy, 98, 2, 71-102.

Scherer, F. M., 1981, Using Linked Patent and R&D Data to Measure Inter-Industry Technology Flows, in Griliches, Z. (ed.), R&D, Patents and Productivity, Chicago, University of Chicago Press for NBER: 417-464.

Song, X., J. B. Jones, 2007, The Effects of Technological Change on Life-Cycle Human Capital Investment, paper presented at the EEA-ESEM Conference, August 2007, Budapest.

Trajtenberg, M., 2002, A Penny for your Quotes: Patent Citations and the Value of Innovations, in Patents, citations, and innovations: a window on the knowledge economy edited by Jaffe, A. B., M. Trajtenberg, MIT Press.

Wooldridge, J., 1997, Quasi-Likelihood Methods for Count Data, in I. Pesaran, M. Hashem (eds.), Handbook of Applied Econometrics, BlackWell, 1997, Chapter 8.

Wooldridge, J., 2002, Econometric Analysis of Cross Section and Panel Data, Cambridge, MIT Press.

Wooldridge, J., 2005, Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. Journal of Applied Econometrics, 20, 39-54.

Data sources

MicroPatents Co., U.S. utility patent database covering the 1979 - 2004 period and U.S. patent citations for years 2003 and 2004.

National Bureau of Economic Research patent citations data-file. CD-ROM included in Patents, citations, and innovations: a window on the knowledge economy edited by Jaffe, A. B., M. Trajtenberg, MIT Press, 2002.

Compustat (North America) Database. Standard & Poor's, 2005.

Consumer Price Index for All Urban Consumers. U.S. Department of Labor: Bureau of Labor Statistics, data downloaded from the Federal Reserve Bank of St. Louis website (http://research.stlouisfed.org).

Tables

Table 1

U.S. patent data set sample size according to industries (1979-2004)

	Industry	Number of firms	Number of patents	Patents / firm
1	Paper and printing	89	16,877	190
2	**Chemicals**	82	40,449	493
3	Rubber and plastics	82	5,045	62
4	Wood and misc.	154	10,310	67
5	Primary metals	63	2,874	46
6	Fabricated metals	98	4,869	50
7	Machinery	261	23,720	91
8	**Electrical machinery**	109	34,006	312
9	Autos	93	20,015	215
10	**Aircrafts and other trans.**	38	20,410	537
11	Textiles and leather	79	1,487	19
12	**Pharmaceuticals**	530	54,681	103
13	Food	77	3,314	43
14	**Computers and inst.**	1,232	251,446	204
15	Oil	32	27,287	853
16	Non-manufacturing	1,457	42,939	29
	Total hi-tech	1,991	400,992	201
	Total non-hi-tech	2,485	158,737	64
	Total	4,476	559,729	125

Notes: We present the number of *patent applications* for each industry. We use the modified SIC classification and hi-tech / non-hi-tech classification proposed by Hall and Mairesse (1996). Hi-tech industry names are written by bold letters. Patent data is obtained from the MicroPatents firm.

Table 2

The 30 largest companies by total number of patent applications between 1979 and 2004

Ranking	Company	Number of patents
(1)	**IBM**	37,094
(2)	**General Electric**	21,117
(3)	**Eastman Kodak**	15,575
(4)	**Motorola**	15,539
(5)	**Hewlett-Packard**	12,855
(6)	**Micron Technology**	11,535
(7)	**Texas Instruments**	11,014
(8)	**Xerox**	10,908
(9)	**Du Pont de Nemours**	10,433
(10)	**Intel**	9,777
(11)	3M	9,354
(12)	**Advanced Micro Devices**	7,945
(13)	AT&T	7,944
(14)	**General Motors**	7,801
(15)	**Lucent Technologies**	7,656
(16)	**Dow Chemical Company**	6,190
(17)	Exxon	6,019
(18)	**Procter & Gamble**	5,983
(19)	**Honeywell**	5,557
(20)	**Merck & Co.**	5,208
(21)	Mobil Oil	5,006
(22)	Shell Oil	4,741
(23)	**Sun Microsystems**	4,456
(24)	Chrysler	4,162
(25)	TRW	4,114
(26)	**United Technologies**	4,066
(27)	Ford Motor	4,043
(28)	Microsoft	3,708
(29)	Phillips Petroleum	3,630
(30)	**Eaton Corporation**	3,393

Notes: Firm names in hi-tech industries are written by bold letters. Patent data is obtained from the MicroPatents firm.

Table 3

Estimation results of the HHG (1984) Poisson model for the new patent data set

Parameter	Basic Poisson	Random effects	Fixed effects
β_0	$0.79^*(0.001)$	$0.81^*(0.009)$	$0.40^*(0.002)$
β_1	$0.04^*(0.001)$	$0.05^*(0.006)$	$0.01^*(0.002)$
β_2	$0.04^*(0.001)$	$0.06^*(0.01)$	$0.02^*(0.002)$
β_3	$0.04^*(0.001)$	$0.11^*(0.009)$	$0.01^*(0.002)$
β_4	$-0.02^*(0.001)$	$0.03^*(0.01)$	$-0.03^*(0.002)$
β_5	$0.04^*(0.001)$	$0.10^*(0.02)$	$-0.02^*(0.003)$
μ_0	$-0.96^*(0.002)$	$-1.49^*(0.007)$	
γ_1	$0.04^*(0.0003)$	$0.07^*(0.001)$	$0.08^*(0.001)$
γ_2	$-0.008^*(0.0001)$	$-0.02^*(0.001)$	$-0.005^*(0.0001)$
γ_3	$0.32^*(0.002)$	$0.002^*(0.004)$	
γ_4	$0.03^*(0.001)$	$0.13^*(0.003)$	
δ		$0.18^*(0.002)$	
log-likelihood	$-20,619.2$	$-5,298.8$	$-7,706.9$

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. The following specification is estimated for $t = 1979, \ldots, 2000$ and $i = 1, \ldots, 4476$:

$$\lambda_{it} = \exp(X_{it}^{'}\beta + \mu_0)\alpha_i$$

$$X_{it}^{'}\beta = \beta_5(L)r_{it} + \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i$$

$$\beta_5(L) = \sum_{i=0}^{5} \beta_i L^i$$

Table 4

Estimation results of the HHG (1984) Poisson model with Wooldridge (2005) extension for the new patent data set

Parameter	Basic Poisson	Random effects	Fixed effects
β_0	$0.85^*(0.002)$	$0.66^*(0.016)$	$0.53^*(0.003)$
β_1	$0.01^*(0.002)$	$0.02(0.012)$	$-0.01^*(0.002)$
β_2	$0.04^*(0.001)$	$0.03^*(0.009)$	$0.02^*(0.002)$
β_3	$0.03^*(0.002)$	$0.06^*(0.013)$	$0.01^*(0.002)$
β_4	$-0.02^*(0.001)$	$-0.003(0.018)$	$-0.03^*(0.002)$
β_5	$0.03^*(0.001)$	$-0.05^*(0.014)$	$0.01^*(0.002)$
κ	$0.001^*(0.000)$	$0.03^*(0.001)$	$0.001^*(0.000)$
μ_0	$-0.91^*(0.002)$	$-1.88^*(0.019)$	
γ_1	$0.05^*(0.0002)$	$0.08^*(0.001)$	$0.10^*(0.001)$
γ_2	$-0.01^*(0.0001)$	$-0.02^*(0.001)$	$-0.01^*(0.0001)$
γ_3	$0.18^*(0.002)$	$0.21^*(0.015)$	
γ_4	$-0.0004(0.001)$	$0.08^*(0.003)$	
γ_5	$0.001^*(0.0000)$	$0.02^*(0.001)$	
δ		$0.23^*(0.002)$	
log-likelihood	-18,425.4	-5,116.3	-7,123.4

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. The following specification is estimated for $t = 1979, \ldots, 2000$ and $i = 1, \ldots, 4476$:

$$\lambda_{it} = \exp(X_{it}'\beta + \mu_0)\alpha_i$$

$$X_{it}'\beta = \beta_5(L)r_{it} + \kappa n_{it-1} + \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1}$$

$$\beta_5(L) = \sum_{i=0}^{5} \beta_i L^i$$

Table 5

Estimation results of the extended dynamic Poisson model

Parameter	Basic Poisson	Random effects	Fixed effects
β_0	$1.06^*(0.002)$	$0.69^*(0.004)$	$0.46^*(0.004)$
β_1	$0.01^*(0.002)$	$0.01^*(0.006)$	$-0.01^*(0.002)$
β_2	$0.04^*(0.002)$	$0.04^*(0.005)$	$0.01^*(0.002)$
β_3	$0.04^*(0.002)$	$0.05^*(0.005)$	$0.01^*(0.002)$
β_4	$-0.02^*(0.002)$	$-0.01(0.008)$	$-0.03^*(0.002)$
β_5	$-0.02^*(0.002)$	$-0.01(0.011)$	$-0.02^*(0.002)$
β_6	$0.01^*(0.002)$	$-0.03^*(0.007)$	$0.02^*(0.002)$
β_7	$0.002(0.002)$	$0.01(0.011)$	$-0.06^*(0.002)$
β_8	$0.01^*(0.002)$	$-0.01(0.014)$	$0.01^*(0.002)$
β_9	$0.02^*(0.001)$	$-0.01(0.009)$	$0.01^*(0.002)$
β_{10}	$0.05^*(0.001)$	$-0.01(0.008)$	$0.03^*(0.002)$
$\bar{\beta}$	$-0.03^*(0.0002)$	$-0.01^*(0.002)$	$0.03^*(0.001)$
κ	$0.001^*(0.000)$	$0.03^*(0.000)$	$0.001^*(0.000)$
ρ_1	$0.82^*(0.003)$	$0.83^*(0.027)$	
ρ_2	$0.88^*(0.002)$	$1.04^*(0.014)$	
ρ_3	$0.22^*(0.005)$	$0.76^*(0.019)$	
ρ_4	$0.26^*(0.008)$	$0.71^*(0.004)$	
ρ_5	$-0.11^*(0.004)$	$0.89^*(0.004)$	
ρ_6	$0.44^*(0.005)$	$0.95^*(0.008)$	
ρ_7	$0.35^*(0.002)$	$0.99^*(0.004)$	
ρ_8	$0.91^*(0.002)$	$1.20^*(0.008)$	
ρ_9	$0.13^*(0.002)$	$1.26^*(0.011)$	
ρ_{10}	$0.36^*(0.002)$	$0.75^*(0.025)$	
ρ_{11}	$-0.06^*(0.006)$	$0.48^*(0.019)$	
ρ_{12}	$0.23^*(0.002)$	$0.73^*(0.004)$	
ρ_{13}	$-0.15^*(0.002)$	$0.72^*(0.003)$	
ρ_{14}	$0.68^*(0.002)$	$0.90^*(0.004)$	
ρ_{15}	$1.12^*(0.002)$	$0.91^*(0.022)$	
μ_0	$-1.39^*(0.004)$	$-2.37^*(0.005)$	
γ_1	$0.05^*(0.0003)$	$0.08^*(0.001)$	$0.11^*(0.001)$
γ_2	$-0.02^*(0.0001)$	$-0.02^*(0.001)$	$-0.02^*(0.0002)$
γ_3	$-0.09^*(0.003)$	$-0.21^*(0.003)$	
γ_4	$0.004^*(0.001)$	$0.07^*(0.003)$	
γ_5	$0.0002^*(0.0000)$	$0.02^*(0.001)$	
δ		$0.24^*(0.002)$	
log-likelihood	-17,066.9	-5,073.7	-7,039.2

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. We estimated the following model for $t = 1979, \ldots, 2000$ and $i = 1, \ldots, 4476$:

$$\lambda_{it} = \exp(X_{it}^{'}\beta + \mu_0)\alpha_i$$

$$X_{it}^{'}\beta = \beta_{10}(L)r_{it} + \kappa n_{it-1} + \bar{\beta}r_{it}^2 + \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1} + \sum_{j=1}^{15}\rho_j I_{ij}$$

$$\beta_{10}(L) = \sum_{i=0}^{10}\beta_i L^i$$

Table 6

Estimation results of the extended dynamic Poisson model with observable R&D spillovers

Parameter	Basic Poisson	Random effects	Fixed effects
β_0	$1.10^*(0.002)$	$0.63^*(0.012)$	$0.41^*(0.003)$
β_1	$-0.03^*(0.002)$	$0.002(0.007)$	$-0.03^*(0.003)$
β_2	$0.03^*(0.002)$	$0.03^*(0.008)$	$0.01^*(0.002)$
β_3	$0.03^*(0.002)$	$0.05^*(0.009)$	$0.002(0.002)$
β_4	$-0.03^*(0.002)$	$-0.01(0.013)$	$-0.04^*(0.002)$
β_5	$0.03^*(0.002)$	$-0.01(0.015)$	$-0.003(0.002)$
β_6	$0.003^*(0.002)$	$-0.03^*(0.013)$	$-0.002(0.002)$
β_7	$0.01^*(0.002)$	$-0.003(0.009)$	$-0.01^*(0.002)$
β_8	$0.02^*(0.002)$	$-0.01(0.007)$	$0.003(0.002)$
β_9	$0.02^*(0.002)$	$-0.01(0.005)$	$0.01^*(0.002)$
β_{10}	$0.05^*(0.002)$	$0.02^*(0.007)$	$-0.001(0.002)$
$\bar{\beta}$	$-0.04^*(0.0003)$	$-0.01^*(0.002)$	$0.02^*(0.001)$
κ	$0.002^*(0.000)$	$0.04^*(0.000)$	$0.001^*(0.000)$
ρ_1	$0.75^*(0.002)$	$0.84^*(0.014)$	
ρ_2	$0.37^*(0.002)$	$1.12^*(0.013)$	
ρ_3	$0.21^*(0.002)$	$0.76^*(0.004)$	
ρ_4	$0.25^*(0.002)$	$0.70^*(0.004)$	
ρ_5	$-0.13^*(0.002)$	$0.86^*(0.009)$	
ρ_6	$0.42^*(0.002)$	$0.93^*(0.018)$	
ρ_7	$0.32^*(0.002)$	$0.98^*(0.003)$	
ρ_8	$0.25^*(0.004)$	$1.26^*(0.015)$	
ρ_9	$0.36^*(0.002)$	$1.25^*(0.007)$	
ρ_{10}	$-0.09^*(0.004)$	$0.83^*(0.006)$	
ρ_{11}	$-0.08^*(0.005)$	$0.47^*(0.014)$	
ρ_{12}	$-0.24^*(0.002)$	$0.81^*(0.011)$	
ρ_{13}	$-0.17^*(0.002)$	$0.68^*(0.004)$	
ρ_{14}	$0.06^*(0.003)$	$1.00^*(0.009)$	
ρ_{15}	$1.21^*(0.002)$	$0.91^*(0.007)$	
continued			

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. We estimated the following model for $t = 1979, \ldots, 2000$ and $i = 1, \ldots, 4476$:

$$\lambda_{it} = \exp(X'_{it}\beta + \mu_0)\alpha_i$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1} + \Psi_{it} + \bar{\beta}r_{it}^2 + \sum_{j=1}^{15} \rho_j I_{ij}$$

$$\Psi_{it} = \beta_{10}(L)r_{it} + \kappa n_{it-1} + r_{it}[\omega_{10}(L)c_{2it}] + r_{it}[\phi_{10}(L)c_{3it}]$$

$$\beta_{10}(L) = \sum_{i=0}^{10} \beta_i L^i, \ \omega_{10}(L) = \sum_{i=0}^{10} \omega_i L^i, \ \phi_{10}(L) = \sum_{i=0}^{10} \phi_i L^i$$

Table 6 (continued)

Estimation results of the extended dynamic Poisson model with observable R&D spillovers

Parameter	Basic Poisson	Random effects	Fixed effects
ω_0	$0.004^*(0.0003)$	$0.001(0.006)$	$0.004^*(0.0002)$
ω_1	$-0.01^*(0.0004)$	$-0.07^*(0.004)$	$-0.01^*(0.0003)$
ω_2	$0.002^*(0.001)$	$-0.13^*(0.007)$	$-0.003^*(0.0004)$
ω_3	$-0.001^*(0.001)$	$-0.05^*(0.006)$	$-0.01^*(0.0004)$
ω_4	$-0.02^*(0.001)$	$0.09^*(0.008)$	$0.00(0.001)$
ω_5	$0.001(0.001)$	$-0.02^*(0.005)$	$0.01^*(0.001)$
ω_6	$0.02^*(0.001)$	$-0.20^*(0.008)$	$0.00(0.002)$
ω_7	$-0.02^*(0.002)$	$0.13^*(0.029)$	$-0.03^*(0.002)$
ω_8	$0.003(0.002)$	$0.74^*(0.006)$	$0.03^*(0.003)$
ω_9	$-0.01^*(0.002)$	$-1.09^*(0.031)$	$0.01^*(0.003)$
ω_{10}	$0.04^*(0.002)$	$-1.13^*(0.072)$	$-0.01^*(0.003)$
ϕ_0	$0.01^*(0.001)$	$0.06^*(0.006)$	$0.01^*(0.0004)$
ϕ_1	$-0.02^*(0.001)$	$-0.08^*(0.006)$	$-0.01^*(0.001)$
ϕ_2	$0.01^*(0.001)$	$-0.03^*(0.005)$	$0.002^*(0.001)$
ϕ_3	$0.03^*(0.001)$	$0.07^*(0.006)$	$0.01^*(0.001)$
ϕ_4	$-0.002(0.001)$	$-0.05^*(0.012)$	$-0.01^*(0.001)$
ϕ_5	$0.01^*(0.002)$	$-0.03^*(0.010)$	$0.0001(0.001)$
ϕ_6	$-0.14^*(0.002)$	$-0.02(0.010)$	$-0.03^*(0.003)$
ϕ_7	$-0.03^*(0.002)$	$0.04^*(0.012)$	$-0.01^*(0.003)$
ϕ_8	$-0.03^*(0.002)$	$-0.89^*(0.027)$	$-0.02^*(0.003)$
ϕ_9	$0.001(0.002)$	$0.26^*(0.081)$	$-0.005(0.004)$
ϕ_{10}	$0.10^*(0.002)$	$-0.14(0.095)$	$-0.01^*(0.004)$
μ_0	$-1.32^*(0.002)$	$-2.39^*(0.005)$	
γ_1	$0.05^*(0.0003)$	$0.08^*(0.001)$	$0.10^*(0.001)$
γ_2	$-0.01^*(0.0001)$	$-0.01^*(0.001)$	$-0.01^*(0.0002)$
γ_3	$0.45^*(0.003)$	$-0.30^*(0.008)$	
γ_4	$0.02^*(0.001)$	$0.07^*(0.003)$	
γ_5	$-0.0004^*(0.000)$	$0.02^*(0.001)$	
δ		$0.25^*(0.003)$	
log-likelihood	-16,009.0	-5,054.2	-6,767.4

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. We estimated the following model for $t = 1979, \ldots, 2000$ and $i = 1, \ldots, 4476$:

$$\lambda_{it} = \exp(X'_{it}\beta + \mu_0)\alpha_i$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1} + \Psi_{it} + \bar{\beta}r_{it}^2 + \sum_{j=1}^{15} \rho_j I_{ij}$$

$$\Psi_{it} = \beta_{10}(L)r_{it} + \kappa n_{it-1} + r_{it}[\omega_{10}(L)c_{2it}] + r_{it}[\phi_{10}(L)c_{3it}]$$

$$\beta_{10}(L) = \sum_{i=0}^{10} \beta_i L^i, \ \omega_{10}(L) = \sum_{i=0}^{10} \omega_i L^i, \ \phi_{10}(L) = \sum_{i=0}^{10} \phi_i L^i$$

Table 7

Estimation results of the extended dynamic Poisson model with latent R&D spillovers

Parameter	Estimation	Parameter	Estimation
β_0	$1.11^*(0.002)$	μ	$0.85^*(0.002)$
β_1	$0.01^*(0.002)$	σ_1	$-0.05^*(0.002)$
β_2	$0.04^*(0.002)$	σ_2	$0.01^*(0.002)$
β_3	$0.04^*(0.001)$	σ_3	$0.01^*(0.002)$
β_4	$-0.01^*(0.001)$	σ_4	$0.01^*(0.002)$
β_5	$-0.02^*(0.002)$	σ_5	$0.04^*(0.002)$
β_6	$0.01^*(0.002)$	σ_6	$0.01^*(0.002)$
β_7	$0.002(0.002)$	σ_7	$-0.003^*(0.002)$
β_8	$0.01^*(0.002)$	σ_8	$0.01^*(0.002)$
β_9	$0.02^*(0.001)$	σ_9	$-0.01^*(0.002)$
β_{10}	$0.06^*(0.001)$	σ_{10}	$0.02^*(0.002)$
κ	$0.001^*(0.000)$	σ_{11}	$0.04^*(0.002)$
$\bar{\beta}$	$-0.04^*(0.0003)$	σ_{12}	$0.02^*(0.002)$
μ_0	$-1.10^*(0.002)$	σ_{13}	$0.02^*(0.002)$
γ_1	$0.05^*(0.0003)$	σ_{14}	$-0.01^*(0.002)$
γ_2	$-0.01^*(0.0001)$	σ_{15}	$-0.01^*(0.002)$
γ_3	$0.17^*(0.002)$	σ_{16}	$0.01^*(0.002)$
γ_4	$-0.003^*(0.001)$		
γ_5	$0.001^*(0.000)$		
loglikelihood	-17,806.1		

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. We estimated the following specification for $t = 1979, \ldots, 2000$, $i = 1, \ldots, 4476$ and $j = 1, \ldots, 16$:

$$\lambda_{it} = \exp(X'_{it}\beta + \mu_0)\alpha_{it}$$

$$X'_{it}\beta = \beta_{10}(L)r_{it} + \kappa n_{it-1} + \bar{\beta}r_{it}^2 + \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1}$$

$$\beta_{10}(L) = \sum_{i=0}^{10} \beta_i L^i$$

$$\alpha_{it} = \exp(\sigma_j l_t^*)$$

$$l_t^* = \mu l_{t-1}^* + \eta_t$$

Table 8

Estimation results of the extended dynamic Poisson model with observable and latent R&D spillovers

Parameter	Variable	Estimation	Parameter	Variable	Estimation
β_0	r_{it}	$1.17^*(0.002)$	μ	l^*_{t-1}	$0.88^*(0.002)$
β_1	r_{it-1}	$-0.02^*(0.002)$	σ_1	l^*_t	$-0.02^*(0.002)$
β_2	r_{it-2}	$0.03^*(0.002)$	σ_2	l^*_t	$-0.01^*(0.002)$
β_3	r_{it-3}	$0.03^*(0.002)$	σ_3	l^*_t	$0.01^*(0.002)$
β_4	r_{it-4}	$-0.03^*(0.002)$	σ_4	l^*_t	$0.02^*(0.002)$
β_5	r_{it-5}	$0.03^*(0.002)$	σ_5	l^*_t	$0.06^*(0.002)$
β_6	r_{it-6}	$0.005^*(0.002)$	σ_6	l^*_t	$0.03^*(0.002)$
β_7	r_{it-7}	$0.01^*(0.002)$	σ_7	l^*_t	$0.01^*(0.002)$
β_8	r_{it-8}	$0.02^*(0.002)$	σ_8	l^*_t	$-0.001(0.002)$
β_9	r_{it-9}	$0.03^*(0.001)$	σ_9	l^*_t	$-0.02^*(0.002)$
β_{10}	r_{it-10}	$0.06^*(0.001)$	σ_{10}	l^*_t	$0.02^*(0.002)$
κ	n_{it-1}	$0.002^*(0.000)$	σ_{11}	l^*_t	$0.05^*(0.002)$
$\bar{\beta}$	r^2_{it}	$-0.05^*(0.0003)$	σ_{12}	l^*_t	$0.01^*(0.002)$
μ_0	-	$-1.04^*(0.002)$	σ_{13}	l^*_t	$0.03^*(0.002)$
γ_1	t	$0.04^*(0.0003)$	σ_{14}	l^*_t	$-0.003^*(0.001)$
γ_2	tr_{it}	$-0.01^*(0.0001)$	σ_{15}	l^*_t	$-0.01^*(0.002)$
γ_3	D_i	$0.12^*(0.002)$	σ_{16}	l^*_t	$0.01^*(0.002)$
γ_4	z_i	$0.01^*(0.001)$			
γ_5	n_{i1}	$0.0002^*(0.000)$			
ω_0	$r_{it}c_{2it}$	$0.002^*(0.0003)$	ϕ_0	$r_{it}c_{3it}$	$0.01^*(0.001)$
ω_1	$r_{it}c_{2it-1}$	$-0.01^*(0.0004)$	ϕ_1	$r_{it}c_{3it-1}$	$-0.02^*(0.001)$
ω_2	$r_{it}c_{2it-2}$	$0.003^*(0.001)$	ϕ_2	$r_{it}c_{3it-2}$	$0.01^*(0.001)$
ω_3	$r_{it}c_{2it-3}$	$-0.003^*(0.001)$	ϕ_3	$r_{it}c_{3it-3}$	$0.03^*(0.001)$
ω_4	$r_{it}c_{2it-4}$	$-0.02^*(0.001)$	ϕ_4	$r_{it}c_{3it-4}$	$-0.001^*(0.001)$
ω_5	$r_{it}c_{2it-5}$	$-0.001^*(0.001)$	ϕ_5	$r_{it}c_{3it-5}$	$0.01^*(0.002)$
ω_6	$r_{it}c_{2it-6}$	$0.03^*(0.001)$	ϕ_6	$r_{it}c_{3it-6}$	$-0.15^*(0.002)$
ω_7	$r_{it}c_{2it-7}$	$-0.02^*(0.002)$	ϕ_7	$r_{it}c_{3it-7}$	$-0.03^*(0.002)$
ω_8	$r_{it}c_{2it-8}$	$-0.02^*(0.002)$	ϕ_8	$r_{it}c_{3it-8}$	$-0.04^*(0.002)$
ω_9	$r_{it}c_{2it-9}$	$-0.003^*(0.004)$	ϕ_9	$r_{it}c_{3it-9}$	$0.01^*(0.002)$
ω_{10}	$r_{it}c_{2it-10}$	$0.07^*(0.003)$	ϕ_{10}	$r_{it}c_{3it-10}$	$0.10^*(0.002)$
log-likelihood	-16,668.0				

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. We estimated the following specification for $t = 1979, \ldots, 2000$, $i = 1, \ldots, 4476$ and $j = 1, \ldots, 16$:

$$\lambda_{it} = \exp(X'_{it}\beta + \mu_0)\alpha_{it}$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1} + \bar{\beta}r^2_{it} + \Psi_{it}$$

$$\Psi_{it} = \beta_{10}(L)r_{it} + \kappa n_{it-1} + r_{it}[\omega_{10}(L)c_{2it}] + r_{it}[\phi_{10}(L)c_{3it}]$$

$$\beta_{10}(L) = \sum_{i=0}^{10} \beta_i L^i, \ \omega_{10}(L) = \sum_{i=0}^{10} \omega_i L^i, \ \phi_{10}(L) = \sum_{i=0}^{10} \phi_i L^i$$

$$\alpha_{it} = \exp(\sigma_j l^*_t)$$

$$l^*_t = \mu l^*_{t-1} + \eta_t$$

Table 9

Response of the average firm: Partial effects for alternative Poisson models with observable and latent R&D spillovers

Variable	LF-Poisson (Table 8)	Variable	Extended Poisson (Table 5)	Basic Poisson (Table 3)
t	0.262	t	0.294	0.275
D_i	1.080	D_i	0.820	1.740
z_i	0.761	z_i	0.611	1.120
n_{i1}	0.001	n_{i1}	0.001	-
n_{it-1}	0.011	n_{it-1}	0.004	-
$r_{it}t$	-0.669	$r_{it}t$	-0.821	-0.365
r_{it}^2	-0.191	r_{it}^2	-0.100	-
r_{it}	5.190	r_{it}	4.854	3.371
$\omega_{10}(L)r_{it}c_{2it}$	0.002	$\omega_{10}(L)r_{it}c_{2it}$	-	-
$\phi_{10}(L)r_{it}c_{3it}$	0.005	$\phi_{10}(L)r_{it}c_{3it}$	-	-
$\sum r_{it}$	4.325	$\sum r_{it}$	3.933	3.006
$\hat{\sigma}_1 l_t^*$	-0.078	I_{i1}	2.332	-
$\hat{\sigma}_2 l_t^*$	-0.001	I_{i2}	1.602	-
$\hat{\sigma}_3 l_t^*$	0.028	I_{i3}	0.992	-
$\hat{\sigma}_4 l_t^*$	0.082	I_{i4}	0.978	-
$\hat{\sigma}_5 l_t^*$	0.280	I_{i5}	-0.388	-
$\hat{\sigma}_6 l_t^*$	0.156	I_{i6}	1.445	-
$\hat{\sigma}_7 l_t^*$	-0.018	I_{i7}	1.297	-
$\hat{\sigma}_8 l_t^*$	-0.006	I_{i8}	1.908	-
$\hat{\sigma}_9 l_t^*$	-0.098	I_{i9}	0.525	-
$\hat{\sigma}_{10} l_t^*$	0.100	I_{i10}	0.164	-
$\hat{\sigma}_{11} l_t^*$	0.233	I_{i11}	0.154	-
$\hat{\sigma}_{12} l_t^*$	0.057	I_{i12}	-0.072	-
$\hat{\sigma}_{13} l_t^*$	0.207	I_{i13}	-0.608	-
$\hat{\sigma}_{14} l_t^*$	-0.021	I_{i14}	1.348	-
$\hat{\sigma}_{15} l_t^*$	-0.025	I_{i15}	2.874	-
$\hat{\sigma}_{16} l_t^*$	0.055	-	-	-

Notes: We computed partial effects for the specifications presented in Tables 3 (first column), 5 (first column) and 8 by computing the partial derivatives of λ_{it} with respect to the explanatory variables. Since the partial effects depend on the explanatory variables, we evaluated the derivatives by computing the averages of λ_{it} and the corresponding variables. The following formulas show the computation of partial effects for the most general latent-factor Poisson specification $\lambda_{it} = \exp(\mu_0 + X_{it}'\beta + \sigma_j l_t^*)$ for $t = 1979, \ldots, 2000$, $i = 1, \ldots, 4476$ and $j = 1, \ldots, 16$:

$$\frac{\partial \hat{\lambda}_{it}}{\partial x_{jt}} = \hat{\lambda}_{it} \frac{\partial(\hat{\mu}_0 + X_{it}'\hat{\beta})}{\partial x_{jt}} = \hat{\beta}_j \hat{\lambda}_{it}$$

$$\frac{\partial \hat{\lambda}_{it}}{\partial l_t^*} = \hat{\sigma}_j \hat{\lambda}_{it}$$

Table 10A

The estimates of average patent intensity, $\overline{\lambda}$ over 1979-2000 by industry

year/indu.	1-16	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1979-2000	4.9	6.5	17.5	2.2	2.1	1.8	1.2	3.3	12.5	10.0	21.3	0.6	4.8	2.1	7.8	22.1	1.4
1979	1.4	1.4	6.0	1.0	0.9	1.4	0.6	1.0	2.9	4.7	12.7	0.4	1.6	1.4	1.4	9.1	0.5
1980	1.9	2.2	8.5	1.0	0.9	1.5	0.7	1.1	8.6	4.6	14.8	0.4	1.8	1.3	2.0	16.2	0.5
1981	2.2	2.2	10.7	1.2	1.0	1.7	0.7	1.3	10.3	5.2	17.2	0.4	2.1	1.4	2.3	20.0	0.6
1982	2.5	2.6	12.4	1.3	1.0	1.6	0.8	1.4	10.9	6.1	19.0	0.4	2.4	1.5	2.8	22.0	0.7
1983	2.3	2.8	11.5	1.3	1.0	1.3	0.8	1.3	9.0	5.2	16.0	0.4	2.2	1.4	2.6	21.9	0.7
1984	2.7	3.2	13.1	1.5	1.1	1.4	0.9	1.5	10.4	6.5	17.8	0.4	2.6	1.4	3.1	23.2	0.9
1985	2.9	4.0	13.8	1.5	1.2	1.4	1.0	1.7	11.6	6.6	17.5	0.4	2.8	1.8	3.2	26.0	1.0
1986	3.0	4.5	15.9	1.6	1.2	1.4	1.0	1.8	11.3	7.2	18.7	0.5	3.0	1.8	3.5	24.2	1.0
1987	3.3	5.1	17.4	1.7	1.4	1.6	1.0	2.0	11.2	7.6	19.6	0.5	3.4	2.0	3.9	22.3	1.1
1988	3.7	6.0	19.1	1.9	1.9	1.8	1.1	2.3	11.6	9.2	23.4	0.5	3.9	2.1	4.6	25.8	1.1
1989	4.7	7.0	23.7	2.3	2.1	2.1	1.4	2.8	17.4	10.9	29.6	0.6	4.7	2.4	5.9	31.1	1.3
1990	5.2	7.2	27.9	2.4	2.2	2.1	1.3	3.1	20.6	11.8	27.8	0.6	5.0	2.3	6.9	29.6	1.4
1991	5.6	7.7	26.0	2.5	2.3	2.2	1.4	3.3	17.2	12.6	27.9	0.7	5.5	2.6	8.2	33.1	1.4
1992	6.1	8.2	24.5	2.6	2.4	2.2	1.5	3.5	15.6	13.2	26.6	0.7	6.0	2.6	9.7	28.3	1.6
1993	6.1	7.8	22.5	2.7	2.7	2.1	1.5	3.8	11.9	13.9	23.7	0.7	6.5	2.7	10.2	25.5	1.7
1994	5.9	8.5	20.9	2.8	2.7	1.8	1.5	4.3	11.1	16.2	22.0	0.7	6.7	2.6	8.7	23.4	2.0
1995	7.2	13.2	21.3	3.1	3.2	2.4	1.8	4.9	13.0	14.5	22.8	0.9	7.9	3.1	12.2	22.7	2.0
1996	7.7	10.3	20.3	3.4	3.2	2.0	1.8	5.6	11.2	16.9	22.6	0.8	8.5	2.7	14.2	22.2	2.0
1997	9.0	13.0	18.9	3.5	3.5	2.3	2.0	6.3	12.3	15.5	24.4	0.9	8.5	2.7	18.0	23.2	2.3
1998	8.5	10.9	16.2	3.1	3.2	1.6	1.5	6.0	12.8	10.7	23.5	0.9	7.0	2.0	17.9	16.2	2.5
1999	7.8	7.6	17.2	3.5	3.2	1.6	1.5	5.9	13.6	10.1	20.7	0.6	6.5	1.9	16.1	9.4	2.2
2000	7.8	8.3	16.7	3.4	3.7	1.7	1.5	7.2	20.6	11.0	20.2	0.7	6.7	1.9	14.6	11.0	2.5

Notes: We computed $\overline{\lambda}$ for the latent-factor Poisson specification presented in Table 8 by computing the following means of $\hat{\lambda}_{it}$:

$$\overline{\lambda} = \frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} \hat{\lambda}_{it}$$
$$\overline{\lambda}_t = \frac{1}{N} \sum_{i=1}^{N} \hat{\lambda}_{it}$$
$$\overline{\lambda}_j = \frac{1}{TN_j} \sum_{t=1}^{T} \sum_{i=1}^{N_j} \hat{\lambda}_{it} \quad \text{for } i \in j \text{ industry}$$
$$\overline{\lambda}_{jt} = \frac{1}{N_j} \sum_{i=1}^{N_j} \hat{\lambda}_{it} \quad \text{for } i \in j \text{ industry}$$

where N_j denotes the number of firms in industry $j = 1, \ldots, 16$. Firms are classified into the following 16 industries: (1) paper and printing, (2) chemicals, (3) rubber and plastics, (4) wood and misc., (5) primary metals, (6) fabricated metals, (7) machinery, (8) electrical machinery, (9) autos, (10) aircrafts and other trans., (11) textiles and leather, (12) pharmaceuticals, (13) food, (14) computers and inst., (15) oil, (16) non-manufacturing. (Hall and Mairesse, 1996)

Table 10B

The estimates of average number of patent applications, $\bar{n}$ over 1979-2000 by industry

year/indu.	1-16	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1979-2000	5.0	7.5	21.1	2.4	2.5	1.9	2.0	3.4	12.7	8.6	21.9	0.8	4.2	1.7	7.8	37.0	1.2
1979	2.5	3.3	18.7	1.3	2.0	2.4	2.5	1.6	11.3	4.4	15.5	0.6	2.2	0.8	2.4	41.6	0.3
1980	2.6	2.9	19.4	1.4	2.1	2.3	1.7	1.8	11.5	3.9	16.3	0.8	2.2	0.8	2.5	44.4	0.3
1981	2.6	3.2	18.6	1.5	2.0	2.4	2.2	1.6	11.2	4.2	17.1	0.8	2.0	1.2	2.5	44.9	0.4
1982	2.8	3.7	19.8	1.9	1.7	2.6	1.6	1.6	12.2	3.9	17.5	0.7	2.1	0.8	2.6	53.7	0.6
1983	2.6	3.5	19.7	1.6	1.5	1.9	1.7	1.5	10.6	4.2	16.2	0.6	1.9	1.0	2.4	46.2	0.6
1984	2.7	4.2	19.1	1.6	1.6	2.3	1.7	1.6	10.8	4.1	16.6	0.6	2.2	0.9	2.6	50.3	0.6
1985	2.9	5.1	21.8	1.8	1.6	2.0	1.8	1.6	11.2	4.4	16.4	0.8	2.4	1.2	2.9	43.9	0.6
1986	3.0	5.4	20.7	1.7	1.4	1.7	1.5	1.9	10.8	4.4	18.3	1.0	2.6	1.2	3.3	37.4	0.6
1987	3.2	5.5	19.6	1.9	1.3	1.5	1.5	2.0	11.4	5.6	25.1	1.0	2.8	1.7	3.5	38.7	0.6
1988	3.6	6.4	22.6	2.0	1.7	1.5	1.5	2.1	12.3	5.9	30.3	0.7	2.9	1.6	4.2	41.4	0.7
1989	3.9	6.1	26.2	1.9	1.8	1.4	1.7	2.5	13.1	5.5	30.8	0.8	3.1	1.9	4.8	41.7	1.0
1990	4.3	7.1	27.1	2.4	1.9	1.4	2.3	2.7	13.3	6.7	24.8	0.6	3.3	1.7	5.6	46.6	1.0
1991	4.5	7.3	24.7	2.1	2.3	1.6	2.1	2.8	13.1	7.5	26.1	0.6	3.5	2.0	6.5	42.3	1.0
1992	4.9	8.1	24.0	2.3	2.5	1.4	2.0	3.2	14.4	8.8	22.3	0.6	3.8	2.0	7.2	40.0	1.2
1993	5.1	9.4	23.1	2.4	2.3	1.3	1.6	3.1	13.6	9.9	23.5	0.8	4.6	2.1	7.4	34.9	1.4
1994	5.9	11.6	21.1	2.6	2.6	2.0	1.9	3.5	13.6	11.2	20.2	0.8	6.1	2.2	9.6	29.7	1.5
1995	7.5	13.9	22.6	3.2	3.1	1.9	1.9	4.5	13.6	12.8	19.9	1.0	10.3	2.2	12.9	28.9	1.6
1996	7.6	12.0	19.1	3.0	3.2	2.2	2.5	5.3	12.2	12.9	22.5	0.7	6.3	1.9	14.9	20.2	1.9
1997	9.3	13.2	21.2	3.7	4.3	2.2	2.6	6.4	12.7	15.2	24.3	0.8	7.6	2.4	18.5	25.1	2.7
1998	9.2	10.1	19.0	3.6	3.8	1.7	3.1	6.7	13.3	17.3	24.4	0.8	6.6	2.2	19.1	23.7	2.6
1999	9.5	11.2	18.6	3.8	5.2	1.7	2.7	7.9	16.2	19.4	27.1	0.8	7.1	3.1	19.0	19.9	2.7
2000	9.1	12.6	16.9	4.8	5.2	1.7	3.1	8.4	17.3	16.0	26.4	0.7	7.1	3.1	17.9	18.4	2.4

Notes: We computed the following means of n_{it}:

$$\bar{n} = \frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} n_{it}$$
$$\bar{n}_t = \frac{1}{N} \sum_{i=1}^{N} n_{it}$$
$$\bar{n}_j = \frac{1}{TN_j} \sum_{t=1}^{T} \sum_{i=1}^{N_j} n_{it} \quad \text{for } i \in j \text{ industry}$$
$$\bar{n}_{jt} = \frac{1}{N_j} \sum_{i=1}^{N_j} n_{it} \quad \text{for } i \in j \text{ industry}$$

where N_j denotes the number of firms in industry $j = 1, \ldots, 16$. Firms are classified into the following 16 industries: (1) paper and printing, (2) chemicals, (3) rubber and plastics, (4) wood and misc., (5) primary metals, (6) fabricated metals, (7) machinery, (8) electrical machinery, (9) autos, (10) aircrafts and other trans., (11) textiles and leather, (12) pharmaceuticals, (13) food, (14) computers and inst., (15) oil, (16) non-manufacturing. (Hall and Mairesse, 1996)

Table 11

Mean N-SSR statistic for two alternative conditional mean functions

| | Box-Cox model | | RESET | |
year	$N-SSR$	p-value	$N-SSR$	p-value
1989	1.45	0.23	1.49	0.47
1990	1.56	0.21	2.98	0.23
1991	1.30	0.25	1.42	0.49
1992	1.13	0.29	2.74	0.25
1993	1.92	0.17	6.44	0.04
1994	0.04	0.84	7.30	0.03
1995	1.09	0.30	2.98	0.23
1996	0.80	0.37	4.00	0.14
1997	1.07	0.30	3.19	0.20
1998	1.37	0.24	1.37	0.50
1999	1.66	0.20	2.32	0.31
2000	5.15	0.02	5.16	0.08

Notes: We present the p-values corresponding to the mean $N-SSR$ test statistic for each year, where the mean is computed over R trajectories of the latent factor. The test statistics corresponding to the Box-Cox type model and the non-linear RESET model are asymptotically χ_1^2 and χ_2^2 random variables, respectively. We exclude the first ten years of our sample (1979-1988) in the conditional mean specification test as our model is dynamic that includes ten lags of R&D expenses and patent citations data.

Table 12

Estimates of ρ_i for various count data models to test the endogeneity of R&D expenditure

A. HHG (1984) Poisson model with fixed effects (see Table 3):

Parameter	Estimation
ρ_1	0.00 (0.001)
ρ_3	-0.15^* (0.009)

B. Extended dynamic Poisson model with fixed effects (see Table 5):

Parameter	Estimation
ρ_1	0.00 (0.000)
ρ_2	0.01 (0.012)
ρ_3	-0.14^* (0.004)

C. Latent-factor dynamic Poisson model (see Table 8):

Parameter	Estimation
ρ_1	0.00 (0.001)
ρ_2	-0.02 (0.020)
ρ_3	-0.03 (0.025)
ρ_4	-0.01 (0.010)
ρ_5	-0.03 (0.028)

Notes: Robust standard errors are reported in parentheses. * = coefficient significant at the 5 percent level. We estimated the following specifications for $t = 1979, \ldots, 2000$, $i = 1, \ldots, 4476$ and $j = 1, \ldots, 16$:

HHG (1984) Poisson model with fixed effects (see Table 3):

$$\lambda_{it} = \exp(X'_{it}\beta + \rho_1\hat{v}_{1it} + \rho_3\hat{v}_{3it})\alpha_i$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \beta_5(L)r_{it}$$

Extended dynamic Poisson model with fixed effects (see Table 5):

$$\lambda_{it} = \exp(X'_{it}\beta + \rho_1\hat{v}_{1it} + \rho_2\hat{v}_{2it} + \rho_3\hat{v}_{3it})\alpha_i$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \bar{\beta}r^2_{it} + \Psi_{it}$$

$$\Psi_{it} = \beta_{10}(L)r_{it} + \kappa n_{it-1}$$

Latent-factor dynamic Poisson model (see Table 8):

$$\lambda_{it} = \exp(\mu_0 + X'_{it}\beta + \sigma_j l^*_t + \rho_1\hat{v}_{1it} + \rho_2\hat{v}_{2it} + \rho_3\hat{v}_{3it} + \rho_4\hat{v}_{4it} + \rho_5\hat{v}_{5it})$$

$$X'_{it}\beta = \gamma_1 t + \gamma_2 tr_{it} + \gamma_3 D_i + \gamma_4 z_i + \gamma_5 n_{i1} + \bar{\beta}r^2_{it} + \Psi_{it}$$

$$\Psi_{it} = \beta_{10}(L)r_{it} + \kappa n_{it-1} + r_{it}[\omega_{10}(L)c_{2it}] + r_{it}[\phi_{10}(L)c_{3it}]$$

$$l^*_t = \mu l^*_{t-1} + \eta_t$$

Legends for the figures

Fig. 1. Total number of patent applications between 1979 and 2000. *Notes*: The figure presents the total number of patent applications for the 1979-2000 period. We employ this time period for the estimation of our econometric models. We also present the fitted quadratic regression line of $\hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$ to show the positive trend of the patent applications.
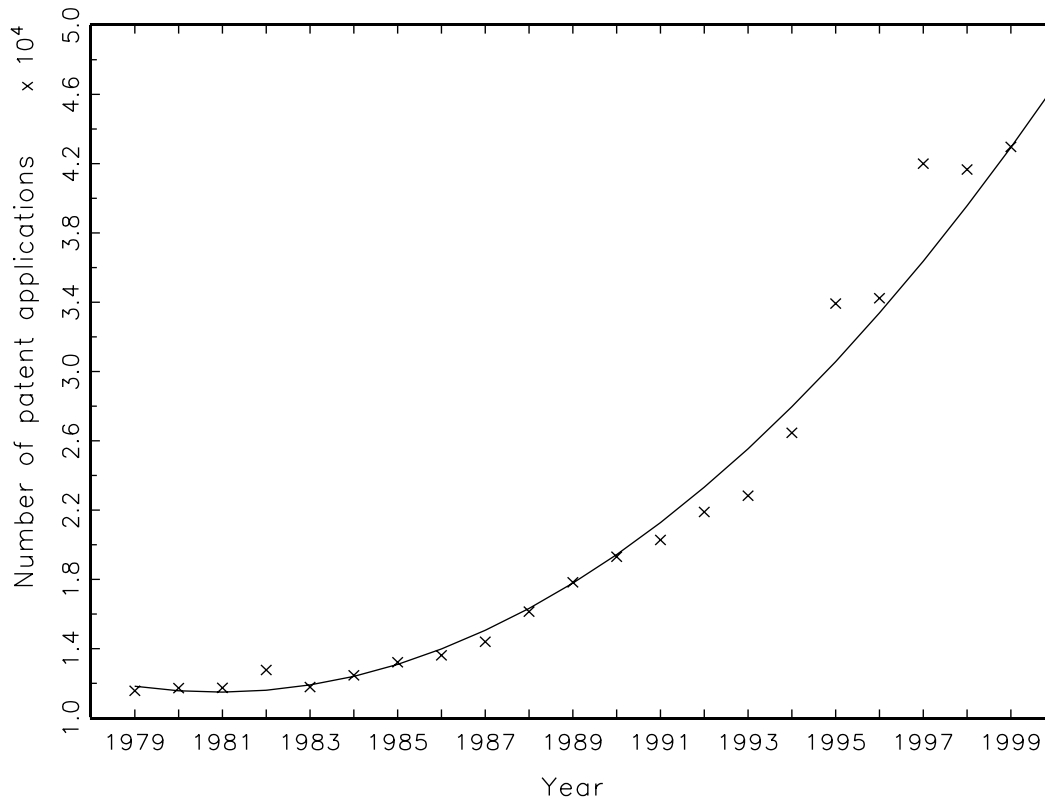
Fig. 2. Patent propensity between 1979 and 2000. *Notes*: The figure presents patent propensity, i.e. mean number of patent applications, n_{it} divided by mean R&D expenses, r_{it} for each year for the 1979-2000 period. We employ this time period for the estimation of our econometric models. We also present the fitted quadratic regression line of $\hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$ to show the positive trend of the patent propensity.
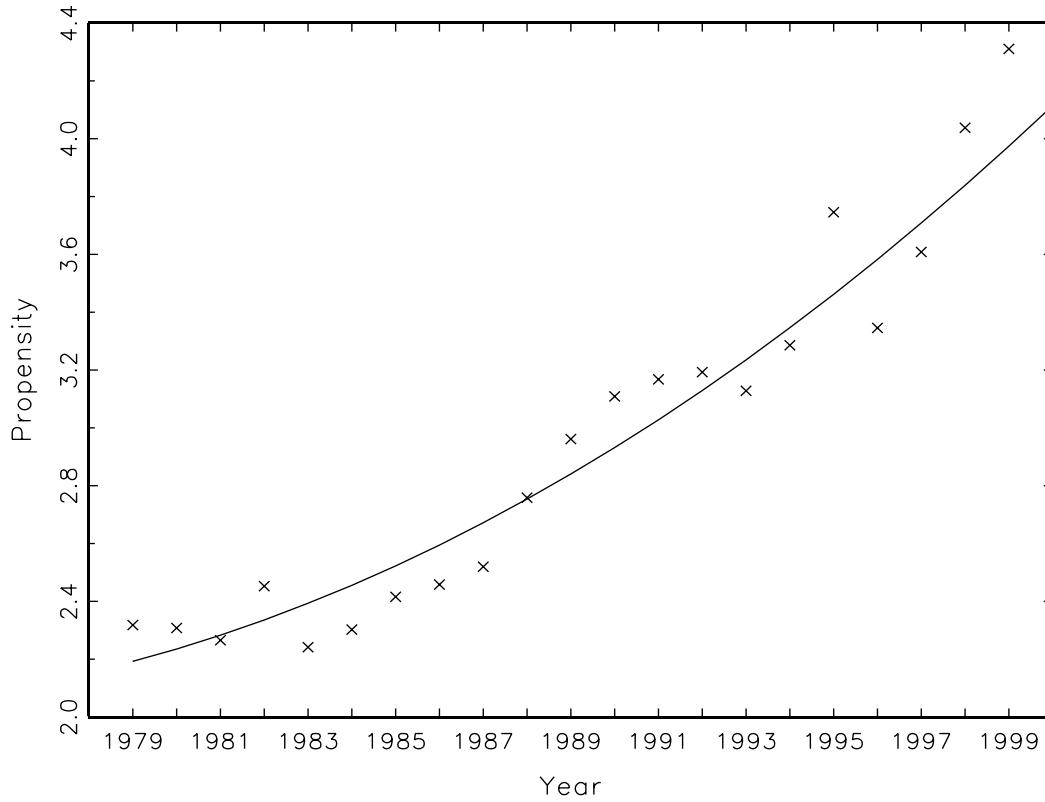
Fig. 3. Mean R&D expenditure versus mean number of patent applications between 1979 and 2000. *Notes*: We also present the fitted quadratic regression line of $\hat{\beta}_0 + \hat{\beta}_1 \bar{r}_{it} + \hat{\beta}_2 \bar{r}_{it}^2$ to show the positive relationship between the variables.

Fig. 4. Application-grant lag empirical distribution. *Notes*: The figure presents the cumulative distribution of the time duration between the patent application date and the publication date measured in years. The distribution is computed using the application-grant lag of patents, which were submitted to the U.S. Patent Office in 1997. (This year is practically not affected by sample truncation bias).

Fig. 5. Citation lag empirical distribution. *Notes*: The figure presents the cumulative distribution of the time duration between the publication dates of citing and cited patents measured in years.

Fig. 6. Average number of citations received by technological category. *Notes*: The figure presents the average number of citations that patents receive from future patents by technological category during the 1979-2005 period.

Fig. 7. Average number of citations received by technological category corrected for sample truncation bias. *Notes*: The figure presents the average number of citations received corrected for sample truncation bias using the *fixed effects* approach of Hall et al (2001). That is for each patent we divide the number of citations received from future patents by the average number of patent citations received in the corresponding technological category and year.

Fig. 8. Observable R&D spillovers by industries. *Notes*: The figure shows the average quality of knowledge that firms receive from past patents by citing them in the patent documents. Patent citations are classified as (1) self-citations, (2) intra-industry citations and (3) inter-industry citations. The quality of knowledge is measured by the number of citations received from future patents (corrected for sample truncation bias). Firms are classified into the following 16 industries: (1) paper and printing, (2) chemicals, (3) rubber and plastics, (4) wood and misc., (5) primary metals, (6) fabricated metals, (7) machinery, (8) electrical machinery, (9) autos, (10) aircrafts and other trans., (11) textiles and leather, (12) pharmaceuticals, (13) food, (14) computers and inst., (15) oil, (16) non-manufacturing. (Hall and Mairesse, 1996)

Fig. 9. Observable R&D spillovers for hi-tech and non-hi-tech firms. *Notes*: The figure shows the average quality of knowledge that firms receive from past patents by citing them in the patent documents. Patent citations are classified as (1) self-citations, (2) intra-industry citations and (3) inter-industry citations. The quality of knowledge is measured by the number of citations received from future patents (corrected for sample truncation bias). Firms are classified into two sectors: (1) non-hi-tech (others) and (2) hi-tech. (Hall and Mairesse, 1996)

Fig. 10. Evolution of R&D expenses for the extended dynamic Poisson model. *Notes*: The figure shows the evolution of R&D expenses modeled by the extended Poisson model presented in Table 5. The figure presents the evolution of $\gamma_2 tr_{it} + \beta_0 r_{it} + \bar{\beta} r_{it}^2$ for the 1979-2000 period.

Fig. 11. Impact of R&D expenditure on patent application intensity. *Notes*: The figure shows the impact of R&D expenditure on patent application intensity for the HHG (1984) basic Poisson model, the extended basic Poisson model, the latent-factor Poisson model with observable spillovers presented in Tables 3, 5 and 8, respectively.

Fig. 12. Impact of R&D expenditure on patent application intensity. *Notes*: The figure shows the impact of R&D expenditure on patent application intensity for the HHG (1984) Poisson model with random effects, the extended Poisson model with random effects, the latent-factor Poisson model with observable spillovers presented in Tables 3, 5 and 8, respectively.

Fig. 13. Impact of R&D expenditure on patent application intensity. *Notes*: The figure shows the impact of R&D expenditure on patent application intensity for the HHG (1984) Poisson model with fixed effects, the extended Poisson model with fixed effects, the latent-factor Poisson model with observable spillovers presented in Tables 3, 5 and 8, respectively.

Fig. 14. Observable intra-industry and inter-industry R&D spillovers. *Notes*: The figure shows the values of observable intra-industry and inter-industry R&D spillover estimates, ω and ϕ, respectively for the latent-factor dynamic Poisson model with observable R&D spillovers presented in Table 8.

Figures

Fig. 1. Total number of patent applications between 1979 and 2000. *Notes*: The figure presents the total number of patent applications for the 1979-2000 period. We employ this time period for the estimation of our econometric models. We also present the fitted quadratic regression line of $\hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$ to show the positive trend of the patent applications.

Fig. 2. Patent propensity between 1979 and 2000. *Notes*: The figure presents patent propensity, i.e. mean number of patent applications, n_{it} divided by mean R&D expenses, r_{it} for each year for the 1979-2000 period. We employ this time period for the estimation of our econometric models. We also present the fitted quadratic regression line of $\hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$ to show the positive trend of the patent propensity.

Fig. 3. Mean R&D expenditure versus mean number of patent applications between 1979 and 2000. *Notes*: We also present the fitted quadratic regression line of $\hat{\beta}_0 + \hat{\beta}_1 \bar{r}_{it} + \hat{\beta}_2 \bar{r}_{it}^2$ to show the positive relationship between the variables.

Fig. 4. Application-grant lag empirical distribution. *Notes*: The figure presents the cumulative distribution of the time duration between the patent application date and the publication date measured in years. The distribution is computed using the application-grant lag of patents, which were submitted to the U.S. Patent Office in 1997. (This year is practically not affected by sample truncation bias).

Fig. 5. Citation lag empirical distribution. *Notes*: The figure presents the cumulative distribution of the time duration between the publication dates of citing and cited patents measured in years.

52

Fig. 6. Average number of citations received by technological category. *Notes*: The figure presents the average number of citations that patents receive from future patents by technological category during the 1979-2005 period.

Fig. 7. Average number of citations received by technological category corrected for sample truncation bias. *Notes*: The figure presents the average number of citations received corrected for sample truncation bias using the *fixed effects* approach of Hall et al (2001). That is for each patent we divide the number of citations received from future patents by the average number of patent citations received in the corresponding technological category and year.
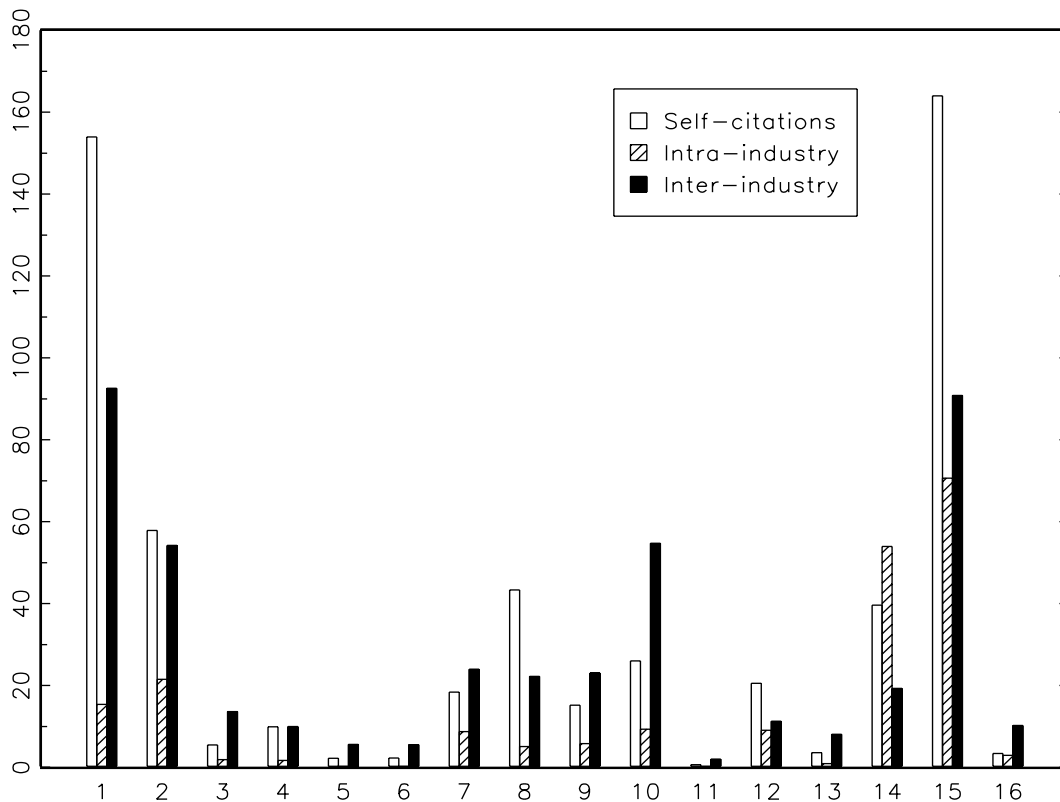
Fig. 8. Observable R&D spillovers by industries. *Notes*: The figure shows the average quality of knowledge that firms receive from past patents by citing them in the patent documents. Patent citations are classified as (1) self-citations, (2) intra-industry citations and (3) inter-industry citations. The quality of knowledge is measured by the number of citations received from future patents (corrected for sample truncation bias). Firms are classified into the following 16 industries: (1) paper and printing, (2) chemicals, (3) rubber and plastics, (4) wood and misc., (5) primary metals, (6) fabricated metals, (7) machinery, (8) electrical machinery, (9) autos, (10) aircrafts and other trans., (11) textiles and leather, (12) pharmaceuticals, (13) food, (14) computers and inst., (15) oil, (16) non-manufacturing. (Hall and Mairesse, 1996)

Fig. 9. Observable R&D spillovers for hi-tech and non-hi-tech firms. *Notes*: The figure shows the average quality of knowledge that firms receive from past patents by citing them in the patent documents. Patent citations are classified as (1) self-citations, (2) intra-industry citations and (3) inter-industry citations. The quality of knowledge is measured by the number of citations received from future patents (corrected for sample truncation bias). Firms are classified into two sectors: (1) non-hi-tech (others) and (2) hi-tech. (Hall and Mairesse, 1996)

Fig. 10. Evolution of R&D expenses for the extended dynamic Poisson model. *Notes*: The figure shows the evolution of R&D expenses modeled by the extended Poisson model presented in Table 5. The figure presents the evolution of $\gamma_2 tr_{it} + \beta_0 r_{it} + \bar{\beta} r_{it}^2$ for the 1979-2000 period.
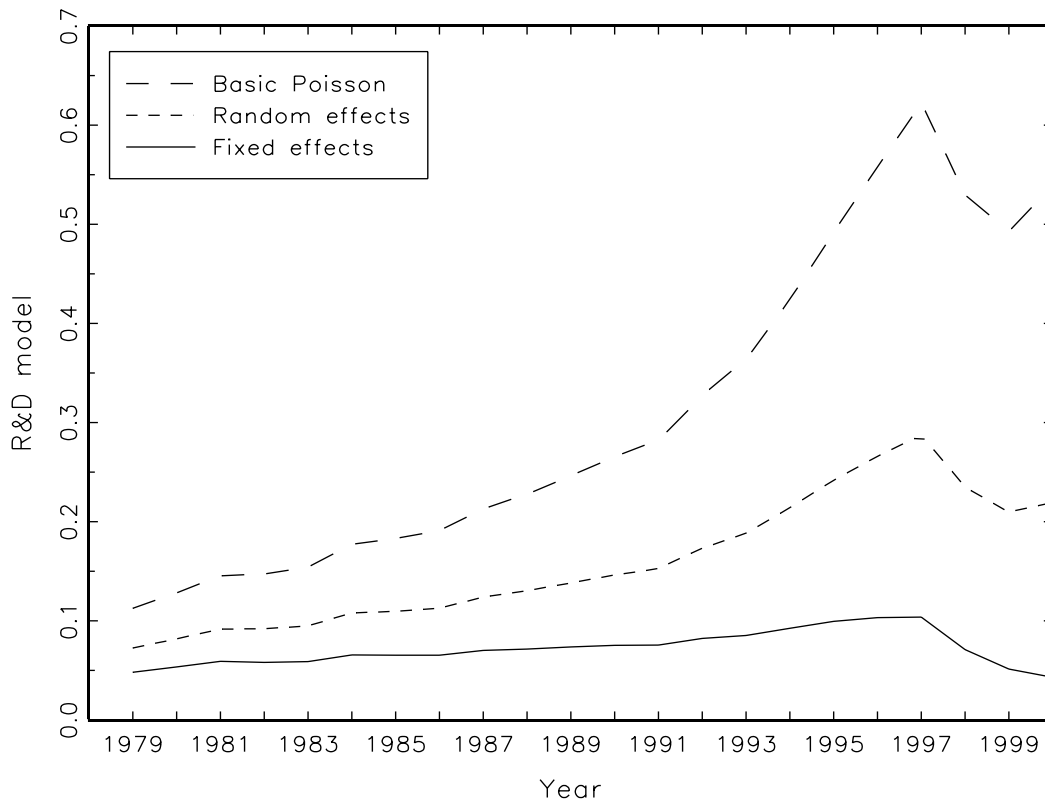
Fig. 11. Impact of R&D expenditure on patent application intensity.

Notes: The figure shows the impact of R&D expenditure on patent application intensity for the HHG (1984) basic Poisson model, the extended basic Poisson model, the latent-factor Poisson model with observable spillovers presented in Tables 3, 5 and 8, respectively.
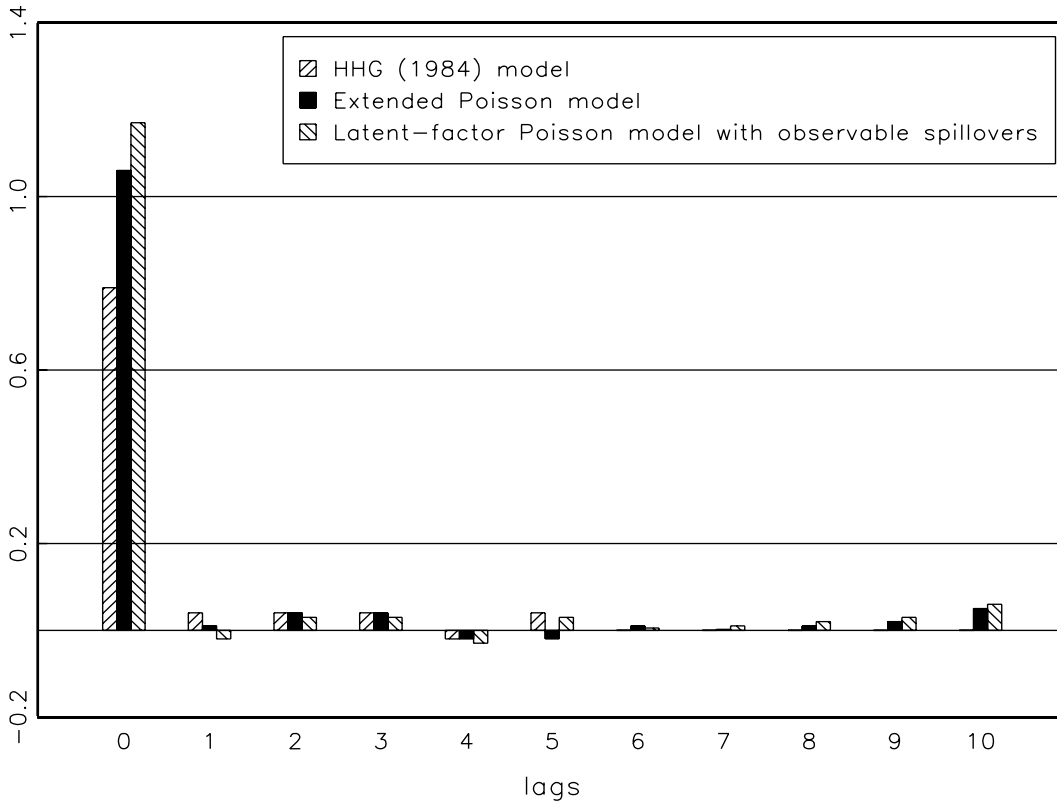
Fig. 12. Impact of R&D expenditure on patent application intensity.
Notes: The figure shows the impact of R&D expenditure on patent application intensity for the HHG (1984) Poisson model with random effects, the extended Poisson model with random effects, the latent-factor Poisson model with observable spillovers presented in Tables 3, 5 and 8, respectively.
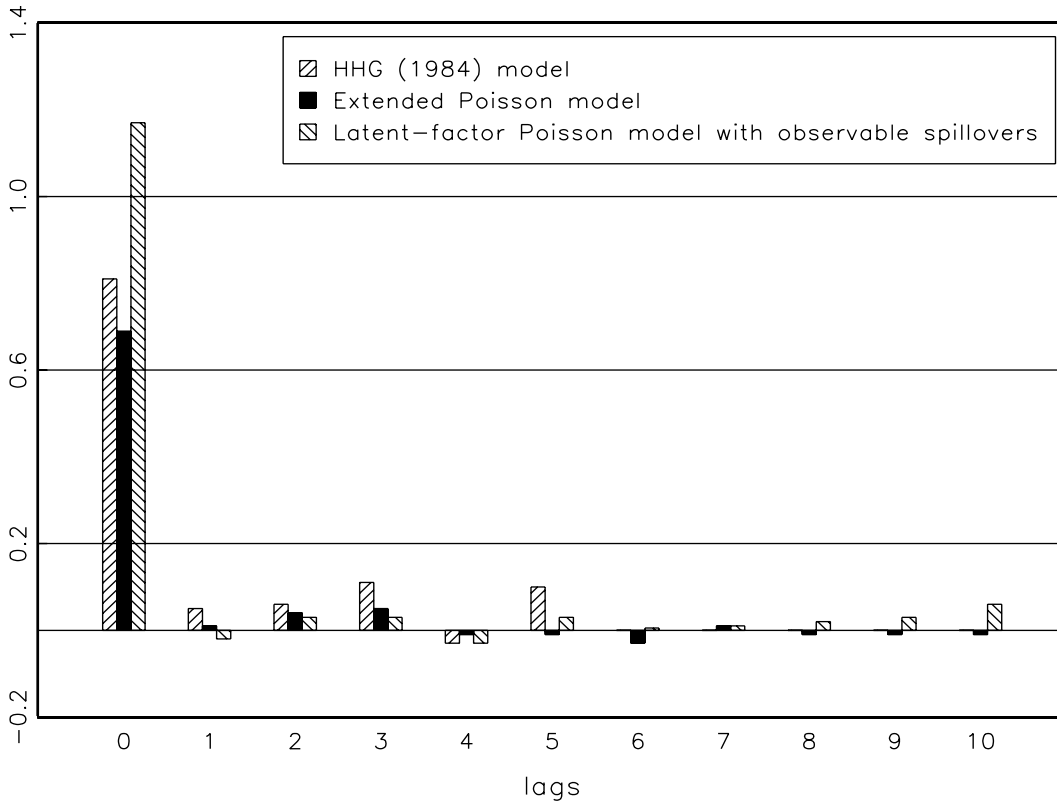
Fig. 13. Impact of R&D expenditure on patent application intensity.

Notes: The figure shows the impact of R&D expenditure on patent application intensity for the HHG (1984) Poisson model with fixed effects, the extended Poisson model with fixed effects, the latent-factor Poisson model with observable spillovers presented in Tables 3, 5 and 8, respectively.
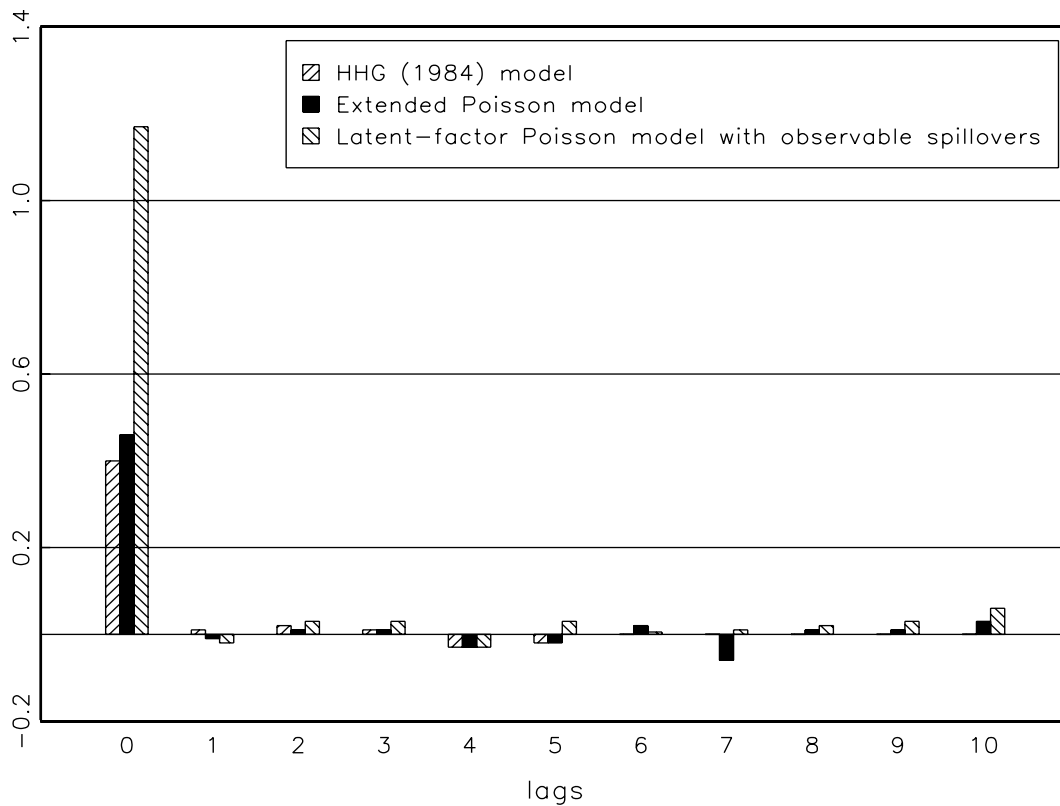
Fig. 14. Observable intra-industry and inter-industry R&D spillovers.

Notes: The figure shows the values of observable intra-industry and inter-industry R&D spillover estimates, ω and ϕ, respectively for the latent-factor Poisson model with observable R&D spillovers presented in Table 8.