UNIVERSIDAD CARLOS III DE MADRID
working papers

Working Paper 12-02
Economic Series
January, 2012

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Patents, Secret Innovations and Firm's Rate of Return: Differential Effects of the Innovation Leader

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*

January 2012

Abstract

This paper studies the dynamic interactions and the spillovers that exist among patent application intensity, secret innovation intensity and stock returns of a well-defined technological cluster of firms. We study the differential behavior when there is an Innovation Leader (IL) and the rest of the firms are Innovation Followers (IFs). The leader and the followers of the technological cluster are defined according to their patent innovation activity (stock of knowledge). We use data on stock returns and patent applications of a panel of technologically related firms of the United States (US) economy over the period 1979 to 2000. Most firms of the technological cluster are from the pharmaceutical-products industry. Interaction effects and spillovers are quantified by applying several Panel Vector Autoregressive (PVAR) market value models. Impulse Response Functions (IRFs) and dynamic interaction multipliers of the PVAR models are estimated. Secret patent innovations are estimated by using a recent Poisson-type patent count data model, which includes a set of dynamic latent variables. We show that firms' stock returns, observable patent intensities and secret patent intensities have significant dynamic interaction effects for technologically related firms. The predictive absorptive capacity of the IL is the highest and this type of absorptive capacity is positively correlated with good firm performance measures. The innovation spillover effects that exist among firms, due to the imperfect appropriability of the returns of the investment in R&D, are specially important for secret innovations and less relevant for observed innovations. The flow of spillovers between followers and the leader is not symmetric being higher from the IL to the IFs.

JEL classification: C15; C31; C32; C33; C41

Keywords : Patent count data model; Stock market value; Secret innovations; Absorptive capacity; Technological proximity; Panel Vector Autoregression (PVAR); Impulse Response Function (IRF); Efficient Importance Sampling (EIS);

*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

During the past decades, innovations protected by patents have played a key role in business strategies. This fact motivated several studies about the determinants of patents and the impact of patents on innovation, firm value and competitive advantage. Five are the usual main motives for firms to patent their inventions: protection from imitation, blocking competitors, technological image and reputation, exchange potential in cooperations and as an internal firm R&D performance indicator. Patents help sustaining competitive 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 Research and Development (R&D) should be reflected in more patent 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 are more valuable to the firm due to increasing damage awards from infringers. Shapiro (2007) notes that patents are playing an increasingly important, and shifting, role in the United States (US) economy: "There is evidence that firms in a number of industries adjusted their strategies in the 1980s and early 1990s in response to changes in the patent system. They began seeking more patents, but not necessarily because they were devoting more resources to R&D" (Shapiro, 2007). The observed increase in the R&D efficiency through the 90's could be due to increases in R&D differentiation, the increase in the number of research fields and technologies and the use of more sophisticated patent strategies due to the increases in competitive pressure through time. Jaffe (1999) mentioned that a multiplicity of explanatory factors flows in and therefore the individual contribution of each must remain unclear.

These findings motivate us to study the determinants of patents and the dynamic interdependence among observed patent intensity, secret innovation intensity and stock returns. We do that by using new Poisson-type count panel data models and Panel Vector Autoregressive (PVAR) econometric models, which control for variables that are observed by the firms but unobserved for the econometrician (latent variables).

The present paper builds on the patent-firm data set and some results of Blazsek and Escribano (2010).[1] The database applied includes 4,476 companies from several manufacturing and services

[1] Application of patent data is motivated by Griliches (1990), who states that the main advantages of patent data are the following: (a) by definition patents are closely related to inventive activity; (b) patent documents are objective since they are produced by an independent patent office and their standards change slowly over time; and (c) patent data are widely available in several countries, over long periods of time, and cover almost every field of innovation.

industries of the US economy for the period 1979 to 2000. Firms of the data set are classified into different technological clusters, where each group includes technologically related firms. We focus on a specific cluster of 111 firms. Most of these firms are in the SIC283 drugs product-market sector.[2] Nevertheless, some companies of the technological cluster belong to other product-market sectors, for example, the computers, chemicals or food industries; see Tables 1 and 2.

The objective of this paper is to learn about the dynamic interaction between patent innovation leaders and patent innovation followers of the technological cluster, by allowing for the presence of secret patent innovations. Secret patent innovations are identified by using a recent Poisson-type patent count data model of Blazsek and Escribano (2010), which includes a set of dynamic latent variables. This patent count data model is estimated by the Maximum Simulated Likelihood (MSL) method, employing the Efficient Importance Sampling (EIS) variance reduction technique of Richard and Zhang (2007).

In the technological cluster analyzed, the permanent Innovation Leader (IL) and the permanent Innovation Followers (IFs) in patent innovation activity are identified. Interaction effects are quantified by applying PVAR specifications to several market value models, estimated by the Quasi Maximum Likelihood (QML) method. Moreover, the Impulse Response Functions (IRFs) and the dynamic multipliers of the PVAR model are also estimated in order to have more precise information on dynamic interaction effects.

The remaining part of this paper is organized as follows. The relevant literature is reviewed in Section 2. The data set and the definitions of patent innovation leaders and followers are given in Section 3. Description of the econometric models and summary of the empirical results are provided in Section 4. Section 5 concludes. Statistical inference procedures of the econometric models are presented in Appendices 1-4.

2. Innovation, competition and market value of firms

In this section, the relevant literature is summarized, relating innovation and R&D with competition and with the market value of firms. The positive differential market value effects between the firms that are innovation leaders and those that are followers are reviewed.

[2] Standard Industry Classification (SIC)

2.1. Innovation and market value

Innovation activity exists since it has a positive impact on the future cash flow and the current value of a company, which motivates owners to promote innovative activity within their firm. As profits on R&D are usually realized during several years in the future, current accounting based net profit is a rather noisy measure of R&D benefits. Therefore, in the economics literature, several papers have investigated the impact of R&D on stock market price, which avoids the problem of timing differential of R&D expenses and the associated future cash flow to equity, since current stock prices are determined by a forward-looking perspective of investors. This approach is also useful for the consideration of different measures of R&D activity that may capture econometrically observable and latent innovations, for example patents and trade secrets, respectively, as investors may be aware of R&D related information hidden from the researcher.

Griliches (1981) constructs a stock of knowledge variable from lagged R&D expenses and the number of patents. He finds a significant positive relationship between market value, R&D expenditure and number of patents for a panel of large US firms for the period 1968 to 1974.

Pakes (1985) focuses on the dynamic relationships among the firm's number of successful patent applications, R&D expenditures and stock market value. Pakes concludes that the events that lead the market to reevaluate the firm are significantly correlated with unpredictable changes in both the R&D and the patents of the firm.

Hall (1993) uses data on US manufacturing firms for the period 1979 to 1991, finding that the stock market valuation of R&D broke down in the mid-1980s.[3]

Lev and Sougiannis (1996) estimate the inter-temporal relation between R&D capital and stock returns of public firms in the US during the period 1975 to 1991. These authors show that R&D capital, defined as the weighted sum of past R&D expenses, is associated with subsequent stock returns.[4]

Blundell et al (1999) employ US firm-level panel data for the period 1972 to 1982. They examine the relationship between surprise innovations and firm performance by using a dynamic panel count data specification. The stock of innovation is constructed from a count of 'technologically significant and commercially important innovations' commercialized by the firm. They find a positive impact of innovation on market value.

[3] Hall et al (2006) have similar findings.
[4] See also Lev et al (2005).

Chan et al (2001) investigate the relationship between R&D capital and stock returns of US firms for the period 1975 to 1995. They define R&D capital, based on the estimates of Lev and Sougiannis (1996), as a weighted sum of contemporaneous and past R&D expenses. Chan et al (2001) show a positive relationship between the R&D to market value variable and abnormal future stock returns. Furthermore, they evidence a delayed association of R&D activity and future excess stock returns, which could be due to a delayed reaction of the stock market or an inadequate adjustment for risk (see Chambers et al, 2002).

Hall et al (2005) investigate the relationship between knowledge stock and market value in the US during the period 1963 to 1995. The knowledge stock variable is constructed from R&D expenses, number of patents and patent citations, capturing the different importance of each patent. They build on Griliches (1981) to estimate Tobin's q equations. Their results show that, in addition to patent counts, patent citations contain important information about stock market value.

Finally, some papers investigated the interaction of firms' market value and R&D for European firm-level data. Hall and Oriani (2006) investigate this interaction for German, French and Italian data. Hall et al (2007) extend the analysis of Hall and Oriani (2006) for 33 European countries. Both papers find country-dependent results about the stock market valuation of R&D activity.

2.2. Innovation and competition

Technological improvements give innovator companies a competitive advantage. Nevertheless, the non-rival nature of knowledge may create a business-stealing effect among competitors as the innovator's effort decreases the cost of competitor firms' subsequent innovations. There is a large literature of economics and strategic management, which differentiates among firms by their R&D and patenting activity to study the implications of a firm's research intensity on its competitors' market value and innovations.[5] The relationship between firms' stock market value and R&D is investigated by recognizing that R&D activities are different among companies. Firms strategically decide to be R&D leaders or followers. Companies that introduce innovative products are R&D leaders, while other firms, who mimic the products of the innovation leaders, are followers. Results in the existing literature suggest that R&D leaders have sustained future profitability.

Caves and Porter (1977) introduce a framework that explains intra-industry profit differentials

[5] See for example the works of Porter (1979, 1980, 1985).

based on pre-commitment to specialized resources such as R&D. Gilbert and Newbery (1982) analyze a model where incremental innovations are awarded to the firm that spends the most on R&D, and they show that the incumbent firm continues to earn monopoly rents. On the other hand, Reinganum (1985) shows that incumbent firms have less incentives to invest in innovation: even though incumbents make more profits in the short-term, entrants are more profitable in the long-term and they overtake incumbents in the long run.

Jaffe (1986) finds evidence of knowledge spillovers[6] by using various indicators of R&D activity. He evidences that firms whose research is in a sector where there is high research intensity, obtain more patents per dollar of R&D, higher accounting profits to R&D and higher market value to R&D than firms in a sector with low R&D intensity.

Caves and Ghemawat (1992) investigate the factors that sustain profit differences across firms within an industry and find that differentiation-related strategies, which include R&D, are more important than cost-related strategies. They find that differentiation related strategies are indicative of research leadership in the product market by introducing new products, services, brands, etc., while cost-related strategies include higher capacity and cost structure advantages.

Jovanovic and MacDonald (1994) point out that innovation and imitation tend to be substitutes. Though, the benefits generated by other firms' R&D efforts depend on the technological differences among firms and the absorptive capacity of the imitator firm. Naturally, these factors create time lags in the adoption of technologies.[7]

Aghion et al (2005) develop a model where competition discourages laggard firms from innovating but encourages neck-and-neck firms to innovate. Due to the effect of competition on the equilibrium industry structure, their model generates an inverted-U shaped relationship between innovation and competition. The paper provides empirical evidence that (1) the average technological distance between innovation leaders and innovation followers increases with competition, and (2) the inverted-U is steeper when industries are more neck-and-neck.

Lev et al (2006) use US data over the period 1975 to 2002. They differentiate between R&D leaders and followers, and compare the stock market valuation of R&D leaders and followers. They show that R&D leaders earn significant future excess returns, while R&D followers only earn average returns.

[6] A spillover of knowledge occurs when a new innovation created by a company is adopted by another firm.

[7] Nabseth and Ray (1974), Mansfield et al (1981), Rogers (1983), and Pakes and Schankerman (1984) report that knowledge spills over gradually, in a dynamic fashion, to other firms.

Finally, Ciftci et al (2011) find that R&D leaders obtain substantial risk-adjusted returns during the first four-five future years. However, these excess returns converge to those of R&D followers afterwards.

3. Data and definitions of innovation leaders and followers

The data applied in this paper have been derived from the general US patent-firm specific data set of Blazsek and Escribano (2010). In Section 3.1, some details of this general data set are discussed. In Section 3.2, the construction of the clusters of technologically related firms is summarized. In Section 3.3, the classification of firms to patent innovation leader and follower groups is presented.

3.1. General US patent-firm data set

The general data set includes 4,476 US firms from several manufacturing and services industries of the US economy for the period 1979 to 2000. These firms published more than 500,000 patents during this period. Blazsek and Escribano (2010) created a data set for these firms based on the recommendations of Hall et al (2001).[8]

The panel data have been collected from several sources. Patent data are obtained from the National Bureau of Economic Research (NBER) and Micro Patents Co. The database includes the United States Patent and Trademark Office (USPTO) patent number, application date, publication date, USPTO patent number of cited patents, 3-digit US technological class and assignee name (company name if the patent was assigned to a firm) for each patent. Furthermore, annual stock returns, collected from the Center for Research on Stock Prices (CRSP), have been downloaded from the Wharton Research Data Service (WRDS). Additional company specific information have been obtained from the Standard & Poor's (S&P) Compustat data files. In particular, the data set includes book value of equity, stock market value, SIC code and R&D expenditure for all firms. Firm-specific accounting data are corrected for inflation by using consumer price index data collected from the US Department of Labor, Bureau of Labor Statistics. Finally, annual data on the S&P500 stock index return obtained from the Compustat data files are also included in the panel data set.

[8] See the details of the data base procedures and a more detailed description of the general US database in Blazsek and Escribano (2010).

3.2. Technological clustering

We use a technology related grouping of all companies of the general US data set. Technology based grouping of firms is preferred to product-market based (for example SIC based) grouping, as under a technology based grouping, the flows of knowledge are expected to be more important. Using an incorrect grouping dilutes the measurement of knowledge spillovers and makes it difficult to identify competitors' effects on firm's innovation activities.

Technological clusters of firms can be formed based on the idea of firms' technological proximity. In the past literature, researchers employed different frameworks to capture technological proximity, which included patent based,[9] productivity based[10] and alternative measures.[11] See the reviews of Mohnen (1996) and Cincera (2005).

The patent based technological proximity measures may be either technological category based or patent citations based measures. In the technological category based approach, the number of patents published by a firm in each technological category is counted, and a vector is formed for each company over the technological category space. Technological proximity of two firms is computed by evaluating the distance of their two vectors.[12] In the patent citations based framework, the technological proximity measures capture the overlap in patent citations between firms. These measures ask how many of the patents that one firm cites are also cited by another firm.[13]

In this paper, we use a patent technological category based proximity measure to assign firms into technologically similar groups. Some important questions related to the clustering procedure are:

(1) **How to define technological categories?** We use 36 technological sub-categories, as suggested by Hall et al (2001). These authors create 36 two digit technological sub-categories from the patent technological classification of USPTO, which contains about 400 technological classes.

(2) **How to choose the clustering algorithm?** We apply Ward's linkage clustering (Ward, 1963)

[9] See Griliches (1979), Scherer (1982), Jaffe (1986, 1988, 1989), Cohen and Levinthal (1989), Stuart and Podolny (1996), Fagerberg et al (1996), Mowery et al (1996), Breschi et al (1998, 2003), Ahuja (2000), Harhoff (2000), Autant-Bernard (2001), Fung (2003), Song et al (2003), Rosenkopf and Almeida (2003), Cantner and Meder (2007), Messeni (2008), Fischer et al (2009) and Lychagin et al (2010).

[10] See Kumar and Russell (2002), Frantzen (2004), Aghion et al (2005), Vandenbussche et el. (2006), Griffith et al (2009) and Islam (2011).

[11] See Adams (1990) and Kaiser (2002).

[12] See Breschi et al (1998) and Benner and Waldfogel (2008) for reviews. See also the papers of Griliches (1979), Scherer (1982), Jaffe (1986), Ahuja (2000), Fung (2003) and Song et al (2003).

[13] See Stuart and Podolny (1996), Mowery et al (1996) and Fischer et al (2009).

to perform technological clustering since several papers comparing different clustering techniques, conclude that Ward's method tends to identify better clusters than other methods.[14]

(3) **How to measure the distance between two patent count vectors over the space of technological categories?**[15] We use the angle distance measure to form technological clusters of firms due to the fact that the angle measure of proximity is purely directional, therefore, it is not directly affected by the length of the technological category count vectors.[16]

(4) **How many groups of technologically similar firms to create?** We group all firms of the US data set into 16 clusters.

The technological clustering framework presented creates a technology related grouping of 16 clusters of the 4,476 US companies of the data set of Blazsek and Escribano (2010). We focus on a specific cluster of $N = 111$ companies.

In order to see the product-market industries of the companies of the technological cluster selected, the SIC and Hall-Mairesse (HM, 1996) based sector classifications of these firms are presented in Tables 1 and 2, respectively. These tables show that the cluster selected mainly includes drugs firms from the SIC283 sector (Table 1) and the Pharmaceuticals product-market related sector (Table 2). Nevertheless, these tables exhibit that the technological cluster includes companies from other product-market sectors as well. For example, the technological cluster includes firms from the Grain mill products (SIC2040), Beverages (SIC2080), Paints (SIC2851), Plastics products (SIC3089) and Electromedical & electrotherapeutic apparatus (SIC3845) industries.

[APPROXIMATE LOCATION OF TABLE 1, TABLE 2]

Finally, Figure 1 shows the evolution of patent application counts and the estimated total patent application intensity, over the period 1979 to 2000, for all firms in the technological cluster.[17] This figure shows an exponential growth of patent applications counts over the sample period. The level of

---

[14] See Kuiper and Fisher (1975) and Jain et al (1986).

[15] In the past literature, several distance measures have been used in the R&D literature. Three popular measure are: (1) Euclidean or L2 distance (see Ahuja, 2000; Rosenkopf and Almeida, 2003); (2) angle between two vectors (see Jaffe, 1986, 1989; Autant-Bernard, 2001; Messeni, 2008; Lychagin et al, 2010) and (3) correlation coefficient.

[16] "The length of the vector depends on the degree of focus or concentration of the firm's research interests. (The length is actually the square root of the Herfindahl index of concentration of the category shares.) Other proximity measures, notably the Euclidean distance between the vector endpoints, are very sensitive to the length." (Jaffe, 1986)

[17] The patent intensity model applied to estimate total patent application intensity is summarized in Section 4.

patent applications per year was about 600 patents in 1979, which increased to about 1,300 patents in 2000. Moreover, we observe a local peak of patent counts in 1995 with a level of about 2,000 patents. This evolution of the number of patents applications is followed remarkably well by our new, flexible and dynamic Poison-type, count panel data model of patent applications intensity (see Section 4 for more details on the model specification and estimation).

[APPROXIMATE LOCATION OF FIGURE 1]

3.3. Definition of the IL and the IFs in the technological cluster

In the R&D literature, alternative definitions of R&D leadership have been proposed. For example, Lev et al (2006) state: "Firms with R&D intensity measure greater than (lesser than or equal to) that of the industry are classified as leaders (followers). These authors measure R&D intensity by using two proxies: the R&D expenditure to sales ratio and the R&D expenditure to market value ratio." Furthermore, Chambers et al (2002) and Ciftci et al (2011) use the R&D capital to sales ratio variable to indicate R&D leadership, where R&D capital is defined as "*R&D assets* is the asset that would have been reported if R&D expenditures were capitalized and amortized over five years beginning in the year after the expenditures were made."

We define the permanent IL firm based on the absolute temporal dominance observed in the evolution of the knowledge stock built up from the citations weighted annual patent counts, $\tilde{c}_{fit}n_{it}$, over the period 1979 to 2000. The knowledge stock variable is computed as follows: $\sum_{s=0}^{t} \tilde{c}_{fis}n_{is}(1-\delta)^{t-s}$, where n_{is} denotes the number of successful patent applications, $\tilde{c}_{fit}$ is the number of citations received from subsequent patents (forward citations, henceforth) corrected for sample truncation bias by the *fixed effects* methodology of Hall et al (2001).[18] We employ the $\tilde{c}_{fit}$ variable as several previous works report that the number of forward citations of patents is an appropriate measure of patent quality (Lanjouw and Schankerman, 1999; Hall et al, 2001). Finally, a depreciation rate, $\delta = 15\%$ is used to account for the decreasing value of past knowledge.[19]

The firm with the highest knowledge stock in every year during the observation period is called the permanent IL of the technological cluster. Other firms in the technological cluster are assigned to the

[18] See Blazsek and Escribano (2010) for further details.

[19] In the R&D literature, several authors use $\delta = 15\%$. See for example Hall (1993) and Hall et al (2005).

permanent IF group.[20]

Table 3 shows the first 20 of the 111 firms of the technological cluster, ranked according to the mean knowledge stock over the period 1979 to 2000. The table shows the mean, computed over the period 1979 to 2000, of the following nine variables: (V1) patent applications count, n_{it}; (V2) forward citations received count, c_{fit}; (V3) forward citations received count corrected for sample truncation bias (see Hall et al, 2001), $\tilde{c}_{fit}$; (V4) knowledge stock, $\sum_{s=0}^{t} \tilde{c}_{fis} n_{is} (1-\delta)^{t-s}$; (V5) log R&D expenses, r_{it}; (V6) log book value, z_{it}; (V7) log stock market value, m_{it}; (V8) log R&D expenses to log sales, r_{it}/s_{it}; and (V9) log R&D expenses to log stock market value, r_{it}/m_{it}. In Table 3, firms are ranked according to the mean knowledge stock (V4). The second column of this table reports the patent innovation leadership cluster for each company and shows that 'Merck & Co., Inc.' (Merck, henceforth) is the permanent IL of the technological cluster for the first six out of the nine indicators considered in Table 3. The last two indicators, (V8) and (V9), are relative measures of R&D and clearly the rankings based on those indicators are very different of the rest, (V1) to (V6), and so will be the firm identified as the IL.

[APPROXIMATE LOCATION OF TABLE 3]

In order to motivate the choice of Merck as the permanent IL, Table 4 presents the evolution of the knowledge stock for the firms with the highest mean (V4) over the period 1979 to 2000, with a clear distance with the second and third firms of the ranking in terms of the stock of knowledge; Eli Lilly and Abbott Laboratories. Notice also that there was a huge increase in the knowledge stock in years 1995 to 1997 of these two firms (Figure 1 shows the aggregate effect in terms on the number of patent applications). Nevertheless, even during those three years, Merck had a stock of knowledge which is more than double, while during the rest of the years the distance is even larger; from three to seven times bigger. Table 4 also exhibits that Merck has the highest knowledge stock in every year among the firms presented, which represent the companies with the highest knowledge stock over the period 1979 to 2000 in the technological cluster (see Table 3). These results support our conclusion that Merck, is the permanent IL of the technological cluster. In addition, Figure 2 shows the number

[20] Using a different clustering procedure, firms were also classified according to patent innovations intensity to Group of Leaders (GL) and Group of Followers (GF) clusters. The GL group was formed by the following six companies: Abbott Laboratories; Bristol-Myers Squibb Co; Eli Lilly and Co; Merck & Co, Inc; Pfizer, Inc; and Warner-Lambert Co. The econometric PVAR models were also estimated for this classification and the results are qualitatively similar and available from the authors.

of patent applications and knowledge stock for the IL (Merck) and the cross-sectional mean knowledge stock of the IFs during the period 1979 to 2000. This figure also supports the selection of Merck as the permanent IL since both variables of Merck are above the mean knowledge stock and the mean number of patent applications of IFs in every year from 1979 until 2000.

[APPROXIMATE LOCATION OF TABLE 4, FIGURE 2]

The variables (V8) and (V9) are considered in this paper in order to check the robustness of the patent innovation leadership clustering procedure of our paper with the classification methods of Chambers et al (2002), Lev et al (2006) and Ciftci et al (2011). These authors employ different R&D intensity measures like; R&D to sales and R&D to market value ratios to detect R&D leadership. Table 5 summarizes the contemporaneous and dynamic cross-correlation coefficients among these variables. The results of the rankings obtained for the variables R&D to sales (V8) and R&D to market value (V9) are not consistent with the clustering method of the present paper, at least, due to the following three reasons. First, the patent IL, according to variables (V8) and (V9), is different from the IL determined by the (V4) variable.[21] Second, the present work implements a technology based and not a product-market based industry classification as Chambers et al (2002), Lev et al (2006) and Ciftci et al (2011). Third, the correlation between market value and (V8)-(V9) are negative (countercyclical), while the correlation coefficients between market value and (V1)-(V4) are positive (procyclical), motivating the choice of variable (V4) for the definition of innovation leadership. The cross-correlation coefficients corresponding to the knowledge stock (V4) are indicated by bold font in Table 5. Furthermore, from the cross-correlations we conclude that R&D and book value are procyclical but lagging with the knowledge stock. The number of patent applications and forward citations received are strongly procyclical and leads the evolution of the stock of knowledge (V4). The two relative measures of R&D intensity, (V8) and (V9), have a low and countercyclical cross-correlation with the stock of knowledge, and with several of the other measures of innovation included in Table 5, supporting our view that (V8) and (V9) are not good indicators to identify ILs.

These facts reflect the intrinsic innovation uncertainty and the uncertainty related to the appropriability of the innovation returns. This uncertainty is reduced once the innovation is patented or the innovation is kept secret.

[21] R&D leadership clustering and correlation results for alternative variables of Chambers et al (2002), Lev et al (2006) and Ciftcy et al (2008) are available from the authors.

[APPROXIMATE LOCATION OF TABLE 5]

4. PVAR models and empirical results

The PVAR models of this paper include three endogenous variables: stock return, y_{it}, log observable patent intensity, $\ln \lambda_{it}^{o}$ and log secret patent intensity, $\ln \lambda_{it}^{*}$ of $i = 1, \ldots, N$ technologically related firms over $t = 0, \ldots, T$ periods. We form the following 3×1 vector from these variables: $Y_{it} = (Y_{1it}, Y_{2it}, Y_{3it})' = (y_{it}, \ln \lambda_{it}^{o}, \ln \lambda_{it}^{*})'$. Moreover, the following deterministic, exogenous and predetermined variables are also included in the PVAR model: time trend, t, S&P500 stock index return, $\overline{y}_t$, firm size measured by log book value of equity, z_{it} and log R&D expenditure over log market value, r_{it-1}/m_{it-1}.

The summary of the econometric models and their estimation results is organized as follows. In Section 4.1, the benchmark innovation and market value Model 1 is presented, which measures the interaction among a firm's stock return, observable and latent patent intensity components without making difference among firms of the panel according to patent innovations leadership. This section also presents the specification of the dynamic Poisson patent count data model, which separates the observable and secret components of firms' patent intensity. In Section 4.2, two extended innovation and market value models are considered: First, Model 2 measures different effects for the IL and for IFs but it does not measure directly the inter-firm interaction effects of innovative activity between the IL and IFs. Second, Model 3 makes difference among firms according to patent innovation leadership and quantifies the impact of the IL on IFs and the effect of IFs on the IL.

4.1. Benchmark innovation and market value model

Model 1. The first innovation and market value model focuses on the interaction among the firm's stock return, observable patent innovations and secret patent innovations. The PVAR model is specified in two systems of three equations as follows:

$$\begin{pmatrix} y_{it} \\ \ln \lambda_{it}^{o} \\ \ln \lambda_{it}^{*} \end{pmatrix} = \begin{pmatrix} \omega_y \\ \omega_o \\ \omega_* \end{pmatrix} + \begin{pmatrix} \rho_y \\ \rho_o \\ \rho_* \end{pmatrix} t + \begin{pmatrix} \beta_y \\ \beta_o \\ \beta_* \end{pmatrix} \overline{y}_t + \begin{pmatrix} \psi_y \\ \psi_o \\ \psi_* \end{pmatrix} z_{it} + \begin{pmatrix} \phi_y \\ \phi_o \\ \phi_* \end{pmatrix} \frac{r_{it-1}}{m_{it-1}} + \begin{pmatrix} \tilde{y}_{it} \\ \ln \tilde{\lambda}_{it}^{o} \\ \ln \tilde{\lambda}_{it}^{*} \end{pmatrix} \quad (1)$$

with

$$\begin{pmatrix} \tilde{y}_{it} \\ \ln \tilde{\lambda}^o_{it} \\ \ln \tilde{\lambda}^*_{it} \end{pmatrix} = \begin{pmatrix} a_{yi} \\ a_{oi} \\ a_{*i} \end{pmatrix} + \begin{pmatrix} \zeta_{11} & \zeta_{12} & \zeta_{13} \\ \zeta_{21} & \zeta_{22} & \zeta_{23} \\ \zeta_{31} & \zeta_{32} & \zeta_{33} \end{pmatrix} \begin{pmatrix} \tilde{y}_{it-1} \\ \ln \tilde{\lambda}^o_{it-1} \\ \ln \tilde{\lambda}^*_{it-1} \end{pmatrix} + \begin{pmatrix} e_{yit} \\ e_{oit} \\ e_{*it} \end{pmatrix}, \tag{2}$$

where the patent intensity, λ_{it} is based on observable economic variables (denoted by λ^o_{it}) and unobservable information for the econometrician (denoted by λ^*_{it}). The PVAR model can be reformulated in a more compact matrix notation as follows:

$$Y_{it} = \gamma X_{it} + \tilde{Y}_{it} = \omega + \rho t + \beta \overline{y}_t + \psi z_{it} + \phi \frac{r_{it-1}}{m_{it-1}} + \tilde{Y}_{it}, \tag{3}$$

$$\tilde{Y}_{it} = a_i + \zeta \tilde{Y}_{it-1} + e_{it} \tag{4}$$

with $e_{it} \sim N(0, \Omega_e)$, $\tilde{Y}_{i0} \sim N(0, \Omega_0)$, $a_i \sim N(0, \Omega_a)$ and $\text{cov}(Y_{i0}, a_i) = \Omega_{0a}$ for $i = 1, \ldots, N$ firms that belong to the certain technologcal cluster, during $t = 0, \ldots, T$ years.

The elements of the $e_{it} = (e_{yit}, e_{oit}, e_{*it},)'$ vector of error terms may be contemporaneously correlated with each other (through Ω_e), but are uncorrelated with their own lagged values and are uncorrelated with all of the right-hand side variables of the regression equation. The $a_i \sim N(0, \Omega_a)$ is a 3×1 random vector of firm-specific *random effects* with covariance matrix Ω_a.[22] The model controls for the initial conditions, $\tilde{Y}_{i0} \sim N(0, \Omega_0)$ by introducing the Ω_0 covariance matrix of $\tilde{Y}_{i0}$ in a short-panel setup.

In the system of equations (3), ω is a 3×1 vector of constant parameters. The ρ, β, ψ and ϕ are 3×1 parameter vectors, which measure the impact of a linear time trend, t, stock index return, $\overline{y}_t$, firm size, z_{it} and lagged R&D expenses to lagged market value, r_{it-1}/m_{it-1}, respectively, on Y_{it}. The first column of the X_{it} matrix is a 3×1 vector of ones (for the constant parameters), while the subsequent columns include the exogenous explanatory variables.

In the system of equations (4), ζ and Ω_e are 3×3 parameter matrices which capture the dynamic and contemporaneous interaction, respectively, among the variables of $\tilde{Y}_{it} = (\tilde{y}_{it}, \ln \tilde{\lambda}^o_{it}, \ln \tilde{\lambda}^*_{it})'$. The $\zeta_{11}, \zeta_{22}, \zeta_{33}$ elements in the diagonal of ζ capture the first-order autoregressive effects. The six off-

[22] An alternative choice for unobservable heterogeneity could be the *fixed effects* specification discussed in Binder et at (2005).

diagonal elements of ζ measure the dynamic impacts of the three endogenous variables. More formally, ζ_{jk} for $j, k = 1, 2, 3$ captures the partial effect of $\tilde{Y}_{kit-1}$ on $\tilde{Y}_{jit}$, keeping the rest of the variables constant. Model 1 is covariance stationary when all eigenvalues of ζ are inside the unit circle. The statistical inference of Model 1 is presented in Appendix 1.

In what follows, the estimation of the observable and latent patent innovation intensities is presented, which are included in the PVAR Models. An extension of the dynamic patent count data model of Blazsek and Escribano (2010) is employed to model the patent innovation intensity (λ_{it}) of $i = 1, \ldots, N$ firms over $t = 0, \ldots, T$ periods. The model includes dynamic latent variables and it can separate patent intensity to observable (λ_{it}^{o}) and secret components (λ_{it}^{*}). These authors show that the latent variables improve the model specification of previous patent count data models.[23]

Conditional on the patent application determinants (R&D, etc.), which are "exogenous" following the exogeneity testing results of Blazsek and Escribano (2010), we want to analyze the dynamic interactions of patent application decisions based on observable and unobservable (secret) innovation determinants at the firm level and their impact on firms' rates of return. Therefore, this panel data model is useful to identify important empirical regularities that have not been considered before, and that are based on important unobserved firm level determinants of innovation like; innovation productivity, absorptive capacity, managerial ability, etc.; see Arora et al (2008).

Before presenting the patent applications count data model specification, let us introduce some notation first. In the patent count model applied, the patent applications count, n_{it} is the endogenous variable. Denote the set of patent counts by $N_{ij} = \{n_{it} : t = 0, \ldots, j\}$ with $0 \leq j \leq T$. Moreover, several exogenous explanatory variables are also considered: log R&D expenditure, r_{it}, firm size measured by the log book value, z_{it}, and number of citations made to previous patents of other firms (backward citations, henceforth) in the same industry, c_{b1it} and in other industries, c_{b2it}.[24] Let $c_{bit} = (c_{b1it}, c_{b2it})'$ denote a 2×1 vector capturing backward citations. Moreover, let Q_{ij} denote the $4N \times j$ data matrix

[23] Blazsek and Escribano (2010) show that the latent variables included in the Poisson model help to solve the potential endogeneity problem of R&D expenses reported by previous authors. Furthermore, they also show that the conditional mean function of patent counts is correctly specified in their model with respect to different specification tests.

[24] See Fung (2005) about different knowledge pools of firms.

of exogenous variables:

$$Q_{ij} = \begin{bmatrix} r_{i0} & r_{i1} & \cdots & r_{ij} \\ z_{i0} & z_{i1} & \cdots & z_{ij} \\ c_{bi0} & c_{bi1} & \cdots & c_{bij} \end{bmatrix} \tag{5}$$

with $0 \leq j \leq T$. Finally, introduce a latent variable denoted by l^*_{it}, which represents the unobservable determinants of patent applications activity. Denote the set of latent variables by $L^*_{ij} = \{l^*_{it} : t = 0, \ldots, j\}$ with $0 \leq j \leq T$.

Similarly to Hausman et al (1984), the conditional distribution of patent application counts is modeled by specifying the conditional hazard function of the point process formed by the patent arrival times. Define the conditional hazard function at instant $\tau \geq 0$ (in continuous time) corresponding to firm i in period t as follows (see Cox and Isham, 1980):

$$\lambda_{it}(\tau) = \lim_{\delta_0 \to 0} \frac{\Pr\{n_{it}(\tau + \delta_0) - n_{it}(\tau) > 0 | N_{it-1}, L^*_{it}, Q_{it}\}}{\delta_0}, \tag{6}$$

where $\delta_0 > 0$ and $n_{it}(\tau)$ is the number of patents of the firm i until instant τ in the period t. The $\lambda_{it}(\tau)$ can be interpreted as the instantaneous probability that firm i has a new patent at the point of time τ in period t given all information available in the beginning of period t. Thus, the conditional hazard represents the patent application intensity of firm i in period t.[25]

In this paper, the conditional hazard is assumed to be constant within each period, therefore, it can be indexed by t as follows: $\lambda_{it} = \lambda_{it}(\tau)$. Due to this assumption, the conditional distribution of patent counts is Poisson distribution with intensity λ_{it}. We denote the conditional density of $n_{it}|(N_{it-1}, L^*_{it-1}, Q_{it})$ as follows:

$$f_t(n_{it}|N_{it-1}, L^*_{it-1}, Q_{it}) = \frac{\exp(-\lambda_{it})\lambda_{it}^{n_{it}}}{n_{it}!}. \tag{7}$$

Furthermore, we specify the log intensity of n_{it} as follows:

$$\ln \lambda_{it} = \ln \lambda^o_{it} + \ln \lambda^*_{it}, \tag{8}$$

[25] Notice that the conditioning set in Equation (6) includes r_{it}, z_{it}, c_{bit} and l^*_{it} for period t. Thus, R&D expenses, patent citations and firm size are exogenous variables in the patent count data model.

where $\ln \lambda_{it}^{o}$ represents the observable component of patent intensity and $\ln \lambda_{it}^{*}$ denotes the secret component of patent intensity. These two components of patent application intensity are formulated as follows. First, the firm-level patent fundamentals observed for the econometrician are given by

$$\ln \lambda_{it}^{o} = \kappa_{0i} n_{i1} + \kappa_{1i} \ln \lambda_{it-1}^{o} + \kappa_{2i} r_{it} + \kappa_{3i} r_{it}^{2} + \kappa_{4i} c_{b1it} r_{it} + \kappa_{5i} c_{b2it} r_{it} + \kappa_{6i} z_{it}, \tag{9}$$

where κ_{0i} controls for initial conditions of firm i and $|\kappa_{1i}| < 1$ measures the AR(1) impact of the observable component of firm i.[26] The κ_{2i} and κ_{3i} parameters capture the linear and quadratic impacts of R&D expenses, respectively. κ_{4i} and κ_{5i} control for the interaction of R&D expenses with intra-industry and inter-industry backward citations, respectively. Finally, κ_{6i} measures the impact of firm size. Second, the patent intensity related to information that is only available for the firm but not observed for the econometrician (secret) is specified as follows:

$$\ln \lambda_{it}^{*} = \mu_{0i} + \sigma_i l_{it}^{*}, \tag{10}$$

where μ_{0i} is a parameter capturing the unobserved firm specific *fixed effects* like managerial ability, etc., σ_i is a real parameter that measures part of the firm's unobserved absorptive capacity by capturing the impact of l_{it}^{*} on patent intensity, i.e. firm level component of innovation productivity,

$$l_{it}^{*} = \mu_i^{*} l_{it-1}^{*} + u_{it} \tag{11}$$

with $u_{it} \sim N(0,1)$ i.i.d., where $|\mu_i^{*}| < 1$ captures the dynamics of unobserved firm level innovation productivity which follows an AR(1) process with positive coefficient indicating the usual persistent behavior of productivity shocks.[27]

In this paper, a univariate count data model is estimated for each firm separately. The statistical inference of the Poisson model is presented in Appendix 2. After the estimation of the count data model's parameters, the filtered estimates of λ_{it}^{*} are computed in order to be included in the PVAR models; see Appendix 3.

[26] This specification is different from Wooldridge (2005), where the n_{it-1} dynamic term is considered in $\ln \lambda_{it}$. In the present work, n_{it} includes both the observable and latent patent intensity. Therefore, $\ln \lambda_{it-1}^{o}$ is included in $\ln \lambda_{it}^{o}$, instead of n_{it-1}, to separate the observable and latent components of patent intensity.

[27] Notice that in the AR(1) specification of Equation (11), the constant is restricted to zero value due to parameter identification reasons.

Empirical results for the patent count data model. The number of parameters estimated in the univariate dynamic Poisson models is high, therefore, they are not fully reported in this paper.

In order to give a general overview of the patent intensity model estimation results, the mean patent innovations intensity estimates and patent counts, over the period 1979 to 2000, are presented in Figure 1. Our patent intensity model is able to track very well the positive trend as well as the behavior through the business cycle, of the number of patent applications per year. Moreover, the evolution of the mean observable and the mean latent patent intensity components, over the same period, is presented on Figure 3. This figure shows that the level of observable and secret patent innovations activity has been similar and approximately constant from 1979 until 1987. However, the observable component of patent intensity increased after 1987. Until 1992, the latent patent innovations intensity component stayed constant but afterwards it also increased. As the overall increase in the level of the observable component was more significant than that of the secret component, the level of the observable component was more than twice as high as that of the latent component in the mid-90's. Nevertheless, in the last years of our sample, a decreasing tendency of both components of patent application intensity can be observed. The observable component decreased after 1997, while the secret component fall after 1996. As a consequence, the level of observable patent innovation intensity is about three times higher than the level of secret patent innovation intensity in 2000.

[APPROXIMATE LOCATION OF FIGURE 3]

The main advantage of the patent count data model of this paper is that it includes a set of latent variables, l_{it}^*. In the remaining part of this subsection, the estimation results of the μ_i^* and σ_i parameters are summarized.

The firm-specific impact of l_{it}^* is measured by the σ_i parameter, interpreted as the unpredicted absorptive capacity of the firm due to secret innovations, see Escribano et al (2009) for an alternative approach. Furthermore, the long-run impact of a unit shock in u_{it} on $\sigma_i l_{it}^*$ can be expressed as follows. First, consider the infinite moving average representation of l_{it}^*:

$$l_{it}^* = \left(\frac{1}{1 - \mu_i^* L} \right) u_{it} = \left[1 + \mu_i^* L + (\mu_i^*)^2 L^2 + \ldots \right] u_{it} = \mu_i^*(L) u_{it}. \tag{12}$$

Substituting this representation of l_{it}^* into Equation (10) we obtain

$$\ln \lambda_{it}^* = \mu_{0i} + \sigma_i \mu_i^*(L) u_{it}. \tag{13}$$

In the long run ($L = 1$), this equation becomes

$$\ln \lambda_{it}^* = \mu_{0i} + \sigma_i \mu_i^*(1) u_{it}. \tag{14}$$

Therefore, the long run impact of secret innovations (u_{it}) on unobserved patent intensity is given by

$$\left. \frac{\partial \ln \lambda_{it}^*}{\partial u_{it}} \right|_{L=1} = \sigma_i \mu_i^*(1) = \frac{\sigma_i}{1 - \mu_i^*}, \tag{15}$$

which depends on two factors: the unpredicted absorptive capacity due to unobserved innovation factors (σ_i) and the unobserved degree of persistence of the secret innovation process (μ_i^*). The interpretation of $1/\sigma_i$ is derived as follows. First, rewrite Equation (10) as:

$$l_{it}^* = \frac{1}{\sigma_i} (\ln \lambda_{it}^* - \mu_i). \tag{16}$$

Then, we can express $1/\sigma_i$ by taking the following derivative of l_{it}^*:

$$\frac{\partial l_{it}^*}{\partial (\ln \lambda_{it}^* - \mu_i)} = \frac{1}{\sigma_i} \tag{17}$$

and it measures the marginal effects of secret patent intensity ($\ln \lambda_{it}^*$) on unobserved innovation productivity (l_{it}^*). Table 6 shows 30 firms from the technological cluster, ranked according to two indicators: (1) their predictive absorptive capacity generated from the innovations kept secret by the i-th firm, $1/\sigma_i$ and (2) their mean log market value, $(1/T) \sum_{t=1}^{T} m_{it}$. One important conclusion comes out of Table 6; Merck is not only the IL firm but it is also the firm with the highest predictive absorptive capacity and with the highest transform rate of secret patent intensity into unobserved innovation productivity, see Equation (18).

[APPROXIMATE LOCATION OF TABLE 6]

Figure 4 presents the estimates of $1/\sigma_i$ for the 60 firms, as a function of the means of the variables

(V1)-(V7) and mean annual stock return, computed over the period 1979 to 2000. The figure exhibits the estimates of the linear regression line fitted to $1/\sigma_i$ and the corresponding R-squared values to inform about the explanatory power of each variable. The first four panels in Figure 4 show that firms with high patent applications count (V1), high patent citations received count (V2, V3) and high patent citations weighted patent applications count (V4) tend to have higher $1/\sigma_i$. Panel five presents that firms with high R&D expenses (V5) have high predictive absorptive capacity. Panel six exhibits that large firms (high book value, V6) have higher predictive absorptive capacity. Finally, the last two panels evidence that firms with high log market value (V7) and high annual stock return also tend to have higher predictive absorptive capacity. In summary, Figure 4 shows that those firms with a high predictive unobserved absorptive capacity, are those that invest more in R&D, receive more forward citations, apply for more patents and have higher stock returns and stock market values. Between 30% to 51% of the variability (R-squared) of the predictive absorptive capacity due to secret innovation is explained by individual firm level performance measures.

[APPROXIMATE LOCATION OF FIGURE 4]

Empirical results for Model 1. Table 7 presents the parameter estimates of Model 1. First, in the system of equations (1), significant positive trends are estimated for stock return ($\hat{\delta}_y = 0.02$), observable patent intensity ($\hat{\delta}_o = 0.46$) and secret patent intensity ($\hat{\delta}_* = 0.18$). Significant and positive β coefficients are measured for the impact of the S&P500 stock index return. The highest $\hat{\beta}_y = 0.72$ coefficient is observed for the stock return and positive $\hat{\beta}_o = 0.09$ and $\hat{\beta}_* = 0.05$ and significant parameters are found for the impact of the S&P500 return on observable patent innovations and secret patent innovations, respectively. The influence of firm size is different for each endogenous variable: (1) smaller companies tend to have higher stock returns than large ones, i.e. $\hat{\psi}_y = -0.15$; (2) large firms have a positive impact on observable patent intensity, i.e. $\hat{\psi}_o = 0.02$; and (3) smaller firms produce more secret innovations than large companies, i.e. $\hat{\psi}_* = -0.06$. Finally, lagged R&D intensity measured by R&D to market value has a significant positive effect on firms' stock market valuation ($\hat{\phi}_y = 3.15$) and on observable patent activity ($\hat{\phi}_o = 0.05$), while it has a significant negative effect on latent patent innovation intensity ($\hat{\phi}_* = -0.15$), which is consistent with the negative cross-correlations between (V9) and the stock of knowledge shown in Table 5.

Second, in the system of market adjusted equations (2), the ζ matrix shows significant and positive

PVAR(1) dynamics of both observable and secret patent intensity components with higher persistence in the latent component. There is a significant positive dynamic impact of observable patent intensity ($\hat{\zeta}_{12} = 0.19$) and a significant positive dynamic impact of secret patent activity ($\hat{\zeta}_{13} = 0.70$) on stock return. Therefore, the results suggest that the dynamic impact of the latent patent intensity component on stock return is much higher than that of the observable component. Furthermore, a positive dynamic interaction is found between the observable and latent intensity components: $\hat{\zeta}_{23} = 0.70$ and $\hat{\zeta}_{32} = 0.21$. The estimates show that the impact of the secret component on the observable component is higher than the opposite effect: $(\hat{\zeta}_{23} = 0.70) > (\hat{\zeta}_{32} = 0.21)$. We find that the PVAR process is covariance stationary (see the maximum modulus of the eigenvalues of $\hat{\zeta}$ in Table 7). Finally, the estimates presented in the Ω_e matrix measure the contemporaneous interaction among the three endogenous variables. The parameter values show a significant positive contemporaneous interaction between shocks in the stock return and shocks in the secret patent intensity ($\hat{\Omega}_{e13} = 0.36$), and between shocks in observable and latent patent innovations activity ($\hat{\Omega}_{e23} = 0.37$). However, a negative contemporaneous covariance is found between shocks in the stock return and observable patent activity ($\hat{\Omega}_{e12} = -0.20$).

[APPROXIMATE LOCATION OF TABLE 7]

IRF of Model 1. In order to interpret the dynamic effects of the parameter estimates of the PVAR model over several lags, the IRFs are also estimated. These represent the impact of a current unit shock in the orthogonal (structural) error terms, $\epsilon_{it} = (\epsilon_{yit}, \epsilon_{oit}, \epsilon_{*it})' = (\sqrt{\Omega_e})^{-1} e_{it}$, on future values of $\tilde{Y}_{it}$. We use the methodology of Cao and Sun (2011), who establish the asymptotic distributions of the IRFs in short PVARs. They derive the estimate of the IRF and show how the corresponding confidence bands can be estimated. Cao and Sun (2011) prove the asymptotic validity of a bootstrap approximation of confidence bands.

The orthogonalized IRF of Model 1 is derived in Appendix 4, where we obtain that

$$\tilde{Y}_{it} = \sum_{j=0}^{\infty} \zeta^j a_i + \sum_{j=0}^{\infty} \zeta^j \sqrt{\Omega_e} \epsilon_{it-j}. \tag{18}$$

Therefore, the IRF is given by the following infinite sequence of the matrices of standardized dynamic multipliers, Θ_j:

$$\Theta_j = (\zeta)^j \sqrt{\Omega_e} \text{ for } j = 0, 1, 2, \ldots, \infty, \tag{19}$$

where $\sqrt{\Omega_e}$, the lower triangular Cholesky matrix of Ω_e, is used to orthogonalize the IRF. The 95% confidence bands around the estimate of Θ_j are obtained by the simulation based method suggested by Cao and Sun (2011). The 97.5% and 2.5% quantiles of Θ_j are estimated by 10,000 Monte Carlo replications of Θ_j.

In the VAR literature, there are different choices for the matrix used to orthogonalize the IRF. The different matrices are determined by the different orders of equations in the system. We have three endogenous variables in the PVAR model, i.e. there are six possible orders of the equations. The Cholesky matrix, $\sqrt{\Omega_e}$ and the standardized dynamic multiplier, Θ_j is different for each order. Nevertheless, the estimation procedure of Model 1, presented in Appendix 1, assumes that the diagonal of $\sqrt{\Omega_e}$ is normalized to ones for parameter identification reasons. This assumption implies a certain order for the equations as only for the order $(y_{it}, \ln\lambda^o_{it}, \ln\lambda^*_{it})$ it is obtained that the diagonal of $\sqrt{\Omega_e}$ contains ones. For different orders of equations, the diagonal of the Cholesky matrix does not equal to ones, therefore, it is not compatible with the normalization assumption of Appendix 1. As a consequence, we apply the order $(y_{it}, \ln\lambda^o_{it}, \ln\lambda^*_{it})$ to orthogonalize the IRF. The economic intuition behind these a priori identification restrictions is given below, after Equation (20).

The following restrictions are imposed on the contemporaneous relations (see Appendix 1):

$$\Theta_0\epsilon_{it} = (\zeta)^0\sqrt{\Omega_e}\epsilon_{it} = \begin{pmatrix} 1 & 0 & 0 \\ \Omega_{e_{21}} & 1 & 0 \\ \Omega_{e_{31}} & \Omega_{e_{32}} & 1 \end{pmatrix} \begin{pmatrix} \epsilon_{yit} \\ \epsilon_{oit} \\ \epsilon_{*it} \end{pmatrix} = \begin{pmatrix} \epsilon_{yit} \\ \Omega_{e21}\epsilon_{yit} + \epsilon_{oit} \\ \Omega_{e31}\epsilon_{yit} + \Omega_{e32}\epsilon_{oit} + \epsilon_{*it} \end{pmatrix}. \tag{20}$$

The intuitive justification of these restrictions is the following: It is clear that the firm's innovation decisions are affected by more than one type of shock. Therefore, for identification reasons, we need to impose certain a priori restrictions. The restrictions we are imposing are based on the presumption "innovations and patent intensity news affect stock returns with a lag, while any variation in stock returns is rapidly affecting innovation decisions". News on observable patent intensity, affect contemporaneously patent intensity decisions based on secret innovations, but not the other way around.

That is, all the contemporaneous relations among the three variables, are associated either to financial information (stock returns) or to observable patent intensity decisions. Secret innovations require more time to be spread within the firm given to warranty the information is confidential. However, when these secret innovations are transmitted internally or spillover other firms, they are expected to have higher and more persistent impacts than observable innovations.

The contemporaneous relationships among the endogenous variables, $\tilde{Y}_{it}$ are calculated as follows. We orthogonalize Equation (4) multiplying each term by $(\sqrt{\Omega_e})^{-1}$:

$$(\sqrt{\Omega_e})^{-1}\tilde{Y}_{it} = (\sqrt{\Omega_e})^{-1}a_i + (\sqrt{\Omega_e})^{-1}\zeta\tilde{Y}_{it-1} + \epsilon_{it}. \tag{21}$$

According to the estimates of $(\sqrt{\Omega_e})^{-1}$, reported in Table 7, the left hand side of Equation (21) is

$$\begin{pmatrix} 1.00 & 0.00 & 0.00 \\ 0.20 & 1.00 & 0.00 \\ -0.45 & -0.44 & 1.00 \end{pmatrix} \begin{pmatrix} \tilde{y}_{it} \\ \ln\tilde{\lambda}^o_{it} \\ \ln\tilde{\lambda}^*_{it} \end{pmatrix} = \begin{pmatrix} \tilde{y}_{it} \\ 0.20\tilde{y}_{it} + \ln\tilde{\lambda}^o_{it} \\ -0.45\tilde{y}_{it} - 0.44\ln\tilde{\lambda}^o_{it} + \ln\tilde{\lambda}^*_{it} \end{pmatrix} \tag{22}$$

Combining Equations (21) and (22), we obtain that the contemporaneous relationships are indicated by the first term of the right hand side of the following equation:

$$\begin{pmatrix} \tilde{y}_{it} \\ \ln\tilde{\lambda}^o_{it} \\ \ln\tilde{\lambda}^*_{it} \end{pmatrix} = \begin{pmatrix} 0.00 \\ -0.20\tilde{y}_{it} \\ 0.45\tilde{y}_{it} + 0.44\ln\tilde{\lambda}^o_{it} \end{pmatrix} + \begin{pmatrix} a_{yi} \\ 0.20a_{yi} + a_{oi} \\ -0.45a_{yi} - 0.44a_{oi} + a_{*i} \end{pmatrix} +$$

$$+ \begin{pmatrix} 0.04\tilde{y}_{it-1} + 0.19\ln\tilde{\lambda}^o_{it-1} + 0.7\ln\tilde{\lambda}^*_{it-1} \\ 0.02\tilde{y}_{it-1} + 0.26\ln\tilde{\lambda}^o_{it-1} + 0.84\ln\tilde{\lambda}^*_{it-1} \\ 0.01\tilde{y}_{it-1} + 0.03\ln\tilde{\lambda}^o_{it-1} + 0.08\ln\tilde{\lambda}^*_{it-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{yit} \\ \epsilon_{oit} \\ \epsilon_{*it} \end{pmatrix}. \tag{23}$$

The infinite Vector Moving Average (VMA) representation of $\tilde{Y}_{it}$ helps to interpret the matrix of standardized dynamic multipliers, Θ_j presented on Figure 5:

$$\tilde{Y}_{it} = (I_3 - \zeta L)^{-1}a_i + \sum_{j=0}^{\infty}\Theta_j\epsilon_{it-j}. \tag{24}$$

In the previous formula, $\Theta_j = \partial \tilde{Y}_{it}/\partial \epsilon_{it-j}$ measures how unit impulses of standardized shocks at time $t-j$ impact $\tilde{Y}_{it}$. More precisely, each element of Θ_j can be expressed as follows:

$$\Theta_j = \begin{pmatrix} \Theta_{j11} & \Theta_{j12} & \Theta_{j13} \\ \Theta_{j21} & \Theta_{j22} & \Theta_{j23} \\ \Theta_{j31} & \Theta_{j32} & \Theta_{j33} \end{pmatrix} = \begin{pmatrix} \frac{\partial \tilde{y}_{it+j}}{\partial \epsilon_{yit}} & \frac{\partial \tilde{y}_{it+j}}{\partial \epsilon_{oit}} & \frac{\partial \tilde{y}_{it+j}}{\partial \epsilon_{*it}} \\ \frac{\partial \ln \tilde{\lambda}^o_{it+j}}{\partial \epsilon_{yit}} & \frac{\partial \ln \tilde{\lambda}^o_{it+j}}{\partial \epsilon_{oit}} & \frac{\partial \ln \tilde{\lambda}^o_{it+j}}{\partial \epsilon_{*it}} \\ \frac{\partial \ln \tilde{\lambda}^*_{it+j}}{\partial \epsilon_{yit}} & \frac{\partial \ln \tilde{\lambda}^*_{it+j}}{\partial \epsilon_{oit}} & \frac{\partial \ln \tilde{\lambda}^*_{it+j}}{\partial \epsilon_{*it}} \end{pmatrix}. \tag{25}$$

A general element, Θ_{jkl} is interpreted as the impact of one unit standardized shock, ϵ_{lit} on the variable Y_{kit+j}.

Figure 5 exhibits for Model 1 the evolution of six components of Θ_j and the corresponding confidence bands for $j = 1, \ldots, 15$ future periods to report decreasing dynamic effects of the orthogonalized error terms on future stock returns, observable patent intensities and secret patent intensities. Figure 5 shows: a) The impacts of secret innovation shocks are the largest. The dynamic effects on stock returns, on observable and on secret patent intensities are of similar positive magnitude with a similar slow path of decline, reaching zero only after 12 years. b) Observable innovation shocks have impacts of lower positive magnitude on stock returns, observable and secret sources of patent intensity and share common high speed of decline, reaching zero only after two or three years.

[APPROXIMATE LOCATION OF FIGURE 5]

Our PVAR models are covariance stationary, therefore, $\lim_{j\to\infty}\Theta_{jkl} = 0$ for $k, l = 1, 2, 3$ and no standard shock, ϵ_{it-j} has long-run impact on $\tilde{Y}_{it}$. However, it is interesting to evaluate the long-run cumulative impact of standardized shocks by considering the following long-run impact matrix:

$$\Theta_j(1) = \begin{pmatrix} \Theta_{11}(1) & \Theta_{12}(1) & \Theta_{13}(1) \\ \Theta_{21}(1) & \Theta_{22}(1) & \Theta_{23}(1) \\ \Theta_{31}(1) & \Theta_{32}(1) & \Theta_{33}(1) \end{pmatrix} = \begin{pmatrix} \sum_{j=0}^{\infty} \Theta_{j11} & \sum_{j=0}^{\infty} \Theta_{j12} & \sum_{j=0}^{\infty} \Theta_{j13} \\ \sum_{j=0}^{\infty} \Theta_{j21} & \sum_{j=0}^{\infty} \Theta_{j22} & \sum_{j=0}^{\infty} \Theta_{j23} \\ \sum_{j=0}^{\infty} \Theta_{j31} & \sum_{j=0}^{\infty} \Theta_{j32} & \sum_{j=0}^{\infty} \Theta_{j33} \end{pmatrix}. \tag{26}$$

In this paper, $\Theta_j = (\zeta)^j \sqrt{\Omega_e}$, therefore, the long-run impact matrix of Model 1 can be computed as follows:

$$\Theta_j(1) = \sum_{j=0}^{\infty} \Theta_j = \sum_{j=0}^{\infty} (\zeta)^j \sqrt{\Omega_e} = (I_3 - \zeta)^{-1} \sqrt{\Omega_e}. \tag{27}$$

The long-run impact matrix, $\Theta_j(1)$, estimated for Model 1 is presented in Table 7. The results show that the structural latent innovation shocks has the highest long-term impact on all variables, followed by, the long-term impact of the structural observable innovation shocks and stock return error terms. In summary, among the innovation factors affecting the firms, the secret innovation kept by the firms is the one affecting more, and positively, the returns of the firms. This important empirical finding is consistent with the results obtained by Pakes (1985), in a different context, where unpredictable increases in patents increases the market value of firms.

4.2. Extended innovation and market value models

In this section, two extensions of the benchmark PVAR model are proposed (Models 2 and 3) and the corresponding estimation results are summarized. Model 2 measures different effects for the IL and for IFs. Nevertheless, it does not measure directly the inter-firm interaction effects of innovative activity between the IL and IFs. Model 3 parameterizes the impact of the IL on IFs and the effect of IFs on the IL.

Model 2. Differential effects of the single IL in innovation. Model 2 considers IL and IF companies of the technological cluster. The model is specified as follows:

$$Y_{it} = \gamma X_{it} + \gamma_{\text{IL}} X_{it} D(i = \text{IL}) + \tilde{Y}_{it}, \tag{28}$$

$$= \left(\omega + \delta t + \beta \overline{y}_t + \psi z_{it} + \phi \frac{r_{it-1}}{m_{it-1}} \right) +$$

$$+ \left(\omega_{\text{IL}} + \delta_{\text{IL}} t + \beta_{\text{IL}} \overline{y}_t + \psi_{\text{IL}} z_{it} + \phi_{\text{IL}} \frac{r_{it-1}}{m_{it-1}} \right) D(i = \text{IL}) + \tilde{Y}_{it}$$

$$\tilde{Y}_{it} = a_i + \zeta \tilde{Y}_{it-1} + \zeta_{\text{IL}} \tilde{Y}_{it-1} D(i = \text{IL}) + e_{it} \tag{29}$$

with $e_{it} \sim N(0, \Omega_e)$, $\tilde{Y}_{i0} \sim N(0, \Omega_0)$, $a_i \sim N(0, \Omega_a)$ and $\text{cov}(Y_{i0}, a_i) = \Omega_{0a}$. The dummy variable $D(i = \text{IL}) = 1$ if $i = \text{IL}$ and zero otherwise. In systems of equations (28) and (29), the γ_{IL} and ζ_{IL} matrices measure the additional impacts of the IL firm not captured by the γ and ζ parameter matrices, respectively. Thus, the IF effects are captured by the estimates of γ and ζ, while the IL effects are measured by the $(\gamma + \gamma_{\text{IL}})$ and $(\zeta + \zeta_{\text{IL}})$ matrices. The interpretation of the components of the $(\gamma + \gamma_{\text{IL}})$ and $(\zeta + \zeta_{\text{IL}})$ matrices coincide with those of the γ and ζ matrices, respectively, for Model

1. Model 2 is covariance stationary when all eigenvalues of both ζ and $(\zeta + \zeta_{\text{IL}})$ are inside the unit circle. Interpretation of the parameters also included in Model 1 is the same as before. The estimation of Model 2 is summarized in Appendix 1.

Empirical results for Model 2. The parameter estimates of Model 2 are presented in Table 8. First, in the system of equations (28), the γ parameter matrix indicates the parameters of the firms that are IFs and similar estimation results are obtained as in Model 1. Therefore, the results reported for Model 1 for the γ parameter are valid for the IFs in Model 2. For the IL firm, the estimates of $(\gamma + \gamma_{\text{IL}})$ for Model 2 indicates the behavior of the leader. Most parameters of the $(\hat{\gamma} + \hat{\gamma}_{\text{IL}})$ matrix are similar to $\hat{\gamma}$. Nevertheless, the impact of firm size on stock returns of the IL is not significant $(\hat{\psi}_y + \hat{\psi}_{\text{IL}y} = -0.12 + 0.12 = 0.00)$, and firm size has a positive effect on the secret patent innovations intensity of the IL $(\hat{\psi}_* + \hat{\psi}_{\text{IL}*} = -0.02 + 0.05 = 0.03)$.

Second, in the system of equations (29), similar estimation results are also obtained for the ζ and Ω_e parameter matrices as in Model 1. Therefore, the results for the ζ parameters reported for Model 1 are valid for IF companies in Model 2, and the results reported for Model 1 for the Ω_e parameters are valid for all companies in the sample. The estimates of $(\zeta + \zeta_{\text{IL}})$ can be interpreted for the IL firm in Model 2. Most but some key parameters of the $(\hat{\zeta} + \hat{\zeta}_{\text{IL}})$ matrix are similar to $\hat{\zeta}$. In particular, lagged stock return has a positive effect on current secret patent innovations intensity $(\hat{\zeta}_{31} + \hat{\zeta}_{31\text{IL}} = 0.10)$ and on current returns $(\hat{\zeta}_{11} + \hat{\zeta}_{11\text{IL}} = 0.23)$ for the IL. Finally, the AR process is found to be stationary for Model 2 since all eigenvalues of $\hat{\zeta}$ and $(\hat{\zeta} + \hat{\zeta}_{\text{IL}})$ are inside the unit circle (see Table 8). However, the persistence on the stock returns of the IL is much higher $(\hat{\zeta}_{11} + \hat{\zeta}_{11\text{IL}} = 0.23$, AR(1) coefficient) than that of the IFs $(\hat{\zeta}_{11} = 0.04)$. The opposite result is obtained for the two (observed and unobserved) patent intensity components. Therefore, the marked adjusted patent intensities of the innovation leader is less predictable than the followers.

[APPROXIMATE LOCATION OF TABLE 8]

IRF of Model 2. We apply the methodology of Cao and Sun (2011) to derive the estimates of the orthogonalized IRFs and the corresponding 95% confidence bands of Model 2. The IRFs are derived in Appendix 4, where we obtain that

$$\tilde{Y}_{it} = \sum_{j=0}^{\infty}[\zeta + \zeta_{\mathrm{IL}} D(i = \mathrm{IL})]^j a_i + \sum_{j=0}^{\infty}[\zeta + \zeta_{\mathrm{IL}} D(i = \mathrm{IL})]^j (\sqrt{\Omega_e})\epsilon_{it-j}. \tag{30}$$

Therefore, the IRFs are given by

$$\Theta_{ij}(\mathrm{IL}) = (\zeta + \zeta_{\mathrm{IL}})^j \sqrt{\Omega_e} \text{ for } i = \mathrm{IL} \text{ and } j = 0, 1, 2, \ldots, \infty, \tag{31}$$

$$\Theta_{ij}(\mathrm{IF}) = \zeta^j \sqrt{\Omega_e} \text{ for } i \in \mathrm{IF} \text{ and } j = 0, 1, 2, \ldots, \infty. \tag{32}$$

Figure 5 exhibits for Model 2 the evolution of six components of Θ_{ij} for $i = \mathrm{IL}$, denoted by $\Theta_{ij}(\mathrm{IL})$ for $j = 1, \ldots, 15$ future periods, to exhibit decreasing dynamic effects of the orthogonalized error terms on future stock returns, observable patent intensities and secret patent intensities. The IRFs corresponding to the IF company are not reported since they are similar to the IRF presented for Model 1 of Figure 5. The IRF of the IL are only similar to the IRF of the IFs for the shocks an returns. The positive reactions to shocks on innovation intensity components are of much lower magnitude than those of IFs and their speed of decline is much faster for the IL than for the rest.

The long-run impact matrix of Model 2 for IL and IF firms can be computed as follows:

$$\Theta_{ij}(1) = (I_3 - \zeta - \zeta_{\mathrm{IL}})^{-1} \sqrt{\Omega_e} \text{ for } i = \mathrm{IL} \text{ and } j = 0, 1, 2, \ldots, \infty \tag{33}$$

and

$$\Theta_{ij}(1) = (I_3 - \zeta)^{-1} \sqrt{\Omega_e} \text{ for } i \in \mathrm{IF} \text{ and } j = 0, 1, 2, \ldots, \infty. \tag{34}$$

The estimates of the long-run impact matrices are presented in Table 8. For the IF firms, the results show that the structural impulse of the latent innovation error term is the largest in terms of the quantitative impact but is the most persistent shock, followed by, the observable patent intensity and stock return shocks for all variables. For the IL company, all long run impacts are also positive but of a much lower magnitude than for the rest of the firms (IFs). The long-run impact of stock return impulses is the highest for $\tilde{y}_{it}$, the impact of observable innovation impulses is the highest for $\ln \tilde{\lambda}^o_{it}$ and the impact of secret innovation impulses is the highest for $\ln \tilde{\lambda}^*_{it}$. The main conclusion we extract from this analysis is that the long run impacts of any of shocks are much larger for innovation follower firms than for the leader.

Model 3. Interactions between the patent innovations of IL and IFs. This specification measures the impact of the IL company on IFs and IFs on the IL firm. The model is formulated as follows:

$$Y_{it} = \gamma X_{it} + \tilde{Y}_{it} = \omega + \delta t + \beta \overline{y}_{it} + \psi z_{it} + \phi \frac{r_{it-1}}{m_{it-1}} + \tilde{Y}_{it}, \tag{35}$$

$$\tilde{Y}_{it} = a_i + \zeta \tilde{Y}_{it-1} + D(i \in \mathrm{IF}) \zeta_{\mathrm{IL}} \tilde{Y}_{\mathrm{IL}t-1} + D(i = \mathrm{IL}) \left(\sum_{k \in \mathrm{IF}} \zeta_{\mathrm{IF}} \tilde{Y}_{kt-1} \right) + e_{it} \tag{36}$$

with $e_{it} \sim N(0, \Omega_e)$, $\tilde{Y}_{i0} \sim N(0, \Omega_0)$, $a_i \sim N(0, \Omega_a)$ and $\mathrm{cov}(Y_{i0}, a_i) = \Omega_{0a}$. The parameters with subindex IL measure the impact of the IL firm on IF companies, while the parameters with subindex IF capture the effects of IF firms on the IL. The interpretation of the parameters also included in Model 1 is the same as before. Model 3 is covariance stationary when all eigenvalues of ζ are inside the unit circle.

Empirical results for Model 3. The parameter estimates of Model 3 are presented in Table 9. First, in the system of equations (35), for the γ parameters similar values are estimated as in the previous PVAR models. Second, in the systems of equations (36), for the ζ and Ω_e parameter matrices similar results are obtained as in Models 1 and 2. The AR process is covariance stationary for Model 3 (see Table 9). Finally, several parameters of the ζ_{IL}, ζ_{IF} matrices, which account for the interaction effects of the IL and IFs, are found to be significantly different from zero. More details about the multiperiod dynamic results of inter-firm interaction are provided in the following subsection.

[APPROXIMATE LOCATION OF TABLE 9]

IRF and dynamic multipliers of Model 3. We employ the methodology of Cao and Sun (2011) to compute the orthogonalized IRFs and their 95% confidence bands of Model 3. In addition, the so-called dynamic multipliers are also computed for Model 3, which account for the dynamic inter-firm effects between the patent IL and IFs. The IRFs and the dynamic multipliers are derived in Appendix

4, where we obtain that

$$\tilde{Y}_{it} = \sum_{j=0}^{\infty} \zeta^j a_i + \sum_{j=0}^{\infty} \zeta^j \zeta_{\text{IL}} \tilde{Y}_{\text{IL}t-1-j} D(i \in \text{IF}) + \tag{37}$$

$$+ \sum_{j=0}^{\infty} \sum_{k \in \text{IF}} \zeta^j \zeta_{\text{IF}} \tilde{Y}_{kt-1-j} D(i = \text{IL}) + \sum_{j=0}^{\infty} \zeta^j (\sqrt{\Omega_e}) \epsilon_{it-j}.$$

Therefore, the IRFs of Model 3 are given by:

$$\Theta_j = (\zeta)^j \sqrt{\Omega_e} \text{ for } j = 0, 1, 2, \ldots, \infty. \tag{38}$$

Moreover, the dynamic multipliers or interaction effects between leaders and followers are given by:

$$\Gamma_j(\text{IL} \to \text{IF}) = (\text{Effects of } \tilde{Y}_{\text{IL}t-j} \text{ on } \tilde{Y}_{it} \text{ for } i \in \text{IF}) = \zeta^j \zeta_{\text{IL}} \text{ for } j = 1, 2, \ldots, \infty, \tag{39}$$

$$\Gamma_j(\text{IF} \to \text{IL}) = (\text{Effects of } \tilde{Y}_{kt-j} \text{ for } k \in \text{IF on } \tilde{Y}_{\text{IL}t}) = \zeta^j \zeta_{\text{IF}} \text{ for } j = 1, 2, \ldots, \infty. \tag{40}$$

Figure 6 exhibits the evolution of twelve components of $\Gamma_j(\text{IL} \to \text{IF})$ and $\Gamma_j(\text{IF} \to \text{IL})$ for $j = 1, \ldots, 15$ future periods to report decreasing dynamic effects of other companies on the firm's future stock returns, observable patent intensities and secret patent intensities. Figure 6 shows the next results for different components of Γ.

The interactions from the IL to the IFs are the following: a) Shocks on observable patent intensities of the IL have low positive impacts on stock returns and patent intensities of IFs that vanish after two years. The effects on secret patent intensities of the IFs are negative but also last only two years. b) Shocks on secret patent intensities of the IL have much larger effects on the IFs and the spillovers last more than 10 years before disappearing. Those innovations that spillover from IL to IFs are positive for stock returns and observables patent intensities (complements) and are negative for secret patent intensities (substitutes). Interactions from the IFs to the IL are of much lower magnitude: a) The impact of observable patent intensities of the IFs on the IL are small and last only one year. Those spillover effects are negative for stock returns and secret patent intensities and positive for observable patent intensity (complements). b) The spillovers from secret patent intensities of the IFs on the IL are more important and last longer (10 years). The signs are the same as the spillovers from IFs to IL on observable patent intensities. Therefore, the dominant spillovers goes from the IL to the IFs and are specially important if they are based on secret patent information decisions of the firm.

The long-run impact matrix of Model 3 can be computed by $\Theta_j(1) = (I_3 - \zeta)^{-1}\sqrt{\Omega_e}$. The estimate of this matrix is presented in Table 9. The results show that the structural latent error term has the highest long-term impact on all variables, followed by the structural observable and stock return error terms.

[APPROXIMATE LOCATION OF FIGURE 6]

5. Summary and conclusions

In this paper, we use dynamic panel data models to identify the interaction among stock return, observable and secret patent application activity for a cluster of technologically related US firms for the 22-year time period of 1979 to 2000. We create the technological cluster by Ward's (1963) method by using firm-level data on patent counts over the technological category space. The technological cluster analyzed includes 111 firms. Most of these firms are from the pharmaceutical product-market sector (SIC283). However, some firms of the cluster operate in different product-market industries, like computers, chemicals or food sectors.

We classify the firms of the technological cluster according to their patent stock of knowledge into patent innovation leader and followers groups. In particular, we identify a single Innovation Leader (IL) firm over the period 1979 to 2000: Merck & Co. As this company preserves its leadership during the whole period, we say that Merck is the permanent IL of the technological cluster. We assign other technologically related firms to the Innovation Follower (IF) cluster.

We estimate several dynamic PVAR market value models by the Quasi Maximum Likelihood (QML) method of Binder et al (2005). In these models, a secret patent application intensity component is included. We model this latent patent intensity component according to the Poisson-type count data framework of Blazsek and Escribano (2010). This patent count data model is estimated by the MSL method, applying the Efficient Importance Sampling (EIS) variance reduction technique of Richard and Zhang (2007).

The parameter estimates, the IRFs and the dynamic multipliers of the PVAR market value models have suggested significant contemporaneous and dynamic intra-firm and inter-firm (spillover) effects among stock returns, observable patent intensity and secret patent intensity in the technological cluster analyzed. The main conclusions are the following:

a) Empirical results considering all firms (Model 1): Innovation shocks based on observables have a small and short (only two years) positive impact on stock returns and on observable and unobservable patent intensities. However, innovations shocks based on secret information have large and long (more than 10 years) positive effects on stock returns and on observable and unobservable patent intensities.

b) Results considering separately IL and IFs (Model 2): The results for IFs are similar to those mentioned before for all firms. However, the behavior of the IL is different. The innovation shocks based on observables are also short (2 years) but smaller than the rest of the firms. This is also valid for the innovation shocks based on secret patent intensities of the IL but it is not true for the stock returns of the IL; the reaction of the stock returns of the IL is similar to the rest of the firms (IFs). In summary, the stock returns reaction of the IL is similar to the IFs but the innovations reactions of the IL are smaller and have shorter memory.

c) Innovation spillovers among innovation leaders and followers (Model 3): The flow of innovation spillovers is larger from the IL to the IFs, than the other way around. The interaction effects of innovation shocks are large if they flow from IL to IFs, specially if the source of innovation is secret. These spillovers are positive between observable patent intensities (innovation complements) and negative between secret patent intensities (innovation substitutes). In both cases, the spillovers last more than 10 years from IL to IFs and around 5 years if the flow goes from IFs to IL. The reaction on stock returns of the IL and the IFs differ in magnitude and sign; the main spillovers again come from the secret innovation component (last 10 years) and are positive going from IL to IFs and are negative, but of similar magnitude, going form IFs to the IL. In summary, the innovation spillovers benefit more IFs than the IL.

Acknowledgments

The authors thank to Luc Bauwens, Andrea Fosfuri, Marco Giarratana, Luis Albériko Gil-Alana, Christian Gouriéroux, Bronwyn Hall, Jerry Hausman, Javier Hualde, Thierry Kamionka, Miklós Koren, Gábor Kőrösi, Germán López, Michel Lubrano, Velayoudom Marimoutou, Antonio Moreno, Mikel Tapia, Jérôme Vandenbussche and other members of CREST, GREQAM, Universidad Carlos III de Madrid and Universidad de Navarra for all support. For helpful discussions and comments on previous versions of this paper the authors thank to participants of MKE Annual Conference, Budapest, December 2009; ETSERN Pamplona Meeting, June 2010; and Summer Workshop in Economics of the Hungarian Academy of Sciences, MTA-KTI, Budapest, July 2010. In this context, we thank Miklós Koren for helpful discussions and feedback. The first author acknowledges the research scholarship of CREST, Paris, and the research financing of the PIUNA project of Universidad de Navarra. Funding from the Bank of Spain Excellence Program is gratefully acknowledged by the second author. Both authors acknowledge funding from MICFIN (ECO2009-08308) network.

Appendix 1: Estimation of PVAR models by QML

This appendix presents the statistical inference applied for the PVAR Models 1, 2 and 3. These models are estimated by the QML method proposed by Binder et al (2005).[28] Definitions of parameters in these equations are presented in Section 4. In order to simplify the notation, let N refer to the sample size of all models in this appendix.

The covariance matrices Ω_e, Ω_0 and Ω_a are parameterized as follows. In order to identify the parameters in these matrices, the diagonals of the $\sqrt{\Omega_e}$, $\sqrt{\Omega_0}$ and $\sqrt{\Omega_a}$ matrices are restricted to ones:

$$\sqrt{\Omega_e} = \begin{bmatrix} 1 & 0 & 0 \\ \Omega_{e_{21}} & 1 & 0 \\ \Omega_{e_{31}} & \Omega_{e_{32}} & 1 \end{bmatrix}, \quad \sqrt{\Omega_0} = \begin{bmatrix} 1 & 0 & 0 \\ \Omega_{0_{21}} & 1 & 0 \\ \Omega_{0_{31}} & \Omega_{0_{32}} & 1 \end{bmatrix}, \quad \sqrt{\Omega_a} = \begin{bmatrix} 1 & 0 & 0 \\ \Omega_{a_{21}} & 1 & 0 \\ \Omega_{a_{31}} & \Omega_{a_{32}} & 1 \end{bmatrix}. \quad \text{(A1.1)}$$

These Cholesky matrices imply the following parameterization of Ω_e, Ω_0 and Ω_a:

$$\Omega_e = \begin{bmatrix} 1 & \Omega_{e_{21}} & \Omega_{e_{31}} \\ \Omega_{e_{21}} & \Omega_{e_{21}}^2 + 1 & \Omega_{e_{21}}\Omega_{e_{31}} + \Omega_{e_{32}} \\ \Omega_{e_{31}} & \Omega_{e_{21}}\Omega_{e_{31}} + \Omega_{e_{32}} & \Omega_{e_{31}}^2 + \Omega_{e_{32}}^2 + 1 \end{bmatrix}, \quad \text{(A1.2)}$$

$$\Omega_0 = \begin{bmatrix} 1 & \Omega_{0_{21}} & \Omega_{0_{31}} \\ \Omega_{0_{21}} & \Omega_{0_{21}}^2 + 1 & \Omega_{0_{21}}\Omega_{0_{31}} + \Omega_{0_{32}} \\ \Omega_{0_{31}} & \Omega_{0_{21}}\Omega_{0_{31}} + \Omega_{0_{32}} & \Omega_{0_{31}}^2 + \Omega_{0_{32}}^2 + 1 \end{bmatrix},$$

$$\Omega_a = \begin{bmatrix} 1 & \Omega_{a_{21}} & \Omega_{a_{31}} \\ \Omega_{a_{21}} & \Omega_{a_{21}}^2 + 1 & \Omega_{a_{21}}\Omega_{a_{31}} + \Omega_{a_{32}} \\ \Omega_{a_{31}} & \Omega_{a_{21}}\Omega_{a_{31}} + \Omega_{a_{32}} & \Omega_{a_{31}}^2 + \Omega_{a_{32}}^2 + 1 \end{bmatrix}.$$

This specification yields symmetric and positive semidefinite covariance matrices.[29] Furthermore, the elements of the Ω_{0a} 3×3 matrix capturing the covariance between a_i and Y_{i0}, are restricted to zeros due to parameter identification reasons.

[28] In the literature, several papers have analyzed likelihood based estimation of dynamic panel data models. See Balestra and Nerlove (1966), Nerlove (1971), Bhargava and Sargan (1983), and Nerlove and Balestra (1996) for dynamic panel data models with *random effects*. See Lancaster (2002), Hsiao et al (2002), Groen and Kleibergen (2003), Bun and Carree (2005), Kruiniger (2008), and Dhaene and Jochmans (2010) for dynamic panel data models with *fixed effects*.

[29] Blanchard and Quah (1989) and Gil-Alana and Moreno (2009) impose similar restrictions.

Let θ denote the vector of the model's parameters. Let $\tilde{\mathcal{Y}} = (\tilde{Y}_1, \ldots, \tilde{Y}_N | X_{i0}, \ldots, X_{iT})$ and let $\tilde{Y}_i = (\tilde{Y}_{i0}, \ldots, \tilde{Y}_{iT})$. The *random effects* QML estimator of the vector of parameters θ is obtained by maximizing the following log-likelihood function:

$$\ln \mathcal{L}(\tilde{\mathcal{Y}}; \theta) = \sum_{i=1}^{N} \ln f(\tilde{Y}_i | X_{i0}, \ldots, X_{iT}) = \sum_{i=1}^{N} -\frac{3(T+1)}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma_\eta| - \frac{1}{2} \omega_i' \Sigma_\eta^{-1} \omega_i, \qquad \text{(A1.3)}$$

where Σ_η is a $3(T+1) \times 3(T+1)$ matrix defined as:

$$\Sigma_\eta = \begin{bmatrix} \Omega_0 & \iota_T' \otimes \Omega_{0a}' \\ \iota_T \otimes \Omega_{0a} & I_T \otimes \Omega_e + \iota_T \iota_T' \otimes \Omega_a \end{bmatrix} \qquad \text{(A1.4)}$$

with ι_T being a $T \times 1$ vector of ones, I_T being a $T \times T$ identity matrix. Furthermore, each element of $\omega_i = (\omega_{i0}, \ldots, \omega_{iT}) = [(a_i + e_{i0}), \ldots, (a_i + e_{iT})]$ is computed as follows:

$$\begin{aligned} &\textit{Model 1}: \quad \omega_{it} = \tilde{Y}_{it} - \zeta \tilde{Y}_{it-1} \\ &\textit{Model 2}: \quad \omega_{it} = \tilde{Y}_{it} - \zeta \tilde{Y}_{it-1} - \zeta_{\text{IL}} \tilde{Y}_{it-1} D(i = \text{IL}) \\ &\textit{Model 3}: \quad \omega_{it} = \tilde{Y}_{it} - \zeta \tilde{Y}_{it-1} - D(i \in \text{IF}) \zeta_{\text{IL}} \tilde{Y}_{\text{IL}t-1} - D(i = \text{IL}) \left(\textstyle\sum_{k \in \text{IF}} \zeta_{\text{IF}} \tilde{Y}_{kt-1} \right). \end{aligned} \qquad \text{(A1.5)}$$

Appendix 2: Estimation of the patent count data model by EIS

The inference procedure of Blazsek and Escribano (2010) is applied for the Poisson model with latent variables. The model is estimated by the MSL method (Gouriéroux and Monfort, 1991), using the EIS technique of Richard and Zhang (2007). The EIS method has been applied for the precise 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). Recall from Equation (7) that the conditional density of $n_{it} | (N_{it-1}, L_{it-1}^*, Q_{it})$ is given by:

$$f_t(n_{it} | N_{it-1}, L_{it-1}^*, Q_{it}) = \frac{\exp(-\lambda_{it}) \lambda_{it}^{n_{it}}}{n_{it}!}. \qquad \text{(A2.1)}$$

Furthermore, the density of the dynamic latent variable l_{it}^* conditional on l_{it-1}^* is given by:

$$f_t^*(l_{it}^* | l_{it-1}^*) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{(l_{it}^* - \mu_i^* l_{it-1}^*)^2}{2} \right]. \qquad \text{(A2.2)}$$

If all latent variables ($l^*_{it} : t = 0, \ldots, T$) were observable then the joint likelihood of a realization ($n_{it}, l^*_{it} : t = 0, \ldots, T$) could be written as follows:

$$\prod_{t=0}^{T} f_t(n_{it}|N_{it-1}, L^*_{it-1}, Q_{it}) f^*_t(l^*_{it}|l^*_{it-1}) = \prod_{t=0}^{T} \frac{\exp(-\lambda_{it})\lambda_{it}^{n_{it}}}{n_{it}!} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(l^*_{it} - \mu^*_i l^*_{it-1})^2}{2}\right]. \quad \text{(A2.3)}$$

However, the L^*_{iT} are not observed. Therefore, we integrate out all latent variables from the likelihood function, with respect to the assumed normal distribution, to get the marginal density of patent counts. Since the number of $\{l^*_{it} : t = 0, \ldots, T\}$ is equal to the number of periods observed, the integrated likelihood function is the following $(T + 1)$-dimensional integral:

$$\mathcal{L} = \int_{\mathbb{R}^{T+1}} \prod_{t=0}^{T} \frac{\exp(-\lambda_{it})\lambda_{it}^{n_{it}}}{n_{it}!} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(l^*_{it} - \mu^*_i l^*_{it-1})^2}{2}\right] dL^*_{iT}. \quad \text{(A2.4)}$$

Rewrite the likelihood of patent counts as follows:

$$\mathcal{L}(N_{iT}, \theta) = \int_{\mathbb{R}^{T+1}} g(N_{iT}, L^*_{iT}|Q_{iT}, \theta) dL^*_{iT} = \int_{\mathbb{R}^{T+1}} \prod_{t=0}^{T} g_t(n_{it}, l^*_{it}|N_{it-1}, L^*_{it-1}, Q_{it}, \theta_t) dL^*_{iT}, \quad \text{(A2.5)}$$

where g is the joint density of (N_{iT}, L^*_{iT}) and θ denotes the vector of parameters of the model.

The major difficulty related to the statistical inference of the model is the precise evaluation of the $(T + 1)$-dimensional integral in $\mathcal{L}$ for given parameter values. This is performed numerically by Monte Carlo (MC) simulation method using the EIS technique (Richard and Zhang, 2007). The EIS procedure is nested into a typical likelihood function maximization procedure. In order to maintain the stability of that procedure, the same set of i.i.d. $N(0, 1)$ random numbers (i.e., *common random numbers*) is used for every set of parameters to estimate the value of the integrated likelihood function (see Richard and Zhang, 2007).

The EIS methodology consists of the following elements. First, an *auxiliary sampler*, m is introduced, which is included in the likelihood function. Then, it is factorized into the product of $(T + 1)$ sequential auxiliary densities, $\{m_t : t = 0, \ldots, T\}$ as follows:

$$\mathcal{L}(N_{iT}, \theta) = \int_{\mathbb{R}^{T+1}} \prod_{t=0}^{T} \frac{g_t(n_{it}, l^*_{it}|N_{it-1}, L^*_{it-1}, Q_{it}, \theta_t)}{m_t(l^*_{it}|L^*_{it-1}, \theta^*_t)} \times m_t(l^*_{it}|L^*_{it-1}, \theta^*_t) dL^*_{iT}, \quad \text{(A2.6)}$$

where θ^*_t denotes the parameters of the i-th auxiliary sampler. Then, the *importance MC estimate* of

$\mathcal{L}(N_{iT}, \theta)$ for given θ^* is:

$$\hat{L}_R(N_{iT}, \theta, \theta^*) = \frac{1}{R} \sum_{r=1}^{R} \prod_{t=0}^{T} \frac{g_t(n_{it}, l^*_{itr} | N_{it-1}, L^*_{it-1r}, Q_{it}, \theta_t)}{m_t(l^*_{itr} | L^*_{it-1r}, \theta^*_t)}, \tag{A2.7}$$

where θ^* is the vector of parameters of the auxiliary sampler defined as the *union* of all θ^*_ts and $\{l^*_{itr} : t = 0, \ldots, T\}$ denotes the r-th trajectory of i.i.d. draws from $\{m_t : t = 0, \ldots, T\}$ and $r = 1, \ldots, R$.

In the application of the EIS method, we use the answers of Richard and Zhang (2007) for the next two questions related to the sequential auxiliary densities: (a) How to choose the distribution for m_t in order to simulate l^*_{itr}? and (b) How to choose the θ^*_t parameters of m_t?

(a) **How to choose the distribution for m_t in order to simulate l^*_{itr}?** Richard and Zhang (2007) suggest defining the auxiliary sampler, m_t with its associated density kernel, k_t:

$$k_t(L^*_{it}, \theta^*_t) = m_t(l^*_{it} | L^*_{it-1}, \theta^*_t) \chi_t(L^*_{it-1}, \theta^*_t), \tag{A2.8}$$

where

$$\chi_t(L^*_{it-1}, \theta^*_t) = \int_{\mathbb{R}} k_t(L^*_{it}, \theta^*_t) dl^*_{it} \tag{A2.9}$$

denotes the t-th integrating constant associated to k_t. Richard and Zhang (2007) suggest choosing k_t as a kernel of the normal distribution. Moreover, we include f^*_t into the auxiliary sampler, m_t, see Bauwens and Hautsch (2006). Therefore, the t-th normal density kernel has the following form:

$$k_t(L^*_{it}, \theta^*_t) = \exp(\theta^*_{1t} l^*_{it} + \theta^*_{2t} (l^*_{it})^2) \times \exp\left[-\frac{(l^*_{it} - \mu^*_i l^*_{it-1})^2}{2} \right], \tag{A2.10}$$

where $\theta^*_t = (\theta^*_{1t}, \theta^*_{2t})$ determines the conditional mean and variance of the t-th auxiliary sampler, m_t. Bauwens and Hautsch (2006) show that the conditional mean, μ_t and conditional variance, π^2_t of the normal auxiliary sampler, m_t are given by:

$$\mu_t = \pi^2_t (\theta^*_{1t} + \mu^*_i l^*_{it-1}), \tag{A2.11}$$

$$\pi^2_t = \frac{1}{1 - 2\theta^*_{2t}}. \tag{A2.12}$$

Therefore, for given parameters of the auxiliary sampler a trajectory of $\{l^*_{it} : t = 0, \ldots, T\}$ can be generated from the following AR(1) process:

$$l^*_{it} = \pi_t^2 \theta^*_{1t} + \pi_t^2 \mu^*_i l^*_{it-1} + \pi_t \eta_{it}, \tag{A2.13}$$

where $\eta_{it} \sim N(0,1)$ are i.i.d. common random numbers.

(b) **How to choose the θ^*_t parameters of m_t?** The EIS methodology relies on the optimal choice of parameters of the auxiliary samplers in the sense that for given m, the variance of $\hat{L}_R(N_{iT}, \theta, \theta^*)$ is minimized, i.e.:

$$\theta^*(N_{iT}, \theta) = \arg\min_{\theta^*} Var[\hat{L}_R(N_{iT}, \theta, \theta^*)]. \tag{A2.14}$$

From Equation (A2.7), one can see that this variance is 'small' if the auxiliary sampler, m_{it} provides a 'good fit' to the g_t function. Expressing the auxiliary sampler by its associated density kernel and integrating constant from (A2.8), m_t may provide a 'good fit' to g_t if

$$\ln g_t(n_{it}, l^*_{it} | N_{it-1}, L^*_{it-1}, Q_{it}, \theta_t) + \ln \chi_t(L^*_{it-1}, \theta^*_{it}) \simeq \ln k_t(L^*_{it}, \theta^*_t). \tag{A2.15}$$

Richard and Zhang (2007) show that if the auxiliary samplers are normal distributions then the MC variance minimization problem stated in Equation (A2.14) can be reduced to a recursive sequence of $(T+1)$ Ordinary Least Squares (OLS) problems, each of the following form (see also Bauwens and Hautsch, 2006):

$$\ln g_t(n_{it}, l^*_{itr} | N_{it-1}, L^*_{it-1r}, Q_t, \theta) + \ln \chi_{t+1}(L^*_{itr}, \hat{\theta}^*_{t+1}) = \theta^*_{t0} + \theta^*_{t1} l^*_{itr} + \theta^*_{t2} (l^*_{itr})^2 + u_{tr} \tag{A2.16}$$

for $t = T, \ldots, 0$, $r = 1, \ldots, R$, $\chi_{T+1}(L^*_{iiT}, \hat{\theta}^*_{T+1}) = 1$ 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 as the $\hat{\theta}^*_{t+1}$ estimates are used to compute the integrating constant for the next, t-th OLS regression.[30] Thus, the regressions are run backwards, i.e. from T

[30] This is based on the permutation of the integrating constants in Equation (A2.7), see Richard and Zhang (2007) for more details.

to 0. 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 this paper, 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 $\{l^*_{itr}\}^T_{t=0}$ from the natural sampler, $N(\mu l^*_{it-1r}, 1)$.

Step 2: For each t (from T to 0), estimate the regression in (A2.16).

Step 3: Given the OLS estimates of θ^* obtained in Step 2, draw R trajectories $\{l^*_{itr}\}^T_{t=0}$ from the auxiliary samplers, $\{m_t\}^T_{t=0}$. Iterate Steps 2 and 3 five times (Richard and Zhang, 2007).

Step 4: The estimate of the likelihood function, $\hat{L}_R$ can be computed as follows. From (A2.8), express m as:

$$m_t(l^*_{it}|L^*_{it-1}, \theta^*_t) = \frac{k_t(L^*_{it}, \theta^*_t)}{\chi_t(L^*_{it-1}, \theta^*_t)}, \tag{A2.17}$$

From (A2.10), one may deduce that the t-th integrating constant is given by:

$$\chi_t(L^*_{it-1}, \theta^*_t) = \sqrt{2\pi\pi^2_t} \times \exp\left[-\frac{(\mu^*_i)^2(l^*_{it-1})^2}{2} + \frac{\mu^2_t}{2\pi^2_t}\right]. \tag{A2.18}$$

Compute the estimation of the likelihood function of formula (A2.7) using formulas (A2.10), (A2.17) and (A2.18).

Appendix 3: Estimation of the secret patent intensity component

One of the variables of each PVAR market value model is the log latent component of patent innovation intensity, $\ln \lambda^*_{it}$. The value of λ^*_{it} is estimated by computing the expectation of the latent patent innovations component conditional on the observable information set, i.e. $E[\lambda^*_{it}|N_{it-1}, Q_t]$. In order to obtain this estimate, all latent variables (l^*_t) are integrated out from the expectation and it can be computed similarly to Bauwens and Hautsch (2006, pp. 460) as follows:

$$E[\lambda^*_{it}|N_{it-1}, Q_t] = \frac{\int_{\mathbb{R}^t} \lambda^*_{it} f^*_t(l^*_t|N_{it-1}, L^*_{t-1}, Q_t) g(N_{it-1}, L^*_{t-1}|Q_{t-1}, \theta_t) dL^*_t}{\int_{\mathbb{R}^{t-1}} g(N_{it-1}, L^*_{t-1}|Q_{t-1}, \theta_t) dL^*_{t-1}}, \tag{A3.1}$$

where g is the density function of $(N_{it-1}, L^*_{t-1}|Q_{t-1}, \theta_t)$. Notice that $f^*_t(l^*_t|N_{it-1}, L^*_{t-1}, Q_t) = f^*_t(l^*_t|l^*_{t-1})$. The high-dimensional integrals in this ratio are approximated numerically by the EIS technique presented in Appendix 2. Finally, the estimates of $E[\lambda^*_{it}|N_{it-1}, Q_t]$ are included in the PVAR models as secret patent innovations variables.

Appendix 4: IRFs and dynamic multipliers of the PVAR models

In this appendix, orthogonalized IRFs and dynamic multipliers of Models 1, 2 and 3 are derived. For all PVAR models, the IRFs are orthogonalized by applying the $(\sqrt{\Omega_\epsilon})^{-1}$ matrix, which gives information on the contemporaneous relationships among the variables corresponding to the non-orthogonalized IRFs. Using the parameters of the PVAR specification of this paper, $(\sqrt{\Omega_\epsilon})^{-1}$ is given by

$$(\sqrt{\Omega_\epsilon})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\Omega_{e21} & 1 & 0 \\ \Omega_{e21}\Omega_{e32} - \Omega_{e31} & -\Omega_{e32} & 1 \end{pmatrix}. \tag{A4.1}$$

Model 1: Multiply Equation (4) by $(\sqrt{\Omega_e})^{-1}$ and express $\tilde{Y}_{it}$ as follows:

$$(\sqrt{\Omega_e})^{-1}\tilde{Y}_{it} = (\sqrt{\Omega_e})^{-1}a_i + (\sqrt{\Omega_e})^{-1}\zeta\tilde{Y}_{it-1} + \epsilon_{it}, \tag{A4.2}$$

$$\tilde{Y}_{it} = (I_3 - \zeta L)^{-1}\sqrt{\Omega_e}(\sqrt{\Omega_e})^{-1}a_i + (I_3 - \zeta L)^{-1}\sqrt{\Omega_e}\epsilon_{it}, \tag{A4.3}$$

where $\epsilon_{it} = (\sqrt{\Omega_e})^{-1}e_{it}$ are orthogonal (structural) error terms. Since $(I_3 - \zeta L)^{-1} = \sum_{j=0}^{\infty}\zeta^j L^j$, then

$$\tilde{Y}_{it} = \sum_{j=0}^{\infty}\zeta^j a_i + \sum_{j=0}^{\infty}\zeta^j\sqrt{\Omega_e}\epsilon_{it-j}, \tag{A4.4}$$

and the IRFs are given by:

$$\Theta_j = \zeta^j\sqrt{\Omega_e} \text{ for } j = 0, 1, 2, \ldots, \infty. \tag{A4.5}$$

Model 2: Multiply Equation (29) by $(\sqrt{\Omega_e})^{-1}$ and express $\tilde{Y}_{it}$ as follows:

$$(\sqrt{\Omega_e})^{-1}\tilde{Y}_{it} = (\sqrt{\Omega_e})^{-1}a_i + (\sqrt{\Omega_e})^{-1}\zeta\tilde{Y}_{it-1} + (\sqrt{\Omega_e})^{-1}\zeta_{\text{IL}}\tilde{Y}_{it-1}D(i = \text{IL}) + \epsilon_{it}, \tag{A4.6}$$

$$\tilde{Y}_{it} = \{I_3 - [\zeta + \zeta_{\text{IL}} D(i = \text{IL})]L\}^{-1} a_i + \{I_3 - [\zeta + \zeta_{\text{IL}} D(i = \text{IL})]L\}^{-1} (\sqrt{\Omega_e}) \epsilon_{it}, \tag{A4.7}$$

where $\epsilon_{it} = (\sqrt{\Omega_e})^{-1} e_{it}$ are orthogonal error terms.

Since $\{I_3 - [\zeta + \zeta_{\text{IL}} D(i = \text{IL})]L\}^{-1} = \sum_{j=0}^{\infty} [\zeta + \zeta_{\text{IL}} D(i = \text{IL})]^j L^j$, then

$$\tilde{Y}_{it} = \sum_{j=0}^{\infty} [\zeta + \zeta_{\text{IL}} D(i = \text{IL})]^j a_i + \sum_{j=0}^{\infty} [\zeta + \zeta_{\text{IL}} D(i = \text{IL})]^j (\sqrt{\Omega_e}) \epsilon_{it-j}, \tag{A4.8}$$

and the IRFs are given by:

$$\Theta_{ij}(\text{IL}) = (\zeta + \zeta_{\text{IL}})^j \sqrt{\Omega_e} \text{ for } i = \text{IL} \text{ and } j = 0, 1, 2, \ldots, \infty, \tag{A4.9}$$

$$\Theta_{ij}(\text{IF}) = \zeta^j \sqrt{\Omega_e} \text{ for } i \in \text{IF} \text{ and } j = 0, 1, 2, \ldots, \infty. \tag{A4.10}$$

Model 3: Multiply Equation (36) by $(\sqrt{\Omega_e})^{-1}$ and express $\tilde{Y}_{it}$ as follows:

$$(\sqrt{\Omega_e})^{-1} \tilde{Y}_{it} = (\sqrt{\Omega_e})^{-1} a_i + (\sqrt{\Omega_e})^{-1} \zeta \tilde{Y}_{it-1} + (\sqrt{\Omega_e})^{-1} \zeta_{\text{IL}} \tilde{Y}_{\text{IL}t-1} D(i \in \text{IF}) + \tag{A4.11}$$

$$+ (\sqrt{\Omega_e})^{-1} \left(\sum_{k \in \text{IF}} \zeta_{\text{IF}} \tilde{Y}_{kt-1} \right) D(i = \text{IL}) + \epsilon_{it},$$

$$\tilde{Y}_{it} = (I_3 - \zeta L)^{-1} a_i + (I_3 - \zeta L)^{-1} \zeta_{\text{IL}} \tilde{Y}_{\text{IL}t-1} D(i \in \text{IF}) + \tag{A4.12}$$

$$+ (I_3 - \zeta L)^{-1} \left(\sum_{k \in \text{IF}} \zeta_{\text{IF}} \tilde{Y}_{kt-1} \right) D(i = \text{IL}) + (I_3 - \zeta L)^{-1} (\sqrt{\Omega_e}) \epsilon_{it},$$

where $\epsilon_{it} = (\sqrt{\Omega_e})^{-1} e_{it}$ are orthogonal (structural) error terms. Since $(I_3 - \zeta L)^{-1} = \sum_{j=0}^{\infty} \zeta^j L^j$, then

$$\tilde{Y}_{it} = \sum_{j=0}^{\infty} \zeta^j a_i + \sum_{j=0}^{\infty} \zeta^j \zeta_{\text{IL}} \tilde{Y}_{\text{IL}t-1-j} D(i \in \text{IF}) + \tag{A4.13}$$

$$+ \sum_{j=0}^{\infty} \sum_{k \in \text{IF}} \zeta^j \zeta_{\text{IF}} \tilde{Y}_{kt-1-j} D(i = \text{IL}) + \sum_{j=0}^{\infty} \zeta^j (\sqrt{\Omega_e}) \epsilon_{it-j},$$

and the IRFs are given by:

$$\Theta_j = \zeta^j \sqrt{\Omega_e} \text{ for } j = 0, 1, 2, \ldots, \infty. \tag{A4.14}$$

Moreover, the dynamic interaction multipliers are given by:

$$\Gamma_j(\text{IL} \to \text{IF}) = (\text{Effects of } \tilde{Y}_{\text{IL}t-j} \text{ on } \tilde{Y}_{it} \text{ for } i \in \text{IF}) = \zeta^j \zeta_{\text{IL}} \text{ for } j = 1, 2, \ldots, \infty, \tag{A4.15}$$

$$\Gamma_j(\text{IF} \to \text{IL}) = (\text{Effects of } \tilde{Y}_{kt-j} \text{ for } k \in \text{IF on } \tilde{Y}_{\text{IL}t}) = \zeta^j \zeta_{\text{IF}} \text{ for } j = 1, 2, \ldots, \infty. \tag{A4.16}$$

Data sources

Center for Research on Stock Prices (CRSP) accessed from the Wharton Research Data Service (WRDS).

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 (available at: http://research.stlouisfed.org).

MicroPatents Co., US Utility Patent Database covering the period 1979 to 2004 and US patent citations for years 2003 and 2004.

National Bureau of Economic Research Patent Citations Data-File. CD-ROM included in: Jaffe, A. B., M. Trajtenberg (Eds.), Patents, Citations, and Innovations: A Window on the Knowledge Economy, MIT Press, 2002.

References

Adams, J. D., 1990, Fundamental stocks of knowledge and productivity growth. Journal of Political Economy, 98, 41, 673-702.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, P. Howitt, 2005, Competition and innovation: an inverted-U relationship. The Quarterly Journal of Economics, May 2005, 701-728.

Ahuja, G., 2000, Collaboration networks, structural holes and innovation: a longitudinal study. Administrative Science Quarterly, 45, 425-455.

Arora, A., M. Ceccagnoli, W. M. Cohen, 2008, R&D and the patent premium. International Journal of Industrial Organization, 26, 1153-1179.

Autant-Bernard, C., 2001, The geography of knowledge spillovers and technological proximity. Economics of Innovation and New Technology, 10, 237-254.

Balestra, P., M. Nerlove, 1966, Pooling cross-section and time series data in the estimation of a dynamic model: the demand for natural gas. Econometrica, 34, 3, 585-612.

Bauwens, L., N. Hautsch, 2006, Stochastic conditional intensity processes. Journal of Financial Econometrics, 4, 3, 450-493.

Benner, M., J. Waldfogel, 2008, Close to you? Bias and precision in patent-based measures of technological proximity. Research Policy, 37, 9, 1556-1567.

Bhargava, A., J. D. Sargan, 1983, Estimating dynamic random-effects models from panel data covering short time periods. Econometrica, 51, 6, 1635-1659.

Binder, M., C. Hsiao, M. H. Pesaran, 2005, Estimation and inference in short panel vector autoregressions with unit roots and cointegration. Econometric Theory, 21, 795-837.

Blanchard, O. J., D. Quah, 1989, The dynamic effects of aggregate demand and supply disturbances. The American Economic Review, 79, 655-673.

Blazsek, S., A. Escribano, 2010, Knowledge spillovers in U.S. patents: a dynamic patent intensity model with secret common innovation factors. Journal of Econometrics, 159, 1, 14-32.

Blundell, R., R. Griffith, J.V. Reenen, 1999, Market share, market value and innovation in a panel of British manufacturing firms. Review of Economic Studies, 66, 529-554.

Breschi, S., F. Lissoni, F. Malerba, 1998, Knowledge proximity and technological diversification. Mimeo, CESPRI.

Breschi, S., F. Lissoni, F. Malerba, 2003, Knowledge relatedness in firm technological diversification. Research Policy, 32, 1, 69-87.

Bun, M. J. G., M. A. Carree, 2005, Bias-corrected estimation in dynamic panel data models. Journal of Business and Economic Statistics, 23, 2, 200-210.

Cantner, U., A. Meder, 2007, Technological proximity and the choice of cooperation partner. Journal of Economic Interaction and Coordination, 2, 45-65.

Cao, B., Y. Sun, 2011, Asymptotic distributions of impulse response functions in short panel vector autoregressions. Journal of Econometrics, 163, 2, 127-143.

Caves, R. E., M. E. Porter, 1977, From entry barriers to mobility barriers: conjectural decisions and contrived deterrence to new competition. Quarterly Journal of Economics, 91, 241-262.

Caves, B. E., P. Ghemawat, 1992, Identifying mobility barriers. Strategic Management Journal, 13, 1-12.

Chambers, D., R. Jennings, R. B. Thompson, 2002, Excess returns to R&D-intensive firms. Review of Accounting Studies, 7, 133-158.

Chan, L., J. Lakonishok, T. Sougiannis, 2001, The stock market valuation of research and development expenditures. Journal of Finance, 56, 2431-2457.

Ciftci, M, B. Lev, S. Radhakrishnan, 2011, Is R&D mispriced or properly risk-adjusted? Journal of Accounting, Auditing and Finance, 26, 1, 81-116.

Cincera, M., 2005, Firms' productivity growth and R&D spillovers: an analysis of alternative technological proximity measures. CEPR Discussion Papers 4894.

Cohen, W. M., D. Levinthal, 1989, Innovation and learning: the two faces of R&D. Economic Journal, 99, 569-596.

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

Dhaene, G., K. Jochmans, 2010, Split-panel jackknife estimation of fixed-effect models, CORE discussion paper, 2010/3.

Escribano A., A. Fosfuri, J. A. Tribó, 2009, Managing external knowledge flows: the moderating role of absorptive capacity. Research Policy, 38, 96-105.

Fagerberg, J., B. Verspagen, C. Marjolein, 1996, Technology, growth and unemployment across European regions. Regional Studies, 31, 5, 457-466.

Fischer, M. M., T. Scherngell, E. Jansenberger, 2009, Geographic localisation of knowledge spillovers: evidence from high-tech patent citations in Europe. The Annals of Regional Science, 43, 839-858.

Frantzen, D., 2004, Technological diffusion and productivity convergence: a study for manufacturing in the OECD. Southern Economic Journal, 71, 2, 352-376.

Fung, M. K., 2003, Technological proximity and co-movements of stock returns. Economics Letters, 79, 131-136.

Fung, M. K., 2005, Are knowledge spillovers driving the convergence of productivity among firms? Economica, 72, 287-305.

Gil-Alana, L. A., A. Moreno, 2009, Technology shocks and hours worked: a fractional integration perspective. Macroeconomic Dynamics, 13, 5, 580-604.

Gilbert, R. J., D. M. G. Newbery, 1982, Preemptive patenting and the persistence of monopoly. The American Economic Review, 72, 514-526.

Gouriéroux, C., A. Monfort, 1991, Simulation based inference in models with heterogeneity. Annales d'Économie et de Statistique, 20/21, 69-107.

Griffith, R., S. Redding, H. Simpson, 2009, Technological catch-up and geographic proximity. Journal of Regional Science, 49, 4, 689-720.

Griliches, Z., 1979, Issues in assessing the contribution of research and development to productivity growth. Bell Journal of Economics, 10, 1, 92-116.

Griliches, Z., 1981, Market value, R&D, and patents. Economics Letters, 7, 183-187.

Griliches, Z., 1990, Patent statistics as economic indicators: a survey. Journal of Economic Literature, 28, 1661-1707.

Groen, J. J. J., F. Kleibergen, 2003, Likelihood-based cointegration analysis in panels of vector error-correction models. Journal of Business and Economic Statistics, 21, 2, 295-318.

Hall, B., 1993, The stock market's valuation of research and development investment during the 1980s. The American Economic Review, 83, 259-264.

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, Elsevier Science, Amsterdam, Netherlands.

Hall, B., A. B. Jaffe, M. Trajtenberg, 2001, The NBER patent citation data file: lessons, insights and methodological tools, NBER Working Paper No. 8498.

Hall, B., A. B. Jaffe, M. Trajtenberg, 2005, Market value and patent citations. The RAND Journal of Economics, 36, 1, 16-38.

Hall, B., D. Czarnitzki, R. Oriani, 2006, The market value of knowledge assets in U.S. and European firms. In: D. Bosworth and E. Webster (Eds.), The Management of Intellectual Property, Edward Elgar Publishing Limited, Cheltenham, UK.

Hall, B., R. Oriani, 2006, Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany and Italy. International Journal of Industrial Organization, 24, 971-993.

Hall, B., G. Thoma, S. Torrisi, 2007, The market value of patents and R&D: evidence from European firms, NBER Working Paper No. 13426.

Harhoff, D., 2000, R&D spillovers, technological proximity, and productivity growth – evidence from German panel data. Schmalenbach Business Review, 52, 238-260.

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.

Hsiao, C., M. H. Pesaran, A. K. Tahmiscioglu, 2002, Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. Journal of Econometrics, 109, 107-150.

Islam, R. 2011, The effects of human capital composition on technological convergence. Journal of Macroeconomics, 33, 3, 465-476.

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., 1989, Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers. Research Policy, 18, 87-97.

Jaffe, A. B., 1999, The US patent system in transition: policy innovation and innovation process. NBER working paper series 7280. National Bureau of Economic Research, Cambridge, MA.

Jain, N. C., A. Indrayan, L. R. Goel, 1986, Monte Carlo comparison of six hierarchical clustering methods on random data. Pattern Recognition, 19, 1, 95-99.

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

Kaiser, U., 2002, Measuring knowledge spillovers in manufacturing and services: an empirical assessment of alternative approaches. Research Policy, 31, 1, 125-144.

Kruiniger, H., 2008, Maximum likelihood estimation and inference methods for the covariance stationary panel AR(1)/unit root model. Journal of Econometrics, 144, 447-464.

Kuiper, F. K., L. Fisher, 1975, A Monte Carlo comparison of six clustering procedures. Biometrica, 31, 777-783.

Kumar, S., R. R. Russell, 2002, Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence. The American Economic Review, 92, 3, 527-548.

Lancaster, T., 2002, Orthogonal parameters and panel data. Review of Economic Studies, 69, 647-666.

Lanjouw, J. O. and M. Schankerman, 1999, The quality of ideas: measuring innovation with multiple indicators, NBER Working Paper No. 7345.

Lev, B., T. Sougiannis, 1996, The capitalization, amortization and value-relevance of R&D. Journal of Accounting and Economics, 21, 107-138.

Lev, B., B. Sarath, T. Sougiannis, 2005, R&D-related reporting biases and their consequences. Contemporary Accounting Research, 22, 4, 977-1026.

Lev, B., S. Radhakrishnan, M. Ciftci, 2006, The stock market valuation of R&D leaders, SSRN Working Paper Series.

Liesenfeld, R., J.-F. Richard, 2003, Univariate and multivariate stochastic volatility models: estimation and diagnostics. Journal of Empirical Finance, 10, 505-531.

Lychagin, S., J. Pinkse, M. E. Slade, J. Van Reenen, 2010. Spillovers in space: does geography matter? NBER Working Paper No. 16188.

Mansfield, E., M. Schwartz, S. Wagner, 1981, Imitation costs and patents: an empirical study. Economic Journal, 91, 907-918.

Messeni, A., 2008, Proximity and knowledge gatekeepers: the case of the Polytechnic University of Turin. Journal of Knowledge Management, 12, 5, 34-51.

Mohnen, P., 1996, R&D externalities and productivity growth. STI Review, 18, 39-66.

Mowery, D. C., J. E. Oxley, B. S. Silverman, 1996, Strategic alliances and inter-firm knowledge transfer. Strategic Management Journal, 17, 77-91.

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

Nerlove, M., 1971, Further evidence on the estimation of dynamic economic relations from a time series of cross-sections. Econometrica, 39, 2, 359-382.

Nerlove, M., P. Balestra, 1996, Formulation and estimation of econometric models for panel data. In: Mátyás, L., Sevestre, P. (Eds.), The Econometrics of Panel Data – A Handbook of of the Theory and Applications, 2nd Revised Edition, Kluwer Academic Publishers, Dordrecht, Netherlands.

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, University of Chicago Press, Chicago, USA.

Pakes, A., 1985, On patents, R&D, and the stock market rate of return. Journal of Political Economy, 93, 2, 390-409.

Porter M. E., 1979, How competitive forces shape strategy. Harvard Business Review, March-April, 137-145.

Porter M. E., 1980, Competitive Strategy, Free Press, New York, USA.

Porter M. E., 1985, Competitive Advantage, Free Press, New York, USA.

Reinganum, J. F., 1985, Innovation and industry evolution. Quarterly Journal of Economics, 100, 81-99.

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, Free Press, New York, USA.

Rosenkopf, L., P. Almeida, 2003, Overcoming local search through alliances and mobility. Management Science, 49, 6, 751-766.

Scherer, F. M., 1982, Inter-industry technology flows and productivity growth. The Review of Economics and Statistics, 64, 4, 627-634.

Shapiro, C., 2007, The design and use of patents, Economics for Management Lecture Series, IESE-Fundación BBVA Lecture.

Song, J., P. Almeida, G. Wu, 2003, Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? Management Science, 49, 4, 351-365.

Stuart, T. E., J. M. Podolny, 1996, Local search and the evolution of technological capabilities. Strategic Management Journal, 17, 21-38.

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

Ward, J. H., 1963, Hierarchical grouping to optimize an objective function. Journal of American Statistical Association, 58, 301, 236-244.

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.

Table 1. SIC based industry diversification in the technological cluster

SIC industry name	SIC	No. firms
Pharmaceutical preparations	2834	47
Biological products (no diagnostic substances)	2836	31
In vitro & in vivo diagnostic substances	2835	7
Perfumes, cosmetics & other toilet preparations	2844	3
Surgical & medical instruments & apparatus	3841	3
Medicinal chemicals & botanical products	2833	2
Wholesale-drugs, proprietaries & druggists' sundries	5122	2
Services-medical laboratories	8071	2
Grain mill products	2040	1
Beverages	2080	1
Chemicals & allied products	2800	1
Soap, detergents, cleaning preparations, perfumes, cosmetics	2840	1
Paints, varnishes, lacquers, enamels & allied prods	2851	1
Agricultural chemicals	2870	1
Plastics products, NEC	3089	1
Electromedical & electrotherapeutic apparatus	3845	1
Wholesale-medical, dental & hospital equipment & supplies	5047	1
Fire, marine & casualty insurance	6331	1
Services-hospitals	8060	1
Services-engineering, accounting, research, management	8700	1
Services-commercial physical & biological research	8731	1
Non-operating establishments	9995	1
Total number of firms		111

Notes: Standard Industry Classification (SIC).

Table 2. Hall and Mairesse (HM, 1996) based industry diversification in the technological cluster

HM industry name	No. firms
Pharmaceuticals	92
Non-manufacturing	10
Computers and inst.	4
Chemicals	2
Food	2
Rubber and plastics	1
Total number of firms	111

Table 3. Patent innovations leadership classification of firms based on mean values from 1979 to 2000

Firm name (SIC)	Cluster	(V1)	(V2)	(V3)	(V4)	(V5)	(V6)	(V7)	(V8)	(V9)
1. **Merck & Co.** (2834)	IL	217.6	1367.5	136.7	147232.4	12.39	8.47	13.20	0.82	1.16
2. Eli Lilly and Co. (2834)	IF	116.0	613.6	58.6	43645.5	12.21	8.15	12.69	0.83	1.17
3. Abbott Laboratories (2834)	IF	97.5	720.8	73.9	40954.3	11.89	7.91	12.67	0.80	1.13
4. Warner-Lambert (2834)	IF	81.7	656.2	61.3	31542.0	10.64	7.23	10.55	0.75	1.29
5. Pfizer, Inc. (2834)	IF	103.0	553.2	49.1	23373.0	12.21	8.32	12.79	0.81	1.16
6. Bristol-Myers Sq. (2834)	IF	69.7	307.4	34.2	11509.1	12.11	8.27	12.95	0.80	1.14
7. American Home (2834)	IF	52.8	330.7	30.7	8396.9	10.82	8.05	11.38	0.72	1.25
8. Alza Co. (2834)	IF	35.5	547.6	40.9	7683.0	8.24	5.28	9.92	0.78	1.01
9. Mallinckrodt, Inc. (2835)	IF	23.5	181.9	16.9	2007.8	9.25	6.92	9.82	0.68	1.11
10. Pharmacia & Upj. (2834)	IF	21.4	45.9	8.5	1922.6	10.96	7.45	10.34	0.79	1.34
11. Church & Dwight (2840)	IF	12.3	83.4	9.9	1537.8	7.96	5.10	9.74	0.66	0.89
12. NeoRx Co. (2835)	IF	6.5	68.4	7.8	500.5	7.25	4.12	7.92	1.07	0.96
13. Alliance Pharma. (2834)	IF	4.2	69.9	6.9	369.8	7.42	4.33	8.83	1.06	0.87
14. Xoma Co. (2836)	IF	6.9	48.7	5.0	329.6	8.11	4.21	8.84	1.15	0.96
15. Enzon, Inc. (2836)	IF	4.2	47.1	6.0	235.9	6.95	4.13	8.79	1.08	0.83
16. Guilford Pharma. (2834)	IF	3.5	23.0	4.9	216.5	6.75	4.20	6.98	0.98	1.01
17. Sugen, Inc. (2836)	IF	4.4	23.8	4.0	216.2	6.50	4.02	6.42	0.95	1.05
18. Inhale Therapeutic (2834)	IF	2.5	31.3	7.0	169.3	6.33	4.17	7.15	0.92	0.97
19. Corvas, Inc. (2836)	IF	3.5	16.3	2.1	122.5	6.58	4.07	7.00	1.00	1.00
20. Molecular Biosyst. (2835)	IF	2.0	57.5	4.8	90.5	6.97	4.23	7.97	1.02	0.93

Notes: Standard Industry Classification (SIC). Innovation Leader (IL). Innovation Follower (IF). The table presents nine variables for 20 out of the 111 firms of the technological cluster analyzed for the period 1979 to 2000. The mean, over the period 1979 to 2000, of the following variables is presented:

(V1) patent applications count: n_{it}
(V2) forward citations received count: c_{fit}
(V3) forward citations received count corrected for sample truncation bias (see Hall et al, 2001): $\tilde{c}_{fit}$
(V4) knowledge stock: $\sum_{s=0}^{t} \tilde{c}_{fis} n_{is} (1-\delta)^{t-s}$
(v5) log R&D expenses: r_{it}
(V6) log book value: z_{it}
(V7) log stock market value: m_{it}
(V8) log R&D expenses to log sales: r_{it}/s_{it}
(V9) log R&D expenses to log stock market value: r_{it}/m_{it}

Firms are ranked according to the mean of (V4). The companies not presented in the table from the technological cluster are assigned to the IF cluster. For seven out of the nine variables considered, Merck is the leader. Table 4 and Figure 2 show that this leadership is preserved for the whole period. We, therefore, call Merck the permanent IL. The rest of the firms from the technological cluster are the permanent IFs.

Table 4. Evolution of the knowledge stock for some firms of the technological cluster

Year	Merck & Co., Inc.	Eli Lilly and Co.	Abbott Laborat.	Warner-Lambert Co.	Pfizer, Inc.	Bristol-Myers Squibb	American Home Products	Alza Co.	Mallinckrodt Spec. Chem. Co., Inc.
1979	**31,215**	4,269	4,467	571	4,804	778	1,087	1,049	233
1980	**51,807**	13,245	6,246	1,411	6,131	1,237	2,269	2,805	331
1981	**65,132**	21,253	6,629	3,498	7,800	1,679	3,108	3,014	558
1982	**73,959**	21,034	7,164	4,159	8,658	1,739	3,135	3,681	621
1983	**79,546**	22,916	7,026	8,318	9,752	2,061	3,422	3,815	795
1984	**83,616**	25,138	6,928	16,406	11,298	2,521	3,433	4,806	925
1985	**89,006**	25,415	6,517	30,040	16,978	3,292	3,961	5,523	1,197
1986	**97,238**	23,248	7,754	43,104	18,579	3,694	5,771	6,628	1,358
1987	**110,944**	20,924	9,491	46,594	20,046	4,022	7,457	7,599	1,627
1988	**108,461**	19,228	12,499	50,251	24,377	4,988	7,516	9,036	1,789
1989	**115,519**	18,659	22,947	51,266	24,662	5,125	8,781	8,867	1,888
1990	**136,414**	18,617	28,476	52,747	29,798	6,483	8,892	9,472	2,513
1991	**168,611**	18,200	41,039	48,972	29,716	6,086	12,374	10,247	2,492
1992	**204,970**	20,146	50,468	44,339	31,188	8,039	13,667	11,131	3,480
1993	**201,721**	33,182	59,326	43,195	29,225	13,129	12,963	10,693	3,525
1994	**213,937**	46,093	70,367	41,515	31,009	15,414	14,425	10,754	3,321
1995	**224,626**	125,948	103,236	42,818	31,387	22,760	14,590	11,092	3,312
1996	**233,309**	112,243	100,738	37,540	29,102	26,825	13,133	9,837	3,078
1997	**246,212**	108,158	100,803	34,863	31,735	28,880	11,966	10,861	2,689
1998	**248,862**	98,847	91,705	33,734	30,075	30,646	10,735	10,660	2,298
1999	**235,728**	86,997	83,547	31,169	42,459	32,501	10,286	9,403	2,475
2000	**218,279**	76,439	73,621	27,414	45,426	31,301	11,762	8,054	3,667

Notes: The table presents the knowledge stock for nine firms presented in Table 3. These firms have the highest ranking in the technological cluster according to the mean knowledge stock. The table shows that the knowledge stock of Merck, indicated by bold numbers, was permanently higher than that of other firms in the technological cluster in every year.

Table 5. Cross-correlation matrices among innovation variables, firm size and market value

	$V1_t$	$V2_t$	$V3_t$	$V4_t$	$V5_t$	$V6_t$	$V7_t$	$V8_t$	$V9_t$
$V1_{t-1}$	0.89	0.72	0.81	**0.86**	0.60	0.70	0.48	-0.09	0.03
$V2_{t-1}$	0.81	0.90	0.88	**0.71**	0.58	0.64	0.40	-0.08	0.11
$V3_{t-1}$	0.88	0.84	0.89	**0.82**	0.60	0.68	0.46	-0.09	0.06
$V4_{t-1}$	**0.79**	**0.55**	**0.69**	**0.95**	**0.43**	**0.51**	**0.38**	**-0.06**	**-0.04**
$V5_{t-1}$	0.62	0.55	0.60	**0.46**	0.90	0.77	0.67	0.03	0.06
$V6_{t-1}$	0.72	0.59	0.66	**0.54**	0.73	0.91	0.62	-0.19	0.02
$V7_{t-1}$	0.48	0.33	0.42	**0.39**	0.60	0.61	0.85	-0.08	-0.47
$V8_{t-1}$	-0.09	-0.08	-0.08	**-0.06**	0.08	-0.18	-0.01	0.54	0.03
$V9_{t-1}$	0.05	0.18	0.11	**-0.01**	0.15	0.06	-0.41	0.08	0.73

	$V1_t$	$V2_t$	$V3_t$	$V4_t$	$V5_t$	$V6_t$	$V7_t$	$V8_t$	$V9_t$
$V1_t$	1.00								
$V2_t$	0.82	1.00							
$V3_t$	0.93	0.95	1.00						
$V4_t$	**0.88**	**0.64**	**0.78**	1.00					
$V5_t$	0.64	0.57	0.62	**0.47**	1.00				
$V6_t$	0.74	0.62	0.68	**0.55**	0.79	1.00			
$V7_t$	0.50	0.36	0.45	**0.41**	0.68	0.67	1.00		
$V8_t$	-0.09	-0.08	-0.09	**-0.06**	0.11	-0.20	-0.04	1.00	
$V9_t$	0.04	0.15	0.09	**-0.03**	0.15	0.01	-0.59	0.09	1.00

	$V1_t$	$V2_t$	$V3_t$	$V4_t$	$V5_t$	$V6_t$	$V7_t$	$V8_t$	$V9_t$
$V1_{t+1}$	0.89	0.81	0.88	**0.79**	0.62	0.72	0.48	-0.09	0.05
$V2_{t+1}$	0.71	0.90	0.84	**0.55**	0.55	0.59	0.33	-0.08	0.18
$V3_{t+1}$	0.81	0.88	0.89	**0.69**	0.60	0.66	0.42	-0.08	0.10
$V4_{t+1}$	**0.86**	**0.71**	**0.82**	**0.95**	**0.46**	**0.54**	**0.39**	**-0.06**	**-0.01**
$V5_{t+1}$	0.60	0.58	0.60	**0.43**	0.90	0.73	0.60	0.08	0.15
$V6_{t+1}$	0.70	0.64	0.68	**0.52**	0.77	0.91	0.61	-0.18	0.06
$V7_{t+1}$	0.48	0.40	0.46	**0.38**	0.68	0.62	0.85	-0.01	-0.41
$V8_{t+1}$	-0.09	-0.08	-0.09	**-0.06**	0.04	-0.19	-0.08	0.54	0.09
$V9_{t+1}$	0.02	0.11	0.06	**-0.04**	0.05	0.02	-0.47	0.03	0.73

Notes: The cross-correlation coefficients are computed for the 111 firms of the technological cluster, for the period 1979 to 2000, among the variables (V1)-(V9) presented in Table 3. The bold numbers correspond to the (V4) knowledge stock variable.

Table 6. Ranking of firms with respect to the predictive absorptive capacity and mean log market value

Ranking	Firm name	$1/\hat{\sigma}_i$	Ranking	Firm name	$(1/T)\sum_{t=1}^{T} m_{it}$
1	**Merck & Co., Inc.**	8.80	1	**Merck & Co., Inc.**	7.90
2	Cypress Bioscience, Inc.	8.42	2	Bristol-Myers Squibb Co.	7.65
3	Bristol-Myers Squibb Co.	6.26	3	Pfizer, Inc.	7.49
4	Pfizer, Inc.	5.91	4	Eli Lilly and Co.	7.39
5	Alza Co.	5.83	5	Abbott Laboratories	7.37
6	Block Drug Co., Inc.	5.80	6	Warner-Lambert Co.	5.25
7	Interferon Sciences, Inc.	4.49	7	Pharmacia & Upjohn AB	5.04
8	Warner-Lambert Co.	4.12	8	Alza Co.	4.62
9	Mallinckrodt Co., Inc.	4.07	9	Mallinckrodt Co., Inc.	4.52
10	Abbott Laboratories	3.69	10	Church & Dwight Co., Inc.	4.45
11	Vical, Inc.	3.30	11	Block Drug Co., Inc.	3.64
12	Cytogen Co.	3.07	12	Xoma Co.	3.54
13	Xoma Co.	2.92	13	Alliance Pharma. Co.	3.53
14	NeoRx Co.	2.78	14	Enzon, Inc.	3.50
15	Eli Lilly and Co.	2.77	15	Cytogen Co.	3.32
16	Pharmacia & Upjohn AB	2.67	16	MedImmune, Inc.	3.01
17	Vion Pharma., Inc.	2.65	17	Chattem, Inc.	2.98
18	Enzon, Inc.	2.01	18	Scios, Inc.	2.96
19	Biomatrix, Inc.	2.00	19	Cypress Bioscience, Inc.	2.70
20	NPS Pharma., Inc.	1.86	20	IDEC Pharma. Co.	2.69
21	Shaman Pharma., Inc.	1.81	21	NeoRx Co.	2.62
22	Valentis, Inc.	1.74	22	The Immune Response Co.	2.51
23	Alliance Pharma. Co.	1.72	23	Agouron Pharma., Inc.	2.41
24	Quigley Company, Inc.	1.68	24	Interferon Sciences, Inc.	2.37
25	Scios, Inc.	1.66	25	Biomatrix, Inc.	2.10
26	Zonagen, Inc.	1.65	26	Vical, Inc.	1.97
27	Celtrix Pharma., Inc.	1.63	27	Biospherics, Inc.	1.86
28	MedImmune, Inc.	1.56	28	Immulogic Pharma. Co.	1.84
29	Magainin Pharma., Inc.	1.51	29	Cytel, Inc.	1.82
30	The Immune Response Co.	1.50	30	Magainin Pharma., Inc.	1.82

Notes: For the firms presented in the table, σ_i is significant at the 5% level. For the firms excluded from the table, σ_i is not significant at the 5% level.

Table 7. Parameter estimates of Model 1: benchmark model

Mean equation

γ'	y_{it}	$\ln\lambda^o_{it}$	$\ln\lambda^*_{it}$
ω'	**-0.10**	**0.47**	**-0.10**
t	**0.02**	**0.46**	**0.18**
$\overline{y}_t$	**0.72**	**0.09**	**0.05**
z_{it}	**-0.15**	**0.02**	**-0.06**
$\frac{r_{it-1}}{m_{it-1}}$	**3.15**	**0.05**	**-0.15**

PVAR(1) effects ($M_\zeta = 0.93$)

ζ'	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
$\tilde{y}_{it-1}$	**0.04**	**0.01**	**0.03**
$\ln\tilde{\lambda}^o_{it-1}$	**0.19**	**0.22**	**0.21**
$\ln\tilde{\lambda}^*_{it-1}$	**0.70**	**0.70**	**0.70**

Covariance effects

Ω_e	Ω_{ey}	Ω_{eo}	Ω_{e*}
Ω_{ey}	**1.00**	**-0.20**	**0.36**
Ω_{eo}	**-0.20**	**1.04**	**0.37**
Ω_{e*}	**0.36**	**0.37**	**1.32**

Matrix of orthogonalization

$(\sqrt{\Omega_e})^{-1}$	Ω_{ey}	Ω_{eo}	Ω_{e*}
Ω_{ey}	**1.00**	0.00	0.00
Ω_{eo}	**0.20**	**1.00**	0.00
Ω_{e*}	**-0.45**	**-0.44**	**1.00**

Long-run impact matrix

$\Theta_j(1)'$	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
ϵ_{yit-j}	4.76	3.52	4.14
ϵ_{oit-j}	8.25	9.28	8.79
ϵ_{*it-j}	11.08	11.08	12.19

Notes: Panel Vector Autoregressive (PVAR). Bold numbers denote parameter significance at the 5 percent level. The M_ζ denotes the maximum modulus of the eigenvalues of ζ.

Model 1. Benchmark market value model:

$$Y_{it} = \gamma X_{it} + \tilde{Y}_{it} = \omega + \rho t + \beta\overline{y}_t + \psi z_{it} + \phi\frac{r_{it-1}}{m_{it-1}} + \tilde{Y}_{it}$$

$$\tilde{Y}_{it} = a_i + \zeta\tilde{Y}_{it-1} + e_{it}$$

The estimates of Ω_0 and Ω_a and are not reported in the table. The second equation is orthogonalized by multiplying each term by $(\sqrt{\Omega_e})^{-1}$, which gives the next expression:

$$(\sqrt{\Omega_e})^{-1}\begin{pmatrix}\tilde{y}_{it}\\ \ln\tilde{\lambda}^o_{it}\\ \ln\tilde{\lambda}^*_{it}\end{pmatrix} = (\sqrt{\Omega_e})^{-1}\begin{pmatrix}a_{yi}\\ a_{oi}\\ a_{*i}\end{pmatrix} + (\sqrt{\Omega_e})^{-1}\begin{pmatrix}\zeta_{11} & \zeta_{12} & \zeta_{13}\\ \zeta_{21} & \zeta_{22} & \zeta_{23}\\ \zeta_{31} & \zeta_{32} & \zeta_{33}\end{pmatrix}\begin{pmatrix}\tilde{y}_{it-1}\\ \ln\tilde{\lambda}^o_{it-1}\\ \ln\tilde{\lambda}^*_{it-1}\end{pmatrix} + \begin{pmatrix}\epsilon_{yit}\\ \epsilon_{oit}\\ \epsilon_{*it}\end{pmatrix}$$

where $\epsilon_{it} = (\epsilon_{yit}, \epsilon_{oit}, \epsilon_{*it})' = (\sqrt{\Omega_e})^{-1}e_{it}$ are orthogonal (structural) error terms. According to the estimates of $(\sqrt{\Omega_e})^{-1}$, the left hand side of the previous equation is

$$\begin{pmatrix}1.00 & 0.00 & 0.00\\ 0.20 & 1.00 & 0.00\\ -0.45 & -0.44 & 1.00\end{pmatrix}\begin{pmatrix}\tilde{y}_{it}\\ \ln\tilde{\lambda}^o_{it}\\ \ln\tilde{\lambda}^*_{it}\end{pmatrix} = \begin{pmatrix}\tilde{y}_{it}\\ 0.20\tilde{y}_{it} + \ln\tilde{\lambda}^o_{it}\\ -0.45\tilde{y}_{it} - 0.44\ln\tilde{\lambda}^o_{it} + \ln\tilde{\lambda}^*_{it}\end{pmatrix}$$

Therefore, the contemporaneous relationships are indicated by the first term of the right hand side of the following equation:

$$\begin{pmatrix}\tilde{y}_{it}\\ \ln\tilde{\lambda}^o_{it}\\ \ln\tilde{\lambda}^*_{it}\end{pmatrix} = \begin{pmatrix}0.00\\ -0.20\tilde{y}_{it}\\ 0.45\tilde{y}_{it} + 0.44\ln\tilde{\lambda}^o_{it}\end{pmatrix} + \begin{pmatrix}a_{yi}\\ 0.20a_{yi} + a_{oi}\\ -0.45a_{yi} - 0.44a_{oi} + a_{*i}\end{pmatrix} +$$

$$+\begin{pmatrix}0.04\tilde{y}_{it-1} + 0.19\ln\tilde{\lambda}^o_{it-1} + 0.70\ln\tilde{\lambda}^*_{it-1}\\ 0.02\tilde{y}_{it-1} + 0.26\ln\tilde{\lambda}^o_{it-1} + 0.84\ln\tilde{\lambda}^*_{it-1}\\ 0.01\tilde{y}_{it-1} + 0.03\ln\tilde{\lambda}^o_{it-1} + 0.08\ln\tilde{\lambda}^*_{it-1}\end{pmatrix} + \begin{pmatrix}\epsilon_{yit}\\ \epsilon_{oit}\\ \epsilon_{*it}\end{pmatrix}.$$

Table 8. Parameter estimates of Model 2: differential effects of the permanent IL

Mean equation (IF effects)

γ'	y_{it}	$\ln\lambda^o_{it}$	$\ln\lambda^*_{it}$
ω'	**-0.10**	**0.44**	**-0.13**
t	**0.02**	**0.46**	**0.18**
$\overline{y}_t$	**0.72**	**0.09**	**0.05**
z_{it}	**-0.12**	**0.05**	**-0.02**
$\frac{r_{it-1}}{m_{it-1}}$	**3.19**	0.03	**-0.14**

Mean equation (IL effects)

γ'_{IL}	y_{it}	$\ln\lambda^o_{it}$	$\ln\lambda^*_{it}$
ω'_{IL}	**0.54**	**3.60**	**1.65**
t	-0.01	-0.05	0.05
$\overline{y}_t$	0.00	**0.09**	**0.10**
z_{it}	**0.12**	0.01	**0.05**
$\frac{r_{it-1}}{m_{it-1}}$	**-0.59**	0.01	**-0.01**

PVAR(1) effects of IF ($M_\zeta = 0.93$)

ζ'	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
$\tilde{y}_{it-1}$	**0.04**	**0.01**	**0.02**
$\ln\tilde{\lambda}^o_{it-1}$	**0.20**	**0.22**	**0.21**
$\ln\tilde{\lambda}^*_{it-1}$	**0.70**	**0.70**	**0.70**

PVAR(1) effects of IL ($M_{\text{IL}} = 0.62$)

$(\zeta' + \zeta'_{\text{IL}})$	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
$\tilde{y}_{it-1}$	**0.23**	**0.06**	**0.10**
$\ln\tilde{\lambda}^o_{it-1}$	**0.21**	**0.07**	**0.11**
$\ln\tilde{\lambda}^*_{it-1}$	**0.67**	**0.23**	**0.34**

Covariance effects

Ω_e	Ω_{ey}	Ω_{eo}	Ω_{e*}
Ω_{ey}	**1.00**	**-0.36**	**0.35**
Ω_{eo}	**-0.36**	**1.13**	**0.31**
Ω_{e*}	**0.35**	**0.31**	**1.31**

Matrix of orthogonalization

$(\sqrt{\Omega_e})^{-1}$	Ω_{ey}	Ω_{eo}	Ω_{e*}
Ω_{ey}	**1.00**	0.00	0.00
Ω_{eo}	**0.36**	**1.00**	0.00
Ω_{e*}	**-0.51**	**-0.44**	**1.00**

Long-run impact matrix (IF)

$\Theta_{ij}(1)'$	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
ϵ_{yit-j}	3.73	2.30	3.03
ϵ_{oit-j}	7.47	8.42	7.84
ϵ_{*it-j}	10.12	10.01	11.01

Long-run impact matrix (IL)

$\Theta_{ij}(1)'$	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
ϵ_{yit-j}	2.00	-0.05	0.82
ϵ_{oit-j}	1.35	1.44	1.10
ϵ_{*it-j}	1.80	0.58	1.88

Notes: Panel Vector Autoregressive (PVAR). Innovation Leader (IL). Innovation Follower (IF). Bold numbers denote parameter significance at the 5 percent level. The M_ζ and M_{IL} denote the maximum modulus of the eigenvalues of ζ and $(\zeta + \zeta_{\text{IL}})$, respectively.

Model 2. Differential effects of the permanent IL:

$$Y_{it} = \gamma X_{it} + \gamma_{\text{IL}} X_{it} D(i = \text{IL}) + \tilde{Y}_{it} = \omega + \delta t + \beta \overline{y}_t + \psi z_{it} + \phi \frac{r_{it-1}}{m_{it-1}} + (\omega_{\text{IL}} + \delta_{\text{IL}} t + \beta_{\text{IL}} \overline{y}_t + \psi_{\text{IL}} z_{it} + \phi_{\text{IL}} \frac{r_{it-1}}{m_{it-1}}) D_i + \tilde{Y}_{it}$$

$$\tilde{Y}_{it} = a_i + \zeta \tilde{Y}_{it-1} + \zeta_{\text{IL}} \tilde{Y}_{it-1} D_i + e_{it}$$

where $D_i = 1$ if $i = \text{IL}$ and zero otherwise. The estimates of Ω_0 and Ω_a and are not reported in the table. The contemporaneous relationships are indicated by the first term of the right hand side of the following equation:

$$\begin{pmatrix} \tilde{y}_{it} \\ \ln\tilde{\lambda}^o_{it} \\ \ln\tilde{\lambda}^*_{it} \end{pmatrix} = \begin{pmatrix} 0.00 \\ -0.36\tilde{y}_{it} \\ 0.51\tilde{y}_{it} + 0.44\ln\tilde{\lambda}^o_{it} \end{pmatrix} + \begin{pmatrix} a_{yi} \\ 0.36a_{yi} + a_{oi} \\ -0.51a_{yi} - 0.44a_{oi} + a_{*i} \end{pmatrix} +$$

$$+ \begin{pmatrix} 0.04\tilde{y}_{it-1} + 0.20\ln\tilde{\lambda}^o_{it-1} + 0.70\ln\tilde{\lambda}^*_{it-1} \\ 0.02\tilde{y}_{it-1} + 0.29\ln\tilde{\lambda}^o_{it-1} + 0.95\ln\tilde{\lambda}^*_{it-1} \\ 0.01\ln\tilde{\lambda}^o_{it-1} + 0.03\ln\tilde{\lambda}^*_{it-1} \end{pmatrix} + \begin{pmatrix} 0.19\tilde{y}_{it-1} + 0.01\ln\tilde{\lambda}^o_{it-1} - 0.03\ln\tilde{\lambda}^*_{it-1} \\ 0.12\tilde{y}_{it-1} - 0.15\ln\tilde{\lambda}^o_{it-1} - 0.48\ln\tilde{\lambda}^*_{it-1} \\ -0.04\tilde{y}_{it-1} - 0.04\ln\tilde{\lambda}^o_{it-1} - 0.14\ln\tilde{\lambda}^*_{it-1} \end{pmatrix} D_i + \begin{pmatrix} \epsilon_{yit} \\ \epsilon_{oit} \\ \epsilon_{*it} \end{pmatrix}.$$

Table 9. Parameter estimates of Model 3: dynamic interaction between IL and IF

Mean equation

γ'	y_{it}	$\ln\lambda^o_{it}$	$\ln\lambda^*_{it}$
ω'	**-0.12**	**0.36**	**0.13**
t	**0.03**	**0.46**	**0.19**
$\overline{y}_t$	**0.71**	**0.10**	**0.05**
z_{it}	**-0.14**	**0.06**	**-0.04**
$\frac{r_{it-1}}{m_{it-1}}$	**3.20**	**0.03**	**-0.13**

PVAR(1) effects ($M_\zeta = 0.93$)

ζ'	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
$\tilde{y}_{it-1}$	**0.04**	0.01	**0.03**
$\ln\tilde{\lambda}^o_{it-1}$	**0.19**	**0.22**	**0.21**
$\ln\tilde{\lambda}^*_{it-1}$	**0.70**	**0.70**	**0.70**

Dynamic effects of IL on IFs

ζ'_{IL}	$\tilde{y}_{\text{IF}t}$	$\ln\tilde{\lambda}^o_{\text{IF}t}$	$\ln\tilde{\lambda}^*_{\text{IF}t}$
$\tilde{y}_{\text{IL}t-1}$	**-0.07**	**-0.10**	**0.04**
$\ln\tilde{\lambda}^o_{\text{IL}t-1}$	-0.02	**-0.03**	**0.02**
$\ln\tilde{\lambda}^*_{\text{IL}t-1}$	**0.12**	**0.19**	**-0.10**

Dynamic effects of IFs on IL

ζ'_{IF}	$\tilde{y}_{\text{IL}t}$	$\ln\tilde{\lambda}^o_{\text{IL}t}$	$\ln\tilde{\lambda}^*_{\text{IL}t}$
$\tilde{y}_{\text{IF}t-1}$	**-0.14**	**0.05**	**-0.02**
$\ln\tilde{\lambda}^o_{\text{IF}t-1}$	**0.06**	-0.01	0.01
$\ln\tilde{\lambda}^*_{\text{IF}t-1}$	**0.06**	**-0.02**	0.00

Covariance effects

Ω_e	Ω_{ey}	Ω_{eo}	Ω_{e*}
Ω_{ey}	**1.00**	**-0.39**	**0.36**
Ω_{eo}	**-0.39**	**1.15**	**0.30**
Ω_{e*}	**0.36**	**0.30**	**1.33**

Matrix of orthogonalization

$(\sqrt{\Omega_e})^{-1}$	Ω_{ey}	Ω_{eo}	Ω_{e*}
Ω_{ey}	**1.00**	0.00	0.00
Ω_{eo}	**0.39**	**1.00**	0.00
Ω_{e*}	**-0.53**	**-0.44**	**1.00**

Long-run impact matrix

$\Theta_j(1)'$	$\tilde{y}_{it}$	$\ln\tilde{\lambda}^o_{it}$	$\ln\tilde{\lambda}^*_{it}$
ϵ_{yit-j}	4.12	2.69	3.49
ϵ_{oit-j}	8.23	9.26	8.78
ϵ_{*it-j}	11.14	11.14	12.26

Notes: Panel Vector Autoregressive (PVAR). Innovation Leader (IL). Innovation Follower (IF). Bold numbers denote parameter significance at the 5 percent level. The M_ζ denotes the maximum modulus of the eigenvalues of ζ.

Model 3. Interactions between the innovation permanent IL and the IFs:

$$Y_{it} = \gamma X_{it} + \tilde{Y}_{it} = \omega + \delta t + \beta\overline{y}_{it} + \psi z_{it} + \phi\frac{r_{it-1}}{m_{it-1}} + \tilde{Y}_{it}$$

$$\tilde{Y}_{it} = a_i + \zeta\tilde{Y}_{it-1} + D(i \in \text{IF})\zeta_{\text{IL}}\tilde{Y}_{\text{IL}t-1} + D(i = \text{IL})\sum_{k\in\text{IF}}\zeta_{\text{IF}}\tilde{Y}_{kt-1} + e_{it}$$

The estimates of Ω_0 and Ω_a and are not reported in the table. The contemporaneous relationships are indicated by the first term of the right hand side of the following equation:

$$\begin{pmatrix} \tilde{y}_{it} \\ \ln\tilde{\lambda}^o_{it} \\ \ln\tilde{\lambda}^*_{it} \end{pmatrix} = \begin{pmatrix} 0.00 \\ -0.39\tilde{y}_{it} \\ 0.53\tilde{y}_{it} + 0.44\ln\tilde{\lambda}^o_{it} \end{pmatrix} + \begin{pmatrix} a_{yi} \\ 0.39a_{yi} + a_{oi} \\ -0.53a_{yi} - 0.44a_{oi} + a_{*i} \end{pmatrix} +$$

$$+ \begin{pmatrix} 0.04\tilde{y}_{it-1} + 0.19\ln\tilde{\lambda}^o_{it-1} + 0.70\ln\tilde{\lambda}^*_{it-1} \\ 0.03\tilde{y}_{it-1} + 0.29\ln\tilde{\lambda}^o_{it-1} + 0.97\ln\tilde{\lambda}^*_{it-1} \\ 0.01\ln\tilde{\lambda}^o_{it-1} + 0.02\ln\tilde{\lambda}^*_{it-1} \end{pmatrix} + D(i \in \text{IF}) \begin{pmatrix} -0.70\tilde{y}_{\text{IL}t-1} - 0.02\ln\tilde{\lambda}^o_{\text{IL}t-1} + 0.12\ln\tilde{\lambda}^*_{\text{IL}t-1} \\ -0.10\tilde{y}_{\text{IL}t-1} - 0.03\ln\tilde{\lambda}^o_{\text{IL}t-1} + 0.19\ln\tilde{\lambda}^*_{\text{IL}t-1} \\ 0.04\tilde{y}_{\text{IL}t-1} + 0.02\ln\tilde{\lambda}^o_{\text{IL}t-1} - 0.10\ln\tilde{\lambda}^*_{\text{IL}t-1} \end{pmatrix} +$$

$$+D(i = \text{IL})\sum_{k\in\text{IF}} \begin{pmatrix} -0.14\tilde{y}_{kt-1} + 0.06\ln\tilde{\lambda}^o_{kt-1} + 0.06\ln\tilde{\lambda}^*_{kt-1} \\ 0.05\tilde{y}_{kt-1} - 0.01\ln\tilde{\lambda}^o_{kt-1} - 0.02\ln\tilde{\lambda}^*_{kt-1} \\ -0.02\tilde{y}_{kt-1} + 0.01\ln\tilde{\lambda}^o_{kt-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{yit} \\ \epsilon_{oit} \\ \epsilon_{*it} \end{pmatrix}.$$

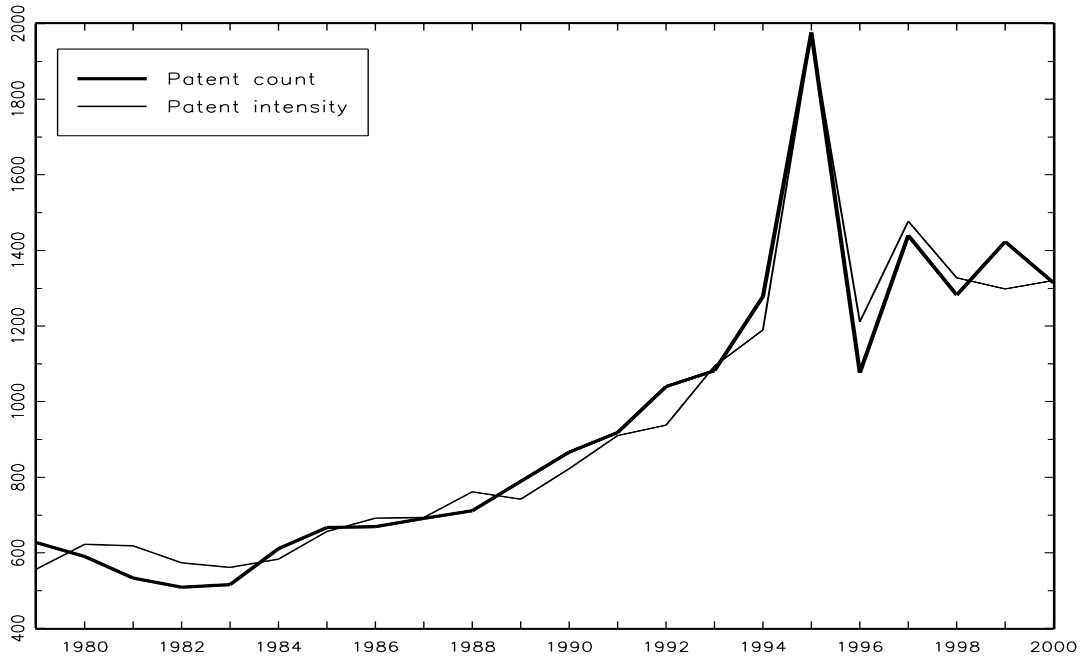

Figure 1. Evolution of total patent application counts and total patent application intensity estimates of firms in the technological cluster over the period 1979 to 2000.
Notes: The figure shows the evolution of $\sum_{i=1}^{N} n_{it}$ and $\sum_{i=1}^{N} \hat{\lambda}_{it}$.

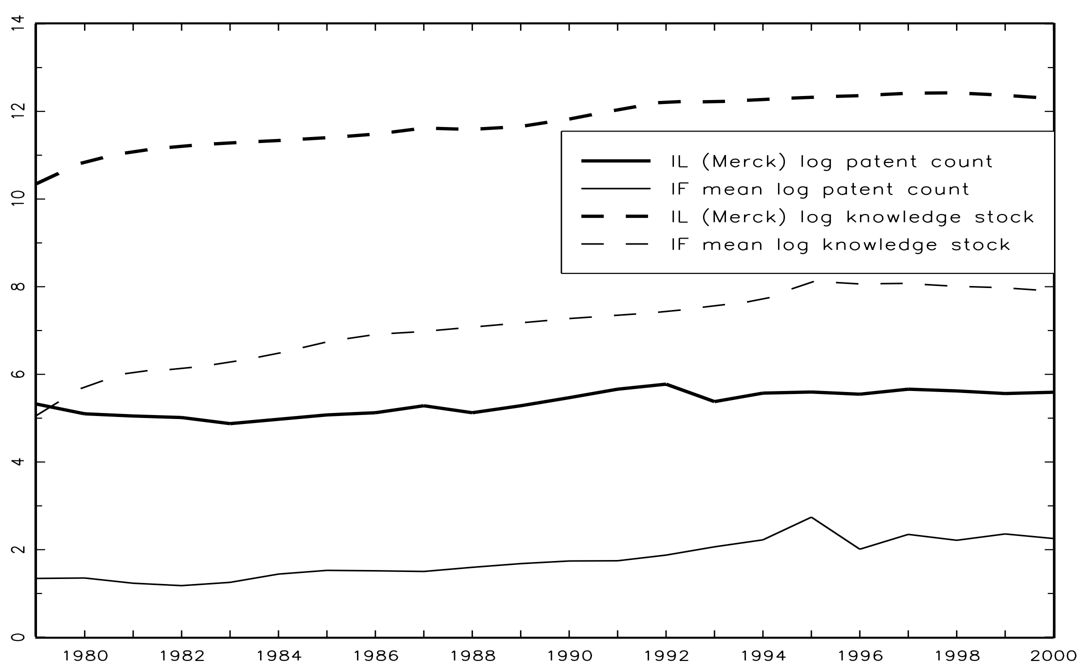

Figure 2. Evolution of the patent application counts and knowledge stock per firm for permanent IL (Merck) and permanent IFs over the period 1979 to 2000.

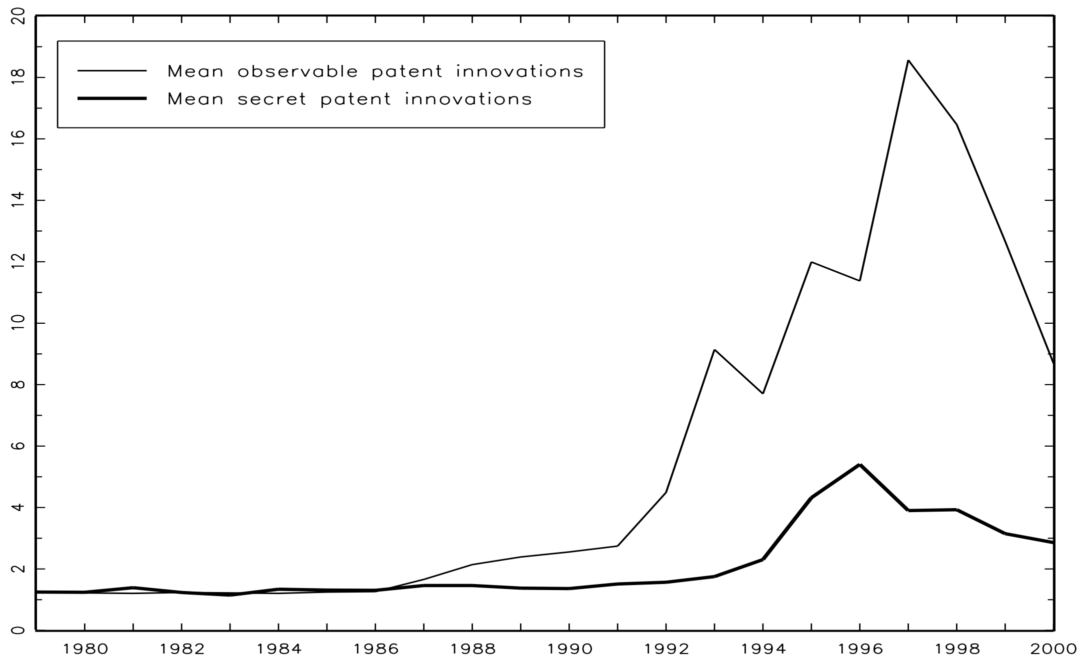

Figure 3. Evolution of mean observable and mean latent components of patent intensity of firms in the technological cluster over the period 1979 to 2000.
Notes: The figure shows the evolution of $(1/N)\sum_{i=1}^{N}\hat{\lambda}_{it}^{o}$ and $(1/N)\sum_{i=1}^{N}\hat{\lambda}_{it}^{*}$.

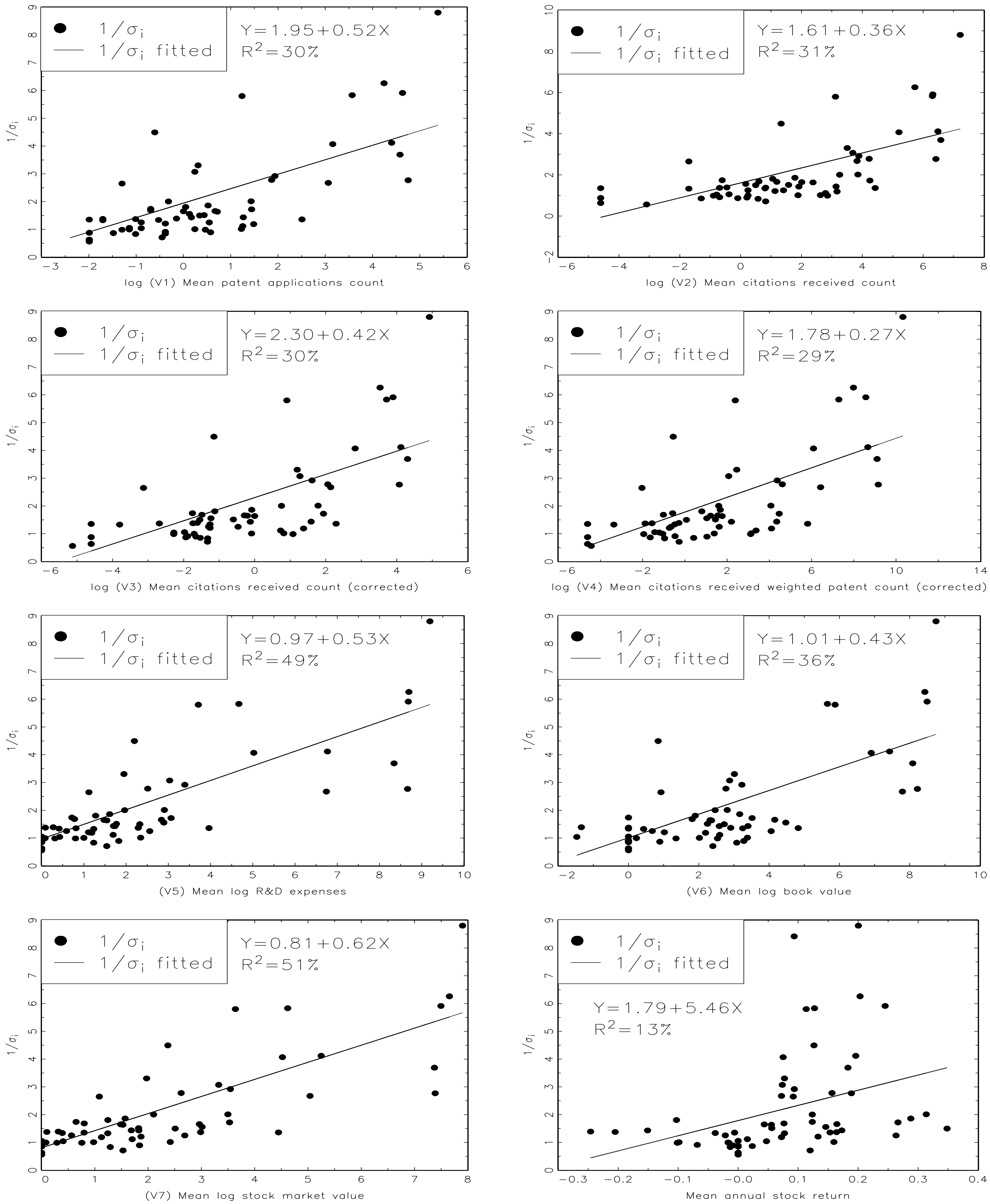

Figure 4. Estimates of predictive absorptive capacity, $1/\sigma_i$ as a function of different variables.

Notes: The figure shows the predictive absorptive capacity, $1/\sigma_i$ for 60 of the 111 firms in the technological cluster for the cases, where the σ_i parameter is significant at the 5% level. See the list of these firms in Table 5. The $1/\sigma_i$ is presented as a function of the variables (V1)-(V7) and the mean annual stock return. See the definitions of the variables (V1)-(V7) in the notes of Table 3.

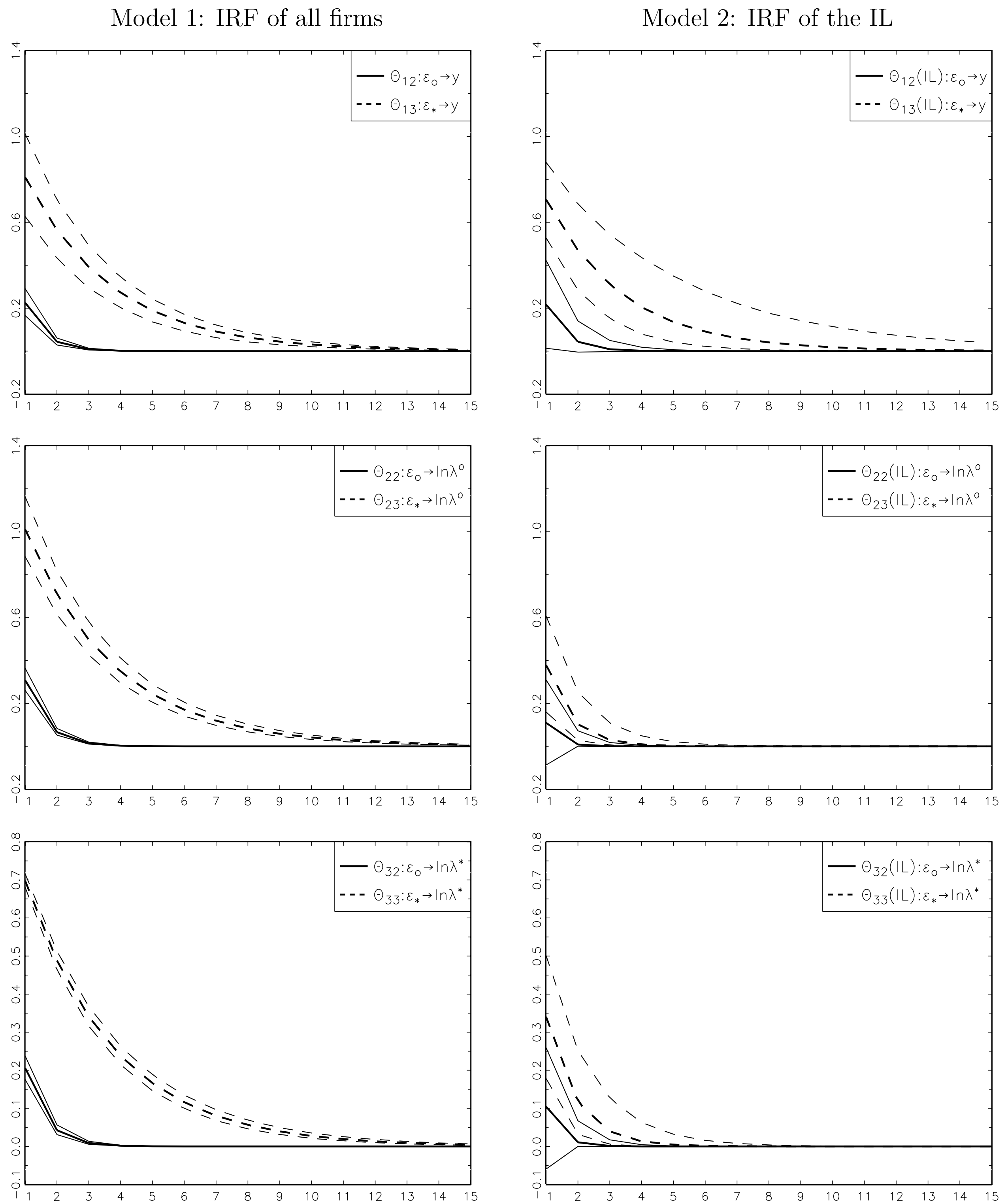

Figure 5. IRF of all firms and IRF of the IL firm in Models 1 and 2, respectively, for $j = 1, \ldots, 15$ leads.

Notes: Impulse Response Function (IRF). Innovation Leader (IL). Θ represents the IRF in Model 1. Θ(IL) represents the IRF of the IL firm in Model 2. The 95% confidence bands are also presented in the figure. $\epsilon_{it} = (\epsilon_{yit}, \epsilon_{oit}, \epsilon_{*it})' = (\sqrt{\Omega_e})^{-1} e_{it}$ are orthogonal (structural) error terms.

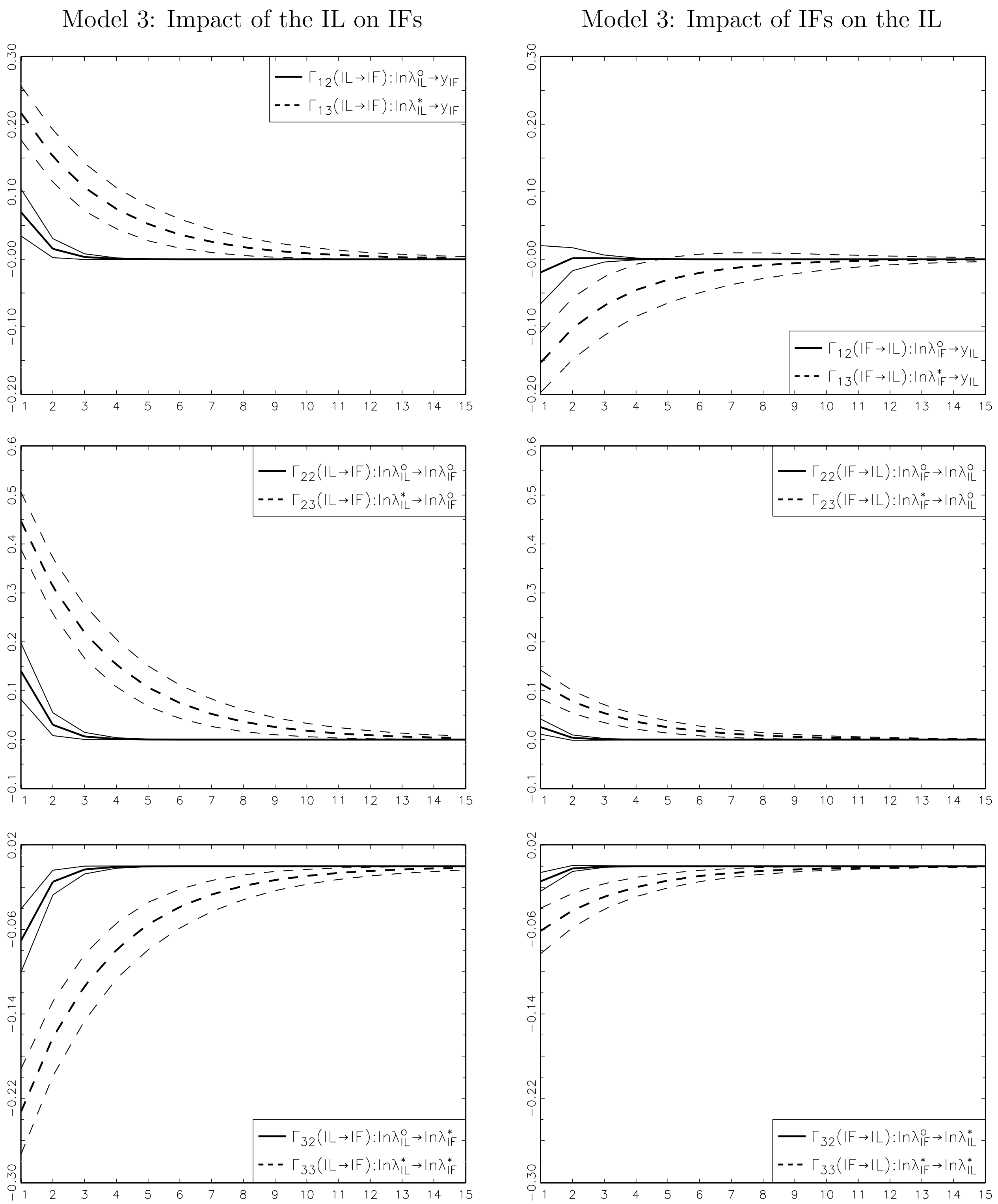

Figure 6. Dynamic multipliers of Model 3 for $j = 1, \ldots, 15$ leads.
Notes: Innovation Leader (IL). Innovation Follower (IF). $\Gamma(\text{IL} \rightarrow \text{IF})$ represents the dynamic interaction of the IL on IF firms. $\Gamma(\text{IF} \rightarrow \text{IL})$ represents the dynamic interactions of IF firms on the IL. The 95% confidence bands are also presented in the figure.