UNIVERSIDAD CARLOS III DE MADRID
working papers

UC3M Working papers
Economics
14-12
June, 2014
ISSN 2340-5031

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

Propensity to patent, R&D and market competition: Dynamic spillovers of innovation leaders and followers

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

[a]*School of Business, Universidad Francisco MarroquÍn, Guatemala*
[b]*Department of Economics, Universidad Carlos III de Madrid, Spain*

June 2014

Abstract

Dynamic interactions among stock return, Research and Development (R&D) expenses, patent applications based on R&D investment, and the propensity to patent are studied in this work for a panel of firms from the United States. The panel includes technologically similar firms, neck-to-neck, mostly from the drugs product-market sector. Firms' propensity to patent is modeled by a dynamic latent-factor patent count data model that separates patented and non patented R&D. Patent innovation leader and follower firms are identified according to their knowledge stock. Significant and positive dynamic spillover effects are obtained among patent application leaders and followers. We observe that neck-to-neck firms in patent innovation activity produce an inverted-U relationship between market competition and innovation. Furthermore, firms' propensity to patent is positively correlated with market competition and there is a positive feedback in both directions. Increasing the degree of competition in the market enhances innovation and patent applications, in order to help firms to appropriate part of the benefits of their R&D investments. On the other hand, firms by increasing their patent applications defend themselves from competitors, trying to improve their market share. However, due to the diffusion of knowledge through patent applications, knowledge spills over to competitors therefore, the degree of competition and innovation increases in the market.

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

Keywords: propensity to patent; competition; technological proximity; patent innovation leaders and followers; latent factor patent count data model; panel vector autoregression; simulated quasi maximum likelihood; efficient importance sampling

* 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@ufm.edu (S. Blazsek), alvaroe@eco.uc3m.es (A. Escribano).

1. Introduction

There are five usual main motives for firms to patent their inventions: protection from imitation, blocking competitors, technological image and reputation, exchange potential in co-operations, and internal firm Research & Development (R&D) performance indicator. Patents help in sustaining competitive advantages by increasing production cost of competitors, by signaling better quality of products, and by serving as barriers to entry.

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. The observed increase in R&D efficiency through the 1990s 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 increases in competitive pressure through time." (Shapiro, 2007, p. 5)

We use patent and firm-specific data of 4,476 US firms from several industries over the period 1979 to 2000. Firms are classified in technological groups according to technological proximity. We focus on a cluster of 111 firms that are mainly from the drugs product-market sector. We identify the patent innovation leaders and followers of the technological cluster according to their knowledge stock. The objective of this work is to learn about dynamic interactions (spillovers) between patent innovation leaders and followers, allowing for different propensity to patent for different firms.

We consider different dynamic measures of innovation activity that may capture patented R&D (i.e., publicly disclosed innovations) and non-patented R&D (i.e., not appropriated R&D or trade secrets). Patented and non-patented R&D are separated by using a latent-factor patent count data model. In this model, propensity to patent is driven by a common latent factor, representing the level of market competition. We call this factor the 'common competitive factor'. We study dynamic R&D spillovers among patent innovation leaders and followers by Panel Vector Autoregression (PVAR) models. The econometric models applied involve variables that are observed by firm managers who choose what proportion of the firm's R&D output to patent or keep secret, but the same variables are not included in the data set available for the econometrician.

As mentioned by Boldrin and Levine (2010): "We should protect not only the property rights of the innovators but also the rights of those who have legitimately obtained a copy of the idea, directly or indirectly, from the original innovator. The former encourages innovation; the later encourages the diffusion, adoption and improvements of innovations" (Boldrin and Levine, 2010, p. 8). In this paper we show that more competition increases the propensity to patent and the R&D investment to enhance firm's absorptive capacity, see Escribano et al. (2008), and therefore more innovation and more patents are produced. This two-way transmission mechanism has an extra dynamic multiplier effect through the spillovers between innovation leaders and followers.

In particular, we find positive dynamic spillover effects between patent innovation leaders and followers, indicating that firms in the technological cluster are neck-to-neck in innovation activity. We also find support for the well-known inverted-U relationship between market competition and innovation of Aghion et al. (2005). Increases in market competition conditions within the technological cluster are related with increases in the propensity to patent, and vice-versa (feed-back). This suggests that pharmaceutical firms reacted to the increasing level of market competition by patenting a significantly higher proportion of their innovation output after 1990, which at the same time increased the diffusion of knowledge among competitors enhancing therefore innovation.

Remaining part of this work is structured as follows. First, we review the existing literature in Section 2. Then, Section 3 presents the data set, technological clustering of firms, and definitions of patent innovation leaders and followers. Section 4 describes the econometric models and summarizes empirical findings. Finally, Section 5 concludes.

2. Literature review

2.1. Firm value and innovation activity

During recent 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.

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. Lev and Sougiannis (1996) estimate the inter-temporal relation between R&D capital and stock returns of public firms in the US, showing that R&D capital is associated with subsequent stock returns; see also Lev et al. (2005). Blundell et al. (1999) examine the relationship between surprise innovations and firm performance by using a dynamic panel count data specification and find a positive impact of innovation on market value of US firms. Chan et al. (2001) show a positive relationship between R&D capital 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. Their results show that in addition to patent counts, patent citations contain important information about stock market value.

2.2. Innovation leaders and followers

Technological improvements give innovator companies a competitive advantage. Nevertheless, the non-rival nature of knowledge may create a business-stealing effect among competitors as innovator's effort decreases the cost of competitor firms' subsequent innovations. Firms strategically decide to be R&D leaders or followers. Companies that introduce innovative products are R&D leaders, while other firms who mimic products of innovation leaders, are followers. 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. Results in the existing literature suggest that R&D leaders have sustained future profitability.

Porter (1979, 1980, 1985) investigates the relationship between firms' stock market value and R&D by recognizing that R&D activities are different among companies. Caves and Porter (1977) introduce a framework that explains intra-industry profit differentials based on precommitment 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 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, benefits generated by other firms' R&D efforts depend on technological differences among firms and the absorptive capacity of the imitator firm. Naturally, these factors create time lags in the adoption of technologies. For example, 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. Aghion et al. (2005) develop a model where competition discourages laggard firms from innovating but encourages neck-to-neck firms to innovate. Due to the effect of competition on equilibrium industry structure, their model generates an inverted-U shaped relationship between innovation and competition. They show that the average technological distance between innovation leaders and innovation followers increases with competition and the inverted-U is steeper when industries are more neck-to-neck. Lev et al. (2006) differentiate between R&D leaders and followers and compare 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. Ciftci et al. (2011) find that R&D leaders obtain substantial risk-adjusted returns during the first four to five future years. However, these excess returns converge to those of R&D followers afterwards.

3. Data

The data set includes 4,476 US firms from several manufacturing and service industries of the US economy over the period 1979 to 2000; $T = 22$ years. These firms published more than 500,000 patents during the sample period. We created the data set based on the recommendations of Hall et al. (2001).

Data have been collected from several sources. Patent data have been obtained from the National Bureau of Economic Research US Patent Citations Data File and MicroPatents Co. The patent database includes the US Patent and Trademark Office (USPTO) patent number, application date, publication date, USPTO patent number of cited patents, three-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, have been downloaded from the Wharton Research Data Service. Additional company specific information has been obtained from the Standard & Poor's Compustat data files. The firm data set includes book value of equity, stock market value, Standard Industry Classification (SIC) code, and R&D expenditure for all firms. Firm-specific accounting data are corrected for inflation by using the US consumer price index, collected from the US Department of Labor, Bureau of Labor Statistics. In the remaining part of this section, we describe the technological clustering procedure applied to US firms and present the definitions of patent innovation leaders and followers.

3.1. Technological proximity

We perform 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 measurement of knowledge spillovers and makes it difficult to identify competitors' effects on firms' 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, productivity-based and alternative measures. Mohnen (1996), Cincera (2005), and Benner and Waldfogel (2008) review the literature on technological proximity.

We use a patent-based proximity measure to classify firms in technologically similar groups. Technological clusters are formed as follows. To each firm, we assign technological categories using the technological classification suggested by Hall et al. (2001). These authors create 36 technological subcategories from the patent technological classification of USPTO that contains about 400 technological classes. We apply Ward's (1963) linkage clustering to perform technological clustering, motivated by Kuiper and Fisher (1975) and Jain et al. (1986). We use the angle distance measure to form technological clusters of firms, which is purely directional, therefore, it is not directly affected by the degree of concentration of the firm's research interests (Jaffe, 1986, p. 986). The technological clustering procedure creates a technology related grouping of 16 clusters of the 4,476 US companies.

We focus on a cluster of $N = 111$ companies. Table 1 shows the product-market industries of firms in the technological cluster according to two industry classifications. First, according to the SIC, the technological cluster includes 87 firms from the SIC283 drugs sector. Second, the modified SIC of Hall and Mairesse (1996) shows that 92 companies of the technological cluster are in the pharmaceutical sector. Nevertheless, Table 1 presents that the technological cluster includes companies from other product-market sectors as well. For example, it includes firms from the Grain mill products

(SIC2040), Beverages (SIC2080), Paints (SIC2851), Plastics products (SIC3089), and Electromedical and electrotherapeutic apparatus (SIC3845) industries.

Fig. 1a) shows the evolution of total patent application count and total patent application intensity, estimated for all firms in the technological cluster over the period 1979 to 2000; see Section 3 for patent application intensity estimation. The figure shows a significant growth of patent applications counts over the sample period. The level of patent applications per year was about 600 patents in 1979, which increased to about 1,300 patents in 2000.

[APPROXIMATE LOCATION OF TABLE 1; FIGURE 1]

3.2 Patent innovation leaders and followers

We define the permanent Innovation Leader (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. The evolution of firm i's knowledge stock is computed by $\sum_{s=0}^{t} \tilde{c}_{f,is}\tilde{P}_{is}(1-\delta)^{t-s}$ for $t = 1, \ldots, T$, where $\tilde{P}_{is}$ denotes the number of successful patent applications and $\tilde{c}_{f,it}$ is the number of citations received from subsequent patents (i.e., forward citations) corrected for sample truncation bias. We use $\tilde{c}_{f,it}$ to weight patent counts since Lanjouw and Schankerman (1999) and Hall et al. (2001) report that the number of forward patent citations is an appropriate measure of patent quality. Nevertheless, more recent patents in the end of the sample have less chance to receive citations from later patents than earlier patents, creating a sample truncation bias for the forward citations count. We correct for this bias by the fixed effects method suggested by Hall et al. (2001). Furthermore, motivated by Hall (1993) and Hall et al. (2005), we use the $\delta = 15\%$ annual depreciation rate to account for the decreasing value of past knowledge.

The firm with the highest knowledge stock in every year during the observation period is called the permanent IL of the technological cluster. We define the dummy variable $D_{it}(i = \text{IL})$ taking the value one if firm i is the permanent IL and zero otherwise. Other firms in the technological cluster are assigned to the permanent Innovation Follower (IF) group. We define the dummy variable $D_{it}(i \in \text{IF})$ taking the value one if firm i is in the permanent IF group and zero otherwise.

Table 2 shows the first 20 firms of the technological cluster, ranked according to (V4) mean knowledge stock over the period 1979 to 2000. The table shows the average of the following variables computed over the sample period: (V1) patent applications count, $\tilde{P}_{it}$; (V2) forward citations received count, $c_{f,it}$; (V3) forward citations received count corrected for sample truncation bias, $\tilde{c}_{f,it}$; (V4) knowledge stock, $\sum_{s=0}^{t} \tilde{c}_{f,is} \tilde{P}_{is}(1-\delta)^{t-s}$; (V5) log R&D expenses, $\tilde{r}_{it}$; (V6) log book value, z_{it}; (V7) log stock market value, m_{it}; (V8) log R&D expenses to log sales, $\tilde{r}_{it}/s_{it}$; (V9) log R&D expenses to log stock market value, $\tilde{r}_{it}/m_{it}$. For seven out of the nine variables considered, Merck & Co., Inc. (Merck, henceforth) is the leader.

Table 3 presents the evolution of knowledge stock for eight firms with the highest mean (V4) over the period 1979 to 2000. The table shows that the knowledge stock of Merck was permanently higher than that of other firms in every year. These results support our conclusion that Merck is the permanent IL of the technological cluster. In addition, Fig. 1b) shows the number of patent applications and knowledge stock for the IL (Merck) and the cross-sectional mean knowledge stock of IF 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 IF in every year from 1979 until 2000. The companies not presented in Table 3 from the technological cluster are assigned to the IF group.

In the R&D literature, different definitions of R&D leadership were also proposed. Lev et al. (2006) measure R&D intensity by two proxies: R&D expenditure to sales and R&D expenditure to market value. Furthermore, Chambers et al. (2002) and Ciftci et al. (2011) indicate R&D leadership by the R&D capital to sales ratio. We consider the variables (V8) R&D to sales and (V9) R&D to market value to check the robustness of the patent innovation leadership clustering procedure of our work with these authors. The results of the rankings obtained for (V8) and (V9) are not consistent with the clustering method of the present study, at least, due to the following reasons. First, 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). Second, the contemporaneous and dynamic cross-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.

In addition, we also classify firms according to their knowledge stock to Group of Leaders (GL) and Group of Followers (GF). We form these groups based on the (V4) mean knowledge stock variable, using Ward's (1963) clustering method. We define the dummy variable $D_{it}(i \in \text{GL})$ taking the value one if firm i is in the GL cluster and zero otherwise. Moreover, we define the dummy variable $D_{it}(i \in \text{GF})$

taking the value one if firm i is in the GF cluster and zero otherwise. The patent innovation leader group is formed by the following eight companies: Merck; Eli Lilly; Abbott Laboratories; Warner-Lambert; Pfizer; Bristol-Myers Squibb; American Home Products; Alza. See the mean knowledge stock and the evolution of knowledge stock of these firms in Tables 2 and 3, respectively.

[APPROXIMATE LOCATION OF TABLES 2-3]

4. Econometric models and empirical results

4.1. Benchmark innovation and market value model

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

The benchmark model of our empirical analysis is Pakes (1985), who studies the stock market valuation, log R&D expenditure, and log patent application count for a panel of 120 US firms over the period 1968 to 1975. Pakes (1985) formulates the following system of three equations to measure the dynamic and simultaneous interaction among stock return, q_{it}; log R&D expenses, r_{it}; log patent application count, $\ln P_{it}$:

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} \epsilon_{it} + \eta_{1it} \\ \sum_{\tau=0}^{\infty} c_{2\tau}\epsilon_{it-\tau} \\ \sum_{\tau=0}^{\infty} c_{3\tau}\epsilon_{it-\tau} + \sum_{\tau=0}^{\infty} b_{3\tau}\eta_{3it-\tau} \end{bmatrix} \tag{4.1}$$

where $\eta_{1it} \sim N(0, \sigma_1^2)$, $\epsilon_{it} \sim N(0, \sigma_2^2)$, and $\eta_{3it} \sim N(0, \sigma_3^2)$ are independent.

Equation (4.1) specifies contemporaneous and dynamic interactions among the endogenous variables according to a restricted Vector Moving Average, VMA(∞) representation. Similar to Pakes (1985, p. 396), in each equation of the system (4.1), we include time effects and a dummy variable controlling for zero patent application count. Furthermore, in the models of Sections 4.3 and 4.4, we also include in each equation the firm size measured by log book value of equity; extending Pakes (1985).

Pakes (1985) is consistent with the empirical results of Fama (1970) and LeRoy and Porter (1981), since he uses the no arbitrage condition to model the one-period stock return as the sum of the excess return and an uncorrelated error term. According to this condition, the noise term process does not allow to make excess returns on the market for those investors who use publicly available information and simple trading rules. An unexpected research-related event shifting the firm's value motivates managers to change the firm's R&D program, and R&D expenses are determined by the weighted sum of current and past excess stock returns, ϵ_{it}. Furthermore, patent applications are influenced by current and past excess returns, ϵ_{it}, and also by current and past values of an i.i.d. adjustment factor representing the propensity to patent, η_{3it} (see Scherer, 1965a, 1965b).

Pakes (1985) assumes that $\epsilon_{it} = \theta q_{it} + v_{it}$ so that v_{it} and q_{it} are orthogonal. Then, Equation (4.1) in VAR(∞) representation with contemporaneous relationships imposed can be written as

$$\begin{aligned} q_{it} &= \epsilon_{it} + \eta_{1it} \\ r_{it} &= c_{20}\theta q_{it} + \zeta_{22}(L) r_{it-1} + c_{20} v_{it} \\ \ln P_{it} &= \gamma_0 r_{it} + \zeta_{32}(L) r_{it-1} + \zeta_{33}(L) \ln P_{it-1} + \eta_{3it} \end{aligned} \tag{4.2}$$

To obtain a VAR(1) in structural form, we assume that the coefficients in Equation (4.2) satisfy that $\zeta_{m\tau} = \zeta_m^\tau$ for $m = 22, 32, 33$. Then, (4.2) can be written in VAR(1) reduced form as

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 \\ 0 & \gamma_0\zeta_{22} + \zeta_{32} & \zeta_{33} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ c_{20}\theta & 1 & 0 \\ c_{20}\theta\gamma_0 & \gamma_0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{it} + \eta_{1it} \\ c_{20} v_{it} \\ \eta_{3it} \end{bmatrix} \tag{4.3}$$

We can also write Equation (4.3) with a vector of standard normal i.i.d. error terms as follows:

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 \\ 0 & \gamma_0\zeta_{22} + \zeta_{32} & \zeta_{33} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_1 & 0 & 0 \\ \tilde{\sigma}_{12} & \tilde{\sigma}_2 & 0 \\ \gamma_0\tilde{\sigma}_{12} & \gamma_0\tilde{\sigma}_2 & \sigma_3 \end{bmatrix} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \end{bmatrix} \tag{4.4}$$

where $(e_{1it}, e_{2it}, e_{3it})' \sim N(0_{3\times1}, I_3)$. Furthermore, $\tilde{\sigma}_1$, $\tilde{\sigma}_{12}$, and $\tilde{\sigma}_2$ can be expressed by the parameters of Equation (4.3). See Appendix A where Equation (4.4) is derived from Equation (4.3). Equation (4.4) is the restricted VAR(1) formulation of Pakes (1985). The diagonal elements of the lower triangular matrix are positive. Therefore, the Cholesky decomposition of the covariance matrix of errors is unique and the covariance matrix of errors is positive definite.

Model 1.—We extend the benchmark model of Pakes (1985) to a restricted PVAR(1) model by considering fixed effects, a_i in Equation (4.4):

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} a_{q,i} \\ a_{r,i} \\ a_{P,i} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 \\ 0 & \gamma_0\zeta_{22} + \zeta_{32} & \zeta_{33} \end{bmatrix}}_{\zeta} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \tilde{\sigma}_1 & 0 & 0 \\ \tilde{\sigma}_{12} & \tilde{\sigma}_2 & 0 \\ \gamma_0\tilde{\sigma}_{12} & \gamma_0\tilde{\sigma}_2 & \sigma_3 \end{bmatrix}}_{\Omega} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \end{bmatrix} \quad (4.5)$$

for $i = 1, \ldots, 111$ firms and $t = 1979, \ldots, 2000$. The spectral radius of ζ, $\rho(\zeta)$ is less than one.

We can also formulate Model 1 in a compact matrix notation. The endogenous variables of the three-dimensional PVAR(1) model are $Y_{it} = (q_{it}, r_{it}, \ln P_{it})'$. In the PVAR equation, fixed effects are denoted by $a_i = (a_{q,i}, a_{r,i}, a_{P,i})'$ and error terms are summarized by $e_{it} = (e_{1it}, e_{2it}, e_{3it})'$. Then, Model 1 can be written as $Y_{it} = a_i + \zeta Y_{it-1} + \Omega e_{it}$. The Impulse Response Function (IRF) matrix, Θ_j is given by $\Theta_j = \zeta^j \Omega$; see Appendix B.

Model 2.—The unrestricted PVAR(1) model with fixed effects is given by

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} a_{q,i} \\ a_{r,i} \\ a_{P,i} \end{bmatrix} + \underbrace{\begin{bmatrix} \zeta_{11}^* & \zeta_{12}^* & \zeta_{13}^* \\ \zeta_{21}^* & \zeta_{22}^* & \zeta_{23}^* \\ \zeta_{31}^* & \zeta_{32}^* & \zeta_{33}^* \end{bmatrix}}_{\zeta^*} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \sigma_1^* & 0 & 0 \\ \sigma_{12}^* & \sigma_2^* & 0 \\ \sigma_{13}^* & \sigma_{23}^* & \sigma_3^* \end{bmatrix}}_{\Omega^*} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \end{bmatrix} \quad (4.6)$$

for $i = 1, \ldots, 111$ firms and $t = 1979, \ldots, 2000$, where $\sigma_1^* > 0$, $\sigma_2^* > 0$, $\sigma_3^* > 0$, and the spectral radius of ζ^*, $\rho(\zeta^*)$ is less than one. We can formulate Model 2 in a compact matrix notation as $Y_{it} = a_i + \zeta^* Y_{it-1} + \Omega^* e_{it}$. The IRF matrix, Θ_j is given by $\Theta_j = (\zeta^*)^j \Omega^*$; see Appendix B.

Estimation results.—Models 1 and 2 are estimated by the Quasi Maximum Likelihood (QML) method (see Hsiao et al. 2002). Table 4 presents the parameter estimates, QML standard errors, and model diagnostic tests. Fig. 2 presents the off-diagonal elements of the IRF matrix until 30 leads. We start with the model diagnostic results. The residual diagnostics part of Table 4 shows average p-values corresponding to: a) t test for $H_0 : E[e_{it}] = 0$ for $i = 1, \ldots, 111$; b) χ^2 test for $H_0 : \mathrm{Var}[e_{it}] = 1$ for $i = 1, \ldots, 111$; c) Ljung and Box (LB, 1978) test for $H_0 : \{e_{it} : t = 1, \ldots, T\}$ are uncorrelated for $i = 1, \ldots, 111$. The table shows that, on average, we are not able to reject H_0 at the 10% level of significance. According to Table 4, the spectral radius of ζ is less than one for both models, i.e., the PVAR(1) is covariance stationary. The likelihood based model selection metrics support Model 2, compared to the nested alternative. The Likelihood Ratio (LR) test shows that Model 2 is superior at any level of significance and the Akaike Information Criterion (AIC) also supports the more general model.

The parameter estimates of the ζ matrix (Model 1) evidence significant Granger causality (Hamilton, 1994) of log R&D expenses on log patent application count, since $\gamma_0\zeta_{22} + \zeta_{32}$ is significant. Furthermore, the IRF analysis of Model 1 presented in Fig. 2 suggests positive dynamic impact of stock return shocks on log R&D leads (Θ_{21}) and also positive impact of stock return shocks on log log patent application count leads (Θ_{31}). These results are similar to the findings of Pakes (1985, pp. 403-404). The estimates of ζ^* (Model 2) show significant Granger causality of log R&D on stock return (ζ^*_{12}), stock return on log R&D (ζ^*_{21}), and log R&D on log patent application count (ζ^*_{32}). The IRF analysis of Model 2 presented in Fig. 2 provides the following evidence. First, stock return shocks have positive effects both on log R&D (Θ_{21}) and log patent application count (Θ_{31}). These effects are more significant for log R&D which are observed contemporaneously and for all lags. For log patent counts, the positive effects start from the second leading year. Second, log R&D expenditure shocks have significant positive impact on both stock return (Θ_{12}) and log patent application count (Θ_{32}) for all leads.

The previous three-dimensional dynamic models are extended in the following sections. In Section 4.2., we propose a dynamic specification for the latent propensity to patent factor by using the latent-factor patent count panel data model of Blazsek and Escribano (2010). Then, in Sections 4.3 and 4.4, we extend Blazsek and Escribano (2012) and use four-dimensional dynamic models for stock return, q_{it}; log R&D expenses, r_{it}; log patent application count, $\ln P_{it}$ (i.e., log patented R&D); log non-patented R&D, $\ln P^{\times}_{it}$.

[APPROXIMATE LOCATION OF TABLE 4; FIGURE 2]

4.2. Patented and non-patented R&D investments

We model the conditional distribution of patent application count, $\tilde{P}_{it}$, by the Poisson distribution with patent intensity parameter, $\lambda_{it} = \tilde{P}^o_{it}\tilde{P}^*_{it}$ (see Hausman et al. 1984). In this model, $\tilde{P}^o_{it}$ is a function of total R&D investment and $\tilde{P}^*_{it} \in (0,1)$ captures propensity to patent. $\tilde{P}^*_{it}$ represents the percentage of the total R&D investment submitted to the patent office. $\tilde{P}^*_{it}$ adjusts total R&D investment for trade secrecy instead of revealing R&D information by patents (e.g., Kahn, 1962; Machlup, 1962) or the firm's absorptive capacity of rents from R&D (e.g., Scherer, 1965b; Arora et al., 2008). If $\tilde{P}^*_{it} \simeq 1$, then $\lambda_{it} \simeq \tilde{P}^o_{it}$ and the total R&D investment is submitted for patents. However, in general, $\lambda_{it} = \tilde{P}^o_{it}\tilde{P}^*_{it}$, i.e., only part of the total R&D investment is submitted to patents. The higher the $\tilde{P}^*_{it}$ is, the higher the firm's absorptive capacity is or the lower portion of innovation productivity is kept secret; see Escribano et al. (2009).

The conditional probability mass function of $\tilde{P}_{it}$ is

$$f(\tilde{P}_{it}|\mathcal{F}_t) = \frac{\exp(-\tilde{P}^o_{it}\tilde{P}^*_{it})(\tilde{P}^o_{it}\tilde{P}^*_{it})^{\tilde{P}_{it}}}{\tilde{P}_{it}!} \tag{4.7}$$

for $i = 1, \ldots, 111$ firms and $t = 1, \ldots, 22$; from 1979 to 2000. The conditional expectation of patent application count is given by the patent intensity parameter, $E(\tilde{P}_{it}|\mathcal{F}_t) = \lambda_{it}$. The conditioning set in the latent-factor count data model is

$$\mathcal{F}_t = \left[(\tilde{P}_{i1}, \tilde{r}_{i1}, c_{i1}, d_{i1}, l^*_1), \ldots, (\tilde{P}_{it-1}, \tilde{r}_{it-1}, c_{it-1}, d_{it-1}, l^*_{t-1}), (\tilde{r}_{it}, c_{it}, d_{it}, l^*_t) : i = 1, \ldots, N\right] \tag{4.8}$$

where $\tilde{r}_{it}$ denotes log R&D expenditure; c_{it} is the number of backward patent citations to other firms' past patents in the technological cluster (technologically related firms); d_{it} is the number of backward patent citations to other firms' past patents not in the technological cluster (not technologically related firms); l^*_t is a latent factor that drives propensity to patent in the technological cluster.

The log total R&D investment function is specified as

$$\begin{aligned} \ln \tilde{P}^o_{it} = \quad & \mu_0 + \gamma_1 t + \gamma_2 t\tilde{r}_{it} + \gamma_3 \tilde{r}^2_{it} + \gamma_4 z_{it} + \gamma_5 \tilde{P}_{i1} + \\ & \textstyle\sum_{k=0}^{q} \beta_k \tilde{r}_{it-k} + \sum_{k=0}^{q} \omega_k c_{it-k}\tilde{r}_{it} + \sum_{k=0}^{q} \phi_k d_{it-k}\tilde{r}_{it} + \sum_{k=1}^{p} \kappa_k \ln \tilde{P}^o_{it-k} \end{aligned} \tag{4.9}$$

where μ_0 is the constant parameter, γ_1 and γ_2 control for linear time trend; γ_3 captures non-linearities in log R&D expenditure; γ_4 measures the impact of firm size (log book value); γ_5 controls for initial conditions; β_k captures distributed lags of log R&D expenses; ω_k and ϕ_k control for the interaction of distributed lags of intra-industry and inter-industry backward citations, respectively, and current log R&D expenses; κ_k controls for AR dynamics. The same specification is used by Blazsek and Escribano (2010).

The propensity to patent component is specified according to the Probit model as $\tilde{P}^*_{it} = \Phi(\mu_i + \sigma_i l^*_t)$, where Φ is the cumulative distribution function (c.d.f.) of the standard normal distribution; μ_i is firm-specific fixed effect; $\sigma_i \in \mathbb{R}$ is a firm-specific scaling parameter for the latent factor since competition in the market affects firms differently; l^*_t is an underlying common factor for propensity to patent in the technological cluster. This common latent factor represents the level of market competition in the technological cluster. We call l^*_t the common competitive factor. The latent factor is specified as

$$l^*_t = \mu^* l^*_{t-1} + u_t \quad \text{with} \quad u_t \sim N(0,1) \text{ i.i.d.} \tag{4.10}$$

where $|\mu^*| < 1$ measures the average persistence of market competition in the technological cluster. The constant term in this equation is restricted to zero and the dynamic parameter μ^* is assumed to be the same for all firms since these are determined by common knowledge on market competition. Moreover, these restrictions also help parameter identification. We extend the specification of Pakes (1985) since the propensity to patent series for firm i, $\{\tilde{P}^*_{it} : t = 1, \ldots, T\}$ are serially correlated and are driven by the common competitive factor of the technological cluster. Nevertheless, $\tilde{P}^*_{it}$ is firm specific due to the level parameter μ_i and the scaling parameter σ_i.

We estimate the latent-factor patent count data model by the simulated QML method, applying the Efficient Importance Sampling (EIS) technique of Richard and Zhang (2007); see Appendix C. See applications of this method in Liesenfeld and Richard (2003), Bauwens and Hautsch (2006), and Blazsek and Escribano (2010).

The parameter estimates and QML standard errors are presented in Table 5. This table shows that the common competitive factor is covariance stationary with dynamic parameter $\mu^* = 0.91$. Fig. 1a) presents the evolution of total patent application counts, $\sum_{i=1}^{111} \tilde{P}_{it}$ and total patent application intensity, $\sum_{i=1}^{111} \lambda_{it}$ for the technological cluster over the period 1979 to 2000. The figure evidences that the latent-factor patent count data model fits well to the patent application count time series for most years. The only exceptions are 1995 and 1996 when there are significant differences between observed patent counts and fitted patent intensity estimates. These outliers, partly, may be due to the implementation of the General Agreement on Tariffs and Trade on 8 June 1995, which influenced effective pharmaceutical patent life (see Grabowski and Vernon, 2000).

We compute the filtered estimates of the common competitive factor by estimating $E[l_t^*|\mathcal{F}_t^o]$ for all t. In this expectation, we condition on the following observable information set:

$$\mathcal{F}_t^o = \Big[(\tilde{P}_{i1}, \tilde{r}_{i1}, c_{i1}, d_{i1}), \ldots, (\tilde{P}_{it-1}, \tilde{r}_{it-1}, c_{it-1}, d_{it-1}), (\tilde{r}_{it}, c_{it}, d_{it}) : i = 1, \ldots, N\Big] \tag{4.11}$$

The estimation of $E[l_t^*|\mathcal{F}_t^o]$ involves the proportion of two high dimensional integrals, each evaluated by the EIS technique; see Appendix D. Evolution of this factor is presented in Fig. 1c), showing $E[l_t^*|\mathcal{F}_t^o]$ over the period 1979 to 2000. Fig. 1c) shows that the common competitive factor decreases until 1990 and jumps to a higher level afterwards. Fig. 1d) presents the evolution of mean percentage of total R&D submitted to the patent office, $\tilde{P}_{it}^*$ over the period 1979 to 2000. The figure shows that during the 1990s, the percentage of patent applications increased from about 7.2% to above 10%. In the years 1996 and 1997, we see outliers for this variable and afterwards the level seems to stabilize above 10%.

Fig. 3 presents first and second order polynomial regression results about the determinants of propensity to patent in the technological cluster. The figure shows fitted values of $(1/22)\sum_{t=1}^{22} \ln \tilde{P}_{it}^*$ for $i = 1, \ldots, 111$ firms, regressed on the variables (V1), (V2), (V5), and (V7) over the period 1979 to 2000. The figure exhibits the estimates of the regression model fitted to mean $\ln \tilde{P}_{it}^*$ and the corresponding R-squared values to inform about the explanatory power of each variable. The first panel shows an inverted-U relationship for the log mean patent application count (V1) and mean log propensity to patent. For most firms in the technological cluster, a higher patent application count is associated with a higher propensity to patent level. Nevertheless, for some firms with high number of patent applications, we see that propensity to patent is relatively low. The second panel shows a linear increasing relation between log mean citations received count (V2) and mean log propensity to patent. Similar to the first panel, we can see that some firms with high number of citations received from subsequent patents, exhibit relatively low propensity to patent. The third panel shows an inverted-U relationship for mean log R&D expenses (V5) and mean log propensity to patent. For most firms in the technological cluster, higher R&D expenses are associated with higher propensity to patent.

Nevertheless, some firms that exhibit high R&D expenditure values have relatively low propensity to patent level. The last panel shows a positive linear relation between mean log stock market value and log propensity to patent. We can see that a high level of propensity to patent is significantly associated with high firm value.

We use the estimates of $\tilde{P}^o_{it}$ and $\tilde{P}^*_{it}$ in the extended PVAR models presented in the following section, to capture the simultaneous and dynamic effects among stock return, log R&D expenses, log patented R&D intensity, and log non-patented R&D intensity. Non-patented R&D intensity is approximated by the estimates of $\tilde{P}^o_{it}(1-\tilde{P}^*_{it})$.

[APPROXIMATE LOCATION OF TABLE 5; FIGURE 3]

4.3. Extended innovation and market value model

We extend the three-dimensional innovation and market values Models 1 and 2, by proposing four-dimensional dynamic models for stock return, q_{it}; log R&D expenses, r_{it}; log patent application count, $\ln P_{it}$ (i.e., patented R&D); log non-patented R&D, $\ln P^{\times}_{it}$.

We start with a Pakes (1985) like restricted VAR(∞) specification (structural form) of a system of four equations, extending Equation (4.2), as follows:

$$\begin{aligned}
q_{it} &= \epsilon_{it} + \eta_{1it} \\
r_{it} &= c_{20}\theta q_{it} + \zeta_{22}(L) r_{it-1} + c_{20} v_{it} \\
\ln P_{it} &= \gamma_0 r_{it} + \zeta_{32}(L) r_{it-1} + \zeta_{33}(L)\ln P_{it-1} + \zeta_{34}(L)\ln P^{\times}_{it-1} + \eta_{3it} \\
\ln P^{\times}_{it} &= \delta_0 r_{it} + \phi_0 \ln P_{it} + \zeta_{42}(L) r_{it-1} + \zeta_{43}(L)\ln P_{it-1} + \zeta_{44}(L)\ln P^{\times}_{it-1} + \eta_{4it}
\end{aligned} \tag{4.12}$$

where $\eta_{1it} \sim N(0,\sigma_1^2)$, $\epsilon_{it} \sim N(0,\sigma_2^2)$, $\eta_{3it} \sim N(0,\sigma_3^2)$, and $\eta_{4it} \sim N(0,\sigma_4^2)$ are independent. Suppose that $\epsilon_{it} = \theta q_{it} + v_{it}$ so that v_{it} and q_{it} are orthogonal. To obtain a VAR(1) in structural form, we assume that the coefficients in Equation (4.12) satisfy that $\zeta_{m\tau} = \zeta^{\tau}_m$ for $m = 22, 32, 33, 34, 42, 43, 44$. Then, (4.12) can be written in VAR(1) reduced form as

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \\ \ln P^{\times}_{it} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 & 0 \\ 0 & \gamma_0\zeta_{22} + \zeta_{32} & \zeta_{33} & \zeta_{34} \\ 0 & (\delta_0 + \gamma_0\phi_0)\zeta_{22} + \phi_0\zeta_{32} + \zeta_{42} & \phi_0\zeta_{33} + \zeta_{43} & \zeta_{44} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \\ \ln P^{\times}_{it-1} \end{bmatrix} +$$

$$
+\begin{bmatrix} 1 & 0 & 0 & 0 \\ c_{20}\theta & 1 & 0 & 0 \\ c_{20}\theta\gamma_0 & \gamma_0 & 1 & 0 \\ c_{20}\theta\delta_0 + c_{20}\theta\gamma_0\phi_0 & \delta_0+\gamma_0\phi_0 & \phi_0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{it}+\eta_{1it} \\ c_{20}v_{it} \\ \eta_{3it} \\ \eta_{4it} \end{bmatrix} \tag{4.13}
$$

We can also write (4.13) with a vector of standard normal i.i.d. error terms as follows:

$$
\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \\ \ln P_{it}^{\times} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 & 0 \\ 0 & \gamma_0\zeta_{22}+\zeta_{32} & \zeta_{33} & \zeta_{34} \\ 0 & (\delta_0+\gamma_0\phi_0)\zeta_{22}+\phi_0\zeta_{32}+\zeta_{42} & \phi_0\zeta_{33}+\zeta_{43} & \zeta_{44} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \\ \ln P_{it-1}^{\times} \end{bmatrix} +
$$

$$
+\begin{bmatrix} \tilde{\sigma}_1 & 0 & 0 & 0 \\ \tilde{\sigma}_{12} & \tilde{\sigma}_2 & 0 & 0 \\ \gamma_0\tilde{\sigma}_{12} & \gamma_0\tilde{\sigma}_2 & \sigma_3 & 0 \\ (\delta_0+\gamma_0\phi_0)\tilde{\sigma}_{12} & (\delta_0+\gamma_0\phi_0)\tilde{\sigma}_2 & \phi_0\sigma_3 & \sigma_4 \end{bmatrix} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \\ e_{4it} \end{bmatrix} \tag{4.14}
$$

where $(e_{1it}, e_{2it}, e_{3it}, e_{4it})' \sim N(0_{4\times 1}, I_4)$. Moreover, $\tilde{\sigma}_1$, $\tilde{\sigma}_{12}$, and $\tilde{\sigma}_2$ can be expressed by the parameters of Equation (4.13). See Appendix A where Equation (4.14) is derived from Equation (4.13). The diagonal elements of the lower triangular matrix are positive. Therefore, the Cholesky decomposition of the covariance matrix of errors is unique and the covariance matrix of errors is positive definite.

Model 3.—We extend the model of Equation (4.14) to a restricted PVAR(1) model by considering fixed effects, a_i as follows:

$$
\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \\ \ln P_{it}^{\times} \end{bmatrix} = \begin{bmatrix} a_{q,i} \\ a_{r,i} \\ a_{P,i} \\ a_{\times,i} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 & 0 \\ 0 & \gamma_0\zeta_{22}+\zeta_{32} & \zeta_{33} & \zeta_{34} \\ 0 & (\delta_0+\gamma_0\phi_0)\zeta_{22}+\phi_0\zeta_{32}+\zeta_{42} & \phi_0\zeta_{33}+\zeta_{43} & \zeta_{44} \end{bmatrix}}_{\zeta} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \\ \ln P_{it-1}^{\times} \end{bmatrix} +
$$

$$
+\underbrace{\begin{bmatrix} \tilde{\sigma}_1 & 0 & 0 & 0 \\ \tilde{\sigma}_{12} & \tilde{\sigma}_2 & 0 & 0 \\ \gamma_0\tilde{\sigma}_{12} & \gamma_0\tilde{\sigma}_2 & \sigma_3 & 0 \\ (\delta_0+\gamma_0\phi_0)\tilde{\sigma}_{12} & (\delta_0+\gamma_0\phi_0)\tilde{\sigma}_2 & \phi_0\sigma_3 & \sigma_4 \end{bmatrix}}_{\Omega} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \\ e_{4it} \end{bmatrix} \tag{4.15}
$$

for $i = 1, \ldots, 111$ firms and $t = 1979, \ldots, 2000$. The spectral radius of ζ, $\rho(\zeta)$ is less than one.

We can formulate Model 3 in a compact matrix notation. The endogenous variables of the four-dimensional PVAR(1) model are $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P_{it}^{\times})'$. In the PVAR equation, fixed effects are denoted by $a_i = (a_{q,i}, a_{r,i}, a_{P,i}, a_{\times,i})'$ and error terms are summarized by $e_{it} = (e_{1it}, e_{2it}, e_{3it}, e_{4it})'$. The Model 3 can be written as $Y_{it} = a_i + \zeta Y_{it-1} + \Omega e_{it}$. The IRF matrix, Θ_j is given by $\Theta_j = \zeta^j \Omega$; see Appendix B.

Model 4.—The unrestricted PVAR(1) model with fixed effects is given by

$$
\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \\ \ln P_{it}^{\times} \end{bmatrix} = \begin{bmatrix} a_{q,i} \\ a_{r,i} \\ a_{P,i} \\ a_{\times,i} \end{bmatrix} + \underbrace{\begin{bmatrix} \zeta_{11}^* & \zeta_{12}^* & \zeta_{13}^* & \zeta_{14}^* \\ \zeta_{21}^* & \zeta_{22}^* & \zeta_{23}^* & \zeta_{24}^* \\ \zeta_{31}^* & \zeta_{32}^* & \zeta_{33}^* & \zeta_{34}^* \\ \zeta_{41}^* & \zeta_{42}^* & \zeta_{43}^* & \zeta_{44}^* \end{bmatrix}}_{\zeta^*} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \\ \ln P_{it-1}^{\times} \end{bmatrix} + \underbrace{\begin{bmatrix} \sigma_1^* & 0 & 0 & 0 \\ \sigma_{12}^* & \sigma_2^* & 0 & 0 \\ \sigma_{13}^* & \sigma_{23}^* & \sigma_3^* & 0 \\ \sigma_{14}^* & \sigma_{24}^* & \sigma_{34}^* & \sigma_4^* \end{bmatrix}}_{\Omega^*} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \\ e_{4it} \end{bmatrix} \tag{4.16}
$$

for $i = 1, \ldots, 111$ firms and $t = 1979, \ldots, 2000$, where $\sigma_1^* > 0$, $\sigma_2^* > 0$, $\sigma_3^* > 0$, $\sigma_4^* > 0$, and the spectral radius of ζ^*, $\rho(\zeta^*)$ is less than one. We can formulate Model 4 in a compact matrix notation as $Y_{it} = a_i + \zeta^* Y_{it-1} + \Omega^* e_{it}$. The IRF matrix, Θ_j is given by $\Theta_j = (\zeta^*)^j \Omega^*$; see Appendix B.

Estimation results.—Models 3 and 4 are estimated by the QML method. Table 6 presents the parameter estimates, QML standard errors, and model diagnostic tests for both models. Figs 4 and 5 present the off-diagonal elements of the IRF matrix until 30 leads for Models 3 and 4, respectively. According to the residual diagnostic tests, on average, we are not able to reject model specification assumptions at the 10% level of significance. Both the LR test and the AIC metric suggest better performance for the more general Model 4. We find that both Models 3 and 4 are covariance stationary.

The IFR figures of Models 3 and 4 show the same effects among q_{it}, r_{it} and $\ln P_{it}$, as Models 1 and 2; see Figs 2, 4 and 5. Therefore, we focus on the IRFs involving non-patented R&D, $\ln P_{it}^{\times}$. Both Figs

4 and 5 show that stock return (Θ_{41}), log R&D expenses (Θ_{42}), and log patent count (Θ_{43}) shocks have positive impact on non-patented R&D. These positive effects can be seen more clearly in the IRFs of Model 4; for Model 3 they are not always significant. Furthermore, Fig. 5 (Model 4) shows that the effects of non-patented R&D shocks on other variables (Θ_{14}, Θ_{24}, and Θ_{34}) are not significant, although, the average IRF is positive.

The four-dimensional dynamic models presented are further extended in the following section. We consider different interaction effects for patent innovation leaders and followers in the technological cluster. This way, we measure the simultaneous and dynamic interaction among patent innovation leaders and followers of stock return, q_{it}; log R&D expenses, r_{it}; log patent application count, $\ln P_{it}$; log non-patented R&D, $\ln P_{it}^{\times}$.

[APPROXIMATE LOCATION OF TABLE 6; FIGURES 4-5]

4.4. Extended model for patent innovation leaders and followers

We extend Models 3 and 4 by considering different contemporaneous and dynamic effects for patent innovation leaders and followers in the technological cluster. Models 5 and 6 use the IL and IF, while Models 7 and 8 use the GL and GF clusters. Models 5 and 7 are the Pakes (1985) like restricted PVAR models; extending Model 3. Models 6 and 8 are unrestricted PVAR models; extending Model 4.

We formulate Models 6 to 8 in a compact matrix notation. The endogenous variables of the four-dimensional PVAR(1) model are $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P_{it}^{\times})'$. In the PVAR equation, fixed effects are denoted by $a_i = (a_{q,i}, a_{r,i}, a_{P,i}, a_{\times,i})'$ and error terms are summarized by $e_{it} = (e_{1it}, e_{2it}, e_{3it}, e_{4it})'$.

Model 5.—Model 5 is formulated as follows:

$$Y_{it} = a_i + \zeta Y_{it-1} + \zeta_{\mathrm{IL}} Y_{\mathrm{IL},t-1} D_{it}(i \in \mathrm{IF}) + \zeta_{\mathrm{IF}} \left(\sum_{k \in \mathrm{IF}} Y_{kt-1} \right) D_{it}(i = \mathrm{IL}) + \Omega e_{it} \tag{4.17}$$

The structure of ζ_{IL} and ζ_{IF} is the same as that of the restricted ζ matrix; see Model 3. The spectral radius of ζ is less than one. Ω is a lower triangular matrix with positive diagonal elements. Appendix B shows that the IRF matrix, Θ_j is $\Theta_j = \zeta^j \Omega$ and the matrices of dynamic interaction multipliers are

$$\Gamma_j(\mathrm{IL} \rightarrow \mathrm{IF}) = (\text{effects of } Y_{\mathrm{IL},t-j} \text{ on } Y_{it} \text{ for } i \in \mathrm{IF}) = \zeta^j \zeta_{\mathrm{IL}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{4.18}$$

$$\Gamma_j(\mathrm{IF} \rightarrow \mathrm{IL}) = (\text{effects of } Y_{k,t-j} \text{ on } Y_{\mathrm{IL},t} \text{ for } k \in \mathrm{IF}) = \zeta^j \zeta_{\mathrm{IF}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{4.19}$$

Model 6.—Model 6 is formulated as follows:

$$Y_{it} = a_i + \zeta^* Y_{it-1} + \zeta^*_{\mathrm{IL}} Y_{\mathrm{IL},t-1} D_{it}(i \in \mathrm{IF}) + \zeta^*_{\mathrm{IF}} \left(\sum_{k \in \mathrm{IF}} Y_{kt-1} \right) D_{it}(i = \mathrm{IL}) + \Omega^* e_{it} \tag{4.20}$$

where ζ^*_{IL} and ζ^*_{IF} are unrestricted as ζ^*; see Model 4. The spectral radius of ζ^* is less than one. Ω^* is a lower triangular matrix with positive diagonal elements. Appendix B shows that the IRF matrix, Θ_j is $\Theta_j = (\zeta^*)^j \Omega^*$ and the matrices of dynamic interaction multipliers are

$$\Gamma_j(\mathrm{IL} \to \mathrm{IF}) = (\text{effects of } Y_{\mathrm{IL},t-j} \text{ on } Y_{it} \text{ for } i \in \mathrm{IF}) = (\zeta^*)^j \zeta^*_{\mathrm{IL}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{4.21}$$

$$\Gamma_j(\mathrm{IF} \to \mathrm{IL}) = (\text{effects of } Y_{k,t-j} \text{ on } Y_{\mathrm{IL},t} \text{ for } k \in \mathrm{IF}) = (\zeta^*)^j \zeta^*_{\mathrm{IF}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{4.22}$$

Model 7.—Model 7 is formulated as follows:

$$Y_{it} = a_i + \zeta Y_{it-1} + \zeta_{\mathrm{GL}} Y_{\mathrm{GL},t-1} D_{it}(i \in \mathrm{GF}) + \zeta_{\mathrm{GF}} \left(\sum_{k \in \mathrm{GF}} Y_{kt-1} \right) D_{it}(i = \mathrm{GL}) + \Omega e_{it} \tag{4.23}$$

The structure of ζ_{GL} and ζ_{GF} is the same as that of the restricted ζ matrix; see Model 3. The spectral radius of ζ is less than one. Ω is a lower triangular matrix with positive diagonal elements. Appendix B shows that the IRF matrix, Θ_j is $\Theta_j = \zeta^j \Omega$ and the matrices of dynamic interaction multipliers are

$$\Gamma_j(\mathrm{GL} \to \mathrm{GF}) = (\text{effects of } Y_{\mathrm{GL},t-j} \text{ on } Y_{it} \text{ for } i \in \mathrm{GF}) = \zeta^j \zeta_{\mathrm{GL}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{4.24}$$

$$\Gamma_j(\mathrm{GF} \to \mathrm{GL}) = (\text{effects of } Y_{k,t-j} \text{ on } Y_{\mathrm{GL},t} \text{ for } k \in \mathrm{GF}) = \zeta^j \zeta_{\mathrm{GF}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{4.25}$$

Model 8.—Model 8 is formulated as follows:

$$Y_{it} = a_i + \zeta^* Y_{it-1} + \zeta^*_{\mathrm{GL}} Y_{\mathrm{GL},t-1} D_{it}(i \in \mathrm{GF}) + \zeta^*_{\mathrm{GF}} \left(\sum_{k \in \mathrm{GF}} Y_{kt-1} \right) D_{it}(i = \mathrm{GL}) + \Omega^* e_{it} \tag{4.26}$$

where ζ^*_{GL} and ζ^*_{GF} are unrestricted as ζ^*; see Model 4. The spectral radius of ζ^* is less than one. Ω^* is a lower triangular matrix with positive diagonal elements. Appendix B shows that the IRF matrix,

Θ_j is $\Theta_j = (\zeta^*)^j \Omega^*$ and the matrices of dynamic interaction multipliers are

$$\Gamma_j(\mathrm{GL} \to \mathrm{GF}) = (\text{effects of } Y_{\mathrm{GL},t-j} \text{ on } Y_{it} \text{ for } i \in \mathrm{GF}) = (\zeta^*)^j \zeta^*_{\mathrm{GL}} \text{ for } j = 0, 1, 2, \ldots, \infty \quad (4.27)$$

$$\Gamma_j(\mathrm{GF} \to \mathrm{GL}) = (\text{effects of } Y_{k,t-j} \text{ on } Y_{\mathrm{GL},t} \text{ for } k \in \mathrm{GF}) = (\zeta^*)^j \zeta^*_{\mathrm{GF}} \text{ for } j = 0, 1, 2, \ldots, \infty \quad (4.28)$$

Estimation results.—Models 5-8 are estimated by the QML method. Tables 7-10 present the parameter estimates, QML standard errors, and model diagnostic tests for Models 5-8, respectively. Figs 6, 9, 12, and 15 present the off-diagonal elements of the IRF matrix until 30 leads for Models 5-8, respectively. Figs 7-8, 10-11, 13-14, and 16-17 present the matrices of dynamic interaction multipliers until 30 leads for Models 5-8, respectively.

According to the residual diagnostic tests presented in Tables 7-10, on average, we are not able to reject model specification assumptions at the 10% level of significance. Both the LR test and the AIC metric suggest better performance for the unrestricted Models 6 and 8, compared to Models 5 and 7, respectively. We find that all models are covariance stationary. The IRFs of the extended models presented in Figs 6, 9, 12, and 15 show similar dynamic effects to the IRFs of Models 3 and 4. In the remaining part of this section, we focus on the dynamic effects among patent innovation leaders and followers for Models 6 and 8, supported by the likelihood-based model selection metrics.

Fig. 10 (Model 6) shows the dynamic interaction multipliers measuring spillovers from the IL (Merck) to IF companies, providing the following results. First, we see positive spillover effects of all IL variables on IF stock return, starting from the first and second leads; see Γ_{11}, Γ_{12}, Γ_{13}, and Γ_{14}. The highest spillover effects are associated with patented and non-patented R&D variables of the IL on the stock market valuation of IF firms. We can also see that IL R&D expenses have contemporaneous negative effect on IF stock return, however, this effect changes to positive from the first lead; see Γ_{12}. Second, we see significant positive dynamic effects of IL non-patented R&D on IF log R&D expenses (Γ_{24}) and IF non-patented R&D activity (Γ_{44}). For both variables, Fig. 10 shows a clear spillover pattern over several years. Third, Γ_{34} shows that IL non-patented R&D has negative contemporaneous effect on IF patent applications, however, from the first lead this effect changes to positive for all subsequent years. Fourth, we observe positive dynamic effects of IL log R&D expenses on IF non-patented R&D; see Γ_{42}. These findings suggest positive dynamic R&D spillovers from the IL (Merck) to IF firms for future year. For stock market and patent count we observe negative contemporaneous effects which change sign for all leading years. Moreover, the results also evidence that IF firms are more influenced by the non-patented R&D of Merck than by its patented R&D activity, emphasizing the importance of measuring non-patented R&D activity.

Fig. 11 (Model 6) exhibits the dynamic interaction multipliers capturing the spillovers from IF firms to Merck, providing the next findings. The results show evidence of positive dynamic effects over future years. First, we see positive effects of all IF variables on Merck stock return, starting from the first and second lead. The highest positive effects are observed from the non-patented R&D of followers; see Γ_{11}, Γ_{12}, Γ_{13}, and Γ_{14}. Second, we find that both patented R&D and non-patented R&D of Merck are influenced positively by all IF variables, starting from the third lead. These results indicate positive R&D spillovers from IF firms to the IL and also show the importance on non-patented R&D activity.

Furthermore, the dynamic multiplier estimates reported in Figs 16 and 17 (Model 8), which measure spillovers between GL and GF firms of the technological cluster, are similar to the dynamic multipliers of Model 6. This suggests that the spillover effects identified between Merck and its followers are robust for different clustering procedure of patent innovation leaders and followers.

Comparing the effects reported in Figs 10 and 11, we can see positive spillovers between IL and IF in both directions. We discuss these results in the context of the competition and innovation model of Aghion et al. (2005). We assess the level of competition by computing the following measure:

$$CO_t = 1 - \frac{1}{111}\sum_{i=1}^{111} LI_{it} = 1 - \frac{1}{111}\sum_{i=1}^{111} \frac{\text{operating profit}_{it} - \text{financial costs}_{it}}{\text{sales}_{it}} \tag{4.29}$$

where LI_{it} is the Lerner Index or price cost margin; see Nickell (1996) and Aghion et al. (2005). High values of this competition measure indicate competitive industry, while low values indicate market power. CO_t is an alternative observable measure of market competition to our latent common competitive factor, l_t^*. The framework of Aghion et al. (2005) provides the following discussion of our results.

First, Fig. 18a) presents that market competition, in general, increases in the technological cluster over the period 1979 to 2000. Aghion et al. (2005) conclude that increasing market competition discourages laggard firms to innovate while encourages neck-to-neck firms to innovate. The positive dynamic spillover effects estimated in both directions in the drugs industry indicate that, in the technological cluster analyzed, firms are neck-to-neck in innovation activity. The estimation results show that Merck has rapid contemporaneous impact on followers, and for the IFs it takes about three years to influence Merck. This finding is consistent with Nasbeth and Ray (1974), Mansfield et al. (1981), Rogers (1983), Pakes and Schankerman (1984), and Jovanovic and MacDonald (1994), who report that innovation spills over gradually to competitors.

Second, Figs 18b) to 18d) present the inverted-U relationship between competition, CO_t and three measures of innovation: total R&D investment, $\tilde{P}^o_{it}$; patented R&D, $\tilde{P}_{it}$; non-patented R&D, $\tilde{P}^o_{it}(1-\tilde{P}^*_{it})$. This result is very similar to Aghion et al. (2005). The figure shows that the maximum level of innovation is achieved at the $95\% - 97\%$ level of competition, which is equivalent to an average $3\% - 5\%$ price cost margin in the drugs industry.

Third, Fig 18e) presents the estimates of the common competitive factor, l^*_t and market competition, CO_t. The figure shows the least squares estimates of the linear regression model, suggesting positive relationship between the common competitive factor and the observable market competition metric. This provides support for the interpretation of l^*_t as level of market competition. The least squares estimates also suggest that the common competitive factor, l^*_t in the technological cluster is driven by the level of market competition.

Finally, Figs 1c) and 18a) present that the common competitive factor and the level of market competition jump simultaneously in 1990. This suggests that drugs firms reacted to the increasing level of market competition by patenting a significantly higher proportion of their innovation output after 1990. This finding supports Shapiro (2007).

[APPROXIMATE LOCATION OF TABLES 7-8; FIGURES 6-18]

5. Summary and conclusions

We study dynamic interactions between patent innovation leaders and patent innovation followers in a technological cluster, by allowing for different propensity to patent for different firms. We use patent and firm-specific data of 4,476 companies from several manufacturing and service 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 similar firms. We study a specific cluster of 111 firms that are mostly from the drugs product-market sector. In the technological cluster analyzed, the permanent IL, permanent IF, GL, and GF in patent innovation activity are identified.

We extend the approach of Pakes (1985) by considering different dynamic measures of innovation activity that may capture patented R&D (i.e., publicly disclosed innovations) and non-patented R&D (i.e., not appropriated R&D or trade secrets). Patented and non-patented R&D is separated by using the latent-factor patent count data model, which estimates different propensity to patent for each firm. In the patent count data model, propensity to patent is driven by a latent common competitive factor, representing the level of market competition. Given the estimates of patented and non-patented R&D, we study dynamic R&D spillovers among patent innovation leaders and followers by PVAR models.

The PVAR estimates support the findings of Pakes (1985) about dynamic effects among stock return, R&D expenses, and patent application counts of U.S. firms. The extended PVAR models evidence that non-patented R&D is an important dynamic determinant of both patented and non-patented R&D activity in the cluster of technologically similar firms. We find positive spillover effects between patent innovation leaders and followers of the technological cluster in both directions. We also find that the level of competition has increased over the period 1979 to 2000. The positive spillover effects indicate that firms are neck-to-neck in innovation activity in the drugs industry. We also evidence an inverted-U relationship between competition and innovation in the drugs technological cluster. According to the inverted-U relation, the maximum level of innovation is achieved at the $95\% - 97\%$ level of competition, which is equivalent to an average $3\% - 5\%$ price cost margin in the drugs industry. Finally, the results evidence that market competition is a possible driver of the common latent factor affecting firms' propensity to patent. This suggests that firms have patented a higher proportion of their innovation output due to the increasing level of competition in the technological cluster during the 1990s.

Acknowledgements

For helpful discussions and comments received we thank Luc Bauwens, Angeles Carnero, Andrea Fosfuri, Marco Giarratana, Bronwyn Hall, Uwe Hassler, Jerry Hausman, Javier Hualde, Miklós Koren, Gábor Kőrösi, and Martin Wagner. Previous versions of this study were presented at ETSERN Pamplona Meeting, June 2010; Summer Workshop in Economics of the Hungarian Academy of Sciences, MTA-KTI, Budapest, July 2010. Funding from the Bank of Spain Excellence Program and the Ministry of Economy and Competitiveness of Spain (research project ECO2012-33427) is gratefully acknowledged by the second author.

Appendix A

Deriving Equation (4.4)

We derive Equation (4.4) from Equation (4.3) as follows. Equation (4.3) is

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 \\ 0 & \gamma_0\zeta_{22} + \zeta_{32} & \zeta_{33} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ c_{20}\theta & 1 & 0 \\ c_{20}\theta\gamma_0 & \gamma_0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{it} + \eta_{1it} \\ c_{20}v_{it} \\ \eta_{3it} \end{bmatrix} \tag{A.1}$$

The first and second error terms are not orthogonal in this system since $\epsilon_{it} = \theta q_{it} + v_{it}$. First, we derive the covariance matrix of the errors in (A.1). The covariance between the first and second errors is

$$\mathrm{Cov}(\epsilon_{it} + \eta_{1it}, c_{20}v_{it}) = \mathrm{Cov}(\theta q_{it} + v_{it} + \eta_{1it}, c_{20}v_{it}) = c_{20}\mathrm{Var}(v_{it}) \tag{A.2}$$

where second equality uses that η_{1it}, q_{it}, and v_{it} are orthogonal. We express v_{it} as

$$v_{it} = \epsilon_{it} - \theta q_{it} = \epsilon_{it} - \theta(\epsilon_{it} + \eta_{1it}) = (1-\theta)\epsilon_{it} - \theta\eta_{1it} \tag{A.3}$$

Taking the variance of this equation, we have

$$\mathrm{Var}(v_{it}) = (1-\theta)^2\mathrm{Var}(\epsilon_{it}) + \theta^2\mathrm{Var}(\eta_{1it}) = \theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2 \tag{A.4}$$

Therefore,

$$\mathrm{Cov}(\epsilon_{it} + \eta_{1it}, c_{20}v_{it}) = c_{20}[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] \tag{A.5}$$

Then, the distribution of errors in (A.1) is

$$\begin{bmatrix} \epsilon_{it} + \eta_{1it} \\ c_{20}v_{it} \\ \eta_{3it} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 + \sigma_2^2 & c_{20}[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & 0 \\ c_{20}[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & c_{20}^2[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}\right] \tag{A.6}$$

We introduce the following notation:

$$\begin{pmatrix} \sigma_1^2+\sigma_2^2 & c_{20}[\theta^2\sigma_1^2+(1-\theta)^2\sigma_2^2] & 0 \\ c_{20}[\theta^2\sigma_1^2+(1-\theta)^2\sigma_2^2] & c_{20}^2[\theta^2\sigma_1^2+(1-\theta)^2\sigma_2^2] & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix} \equiv \begin{pmatrix} \dot{\sigma}_1^2 & \dot{\sigma}_{12} & 0 \\ \dot{\sigma}_{12} & \dot{\sigma}_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix} \tag{A.7}$$

The Cholesky matrix of the error covariance matrix is

$$\begin{pmatrix} \dot{\sigma}_1^2 & \dot{\sigma}_{12} & 0 \\ \dot{\sigma}_{12} & \dot{\sigma}_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}^{1/2} = \begin{pmatrix} \dot{\sigma}_1 & 0 & 0 \\ \frac{\dot{\sigma}_{12}}{\dot{\sigma}_1} & \sqrt{\dot{\sigma}_2^2-\frac{\dot{\sigma}_{12}^2}{\dot{\sigma}_1^2}} & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \tag{A.8}$$

Using this Cholesky matrix, we rewrite Equation (A.1) to make the error vector orthogonal:

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 \\ 0 & \gamma_0\zeta_{22}+\zeta_{32} & \zeta_{33} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \end{bmatrix} +$$

$$+ \begin{bmatrix} \dot{\sigma}_1 & 0 & 0 \\ c_{20}\theta\dot{\sigma}_1+\frac{\dot{\sigma}_{12}}{\dot{\sigma}_1} & \sqrt{\dot{\sigma}_2^2-\frac{\dot{\sigma}_{12}^2}{\dot{\sigma}_1^2}} & 0 \\ \gamma_0\left(c_{20}\theta\dot{\sigma}_1+\frac{\dot{\sigma}_{12}}{\dot{\sigma}_1}\right) & \gamma_0\sqrt{\dot{\sigma}_2^2-\frac{\dot{\sigma}_{12}^2}{\dot{\sigma}_1^2}} & \sigma_3 \end{bmatrix} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \end{bmatrix} \tag{A.9}$$

where $(e_{1it}, e_{2it}, e_{3it})' \sim N(0_{3\times1}, I_3)$. We simplify the notation in Equation (A.9) by defining new parameters:

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 \\ 0 & \gamma_0\zeta_{22}+\zeta_{32} & \zeta_{33} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \end{bmatrix} + \begin{bmatrix} \tilde{\sigma}_1 & 0 & 0 \\ \tilde{\sigma}_{12} & \tilde{\sigma}_2 & 0 \\ \gamma_0\tilde{\sigma}_{12} & \gamma_0\tilde{\sigma}_2 & \sigma_3 \end{bmatrix} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \end{bmatrix} \tag{A.10}$$

where $\tilde{\sigma}_1$, $\tilde{\sigma}_{12}$, and $\tilde{\sigma}_2$ can be expressed by the original parameters based on Equations (A.7) and (A.9).

Deriving Equation (4.14)

We derive Equation (4.14) from Equation (4.13) as follows. Equation (4.13) is

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \\ \ln P_{it}^{\times} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 & 0 \\ 0 & \gamma_0\zeta_{22} + \zeta_{32} & \zeta_{33} & \zeta_{34} \\ 0 & (\delta_0 + \gamma_0\phi_0)\zeta_{22} + \phi_0\zeta_{32} + \zeta_{42} & \phi_0\zeta_{33} + \zeta_{43} & \zeta_{44} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \\ \ln P_{it-1}^{\times} \end{bmatrix} +$$

$$+ \begin{bmatrix} 1 & 0 & 0 & 0 \\ c_{20}\theta & 1 & 0 & 0 \\ c_{20}\theta\gamma_0 & \gamma_0 & 1 & 0 \\ c_{20}\theta\delta_0 + c_{20}\theta\gamma_0\phi_0 & \delta_0 + \gamma_0\phi_0 & \phi_0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{it} + \eta_{1it} \\ c_{20}v_{it} \\ \eta_{3it} \\ \eta_{4it} \end{bmatrix} \tag{A.11}$$

The first and second error terms are not orthogonal in this system since $\epsilon_{it} = \theta q_{it} + v_{it}$. We derive the distribution of errors according to (A.2)-(A.6) and obtain

$$\begin{bmatrix} \epsilon_{it} + \eta_{1it} \\ c_{20}v_{it} \\ \eta_{3it} \\ \eta_{4it} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 + \sigma_2^2 & c_{20}[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & 0 & 0 \\ c_{20}[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & c_{20}^2[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix} \right] \tag{A.12}$$

We introduce the following notation:

$$\begin{pmatrix} \sigma_1^2 + \sigma_2^2 & c_{20}[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & 0 & 0 \\ c_{20}[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & c_{20}^2[\theta^2\sigma_1^2 + (1-\theta)^2\sigma_2^2] & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix} \equiv \begin{pmatrix} \dot\sigma_1^2 & \dot\sigma_{12} & 0 & 0 \\ \dot\sigma_{12} & \dot\sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix} \tag{A.13}$$

The Cholesky matrix of the error covariance matrix is

$$\begin{pmatrix} \dot\sigma_1^2 & \dot\sigma_{12} & 0 & 0 \\ \dot\sigma_{12} & \dot\sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix}^{1/2} = \begin{pmatrix} \dot\sigma_1 & 0 & 0 & 0 \\ \frac{\dot\sigma_{12}}{\dot\sigma_1} & \sqrt{\dot\sigma_2^2 - \frac{\dot\sigma_{12}^2}{\dot\sigma_1^2}} & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{pmatrix} \tag{A.14}$$

Using this Cholesky matrix, we rewrite Equation (A.11) to make the error vector orthogonal:

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \\ \ln P_{it}^{\times} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 & 0 \\ 0 & \gamma_0\zeta_{22}+\zeta_{32} & \zeta_{33} & \zeta_{34} \\ 0 & (\delta_0+\gamma_0\phi_0)\zeta_{22}+\phi_0\zeta_{32}+\zeta_{42} & \phi_0\zeta_{33}+\zeta_{43} & \zeta_{44} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \\ \ln P_{it-1}^{\times} \end{bmatrix} +$$

$$+ \begin{bmatrix} \dot{\sigma}_1 & 0 & 0 & 0 \\ c_{20}\theta\dot{\sigma}_1 + \frac{\dot{\sigma}_{12}}{\dot{\sigma}_1} & \sqrt{\dot{\sigma}_2^2 - \frac{\dot{\sigma}_{12}^2}{\dot{\sigma}_1^2}} & 0 & 0 \\ \gamma_0\left(c_{20}\theta\dot{\sigma}_1 + \frac{\dot{\sigma}_{12}}{\dot{\sigma}_1}\right) & \gamma_0\sqrt{\dot{\sigma}_2^2 - \frac{\dot{\sigma}_{12}^2}{\dot{\sigma}_1^2}} & \sigma_3 & 0 \\ (\delta_0+\gamma_0\phi_0)\left(c_{20}\theta\dot{\sigma}_1 + \frac{\dot{\sigma}_{12}}{\dot{\sigma}_1}\right) & (\delta_0+\gamma_0\phi_0)\sqrt{\dot{\sigma}_2^2 - \frac{\dot{\sigma}_{12}^2}{\dot{\sigma}_1^2}} & \phi_0\sigma_3 & \sigma_4 \end{bmatrix} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \\ e_{4it} \end{bmatrix} \tag{A.15}$$

where $(e_{1it}, e_{2it}, e_{3it}, e_{4it})' \sim N(0_{4\times 1}, I_4)$. We simplify the notation in Equation (A.15) by defining new parameters:

$$\begin{bmatrix} q_{it} \\ r_{it} \\ \ln P_{it} \\ \ln P_{it}^{\times} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{22} & 0 & 0 \\ 0 & \gamma_0\zeta_{22}+\zeta_{32} & \zeta_{33} & \zeta_{34} \\ 0 & (\delta_0+\gamma_0\phi_0)\zeta_{22}+\phi_0\zeta_{32}+\zeta_{42} & \phi_0\zeta_{33}+\zeta_{43} & \zeta_{44} \end{bmatrix} \begin{bmatrix} q_{it-1} \\ r_{it-1} \\ \ln P_{it-1} \\ \ln P_{it-1}^{\times} \end{bmatrix} +$$

$$+ \begin{bmatrix} \tilde{\sigma}_1 & 0 & 0 & 0 \\ \tilde{\sigma}_{12} & \tilde{\sigma}_2 & 0 & 0 \\ \gamma_0\tilde{\sigma}_{12} & \gamma_0\tilde{\sigma}_2 & \sigma_3 & 0 \\ (\delta_0+\gamma_0\phi_0)\tilde{\sigma}_{12} & (\delta_0+\gamma_0\phi_0)\tilde{\sigma}_2 & \phi_0\sigma_3 & \sigma_4 \end{bmatrix} \begin{bmatrix} e_{1it} \\ e_{2it} \\ e_{3it} \\ e_{4it} \end{bmatrix} \tag{A.16}$$

where $\tilde{\sigma}_1$, $\tilde{\sigma}_{12}$, and $\tilde{\sigma}_2$ can be expressed by the original parameters based on Equations (A.13) and (A.15).

Appendix B

Models 1-4

We use the following general notation for Models 1 to 4:

$$Y_{it} = a_i + \zeta Y_{it-1} + \Omega e_{it} \tag{B.1}$$

where Y_{it} is a $K \times 1$ vector of the endogenous variables; a_i is a $K \times 1$ vector of fixed effects; Ω is a lower triangular matrix with positive elements in its diagonal; $e_{it} \sim N(0_{K\times 1}, I_K)$ is the vector of error terms. We rewrite (B.1) as

$$(I_K - \zeta L)Y_{it} = a_i + \Omega e_{it} \tag{B.2}$$

$$Y_{it} = (I_K - \zeta L)^{-1} a_i + (I_K - \zeta L)^{-1}\Omega e_{it} \tag{B.3}$$

$$Y_{it} = \sum_{j=0}^{\infty} \zeta^j a_i + \sum_{j=0}^{\infty} \zeta^j \Omega e_{it-j} \tag{B.4}$$

By taking derivatives, we get the IRF matrices

$$\Theta_j = \frac{\partial Y_{it+j}}{\partial e_{it}} = \zeta^j \Omega \text{ for } j = 0, 1, 2, \ldots, \infty \tag{B.5}$$

Models 5-6

We use the following general notation for Models 5 and 6:

$$Y_{it} = a_i + \zeta Y_{it-1} + \zeta_{\mathrm{IL}} Y_{\mathrm{IL},t-1} D_{it}(i \in \mathrm{IF}) + \zeta_{\mathrm{IF}} \left(\sum_{k\in \mathrm{IF}} Y_{k,t-1} \right) D_{it}(i = \mathrm{IL}) + \Omega e_{it} \tag{B.6}$$

where Y_{it} is a 4×1 vector of the endogenous variables; a_i is a 4×1 vector of fixed effects; Ω is a lower triangular matrix with positive elements in its diagonal; $e_{it} \sim N(0_{4\times 1}, I_4)$ is the vector of error terms. We rewrite (B.6) as

$$\begin{aligned} Y_{it} &= (I_4 - \zeta L)^{-1} a_i + (I_4 - \zeta L)^{-1} \zeta_{\mathrm{IL}} Y_{\mathrm{IL},t-1} D(i \in \mathrm{IF}) + \\ &+ (I_4 - \zeta L)^{-1} \left(\sum_{k \in \mathrm{IF}} \zeta_{\mathrm{IF}} Y_{k,t-1} \right) D(i = \mathrm{IL}) + (I_4 - \zeta L)^{-1} \Omega e_{it} \end{aligned} \tag{B.7}$$

$$Y_{it} = \sum_{j=0}^{\infty} \zeta^j a_i + \left(\sum_{j=0}^{\infty} \zeta^j \zeta_{\mathrm{IL}} Y_{\mathrm{IL},t-1-j} \right) D(i \in \mathrm{IF}) +$$

$$+\left(\sum_{j=0}^{\infty}\sum_{k\in\mathrm{IF}}\zeta^j\zeta_{\mathrm{IF}}Y_{k,t-1-j}\right)D(i=\mathrm{IL})+\sum_{j=0}^{\infty}\zeta^j\Omega e_{it-j} \tag{B.8}$$

By taking derivatives, we get the IRF matrices

$$\Theta_j=\zeta^j\Omega \text{ for } j=0,1,2,\ldots,\infty \tag{B.9}$$

and the dynamic interaction multiplier matrices

$$\Gamma_j(\mathrm{IL}\rightarrow\mathrm{IF})=(\text{effects of } Y_{\mathrm{IL},t-j} \text{ on } Y_{it} \text{ for } i\in\mathrm{IF})=\zeta^j\zeta_{\mathrm{IL}} \text{ for } j=0,1,2,\ldots,\infty \tag{B.10}$$

$$\Gamma_j(\mathrm{IF}\rightarrow\mathrm{IL})=(\text{effects of } Y_{k,t-j} \text{ on } Y_{\mathrm{IL},t} \text{ for } k\in\mathrm{IF})=\zeta^j\zeta_{\mathrm{IF}} \text{ for } j=0,1,2,\ldots,\infty \tag{B.11}$$

Models 7-8

We use the following matrix notation for Models 7 and 8:

$$Y_{it}=a_i+\zeta Y_{it-1}+\zeta_{\mathrm{GL}}Y_{\mathrm{GL},t-1}D_{it}(i\in\mathrm{GF})+\zeta_{\mathrm{GF}}\left(\sum_{k\in\mathrm{GF}}Y_{k,t-1}\right)D_{it}(i=\mathrm{GL})+\Omega e_{it} \tag{B.12}$$

where Y_{it} is a 4×1 vector of the endogenous variables; a_i is a 4×1 vector of fixed effects; Ω is a lower triangular matrix with positive elements in its diagonal; $e_{it}\sim N(0_{4\times 1},I_4)$ is the vector of error terms. We rewrite (B.12) as

$$Y_{it}=(I_4-\zeta L)^{-1}a_i+(I_4-\zeta L)^{-1}\zeta_{\mathrm{GL}}Y_{\mathrm{GL},t-1}D(i\in\mathrm{GF})+$$

$$+(I_4-\zeta L)^{-1}\left(\sum_{k\in\mathrm{GF}}\zeta_{\mathrm{GF}}Y_{k,t-1}\right)D(i=\mathrm{GL})+(I_4-\zeta L)^{-1}\Omega e_{it} \tag{B.13}$$

$$Y_{it}=\sum_{j=0}^{\infty}\zeta^j a_i+\left(\sum_{j=0}^{\infty}\zeta^j\zeta_{\mathrm{GL}}Y_{\mathrm{GL},t-1-j}\right)D(i\in\mathrm{GF})+$$

$$+\left(\sum_{j=0}^{\infty}\sum_{k\in\mathrm{GF}}\zeta^j\zeta_{\mathrm{GF}}Y_{k,t-1-j}\right)D(i=\mathrm{GL})+\sum_{j=0}^{\infty}\zeta^j\Omega e_{it-j} \tag{B.14}$$

By taking derivatives, we get the IRF matrices

$$\Theta_j = \zeta^j \Omega \text{ for } j = 0, 1, 2, \ldots, \infty \tag{B.15}$$

and the dynamic interaction multiplier matrices

$$\Gamma_j(\text{GL} \to \text{GF}) = (\text{effects of } Y_{\text{GL},t-j} \text{ on } Y_{it} \text{ for } i \in \text{GF}) = \zeta^j \zeta_{\text{GL}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{B.16}$$

$$\Gamma_j(\text{GF} \to \text{GL}) = (\text{effects of } Y_{k,t-j} \text{ on } Y_{\text{GL},t} \text{ for } k \in \text{GF}) = \zeta^j \zeta_{\text{GF}} \text{ for } j = 0, 1, 2, \ldots, \infty \tag{B.17}$$

Appendix C

The Poisson-type patent count data model involving the common competitive latent factor is estimated by the simulated QML method (Gouriéroux and Monfort, 1991). The likelihood function is evaluated by using the EIS technique of Richard and Zhang (2007).

The conditional density of l_t^* is given by

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

The likelihood of a realization $(\tilde{P}, L^*) \equiv (\tilde{P}_{it}, l_t^* : t = 1, \ldots, T; i = 1, \ldots, N)$ is

$$\prod_{i=1}^{N}\prod_{t=1}^{T} f(\tilde{P}_{it}|\mathcal{F}_t) f^*(l_t^*|l_{t-1}^*) = \prod_{i=1}^{N}\prod_{t=1}^{T} \left\{ \frac{\exp(-\tilde{P}_{it}^o \tilde{P}_{it}^*)(\tilde{P}_{it}^o \tilde{P}_{it}^*)^{\tilde{P}_{it}}}{\tilde{P}_{it}!} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(l_t^* - \mu^* l_{t-1}^*)^2}{2}\right] \right\} \tag{C.2}$$

The likelihood of patent counts is obtained by integrating out all latent variables from (C.2):

$$\mathcal{L}(\tilde{P}|\mathcal{F}_e; \theta) = \int_{\mathbb{R}^T} \prod_{i=1}^{N}\prod_{t=1}^{T} \frac{\exp(-\lambda_{it})\lambda_{it}^{\tilde{P}_{it}}}{\tilde{P}_{it}!} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(l_t^* - \mu^* l_{t-1}^*)^2}{2}\right] dL^* \tag{C.3}$$

where θ denotes the vector of parameters of the model and

$$\mathcal{F}_e = \{\tilde{r}_{it}, c_{it}, d_{it} \text{ for } i = 1, \ldots, N; t = 1, \ldots, T\} \tag{C.4}$$

is the information set generated by the exogenous variables. We represent the likelihood of patent counts with the following compact notation:

$$\mathcal{L}(\tilde{P}|\mathcal{F}_e;\theta) = \int_{\mathbb{R}^T} \prod_{i=1}^{N}\prod_{t=1}^{T} g(\tilde{P}_{it}, l_t^*|\mathcal{F}_t;\theta)dL^* \tag{C.5}$$

where g is the joint density of $(\tilde{P}_{it}, l_t^*)$. We introduce the auxiliary sampler, m and include it in the likelihood function, as follows:

$$\mathcal{L}(\tilde{P}|\mathcal{F}_e;\theta,\theta^*) = \int_{\mathbb{R}^T} \prod_{i=1}^{N}\prod_{t=1}^{T} \frac{g(\tilde{P}_{it}, l_t^*|\mathcal{F}_t;\theta)}{m(l_t^*|l_{t-1}^*;\theta_t^*)} \times m(l_t^*|l_{t-1}^*;\theta_t^*)dL^* \tag{C.6}$$

where $\theta^* = (\theta_1^*, \ldots, \theta_T^*)$ denotes the parameters of the auxiliary sampler. The parameters of the auxiliary sampler are different in each period, however, the functional form of the sampler is constant over time. The importance Monte Carlo estimate of $\mathcal{L}(\tilde{P}|\mathcal{F}_e;\theta,\theta^*)$ for given θ^* is

$$\hat{\mathcal{L}}_R(\tilde{P}|\mathcal{F}_e;\theta,\theta^*) = \frac{1}{R}\sum_{r=1}^{R}\prod_{i=1}^{N}\prod_{t=1}^{T} \frac{g(\tilde{P}_{it}, l_{tr}^*|\mathcal{F}_t;\theta)}{m(l_{tr}^*|l_{t-1r}^*;\theta_t^*)} \tag{C.7}$$

where $\{l_{tr}^* : t = 1, \ldots, T\}$ denotes the r-th trajectory of i.i.d. draws from m. Richard and Zhang (2007) suggest defining the auxiliary sampler, m with its density kernel, k:

$$k(l_t^*, l_{t-1}^*;\theta_t^*) = m(l_t^*|l_{t-1}^*;\theta_t^*)\chi(l_{t-1}^*;\theta_t^*) \tag{C.8}$$

where

$$\chi(l_{t-1}^*;\theta_t^*) = \int_{\mathbb{R}} k(l_t^*, l_{t-1}^*;\theta_t^*)dl_t^* \tag{C.9}$$

denotes the integrating constant associated to k. Richard and Zhang (2007) suggest choosing k as a kernel of the normal distribution. Moreover, we include f^* into the auxiliary sampler, m; as suggested by Bauwens and Hautsch (2006). The normal density kernel is given by

$$k(l_t^*, l_{t-1}^*;\theta_t^*) = \exp\left[\theta_{1t}^* l_t^* + \theta_{2t}^*(l_t^*)^2\right] \times \exp\left[-\frac{(l_t^* - \mu^* l_{t-1}^*)^2}{2}\right] \tag{C.10}$$

where $\theta_t^* = (\theta_{1t}^*, \theta_{2t}^*)$ determines the conditional mean and variance of the auxiliary sampler for period

t. The conditional mean, μ_t and conditional variance, π_t^2 of the normal auxiliary sampler, m are

$$\mu_t = \pi_t^2(\theta_{1t}^* + \mu^* l_{t-1}^*) \tag{C.11}$$

$$\pi_t^2 = \frac{1}{1 - 2\theta_{2t}^*} \tag{C.12}$$

A trajectory of $\{l_t^* : t = 1, \ldots, T\}$ can be generated from the auxiliary sampler, as follows:

$$l_t^* = \mu_t + \pi_t \eta_t \tag{C.13}$$

where $\eta_t \sim N(0, 1)$ are i.i.d. common random numbers. Richard and Zhang (2007) suggest using the same set of random numbers (i.e., common random numbers) for every iteration of the maximum likelihood procedure. In the EIS method, the parameters of the auxiliary sampler minimize the variance of the Monte Carlo estimator of the likelihood function:

$$\theta^* = \arg\min_{\theta^*} \mathrm{Var}\left[\hat{\mathcal{L}}_R(\tilde{P}|\mathcal{F}_e; \theta, \theta^*)\right] = \arg\min_{\theta^*} \mathrm{Var}\left[\frac{1}{R}\sum_{r=1}^{R}\prod_{i=1}^{N}\prod_{t=1}^{T}\frac{g(\tilde{P}_{it}, l_{tr}^*|\mathcal{F}_t; \theta)}{m(l_{tr}^*|l_{t-1r}^*; \theta_t^*)}\right] \tag{C.14}$$

This variance is minimized by choosing such values for θ_t^* for which there is a good fit between g and m. To achieve this, Richard and Zhang (2007) suggest solving the minimization problem of (C.14) by estimating a recursive sequence of Ordinary Least Squares (OLS) problems, each of the following form:

$$\ln g(\tilde{P}_{it}, l_{tr}^*|\mathcal{F}_t; \theta) + \ln \chi(l_t^*; \hat{\theta}_{t+1}^*) = \theta_{0t}^* + \theta_{1t}^* l_{tr}^* + \theta_{2t}^*(l_{tr}^*)^2 + u_{tr} \text{ with } r = 1, \ldots, R \tag{C.15}$$

for $t = T, \ldots, 1$, $\chi(l_T^*, \hat{\theta}_{T+1}^*) = 1$ and $\hat{\theta}_{t+1}^*$ is the OLS estimate of θ_{t+1}^*. These regressions are run backwards, from T to 1 and the sample size of each regression is equal to the number of trajectories drawn, R. In our estimation, we choose $R = 50$.

The right side of Equation (C.15) includes the log kernel of the auxiliary sampler. Normal distribution is used for the auxiliary sampler since the log kernel of the normal distribution is a second order polynomial, therefore, its parameters can be estimated by OLS. The EIS technique involves the estimation of a large number of auxiliary sampler parameters over the maximum likelihood iterations. Therefore, it is essential to estimate these parameters very fast to make the EIS procedure feasible. The OLS estimation provides the auxiliary sampler parameter estimates rapidly, making the EIS based

QML estimation feasible in practice.

We can summarize the EIS method as follows:

- Step 1: We draw $R = 50$ trajectories $\{l^*_{tr}\}_{t=1}^{T}$ from the distribution $N(\mu^* l^*_{t-1r}, 1)$.
- Step 2: For each $t = T, \ldots, 1$, we estimate by OLS (C.15) to get the parameters of m.
- Step 3: We draw $R = 50$ trajectories $\{l^*_{tr}\}_{t=1}^{T}$ from the auxiliary samplers.
- We iterate Steps 2 and 3 five times.
- Step 4: We estimate the value of the likelihood function according to (C.7).

With these steps, the likelihood maximization procedure shows proper convergence to the optimum and we can compute the QML standard errors by the sandwich estimator without numerical problems.

Appendix D

To approximate the value of the common competitive factor, l^*_t, we compute its filtered estimates, $E[l^*_t|\mathcal{F}^o_t]$, conditioning on the observable information set

$$\mathcal{F}^o_t = \Big[(\tilde{P}_{i1}, \tilde{r}_{i1}, c_{i1}, d_{i1}), \ldots, (\tilde{P}_{it-1}, \tilde{r}_{it-1}, c_{it-1}, d_{it-1}), (\tilde{r}_{it}, c_{it}, d_{it}) : i = 1, \ldots, N\Big] \tag{D.1}$$

The conditional expectation of l^*_t is

$$E[l^*_t|\mathcal{F}^o_t] = \int_{\mathbb{R}} l^*_t h(l^*_t|\mathcal{F}^o_t) dl^*_t \tag{D.2}$$

where h denotes the conditional density of l^*_t. We introduce $\tilde{P}_t = (\tilde{P}_{is} : i = 1, \ldots, N; s = 1, \ldots, t)$ and $L^*_t = (l^*_s : s = 1, \ldots, t)$. We compute h as follows:

$$h(l^*_t|\mathcal{F}^o_t) = \frac{\dot{g}(\tilde{P}_{t-1}, l^*_t|\mathcal{F}^o_t)}{f(\tilde{P}_{t-1}|\mathcal{F}^o_t)} = \tag{D.3}$$

$$= \frac{\int_{\mathbb{R}^{t-1}} \tilde{g}(\tilde{P}_{t-1}, l^*_t, L^*_{t-1}|\mathcal{F}^o_t) dL^*_{t-1}}{\int_{\mathbb{R}^{t-1}} \ddot{g}(\tilde{P}_{t-1}, L^*_{t-1}|\mathcal{F}^o_t) dL^*_{t-1}} =$$

$$= \frac{\int_{\mathbb{R}^{t-1}} \ddot{g}(\tilde{P}_{t-1}, L^*_{t-1}|\mathcal{F}^o_t) f^*(l^*_t|\tilde{P}_{t-1}, L^*_{t-1}, \mathcal{F}^o_t) dL^*_{t-1}}{\int_{\mathbb{R}^{t-1}} \ddot{g}(\tilde{P}_{t-1}, L^*_{t-1}|\mathcal{F}^o_t) dL^*_{t-1}} =$$

$$= \frac{\int_{\mathbb{R}^{t-1}} f^*(l_t^*|l_{t-1}^*)\ddot{g}(\tilde{P}_{t-1}, L_{t-1}^*|\mathcal{F}_t^o)dL_{t-1}^*}{\int_{\mathbb{R}^{t-1}} \ddot{g}(\tilde{P}_{t-1}, L_{t-1}^*|\mathcal{F}_t^o)dL_{t-1}^*}$$

where $\dot{g}$, f, $\tilde{g}$, $\ddot{g}$, and f^* are conditional density functions of the corresponding random variables or vectors. Substituting Equation (D.3) into (D.2), and using the fact that the denominator in (D.3) is not a function of l_t^*, we obtain that

$$E[l_t^*|\mathcal{F}_t^o] = \frac{\int_{\mathbb{R}^t} l_t^* f^*(l_t^*|l_{t-1}^*)\ddot{g}(\tilde{P}_{t-1}, L_{t-1}^*|\mathcal{F}_t^o)dL_t^*}{\int_{\mathbb{R}^{t-1}} \ddot{g}(\tilde{P}_{t-1}, L_{t-1}^*|\mathcal{F}_t^o)dL_{t-1}^*} \tag{D.4}$$

where the joint density, $\ddot{g}$ is given by

$$\ddot{g}(P_{t-1}, L_{t-1}^*|\mathcal{F}_t^o) = \prod_{i=1}^{N}\prod_{s=1}^{t-1} \frac{\exp(-\lambda_{is})\lambda_{is}^{\tilde{P}_{is}}}{\tilde{P}_{is}!} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(l_s^* - \mu^* l_{s-1}^*)^2}{2}\right] \tag{D.5}$$

The high-dimensional integrals in Equation (D.4) are estimated by the EIS technique.

References

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P., 2005. Competition and innovation: An inverted-U relationship. The Quarterly Journal of Economics 120, 701-728.

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

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

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

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

Blazsek, S., Escribano, A., 2012. Patents, secret innovations and firm's rate of return: Differential effects of the innovation leader. Working Paper 12-02, Economic Series, Universidad Carlos III de Madrid.

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

Boldrin, M., Levine, D.K., 2010. Against Intellectual Monopoly. Cambridge University Press.

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

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

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

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

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

Cincera, M., 2005. Firms' productivity growth and R&D spillovers: An analysis of alternative technological proximity measures, CEPR Discussion Papers No. 4894.

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

Fama, E. F., 1970. Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25, 383-417.

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

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

Grabowski, H.G., Vernon, J.M., 2000. Effective patent life in pharmaceuticals. International Journal of Technology Management 19, 98-120.

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

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., Mairesse, J., 1996. Estimating the productivity of research and development in French and United States manufacturing firms. In: Van Ark, B., Wagner, K. (Eds.), International Productivity Differences, Measurement and Explanations. Elsevier Science, Amsterdam.

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

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

Hamilton, J.D., 1994. Time series analysis. Princeton University Press, Princeton, NJ.

Hausman, J., Hall, B., Griliches, Z., 1984. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52, 909-938.

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

Jaffe, A.B., 1986. Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. The American Economic Review 76, 984-1001.

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

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

Kahn, A. E., 1962. The role of patents. In: John P. Miller (ed.) Competition, Cartels and Their Regulation. North-Holland, Amsterdam.

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

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

LeRoy, S. F., Porter, R. D., 1981. The present-value relation: Tests based on implied variance bounds. Econometrica 49, 555-74.

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

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

Lev, B., Radhakrishnan, S., Ciftci, M., 2006. The stock market valuation of R&D leaders. NYU Working Paper No. 2451/27475.

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

Ljung, G., Box., G., 1978. On a measure of lack of fit in time-series models. Biometrika 65, 297-303.

Machlup, F., 1962. An economic review of the patent system. In: John P. Miller (ed.) Competition, Cartels and Their Regulation. North-Holland, Amsterdam.

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

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

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

Nickell, S., 1996. Competition and Corporate Performance. Journal of Political Economy 4, 724-746.

Pakes, A., Schankerman, M., 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.

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

Porter, M.E., 1979. How competitive forces shape strategy. Harvard Business Review 57, 137-145.

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

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

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

Richard, J.-F., Zhang, W., 2007. Efficient high-dimensional importance sampling. Journal of Econometrics 141, 1385-1411.

Rogers, E.M., 1983. Diffusion of Innovations. Free Press, New York.

Scherer, F. M., 1965a. Corporate inventive output, profits, and growth. Journal of Political Economy 73, 290-297.

Scherer, F. M., 1965b. Firm size, market structure, opportunity, and the output of patented inventions. American Economic Review 55, 1097-1125.

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

Vandenbussche, J., Aghion, P., Meghir, C., 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, 236-244.

Table 1
Product market based industry classification in the technological cluster.

SIC industry name	SIC	K	HM industry name	K
Pharmaceutical preparations	2834	47	Pharmaceuticals	92
Biological products (no diagnostic substances)	2836	31	Non-manufacturing	10
In vitro and in vivo diagnostic substances	2835	7	Computers and inst.	4
Perfumes, cosmetics and other toilet preparations	2844	3	Chemicals	2
Surgical and medical instruments, and apparatus	3841	3	Food	2
Medicinal chemicals and botanical products	2833	2	Rubber and plastics	1
Wholesale-drugs, proprietaries and druggists' sundries	5122	2		
Services-medical laboratories	8071	2		
Grain mill products	2040	1		
Beverages	2080	1		
Chemicals and allied products	2800	1		
Soap, detergents, cleaning preparations, perfumes, cosmetics	2840	1		
Paints, varnishes, lacquers, enamels and allied prods	2851	1		
Agricultural chemicals	2870	1		
Plastics products, NEC	3089	1		
Electromedical and electrotherapeutic apparatus	3845	1		
Wholesale-medical, dental and hospital equipment, and supplies	5047	1		
Fire, marine and casualty insurance	6331	1		
Services-hospitals	8060	1		
Services-engineering, accounting, research, management	8700	1		
Services-commercial physical and biological research	8731	1		
Non-operating establishments	9995	1		
Total number of firms		111		111

Notes: Standard Industry Classification (SIC). Number of firms (K). Hall and Mairesse (HM, 1996) classification.

Table 2
Patent innovations leadership classification of firms.

Firm name (SIC)	Cluster		(V1)	(V2)	(V3)	(V4)	(V5)	(V6)	(V7)	(V8)	(V9)
1. Merck (2834)	IL	GL	217.6	1367.5	136.7	147232.4	12.39	8.47	13.20	0.82	1.16
2. Eli Lilly (2834)	IF	GL	116.0	613.6	58.6	43645.5	12.21	8.15	12.69	0.83	1.17
3. Abbott Lab. (2834)	IF	GL	97.5	720.8	73.9	40954.3	11.89	7.91	12.67	0.80	1.13
4. Warner-Lambert (2834)	IF	GL	81.7	656.2	61.3	31542.0	10.64	7.23	10.55	0.75	1.29
5. Pfizer (2834)	IF	GL	103.0	553.2	49.1	23373.0	12.21	8.32	12.79	0.81	1.16
6. Bristol-Myers (2834)	IF	GL	69.7	307.4	34.2	11509.1	12.11	8.27	12.95	0.80	1.14
7. Am. Home Prod. (2834)	IF	GL	52.8	330.7	30.7	8396.9	10.82	8.05	11.38	0.72	1.25
8. Alza (2834)	IF	GL	35.5	547.6	40.9	7683.0	8.24	5.28	9.92	0.78	1.01
9. Mallinckrodt (2835)	IF	GF	23.5	181.9	16.9	2007.8	9.25	6.92	9.82	0.68	1.11
10. Pharmacia & U. (2834)	IF	GF	21.4	45.9	8.5	1922.6	10.96	7.45	10.34	0.79	1.34
11. Church & Dwight (2840)	IF	GF	12.3	83.4	9.9	1537.8	7.96	5.10	9.74	0.66	0.89
12. NeoRx (2835)	IF	GF	6.5	68.4	7.8	500.5	7.25	4.12	7.92	1.07	0.96
13. Alliance Pharma. (2834)	IF	GF	4.2	69.9	6.9	369.8	7.42	4.33	8.83	1.06	0.87
14. Xoma (2836)	IF	GF	6.9	48.7	5.0	329.6	8.11	4.21	8.84	1.15	0.96
15. Enzon (2836)	IF	GF	4.2	47.1	6.0	235.9	6.95	4.13	8.79	1.08	0.83
16. Guilford Pharma. (2834)	IF	GF	3.5	23.0	4.9	216.5	6.75	4.20	6.98	0.98	1.01
17. Sugen (2836)	IF	GF	4.4	23.8	4.0	216.2	6.50	4.02	6.42	0.95	1.05
18. Inhale Therap. (2834)	IF	GF	2.5	31.3	7.0	169.3	6.33	4.17	7.15	0.92	0.97
19. Corvas (2836)	IF	GF	3.5	16.3	2.1	122.5	6.58	4.07	7.00	1.00	1.00
20. Molecular Bios. (2835)	IF	GF	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); Group of Leaders (GL); Group of Followers (GF). The table presents nine variables for 20 out of the 111 firms of the technological cluster analyzed for the period 1979 to 2000. The following variables are presented:

(V1) mean patent applications count, $(1/T)\sum_{t=1}^{T}\tilde{P}_{it}$
(V2) mean forward citations received count, $(1/T)\sum_{t=1}^{T}c_{f,it}$
(V3) mean forward citations received count corrected for sample truncation bias, $(1/T)\sum_{t=1}^{T}\tilde{c}_{f,it}$
(V4) mean knowledge stock, $(1/T)\sum_{t=1}^{T}\sum_{s=0}^{t}\tilde{c}_{f,is}\tilde{P}_{is}(1-\delta)^{t-s}$
(V5) mean ln R&D expenses, $(1/T)\sum_{t=1}^{T}\tilde{r}_{it}$
(V6) mean ln book value, $(1/T)\sum_{t=1}^{T}z_{it}$
(V7) mean ln stock market value, $(1/T)\sum_{t=1}^{T}m_{it}$
(V8) mean ln R&D expenses to log sales, $(1/T)\sum_{t=1}^{T}\tilde{r}_{it}/s_{it}$
(V9) mean ln R&D expenses to log stock market value, $(1/T)\sum_{t=1}^{T}\tilde{r}_{it}/m_{it}$

Table 3
Evolution of the knowledge stock for firms in the group of patent innovation leaders.

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

Table 4
Parameter estimates and model diagnostics for Models 1 and 2.

Model 1				Model 2			
ζ matrix		*Residuals diagnostics*		*ζ matrix*		*Residuals diagnostics*	
γ_0	$0.05^{*}(0.032)$	t test	p-value	ζ^*_{11}	$-0.06^{**}(0.029)$	t test	p-value
		Mean e_{1it}	1.00	ζ^*_{12}	$0.05^{***}(0.016)$	Mean e_{1it}	1.00
		Mean e_{2it}	1.00	ζ^*_{13}	$-0.01(0.029)$	Mean e_{2it}	1.00
		Mean e_{3it}	1.00	ζ^*_{21}	$0.09^{***}(0.018)$	Mean e_{3it}	1.00
ζ_{22}	$0.78^{***}(0.032)$	χ^2 test	p-value	ζ^*_{22}	$0.77^{***}(0.032)$	χ^2 test	p-value
		Var e_{1it}	0.12	ζ^*_{23}	$0.02(0.025)$	Var e_{1it}	0.12
		Var e_{2it}	0.18	ζ^*_{31}	$-0.01(0.033)$	Var e_{2it}	0.17
ζ_{32}	$0.05(0.032)$	Var e_{3it}	0.13	ζ^*_{32}	$0.09^{***}(0.023)$	Var e_{3it}	0.12
ζ_{33}	$0.23^{***}(0.045)$	LB test	p-value	ζ^*_{33}	$0.23^{***}(0.045)$	LB test	p-value
$\rho(Z)$	0.777	LB e_{1it}	0.50	$\rho(Z)$	0.780	LB e_{1it}	0.52
Cholesky matrix, Ω		LB e_{2it}	0.49	*Cholesky matrix, Ω*		LB e_{2it}	0.50
$\tilde{\sigma}_1$	$0.63^{***}(0.067)$	LB e_{3it}	0.33	σ^*_1	$0.62^{***}(0.067)$	LB e_{3it}	0.33
$\tilde{\sigma}_2$	$0.57^{***}(0.036)$	*Model diagnostics*		σ^*_2	$0.57^{***}(0.036)$	*Model diagnostics*	
σ_3	$0.74^{***}(0.025)$	LL	-7151	σ^*_3	$0.74^{***}(0.025)$	LL	-7132
$\tilde{\sigma}_{12}$	$0.03^{***}(0.011)$	LR	37.85	σ^*_{12}	$0.04^{***}(0.011)$		
		LR p-value	0.000	σ^*_{13}	$-0.01(0.019)$		
		AIC	15002	σ^*_{23}	$0.03^{*}(0.019)$	AIC	14978

Notes: Model 1 is $Y_{it} = a_i + \zeta Y_{it-1} + \Omega e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it})'$; Model 2 is $Y_{it} = a_i + \zeta^* Y_{it-1} + \Omega^* e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it})'$. Ljung-Box (LB); Log Likelihood (LL); Likelihood Ratio (LR); Akaike Information Criterion (AIC). $\rho(\zeta)$ denotes the spectral radius of ζ. *, **, and *** denote parameter significance at the 10%, 5%, and 1% levels, respectively. QML standard errors are reported in parentheses. For each error term, we report average of p-values computed over $i = 1, \ldots, N$ for the following tests: t test for $H_0 : E[e_{it}] = 0$; χ^2 test for $H_0 : \mathrm{Var}[e_{it}] = 1$; LB test for $H_0 : \{e_{it} : t = 1, \ldots, T\}$ are uncorrelated. The LB test is performed for 5 lags.

Table 5
Parameter estimates of the latent-factor patent count data model.

μ_0	$0.50^{***}(0.090)$	β_0	$0.63^{***}(0.016)$	ω_0	$0.00^{**}(0.000)$	ϕ_0	$0.00(0.000)$
$\gamma_1\ t$	$0.14^{***}(0.002)$	β_1	$-0.04^{***}(0.009)$	ω_1	$0.00(0.000)$	ϕ_1	$0.00(0.000)$
$\gamma_2\ tr_{it}$	$-0.01^{***}(0.001)$	β_2	$0.02^{***}(0.007)$	ω_2	$0.00(0.000)$	ϕ_2	$0.00(0.000)$
$\gamma_3\ r_{it}^2$	$-0.04^{***}(0.003)$	β_3	$-0.03^{***}(0.004)$	ω_3	$0.00(0.001)$	ϕ_3	$0.00(0.000)$
$\gamma_4\ z_{it}$	$0.01^{***}(0.001)$	β_4	$-0.03^{***}(0.002)$	ω_4	$0.00(0.001)$	ϕ_4	$0.00(0.000)$
$\gamma_5\ \tilde{P}_{i1}$	$0.05^{***}(0.001)$	β_5	$0.03^{***}(0.005)$	ω_5	$0.00(0.001)$	ϕ_5	$0.00(0.001)$
κ_1	$0.00(0.000)$	β_6	$-0.01^{***}(0.002)$	ω_6	$0.00(0.004)$	ϕ_6	$0.00(0.001)$
μ^*	$0.91^{***}(0.007)$	β_7	$-0.01^{***}(0.002)$	ω_7	$0.00(0.005)$	ϕ_7	$0.00(0.004)$
		β_8	$-0.01^{***}(0.002)$	ω_8	$0.01(0.005)$	ϕ_8	$0.00(0.007)$
		β_9	$-0.01(0.006)$	ω_9	$0.00(0.006)$	ϕ_9	$0.00(0.007)$
		β_{10}	$0.02^{***}(0.008)$	ω_{10}	$0.01(0.007)$	ϕ_{10}	$-0.01(0.013)$

Notes: The latent-factor patent count data model for patent count intensity is $\lambda_{it} = \tilde{P}_{it}^o \tilde{P}_{it}^*$ where $\ln \tilde{P}_{it}^o = \mu_0 + \gamma_1 t + \gamma_2 t\tilde{r}_{it} + \gamma_3 \tilde{r}_{it}^2 + \gamma_4 z_{it} + \gamma_5 \tilde{P}_{i1} + \sum_{k=0}^{10} \beta_k \tilde{r}_{it-k} + \sum_{k=0}^{10} \omega_k c_{it-k}\tilde{r}_{it} + \sum_{k=0}^{10} \phi_k d_{it-k}\tilde{r}_{it} + \kappa_1 \ln \tilde{P}_{it-1}^o$ and $\ln \tilde{P}_{it}^* = \ln \Phi(\mu_i + \sigma_i l_t^*)$. We do not report the estimates of μ_i and σ_i. *, **, and *** denote parameter significance at the 10%, 5%, and 1% levels, respectively. QML standard errors are reported in parentheses.

Table 6
Parameter estimates and model diagnostics for Models 3 and 4.

Model 3				Model 4			
ζ matrix		*Residuals diagnostics*		*ζ matrix*		*Residuals diagnostics*	
γ_0	$0.06^{*}(0.032)$	t test	p-value	ζ^*_{11}	$-0.06^{**}(0.029)$	t test	p-value
δ_0	$0.25^{***}(0.034)$	Mean e_{1it}	1.00	ζ^*_{12}	$0.04^{***}(0.015)$	Mean e_{1it}	1.00
ϕ_0	$0.01^{**}(0.005)$	Mean e_{2it}	1.00	ζ^*_{13}	$-0.01(0.030)$	Mean e_{2it}	1.00
		Mean e_{3it}	1.00	ζ^*_{14}	$0.00(0.016)$	Mean e_{3it}	1.00
		Mean e_{4it}	1.00	ζ^*_{21}	$0.09^{***}(0.018)$	Mean e_{4it}	1.00
ζ_{22}	$0.77^{***}(0.033)$	χ^2 test	p-value	ζ^*_{22}	$0.77^{***}(0.039)$	χ^2 test	p-value
		Var e_{1it}	0.12	ζ^*_{23}	$0.02(0.025)$	Var e_{1it}	0.12
		Var e_{2it}	0.18	ζ^*_{24}	$-0.01(0.120)$	Var e_{2it}	0.17
		Var e_{3it}	0.13	ζ^*_{31}	$-0.01(0.034)$	Var e_{3it}	0.12
ζ_{32}	$0.05(0.033)$	Var e_{4it}	0.10	ζ^*_{32}	$0.09^{***}(0.024)$	Var e_{4it}	0.10
ζ_{33}	$0.23^{***}(0.045)$	LB test	p-value	ζ^*_{33}	$0.23^{***}(0.045)$	LB test	p-value
ζ_{34}	$0.02(0.024)$	LB e_{1it}	0.50	ζ^*_{34}	$0.02(0.028)$	LB e_{1it}	0.51
		LB e_{2it}	0.48	ζ^*_{41}	$0.03^{***}(0.008)$	LB e_{2it}	0.49
ζ_{42}	$-0.19^{***}(0.032)$	LB e_{3it}	0.34	ζ^*_{42}	$0.00(0.016)$	LB e_{3it}	0.34
ζ_{43}	$0.01(0.008)$	LB e_{4it}	0.10	ζ^*_{43}	$0.01(0.011)$	LB e_{4it}	0.14
ζ_{44}	$0.32^{***}(0.072)$	*Model diagnostics*		ζ^*_{44}	$0.31^{***}(0.085)$	*Model diagnostics*	
$\rho(Z)$	0.774	LL	-6976	$\rho(Z)$	0.778	LL	-6956
Cholesky matrix, Ω		LR	39.66	*Cholesky matrix, Ω*			
$\tilde{\sigma}_1$	$0.63^{***}(0.000)$	LR p-value	0.000	σ^*_1	$0.62^{***}(0.067)$		
$\tilde{\sigma}_2$	$0.58^{***}(0.000)$	AIC	14894	σ^*_2	$0.57^{***}(0.036)$	AIC	14876
σ_3	$0.74^{***}(0.000)$			σ^*_3	$0.74^{***}(0.025)$		
σ_4	$0.23^{***}(0.000)$			σ^*_4	$0.22^{***}(0.016)$		
$\tilde{\sigma}_{12}$	$0.03^{***}(0.002)$			σ^*_{12}	$0.04^{***}(0.011)$		
				σ^*_{13}	$-0.01(0.019)$		
				σ^*_{23}	$0.03^{*}(0.018)$		
				σ^*_{14}	$0.01^{***}(0.005)$		
				σ^*_{24}	$0.14^{***}(0.016)$		
				σ^*_{34}	$0.01^{**}(0.004)$		

Notes: Model 3 is $Y_{it} = a_i + \zeta Y_{it-1} + \Omega e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P^{\times}_{it})'$; Model 4 is $Y_{it} = a_i + \zeta^* Y_{it-1} + \Omega^* e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P^{\times}_{it})'$. Ljung-Box (LB); Log Likelihood (LL); Likelihood Ratio (LR); Akaike Information Criterion (AIC). $\rho(\zeta)$ denotes the spectral radius of ζ. *, **, and *** denote parameter significance at the 10%, 5%, and 1% levels, respectively. QML standard errors are reported in parentheses. For each error term, we report average of p-values computed over $i = 1, \ldots, N$ for the following tests: t test for $H_0 : E[e_{it}] = 0$; χ^2 test for $H_0 : \mathrm{Var}[e_{it}] = 1$; LB test for $H_0 : \{e_{it} : t = 1, \ldots, T\}$ are uncorrelated. The LB test is performed for 5 lags. The LR test is performed for Models 3 and 4; see Table 7.

Table 7
Parameter estimates and model diagnostics for Model 5.

ζ matrix		*ζ_{IL} matrix*		*Residuals diagnostics*		*Model diagnostics*	
γ_0	$0.08^{**}(0.033)$	$\zeta_{\mathrm{IL},22}$	$0.27^{***}(0.062)$	t test	p-value	LL	-6716
δ_0	$0.22^{***}(0.033)$	$\zeta_{\mathrm{IL},32}$	$-0.74^{***}(0.148)$	Mean e_{1it}	1.00	LR	238.63
ϕ_0	$0.02^{***}(0.005)$	$\zeta_{\mathrm{IL},33}$	$0.28(0.222)$	Mean e_{2it}	1.00	LR p-value	0.000
ζ_{22}	$0.78^{***}(0.029)$	$\zeta_{\mathrm{IL},34}$	$-0.42^{*}(0.223)$	Mean e_{3it}	1.00	AIC	14402
ζ_{32}	$0.04(0.033)$	$\zeta_{\mathrm{IL},42}$	$0.16^{***}(0.044)$	Mean e_{4it}	1.00		
ζ_{33}	$0.17^{***}(0.047)$	$\zeta_{\mathrm{IL},43}$	$-0.10(0.087)$	χ^2 test	p-value		
ζ_{34}	$0.05^{**}(0.026)$	$\zeta_{\mathrm{IL},44}$	$0.17^{***}(0.046)$	Var e_{1it}	0.16		
ζ_{42}	$-0.18^{***}(0.035)$	*ζ_{IF} matrix*		Var e_{2it}	0.16		
ζ_{43}	$0.02^{**}(0.009)$	$\zeta_{\mathrm{IF},22}$	$0.01^{***}(0.003)$	Var e_{3it}	0.16		
ζ_{44}	$0.39^{***}(0.071)$	$\zeta_{\mathrm{IF},32}$	$0.01(0.007)$	Var e_{4it}	0.10		
$\rho(Z)$	0.783	$\zeta_{\mathrm{IF},33}$	$-0.01^{***}(0.003)$	LB test	p-value		
Cholesky matrix, Ω		$\zeta_{\mathrm{IF},34}$	$-0.04(0.022)$	LB e_{1it}	0.50		
$\tilde{\sigma}_1$	$0.63^{***}(0.067)$	$\zeta_{\mathrm{IF},42}$	$0.00(0.003)$	LB e_{2it}	0.51		
$\tilde{\sigma}_2$	$0.57^{***}(0.036)$	$\zeta_{\mathrm{IF},43}$	$0.00(0.004)$	LB e_{3it}	0.35		
σ_3	$0.72^{***}(0.024)$	$\zeta_{\mathrm{IF},44}$	$0.01(0.012)$	LB e_{4it}	0.22		
σ_4	$0.21^{***}(0.014)$						
$\tilde{\sigma}_{12}$	$0.03^{***}(0.012)$						

Notes: Model 5 is $Y_{it} = a_i + \zeta Y_{it-1} + \zeta_{\mathrm{IL}} Y_{\mathrm{IL},t-1} D_{it}(i \in \mathrm{IF}) + \zeta_{\mathrm{IF}} \sum_{k \in \mathrm{IF}} Y_{kt-1} \; D_{it}(i = \mathrm{IL}) + \Omega e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P^{\times}_{it})'$. Ljung-Box (LB); Log Likelihood (LL); Likelihood Ratio (LR); Akaike Information Criterion (AIC). $\rho(\zeta)$ denotes the spectral radius of ζ. *, **, and *** denote parameter significance at the 10%, 5%, and 1% levels, respectively. QML standard errors are reported in parentheses. For each error term, we report average of p-values computed over $i = 1, \ldots, N$ for the following tests: t test for $H_0 : E[e_{it}] = 0$; χ^2 test for $H_0 : \mathrm{Var}[e_{it}] = 1$; LB test for $H_0 : \{e_{it} : t = 1, \ldots, T\}$ are uncorrelated. The LB test is performed for 5 lags. The LR test is performed for Models 5 and 6; see Table 8.

Table 8
Parameter estimates and model diagnostics for Model 6.

ζ matrix		*ζ_{IL} matrix*		*Residuals diagnostics*	
ζ^*_{11}	$-0.07^{**}(0.033)$	$\zeta^*_{\mathrm{IL},11}$	$0.06(0.103)$	t test	p-value
ζ^*_{12}	$0.04(0.039)$	$\zeta^*_{\mathrm{IL},12}$	$-0.54^{***}(0.138)$	Mean e_{1it}	1.00
ζ^*_{13}	$-0.01(0.026)$	$\zeta^*_{\mathrm{IL},13}$	$-3.47(9.555)$	Mean e_{2it}	1.00
ζ^*_{14}	$0.00(0.133)$	$\zeta^*_{\mathrm{IL},14}$	$0.87^{***}(0.197)$	Mean e_{3it}	1.00
ζ^*_{21}	$0.08^{**}(0.036)$	$\zeta^*_{\mathrm{IL},21}$	$-0.12(0.092)$	Mean e_{4it}	1.00
ζ^*_{22}	$0.75^{***}(0.021)$	$\zeta^*_{\mathrm{IL},22}$	$0.05(0.141)$	χ^2 test	p-value
ζ^*_{23}	$0.06^{**}(0.026)$	$\zeta^*_{\mathrm{IL},23}$	$-4.45(11.771)$	Var e_{1it}	0.14
ζ^*_{24}	$0.07^{*}(0.038)$	$\zeta^*_{\mathrm{IL},24}$	$1.12^{***}(0.193)$	Var e_{2it}	0.17
ζ^*_{31}	$0.02(0.029)$	$\zeta^*_{\mathrm{IL},31}$	$-0.17^{*}(0.100)$	Var e_{3it}	0.17
ζ^*_{32}	$0.09^{***}(0.035)$	$\zeta^*_{\mathrm{IL},32}$	$-0.10(0.100)$	Var e_{4it}	0.10
ζ^*_{33}	$0.19^{***}(0.019)$	$\zeta^*_{\mathrm{IL},33}$	$4.20(11.215)$	LB test	p-value
ζ^*_{34}	$0.07(0.111)$	$\zeta^*_{\mathrm{IL},34}$	$-0.77^{***}(0.195)$	LB e_{1it}	0.50
ζ^*_{41}	$0.02(0.018)$	$\zeta^*_{\mathrm{IL},41}$	$-0.05^{*}(0.032)$	LB e_{2it}	0.49
ζ^*_{42}	$-0.02^{**}(0.008)$	$\zeta^*_{\mathrm{IL},42}$	$0.22^{***}(0.056)$	LB e_{3it}	0.33
ζ^*_{43}	$0.04^{***}(0.012)$	$\zeta^*_{\mathrm{IL},43}$	$-1.32(3.376)$	LB e_{4it}	0.25
ζ^*_{44}	$0.41^{***}(0.012)$	$\zeta^*_{\mathrm{IL},44}$	$0.36^{***}(0.076)$	*Model diagnostics*	
$\rho(Z)$	0.763	*ζ_{IF} matrix*		LL	−6597
Cholesky matrix, Ω		$\zeta^*_{\mathrm{IF},11}$	$0.00(0.253)$	AIC	14221
σ^*_{1}	$0.62^{***}(0.005)$	$\zeta^*_{\mathrm{IF},12}$	$-0.01(0.337)$		
σ^*_{2}	$0.55^{***}(0.009)$	$\zeta^*_{\mathrm{IF},13}$	$0.00(0.140)$		
σ^*_{3}	$0.72^{***}(0.012)$	$\zeta^*_{\mathrm{IF},14}$	$0.00(0.711)$		
σ^*_{4}	$0.21^{***}(0.002)$	$\zeta^*_{\mathrm{IF},21}$	$0.00(0.959)$		
σ^*_{12}	$0.02(0.028)$	$\zeta^*_{\mathrm{IF},22}$	$0.00(0.406)$		
σ^*_{13}	$0.00(0.023)$	$\zeta^*_{\mathrm{IF},23}$	$0.00(0.252)$		
σ^*_{23}	$0.06^{**}(0.026)$	$\zeta^*_{\mathrm{IF},24}$	$0.01(1.255)$		
σ^*_{14}	$0.01(0.018)$	$\zeta^*_{\mathrm{IF},31}$	$0.00(1.180)$		
σ^*_{24}	$0.13^{***}(0.006)$	$\zeta^*_{\mathrm{IF},32}$	$0.00(1.363)$		
σ^*_{34}	$0.01(0.013)$	$\zeta^*_{\mathrm{IF},33}$	$0.00(1.523)$		
		$\zeta^*_{\mathrm{IF},34}$	$0.01(2.708)$		
		$\zeta^*_{\mathrm{IF},41}$	$0.00(0.307)$		
		$\zeta^*_{\mathrm{IF},42}$	$0.00(0.104)$		
		$\zeta^*_{\mathrm{IF},43}$	$0.00(0.044)$		
		$\zeta^*_{\mathrm{IF},44}$	$0.01(0.339)$		

Notes: Model 6 is $Y_{it} = a_i + \zeta^* Y_{it-1} + \zeta^*_{\mathrm{IL}} Y_{\mathrm{IL},t-1} D_{it}(i \in \mathrm{IF}) + \zeta^*_{\mathrm{IF}} \sum_{k \in \mathrm{IF}} Y_{kt-1} \; D_{it}(i = \mathrm{IL}) + \Omega^* e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P^{\times}_{it})'$. Ljung-Box (LB); Log Likelihood (LL); Akaike Information Criterion (AIC). $\rho(\zeta)$ denotes the spectral radius of ζ. *, **, and *** denote parameter significance at the 10%, 5%, and 1% levels, respectively. QML standard errors are reported in parentheses. For each error term, we report average of p-values computed over $i = 1, \ldots, N$ for the following tests: t test for $H_0 : E[e_{it}] = 0$; χ^2 test for $H_0 : \mathrm{Var}[e_{it}] = 1$; LB test for $H_0 : \{e_{it} : t = 1, \ldots, T\}$ are uncorrelated. The LB test is performed for 5 lags.

Table 9
Parameter estimates and model diagnostics for Model 7.

ζ matrix		*ζ_{GL} matrix*		*Residuals diagnostics*		*Model diagnostics*	
γ_0	0.07(0.045)	$\zeta_{\mathrm{GL},22}$	$0.01^{***}(0.005)$	t test	p-value	LL	-6718
δ_0	$0.23^{***}(0.008)$	$\zeta_{\mathrm{GL},32}$	$-0.03^{***}(0.006)$	Mean e_{1it}	1.00	LR	183.78
ϕ_0	0.02(0.017)	$\zeta_{\mathrm{GL},33}$	0.10(0.151)	Mean e_{2it}	1.00	LR p-value	0.000
ζ_{22}	$0.77^{***}(0.021)$	$\zeta_{\mathrm{GL},34}$	0.01(0.018)	Mean e_{3it}	1.00	AIC	14406
ζ_{32}	0.06(0.046)	$\zeta_{\mathrm{GL},42}$	$0.01^{***}(0.002)$	Mean e_{4it}	1.00		
ζ_{33}	$0.16^{***}(0.018)$	$\zeta_{\mathrm{GL},43}$	$-0.02(0.045)$	χ^2 test	p-value		
ζ_{34}	0.05(0.092)	$\zeta_{\mathrm{GL},44}$	0.00(0.004)	Var e_{1it}	0.16		
ζ_{42}	$-0.19^{***}(0.008)$	*ζ_{GF} matrix*		Var e_{2it}	0.16		
ζ_{43}	0.02(0.013)	$\zeta_{\mathrm{GF},22}$	$0.01^{*}(0.007)$	Var e_{3it}	0.17		
ζ_{44}	$0.39^{***}(0.005)$	$\zeta_{\mathrm{GF},32}$	0.00(0.901)	Var e_{4it}	0.10		
$\rho(Z)$	0.766	$\zeta_{\mathrm{GF},33}$	0.01(0.783)	LB test	p-value		
Cholesky matrix, Ω		$\zeta_{\mathrm{GF},34}$	0.01(3.419)	LB e_{1it}	0.50		
$\tilde{\sigma}_1$	$0.63^{***}(0.004)$	$\zeta_{\mathrm{GF},42}$	$-0.01(0.046)$	LB e_{2it}	0.50		
$\tilde{\sigma}_2$	$0.57^{***}(0.008)$	$\zeta_{\mathrm{GF},43}$	0.01(0.045)	LB e_{3it}	0.36		
σ_3	$0.72^{***}(0.012)$	$\zeta_{\mathrm{GF},44}$	0.04(0.124)	LB e_{4it}	0.20		
σ_4	$0.21^{***}(0.002)$						
$\tilde{\sigma}_{12}$	0.03(0.027)						

Notes: Model 7 is $Y_{it} = a_i + \zeta Y_{it-1} + \zeta_{\mathrm{GL}} Y_{\mathrm{GL},t-1} D_{it}(i \in \mathrm{GF}) + \zeta_{\mathrm{GF}} \sum_{k\in\mathrm{GF}} Y_{kt-1}\; D_{it}(i = \mathrm{GL}) + \Omega e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P^{\times}_{it})'$. Ljung-Box (LB); Log Likelihood (LL); Likelihood Ratio (LR); Akaike Information Criterion (AIC). $\rho(\zeta)$ denotes the spectral radius of ζ. *, **, and *** denote parameter significance at the 10%, 5%, and 1% levels, respectively. QML standard errors are reported in parentheses. For each error term, we report average of p-values computed over $i = 1, \ldots, N$ for the following tests: t test for $H_0 : E[e_{it}] = 0$; χ^2 test for $H_0 : \mathrm{Var}[e_{it}] = 1$; LB test for $H_0 : \{e_{it} : t = 1, \ldots, T\}$ are uncorrelated. The LB test is performed for 5 lags. The LR test is performed for Models 7 and 8; see Table 10.

Table 10
Parameter estimates and model diagnostics for Model 8.

ζ matrix		*ζ_{GL} matrix*		*Residuals diagnostics*	
ζ^*_{11}	$-0.05(0.032)$	$\zeta^*_{\mathrm{GL},11}$	$-0.02(0.012)$	t test	p-value
ζ^*_{12}	$0.05(0.039)$	$\zeta^*_{\mathrm{GL},12}$	$0.00(0.006)$	Mean e_{1it}	1.00
ζ^*_{13}	$-0.01(0.026)$	$\zeta^*_{\mathrm{GL},13}$	$-0.06(0.213)$	Mean e_{2it}	1.00
ζ^*_{14}	$0.00(0.128)$	$\zeta^*_{\mathrm{GL},14}$	$0.00(0.018)$	Mean e_{3it}	1.00
ζ^*_{21}	$0.10^{***}(0.035)$	$\zeta^*_{\mathrm{GL},21}$	$-0.02^{*}(0.013)$	Mean e_{4it}	1.00
ζ^*_{22}	$0.77^{***}(0.022)$	$\zeta^*_{\mathrm{GL},22}$	$-0.02^{***}(0.007)$	χ^2 test	p-value
ζ^*_{23}	$0.04(0.026)$	$\zeta^*_{\mathrm{GL},23}$	$-0.92(1.323)$	Var e_{1it}	0.15
ζ^*_{24}	$0.07^{*}(0.038)$	$\zeta^*_{\mathrm{GL},24}$	$0.06^{***}(0.018)$	Var e_{2it}	0.17
ζ^*_{31}	$0.02(0.030)$	$\zeta^*_{\mathrm{GL},31}$	$-0.01(0.010)$	Var e_{3it}	0.17
ζ^*_{32}	$0.11^{***}(0.034)$	$\zeta^*_{\mathrm{GL},32}$	$-0.02^{***}(0.008)$	Var e_{4it}	0.10
ζ^*_{33}	$0.17^{***}(0.019)$	$\zeta^*_{\mathrm{GL},33}$	$0.34(0.534)$	LB test	p-value
ζ^*_{34}	$0.06(0.108)$	$\zeta^*_{\mathrm{GL},34}$	$0.00(0.022)$	LB e_{1it}	0.52
ζ^*_{41}	$0.02(0.019)$	$\zeta^*_{\mathrm{GL},41}$	$-0.01(0.005)$	LB e_{2it}	0.50
ζ^*_{42}	$-0.01(0.009)$	$\zeta^*_{\mathrm{GL},42}$	$0.01^{*}(0.004)$	LB e_{3it}	0.34
ζ^*_{43}	$0.03^{***}(0.013)$	$\zeta^*_{\mathrm{GL},43}$	$-0.19(0.275)$	LB e_{4it}	0.22
ζ^*_{44}	$0.41^{***}(0.013)$	$\zeta^*_{\mathrm{GL},44}$	$0.01^{**}(0.007)$	*Model diagnostics*	
$\rho(Z)$	0.781	*ζ_{GF} matrix*		LL	-6626
Cholesky matrix, Ω		$\zeta^*_{\mathrm{GF},11}$	$0.00(1.935)$	AIC	14280
σ^*_{1}	$0.62^{***}(0.005)$	$\zeta^*_{\mathrm{GF},12}$	$0.00(0.528)$		
σ^*_{2}	$0.55^{***}(0.008)$	$\zeta^*_{\mathrm{GF},13}$	$0.00(1.379)$		
σ^*_{3}	$0.72^{***}(0.012)$	$\zeta^*_{\mathrm{GF},14}$	$0.00(4.218)$		
σ^*_{4}	$0.21^{***}(0.002)$	$\zeta^*_{\mathrm{GF},21}$	$0.00(0.271)$		
σ^*_{12}	$0.03(0.027)$	$\zeta^*_{\mathrm{GF},22}$	$-0.01(0.180)$		
σ^*_{13}	$-0.01(0.023)$	$\zeta^*_{\mathrm{GF},23}$	$0.01(0.504)$		
σ^*_{23}	$0.05^{*}(0.025)$	$\zeta^*_{\mathrm{GF},24}$	$0.05(1.474)$		
σ^*_{14}	$0.01(0.017)$	$\zeta^*_{\mathrm{GF},31}$	$0.00(0.822)$		
σ^*_{24}	$0.13^{***}(0.006)$	$\zeta^*_{\mathrm{GF},32}$	$0.00(0.456)$		
σ^*_{34}	$0.01(0.013)$	$\zeta^*_{\mathrm{GF},33}$	$0.01(1.060)$		
		$\zeta^*_{\mathrm{GF},34}$	$0.02(3.451)$		
		$\zeta^*_{\mathrm{GF},41}$	$0.00(0.537)$		
		$\zeta^*_{\mathrm{GF},42}$	$0.00(0.041)$		
		$\zeta^*_{\mathrm{GF},43}$	$0.01(0.140)$		
		$\zeta^*_{\mathrm{GF},44}$	$0.05(0.298)$		

Notes: Model 8 is $Y_{it} = a_i + \zeta^* Y_{it-1} + \zeta^*_{\mathrm{GL}} Y_{\mathrm{GL},t-1} D_{it}(i \in \mathrm{GF}) + \zeta^*_{\mathrm{GF}} \ \sum_{k \in \mathrm{GF}} Y_{kt-1} \ D_{it}(i = \mathrm{GL}) + \Omega^* e_{it}$ with $Y_{it} = (q_{it}, r_{it}, \ln P_{it}, \ln P^{\times}_{it})'$. Ljung-Box (LB); Log Likelihood (LL); Akaike Information Criterion (AIC). $\rho(\zeta)$ denotes the spectral radius of ζ. *, **, and *** denote parameter significance at the 10%, 5%, and 1% levels, respectively. QML standard errors are reported in parentheses. For each error term, we report average of p-values computed over $i = 1, \ldots, N$ for the following tests: t test for $H_0 : E[e_{it}] = 0$; χ^2 test for $H_0 : \mathrm{Var}[e_{it}] = 1$; LB test for $H_0 : \{e_{it} : t = 1, \ldots, T\}$ are uncorrelated. The LB test is performed for 5 lags.

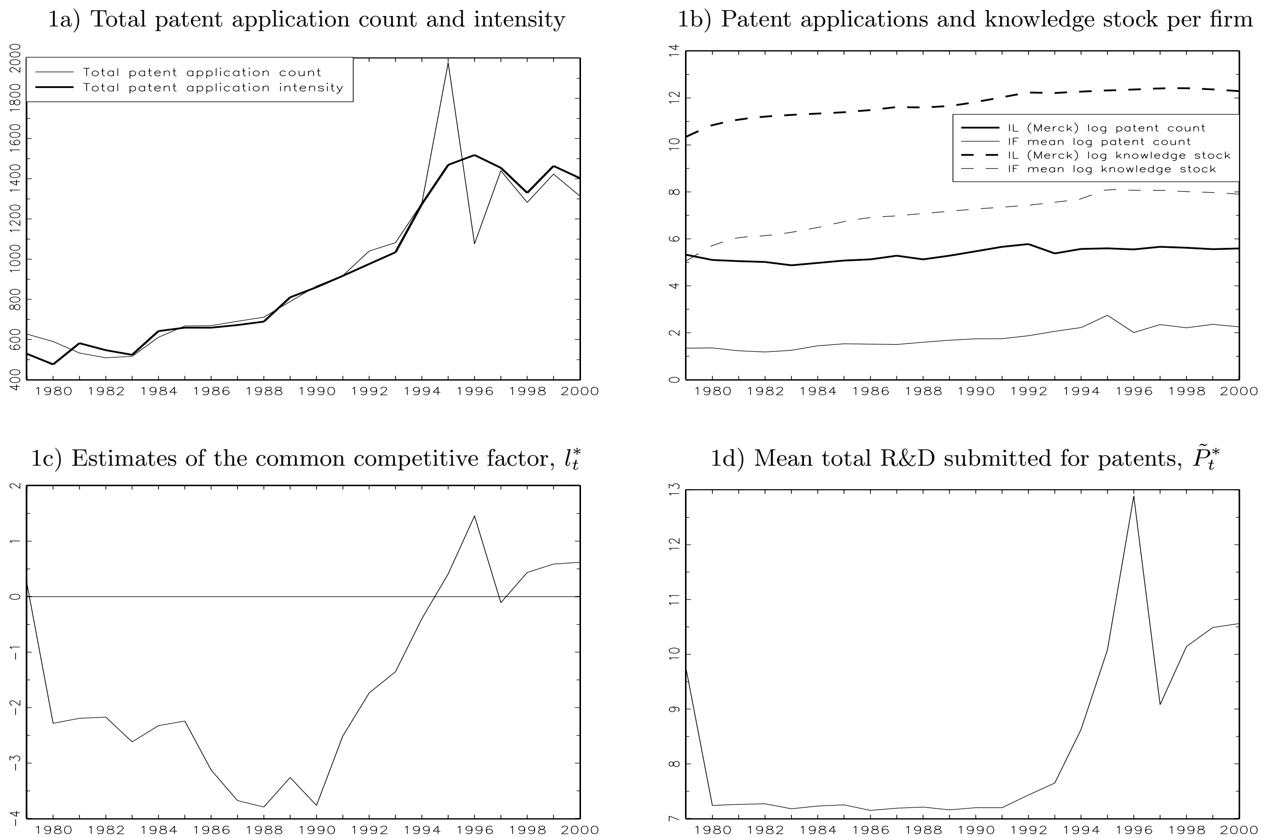


Fig. 1. Patent applications, patent intensity, and propensity to patent.
Notes: 1a) shows the evolution of $\sum_{i=1}^{111} \tilde{P}_{it}$ and $\sum_{i=1}^{111} \lambda_{it}$; 1b) shows patent application count and knowledge stock per firm for IL and IF; 1c) shows the estimates of $E[l_t^* | \mathcal{F}_t^o]$ for all t; 1d) shows the evolution of $(1/111) \sum_{i=1}^{111} \tilde{P}_t^*$, in percentage.

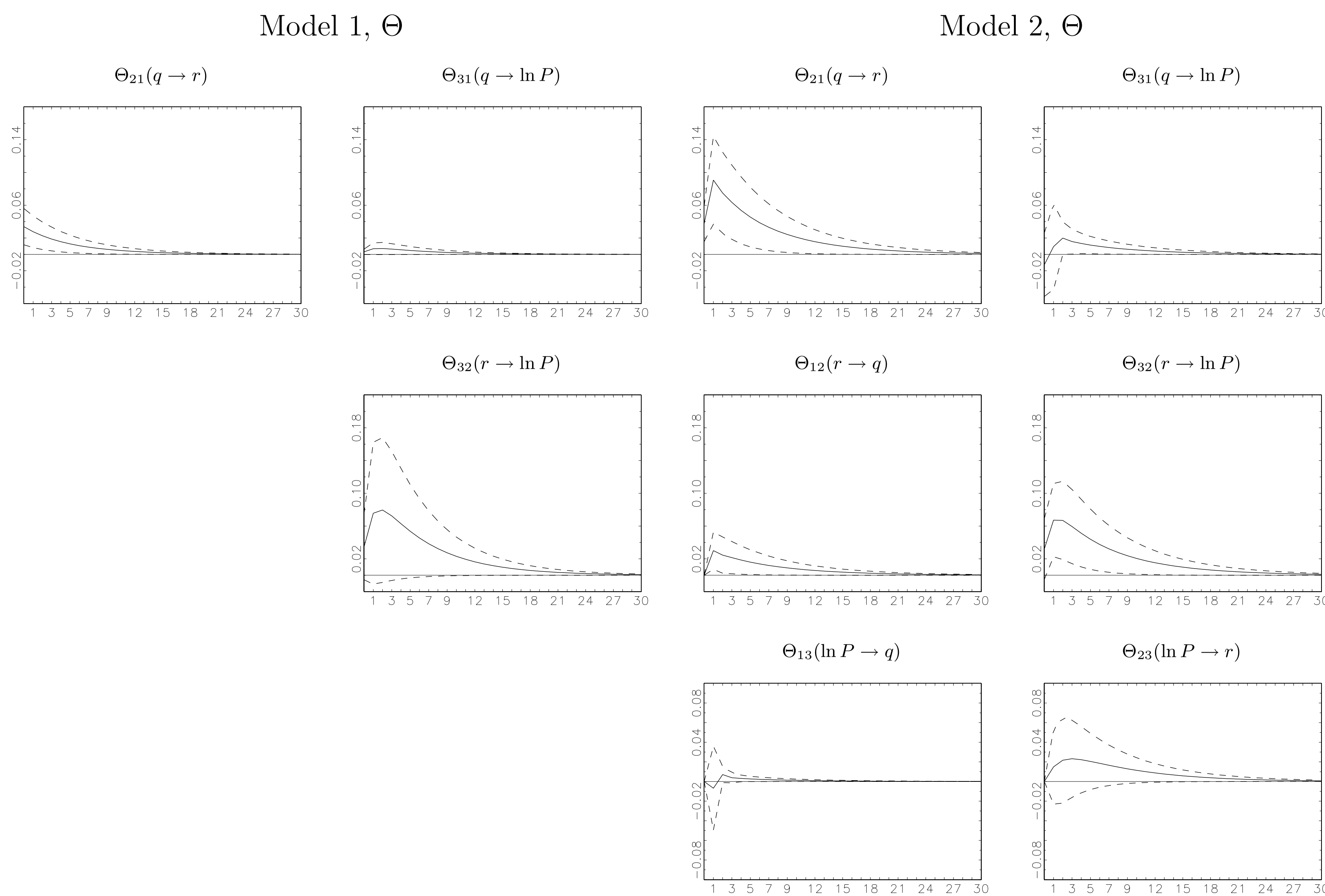


Fig. 2. Impulse response function, Θ_j for Models 1 and 2 for $j = 0, \ldots, 30$ leads.

Notes: Each figure shows $\overline{\Theta}_j$ and the confidence band defined by $\overline{\Theta}_j \pm 2\sigma(\Theta_j)$. The figure does not show the diagonal elements of Θ_j.

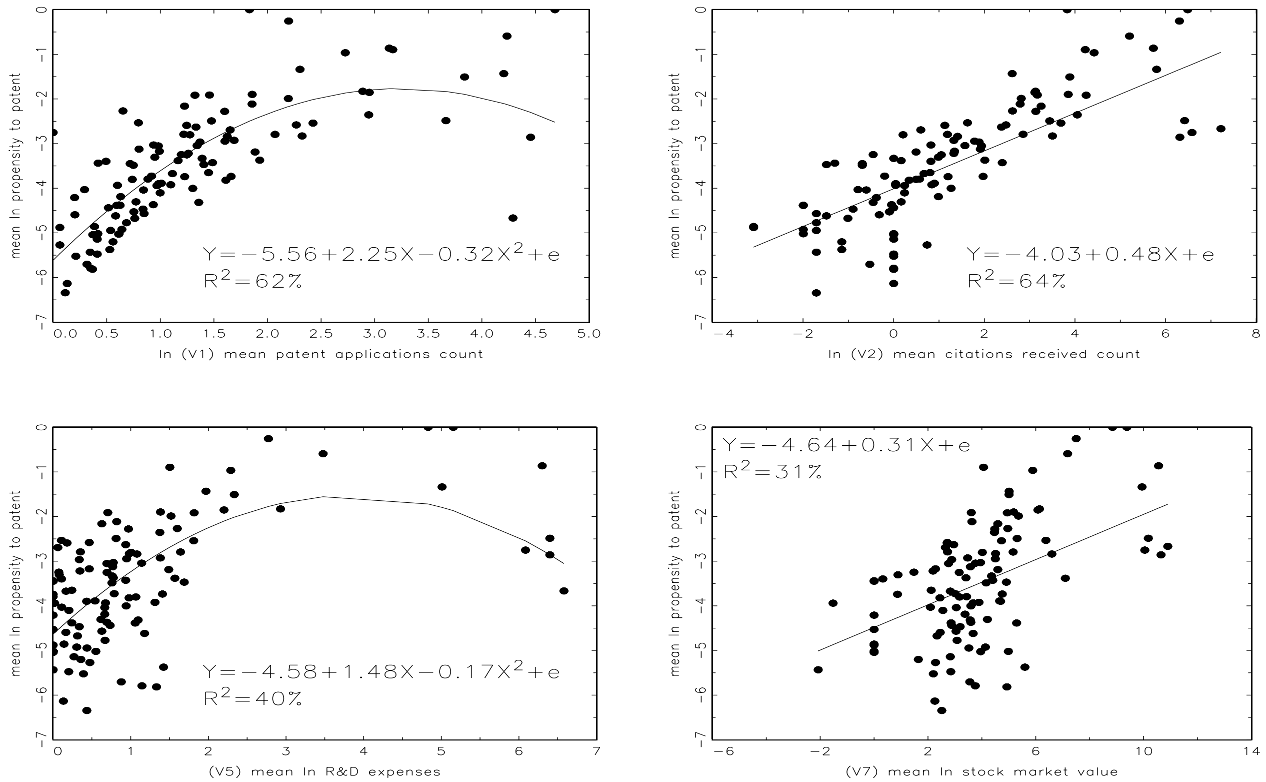


Fig. 3. Determinants of propensity to patent.

Notes: The figure shows the fitted values of $(1/22)\sum_{t=1}^{22} \ln \tilde{P}_{it}^{*}$ for $i = 1, \ldots, 111$. First and second order polynomial regressions are estimated by least squares. The definition of each explanatory variable is presented in Table 2.

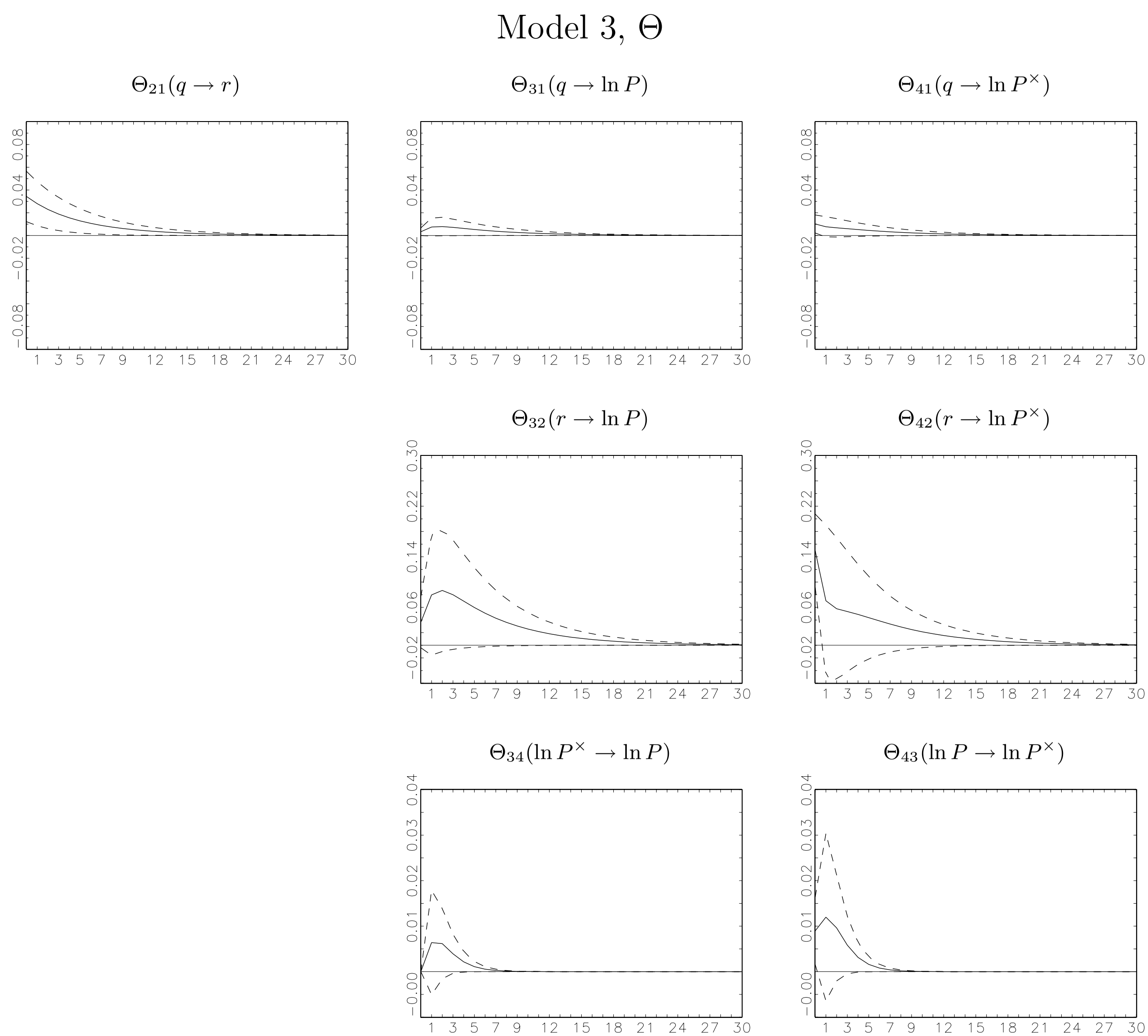


Fig. 4. Impulse response function, Θ_j for Model 3 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Theta}_j$ and the confidence band defined by $\overline{\Theta}_j \pm 2\sigma(\Theta_j)$. The figure does not show the diagonal elements of Θ_j.

Model 4, Θ

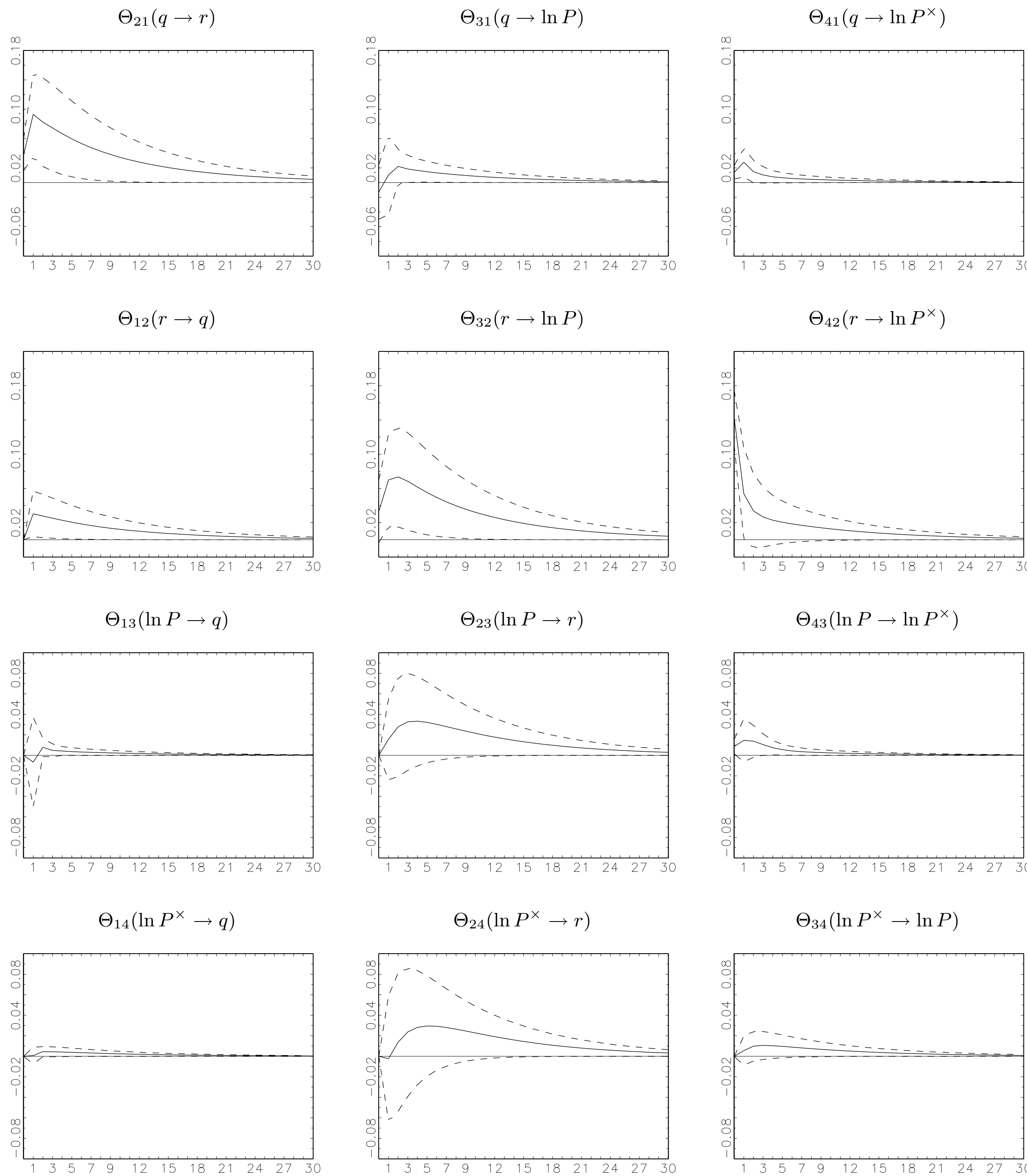


Fig. 5. Impulse response function, Θ_j for Model 4 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Theta}_j$ and the confidence band defined by $\overline{\Theta}_j \pm 2\sigma(\Theta_j)$. The figure does not show the diagonal elements of Θ_j.

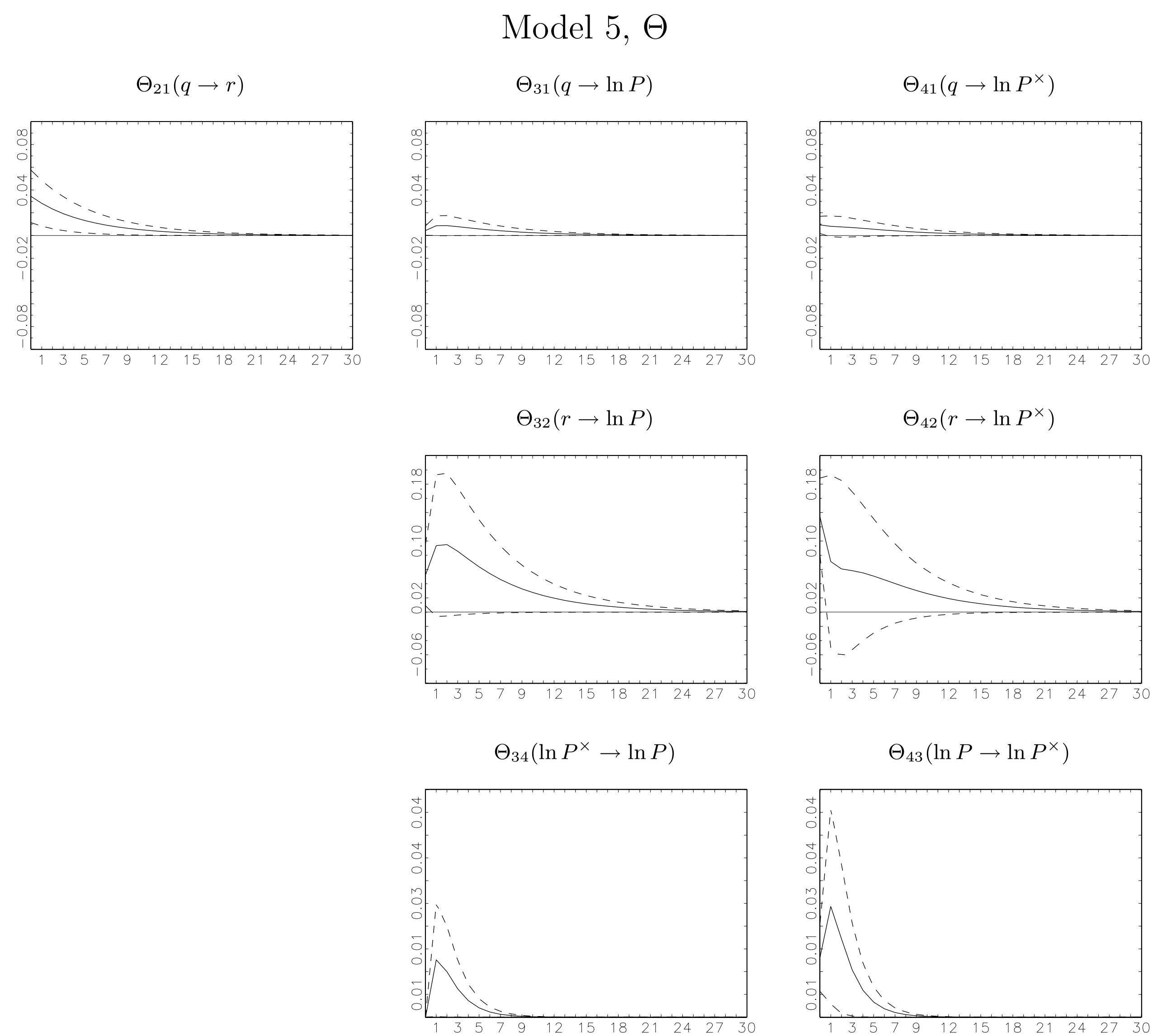


Fig. 6. Impulse response function, Θ_j for Model 5 for $j = 0, \ldots, 30$ leads.

Notes: Each figure shows $\overline{\Theta}_j$ and the confidence band defined by $\overline{\Theta}_j \pm 2\sigma(\Theta_j)$. The figure does not show the diagonal elements of Θ_j.

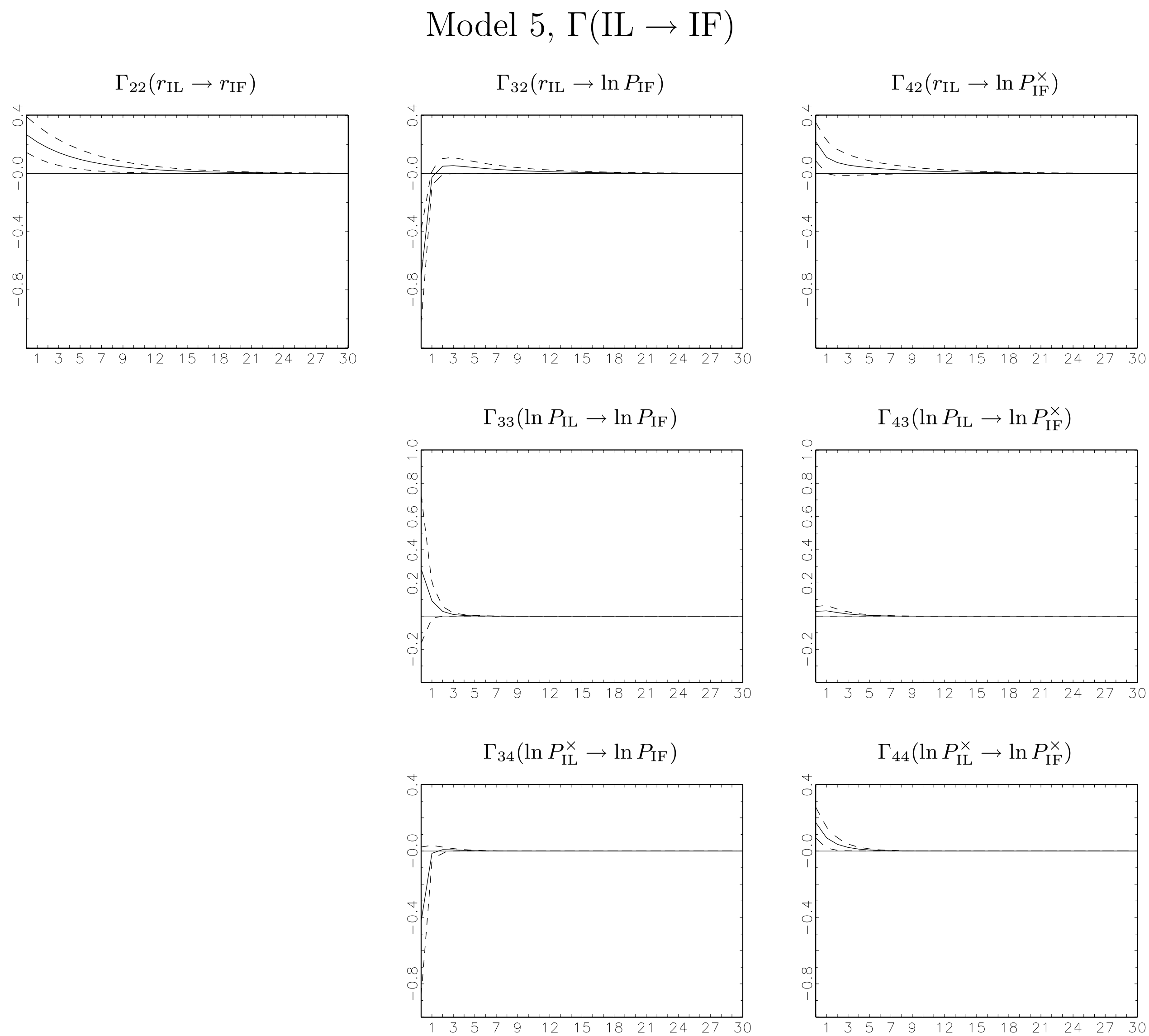


Fig. 7. Dynamic multiplier, $\Gamma_j(\mathrm{IL} \rightarrow \mathrm{IF})$ for Model 5 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

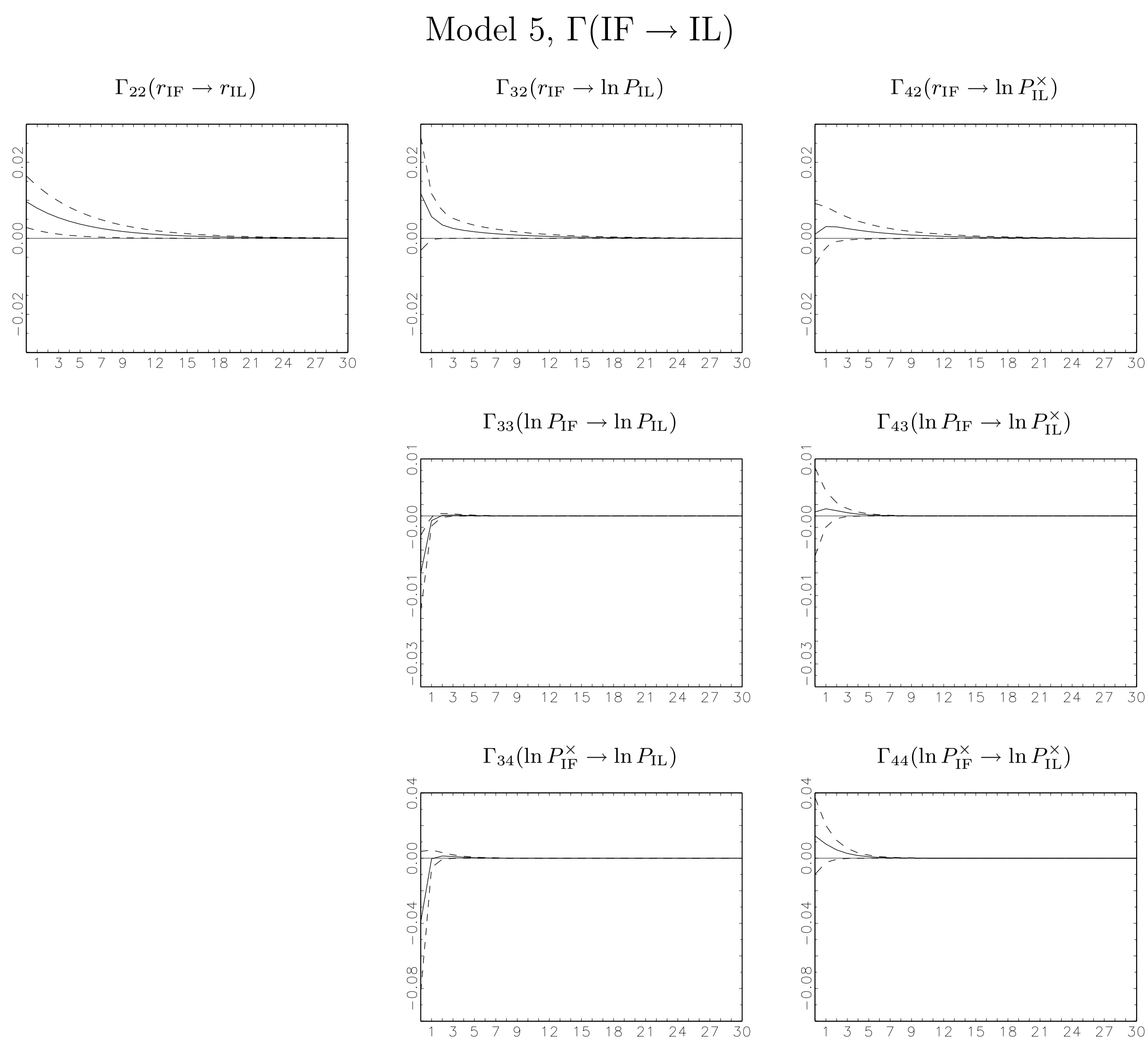


Fig. 8. Dynamic multiplier, $\Gamma_j(\mathrm{IF} \to \mathrm{IL})$ for Model 5 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

Model 6, Θ

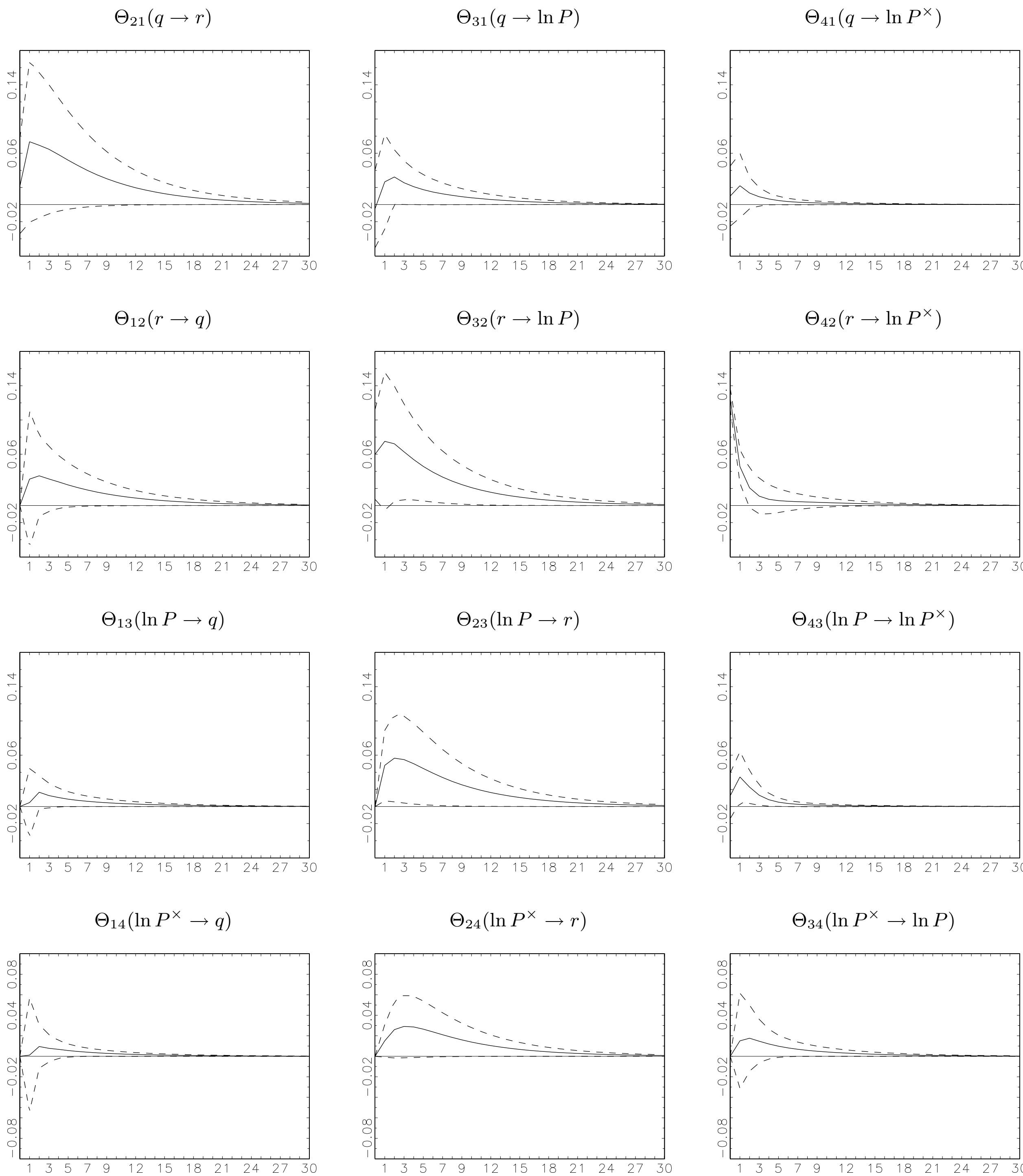


Fig. 9. Impulse response function, Θ_j for Model 6 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Theta}_j$ and the confidence band defined by $\overline{\Theta}_j \pm 2\sigma(\Theta_j)$. The figure does not show the diagonal elements of Θ_j.

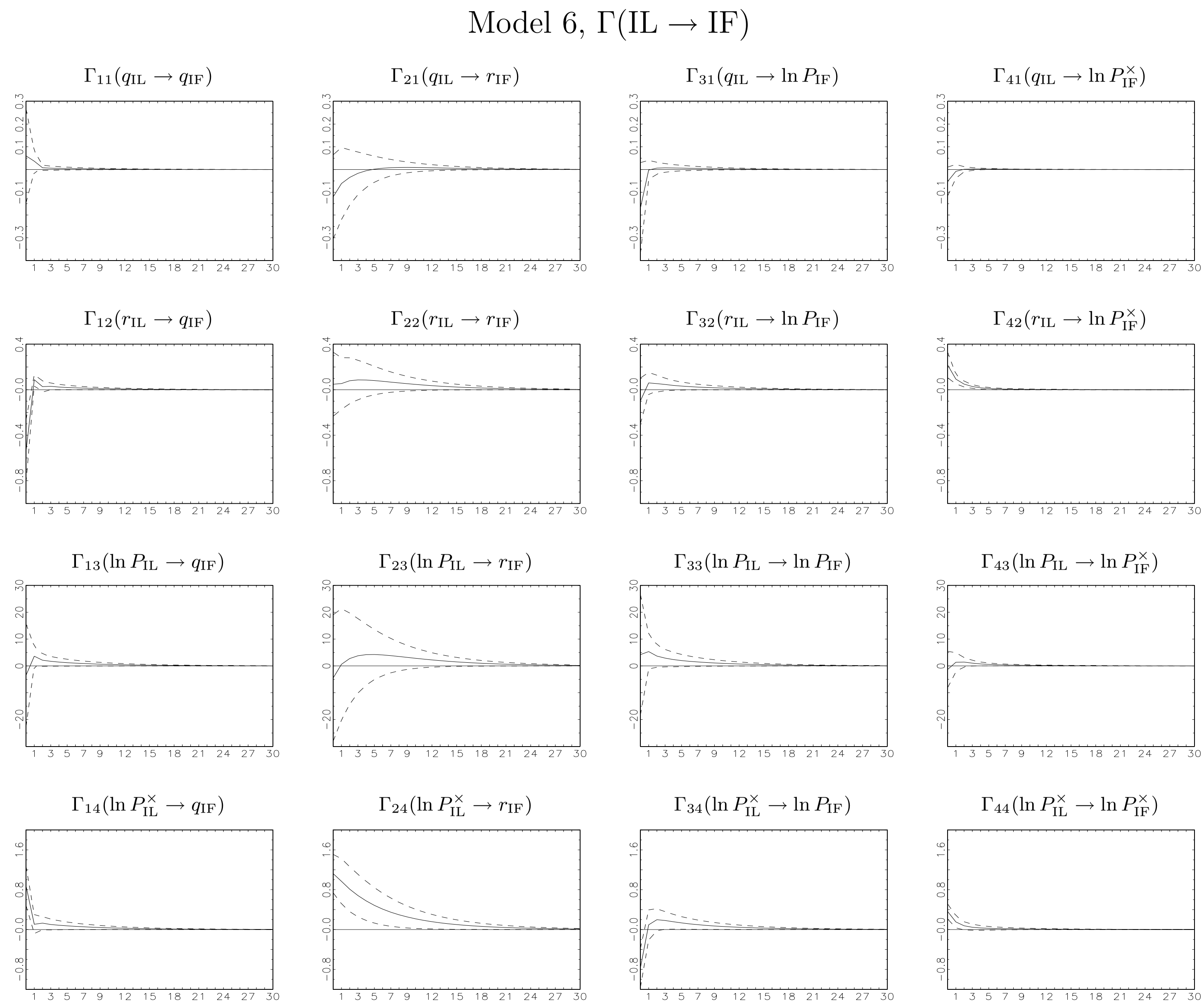


Fig. 10. Dynamic multiplier, $\Gamma_j(\mathrm{IL} \to \mathrm{IF})$ for Model 6 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

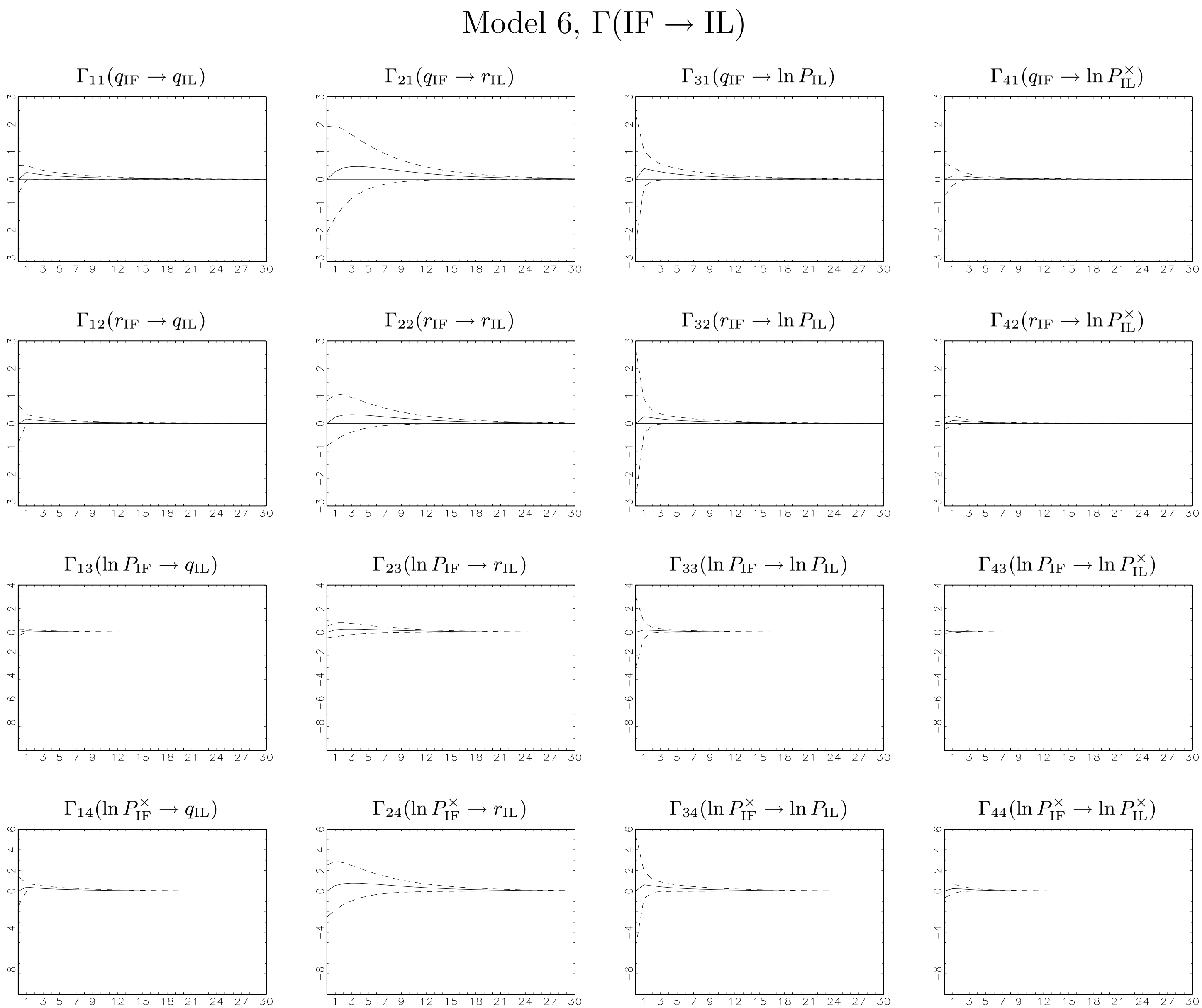


Fig. 11. Dynamic multiplier, $\Gamma_j(\mathrm{IF} \to \mathrm{IL})$ for Model 6 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

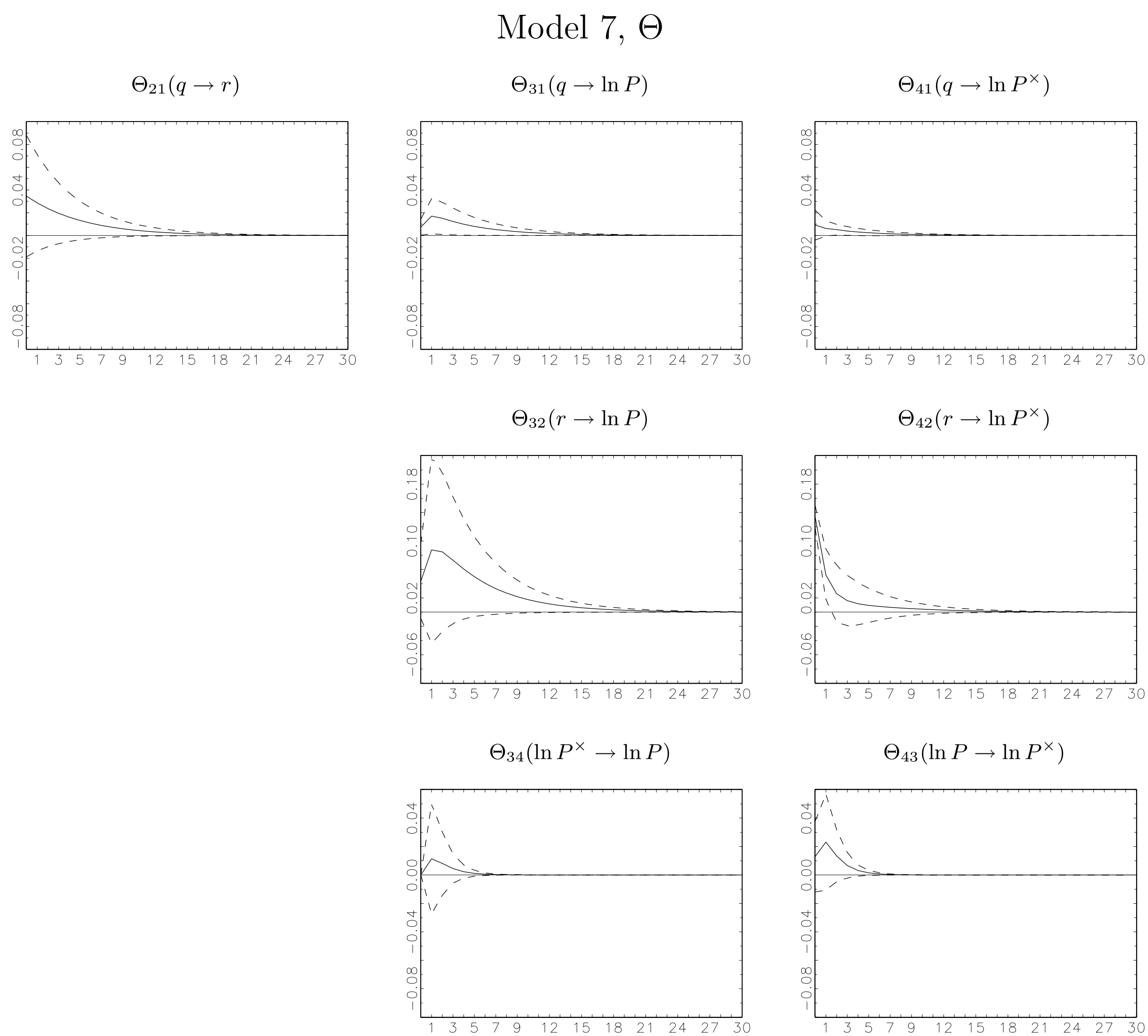


Fig. 12. Impulse response function, Θ_j for Model 7 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Theta}_j$ and the confidence band defined by $\overline{\Theta}_j \pm 2\sigma(\Theta_j)$. The figure does not show the diagonal elements of Θ_j.

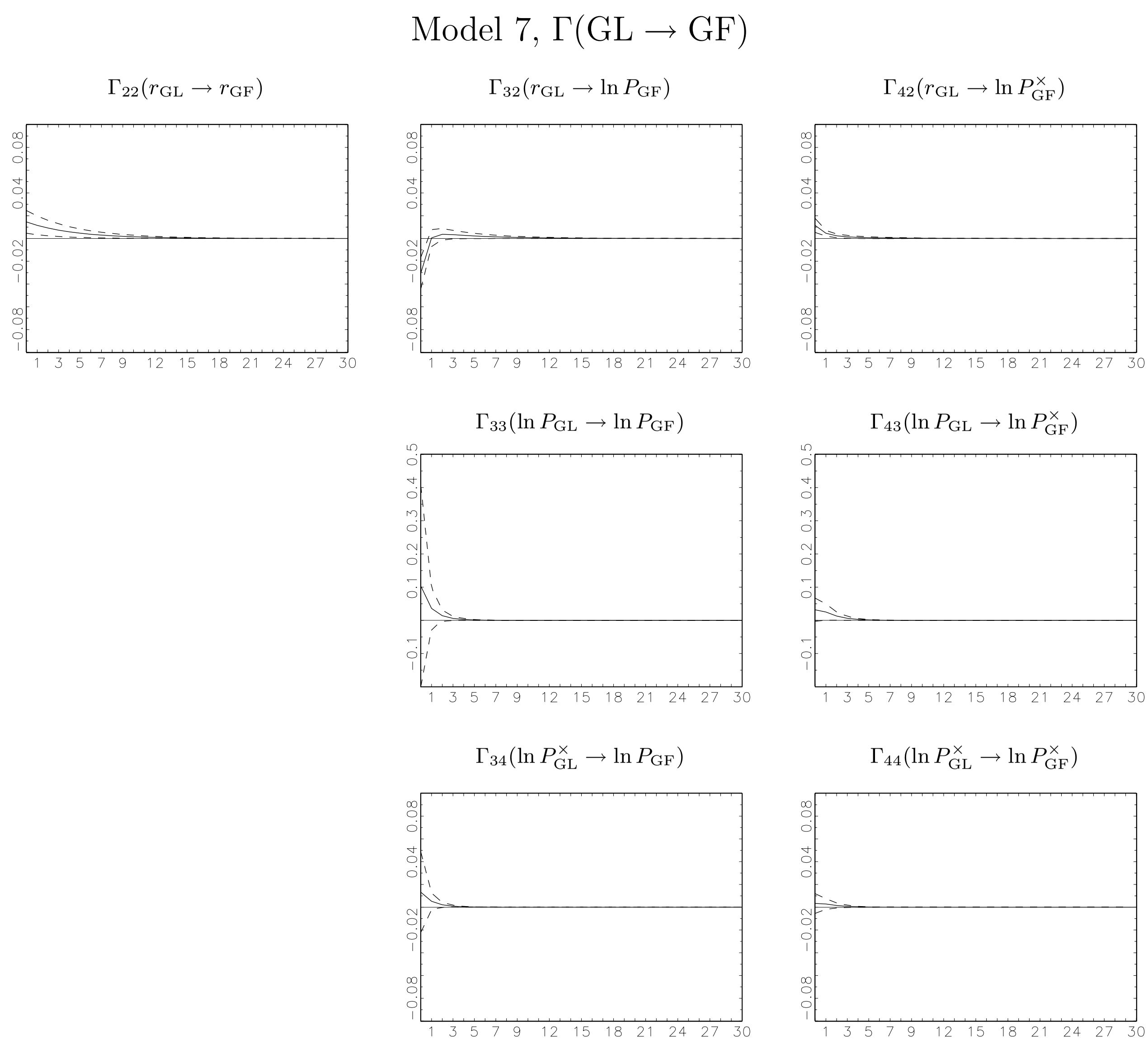


Fig. 13. Dynamic multiplier, $\Gamma_j(\mathrm{GL} \to \mathrm{GF})$ for Model 7 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

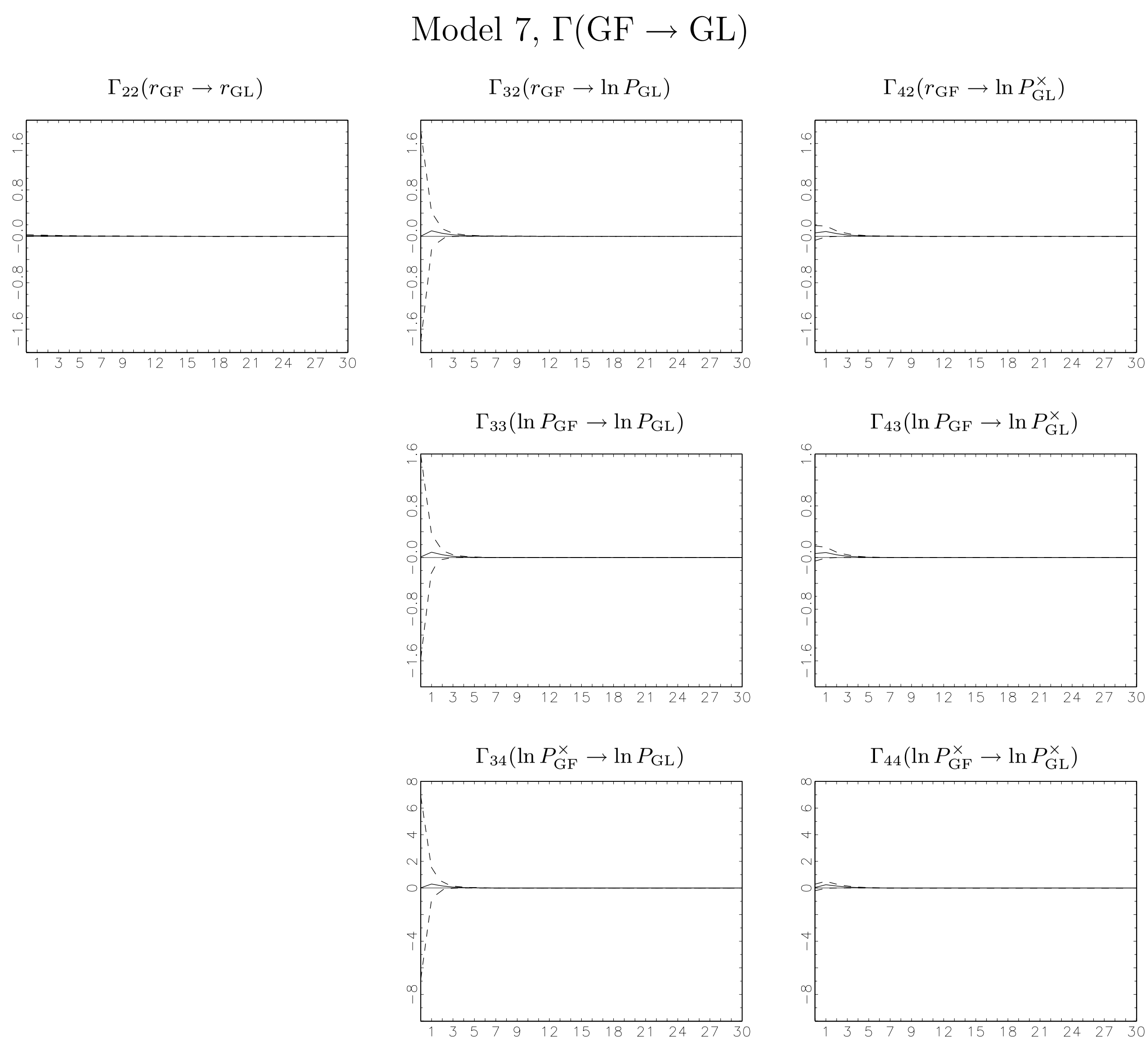


Fig. 14. Dynamic multiplier, $\Gamma_j(\mathrm{GF} \rightarrow \mathrm{GL})$ for Model 7 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

Model 8, Θ

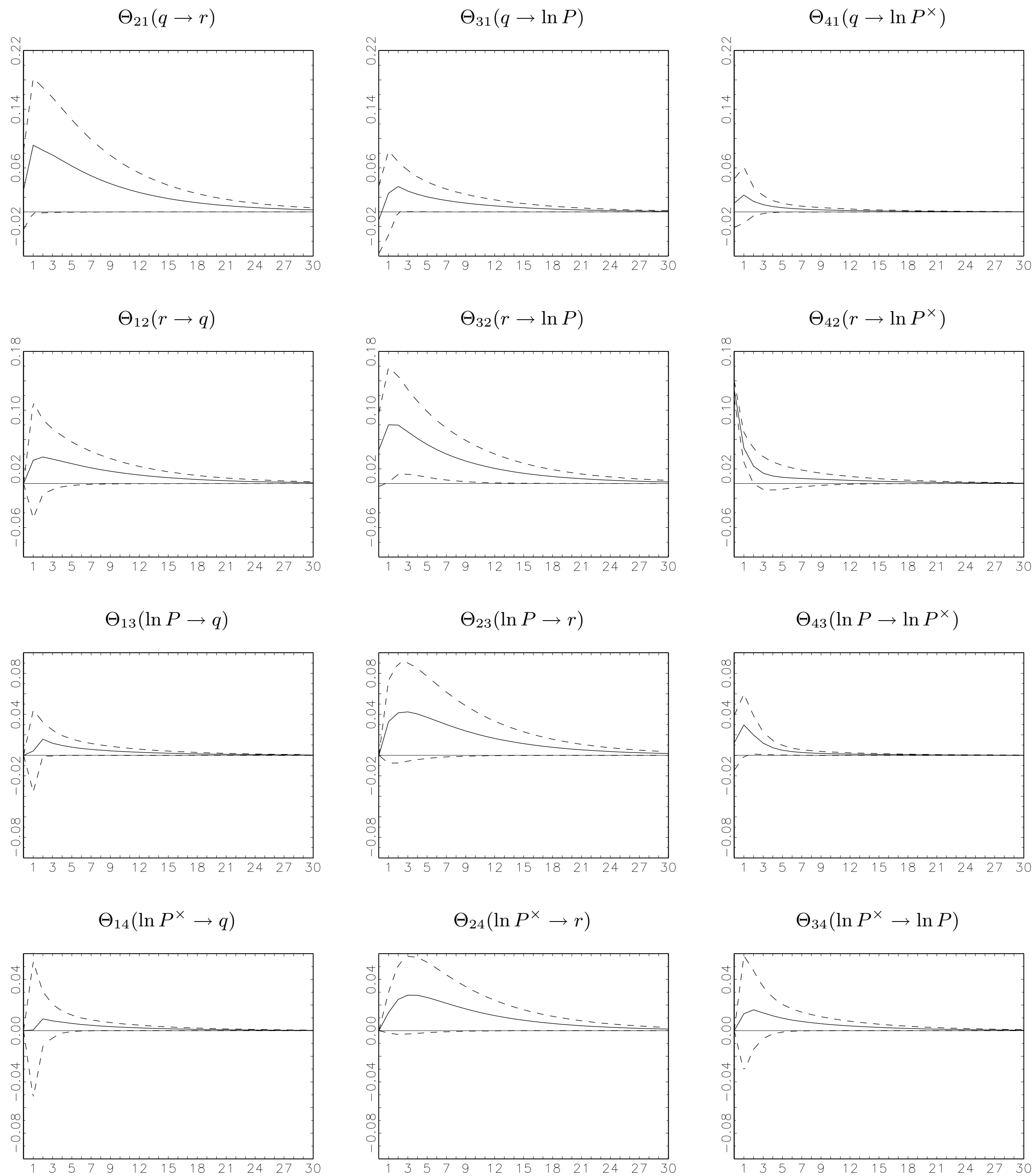


Fig. 15. Impulse response function, Θ_j for Model 8 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Theta}_j$ and the confidence band defined by $\overline{\Theta}_j \pm 2\sigma(\Theta_j)$. The figure does not show the diagonal elements of Θ_j.

Model 8, $\Gamma(\mathrm{GL} \rightarrow \mathrm{GF})$

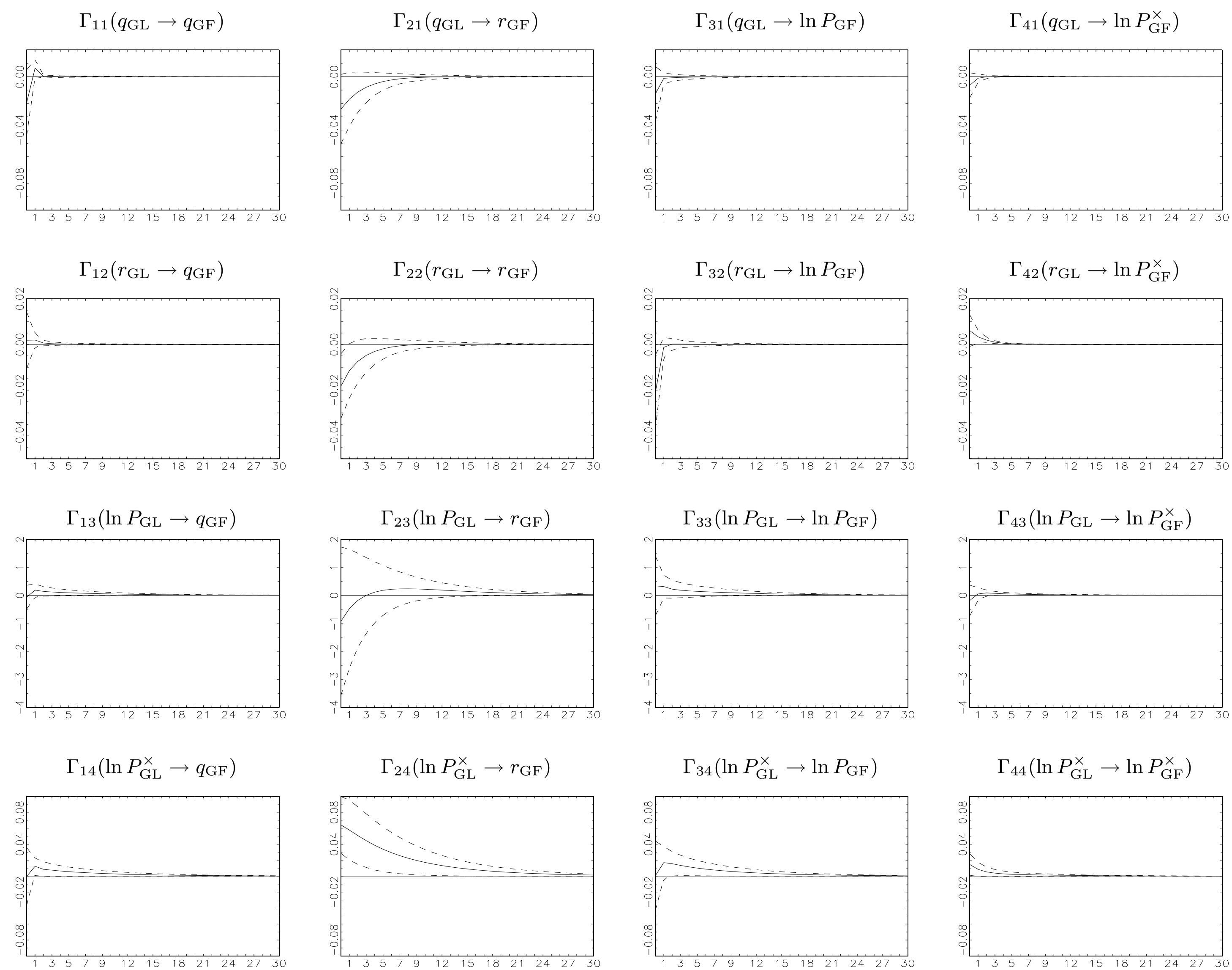


Fig. 16. Dynamic multiplier, $\Gamma_j(\mathrm{GL} \rightarrow \mathrm{GF})$ for Model 8 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

Model 8, $\Gamma(\mathrm{GF} \rightarrow \mathrm{GL})$

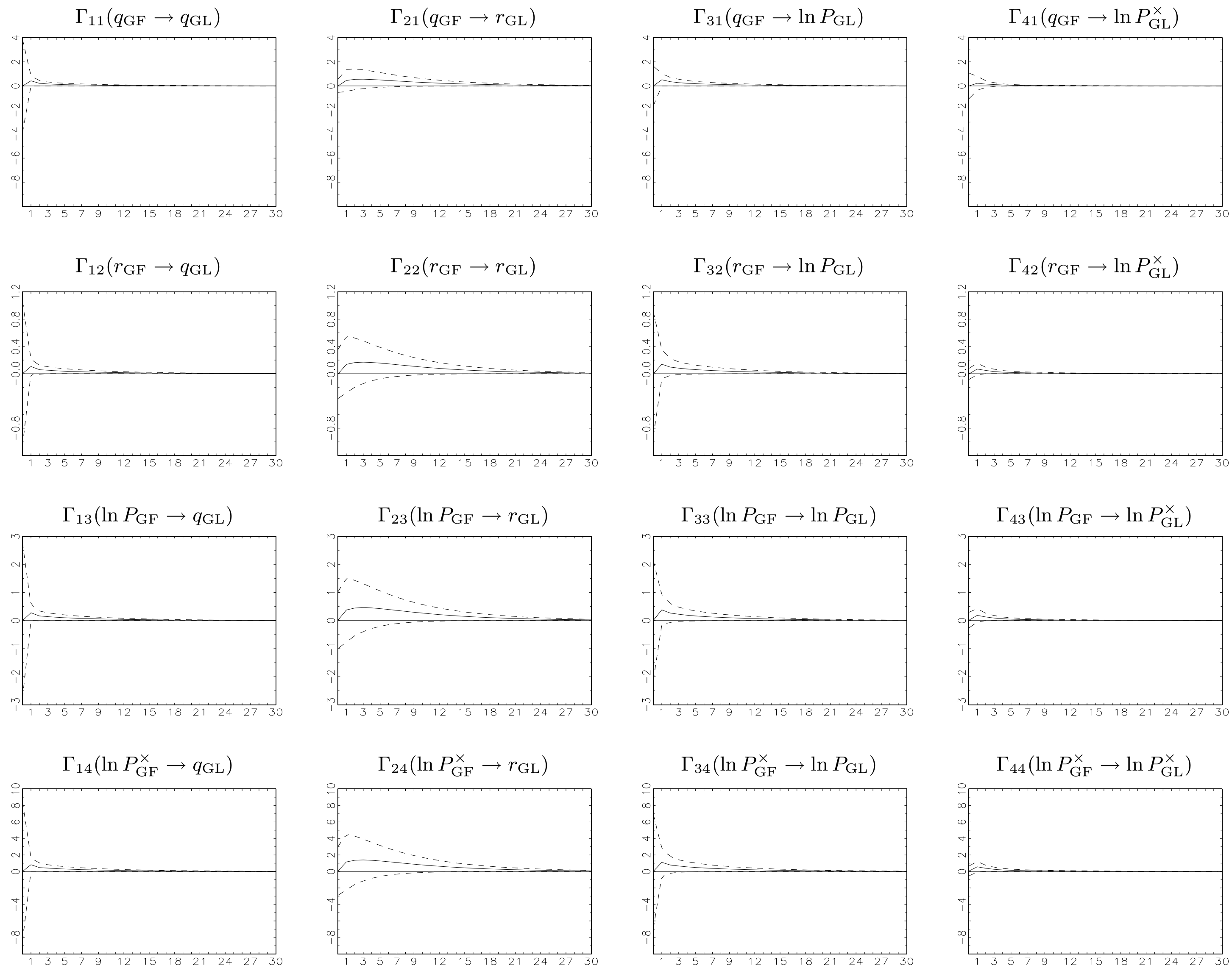
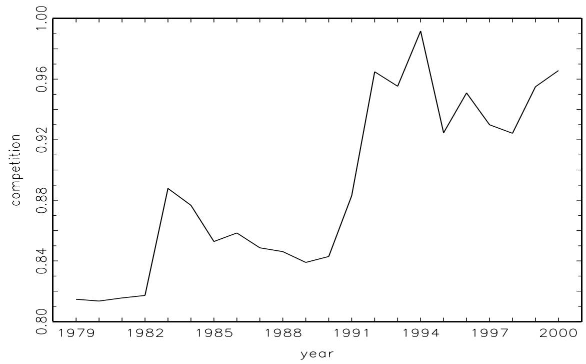
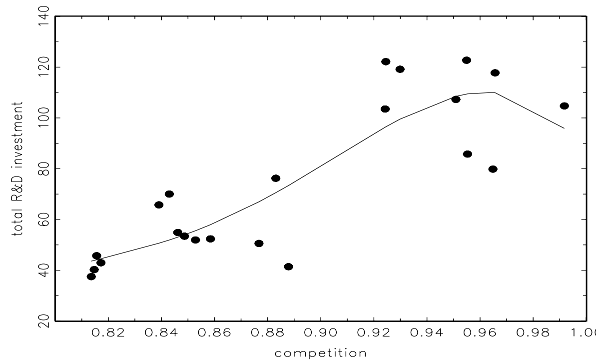


Fig. 17. Dynamic multiplier, $\Gamma_j(\mathrm{GF} \rightarrow \mathrm{GL})$ for Model 8 for $j = 0, \ldots, 30$ leads.
Notes: Each figure shows $\overline{\Gamma}_j$ and the confidence band defined by $\overline{\Gamma}_j \pm 2\sigma(\Gamma_j)$.

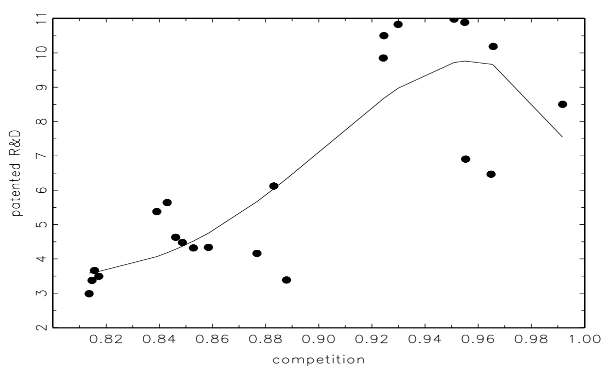
18a) Competition over the period 1979 to 2000



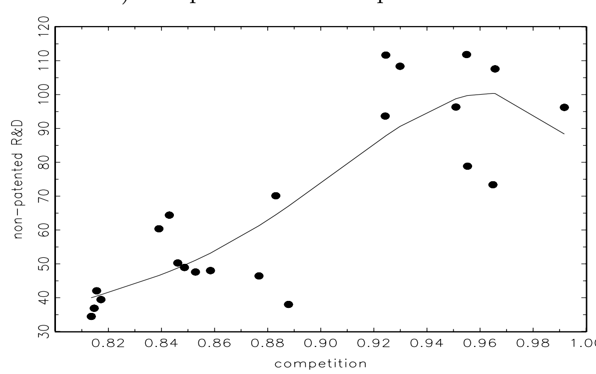
18b) Competition and total R&D investment



18c) Competition and patented R&D



18d) Competition and non-patented R&D



18e) Common competitive factor and competition

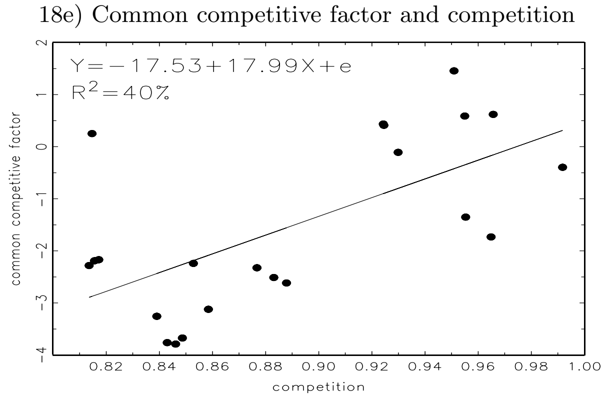


Fig. 18. Competition and innovation.

Notes: Competition is CO_t; total R&D investment is $(1/111)\sum_{i=1}^{111}\tilde{P}_{it}^{o}$; patented R&D is $(1/111)\sum_{i=1}^{111}\tilde{P}_{it}$; non-patented R&D is $(1/111)\sum_{i=1}^{111}\tilde{P}_{it}^{o}(1-\tilde{P}_{it}^{*})$; common competitive factor is l_t^{*}. Panels 18b) to 18d) present the fourth-order polynomial regression least squares estimates of the dependent variable over the period $t = 1979, \ldots, 2000$.