An improved theoretical ground for the linear feedback model and a new indicator*

Yoshitsugu Kitazawa (Kyushu Sangyo University)[a]

December 14, 2012

Abstract

This paper describes a lucid theoretical ground for the linear feedback model proposed by Blundell et al. (2002) and further proposes the indicator on the initial knowledge storage in the framework of the linear feedback model. The values of the indicator are calculated with the estimation results conducted by Blundell et al. (2002). Further, the GMM estimations of the linear feedback model are conducted by using the stationarity moment conditions customized to needs of count panel data, in order to calculate the values of the indicator.

JEL classification: C23; C25; O30

Keywords: linear feedback model; knowledge production; initial knowledge storage; patents-R&D relationship; GMM

1. Introduction

Blundell et al. (2002) propose the linear feedback model (hereafter, LFM) with the aim of incorporating dynamics into the patent generating equation (which is of the count

* Discussion Paper Series, Faculty of Economics, Kyushu Sangyo University Discussion Paper, December 2012, No. 58 Errata information: http://www.ip.kyusan-u.ac.jp/J/kitazawa/ERRATA/errata_lfmni.html

[a] Department of Economics, Kyushu Sangyo University, 3-1 Matsukadai 2-chome, Higashi-ku, Fukuoka, 813-8503, Japan. E-mail: kitazawa@ip.kyusan-u.ac.jp

panel data specification).[1] The treatment of the lagged count variables in the LFM makes us dexterously avoid the problems associated with the colossal-integer-valued and zero-valued dependent variables.[2] Blundell et al. (2002) also give a pioneering economic interpretation for the LFM. However, according to their interpretation, the infinite time span is assumed (which might be supposedly tolerable if we use the pre-sample mean (PSM) estimator proposed by Blundell et al. (1999, 2002) to be hereafter described) and it is felt that the assumption and interpretation for the disturbance are nebulously described.[3] The former has a conflict with the specification in short-run time span in that the model may change in long-run span, while as for the latter it is meaningful in terms of consolidating the theoretical ground of the LFM to make the haziness of the assumption and interpretation clear. To circumvent these problems, we propose a new theoretical ground for the LFM, acceding to their interpretation. In addition, we propose the indicator roughly measuring how large the initial knowledge storage procured beforehand is in comparison with the knowledge production level. We can calculate this indicator by using the persistence parameter in the LFM, which we estimate employing panel data set with span being short. More specifically, if we just obtain the estimate of the persistence parameter using short

[1] The LFM grows out of the integer-valued autoregressive model for the time series Poisson count model developed by Al-Osh and Alzaid (1987), McKenzie (1988), Alzaid and Al-Osh (1987), and Gin-Guan and Yuan (1991). Not to mention Blundell et al. (2002), the estimations of the LFM are conducted by Cincera (1997), Prabhu et al. (2005), Salomon and Shaver (2005), Uchida and Cook (2007), Abdelmoula and Bresson (2008), Gurmu and Pérez-Sebastián (2008), Lucena (2011), Blume-Kohout (2012), and Gallié and Legros (2012) on the patents and/or innovations production. In addition, Damijan et al. (2007) applies the LFM to the analysis about the decision on number of foreign affiliates by the manufacturing firms.

[2] The alternative specifications avoiding the problems are proposed by Crépon and Duguet (1997) and Blundell et al. (1999).

[3] Looking at Blundell et al. (2002), Abdelmoula and Bresson (2008), and Windmeijer (2008), the relationship between the disturbance ε_{it} in the patent production and the disturbance u_{it} in the LFM seems to be hazy in light of their assumptions.

panel, we can calculate the rate of the initial knowledge storage to the knowledge production level in the stationary state. The values of the indicator are calculated by using the estimation results by Blundell et al. (2002), which are obtained by using the patent-R&D panel data with respect to US firms. Further, the values of the indicator are calculated by using the GMM estimation results incorporating not only the conventional moment conditions proposed by Chamberlain (1992) and Wooldridge (1997) but also the moment condition newly developed by Kitazawa (2007).

The rest of the paper is as follows. Section 2 describes the new theoretical ground for the linear feedback model in the framework of time series specification. Section 3 presents the new indicator. Section 4 extends the theoretical ground to the framework of panel data specification and calculates the indicator by using the estimation results by Blundell et al. (2002). Section 5 calculates the indicator after estimating the patent generating function, where the GMM estimations are conducted incorporating the conventional and new-look moment conditions for an US patents-R&D panel dataset. Section 6 concludes.

2. New theoretical ground for the LFM

In this section, we propose an accomplished theoretical ground for the LFM, which constructs the LFM more compatibly than the economic interpretation by Blundell et al. (2002). The model specification becomes possible in the short-run time span by incorporating the notion of the initial knowledge storage to be described, while the new interpretation on the disturbances is presented, which will explain their role without contradiction. In this section, the LFM is presented in the time series specification for simplicity. The extension to panel data specification is presented in section 4.

We consider that the firm's patents at time t (y_t) for $t = 1,\ldots,T$ are generated by the sum of the deterministic knowledge stock of the firm at time t (F_t), the contingent knowledge stock of the firm at time t (V_t) and the initial knowledge storage depreciated at time t (S_t):

$$y_t = F_t + V_t + S_t \ . \tag{2.1}$$

The deterministic knowledge stock at time t is represented as the following distributed lag of the current and past depreciated deterministic knowledge production functions:

$$F_t = \sum_{s=0}^{t-1} (1-\delta_f)^s f(X_{(1)t-s}, \ldots, X_{(K)t-s}, C) \ , \tag{2.2}$$

where $f(X_{(1)t-s}, \ldots, X_{(K)t-s}, C)$ is the knowledge production function at time $t-s$ for $s = 0,\ldots,t-1$, with the given K time-varying positive-valued inputs being $X_{(1)t},\ldots,X_{(K)t}$ and the time-unvarying positive-valued input being C, and the depreciation rate for the deterministic knowledge is δ_f .

The contingent knowledge stock at time t is represented as the following distributed lag of the current and past depreciated incidental noises:

$$V_t = \sum_{s=0}^{t-1} (1-\delta_v)^s v_{t-s} \ , \tag{2.3}$$

where v_{t-s} is the incidental noise at time $t-s$ for $s = 0,\ldots,t-1$ and the depreciation rate for the incidental noise is δ_v. The incidental noise is positive if the incident is windfall, while negative if the incident is mischance, and its mean is postulated to be zero. We regard the incidental noise as being a stochastic variable and interpret it as being a kind of knowledge.[4] Allowing for the uncorrelated structures between I_0 and v_t and between C and v_t, the serially uncorrelated disturbances, and the predetermined inputs, we assume that

$$\mathrm{E}[v_t \mid I_0,\, C,\, v^{t-1},\, X_{(1)}^t, \ldots, X_{(K)}^t] = 0\,, \tag{2.4}$$

where $v^{t-1} = (v_0, \ldots, v_{t-1})$ with v_0 being empty and $X_{(k)}^t = (X_{(k)1}, \ldots, X_{(k)t})$ for $k = 1, \ldots, K$.[5]

The initial knowledge storage depreciated at time t is represented as follows:

$$S_t = (1 - \delta_I)^{t-1} I_0 \;\;, \tag{2.5}$$

where I_0 is the initial knowledge storage and δ_I is the depreciation rate for the initial knowledge storage. We define I_0 as the knowledge storage before the period subject to the study on the patent production. We consider that the firm can garner the initial knowledge storage by means of some sort of swift technology transfer as well as by persevering in accumulating the knowledge stock over many years.

If all depreciation rates are same:

$$\delta = \delta_f = \delta_v = \delta_I \;\;, \tag{2.6}$$

equation (2.1) is written as

---

[4] We define that the negative knowledge impedes the patent production, while the positive knowledge promotes it. We consider that if the negative knowledge is endowed for the firm, the noise is negative, while if the positive knowledge is endowed for the firm, the noise is positive. We consider that the negative knowledge is not deterministically achieved.

[5] The implication of the predetermined inputs is that the incidental noise at time t is uncorrelated with the transformations of the inputs up to time t, but correlated with those after time t. An illustration of the validity of assuming the predetermined inputs is that the increase of the patents production due to the windfall incident would provide impetus for the R&D in future, which would results in the increase of the R&D expenditures.

$$y_t = \sum_{s=0}^{t-1}(1-\delta)^s f(X_{(1)t-s}, \ldots, X_{(K)t-s}, C) + \sum_{s=0}^{t-1}(1-\delta)^s v_{t-s} + (1-\delta)^{t-1} I_0 . \quad (2.7)$$

Accordingly, subtracting (2.7) dated $t-1$ multiplied by $1-\delta$ from (2.7) dated t, we can obtain the following dynamic patent production function:

$$y_t = (1-\delta) y_{t-1} + f(X_{(1)t}, \ldots, X_{(K)t}, C) + v_t \;, \quad (2.8)$$

Further, specifying the knowledge production function dated t as the following Cobb-Douglas type production function:

$$f(X_{(1)t}, \ldots, X_{(K)t}, C) = C \prod_{k=1}^{K} X_{(k)t}^{\beta_{(k)}} \;, \quad (2.9)$$

where $\beta_{(k)}$ is the parameter corresponding to the input $X_{(k)t}$, and letting $\gamma = 1-\delta$, $x_{(k)t} = \log X_{(k)t}$ and $c = \log C$, the LFM is generated:

$$y_t = \gamma\, y_{t-1} + \exp\left(c + \sum_{k=1}^{K} \beta_{(k)}\, x_{(k)t} \right) + v_t \;, \quad (2.10)$$

where it follows from (2.4) that

$$\mathrm{E}[v_t \mid y_1, c, v^{t-1}, x_{(1)}^t, \ldots, x_{(K)}^t] = 0 \,, \quad (2.11)$$

with $x_{(k)}^t = (x_{(k)1}, \ldots, x_{(k)t})$ for $k = 1, \ldots, K$, by noting that the natural logarithm function is bijective for $\mathrm{R}^+ \to \mathrm{R}$ with R^+ and R being the positive real numbers and the real numbers respectively.

An implication based on this theoretical ground is that once the firm procures a massive amount of the initial knowledge storage by some sort of technology transfer, the firm can generate the technological innovations as typified by the patents in abundance, even if the knowledge production level of the firm is meager, as is seen from equation

(2.7).[6] This possibility is intensified if the depreciation rate is small. Further, since the input factors are assumed to be predetermined as is seen from (2.4), we can have a scenario in which the innovations to be generated affect the input factors and therefore the firm could generate the innovations successively.

3. New indicator

We assume that the expected level of the knowledge production function in the stationary state is

$$\mathrm{E}[f(X_{(1)t},\ldots,X_{(K)t},C)] = H\,, \tag{3.1}$$

where stochastic factors can be took into consideration in the inputs $X_{(1)t},\ldots,X_{(K)t}$. Then, the expected number of patents issued in the stationary state is obtained from the dynamic patent production function (2.8) as follows:

$$\mathrm{E}[y_t] = (1/\delta)\,H\,. \tag{3.2}$$

Taking expectation of (2.7) and allowing for (3.1) and (3.2), it follows that

$$(1/\delta)H = H\sum_{s=0}^{t-1}(1-\delta)^s + (1-\delta)^{t-1} I_0^E\,, \tag{3.3}$$

where I_0^E is the mean of the probability distribution that underlies the (realized) initial knowledge storage I_0 (i.e. $I_0^E = \mathrm{E}[I_0]$ implying what we call the background

[6] A typical example is the furious advancement of rocket technologies by the US and USSR after World War II. They raced to develop new rocket technologies each other after expropriating German rocket technologies and immigrating German human resources involved in the technologies into their lands. Their achievement of immense amount of innovations in the field of the rocket technology can be said to be due to the acquirement of a volume of exquisite German technologies as their initial knowledge storages. In the empirical literature, Park and Lee (2006) find that the technological catching-up is more likely to happen in the technological classes with more initial stock of knowledge.

measured by the mean, that gives birth to I_0).

Introducing $\sum_{s=0}^{t-1}(1-\delta)^s = (1-(1-\delta)^t)/\delta$ for (3.3), we can write the rate of the initial knowledge storage to the knowledge production level in the stationary state, as a function of the depreciation rate. That is,

$$\omega = I_0^E / H = (1/\delta) - 1 \quad . \tag{3.4}$$

The index ω roughly measures how large the initial knowledge storage is, compared to the current knowledge production level. We can see from equation (3.4) that the lower depreciation rate results in the larger initial knowledge storage relative to the expected knowledge production level for each period.

Equation (3.4) is also written as follows by using $\gamma = 1-\delta$:

$$\omega = (1/(1-\gamma)) - 1 \quad . \tag{3.5}$$

Judging from equation (3.5), the larger persistence of the patent production (the larger γ) implies that the patents are generated from the larger initial knowledge storage relative to the expected knowledge production level for each period, if the count variables of the patents are stationary. In addition, the relationship between ω and γ is displayed in Figure 1, which shows that the larger γ is also associated with the steeper rise of ω. Simply stated, the higher persistence of patent production is raised by the considerably larger size of the initial knowledge storage relative to the knowledge production level, when the number of patents y_t is in the stationary state.

4. Extension to panel data

In panel data specification, the variables y_t, F_t, V_t, S_t, $x_{(k)t}$ for $k = 1,\ldots,K$, v_t, and I_0 in previous sections are replaced by y_{it}, F_{it}, V_{it}, S_{it}, $x_{(k)it}$ for

$k=1,\ldots,K$, v_{it} , and I_{i0} respectively, where index i denotes the firm i with $i=1,\ldots,N$ and N being number of firms. Further, C and c are replaced by C_i and c_i , both of which control for the individual heterogeneity. Using the replaced variables, the following panel data version LFM is obtained instead of (2.10):

$$y_{it} = \gamma\, y_{i,t-1} + \exp\left(c_i + \sum_{k=1}^{K} \beta_{(k)}\, x_{(k)it} \right) + v_{it}, \qquad \text{for } t=2,\ldots,T, \quad (4.1)$$

where the parameters in the Cobb-Douglas production function $\beta_{(k)}$ for $k=1,\ldots,K$ and the persistence parameter γ (and therefore the depreciation rate δ) are common among all firms and time periods and c_i is regarded as the fixed effect. Accordingly, the assumption (2.11) is replaced by

$$\mathrm{E}[v_{it} \mid y_{i1}, c_i, v_i^{t-1}, x_{(1)i}^t,\ldots,x_{(K)i}^t] = 0, \qquad \text{for } t=2,\ldots,T, \quad (4.2)$$

where $v_i^{t-1} = (v_{i1},\ldots,v_{i,t-1})$ with v_{i0} being empty and $x_{(k)i}^t = (x_{(k)i1},\ldots,x_{(k)it})$ for $k=1,\ldots,K$. In addition to the replacement of variables tailored to the panel data specification as is described above, the variable H and I_0^E in previous section is replaced by H_i and I_{i0}^E . In this panel data specification, ω (i.e. the rate of the initial knowledge storage to the knowledge production level in the stationary state) is common among all firms, as is seen from (3.4).

The panel data that we often confront is of specification with number of firms N being large and time dimension T being small (where asymptotics rely on N instead of T) Since regarding the variables as being (approximately) stationary in short-run span is tolerably reasonable, we can say that the calculation of ω after analyzing the panel data is fairly meaningful.

We calculate ω by using the estimation results by Blundell et al. (2002). The results are obtained by using the patent-R&D panel data of Hall et al. (1986) with respect to 407 US firms from 1972 to 1979. Their specification employs the R&D expenditure as the only input factor, which corresponds to equation (4.1) with $K = 1$ and $x_{(1)it}$ being the natural logarithm of the R&D expenditure at time t for firm i. In this case, we adjust the expression, by setting $x_{it} = x_{(1)it}$ (and accordingly $X_{it} = X_{(1)it}$) and $\beta = \beta_{(1)}$. Finally, the specifications (4.1) and (4.2) reduce to

$$y_{it} = \gamma\, y_{i,t-1} + \exp\left(c_i + \beta\, x_{it}\right) + v_{it}, \qquad \text{for } t = 2,\ldots,T, \quad (4.3)$$

$$\mathrm{E}[v_{it} \mid y_{i1}, c_i, v_i^{t-1}, x_i^t] = 0, \qquad \text{for } t = 2,\ldots,T, \quad (4.4)$$

where $x_i^t = (x_{i1},\ldots,x_{it})$.

Table 1 displays the estimation results by Blundell et al. (2002) and the values of ω calculated from their results (i.e. the estimated values of γ). Some Monte Carlo experiments conducted by Blundell et al. (2002) and Kitazawa (2007) show that the PSM (pre-sample mean) estimator proposed by Blundell et al. (1999, 2002) has some preferable small sample properties if some assumptions are satisfied, while the Level and WG (within group) mean scaling estimators are biased upward and downward respectively.[7] The values of ω calculated by using the PSM estimate is 5.289, while those calculated by using the Level and WG estimates are 8.174 and 0.704. According to the PSM estimate, we can say that in the US firms, the background of the initial

[7] The PSM estimator, whose origin can be traced to Blundell et al. (1995), requires the long pre-sample histories of dependent variables for its consistency. However, the Monte Carlo evidences exhibit that the small sample performance of the PSM estimator is preferable even if the number of pre-sample periods is small. The Level and WG estimators are inconsistent for this specification. The WG estimator is equivalent to the Poisson conditional maximum likelihood estimator proposed by Hausman et al. (1984).

knowledge storage procured before the estimation period is about five times the size of the expected knowledge production level in each year in the estimation period. It seems that the values of ω calculated by the Level and WG estimates highlight the upward and downward biases of their respective estimators.

5. Empirical analysis using GMM

Further, we estimate the LFM (4.3) with (4.4) and then calculate the values of ω based on the estimation results for another count panel data set, where the GMM estimators proposed by Hansen (1982) are used incorporating the new moment conditions proposed by Kitazawa (2007). The dataset is the US one on numbers of patents and R&D expenditures, which is used in Hall et al. (1986) and is composed of 346 firms covering the period 1970-79.[8] The number of firms and the covered period are different from those used in Blundell et al. (2002).

Under the situation where number of firms is large but time dimension is small, the GMM is very useful for obtaining the consistent estimates of the parameters of interest when the explanatory variables are considered to be predetermined for panel data specification with fixed effects.[9] We can construct the moment conditions used for the GMM estimations, based on the assumptions (4.4).

The three types of moment conditions are used for the GMM estimations. The first

[8] The following site displays the zipped file containing the dataset as of December 2012:

http://emlab.berkeley.edu/users/bhhall/pub/data/readme.htm

After unzipping the "patrhgh.zip" file, we can find the "patr7079.dat" file.

[9] An alternative to the GMM estimator is the empirical likelihood (EL) estimator proposed by Owen (1988, 1990, 1991, 2001) and developed by Qin and Lawless (1994), which is expected to have the potential of improving the poor small sample behavior controversial in using the GMM estimator. Hsueh and Lee (2012) utilize the EL estimator to estimate the patent-R&D relationship by using the second dataset of Hall et al. (1986), which is used in the analysis in this section.

is the conventional moment conditions for count panel data model:

$$\mathrm{E}[y_{is}\ (\ \exp(-\beta\,\Delta x_{it})(y_{it}-\gamma\,y_{i,t-1})-(y_{i,t-1}-\gamma\,y_{i,t-2})\ )]=0\,,$$

$$\text{for}\ \ s=1,\ldots,t-2\,;\ \ t=3,\ldots,T\,, \tag{5.1}$$

$$\mathrm{E}[x_{is}\ (\ \exp(-\beta\,\Delta x_{it})(y_{it}-\gamma\,y_{i,t-1})-(y_{i,t-1}-\gamma\,y_{i,t-2})\ )]=0\,,$$

$$\text{for}\ \ s=1,\ldots,t-1\,;\ \ t=3,\ldots,T\,, \tag{5.2}$$

both of which are the moment conditions based on the quasi-differencing transformation by Chamberlain (1992) and Wooldridge (1997).[10] The second is the additional moment conditions nonlinear with respect to γ:

$$\mathrm{E}[(\ \exp(-\beta\,\Delta x_{i,t-1})(y_{i,t-1}-\gamma\,y_{i,t-2})-(y_{i,t-2}-\gamma\,y_{i,t-3})\ )\ \ \exp(-\beta\,x_{it})\ (y_{it}-\gamma\,y_{i,t-1})]=0\,,$$

$$\text{for}\ \ t=4,\ldots,T\,. \tag{5.3}$$

The third is the stationarity moment conditions

$$\mathrm{E}[\Delta y_{i,t-1}\ \ \exp(-\beta\,x_{it})\ (y_{it}-\gamma\,y_{i,t-1})]=0\,,\qquad \text{for}\ \ t=3,\ldots,T\,, \tag{5.4}$$

$$\mathrm{E}[\Delta x_{it}\ \ \exp(-\beta\,x_{it})\ (y_{it}-\gamma\,y_{i,t-1})]=0\,,\qquad \text{for}\ \ t=2,\ldots,T\,, \tag{5.5}$$

both of which are available when the moment generating function for x_{it} is stationary and y_{it} is mean-stationary. The second and third types of moment conditions are proposed by Kitazawa (2007).[11] The conventional, additional nonlinear, and

[10] The moment conditions (5.1) are an extension by Blundell et al. (2002), which are in the framework of the moment conditions based on the quasi-differencing transformation. In the empirical analyses other than those for the LFM, the moment conditions based on the quasi-differencing transformation are also used: Montalvo (1997), Kim and Marschke (2005), and Agarwal et al. (2009), etc.

[11] One type of moment conditions analogous to (5.3) is proposed by Windmeijer (2000) for the case of endogenous explanatory variables, while another type is firstly proposed by Crépon and Duguet (1997) and its variant is subsequently proposed by Kitazawa (2007) for the case of strictly exogenous explanatory variables. In addition, it should be noted that the explanatory variables need to have both negative and positive values when using the moment conditions (5.3), (5.4), and (5.5) (see Wooldridge, 1997 and Windmeijer, 2000).

stationarity moment conditions correspond to the standard moment conditions proposed by Holtz-Eakin et al. (1988) and Arellano and Bond (1991), the additional nonlinear moment conditions proposed by Ahn (1990) and Ahn and Schmidt (1995), and the stationarity moment conditions proposed by Arellano and Bover (1995) and discussed by Ahn and Schmidt (1995) and Blundell and Bond (1998) in the context of the ordinary dynamic panel data model, respectively.[12]

The three GMM estimators using the moment conditions above are investigated for the US dataset: the GMM(qd) estimator using the conventional moment conditions (5.1) and (5.2), the GMM(pr) estimator using the additional nonlinear moment conditions (5.3) together with (5.1) and (5.2), and the GMM(sa) estimator using the stationarity moment conditions (5.4) and (5.5) together with (5.1) and (5.2). Some Monte Carlo experiments carried out by Kitazawa (2007) show that small sample performances of the GMM(pr) and GMM(sa) estimators are superior to that of the GMM(qd) estimator.

We show the estimation results for the span 1972-77 in the dataset. Some descriptive statistics in this span are shown in Table 2. We can say that the basic statistics are not totally different from those used in Blundell et al. (2002). The estimation results are shown in Table 3. Allowing for the upward and downward biases pertaining to the Level and WG estimates respectively, we could consider that the desirable estimates for γ and β lie between 0.914 and 0.325 and between 0.686 and 0.345 respectively. Although the lm1, lm2, and Sargan test statistics say that the moment conditions originating the three types of GMM estimates are all valid, only the GMM(sa) estimates (0.603, 0.617) stay within these realms. We may say that the

[12] Kitazawa (2007) derives these moment conditions from the covariance structure for disturbances in the LFM, expanding the idea of constructing the efficient sets of moment conditions proposed by Ahn (1990) and Ahn and Schmidt (1995) in the context of the dynamic panel data model.

GMM(qd) estimates are admittedly afflicted with the small sample downward bias and the GMM(pr) estimates presumably suffer from some sort of small sample bias. The GMM(sa) estimates are comparatively close to the PSM estimates obtained by Blundell et al. (2002), although the span and the number of firms are different. However, the value of ω calculated by using the GMM(sa) estimate of γ is 1.518, which is much smaller than that calculated by using the PSM estimate.

6. Conclusion

In this paper, we proposed the new theoretical ground for the linear feedback model, which is said to be designed more adroitly than that described by Blundell et al. (2002) despite inheriting their framework. In addition, we proposed the indicator roughly measuring the ratio of the initial knowledge storage to the knowledge production level for each period. Then, we obtained the values of the indicator for the US firms by using the preferable estimation result by Blundell et al. (2002) and by using the GMM estimation result incorporating the stationarity moment conditions proposed by Kitazawa (2007).

References

Abdelmoula, M., Bresson, G (2008). Spatial and technological spillovers in European patenting activities: a dynamic count panel data model. Annales d'Économie et de Statistique 87/88, 167-194.

Agarwal, R., Ganco, M., Ziedonis R.H. (2009). Reputations for toughness in patent enforcement: implications for knowledge spillovers via inventor mobility. Strategic Management Journal 30, 1349–1374.

Ahn, S.C. (1990). Three essays on share contracts, labor supply, and the estimation of models for dynamic panel data. Unpublished Ph.D. dissertation (Michigan State University, East Lansing, MI).

Ahn, S.C., Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. Journal of Econometrics 68, 5-27.

Al-Osh, M.A., Alzaid, A.A. (1987). First-order integer valued autoregressive (INAR(1)) process. Journal of Time Series Analysis 8, 261-275.

Alzaid, A.A., Al-Osh, M.A. (1990). Integer valued pth order autoregressive structure (INAR(1)) process. Journal of Applied Probability 27, 314-324.

Arellano, M., Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277-297.

Arellano, M., Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68, 29-51.

Blume-Kohout, M.E. (2012). Does targeted, disease-specific public research funding influence pharmaceutical innovation? Journal of Policy Analysis and Management 31, 641–660

Blundell, R., Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115-143.

Blundell, R., Griffith, R., Van Reenen, J. (1995). Dynamic count data models of technological innovation. Economic Journal 105, 333-344.

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.

Blundell, R., Griffith, R., Windmeijer, F. (2002). Individual effects and dynamics in count data models. Journal of Econometrics 108, 113-131.

Chamberlain, G. (1992). Comment: sequential moment restrictions in panel data. Journal of Business and Economic Statistics 10, 20-26.

Cincera, M. (1997). Patents, R&D, and technological spillovers at the firm level: some evidence from econometric count models for panel data. Journal of Applied Econometrics 12, 265-280.

Crépon, B., Duguet, E. (1997). Estimating the innovation function from patent numbers: GMM on count panel data. Journal of Applied Econometrics 12, 243-263.

Damijan, J. P., Polanec, S., Prašnikar, J. (2007). Outward FDI and productivity: micro-evidence from Slovenia. World Economy 30, 135-155.

Gallié, E.M., Legros, D. (2012). Firms' human capital, R&D and innovation: a study on French firms. Empirical Economics 43, 581–596.

Gurmu, S., Pérez-Sebastián, F. (2008). Patents, R&D and lag effects: evidence from flexible methods for count panel data on manufacturing firms. Empirical Economics 35, 507-526.

Hall, B.H., Griliches, Z., Hausman, J.A. (1986). Patents and R and D: is there a lag? International Economic Review 27, 265-283.

Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. Econometrica 50, 1029-1054.

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

Holtz-Eakin, D., Newey, W., Rosen, H. (1988). Estimating vector autoregressions with panel data. Econometrica 56, 1371-1395.

Hsueh, S.P., Lee, W.M. (2012). A revisit to the relationship between patents and R&D using empirical likelihood estimation. Economics Bulletin 32, 1208-1214

Jin-Guan, D., Yuan, L. (1991). The integer-valued autoregressive (INAR(p)) model. Journal of Time Series Analysis 12, 129-142.

Kim, J., Marschke, G. (2005). Labor mobility of scientists, technological diffusion and the firm's patenting decision. The RAND Journal of Economics 36, 298-317.

Kitazawa, Y. (2007). Some additional moment conditions for a dynamic count panel data model. Kyushu Sangyo University, Faculty of Economics, Discussion Paper, No. 29. (Paper presented at the 2008 International Symposium on Econometric Theory and Applications, Seoul National University, Korea.)

http://www.ip.kyusan-u.ac.jp/keizai-kiyo/dp29.pdf

http://www.ip.kyusan-u.ac.jp/J/kitazawa/SEMINARS/korea_u080307.pdf (revised)

Lucena, A. (2011). The organizational designs of R&D activities and their performance implications: empirical evidence for Spain. Industry and Innovation 18, 151-176.

McKenzie, E. (1988). Some ARMA models for dependent sequences of poisson counts. Advances in Applied Probability 20, 822-835.

Montalvo, J.G. (1997). GMM estimation of count-panel-data models with fixed effects and predetermined instruments. Journal of Business and Economic Statistics 15, 82-89.

Owen, A.B. (1988). Empirical likelihood ratio confidence intervals for a single functional. Biometrika 75, 237-249.

Owen, A.B. (1990). Empirical likelihood confidence regions. Annals of Statistics 18, 90-120.

Owen, A.B. (1991). Empirical likelihood for linear models. Annals of Statistics 19, 1725-1747.

Owen, A.B. (2001). Empirical likelihood. London, Chapman and Hall.

Park, K.H., Lee, K. (2006). Linking the technological regime to the technological catch-up: analyzing Korea and Taiwan using the US patent data. Industrial and Corporate Change 15, 715-753.

Prabhu, J.C., Chandy, R.K., Ellis, M.E. (2005). The impact of acquisitions on innovation: poison pill, placebo, or tonic? Journal of Marketing 69, 114-130.

Qin, J., Lawless, J. (1994). Empirical likelihood and general estimating equations. Annals of Statistics 22, 300–325.

Salomon, R.M., Shaver, J.M. (2005). Learning by exporting: new insights from examining firm innovation. Journal of Economics and Management Strategy 14, 431-460.

Uchida, Y., Cook, P. (2007). Innovation and market structure in the manufacturing sector: an application of linear feedback models. Oxford Bulletin of Economics and Statistics 69, 557-580.

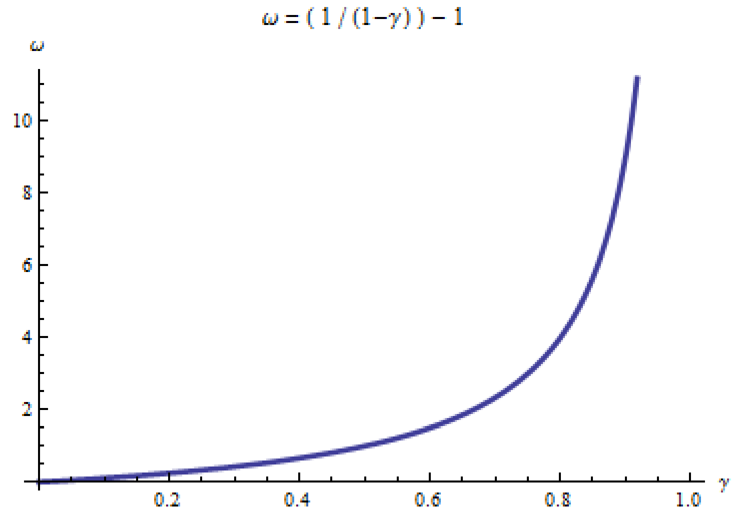
Windmeijer, F. (2000). Moment conditions for fixed effects count data models with endogenous regressors. Economics Letters 68, 21-24.

Windmeijer, F. (2008). GMM for panel count data models. In: László Mátyás and Patrick Sevestre (eds.), The econometrics of panel data. Fundamentals and recent developments in theory and practice, Springer.

Wooldridge, J.M. (1997). Multiplicative panel data models without the strict exogeneity assumption. Econometric Theory 13, 667-678.

Figure 1

Relationship between ω and γ



Note: The graph represents the relationship between the background of initial knowledge storage measured by the mean relative to the expected knowledge production level for each period and the persistence parameter in the LFM, which is equation (3.5).

Table 1

Replication of the estimation results by Blundell et al. (2002) and calculation results of ω

	LEVEL	WG	PSM
γ	0.891	0.413	0.841
β	0.898	0.342	0.506
ω	8.174	0.704	5.289

Notes: (1) Replication is conducted from Table 5 in Blundell et al. (2002). (2) The seven years before the estimation period are used for the calculation of the pre-sample mean. (3) The results by the GMM estimators based on the quasi-differenced transformation proposed by Chamberlain (1992) and Wooldridge (1997) are ruled out, since their estimates of γ are negative suffering from poor small sample performances characteristic of the GMM estimators.

Table 2

Descriptive statistics for the data of patents and R&D

Number of firms	346	346
span	1972-77	1972-77
	Patents	ln R&D
Mean	36.637	0.000
S.D.	75.054	1.960
Minimum	0	−5.062
Maximum	595	5.852
Median	6	−0.271
Proportion of zeros	0.173	
Proportion of positives		0.455

Notes: ln R&D is transformed in deviation from the overall mean (see Windmeijer, 2000).

Table 3

Results for the linear feedback model

N=346 T=6 (span 1972-77)

	γ *t value*	β *t value*	Sargan (df) *p value*	lm1 lm2	ω
Level	0.914	0.686			10.588
	30.765	*13.139*			
WG	0.325	0.345			0.481
	1.177	*0.157*			
GMM(qd)	0.373	0.197	24.60 (22)	−4.594	0.596
	5.300	*1.029*	*0.316*	1.237	
GMM(pr)	0.388	1.013	24.81 (25)	−4.680	0.635
	7.882	*5.213*	*0.473*	1.443	
GMM(sa)	0.603	0.617	38.35 (31)	−5.578	1.518
	13.068	*9.593*	*0.170*	1.509	

Notes: (1) The GMM estimators use the lagged dependent variables dated $t-2$ and before and the lagged explanatory variables dated $t-1$ and before as the instruments for the quasi-differenced equations dated t in the span. (2) No time dummy is included. (3) Sargan is the test statistic of overidentifying restrictions and df is its degree of freedom. (4) The lm1 and lm2 are the test statistics of first-order and second-order serial correlations in quasi-differenced residuals. (5) Results by the two-step estimation are shown. It may be that the t-values for the GMM estimates are a little biased upwards, in light of the Monte Carlo results carried out by Windmeijer (2008) and Kitazawa (2007).