The R&D And Productivity Relationship Of Korean Listed Firms

By: Almas Heshmati and **Hye-Sung Kim**

Abstract

This study analyzes the relationship between R&D investment and the productivity of Korean R&D-engaged firms. An interdependent chain of equations including the propensity to invest, R&D investment and productivity are estimated in a multi-step procedure accounting for selectivity and simultaneity biases. Using Korean firm level panel data of listed firms from 1986 to 2002, we find four main empirical results. First, there is a two-way causal relationship between R&D investment and productivity for Korean listed firms. Second, Chaebol firms were associated with lower R&D growth as well as lower labor productivity growth in comparison to non-Chaebol firms. Third, there was a substantial reduction in growth rates both in R&D investment and labor productivity in 1997-1998, immediately following the Asian financial crisis. Fourth, considering the positive feedback effect from productivity growth to R&D growth, a decrease in R&D investment growth after the Asian financial crisis should have been harmful by further decreasing productivity growth.

Heshmati A, **Kim H.** The R&D and productivity relationship of Korean listed firms. Journal of Productivity Analysis. 2011;36(2):125-142. Publisher version of record available at: https://www.jstor.org/stable/23883831

The R&D and productivity relationship of Korean listed firms

Almas Heshmati · Hyesung Kim

Abstract This study analyzes the relationship between R&D investment and the productivity of Korean R&D-engaged firms. An interdependent chain of equations including the propensity to invest, R&D investment and productivity are estimated in a multi-step procedure accounting for selectivity and simultaneity biases. Using Korean firm level panel data of listed firms from 1986 to 2002, we find four main empirical results. First, there is a two-way causal relationship between R&D investment and productivity for Korean listed firms. Second, *Chaebol* firms were associated with lower R&D growth as well as lower labor productivity growth in comparison to non-*Chaebol* firms. Third, there was a substantial reduction in growth rates both in R&D investment and labor productivity in 1997-1998, immediately following the Asian financial crisis. Fourth, considering the positive feedback effect from productivity growth to R&D growth, a decrease in R&D investment growth after the Asian financial crisis should have been harmful by further decreasing productivity growth.

Keywords R&D investment · Productivity · Selectivity · Chaebol · Simultaneity · Asian financial crisis

JEL classification C33 · E22 · L60 · O32

A. Heshmati (✉)
Department of Food and Resource Economics, College of Life Sciences and Biotechnology, Korea University, East Building Room #217, Anam-dong Seongbuk-gu, Seoul 136-701, Korea
e-mail: heshmati@korea.ac.kr

H. Kim
Department of Political Science, University of Rochester, Harkness Hall, Rochester, NY 14627-0146, USA
e-mail: kim.hyesung@gmail.com

1 Introduction

South Korea has achieved dramatic improvements in its technological capability over the last three decades.[1] R&D investment, an important indicator of investment in technology, has also grown rapidly and steadily since the mid-1990s. According to some R&D-based growth models (Jones 1995; Romer 1990; Grossman and Helpman 1991a, b; Aghion and Howitt 1992, 2009), technological change is critical for economic growth, with much of the change deriving from the R&D efforts of profit-maximizing agents.[2] From a long-term perspective, it is difficult to refute the fact that R&D investment is a key factor in determining the level of knowledge of an economy that can shift the steady-state growth equilibrium upwards. In short, R&D investment is an important factor in determining long-term productivity and growth rates. Korea is no exception, and the improvement in its R&D capability has been an important driving factor of the country's economic growth over the past few decades.

Over the past half-century, South Korea has undergone rapid economic growth and development. In 1962 it was an unremarkable low-income country with a nominal per capita GNI of $110. By 1996, just before the Asian financial crisis, Korea's per capita GNI exceeded $12,000 and in 2007 it was almost $20,000.[3] Key factors thought to have been crucial for such rapid development include cheap, disciplined

[1] A noticeable increase in R&D output occurred during the 1990s. See Sakakibara and Cho (2002).

[2] It can be implied that subsidies to R&D—and possibly other government policies as well—can affect economic growth rates in the long-term.

[3] GNI per capita, Atlas method (current US$), World Development Indicators 2008, The World Bank.

labour, education, an effective government-led industrialization policy and an aggressive export-oriented economic policy. Enormous effort by both government and private firms to build up the country's scientific and technological capacities could be considered one of the key factors in realizing the vision of Korean growth and development. The essence of Korea's remarkable economic growth lies in the increase in its stock of useful knowledge and continuous innovation efforts.

However, some commentators are pessimistic about Korea's government-led industrialization policy, pointing out its strong interventionist nature. Korea's government-led economic management system has been the target of criticism, and recently has even been blamed as one of the main factors that contributed to the Asian financial crisis in 1997. The relatively rapid recovery from the Asian financial crisis and current global economic crisis are strong indications of the effectiveness of the Korean government in crisis management. In carrying out the country's industrial policy, the Korean government selected key industries to nurture, creating entry and exit barriers to protect and support them, as well as providing them with financial and tax benefits. *Chaebols* enjoyed an enormous advantage from such support.[4] Arguably, the *Chaebols*, must also have been related closely to policies that were aimed at building up technological capability.

In fact, until the 1980s, R&D activities in the private sector were carried out within larger-sized firms, mostly *Chaebols*. Although small- and medium-sized enterprises (SMEs) began strengthening their R&D efforts in the 1990s, the core of private sector R&D remained with large-sized firms or *Chaebols*, at least until 1997. Recent empirical results (Oh et al. 2008, 2009) show evidence of size/conglomerate membership effects on productivity growth. Thus, when discussing Korea's technology-based industrial policies, the performance of *Chaebols* and their contributions, both good and bad, cannot be avoided. Throughout this paper, by assuming that R&D is an important contributor to both economic growth at the aggregate level and productivity at the firm level, the difference between *Chaebol* and non-*Chaebol* firms will be examined.

This study empirically examines the relationship between R&D investment and productivity at the firm level. In examining the R&D and productivity relationship, important characteristics of Korean firms such as whether they are categorized as *Chaebol* or *non-Chaebol*, whether they were affected by the Asian financial crisis, and whether they are a large or small sized firm will be carefully taken into consideration. Methodologically, an interdependent chain of equations, including the propensity to invest in R&D, R&D investment, and productivity will be specified. To carry out the empirical analysis of the interdependent chain of equations, a multi-step procedure is used to take care of selectivity and simultaneity biases. First, the propensity towards R&D investments and the R&D investment intensity equations are specified and estimated by the Heckman 2-step method. This step of analysis helps determine the probability of engaging in R&D activities and the intensity of investment in R&D. Second, the relationship between R&D investments and productivity is specified as a system of equations and estimated by an instrumental 2SLS approach and 3SLS method. In this step of the analysis, we control for both sources of potential biases (sample selection and simultaneity equation bias). The empirical analysis is based on Korean firm level panel data of listed firms from 1986 to 2002.

The rest of this paper is organized as follows: Sect. 2 provides a background for R&D activities of Korean listed firms, which is followed by a literature review in Sect. 3. The empirical model is specified in Sect. 4. In Sect. 5, the estimation method is presented, followed by the data and variable descriptions in Sect. 6. Section 7 presents the analysis of the results. In Sect. 8, guidelines for future research are discussed. Section 9 concludes this study.

2 R&D activities of Korean listed firms

In the immediate aftermath of the 1997 Asian financial crisis, the amount of R&D investment declined drastically. The nominal R&D expenditure, which stood at 1218.6 billion won in 1997, had declined to 1,133.7 billion won in 1998. This decline was magnified by the drastic depreciation of the won. However, the impact of the Asian financial crisis on R&D investment seems to have been only

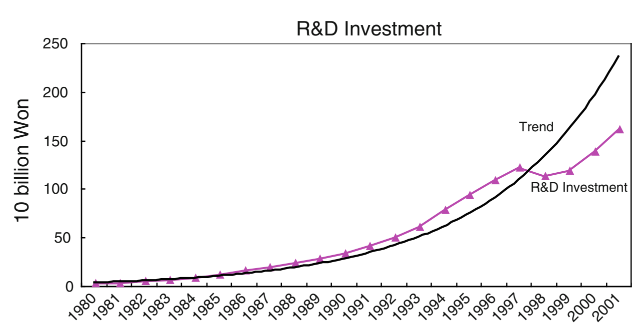

Fig. 1 R&D investment and its trend, Korea 1960–2002

[4] According to the definition by the Korean Fair Trade Commission (KFTC), a chaebol or business group refers to a group of companies that holds more than 30% of its shares owned by some particular individual or by companies governed by those individuals. Since 1987, the KFTC has identified and listed business groups each year.

temporary with an increase in R&D investment in 1999–1,192.2 billion won. (See Fig. 1 for Korea's R&D investment trend).

The size of the reduction in R&D investment also differed between government and private sources. (See Figs. 2 and 3 for R&D investment and its change by sources). Table 1 provides information on government and private R&D expenditures from 1980 to 2001, such as the contributions of each sector and their growth rates. Government R&D expenditures in 1997 and 1998 were 285.1 and 305.2 billion won respectively, which indicates that the government did not reduce its level of R&D expenditures as a result of the Asian financial crisis, although growth seemed to be slowing down. On the other hand, R&D expenditures by the private sector in 1997 and 1998 were 933.5 and 828.5 billion won, respectively, over an 11% decrease. In terms of real values, the decrease in private R&D expenditures in 1998 becomes much more pronounced, falling almost 15%. In addition, government R&D expenditures also decreased approximately 11% in real terms, even though in nominal terms it remained more or less the same (see Table 1 for details). Since it is real growth and not nominal growth in R&D investment that affects productivity, and given that there was a significant reduction in real R&D investment, an important question arises—whether such a significant reduction resulted in a productivity decline in the post-crisis era or whether such an impact was only transitory.

If the growth of technological capability has, in fact, been an important determinant of Korea's growth spurt in the past decades, a significant and rapid decrease in R&D investment or a slowdown in R&D growth may be harmful for the economy. Shin (2004) reported that Korea's GDP growth rate slowed down rapidly after the 1997 financial crisis, and such a slowdown has been attributed to the reduction in R&D investment, especially in the private sector.

However, one of the most notable changes in R&D activities in Korea, especially in the private sector, derives

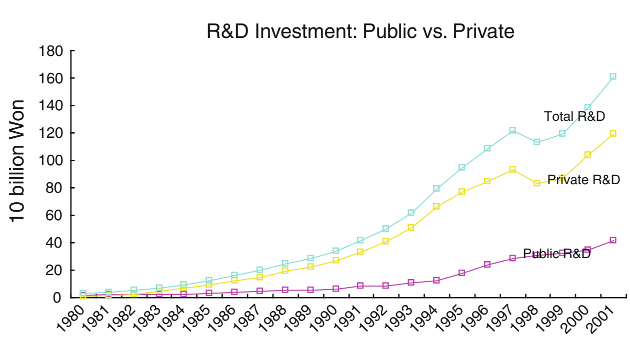

Fig. 2 Development of total R&D in Korea and its private and public components, 1980–2002

from major corporate restructuring of large business groups. Since the 1980s, R&D activities in the private sector were led largely by the large-sized firms. Although SMEs had started to strengthen their R&D efforts in the 1990s by founding research institutes and expanding R&D expenditures, the core of the private R&D was still the large-sized firm, at least until the 1997 Asian financial crisis. For example, in 1997, of the total R&D investment from 2,500 private firms, the top 20 firms accounted for 60% of the total investment. This proportion decreased to 55% in 2000, but it still remained significant. The drop, however, can be attributed to the restructuring of large-sized firms as they cut back on R&D investment or the number of R&D staff in response to the crisis. In contrast, R&D activities by SME firms have increased significantly. R&D investments by SMEs in 2000 almost doubled compared to the number of R&D investments in 1997, while the increase in R&D investment by large-sized firms remained at about 3% for the two periods. The same pattern of contrast between SMEs and large-sized firms is noted in Oh et al. (2008, 2009), Kang and Heshmati (2008) and in a number of Research Papers (for details, refer to Seo (2002)).

Seo (2002) provides a summary of the R&D activities of Korean manufacturing firms that hired at least five employees according to industry-based data for the year 2000. He reported that there were a total of 94,940 Korean manufacturing firms in 2000. The number of firms with fewer than 19 employees, which was the largest proportion of firms, was approximately 20,000. In contrast, only 155 firms had employees over 1,000. 2,908 of the total 94,940 manufacturing firms, a mere 3%, engaged in R&D activities. In particular, among those with fewer than 19 employees, only 548 firms, or less than 1%, were engaged in R&D activities. As employment size increased, the proportion of R&D-active firms also increased, with 55% of firms with 300–499 employees, 78% of firms with 500–999 employees, and 95% of firms with over 1,000 employees engaged in R&D activities. Note, however, that the sample of this study is restricted only to listed firms, and hence a large proportion of firms, approximately 80%, are engaged in R&D activities. Also, approximately 80% are large-sized firms.

One common measure of R&D (for the purpose of comparing R&D sizes) is R&D intensity. In earlier literature, firm size and market structure have been commonly viewed as determining factors of R&D intensity. According to Galbraith (1957) and Demsetz (1969), large-sized firms may be less vulnerable to risks associated with R&D investment and are more likely to use technological innovations as a means of protecting monopoly profits. Furthermore, large firms tend to use technological knowledge gained from technological innovations as a means of

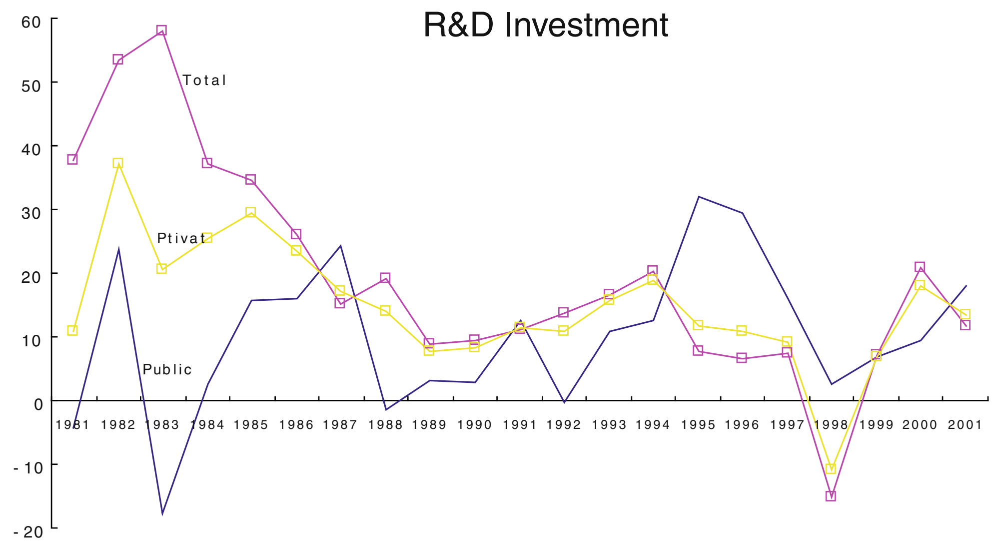

Fig. 3 Changes in R&D investment in Korea and its decomposition, 1980–2002

Table 1 R&D investment in Korea, its distribution and growth rate, 1980–2001, 11,596 observations, unit: 10 billion Won

R&D investment				Contribution		R&D investment growth*		
Government		Private	Total	Government	Private	Government	Private	Total
1980	1.80	1.03	2.83	63.70	36.30	–	–	–
1981	2.02	1.65	3.67	55.00	45.00	−4.40	37.60	10.90
1982	2.64	2.69	5.33	49.60	50.40	23.70	53.50	37.10
1983	2.31	4.51	6.82	33.90	66.10	−17.60	58.10	20.60
1984	2.52	6.56	9.07	27.70	72.30	2.70	37.20	25.50
1985	3.07	9.30	12.37	24.80	75.20	15.80	34.70	29.50
1986	3.74	12.33	16.07	23.30	76.70	16.00	26.00	23.50
1987	4.90	14.95	19.85	24.70	75.30	24.30	15.10	17.20
1988	5.23	19.31	24.54	21.30	78.70	−1.50	19.20	14.10
1989	5.75	22.42	28.17	20.40	79.60	3.10	8.80	7.60
1990	6.51	26.99	33.50	19.40	80.60	3.00	9.50	8.20
1991	8.16	33.43	41.58	19.60	80.40	12.50	11.20	11.50
1992	8.79	41.11	49.89	17.60	82.40	−0.40	13.70	10.90
1993	10.39	51.14	61.53	16.90	83.10	11.00	16.70	15.70
1994	12.60	66.35	78.95	16.00	84.00	12.50	20.30	19.00
1995	17.81	76.60	94.41	18.90	81.10	31.90	7.70	11.60
1996	23.98	84.80	108.78	22.00	78.00	29.50	6.50	10.90
1997	28.51	93.35	121.86	23.40	76.60	15.90	7.30	9.20
1998	30.52	82.85	113.37	26.90	73.10	2.50	−15.00	10.90
1999	32.03	87.19	119.22	26.90	73.10	6.80	7.10	7.00
2000	34.52	103.97	138.49	24.90	75.10	9.40	21.00	17.90
2001	41.87	119.23	161.11	26.00	74.00	18.10	11.70	13.30
1981–1989	–	–	–	26.40	73.60	6.90	32.30	20.70
1990–2001	–	–	–	21.50	78.50	12.70	9.80	10.40
1981–2001	–	–	–	25.70	74.30	10.20	19.40	14.80

* Real growth rate is calculated after being deflated by using GDP deflator

Survey on research and development activities for science and technology, Ministry of Science and Technology 2002 protecting themselves from imitation by other firms. On the other hand, Williamson (1965), Arrow (1974), and Buxton (1985) suggest that as a result of inefficiency associated with monopolistic markets, the innovation process of a monopolist is less active than in competitive firms and that the competitive market structure encourages firms to engage more actively in technological innovations. Some argue that SMEs are more likely to engage in innovations rather than large-sized firms because SMEs are more efficient and faster in decision-making given their simpler and more flexible management structure.

3 Literature review

3.1 R&D and productivity

The idea of how R&D efforts influence productivity levels of firms has received much attention since the pioneering work by Griliches (1979). He later analyzed the data set (the NSF-Census match) containing information on R&D expenditures, sales, employment and other details for approximately 1,000 of the largest manufacturing firms in the US during 1957–1977 (Griliches 1985, 1986). The estimation was conducted in the standard production function framework, augmented by the addition of R&D, capital, and mixed variables.

The findings from his study suggest that R&D continued to contribute to productivity growth in US manufacturing in the 1970s. For a comprehensive survey on a wide range of theoretical and econometric issues dealing with R&D, one can refer to Griliches (1995). Another important piece of work in this area is Griliches and Mairesse (1984), where they estimate an R&D-augmented production function, both in terms of levels and differences, using panel data of 133 large US firms from 1966 to 1977. They find a strong R&D productivity link in the cross-sectional dimension, but the relationship collapses in the time series dimension, which is due to the problems of simultaneity and measurement error. In their study, estimates of the output elasticity of R&D vary depending on model specification, namely from 0.05 to 0.30. A similar range of R&D output elasticities were obtained by Griliches (1986) who computed a marginal product for R&D in US manufacturing as high as 0.62. Cuneo and Mairesse (1984) and Hall and Mairesse (1995) find R&D output elasticities that lie in a range similar to the results found in Griliches and Mairesse (1984).

Seo (2002), who used Korean data, performs a detailed study at the firm level to test empirically whether there was a structural change in R&D activities at the firm level before and after the crisis, how the role of R&D activities has changed in influencing productivity in the context of a model where production factors adjust over time to external shocks, and how and to what extent R&D activities have contributed to labor productivity.

Oh et al. (2008) analyzed TFP growth in Korean manufacturing industries from 1993 to 2003 using both parametric and non-parametric methods The TFP growth rate is decomposed into different components. By classifying the results by industrial sector, time period, geographic location, size class and technology levels they find systematic heterogeneity across a number of time invariant firm characteristics and discuss the underlying causal factors. In another study that Oh et al. (2009) conducted, a comparative analysis on Korean manufacturing plants by size of plant and sources of TFP growth are decomposed into entry, exit, and survival effects of plants, focusing on the pre- and post- Asian crisis periods. Additional survival analyses investigate internal and external determinants of the survival of plants. The results indicate that the exit of SMEs with higher productivity is becoming problematic in the post-crisis period. The improvements in LSEs after the crisis appeared to occur generally in high-technology industrial sectors; SMEs in low-technology industries are suffering from a sluggish market selection process.

3.2 Recent econometric approaches to innovation studies

A growing volume of literature has addressed the analysis of factors influencing the innovative activities of firms over the past two decades. For the details of a recent survey, see Cohen and Levin (1989). Much of this literature focuses merely on an examination of the determinants of the innovative activity of R&D performing firms. However, not all firms are engaged in R&D activity; rather, this activity is restricted to a small proportion of firms. This feature is more common in developing countries.

Studies restricted to samples of firms engaged in formal R&D are, therefore, prone to selection bias. Crepon et al. (1998) and Lööf and Heshmati (2002) have taken into account this selection bias in their studies. In both studies where R&D investment is defined as innovation input and patent count, and innovation share of sales is taken to be innovation output, an innovation equation and a productivity equation are treated as a system. Innovation input is endogenous in the innovation output equation and innovation output is endogenous in the productivity equation. Since there is a major issue of simultaneity, which is likely to interact with selectivity, they take into account both sources of bias. Crepon et al. (1998) show that conventional OLS regressions provide a negligible estimated elasticity of productivity with respect to innovation output because of measurement errors. To deal with error-in-variable problems, instrumental variables are introduced, and the innovation output elasticity is found to increase from about 0 to about 0.30–0.40.

The empirical framework of this study is similar to that of Crepon et al. (1998) and Lööf and Heshmati (2002) in the sense that it takes into account selection bias by utilizing a Probit analysis of innovativeness, a system of innovation input growth and productivity growth equations. However, in contrast to both studies, this study uses innovation input measured by R&D investments as an indicator of innovation rather than innovation output. Thus, this study investigates the link between innovation input measured by R&D investments and the productivity chain, instead of the innovation input-innovation output-and productivity chain. By doing so, it is not possible to capture the impact of innovation output represented by number of patents on productivity, but in fact, the proportion of Korean listed firms that are involved in patenting activities is very small, which implies that patenting activities might not be an accurate representation and a good measure of innovation output. A similar approach using Italian manufacturing firm-level data, where only R&D input is considered, is found in Medda et al. (2005, 2006) while an approach that links R&D spending, process versus product innovation and productivity is found in (Parsi et al. 2006).

Another possibility for adopting innovation output in the study is to use the proportion of innovation sales in the total sales, which can be obtained from the Community Innovation Survey (CIS) data. Using the CIS data will allow us to examine the innovation input–output-productivity link. However, in this case, since the CIS data is cross-sectional, the pattern or impact of innovation over time cannot be observed, which may have significant implications in the case of Korean listed firms. Also, a structural break cannot be detected using the cross-sectional data. Thus, in this study, the link between R&D investments and productivity over the 17 year period between 1986 and 2002 will be examined, taking into account selection and simultaneity biases, which will be adopted from Crepon et al. (1998) and Lööf and Heshmati (2002).[5]

4 Empirical model

This section contains two parts. First, the propensity towards R&D investment and firms' R&D intensity are specified in the generalized tobit model. Second, the R&D investment-productivity relationship is specified in a system of equations and estimated with 2SLS and 3SLS estimation methods.

4.1 R&D investment equation

The probability and intensity of R&D investments will be conducted in the framework of a generalized tobit model (Heckman 1976, 1979) with two equations: the first equation accounts for the fact that the firm is engaged in research activities, and the second one for the magnitude or intensity of these activities.

Let us assume that there exists a latent dependent variable g^*_{it} for the firm i given by the following equation:

$$g^*_{it} = \beta^1_0 + \sum_j \beta^1_j x^1_{jit} + \varepsilon^1_{it} \tag{1}$$

where β is a vector of unknown parameters, x^1_j is a vector of determinants of the incidence of engagement in R&D, and the subscripts j, i, and t denote determinants, firm and time periods respectively. We observe that the firm invests in R&D if g^*_{it} is positive; that is, the dependent variable is binary, namely 0 or 1, which indicates whether or not the firm has a positive R&D investment.

Now, a latent or true intensity of research k^*_{it} for firm i in period t is determined by the following equation:

$$k^*_{it} = \beta^2_0 + \sum_j \beta^2_j x^2_{jit} + \varepsilon^2_{it} \tag{2}$$

where x^2_j is the vector of determinants of R&D investment, β^2_j is the corresponding unknown coefficients vector, and ε^2_{it} is a disturbance that summarizes omitted determinants and other sources of unobserved heterogeneity. If the firm engages in R&D investments (that is, if g^*_{it} is larger than 0), k^*_{it} is equal to k_{it}, which is the actual R&D investments of firm i in period t. Note that the explanatory variables in the two Eqs. (1) and (2), do not have to be the same. However, without a good a priori reason to do otherwise, let us assume $x^1_j = x^2_j$ (Crepon et al. 1998) in both equations. Finally, because k^*_{it} is only observable when g^*_{it} is larger than the industry threshold, the relation can be estimated using a generalized tobit model. However, since the two equations are specified using the same explanatory variables, a tobit approach produces the same results.

4.2 System of R&D and productivity equations

The question of whether R&D investment drives a firm's productivity has been studied for a long time. The theoretical framework for this study is an augmented production function with standard input variables, namely capital, labor, and R&D investment. The equation to be estimated

[5] In constructing the dataset to be used, the listed firms between 1986 and 2002 were selected. However, the observations in 1986 were dropped in estimating the system of equations because transforming data into annual growth rate terms required creating lag variables. Hence, the length of period examined in the empirical analysis for the system of equations was 16 years.

is written in per-employee terms, which is represented by the following equation:

$$q_{it} = \beta_0 + \sum_m \beta_m x_{mit} + \beta_k k_{it} + \varepsilon_{it} \quad (3)$$

where the lower-case letters denote the logarithm of variables. The left-hand variable, q_{it}, is labor productivity, defined as the logarithm of value-added output per employee, and x is an m ($m = 2$) vector of the factors of productivity other than R&D investment, including the logarithm of physical capital per employee and the logarithm of labor. The variable k_{it} is the logarithm of R&D investments per employee, β_m is the elasticity of labor productivity with respect to a vector of inputs, β_k is the elasticity of labor productivity with respect to change in R&D investment per employee, and ε_{it} is the random error term.

What has been often ignored in this specification, however, is the feedback effect from productivity to R&D investment (β_q). This study therefore tries to capture this feedback effect to see if R&D investments increase with productivity ($\beta_q > 0$) in addition to examining the R&D effects of productivity (β_k). More specifically, the model is represented by the following equations:

$$k^*_{it} = \beta^3_0 + \beta_q q_{it} + \sum_j \beta^3_j x^3_{jit} + \varepsilon^3_{it} \quad (4)$$

$$q_{it} = \beta^4_0 + \beta_k k_{it} + \sum_m \beta^4_m x^4_{mit} + \varepsilon^4_{it} \quad (5)$$

where q represents labor productivity (defined as log-value added output per employee) and x^4 is an m vector of the determinants of productivity other than R&D investments, including the logarithm of physical capital per employee and the logarithm of labor defined as the number of employees. The x^3 vector of determinants of R&D investment are the same as x^2 mentioned above, and k_{it} refers to the logarithm of R&D investments per employee.

The main interest of the empirical analysis of this paper is to see whether the slowdown in R&D growth has, in fact, been harmful for the economy by reducing labor productivity growth, and therefore it is more appropriate to transform the level variables into an annual growth rate term.

Transforming q, k, and x^4_m in Eqs. (4) and (5) into an annual growth rate gives us:

$$\dot{k}^*_{it} = \beta^3_0 + \beta_q \dot{q}_{it} + \sum_j \beta^3_j x^3_{jit} + \varepsilon^3_{it} \quad (6)$$

$$\dot{q}_{it} = \beta^4_0 + \beta_k \dot{k}_{it} + \sum_m \beta^4_m \dot{x}^4_{mit} + \varepsilon^4_{it} \quad (7)$$

By estimating all variables in annual growth terms in Eq. (7), we are assuming that the impact on labor productivity growth of the factors not included in the equation is negligible while the impact of the factors included in the equation is substantial. Also, estimating the equation in terms of growth rate allows us to control for unobserved firm-specific but time-invariant effects, which might differ substantially. These are eliminated following the growth rate transformation of the variables.

5 Estimation method

As mentioned previously, the number of firms that engage in R&D or innovation activities across the whole manufacturing sector is very small.[6] Since all of the firms that are not engaged in R&D activities have 0 in terms of R&D expenditures (if $k^* \leq 0$, k is observed as 0), such a sample of data is left-censored. Since only a certain fraction of firms are engaged in formal R&D activities, if an ordinary least squares (OLS) estimation is applied to estimate the model, the selection rule of R&D engagement, or more formally, the propensity for R&D investment is ignored. Thus the OLS estimates are biased and inconsistent.

In order to correct for possible sample selection bias the estimation consists of two separate parts. First, to examine the true intensity of R&D investments, a Heckman (1979) two-stage procedure (often called the Heckit model) is applied to estimate the generalized tobit model, which has been specified in Eqs. (1) and (2). Second, to investigate the R&D growth effect on productivity growth and the productivity growth effect on R&D, a system of equations is simultaneously estimated, which is specified by Eqs. (6) and (7). The 2SLS and 3SLS estimation methods are applied to estimate the system of equations.

5.1 R&D investment equation: Heckman's two-stage estimation

In Sect. 4.1 we specified the R&D investment equation in the framework of a generalized tobit model (Eqs. 1 and 2). Let us recall that Eq. (1) contains some selection (or decision) criteria. In this R&D decision equation, or selection equation, the explanatory variables that are included to capture their impacts on the likelihood of engagement in R&D are outsourcing, profitability, indebtedness, and firm size. To control for the time (technology) and industry effects, trend and industry dummies are also included.

Heckman (1976, 1979) has devised a simple two-stage estimation process that yields consistent estimates. In the first stage, we estimate the corrective term, the inverse

[6] Considering the significance and contributions of listed firms, the focus of this study is on listed manufacturing firms. A large proportion of these firms are engaged in R&D activities.

mills ratio λ_{it}, by utilizing a probit model (see Eq. 1). In the second stage of the estimation, the estimated $\hat{\lambda}_{it}$ is added as an additional explanatory variable in the OLS regression of Eq. (2), that is:

$$k_{it} = \beta_0^2 + \sum_j \beta_j^2 x_{jit}^2 + \hat{\lambda}_{it} + \varepsilon_{it}^2$$

5.2 System of R&D and productivity growth: alternative estimation methods

In Sect. 4.2, the R&D investment growth and productivity growth equations were specified as a system in Eqs. (6) and (7). In this model's specification, we encounter two sources of potential bias if we conduct an OLS estimation method. The first source is sample selection bias, which was discussed in Sect. 3.1. The second source of bias is simultaneous equation bias, which arises from the fact that regressors of each equation might fail to be independent variables because they are dependent variables in a simultaneous system.

In Eq. (6), the R&D investment growth equation, labor productivity growth $(\dot{q})$ is a dependent or endogenous variable since $\dot{q}$ is a function of $\dot{k}$, and growth in R&D investments per employee can be found in Eq. (7). Thus, $\dot{q}$ by Eq. (6) is a function of ε_{it}^3 and therefore labor productivity growth $\dot{q}$ depends on ε_{it}^3. Similarly, growth in R&D investments per employee, $\dot{k}$, depends on ε_{it}^4. Using the OLS estimation method to estimate these equations separately produces biased estimates. One solution to this problem is to replace productivity growth and R&D investment growth on the right-hand side of the two equations with their predicted values.

Taking into account the two possible sources of bias described above, the system of equations to be estimated can be specified as follows:

$$\dot{k}_{it} = \beta_0^3 + \beta_q^3 \hat{\dot{q}}_{it} + \sum_j \beta_j^3 x_{jit}^3 + \varepsilon_{it}^3 \quad (8)$$

$$\dot{q}_{it} = \beta_0^4 + \beta_k^4 \hat{\dot{k}}_{it}^4 + \beta_\lambda^4 \hat{\lambda}_{it}^4 + \sum_m \beta_m^4 x_{mit}^4 + \varepsilon_{it}^4 \quad (9)$$

where $\hat{\dot{k}}$, $\hat{\dot{q}}$ and $\hat{\lambda}$ denote the predicted values of R&D per employment growth, productivity growth, and an estimated inverse Mills ratio (from Eq. 3), reflecting differences in the probability of a firm engaging in R&D investments, respectively. Note also that the dependent variable $\dot{k}$ in Eq. (8) is growth in actual or observed R&D investment per employee, and that x_j^3 and x_m^4 are j and m vectors of variables explaining variations in R&D investment per employee and labor productivity of firm i in period t. In addition to the estimation of R&D growth effects on labor productivity growth (β_k), the system accounts for feedback from labor productivity growth in relation to R&D investment growth (β_q) and also corrects for possible selection bias (β_λ).

The system of equations in (8) and (9) can be estimated in several alternative ways described below.

1. OLS can be used in estimating each single equation separately, but as previously discussed, this method is not advised because it neglects the interdependence of the R&D investment growth and labor productivity growth equations. Hence, simultaneity bias cannot be corrected when OLS method is employed.
2. Two-stage least squares (2SLS) is an instrumental variable approach where instruments are used instead of the dependent variables on the right-hand side of the equations. An instrumental regression is a regression of the dependent regressors on a set of instrumental variables, which can be any independent variables useful for predicting the dependent regressors. Then the predicted values computed in this first step are included in the second stage, which is estimated by OLS estimation. Here the problem is finding suitable instruments. The instruments must be: (i) exogenous, (ii) highly correlated with the dependent variables, but (iii) not correlated with the residuals in each equation, and finally, (iv) the measurement error might be large due to the use of mediocre instruments.
3. A third alternative is two-stage least squares (2SLS) with lag-dependent variables. This method involves the use of lag-dependent variables as instruments for R&D investment (k_{it-1}) and productivity (q_{it-1}). Here the disadvantage is the loss of observations due to use of lag values. One needs also to determine how many lags to use and to compute short and long-run elasticities.
4. A fourth alternative is the generalized methods of moment (GMM). This method, however, is mainly applied to cases with a single equation with lag dependent variable on the right hand side rather than a system of equations. The intensive use of instruments in a case with a short time period makes this method less suitable. Thus, the GMM method is more suitable in simple dynamic models where the lag value is all that one uses to explain the variation in the dependent variable.

5.3 Three-stage least squares (3SLS) estimation method

Another way to estimate Eqs. (7) and (8) is to apply a three-stage least squares (3SLS) estimation method. The 3SLS estimation method can be applied to account for interdependence and feedback effects. Here, the method uses predicted R&D investment and productivity values on the right-hand side. More precisely, 3SLS can be applied when the equation system is simultaneous by combining the 2SLS and seemingly unrelated regression (SUR) methods to take into account both dependent regressors and cross-equation correlation of the errors.

The 2SLS approach is a single-equation estimator, which may be used to estimate any identified structural equation recursively. In order to estimate each identified equation of a complete structural model, a system estimator estimates all identified parameters of a model jointly. Thus, the 3SLS is the system version of 2SLS, which allows for the possibility of a contemporaneous correlation between the disturbances in different structural equations. The identified structural equations are first estimated by 2SLS, and the resultant residuals are used to estimate the disturbance covariance matrix, which is then used to estimate all identified structural parameters jointly. If the estimation process is iterated rather than stopped at the third stage, the estimates will converge to the full information maximum likelihood (FIML) estimates of the structural model. System methods of estimation are, in principle, more efficient than single-equation methods, provided the system specification is correct.

In the 3SLS approach, there are two endogenous variables. These are growth rate of per capita R&D and of labor productivity. The expected value of the endogenous variable is obtained from the first stage of a 2SLS method conducted using a linear model. To correct for the selection problem that not all firms do R&D, the Inverse Mills Ratio from the Probit is included as an explanatory variable in the estimation. There are two main concerns in the use of such estimation procedure.

The first is that the use of a predicted variable from a different estimation may not allow for a proper inference procedure since we cannot know how the coefficient is distributed and how to calculate its standard errors; but the 3SLS is more efficient than the 2SLS. This is true when: (i) the explanatory and excluded variables are exogenous, and (ii) there are no predicted regressors in the model. There might be doubt that the explanatory variables are truly exogenous, unless we believe that these variables are not all simultaneously determined with productivity and R&D activity. One standard method is to use lag of these variables. The standard errors also could be bootstrapped, which is a typical approach when estimated variables are used.

The second is that the data are organized as a panel, but the estimation method used does not fully take advantage of this aspect. Therefore it is important to mention, for instance, how one takes into account possible temporal correlations in the data. In this regard, it is particularly interesting to look at the results of an instrumental variable panel approach that takes into account the above comments. This would mean adding to the 3SLS analysis two separate 2SLS, one for each dependent variable, using a fixed effect method that should correct for possible simultaneity problems resulting from, e.g., good managerial skills. In addition, this would also allow us to conduct tests for over-identifying restrictions on the instruments used.

In the current case, we model the relationship between R&D investment growth and productivity growth as a system. In order to take into account both dependent regressors and cross-equation correlation of the errors, the 3SLS result is applied and will be reported in Sect. 7. Inclusion of an estimated inverted mills ratio and the predicted values of growth in R&D and productivity may affect the standard errors. The 3SLS method is more efficient than any of the other alternative estimation methods discussed above. Thus, the magnitude of the biased standard errors may be very small. Despite this advantage, in cases like this one should use ideally robust standard errors.

6 The data

6.1 The unbalanced sample

The information on the firms' current accounts and balance sheets used to construct the data in this study comes primarily from a firm-level database KIS compiled by the Korea Information Service. The KIS database contains listed companies' financial information collected from companies' annual reports from 1980 to 2002. Out of all listed companies contained in the raw data, the sample of listed firms is extracted over 17 years from 1986 to 2002.

A second source of data used providing information on the 30 largest business groups and their affiliated firms is from the Korea Fair Trade Commission (henceforth KFTC). Each year, KFTC reports the 30 largest business groups and firms that are affiliated with such groups. Note that these firms vary year by year. Especially after the crisis there was a significant change in this affiliation. Also, the definition of *Chaebol* firms no longer exists after 2001. Hence, following the definition of the large business group by KFTC, a *Chaebol* dummy with a value of 1 is given to those firms whose total assets exceed 2 trillion won in 2001 and 2002, and 0 otherwise. Thus, the *Chaebol* dummy indicates a firm's ranking as one of the 30 largest business groups.

Combining the information extracted from KIS to KFTC reports, we constructed a firm level unbalanced panel dataset of listed companies from 1986 to 2002, which includes both firms that entered (went public) and exited

Table 2 Summary statistics of the Korean data, 1980–2001, 11,596 observations

Label	Mean	SD	Minimum	Maximum
Value added	303,391,605	1,656,927,647	−1,666,242,661	40,751,552,376
Physical capital	164,068,119	1,068,705,075	0	42,993,113,538
Labor	1,650	4,354	0	60,898
R&D investment	6,010,574	89,011,070	0	6,165,120,332
Outsourcing	23,974,514	114,053,124	0	2,699,956,436
Profitability	81,559,333	407,226,809	−621,917,145	14,072,299,793
Indebtedness	0.279	0.284	0	8.32
Investment	93,373,524	477,795,486	6083	24,074,246,008
Manufacturing cost/sales ratio	0.712	0.378	−2.017	1.000
Dividend	3463263.024	18,666,785.230	0	89,16,97,586
Chaebol dummy	0.200	0.402	0.000	1.000
Crisis	0.115	0.319	0.000	1.000
Size1	0.040	0.197	0.000	1.000
Size2	0.109	0.312	0.000	1.000
Size3	0.308	0.462	0.000	1.000
Size4	0.215	0.411	0.000	1.000
Size5	0.328	0.469	0.000	1.000
Intangible asset/total assets	0.011	0.044	−0.824	0.954

(were delisted) during the study period. A total of 175 companies were delisted during the study period and a significant proportion of these were due to the Asian financial crisis.

All monetary variables are expressed in fixed 2,000 prices by dividing each by the producer price index with 2000 as the base year. In the data management steps, extreme values were excluded as well as observations that report sales as missing. After all these steps, we are left with 11,596 firm-year observations representing the unbalanced sample of listed firms from 1986 to 2002. Summary statistics of the data and correlation matrix of the variables used are shown in Tables 2 and 3.

6.2 Variable descriptions

6.2.1 Dependent variables

6.2.1.1 Labor productivity In estimating a production function, the two most commonly used measures of productivity are Total Factor Productivity (TFP) and Labor Productivity.[7] When using TFP for estimations, various difficulties occur. These include how to measure productivity and how accurately we can measure the contribution of capital and labor to output growth. This study uses labor productivity. Since the number of hours worked is not readily available to be used in this empirical study, real value-added divided by the number of full time equivalent employees was used as a measure of labor productivity instead. Much of the existing literature, including Seo (2002), uses this variable as a measure of labor productivity.

6.2.1.2 R&D investments The R&D investments variable contains six components, five from the income statement and one from the statement of cost of manufactured goods, all of which are extracted from the KIS database. These include research costs, normal R&D costs, development expenses, amortization of R&D costs, and investment of R&D obtained from the income statement, and research and ordinary development expenses taken from the statement of the cost of manufactured goods. R&D investment is defined as the log of the sum of these six investment components.

6.2.2 Explanatory variables

The explanatory variables in the R&D decision Eq. (1) and R&D Investment Amount Eqs. (2 and 4) need not be the same. Crepon et al. (1998) used the same set of variables $\left(x_j^1 = x_j^2\right)$ in both equations since there is no a priori reason to do otherwise, and this study also includes the same set of explanatory variables in both equations.

A firm's decision to invest in R&D activity can be influenced by various factors. The question of what determines a firm's R&D expenditure decision can be examined

[7] TFP is the weighted average ratio of total value added for all factors of production (labor and capital).

Table 3 Correlation matrix, $n = 11{,}596$ observations, p-value in parenthesis

Pearson correlation coefficients, $N = 11{,}596$
Prob > |r| under H0: Rho = 0

	Trend	Value added	Capital	Labor	R&D investments	Out-sourcing
Trend	1.0000					
Value added	0.0968 (0.0001)	1.0000				
Capital	0.0736 (0.0001)	0.3350 (0.0001)	1.0000			
Labor	−0.0290 (0.0018)	0.3472 (0.0001)	0.5753 (0.0001)	1.0000		
R&D investments	0.0567 (0.0001)	0.2612 (0.0001)	0.3152 (0.0001)	0.4140 (0.0001)	1.0000	
Outsourcing	0.0945 (0.0001)	0.3227 (0.0001)	0.0665 (0.0001)	0.2338 (0.0001)	0.1378 (0.0001)	1.0000

from two different perspectives: the first concerns economic factors, and the other technological opportunity. Both economic and technological factors can influence a firm's R&D investment decision.

For the economic factors, the following explanatory variables are taken into account: (i) those included in the literature to check the consistency of our data against others, including firm size and industry dummies, and (ii) those characterizing the Korean firms' specific patterns. For example, a high-debt ratio is a well-known characteristic of Korean firms, especially during the industrial development period. With this technique, we can examine whether and to what extent these Korea-specific characteristics have contributed to the R&D patterns of firms in addition to the conventional firm-specific factors found in the literature such as firm size. The expected effects are given in parenthesis.

6.2.2.1 Firm size (+) In the literature, firm size has been included as a determinant of innovative activity. It is argued that larger firms are more capable of mobilizing resources and exploiting economies of scale to capture the maximum amount of benefits from innovative activities. However, the empirical results from various studies of the relationship between firm size and R&D investment are not consistent (Cohen and Levin 1989). Some studies found firm size to have a significant positive influence on R&D intensity (Lall 1983), while others concluded that an increase in size led to an increase in the R&D expenditure, but by less than would be proportionate (Katrak 1989, 1990).

Seo (2002) provides a summary of the R&D activities of Korean manufacturing firms for the year 2000. Seo finds a positive relationship between size of employment and their engagement in R&D activities. Since the sample of this study is restricted to listed firms, a large proportion of the firms are engaged in R&D activities. Hence, this study also includes firm size dummies to investigate whether and to what extent there exists a positive relationship between firm size and R&D investment in the case of Korean listed firms during the period between 1986 and 2002. In this study, firms are categorized as very small (less than 100), small (100–199), medium (200–499), large (500–999), and very large (1,000 or more employees).

6.2.2.2 Outsourcing (±) Outsourcing is defined as the log of manufacturing costs spent outside of firms. It is possible that outsourcing includes not only production, but also R&D activities. More technology-oriented and globalized corporate organizations and advanced IT experts have allowed firms to pursue outsourcing strategically. See Heshmati (2003, 2009), for the effect of outsourcing on innovation input as well as output. Following Heshmati's approach, we also include outsourcing in the estimation.

6.2.2.3 Intangible capital share (±) The ratio of the intangible capital to the total sum of tangible and intangible capital is included in the specification of the R&D investment equation. It will help to establish whether intangible capital intensity is positively related to the propensity to invest in R&D and the extent of R&D investments. The ratio of intangible assets to the sum of tangible and intangible assets is used as a measure of the Intangible Capital Share variable.

6.2.2.4 Investment (−), dividend (−), profitability (+) Investment is defined as the log of total investment assets, Dividend is the log of dividends, and Profitability is the log of gross profit. Some literature recognized that since R&D is a high-risk investment, external financing is expensive, and therefore internal cash flow is preferred to external borrowing (Hall 2002). Lee (1995) and Kong and Kim (2000) also found that an increase in internal cash flows is likely to stimulate investment in R&D. It is likely that the more profitable a firm is, the larger the internal cash flows within the firm. By contrast, Investment and Dividend are proxies of external financing and they are likely to have a negative effect on the propensity for R&D investments.

6.2.2.5 Debt ratio (±) The variable representing indebtedness, Debt Ratio, is measured by the sum of short-term and long-term debt divided by total assets. A high debt ratio is a known characteristic of Korean firms. In particular, large-sized and profitable firms or *Chaebols* were able to access favorable lending schemes quite easily, especially from banks at low cost, encouraging them to keep borrowing from banks. It is well known that *Chaebol* firms are more highly leveraged than non-*Chaebol* firms (see Kim et al. 2006).

6.2.2.6 Chaebol dummies (±) Following KFTC's definition of a large business group, a *Chaebol* dummy with value 1 is given to those firms which belong to the 30 largest business groups in Korea, and a value of 0 is given to all other firms. For the years 2001 and 2002 where the business grouping seized a *Chaebol* dummy of 1 is given to those firms whose total assets exceed 2 trillion won and 0 otherwise.

6.2.2.7 Industry dummies (±) Rosenberg (1974) and Scherer (1965) emphasize the importance of technological opportunity, i.e., the varying degree of ease of innovation across technological fields, in determining the R&D expenditure of firms. To control for the variation in technological opportunity, we introduce 22 industry dummy variables in the model. A two-digit Korea Standard Industry Classification was used (see Appendix).

6.2.2.8 Financial crisis (−) The Crisis dummy 1 indicates the years 1997 and 1998. By including crisis dummies, we capture a temporary negative shock to R&D investments. As shown in Table 1 and Figs. 1, 2 and 3, a negative sign is expected.

6.2.2.9 Time trend (±) In order to capture the unobserved changes over time which are common to all sample firms, a time trend is included in the model specification.

7 Analysis of the results

7.1 Determinants of the R&D investment decision

Table 4 shows the result from the Heckman two-step estimation of the R&D investment model which is a combination of probit and OLS models. The first stage is estimated using a random effects probit model.

According to the result from the first stage probit equation for the Korean listed firms the probability of engaging in R&D increases with outsourcing, profitability, dividend, investment assets and manufacturing cost to sales ratio at the <1% level of significance between 1986 and 2002. The probability also increases with the debt ratio but only at the <10% significance level. The coefficient of Crisis was negative as expected but was not statistically significant. The probability of engagement in R&D decreases over time at the <1% level of significance. *Chaebol* affiliation does not seem to influence the probability of engaging in R&D activities.

Table 4 R&D investment equation, (1st step: random effects probit estimation)

Parameter	Definition	Estimate	Std error
Intercept	Intercept	−1.9233***	0.4171
Loutsour	Log(outsourcing)	0.0456***	0.0049
Lprofit	Log(profit)	0.0025	0.0070
DEratio	Debt ratio	0.0277	0.0694
Ldividend	Log(dividend)	0.0251***	0.0038
Linvest	Log(investment)	0.1618***	0.0275
Manufac	Manufacturing cost/sales	0.2431***	0.0749
Crisis	Crisis dummy	−0.1089*	0.0644
Chaebol	Chaebol dummy	−0.1217	0.1025
Trend	Trend	−0.3431***	0.0054
S2	Small size	0.2689***	0.0750
S3	Medium size	0.3209***	0.0944
S4	Large size	0.6401***	0.1202
S5	Very large size	1.0836***	0.1637

No. of observations: 11,596

Out of 11,596 observations, 9,259 observations were associated with positive R&D investment. The rest of the observations, that is 2,336 observations, were the case in which R&D investments were reported as 0 or not reported at all

*** and * indicate statistical significance at <1 and 5% level, respectively. (same for the other tables.)

In the analysis, industry effects were also controlled by including industry dummies

Regarding firm size, it is found that overall, the likelihood of engaging in R&D increases with size. This result is consistent with the 'Stylized Fact 1' presented in Cohen and Klepper (1996, p. 928), which states that the likelihood of a firm reporting positive R&D effort increases with firm size. The probability of engaging in R&D investment also differs across the industry, showing consistency with the hypothesis that the R&D investment probability would differ depending on the different technological opportunities of firms and industries.

7.2 The R&D investment intensity

We now turn to the R&D investment equation (second stage OLS estimation). For sensitivity analysis we first estimated the R&D equation with a simple OLS, ignoring

Table 5 Panel data instrumental approach (2nd step: GR&D and GY equations)

Variable	Definition	gR&D parameter estimate	gR&D bootstrapped std error	gY Parameter estimate	gY bootstrapped std error
Intercept	Intercept	378.9798	244.6997	−4.9157*	3.0380
gY	Growth in output	5.7879**	2.7943	–	–
gR&D	Growth in R&D investment	–	–	0.1292**	0.0659
gK	Growth in capital	–	–	11.5792	9.6323
gL	Growth in labor	–	–	2.4920	3.1956
Mratio	Inverse mills Ratio	−247.6348**	121.3816	–	–
Loutsour	Log(outsourcing)	−2.8497	2.0244	–	–
Lprofit	Log(profit)	−3.1389**	1.5672	–	–
Ldividend	Log(dividend)	−1.5446	2.4718	–	–
DEratio	Debt ratio	−13.7085	12.2825	–	–
Linvest	Log(investment)	−14.1458	16.7845	–	–
Manufac	Manufacturing cost/sales	−9.9784	30.2957	–	–
Chaebol	Chaebol dummy	53.6270***	16.0710	−3.6926	2.9023
Crisis	Crisis dummy	0.1061	6.6145	−0.8354	1.0217
Trend	Trend	2.1987	2.8091	0.1277	0.2094
s2	Small size	−21.9894*	13.4712	−0.0378	0.8966
s3	Medium size	−0.5101	46.0904	−2.6550	3.2064
s4	Large size	−61.6666**	26.6569	3.3116	3.0504
s5	Very large size	−49.9176	38.6608	1.6911	4.2708
R2	Adjusted R2	0.7402		0.8236	

In the analysis, industry effects were also controlled by including industry dummies. The prefix g indicates growth rate

***, ** and * indicate statistical significance at 1%, 5% and 10% levels, respectively

the endogeneity problem.[8] A selection variable, which is labeled the inverse mills ratio (Mratio), was included to correct for the sample selection bias. Note that the coefficient—9.1038 was statistically significant at the <1% significance level suggesting the necessity to correct for overestimation of effects as a result of sample selection problem.

The dependent variable is the log of R&D investments. The estimates of continuous variables represent the elasticities of R&D investments. In the R&D investments equation, positive and significant elasticities are found with respect to profitability and investment assets at the 1% significance level, intangible asset intensity at the 5% significance level and dividend at the 10% significance level. *Chaebol* firms are found to invest more in R&D activities than non-*Chaebol* firms once they are engaged in R&D activities. That is, although *Chaebol* affiliation does not influence a firm's decision to conduct R&D activities, amongst firms conducting R&D, *Chaebol* firms seem to invest more in R&D than non-*Chaebol* firms.

A positive and at the <1% level of statistically significant time Trend coefficient implies that firms were increasing R&D investments over time between 1986 and 2002. The coefficient of the Crisis dummy was negative as expected, indicating that firms reduced the amount of R&D expenditures during the crisis period, but was not statistically significantly different from zero. The insignificant effects might be due to the fact that the reduction in R&D investment after the crisis is already captured by the time trend variable. The two effects are confounded and not easily separated. It is also found that the extent of R&D investment differs by industry classification as the industries differ by their technological opportunity. Excluding industry dummies from the estimation model reduces the explanatory power by approximately 4%.

In accounting for endogeneity and selection problems, two different estimation methods of panel data instrumental approach and three stages least squares method, are employed to estimate the second stage—the system of equations. In both cases robust bootstrap standard errors are reported. The instrumental variable approach results are reported in Table 5.

The selection variable, the coefficient of Mratio in the R&D growth equation, is statistically significant

[8] Due to limited space these results are not reported here but can be obtained from the authors upon request.

(−247.6348) at the less than 5% level. The result suggest that in estimation of R&D investment it is necessary to account for differences in probably of being an R&D investor among the firms by correcting for the selection bias. R&D growth is found to positively affect (0.1292) growth in labor productivity. None of the remaining variables including growth in Capital, growth in Labor, Chaebol, Crisis, time Trend or Size dummies are statistically different than zero.

In accordance with our expectation, we find a positive and at the less than 5% level significant feedback effect from growth in labor productivity (5.7879) on R&D investment. In addition, we unexpectedly find that retained profit reduces the level of R&D investment. The coefficient of Chaebol dummy is highly significant and positive, confirming the higher propensity of the conglomerates to be involved in R&D investment. The significant and negative coefficients of small and large sizes suggest a negative association between R&D and size of firms. The remaining determinants were found to be insignificant.

The results described above show that the results may differ according to the way the dependent variables are specified; namely, in level or growth rates. In the level specification most of determinants of R&D are statistically significant and with expected signs. These turned out to be insignificant when we study growth rate in R&D. A variable identified as a determinant of R&D level does not necessarily serve the same function when considering the growth in R&D. Thus, to be consistent with the mainstream, theoretically and methodologically proper specification and estimation of models may generate redundant parameters and cause difficulties in their interpretation.

7.3 System of R&D investment and productivity growth equations

For sensitivity analysis, the system of two equations, the level of R&D investment and growth in labor productivity are again estimated as a system using 3SLS. The estimation result (not reported here) shows the positive and significant feedback effects from productivity growth to R&D investment growth. The Inverse Mills Ratio was not significant in the R&D investment growth equation.

As expected, labor productivity growth was found to be an important contributor to R&D investment growth per employee at the <1% significance level. *Outsourcing* was found to be positively correlated with R&D investment growth at the 5% significance level and Profitability, measured by the logarithm of profits, was positively correlated with R&D investment at <1% significance level. Firms with a higher debt ratio showed higher growth rates in R&D investment. Regarding the *Trend* variable, R&D Investment growth did not change significantly over time. The significant, negative sign of the Crisis dummy at the <1% significant level implies the decrease in R&D growth in 1997–1998. No systematic pattern was found with respect to firm size.

Let us now turn to the estimation result for the labor productivity growth equation. As expected, R&D is found to be a positive and significant contributor to labor productivity growth at the <1% significance level. Growth in physical capital per employee is found to have positively contributed (at the 5% significance level) to productivity growth as expected. Firm size and labor productivity growth do not show any systematic relationship. Regarding the Crisis dummy, there was a significant slowdown in labor productivity growth during the crisis period of 1997–1998. The dummy variable is defined so as to capture only the transitory effect rather than the long term negative shift effect of the crisis.

Table 6 reports the three stage least squares estimation result of a system that includes two equations: the growth in R&D investment per employee and the labor productivity growth.

In a comparison of the instrumental and 3SLS approaches based on the R&D and labor productivity growth rate specifications we find that the two methods produce similar results, in which there is a positive effect of growth in R&D investment on growth in labor productivity (0.1281) and also a positive feedback effect from growth in labor productivity on growth in R&D (5.7680). Only two other indicators are affecting the growth in R&D investment: dividend and the Chaebol dummy. The strongest effect is attributed to firms belonging to Chaebol conglomerates. Despite the few significant parameters, the interdependent variables and Chaebol explain 74% of variation in R&D growth and 80% of variation in labor productivity growth.

Similar to the instrumental 2SLS approach, the 3SLS results differ greatly depending on the way the dependent variables are specified (level or growth rates). In the level specification most of determinants of R&D investment and labor productivity are statistically significant and with expected signs. Most of these turned out to be insignificant when we study growth rate in R&D. Again there is a tradeoff between methodologically a better specification and estimation of the models and interpretation difficulties. Estimation based on level of the R&D intensity variable is preferred to its growth rate specification.

As we explained in the literature review and data sections, we expected a positive association between the size of firms and their R&D engagement. Chaebol firms with high total asset values are among the large firms. Our expectation was that the financial crisis had a negative temporary shock on R&D investment. Our results showed that the crisis had a negative effect on the propensity to invest but an insignificant effect on the level or growth rate

Table 6 Three stage least-square approach (2nd step: gR&D and gY equations)

Variable	Definition	gR&D parameter estimate	gR&D bootstrapped standard error	gY parameter estimate	gY bootstrapped standard error
Intercept	Intercept	32.6881	86.8871	−3.5829	4.8748
gY	Growth in output	5.7860***	0.0397	–	–
gR&D	Growth in R&D investment	–	–	0.1281***	0.0521
gK	Growth in capital	–	–	11.8729	13.8216
gL	Growth in labor	–	–	2.5298	3.0845
Mratio	Inverse mills ratio	6.9368	54.3484	–	–
Loutsour	Log(outsourcing)	0.2823	0.3408	–	–
Lprofit	Log(profit)	0.9324	0.7429	–	–
Ldividend	Log(dividend)	0.4889*	0.2794	–	–
DEratio	Debt ratio	−5.6292	8.7509	–	–
Linvest	Log(investment)	−3.9856	5.6551	–	–
Manufac	Manufacturing cost/sales	3.8396	8.5421	–	–
Chaebol	Chaebol dummy	43.2750**	19.3962	−4.8164	3.7211
Crisis	Crisis dummy	−7.7903	20.4509	0.0725	3.2603
Trend	Trend	0.6379	1.7204	0.1111	0.2934
s2	Small size	1.8321	22.5822	−0.5792	2.9931
s3	Medium size	21.3992	22.3160	−2.5650	2.9045
s4	Large size	13.0167	23.2409	−1.2889	2.9560
s5	Very large size	−5.9452	26.4976	1.5172	3.7003
R2	Adjusted R2	0.7404		0.7991	

In the analysis, industry effects were also controlled by including industry dummies. The prefix g indicates growth rate

***, ** and * indicate statistical significance at 1%, 5% and 10% levels, respectively

of R&D investment. The positive Chaebol effect on growth in R&D investment might be underestimated due to the size effect. Given that a firm survives a crisis, in the short term a negative shock has small impact on a firm's R&D investment behavior and it strengthens its survival both during and after the crisis period.

8 Guidelines for future research

Some of the guidelines for further research can be summarized as follows: first, due to lack of data, some important variables are excluded from this study such as patents, which is a common indicator of R&D output (or innovation output). In the literature, some researchers suggest that innovation output, not innovation input, causes productivity to increase. Crepon et al. (1998) and Lööf and Heshmati (2002) estimated a system of four equations R&D investment propensity, innovation input, innovation output, and productivity growth to examine this possibility.

In contrast to Crepon et al. (1998) and Lööf and Heshmati (2002, 2006), where innovation input, innovation output, and productivity link were estimated, this study estimates the innovation input and productivity link instead. The major reason for excluding the patents variable is the crudeness of the data. There are many missing values and it is necessary to combine different dataset sources, which requires careful data management and also running the risk of measurement error and therefore incorrect results. The use of the number of patents as a measure of innovation output is very crude as it does not account for the impact of the patents. Any future innovation study linking innovation input–output-productivity growth with high quality data will be informative.

The implications of the use of samples in the analysis and making inferences is also worth noting. One should interpret the estimation result carefully since the sample of this analysis consists of listed companies. Listed firms may be more "R&D-inclined" than non-listed firms, although this is a testable issue. That is, there are more than 79.8% of R&D-engaged firms in our sample, which greatly exceeds the proportion of R&D-engaged firms in the population of companies. These firms account for the bulk of R&D in Korea.

Also, being listed may be a significant determinant in investing in R&D. That is, the sample of listed companies does not represent all companies. Hence, it is important to keep in mind the special characteristics of the sample. Therefore, conducting empirical analyses using samples with different characteristics, i.e., non-listed companies and comparing the results with listed companies by using matching techniques is desirable. The objective here is not to compare the effectiveness of R&D investment by listed versus non-listed firms but rather to measure the effect of R&D on the productivity of a firm in general.

9 Summary and conclusions

In this study, a production function framework has been adopted to analyze the relationships between R&D investment and labor productivity growth using Korea's firm-level panel data, which contains listed firms for the period between 1986 and 2002. The data is unbalanced due to the entry and exit of firms being listed or delisted. An interdependent chain of equations, including the propensity to invest, R&D investment, and productivity growth, has been estimated in a multi-step procedure accounting for the selection and simultaneity biases inherent in the estimation of such models.

The modeling consisted of two parts. The propensity towards R&D investment and the firms' R&D intensity were specified in the generalized tobit model. The (growth in) R&D investment and productivity growth relationship was specified in a system of equations. In order to estimate two parts of the model, the estimation contained two steps. First, to examine the true intensity of R&D investments, a Heckman two-step procedure was applied to estimate the generalized tobit model while a sample selection bias was successfully corrected. Second, to investigate the R&D growth effect of productivity growth and the productivity growth effect of R&D growth, a system of equations was estimated simultaneously by applying a panel data instrumental approach and the 3SLS estimation method.

Important findings of the study are as follows. First, based on the instrumental approach and 3SLS estimation results, it was found that growth in R&D investment per employee significantly and positively influences labor productivity growth. According to the estimation result, the impact of growth in R&D per employee on labor productivity growth was much greater than the impact of growth in capital per employee, indicating the importance of R&D investment's contribution to labor productivity growth in the sample of Korean listed firms.

Second, it was found that the feedback effect of labor productivity growth to per employee R&D growth was large and significant over the entire sample period. This indicates that modeling the relationship between R&D growth and labor productivity growth as a one way causation produces misleading results. The significant feedback effect from labor productivity growth to R&D investment growth found by estimating a system of equations suggests that growth in labor productivity can also cause growth in R&D investment, which is a contributing factor to the labor productivity growth of firms.

Interestingly, *Chaebol* firms were associated with lower R&D growth as well as lower labor productivity growth in comparison to non-*Chaebol* firms. During the crisis period (1997–1998), there was a significant decrease in R&D investment growth. Since the growth of R&D investments has been an important determinant of labor productivity growth for Korean listed firms, it can be argued that a significant and rapid decrease in R&D investment or a slowdown in R&D growth caused by the Asian financial crisis would have harmed productivity growth and the economy. Considering the feedback effect from productivity growth to R&D growth, the effect of such a reduction should have been even more harmful to the economy since a decrease in productivity growth caused by R&D growth would have negatively influenced R&D growth, and in turn, would have led to a further decrease in productivity growth.

Acknowledgments The authors would like to thank the Associate Editor of the journal and an anonymous referee for their very useful comments and suggestions on two earlier versions of this paper.

Appendix

See Table 7.

Table 7 Industry classification

Industry code	Industry dummy	Industry
05,000	1	Fishing
10,000	1	Mining of coal, crude petroleum and natural resources
11,000	1	Mining of metal ores
15,000	2	Manufacture of food products and beverage
16,000	3	Manufacture of tobacco products
17,000	4	Manufacture of textiles, except sewn wearing

Table 7 continued

Industry code	Industry dummy	Industry
18,000	5	Manufacture of sewn wearing apparel and fur articles
19,000	6	Tanning and dressing of leather, manufacturing of luggage and footwear
20,000	7	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
21,000	8	Manufacture of pulp, paper and paper products
22,000	9	Publishing, printing and reproduction of recorded media
23,000	10	Manufacture of coke, refined petroleum products and nuclear fuel
24,000	11	Manufacture of chemicals and chemical products
25,000	12	Manufacture of rubber and plastic production
26,000	13	Manufacture of other non-metallic mineral products
27,000	14	Manufacture of basic metals
28,000	15	Manufacture of fabricated metal products
29,000	16	Manufacture of other machinery and equipment
30,000	17	Manufacture of computers and office machinery
31,000	18	Manufacture of electrical machinery and furniture
32,000	19	Manufacture of electronic components, radio, television, and communication equipment and apparatuses
33,000	20	Manufacture of medical, precision and optical instruments, watches and clocks
34,000	21	Manufacture of motor vehicles, trailers and semitrailers
35,000	22	Manufacture of other transport equipment
36,000	23	Manufacture of furniture; manufacturing of articles n.e.c.
40,000	24	Electricity, gas, steam and hot water supply
45,000	24	General construction
50,000	24	Sale of motor vehicles and motorcycles;
51,000	24	Wholesale trade and commission trade, except of motor vehicles and motorcycles
52,000	24	Retail trade, except motor vehicles and motorcycles
60,000	24	Land transport; transport via pipelines
61,000	24	Water transport
62,000	24	Air transport
63,000	24	Supporting and auxiliary transport activities of travel agencies
64,000	24	Post and telecommunications
72,000	24	Computer and related activities
74,000	24	Professional, scientific and technical services
75,000	24	Business support services
87,000	24	Motion picture, broadcasting and performing arts industries

Based on Korea standard industry classification

References

Aghion P, Howitt P (1992) A model of growth through creative destruction. Econometrica 60:323–351

Aghion P, Howitt P (2009) The economics of growth. The MIT Press, Massachusetts

Arrow KJ (1974) Economic welfare and the allocation of resources for invention. In: Arrow KJ (ed) Essays in the theory of risk bearing. North Holland, Amsterdam

Buxton W (1985) Talcott parsons and the capitalist nation-state. University of Toronto Press, Toronto

Cohen W, Klepper S (1996) A reprise of size and R&D. Econ J 106:925–951

Cohen WM, Levin RC (1989) Empirical studies of innovation and market structure. In: Schmalensee R, Willig RD (eds) Handbook of industrial organization, 2nd edn. Elsevier Science Publishers B.V, Amsterdam, pp 1060–1107

Crepon B, Duget E, Mairesse J (1998) Research, innovation, and productivity: an econometric analysis at the firm level. Econ Innov New Technol 7(2):115–158

Cuneo P, Mairesse J (1984) Productivity and R&D at the firm level in French manufacturing. In: Griliches Z (ed) R&D, patents,

and productivity. University of Chicago Press, Chicago, IL, pp 21–54
Demsetz H (1969) Information and efficiency: another viewpoint. J Law Econ 12:1–22
Galbraith J (1957) American capitalism: the concept of countervailing power. M.E. Sharpe, White Plains, NY
Griliches Z (1979) Issues in assessing the contribution of research and development to productivity growth. Bell J Econ 10(1):92–116
Griliches Z (1985) Data and econometricians—the uneasy alliance. Am Econ Rev 75(2):196–200
Griliches Z (1986) Productivity, R&D and basic research at the firm level in the 1970s. Am Econ Rev 76(1):141–154
Griliches Z (1995) R&D and productivity: econometric results and measurement issues. In: Stoneman P (ed) Handbook of the economics of innovation and technological change. Basil Blackwell, Oxford, pp 52–89
Griliches Z, Mairesse J (1984) Productivity and R&D at the firm level. In: Griliches Z (ed) R&D, patents, and productivity. University of Chicago Press, Chicago, IL, pp 339–374
Grossman GM, Helpman E (1991a) Innovation and growth in the global economy. MIT Press, Cambridge, MA
Grossman GM, Helpman E (1991b) Quality ladders in the theory of growth. Rev Econ Stud 58:43–61
Hall BH (2002) The financing of research and development. Oxford Rev Econ Policy 18(1):35–51
Hall BH, Mairesse J (1995) Exploring the relationship between R&D and productivity in French manufacturing firms. J Econ 65:263–293
Heckman J (1976) The common structure of statistical model of truncation, sample selection and limited dependent variables and a simple estimator for such models. J Econ Soc Meas 5(4): 475–492
Heckman J (1979) Sample selection bias as a specification error. Econometrica 47(1):153–161
Heshmati A (2003) Productivity growth, efficiency and outsourcing in manufacturing and service industries. J Econ Surv 17(1):79–112
Heshmati A (2009) A generalized knowledge production function. IFCAI J Ind Econ 6(1):7–39
Jones CI (1995) R&D-based models of economic growth. J Polit Econ 103(4):759–784
Kang JW, Heshmati A (2008) The Effects of credit guarantees on survival and performance of SMEs in Korea. Small Bus Econ 31(4):445–462
Katrak H (1989) Imported technologies and R&D in a newly industrializing country: the experience of Indian enterprises. J Dev Econ 31:123–139
Katrak H (1990) Imports of technology and the technological effort of Indian enterprises. World Dev 18:371–381
Kim H-S, Heshmati A, Aoun D (2006) Dynamics of capital structure: the case of Korean listed manufacturing companies. Asian Econ J 20(3):275–303
Kong M, Kim B (2000) R&D investment of firm and cash flows. Mid Small Sized Firm Stud 22(2):111–135 (In Korean)
Lall S (1983) Determinants of R&D in an LDC: the indian engineering Industry. Econ Lett 37(3):379–383
Lee BK (1995) The effects of private and public R&D and technology policy. Korea Econ Res Inst (Working Paper)
Lööf H, Heshmati A (2002) Knowledge capital and performance heterogeneity: a firm-level innovation study. Int J Production Econ 76:61–85
Lööf H, Heshmati A (2006) On the relationship between innovation and performance: a sensitivity analysis. Econ Innov New Technol 15(4/5):317–344
Medda G, Piga C, Siegel DS (2005) University R&D and firm productivity: evidence from Italy. J Technol Transf 30(1–2): 199–205
Medda G, Piga C, Siegel DS (2006) Assessing the returns to collaborative research: firm-level evidence from Italy. Econ Innov New Technol 15(1):37–50
Ministry of Science and Technology (2002) Surveys on research and development activities for science and technology. Ministry of Science and Technology, Seoul. (In Korean)
Oh I, Lee JD, Heshmati A (2008) Total factor productivity in Korean manufacturing industries. Glob Econ Rev 37(1):23–50
Oh I, Heshmati A, Baek C, Lee JD (2009) Comparative analysis of firm dynamics by size: the Korean manufacturing. Jpn Econ Rev 60(4):512–538
Parisi M, Schiantarelli F, Sembenelli A (2006) Productivity, innovation creation and absorption, and R&D. Microevidence Italy Eur Econ Rev 50:2037–2061
Romer PM (1990) Endogenous technological change. J Polit Econ 98(5):102–271
Rosenberg N (1974) Science, invention, and economic growth. Econ J 84:90–108
Sakakibara M, Cho D-S (2002) Cooperative R&D in Japan and Korea: a comparison of industrial policy. Res Policy 31(5): 673–692
Scherer FM (1965) Firm size, market structure, opportunity, and the output of patented inventions. Am Econ Rev 55:1097–1125
Seo J-H (2002) The structural change in R&D activities of Korean firms. Korea Development Instit Policy Res Series 2002(08) (In Korean)
Shin TY (2004) The contribution of R&D investment to economic growth. Sci Technol Policy Inst (Working Paper) (In Korean)
Williamson OE (1965) Innovation and market structure. J Polit Econ 73:67–73