Intangibles and Performance

Wendy Li (U.S. Bureau of Economic Analysis)

Paper Prepared for the IARIW 33rd General Conference

Rotterdam, the Netherlands, August 24-30, 2014

First Poster Session

Time: Monday, August 25, Late Afternoon

Intangible and Performance

Wendy C.Y. Li
U.S. Bureau of Economic Analysis (BEA)
1441 L St., Washington, DC 20230
Wendy.Li@bea.gov

Date: July 30, 2014

Abstract

We develop a forward-looking profit model and use non-linear GMM technique to estimate the depreciation rates of business R&D capital. Using data from Compustat between 1987 and 2010, we estimate the model to derive constant industry-specific R&D depreciation rates. The estimates are the first complete set of R&D depreciation rates for major U.S. high-tech industries. They align with the main conclusions from recent studies that the rates are in general higher than the traditionally assumed 15 percent and vary across industries.

1. Model

The premise of our model is that business R&D capital depreciates because its contribution to a firm's profit declines over time. R&D capital generates privately appropriable returns; thus, it depreciates when its appropriable return declines over time. The expected R&D depreciation rate is a necessary and important component of a firm's R&D investment model. A firm pursuing profit maximization will invest in R&D optimally such that the marginal benefit equals the marginal cost. That is, in each period *i*, a firm will choose an R&D investment amount to maximize the net present value of the returns to R&D investment:

$$\max_{RD_i} \pi_i = -RD_i + \sum_{j=0}^{J-1} \frac{q_{i+j+d} I(RD_i)(1-\delta)^j}{(1+r)^{j+d}} \quad , \qquad (1)$$

where RD_i is the R&D investment amount in period *i*, q_i is the sales in period *i*, $I(RD_i)$ is the increase in profit rate due to R&D investment RD_i, δ is the R&D depreciation rate, and *d* is the gestation lag and is assumed to be an integer which is equal to or greater than 0. Period *i*'s R&D investment RD_i will contribute to the profits in later periods, i.e., $i+d, i+d+1, \ldots, i+d+(J-1)$, but at a geometrically declining rate. J should be large enough to cover at least the length of the service lives of R&D assets.[1] *r* is the cost of capital.

Note that when a firm decides the amount of R&D investment for period *i*, the sales *q* for periods later than *i* are not available but can be forecast. In this study past sales are used to forecast the future sales that will be included in the estimation of the depreciation rate. The forecast is based on autoregressive (AR) model of the differenced log sales series. For the various types of industrial data included in this study, the optimal order of the AR model as identified by the Akaike Information Criterion [Mills, 1990] is found to range from 0 to 2. To maintain consistency throughout the study, AR(1) is used to forecast future sales.

The forecast error of the AR model will also affect the estimation of the depreciation rate. To examine this effect, we performed a Monte Carlo calculation with 1000 replications. In each replication, the forecast error of AR(1) at k steps ahead, $\sum_{i=1}^{k} a_1^{k-i}\varepsilon_{t+i}$, was calculated with $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ where σ was obtained by AR estimation. This error is then

[1] J is not necessarily the length of the service lives of R&D assets. J can be ∞ in theory, but in practice any sufficiently large value can be used in calculations. We have confirmed that, with J greater than the service lives of R&D assets, the derived depreciation rates are very stable when we vary the number of J in small increments. In the analysis presented later, we have found that, with the same values of d and J, δ is different across industries.

added to the forecast values based on the AR(1) model. For every industry included in this study, the 1000 estimates of the depreciation rate exhibit a Gaussian distribution.

In the following, the predicted sales in period *i* is denoted as $\hat{q}_i$. In addition, the choice of *J* must cover the duration of R&D assets' contribution to a firm's profit. In this study, we use 20 for *J* except for the pharmaceutical industry where *J* = 40 is used due to the longer product life cycle.[2]

To derive the optimal solution, we define $I(RD)$ as a concave function:

$$I[RD] = I_\Omega\left(1 - \exp\left[\frac{-RD}{\theta}\right]\right) \qquad (2)$$

$I'(RD) > 0$ and $I''(RD) < 0$. And, $\frac{dI}{dRD} = I_\Omega \times e^{\frac{-RD}{\theta}}$ *where* $\frac{dI}{dRD} = I_\Omega$ *when* $RD = 0$.

$I(\mathrm{RD}) \to I_\Omega$ when $RD \to \infty$. The functional form of $I(RD)$ has very few parameters but still gives us the required concave property to derive the optimality condition, an approach adopted by Cohen and Klepper (1996). The model incorporates the assumption of decreasing returns to scale in R&D investments, which is more realistic than the traditional assumption of constant returns to scale. In addition, the model implicitly assumes that innovation is incremental. I_Ω is the upper bound of increase in profit rate due to R&D investments. And, θ defines the investment scale for increases in *RD*. That is, θ can indicate how fast the R&D investment helps a firm achieve a higher profit rate.

From Figure 1, we can see that, for example, when *RD*, the current-period R&D investment amount, equals to θ, the increase in profit rate due to this investment will reach $0.64 I_\Omega$. When *RD* equals to 2θ, the increase in profit rate due to this investment will reach $0.87 I_\Omega$. The value of θ can vary from industry to industry; that is, we expect to see different industries have different R&D investment scales.

The industry data show that the average R&D investment in some industries increases greatly over a period of two decades, and therefore we expect that the investment scale needed to achieve the same increase in profit rate should grow accordingly. For this reason we model the time-dependent feature of θ by $log\theta_t(\theta_{2000}, \alpha) = log\theta_{2000} + \alpha(t - 2000)$, where θ_{2000} is the value of θ in year 2000. The coefficient α is estimated by linear regression of $\log(RD_i) = c + \alpha t$ for each industry. Note that *c* is a constant.

The R&D investment model becomes:

---

[2] *J* is chosen based on the tests that increasing *J* does not change the optimal solutions.

$$\begin{aligned}\pi_i &= -RD_i + \sum_{j=0}^{J-1} \frac{\hat{q}_{i+j+d} I(RD_i)(1-\delta)^j}{(1+r)^{j+d}} \\ &= -RD_i + I_\Omega \left[1 - \exp\left(-\frac{RD_i}{\theta_i(\theta_{2000}, \alpha)}\right)\right] \sum_{j=0}^{J-1} \frac{\hat{q}_{i+j+d}(1-\delta)^j}{(1+r)^{j+d}}\end{aligned} \quad (3)$$

The optimal condition is met when $\partial \pi_i / \partial RD_i = 0$, that is,

$$\frac{\theta_i(\theta_{2000}, \alpha)}{I_\Omega \exp\left(-\frac{RD_i}{\theta_i(\theta_{2000}, \alpha)}\right)} = \sum_{j=0}^{J-1} \frac{\hat{q}_{i+j+d}(1-\delta)^j}{(1+r)^{j+d}} \quad (4),$$

and through this equation we can estimate the depreciation rate δ.

2. Industry-Level Analysis – Constant R&D Depreciation Rates

As a first step in our empirical analysis, we estimate the constant R&D depreciation rates based on two datasets. One is the firm-level Compustat dataset from 1989 to 2008 and the other is the industry-level BEA-NSF dataset from 1987 to 2007. We take the average values of annual sales and R&D investment in each industry from Compustat for estimation. The Compustat dataset contains firm-level sales and R&D investments for SIC-based nine industries. Their corresponding SIC codes and numbers of firms are listed in Table 1. As to the BEA-NSF data, we use the annual industry output and R&D investment in each industry for estimation.

The value of I_Ω can be inferred from the BEA annual return rates of all assets for non-financial corporations. As Jorgenson and Griliches (1967) argue, in equilibrium the rates of return for all assets should be equal to ensure no arbitrage, and so we can use a common rate of return for both tangibles and intangibles (such as R&D assets). For simplicity, we use the average return rates of all assets for non-financial corporations during 1987-2008, 8.9 percent, for I_Ω. In addition, in equilibrium the rate of returns should be equal to the cost of capital. Therefore, we use the same value for r.

We use Equation (4) as the model to estimate the R&D depreciation rate from the data. As $I_\Omega = r = 0.089$, and as RD_i and q_i can be known from data, the only unknown parameters in the equation are δ and θ.

We use a 2-year gestation lag, which is consistent with the finding in Pakes and Schankerman (1984) who examined 49 manufacturing firms across industries and reported that gestation lags between 1.2 and 2.5 years were appropriate values to use (see also Hall and Hayashi 1989). In addition, according to the recent U.S. R&D survey conducted by BEA, Census Bureau and National Science Foundation (NSF) in 2010, the average gestation lag is 1.94 years for all industries. [3]

As mentioned previously, the value of J is set to be 40 for pharmaceuticals and 20 for other industries. Table 2 shows the two sets of the industry-specific constant R&D depreciation rates based on the Compustat company-based data and the BEA-NSF establishment-based data. The values based on the two datasets are plausible for most industries. For example, the pharmaceutical industry has the lowest R&D depreciation rates in both sets of estimates, which may reflect the fact that R&D resources in pharmaceuticals are more appropriable than in other industries due to effective patent protection and other entry barriers. Compared with the pharmaceutical industry, the computers and peripheral industry has the higher R&D depreciation rate, which is consistent with the industry's observations that, the industry has adopted a higher degree of global outsourcing to source from few global suppliers (Li, 2008). In addition, the module design and efficient global supply chain management has made the industry products introduced like commodities, which have shorter product life cycles.

We simplify and estimate the model with the nonlinear generalized method of moments (GMM) approach. In Equation (4),

$$\frac{\theta_i}{I_\Omega \exp\left[-\frac{RD_i}{\theta_i}\right]} = \sum_{j=0}^{J-1} \frac{q_{i+j+d}(1-\delta)^j}{(1+r)^{j+d}}. \tag{5}$$

Define $f(RD;\theta,\delta) \equiv \frac{\theta_i}{I_\Omega \exp[-\frac{RD_i}{\theta_i}]} - \sum_{j=0}^{J-1} \frac{q_{i+j+d}(1-\delta)^j}{(1+r)^{j+d}}$, and $\theta_i \equiv \theta_0(1+G)^i$, where G is the growth rate of θ_i. Still describing the exponential growth, θ_i is now written as $\theta_0(1+G)^i$. To reduce the number of parameters, we estimate G by fitting the data of R&D investment to the equation, $RD_i = RD_0(1+G)^i$. Therefore, $\frac{\theta_i}{I_\Omega \exp[-\frac{RD_i}{\theta_i}]} = \frac{\theta_0(1+G)^i}{I_\Omega \exp[-\frac{RD_i}{\theta_0(1+G)^i}]}$ or $\frac{(1+G)^i}{I_\Omega} \cdot \theta_0 \cdot \exp[\frac{RD_i}{\theta_0(1+G)^i}]$. In addition, by changing the range of j from $[0, J]$ to $[0, \infty)$, we get:

[3] The average gestation lag is based the responses from 6,381 firms across 38 industries in the NSF 2010 Business R&D and Innovation Survey (BRDIS).

$$\sum_{j=0}^{J-1}\frac{q_{i+j+d}(1-\delta)^j}{(1+r)^{j+d}}=\sum_{j=0}^{\infty}\frac{q_{i+j+d}(1-\delta)^j}{(1+r)^{j+d}}=\frac{q_i(1+g)^d}{(1+r)^{d-1}(r+\delta-g+g\delta)}$$

$$\cong\frac{q_i(1+g)^d}{(1+r)^{d-1}(r+\delta-g)}, \qquad (6)$$

with the assumption that *d = 1* and $q_{i+j}\equiv q_i(1+g)^j$.

We can define the nonlinear residual as:

$$\varepsilon_i\equiv\frac{(1+G)^i}{I_\Omega}\cdot\theta_0\exp\left[\frac{RD_i}{\theta_0(1+G)^i}\right]-\frac{q_i(1+g)^d}{(1+r)^{d-1}(r+\delta-g)}, \qquad (7)$$

and choose the instrument variables as $z_i=[1\ RD_{i-1}\ q_{i-1}]'$. The choice of instrument variables is based on the model assumption that, in a forward-looking profit model, the previous R&D investments and sales will not affect the decision of R&D investments in the current period and the future sales related to current period's investment decision.

Let $m(\Theta)\equiv z_i\varepsilon_i$ and $s(\Theta)\equiv E(m(\Theta)m(\Theta)')$.

The corresponding analog sample moments are:

$$\widehat{m(\Theta)}=\frac{1}{n-1}\sum_{i=2}^{n}\begin{bmatrix}\frac{(1+G)^i}{I_\Omega}\cdot\theta_0\exp\left[\frac{RD_i}{\theta_0(1+G)^i}\right]-\frac{q_i(1+g)^d}{(1+r)^{d-1}(r+\delta-g)}\\ RD_i\cdot\quad\frac{(1+G)^i}{I_\Omega}\cdot\theta_0\exp\left[\frac{RD_i}{\theta_0(1+G)^i}\right]-\frac{q_i(1+g)^d}{(1+r)^{d-1}(r+\delta-g)}\\ q_i\cdot\frac{(1+G)^i}{I_\Omega}\cdot\theta_0\exp\left[\frac{RD_i}{\theta_0(1+G)^i}\right]-\frac{q_i(1+g)^d}{(1+r)^{d-1}(r+\delta-g)}\end{bmatrix}, \qquad (8)$$

and $\widehat{s(\Theta)}=\frac{1}{n-1}\sum_{i=2}^{n}m(\widehat{\Theta})m(\widehat{\Theta})'$.

Define $\Theta\equiv[\Theta_1\ \ \Theta_2]'\equiv[\theta_0\ \ \delta]'$ and

$$M(\Theta)\equiv\frac{\partial m(\Theta)}{\partial\Theta}=\begin{bmatrix}\frac{\partial m_1}{\partial\Theta_1} & \frac{\partial m_1}{\partial\Theta_2}\\ \vdots & \vdots\\ \frac{\partial m_5}{\partial\Theta_1} & \frac{\partial m_5}{\partial\Theta_2}\end{bmatrix}. \qquad (9)$$

The corresponding sample analog is:

$$\widehat{M(\Theta)} \equiv \frac{1}{n-1}\sum_{i=2}^{n}\begin{bmatrix} m_{11}(\widehat{\Theta}) & m_{12}(\widehat{\Theta}) \\ \vdots & \vdots \\ m_{51}(\widehat{\Theta}) & m_{52}(\widehat{\Theta}) \end{bmatrix}. \quad (10)$$

Note that GMM estimators are asymptotic normal: $\widehat{\Theta} \sim N\left(\Theta, \frac{V}{n-1}\right)$ where $V = [M^{'} \ s^{-1}M]^{-1}$.

To derive the optimal solution for Θ, we solve the following optimization problem by using iterative GMM estimation approach with the initial weight matrix as an identity matrix:

$$\widehat{\Theta} = \underset{\Theta}{argmin}\, m(\Theta)^{'} \widehat{W}_K m(\Theta) \quad (11)$$

We continue the iterative operations until the change of the value of the objective function is stabilized.

Table 1 is the comparison of the estimations based on nonlinear least square approach (NLLS) with 2-year gestation lag and nonlinear GMM approach with 1-year gestation lag. We can see that most of the nonlinear GMM estimates and standard errors are higher than those of the NLLS ones across industries.

Table 1: Depreciation Rates of Business R&D Capital

Industry	**Nonlinear Least Squares d = 2**	**Nonlinear GMM d = 1**
Computers and peripheral equipment	41% (1%)	36% (15%)
Software	24% (0.3%)	31% (4%)
Pharmaceutical	10% (0.4%)	18% (32%)
Semiconductor	27% (2%)	26% (5%)
Aerospace	21% (3%)	4% (84%)
Communication equipment	31% (2%)	40% (17%)
Computer system design	43% (0.5%)	87% (70%)
Motor vehicles, bodies and trailers, and parts	28% (1%)	31% (26%)
Navigational, measuring, electromedical, and control instruments	26% (0.6%)	28% (8%)
Scientific research and development	16% (0.3%)	34% (5%)

Notes: 1. *d* refers to the gestation lag of a typical R&D investment. 2. In general, the nonlinear GMM estimates have higher standard errors than those associated with the nonlinear lease square estimates.

Figure 1: The Concavity of $I(RD)$