

**Explaining Cross-sectional Differences in  
Market-to-Book Ratios  
in the Pharmaceutical Industry**

Philip Joos<sup>1</sup>

William E. Simon Graduate School of Business Administration

University of Rochester

Rochester, NY 14627

philjoos@simon.rochester.edu

February 25, 2002

<sup>1</sup>This paper is based on my dissertation completed at the Graduate School of Business of Stanford University. I thank my reading committee, Mary Barth, Peter Reiss and especially Bill Beaver (chair) for their helpful comments and suggestions. I appreciate the comments of seminar participants at Columbia University, Duke University, Harvard University, MIT, the University of California at Berkeley, the University of Chicago, the University of Michigan, the University of Rochester, and the University of Washington at Seattle. Finally, I thank my wife, Margot Baetens, for her help with the pharmaceutical data collection, as well as for explaining the medical terminology.

## **Abstract**

This study examines how accounting information is valued in a research-intensive industry, the pharmaceutical industry. Investors are often concerned that the conservative accounting treatment of research and development (R&D) investments is causing a reporting bias, resulting in accounting summary numbers being less value-relevant for research-intensive firms. The paper examines how accounting return on equity (ROE) and R&D are valued for a sample of research-intensive companies and relate the valuation coefficients to both economic and accounting recognition determinants. The empirical results support the four hypotheses on the effects of competitive structure, R&D success, and growth in R&D investments on the ROE and R&D valuation coefficients.

# 1 Introduction.

This paper examines how accounting information is valued in a research-intensive industry, the pharmaceutical industry. I consider economic context as well as the conservative accounting treatment of research and development investments as the major determinants of cross-sectional differences in market-to-book ratios. Investors are often concerned that the conservative accounting treatment of research and development (R&D) investments is causing a reporting bias, resulting in accounting numbers being less value-relevant for research-intensive firms (Lev and Sougiannis 1996). I explicitly address the valuation of earnings, book value of equity, and R&D investments in an accounting based valuation model, and relate differences in valuation coefficients to competitive structure and accounting conservatism proxies. Prior literature (such as Amir and Lev (1996)) focuses on the statistical significance of proxies for intangibles by linearly adding these to a valuation model that includes book value of equity and earnings.<sup>1</sup> In this paper, I argue that nonaccounting variables related to competitive structure do not enter the valuation model as separate terms, but rather affect the coefficients on the accounting variables in the valuation equation. The challenge is to formulate a theoretical valuation model in which one can show the nonlinear effects of the economic *primitives* on the valuation of the accounting variables. A similar approach is taken by Biddle et al. (2001), who relate the difference between market value and book value of equity to current residual income, and examine the convexity (nonlinearity) of the the slope coefficient on residual income.

I focus on a particular industry - the pharmaceutical industry - for four reasons. First, I want to overcome several shortcomings of broad cross-sectional studies. In particular, competitive structure variables (such as barriers-to-entry, concentration, and market share) contain much measurement error when calculated on a large variety of firms in different industries. By focusing on a single industry, I hope to derive less noisy competitive structure variables. Second, I want to study an industry with high levels and large enough cross-sectional variation in market-to-book ratios. Third, I have two important contextual supplemental disclosures available over a long period to compute competitive proxies: drug information from the Food and Drug Administration (FDA) and patent information from the U.S. Patent Office. Finally,

---

<sup>1</sup>Amir and Lev (1996) use market penetration, number of subscribers, total population (POPS) as nonfinancial measures for intangible capital.

relatively large R&D investments are an important feature of the pharmaceutical industry. However, U.S. accounting standards require firms to expense these investments rather than to capitalize them. R&D recognition and the valuation consequences for earnings, book value and R&D expenses are important research issues.

Consistent with Palepu et al. (2000) and reasons explained later in the paper, I deflate the valuation equation by book value of equity. The resulting valuation equation puts the market-to-book ratio as a linear function of book return on equity (*ROE*) and R&D expenses deflated by book value (*RDBV*). The R&D term is the distinctive feature of the model when compared to other residual income valuation specifications in the recent accounting literature. I focus on R&D because it is a key element in the production function of pharmaceutical companies and because it is the major source of accounting conservatism in this industry. R&D enters the valuation equation through an economic durability parameter that represents the degree of success in R&D investments, in other words, the economic depreciation of R&D capital.

In the empirical design, I use proxies for growth in R&D investments, abnormal profitability persistence and economic depreciation of R&D capital. I will refer to these as *value driver proxies*. Two of the empirical measures capture aspects of the competitive structure of the drug industry: firm type (pioneering versus generic drug firms) and a firm's overall therapeutic market share. The two competition proxies are expected to affect abnormal profitability persistence, and therefore the valuation multiples on *ROE* and *RDBV*. The proxy for depreciation of R&D capital is based on patent information. The number of patents granted to a firm is an output-oriented measure of innovation, in contrast with the R&D expenditures that are input oriented. Firms that are more successful in conducting R&D will have more patents per dollar of invested R&D than firms that are less successful. The patent variable affects the economic depreciation rate of R&D capital since successful R&D is expected to generate relatively more future benefits (e.g., more revenues or higher cost-efficiency) than unsuccessful R&D. As a consequence, the patent variable also affects R&D reporting bias in *ROE*. Finally, I have two R&D growth proxies, one shortterm and one medium term growth variable.

There are four hypotheses in this paper. Two hypotheses relate to the effect of compet-

itive structure on the valuation coefficients of *ROE* and R&D, one relates to the valuation effect of differential patent output per dollar of R&D, and the last hypothesis relates to R&D growth effects on the valuation model. I predict that pioneering drug firms (versus generic drug manufacturers) and firms with higher overall therapeutic market shares have higher valuation multiples on *ROE* and R&D. I also predict that firms with a higher number of patents per R&D dollar have a higher valuation coefficient on R&D. Finally, growth in R&D investments combined with conservative accounting (the fully expensing of R&D) affect earnings and book value of equity.<sup>2</sup> I predict that higher growth in R&D investments results in higher valuation coefficients on *ROE* and R&D. However, the last prediction depends on whether the level of growth exceeds a specific threshold discussed later in the paper.

I test the four hypotheses on a sample of 35 drug manufacturers that operate in the pharmaceutical preparation industry over the period 1975-1998. I also include information on their wholly owned subsidiaries. The pooled cross-sectional time series sample has 15 generic and 20 pioneering parent drug firms. I match data obtained from the FDA and the U.S. Patent Office with financial information. The regression model explaining differences in the market-to-book ratio includes an intercept, *ROE* and scaled R&D as explanatory variables. The effect of the competitive structure, patent and growth variables on the coefficients in the valuation equation is studied by defining dummy variables for the value driver proxies that interact with the intercept, *ROE* and R&D variable. As a result, the three valuation coefficients can vary with the level of the four value driver proxies.

Not only do I find that R&D should be included as a separate term in the valuation model for research intensive firms, but also that competitive structure, R&D success and growth in R&D investments affect the valuation multiples of *ROE* and R&D in predicted ways. I also show that accounting loss firms, representing 11% of the sample, have much lower valuation multiples on *ROE* and R&D than profit firms, suggesting a lower value relevance of both *ROE* and R&D for these firms.

Section 2 explains some industry specific institutional terminology and describes the drug

---

<sup>2</sup>Chapter 17 in Penman (2001) illustrates the effect of investment growth on earnings, book value of equity, market-to-book ratio, price-earnings ratio under conservative, neutral and liberal accounting.

development and patent process. The next section derives a valuation model. Section 4 develops the hypotheses. Section 5 describes the variables used in the empirical part of the study. Section 6 presents the empirical regression equations and empirical tests. Section 7 presents the data to estimate the valuation model. Section 8 reports the regression results for the pooled sample. Section 9 presents a summary and conclusions.

## 2 Pharmaceutical Industry.

### 2.1 The Drug Development Process.

Pharmaceutical manufacturers fall into two main categories: pioneering and generic drug firms (Scott Morton 1998). I discuss several aspects of both types of manufacturers, and in particular elements that describe the competition between and among these firms.

**Pioneering drug firms** undertake R&D to discover new drugs and bring them to market under a brand name. Examples of pioneering drug firms are Abbott, Alza, Merck, and Pfizer. The entire process of R&D, production, and quality control is regulated and monitored by FDA. A pioneering firm must have an approved *New Drug Application* (NDA) to sell a new product. Pharmaceutical R&D is composed of fairly standardized steps (PhRMA 1999, 27): discovery, preclinical tests, investigational new drug (IND) submission to FDA, three clinical trials, NDA submission, and postmarketing testing. On average it took 14.9 years for drugs approved between 1990 and 1996 to be discovered, tested and approved. The R&D investments made in this long period are on average allocated as follows: 40% for preclinical activities, 30% for clinical trials, and 20% for IND/NDA approvals, manufacturing and quality control. The Boston Consulting Group estimated the pre-tax cost of developing a drug introduced in 1990 to be 500 million dollars, including the cost of research failures as well as interest costs over the entire R&D period. Dimasi et al. (1995) estimate that on average only 3 out of 10 drugs are able to recoup their R&D investments from their sales.

The large time lag and the magnitude of R&D investments coupled with GAAP required expensing of R&D makes R&D the single most important determinant of *MTB* in the pharmaceutical industry. I present a formal model of this in appendix B.

The second category contains **generic drug manufacturers**. A generic drug firm is an “imitator” firm that submits an *Abbreviated New Drug Applications* (ANDA) to FDA. The ANDA demonstrates that the generic drug is “bioequivalent” to the original branded drug (manufactured by a pioneering drug firm). Alpharma, Mylan and Teva are examples of generic drug manufacturers. Cimetidine is an example of a generic copy of Tagamet, an anti-ulcer drug pioneered by Smithkline in 1977. Several generic firms introduced Cimetidine starting from 1994 (e.g., Mylan in 1994, Teva and Zenith in 1995), after the patent protecting Tagamet expired. Although manufacturing a generic drug and filing an ANDA is much less costly than discovering and developing a new drug, entry costs are nevertheless significant for a generic firm (facing price competition and high sunk costs) (Scott Morton 1999).<sup>3</sup> The ANDA process takes on average about 18 months from first submission of the application to final granting of permission (Scott Morton 1998). A generic firm can enter a therapeutic market when a patent protecting a pioneering drug expires. Typically, a patent expires after a drug has been on the market for 12 years (Grabowski and Vernon 1994).<sup>4</sup> The price of a generic drug is on average between 40% and 60% lower than the original pioneering drug. As a result, the market share of the pioneering drug significantly shrinks in the post-expiration period. A way for a pioneering drug firm to deter entry and to keep a high market share even after patent expiration is to build up brand loyalty. That is, a pioneering firm heavily promotes its products (visits to physicians and pharmacists, advertising in medical journals, or direct mail) and hopes that physicians will keep prescribing its products even after generic entry occurred (Hurwitz and Caves 1988).

I develop two value driver proxies from FDA data. One variable distinguishes generic firms

---

<sup>3</sup>Market entry by generics was relatively limited prior to the 1984 *Drug Price Competition and Patent Term Restoration Act* (Hatch-Waxman Act) because of costly FDA requirements that had to be met by the generic products. That is, generic drugs often would have to duplicate many of the pioneer’s tests to gain market approval after patent expiration. The act facilitated the entry of generic drugs after patent expiration by requiring only an “abbreviated” application procedure from the generic firm while it also restored part of the patent life lost during the premarket regulatory process of pioneering drug introductions (Grabowski and Vernon 1992). In addition, FDA inspects the equipment the firm plans to use to manufacture the generic drug, and also inspects early batches of the drug. The Congressional Budget Office provides a detailed report on the effects of increased competition from generic drugs on the prices and returns in the pharmaceutical industry after the 1984 Act (CBO 1998).

<sup>4</sup>Before 1984, it took on average 3 years between the patent expiration date and generic entry with a 40% probability that generic entry occurs. After 1984, it took on average 1.2 months, with a 91% probability that generic entry occurs (CBO 1998, 67). The next section discusses the patent issue in more detail.

from pioneering firms, and another focuses on a firm's market share within its therapeutic markets.

## 2.2 Patents.

R&D expenditures are an *ex ante* cost or input measure and are not in themselves intangible assets. However, market-to-book reflects the market's expected success of current and future R&D expenditures. The key determinant of success for a pharmaceutical firm is its ability to bring out a succession of new products that have a significant market impact. Performance in the development of successful new products is the key to understanding sales, earnings persistence and valuation. While patents are clearly only one possible measure of success in R&D, NDAs, as well as sales and market shares are alternative measures of research output. However, patents are *early* indicators of R&D success since patent applications are usually submitted early in the innovation process. I relate the annual number of filed (and eventually granted) patents to past R&D investments as an empirical proxy for economic depreciation of R&D capital.

## 3 Valuation Model.

In this section, I derive an expression for the market-to-book (*MTB*) ratio as a linear function of book return on equity and R&D investments (scaled by book value of equity). The R&D term enters the valuation model because of the reporting bias in earnings and book value of equity due to the fully expensing of R&D under current GAAP as opposed to capitalizing R&D investments and subsequently depreciating R&D capital. I formalize the reporting bias in appendix B. I refer to appendix A for an overview of the definition of the variables and parameters used in the paper.

Ohlson (1995) shows that the dividend discount model can be rewritten in terms of book value and expected future abnormal earnings. I prefer the use of a scaled residual earnings valuation model to an unscaled version for two reasons. First, book return on equity (*ROE*) is a central valuation measure (Palepu et al. 2000).<sup>5</sup> Second, there is a statistical basis for using

---

<sup>5</sup>I also estimate an unscaled valuation equation with book value and earnings as separate variables as a

*ROE* as valuation measure. In particular, abnormal return on equity under R&D capitalization ( $CROE^a$ ) is more likely to be a stationary series than unscaled abnormal earnings adjusted for R&D capitalization ( $CNI^a$ ), where  $CROE^a$  is defined as  $CROE - r$ , and  $CNI^a$  is equal to  $NI_t - r BV_{t-1}$  (see table A). Qi et al. (2000) provide empirical evidence on the nonstationarity of unscaled abnormal earnings, and also show book value and abnormal earnings do not cointegrate with market value.<sup>6</sup> Similar to Bernard (1994), I express market value ( $MV$ ) scaled by book value under R&D capitalization as a function of abnormal return on equity under R&D capitalization ( $CROE^a$ ) and growth in book value of equity ( $CBV$ ):

$$\begin{aligned} \frac{MV_t}{CBV_t} &= 1 + \sum_{\tau=1}^{\infty} \frac{1}{(1+r)^\tau} E_t \left[ CROE_{t+\tau}^a \frac{CBV_{t+\tau-1}}{CBV_t} \right], \\ &= 1 + \sum_{\tau=1}^{\infty} \frac{(1+k)^{\tau-1}}{(1+r)^\tau} E_t [CROE_{t+\tau}^a], \end{aligned} \quad (1)$$

where  $k$  is the non-stochastic growth rate of book value of equity under R&D capitalization ( $k = \frac{CBV_{t+1}}{CBV_t} - 1$ ),  $CROE^a (= CROE - r)$  is the abnormal book return on equity after adjusting for R&D capitalization, and  $r$  is the non-stochastic discount rate for future expected dividends.

Similar to Ohlson (1995), I transform the infinite sum in eq.(1) into an empirically measurable expression by specifying a linear information dynamic for  $CROE^a$ . Similar to Geroski (1990) who models excess economic profits, I assume abnormal  $CROE^a$  follows an AR(1) process:

$$\begin{aligned} CROE_{t+1}^a &= \omega CROE_t^a + \varepsilon_{t+1}, \\ CROE_{t+1} &= (1 - \omega)r + \omega CROE_t + \varepsilon_{t+1}, \end{aligned} \quad (2)$$

where  $\varepsilon$  is an i.i.d. disturbance term with mean zero. In other words, eq.(2) formalizes the idea that competition eventually drives all profit rates to a competitive level  $r$ . The speed of convergence to  $r$  is reflected in  $\omega$ . The above assumption does not incorporate an *other information* parameter as in Ohlson (1995). The *other information* will enter the valuation framework in another way, as I show below. Note that I make an explicit stochastic assumption on  $CROE^a$

---

sensitivity check. The conclusions are not qualitatively changed.

<sup>6</sup>Alternatively, I could transform unscaled  $CNI^a$  into a stationary series through some order of differencing. However, the necessary order of differentiation is unknown and could be quite high. O'Hanlon (1997) demonstrates in his appendix that the presence of explosive properties of unscaled abnormal earnings could even prevent the transformation of the abnormal earnings series into a stationary series through differencing. If R&D is not the only source of biased accounting recognition of economic income, then the mean  $CROE$  could deviate positively from the cost of capital  $r$  and might cause unscaled  $CNI^a$  to exhibit explosive properties.

and *not* on reported  $ROE^a$ , since  $CROE$  is assumed to be closer to the conceptual economic profitability rate than  $ROE$ . Modeling the dynamic of  $ROE^a$  is more complicated since  $ROE$  is affected by typical characteristics of the accounting model, such as conservatism or delayed recognition (in particular, expensing of R&D).<sup>7</sup> At this point, I note that the assumption in eq.(2) is the most essential difference with the Ohlson (1995) model, which assumes a stochastic relation on unscaled abnormal earnings.

The rate of convergence of  $CROE$  ( $\omega$ ) to the cost of capital ( $r$ ) is hypothesized to be a function of elements such as market competition (within a therapeutic market), barriers-to-entry, and relative cost efficiency compared to other firms (Mueller 1990). These elements are part of *other information* available at time  $t$  beyond accounting information. In the empirical analysis, I explicitly study the effect of *other information* on the persistence parameter ( $\omega$ ) and other valuation parameters (see below).

Using assumption (2) on the stochastic process of  $CROE^a$ , I rewrite valuation equation (1) as follows:

$$\frac{MV_t}{CBV_t} = 1 + \frac{\omega}{(1+r) - \omega(1+k)} CROE_t^a, \quad (3)$$

where  $\omega(1+k) < 1+r$  is assumed. Higher abnormal return on equity persistence,  $\omega$ , results in a higher valuation coefficient on  $CROE^a$ .<sup>8</sup> Since current GAAP does not allow capitalization of R&D expenses, I write the above valuation relation in terms of observable  $BV$  and  $ROE$  assuming the growth rate of book value of equity under R&D capitalization ( $CBV$ ) is equal to the growth rate of R&D investments (i.e.,  $k = g$ ):

$$\begin{aligned} \frac{MV_t}{BV_t} &= 1 - \frac{\omega}{(1+r) - \omega(1+g)} r + \frac{\omega}{(1+r) - \omega(1+g)} ROE_t + \frac{(1+r)(1+g)(1-\omega)\lambda}{((1+r) - \omega(1+g))(1+g-\lambda)} \frac{RD_t}{BV_t} \\ &= \Lambda_0 + \Lambda_1 ROE_t + \Lambda_2 \frac{RD_t}{BV_t}, \end{aligned} \quad (4)$$

The above equation expresses the market-to-book ratio as a sum of one minus a term involving the cost of capital ( $r$ ) plus a term with reported  $ROE$  plus a term with scaled R&D ( $\frac{RD_t}{BV_t}$ ).

<sup>7</sup>Van Breda (1981) suggests to treat accounting rates of return as the output from an accounting “filter”, the input to which are economic events. He shows that it is easier to understand the dynamics of the input process than to understand the filtered output.

<sup>8</sup>In unreported analyses, I find empirical evidence supporting the AR(1) assumption of  $CROE^a$ .

Since the dependent variable and both  $ROE$  and  $\frac{RD_t}{BV_t}$  are observable, the above equation can be estimated.

I make several predictions on the effects of three key parameters ( $\omega$ ,  $g$ , and  $\lambda$ ) on the valuation coefficients  $\Lambda_0$ ,  $\Lambda_1$  and  $\Lambda_2$ . I summarize the partial derivatives in the table below, where  $R = 1 + r$  and  $G = 1 + g$ :

Parameter	Intercept ( $\Lambda_0$ )	$ROE$ ( $\Lambda_1$ )	$RDBV$ ( $\Lambda_2$ )
$\omega$	$\frac{\partial \Lambda_0}{\partial \omega} = \frac{-R(R-1)}{(R-\omega G)^2}$	$\frac{\partial \Lambda_1}{\partial \omega} = \frac{R}{(R-\omega G)^2}$	$\frac{\partial \Lambda_2}{\partial \omega} = \frac{RG(G-R)\lambda}{(R-\omega G)^2(G-\lambda)}$
$\lambda$	$\frac{\partial \Lambda_0}{\partial \lambda} = 0$	$\frac{\partial \Lambda_1}{\partial \lambda} = 0$	$\frac{\partial \Lambda_2}{\partial \lambda} = \frac{RG^2(1-\omega)}{(R-\omega G)(G-\lambda)^2}$
$g$	$\frac{\partial \Lambda_0}{\partial G} = \frac{-\omega^2(R-1)}{(R-\omega G)^2}$	$\frac{\partial \Lambda_1}{\partial G} = \frac{\omega^2}{(R-\omega G)^2}$	$\frac{\partial \Lambda_2}{\partial G} = \frac{R\lambda(1-\omega)[\omega G^2 - R\lambda]}{(R-\omega G)^2(G-\lambda)^2}$

I briefly discuss the effects of the three parameters  $\omega$ ,  $\lambda$  and  $g$  on the valuation coefficients  $\Lambda_0$ ,  $\Lambda_1$  and  $\Lambda_2$ .

First, higher  $CROE^a$  persistence ( $\omega$ ) results in a lower intercept, a higher valuation coefficient on  $ROE$  and an effect on the R&D coefficient that is positive when  $g > r$  and negative when  $g < r$ , ceteris paribus. Written in mathematical notation:  $\frac{\partial \Lambda_0}{\partial \omega} < 0$ ,  $\frac{\partial \Lambda_1}{\partial \omega} > 0$  and  $\text{sign}(\frac{\partial \Lambda_2}{\partial \omega}) = \text{sign}(g - r)$ . Thus, when a firm experiences lower competition in its therapeutic drug markets and therefore has a higher  $\omega$ , its current level of  $ROE$  will have a higher valuation multiple. From inspecting the partial derivatives, one notices that growth in R&D investments determines the magnitude of the  $\omega$  effect on the valuation multiples: higher growth increases the effect of  $\omega$  on both the  $ROE$  and R&D valuation multiple.

Second,  $\Lambda_2$  is the only coefficient that depends on  $\lambda$ , the durability of R&D capital. Higher R&D durability resulting from more successful R&D investments shows up in a higher valuation coefficient on  $RDBV$ , ceteris paribus ( $\frac{\partial \Lambda_2}{\partial \lambda} > 0$ ).

Third, higher growth in R&D ( $g$ ) increases the valuation coefficient on  $ROE$ , and reduces the intercept ( $\Lambda_0$ ), ceteris paribus:  $\frac{\partial \Lambda_0}{\partial G} < 0$ ,  $\frac{\partial \Lambda_1}{\partial G} > 0$ . The growth effect on the R&D valuation multiple depends on the sign of  $\omega G^2 - R\lambda$ , or whether or not  $G > \sqrt{\frac{R\lambda}{\omega}}$ . If a firm's R&D growth  $G$  is above the  $\sqrt{\frac{R\lambda}{\omega}}$  threshold, then growth has a positive effect on the R&D multiple. For example, a firm with a cost of capital of 10%, abnormal profitability persistence of 0.7, and

R&D durability of 0.9 has a cutoff value of 1.189. So, if the firm's annual growth rate falls below 18.9%, then an increase in growth will have a negative effect on the R&D valuation multiple, as long as the firm's growth rate remains below 18.9%. I will discuss the above predictions in more detail in the hypotheses section.

To summarize, I propose an accounting-based valuation model in which growth, R&D economic durability, and profitability persistence affect the valuation of book return on equity and R&D investments in a nonlinear fashion.

## 4 Hypotheses.

This section discusses testable hypotheses based on the valuation model in (4). The predictions in the hypotheses are based on the partial derivatives of the valuation coefficients  $\Lambda_0$ ,  $\Lambda_1$  and  $\Lambda_2$  with respect to  $\omega$ ,  $\lambda$  and  $g$  (see previous section). The first two hypotheses make predictions on the effect of competition ( $\omega$ ) on the valuation coefficients. The third hypothesis focuses on the economic duration parameter  $\lambda$  in the *RDBV* valuation coefficient. Finally, the fourth hypothesis states a prediction on the effect of R&D growth on the coefficients in the valuation model. I do not formulate a hypothesis on  $r$ , the cost of capital, but assume  $r$  is the same across firms.

The first two hypotheses relate to the  $\omega$  parameter in the valuation model. Firms with more persistent abnormal profitability (*CROE*<sup>a</sup>) - a higher  $\omega$  - will have a higher valuation coefficient on *ROE* and lower intercept, assuming  $\omega$  is uncorrelated with  $g$ ,  $r$ , and  $\lambda$ . The effect of  $\omega$  on the *RDBV* valuation coefficient is expected to be positive if the firm has a growth rate greater than its cost of capital, i.e.,  $g > r$ . Abnormal profitability is negatively related to *competition*. High competition within a specific therapeutic market (e.g., cardiovascular drugs) could quickly drive profit rates to a longterm competitive level. Consistent with the IO literature, I view competition as a primary determinant of  $\omega$ . Effects of competition on  $g$  and  $\lambda$  are less clear. Competition in a pharmaceutical market takes three forms: among brand name drugs, between brand name drugs and generic substitutes, and among generic versions of the same drug. The first hypothesis focuses on the distinction between generic firms and

pioneering drug firms. As explained in section 2.1, generic entry in a therapeutic market occurs when patents of a pioneering drug expire, and when expected future payoffs are high enough to make generic entry profitable. Grabowski and Vernon (1994) report an average decline of a pioneering drug's revenue of 30%, 21%, and 12% respectively in the first three years after generic entry. Generic versions of a drug are sold on average 40% cheaper than the original pioneering drug at retail prices when 1 to 10 generic firms enter a market, and 60% cheaper when more than 10 generic firms sell copies of a given pioneering drug (CBO 1998, 33). Overall, generic firms are more likely to experience higher immediate competition than pioneering drug firms because they cannot use entry barriers such as patents. For example, 12 generic versions for Tagamet (a pioneering anti-ulcer drug marketed by Smithkline since 1977) were introduced between May 1994 and February 1997 after the patent on Tagamet expired in May 1994. The hypothesis below is derived from the table with partial derivatives in section 3.

**Hypothesis 1:** *Generic firms have less abnormal profitability persistence  $\omega$  than pioneering firms. As a result of a lower persistence, the valuation coefficient on ROE ( $\Lambda_1$ ) in eq.(4) will be lower and the intercept ( $\Lambda_0$ ) will be higher for generic firms compared to those valuation coefficients of pioneering drug firms, ceteris paribus. Pioneering firms will have a higher valuation multiple on R&D ( $\Lambda_2$ ) relative to generic firms, only if their growth rate of R&D investments exceeds their cost of capital.*

Rather than simply differentiating between generic and pioneering drug firms, the next hypothesis focuses on another dimension of competitive structure, in particular *market share*. Pioneering drugs often compete within the same therapeutic market with other pioneering and generic drugs. Pioneering drug competition happens among NDAs, where the first marketed drug is called *breakthrough* drug and the others *me-too* drugs. Me-too drugs are new molecular entities (pioneering drugs) that are similar, but not identical, in molecular structure and mechanism to the original or breakthrough new molecular entities.<sup>9</sup> At first glance, the overall

---

<sup>9</sup>Consider the example of *Tagamet*, a pioneering anti-ulcer drug that was introduced in 1977 by Smithkline (therapeutic code 0874 in appendix D). Tagamet was the first drug to relieve ulcers by blocking the histamine 2 (H2) receptors in the lining of the stomach from stimulating acid production by the parietal cells. Such treatment is generally superior to antacids, which only neutralize stomach acid. Seven years after Tagamet became available, Glaxo-Wellcome introduced *Zantac*, which became the largest-selling drug in the world. Two additional H2 antagonists, *Pepcid* and *Axid*, were marketed by Merck (1986) resp. Eli Lilly (1988). Four slightly different drugs using the same therapeutic mechanism (blocking the H2 receptor) were all patentable, and the

pharmaceutical industry does not appear to be highly concentrated, but when narrower therapeutic markets are considered, concentration (or competition) varies widely. A firm's market share reflects the results of competitive forces. Market share is a measure for a firm's market power within a therapeutic market and is affected by quality differences, patent positions (for pioneering firms), scale-related efficiencies (Allen and Hagin 1989), and price discrimination (through brand name loyalty). Higher market power leads to higher persistence of economic profitability. Therefore, market share directly affects the  $\omega$  parameter in the valuation model.

**Hypothesis 2:** *Firms with a large market share will have higher abnormal profitability persistence  $\omega$  than firms with small market shares. As a result,  $\Lambda_1$  in eq.(4) will be higher and  $\Lambda_0$  will be lower for firms with higher market shares compared to firms with lower market shares, ceteris paribus. The R&D valuation multiple will only be higher for firms with larger market shares if the growth rate of these firms is greater than their cost of capital, ceteris paribus.*

The third hypothesis is related to the R&D process, in particular to elements affecting  $\lambda$ . The coefficient on  $RDBV$  in eq.(4) might vary across pharmaceutical firms due to cross-sectional differences in  $\lambda$ , even though all firms could have high degrees of R&D intensity ( $RDBV$ ). As explained before, an essential activity in the industry is the generation of R&D *knowledge*. The value of the knowledge investments is reflected in the durability parameter  $\lambda$ . Firms that are more effective in generating knowledge will have higher values for  $\lambda$ . As explained in section 2.2, an important output measure of knowledge is a firm's patent activity. More successful R&D should be valued more by the capital market than less successful R&D. Patent intensity measures R&D success by relating patents to previous R&D investment. The hypothesis below is derived from the table with partial derivatives in section 3.

**Hypothesis 3:** *Firms with a higher number of patents per dollar invested in R&D will have a higher economic durability of R&D capital ( $\lambda$ ) and hence a higher coefficient on scaled R&D ( $\Lambda_2$ ) in eq.(4) than firms with less patents per dollar R&D, ceteris paribus.*

Finally, the fourth hypothesis addresses growth in R&D expense, i.e. the  $g$  parameter. The growth parameter affects the three valuation coefficients in eq.(4). The current level of *ROE* breakthrough drug had only six years of market exclusivity before being challenged by a competitor (CBO 1998, 19).

will have a higher valuation coefficient for high growth firms compared to low growth firms. Lev et al. (1999) address the relation between R&D growth and *ROE* in their table 3. Although they find that high R&D growth firms tend to have high market-to-book ratios, they do not make specific predictions on the *ROE* valuation coefficient as in this study. The  $g$  parameter also affects the *RDBV* valuation coefficient ( $\Lambda_2$ ) in a non-linear way. The effect of growth on the R&D valuation multiple is only positive if the following growth condition is satisfied:  $G > \sqrt{\frac{R\lambda}{\omega}}$ . I explained this condition in more detail in section 3.

**Hypothesis 4:** *Firms with a higher growth rate in R&D expense ( $g$ ) will have a higher valuation coefficient on current *ROE* ( $\Lambda_1$ ) and on scaled R&D ( $\Lambda_2$ ), but a lower intercept ( $\Lambda_0$ ), ceteris paribus. The prediction on the R&D multiple only holds when the growth condition  $G > \sqrt{\frac{R\lambda}{\omega}}$  is satisfied, and is reversed otherwise.*

## 5 Variables.

I use two sets of variables in this study: the financial variables *ROE* and *RDBV* being valued in the valuation model, and the value driver proxies affecting the valuation coefficients on *ROE* and *RDBV*. All the variables are measured at a firm level, where a firm is defined as the set of a parent and its majority owned subsidiaries. I provide a more detailed discussion on this issue in section 7.1. Appendix A lists all variables used in this paper.

The first set of variables is related to the financial variables in the valuation model and are calculated from the Compustat database (see section 7). The dependent variable is the market-to-book ratio, defined as the ratio of a firm's market value at the end of the fiscal year (Compustat item199  $\times$  item25) to its book value of common equity (item60). The return on equity (*ROE*) is defined as the ratio of current year income before extraordinary items (item18) to the previous year book value of equity (item60). Similar to Dechow et al. (1999), I eliminate extraordinary items from the numerator of *ROE* because the nonrecurring components of net income are not persistent over time. It is precisely the persistent (or recurring) components of *ROE* that affect the valuation coefficient on *ROE* (Ohlson 1999). The last financial variable is deflated R&D investments *RDBV* (item46 over item60).

In the second set of variables, I develop proxies for the parameters  $\omega$ ,  $\lambda$  and  $g$  in eq.(4). The value driver proxies do not enter the regression model as simple linear conditioning variables, as done in most of the the prior intangibles literature. Rather, my value driver proxies affect the coefficients on the intercept,  $ROE$  and R&D in the valuation model. In particular, I consider four variables.

The first variable relates to hypothesis 1.  $DFIRM$  is a dummy variable that makes the distinction between generic and pioneering drug manufacturers. In section 2.1, I described the economic differences between the two types of firms. Often, pioneering drug firms also market ANDAs, or have subsidiaries that are generic producers.<sup>10</sup> I classify a firm  $i$  as “pioneering” if the majority of the drugs it sells in a specific year  $t$  are pioneering drugs (NDAs):

$$DFIRM_{it} = \begin{cases} 1 & \text{if } FIRM_{it} = \frac{\#NDA_{it}}{\#NDA_{it} + \#ANDA_{it}} > 0.5 \quad (\text{pioneering}) \\ 0 & \text{if } FIRM_{it} = \frac{\#NDA_{it}}{\#NDA_{it} + \#ANDA_{it}} \leq 0.5 \quad (\text{generic}) \end{cases} \quad (5)$$

That is, for each sample firm-year observation I calculate the total number of the firm’s NDAs that are still on the market versus the number of ANDAs (i.e. generic drugs). I cumulate the new drugs introduced in a particular year across all years in a firm’s life to get a stock variable on the drug portfolio at each point in time. I subtract the withdrawn or discontinued drugs from the drug portfolio. If the proportion of NDAs in the drug portfolio is larger than the proportion of ANDAs, then I classify the firm as a *pioneer* in that year. The  $DFIRM$  dummy variable is equal to one if a firm is classified as a pioneering drug firm, and zero otherwise. By using counts of drugs I assume that all drugs are of equal economic importance. Ideally, I would base  $DFIRM$  on a breakdown of a firm’s total sales (or gross profits) into sales (gross profits) from generic products and from pioneering drugs. However, this information is not available to me. As a sensitivity test, I compare the company assignment by  $DFIRM$  to the *business description* section in Moody’s Industrial Manual. Specifically, the assignment of each sample firm in appendix C corresponds with the business description in the Moody’s Manual. The assignment does not differ from the one I use.

---

<sup>10</sup>Although the same company rarely produces both a pioneering drug and its generic copy, some generic manufacturers are subsidiaries of pioneering firms. Most generic subsidiaries do not produce copies of their parent company’s drugs (CBO 1998, 34).

The second variable also relates to  $\omega$  and measures a firm's overall average market share in its therapeutic markets. The previous proxy only distinguished between pioneering and generic firms as a competition measure affecting  $\omega$ . The current variable incorporates all sorts of competition within a therapeutic market. That is, between (pioneering) breakthrough drugs, generic substitutes and *me-too* drugs. A firm's market share reflects the results of that competition. Ideally, I want to measure market share as a firm's dollar amount of sales in a therapeutic market divided by the total sales of all firms active in that market. Relating a firm's total sales number to the total sales of all sample firms does not correspond with the theoretical notion underlying market share: a firm should be compared to its rivals in a particular therapeutic market.<sup>11</sup> I would need to have individual drug sales (a dollar amount) for the period 1975-1998 for all my sample firms and their subsidiaries. Individual drug sales data from *IMS Health* (a large pharmaceutical data-collection firm) Drugstore and Hospital Audits were used by Taylor (1999) and Scott Morton (1999) among others. However, it is not feasible to obtain a comprehensive dataset with yearly sales figures for all the drugs in my dataset. As an alternative, I define the average market share of firm  $i$  in year  $t$  as follows:<sup>12</sup>

$$MS_{it} = \frac{1}{J_{it}} \sum_{j=1}^{J_{it}} \frac{(\#NDA_{ijt} + \#ANDA_{ijt})}{\sum_{i=1}^{N_{jt}} (\#NDA_{ijt} + \#ANDA_{ijt})}, \quad (6)$$

where  $\#NDA_{ijt}$  (resp.  $\#ANDA_{ijt}$ ) is the stock of NDAs (ANDAs) of firm  $i$  in market  $j$  in year  $t$ ,  $J_{it}$  is the total number of therapeutic markets in which firm  $i$  is active in year  $t$ , and  $N_{jt}$  is the total number of firms active in market  $j$  in year  $t$ . The  $MS$  proxy is an average of

---

<sup>11</sup>For example, Genentech is a focused pioneering drug firm that concentrates its R&D on human growth hormone deficiencies (hGHD; code 1042 in appendix D). Genentech introduced its rDNA engineered drug *Nutropin* in 1993, 14 years after its scientists first cloned the human growth hormone. Genentech's direct rivals and their drugs are: Tap Pharmaceuticals, a joint venture between Abbott and Takeda (*Lupron*, 1985), Lilly (*Humatrope*, 1987), Pharmacia & Upjohn (*Genotropin*, 1987), Bio-Technology General (*Bio-Tropin*, 1995), and Ares-Serono (*Saizen*, 1997). Genentech's ability to earn an abnormal profitability rate to sustain its persistence ( $\omega$ ) is directly related to the competitive forces in the hGHD market, and not so much to the entire pharmaceutical industry.

<sup>12</sup>The definition assumes all drugs are of equal importance in terms of sales. Grabowski and Vernon (1994) provide evidence of a highly skewed distribution of sales which exists for new drug introductions. Peak sales of the top decile are several times that of the next ranked decile. Furthermore, mean sales are significantly greater than median sales. The distribution of individual drug sales clearly reflects the blockbuster drug influence, i.e. only 10% of all NDAs have an average (after tax) present value over their drug life cycle of 1 billion dollar. Seventy percent of the new drugs do not exceed the average present value of their R&D costs (200 million dollar).

a firm's market shares in a particular year across all therapeutic markets in which it operates. The denominator in the above definition represents the total number of drugs in the same therapeutic class, i.e. the population of drugs for that particular class. Similar to Scott Morton (1999), I define therapeutic markets based on the National Drug Code (NDC) classification provided by the FDA (see appendix D for an overview of the 117 therapeutic markets). The same classification scheme is used in *Mosby's GenRx*, an influential medical reference used by physicians. For drugs without a therapeutic class in the NDC list, I use the *Merck Index* to assign a therapeutic class. One drug can be used for several illnesses and can therefore be classified in more than one class. The NDC classification is very detailed, and drugs within each category can be considered as competitors. This feature is desirable for the purpose of this study. There are several ways to define therapeutic markets, usually at a more aggregate level. Egan et al. (1982) and Matraves (1999) provide a more detailed discussion. As a result of the narrow market definitions, small firms could have large values on  $MS$  if they have a dominant position in small niche markets. Large firms, on the other hand, could have small  $MS$  values if they are only active in very large therapeutic markets with many other competitors. Notice that  $MS$  is an indicator of a firm's market power, and therefore directly affects  $\omega$  (see hypothesis 2): firms with a dominant market position are able to sustain a high level of abnormal economic profitability and therefore have high market-to-book ratios.

The third variable relates to the R&D economic durability parameter  $\lambda$  in the patent hypothesis (Hypothesis 3). Patent intensity,  $PAT$ , is measured as the ratio of the number of patent applications at the US Patent Office in a particular year to the three-year backward moving average R&D investment:

$$PAT_{it} = \frac{\# \text{ Patent applications }_{it}}{\frac{1}{3} \sum_{\tau=0}^2 RD_{it-\tau}} . \quad (7)$$

That is, patents are the result of past research efforts (here the most recent three years), and are therefore deflated by average R&D investments. Notice that the numerator of  $PAT$  is a count variable, and the denominator is a dollar amount. This type of deflation is common in the patent literature, for example Ben-Zion (1984) deflates by book value of equity and Hall et al. (1999) deflate by total assets. I use patent applications instead of patent grants in the

numerator, since the patent application date is a better indication of the time of innovation than the issuance date (Hall et al. 1999). The ratio is a measure of the innovation and R&D success of a firm. Only eventually granted patents appear in the US patent database. A patent truncation bias at the end of the sample period occurs due to the lag between application and grant.<sup>13</sup> That is, in the final sample years not all the grants for patents applied for appear. Similar to Hall et al. (1999), I adjust the patent count variable between 1993 and 1998 as follows:

$$\# \text{ Patent applications}_{it}^{adj} = \frac{\# \text{ Patent applications}_{it}}{\sum_{\tau=0}^{98-t} f_{i\tau}} \quad 93 < t < 98$$

where the numerator on the right hand side is the unadjusted number of patent applications of firm  $i$  in year  $t$ , and  $f_{i\tau}$  is the firm-specific average proportion of patent applications granted  $\tau$  years after the application year  $t$ . The denominator of the adjustment formula represents the firm's historical cumulative frequency of patents granted  $\tau$  years after application. Since the cumulative frequency is close to one for  $\tau = 6$ , the adjustment is only performed for the final six sample years (1993-1998). For example, Merck has the following frequency  $f_{\tau}$  pattern: 7.97% for  $\tau = 1$  (i.e., the percentage of patent applications that are granted within the first year of application), 59.14% for  $\tau = 2$ , 23.56% for  $\tau = 3$ , 6.39% for  $\tau = 4$ , 1.6% for  $\tau = 5$ , and finally 0.47% for  $\tau = 6$ . Thus, on average 99.13% of the patent applications are granted within six years after application.

Malerba et al. (1997) give some caveats on the use of patent count data. First, not all innovations are patented by firms. Secrecy, particularly in the case of process innovations, is sometimes a more effective appropriation mechanism (Encaoua et al. 1998). However, in the pharmaceutical industry, most pioneering firms file their patents early in the drug innovation process (around the time they file an IND to FDA). Second, patents do not grant complete monopoly power in the pharmaceutical industry (CBO 1998). The reason is that firms can frequently discover and patent several different drugs that use the same basic mechanism to treat an illness. That is the case for me-too drugs (see above).<sup>14</sup> Third, patents cannot be

<sup>13</sup>The median firm-specific application-grant lag in the sample varies between 16 and 41 months.

<sup>14</sup>For example, Alza mentions the following in its 1997 annual report: "Although a patent has a statutory presumption of validity in the U.S., the issuance of a patent is not conclusive as to such validity or as to the enforceable scope of the claims of the patent. There can be no assurance that patents of Alza will not be

distinguished in terms of relevance without specific analyses on patent renewals or patent citations. Patents have an extremely skewed distribution of private patent values, meaning that only a small fraction of patents has a high value (Hall et al. 1999). As a result, a simple patent count is expected to be a noisy measure of a firm’s R&D success.

The fourth and last variable relates to the  $g$  parameter in hypothesis 4. I define R&D investment growth,  $RDG$ , as the yearly growth in R&D expenses:

$$RDG_{it} = \frac{RD_{it}}{RD_{it-1}} - 1, \quad (8)$$

where  $RD_{it}$  is firm  $i$ ’s R&D expense in year  $t$ . The  $RDG$  variable is a short-term growth measure, and might not fully be consistent with the constant  $g$  assumption in the valuation model. I therefore define an alternative (annualized) R&D growth measure that spans a longer time window:

$$LTRDG_{it} = \sqrt[3]{\frac{\frac{1}{3} \sum_{\tau=0}^2 RD_{it-\tau}}{\frac{1}{3} \sum_{\tau=3}^5 RD_{it-\tau}}} - 1, \quad (9)$$

where the numerator is the three year (moving) average of R&D investments in period  $[t - 2, t]$ , and the denominator is the average R&D investment in period  $[t - 5, t - 3]$ . The expression within the root is therefore a three year growth, and is annualized by taking the proper root.

## 6 Econometric Analysis.

In this section, I present the empirical specification of the  $ROE$ -based valuation model discussed in section 3. A first empirical specification for the valuation model in eq.(4) is:

$$MTB_t = \alpha + \beta ROE_t + \gamma RDBV_t + \varepsilon_t \quad (10)$$

where  $MTB$  is the market-to-book ratio,  $RDBV$  is the ratio of R&D to  $BV$  (or scaled R&D), and  $\varepsilon$  is a stochastic disturbance assumed to be I.I.D. I could estimate the above equation in

---

successfully challenged in the future. In some cases, third parties have initiated reexamination by the Patent and Trademark Office of patents issued to Alza, and have opposed Alza patents in other jurisdictions. The validity and enforceability of Alza patents after their issuance have also been challenged in litigation. If the outcome of such litigation is adverse to Alza, third parties may then be able to use the invention covered by the patent, in some cases without payment.”

different ways by making different assumptions on the variance-covariance matrix of the disturbance term. A firm-specific estimation of the above regression model suffers from inefficiency in the valuation parameters due to the short time-series since there are only between 8 and 24 observations per firm available. To make better use of the cross-sectional variation in the dataset, I pool the firm data into one cross-section time series sample and study the effects of  $\omega$ ,  $\lambda$  and  $g$  on that pooled sample (see below).

One characteristic of the pharmaceutical industry is the high frequency of negative earnings. In the final sample, 11.5% of the firm-year observations represent accounting losses, and 22 out of 35 firms had at least one year with negative earnings. Hayn (1995), Collins et al. (1999), and Leibowitz (1999) report differences in valuation coefficients between positive and negative earnings.<sup>15</sup> Similar to these studies, I differentiate between firms with positive and negative earnings, and allow the intercept ( $\alpha$ ), *ROE* coefficient ( $\beta$ ), and *RDBV* coefficient ( $\gamma$ ) to vary across positive and negative earnings firms through the indicator variable  $DNEG_t$ , which is equal to 1 if a firm has negative earnings in year  $t$ , and 0 otherwise.

I study the effects of the four value driver proxies *FIRM*, *MS*, *PAT* and *RDG* (or, *LTRDG*) on the valuation coefficients  $\alpha$ ,  $\beta$  and  $\gamma$  in eq.(10) by defining a dummy variable for each value driver proxy, and allowing each valuation coefficient to vary with these dummies. That results in  $2^5 = 32$  possible values on the intercept, *ROE* and *RDBV* valuation coefficient, since there are four different value driver (dummy) variables and the *DNEG* dummy :

$$\begin{aligned}
MTB_t = & \alpha_0(1 - DNEG_t) + \sum_{i=1}^4 \alpha_i DUMMY_{it} \times (1 - DNEG_t) \\
& + \alpha_5 DNEG_t + \sum_{i=1}^4 \alpha_{i+5} DUMMY_{it} \times DNEG_t \\
& + \beta_0(1 - DNEG_t) \times ROE_t + \sum_{i=1}^4 \beta_i DUMMY_{it} \times (1 - DNEG_t) \times ROE_t \\
& + \beta_5 DNEG_t \times ROE_t + \sum_{i=1}^4 \beta_{i+5} DUMMY_{it} \times DNEG_t \times ROE_t \\
& + \gamma_0(1 - DNEG_t) \times RDBV_t + \sum_{i=1}^4 \gamma_i DUMMY_{it} \times (1 - DNEG_t) \times RDBV_t \\
& + \gamma_5 DNEG_t \times RDBV_t + \sum_{i=1}^4 \gamma_{i+5} DUMMY_{it} \times DNEG_t \times RDBV_t + \varepsilon_t,
\end{aligned} \tag{11}$$

---

<sup>15</sup>In particular, Leibowitz (1999) suggests three reasons for differences between positive and negative *ROEs* and their relation with the market-to-book ratio. First, negative earnings could result from conservative accounting, making the relation between market-to-book and negative *ROE* significantly negative. Second, negative earnings as a result of negative transitory items that do not represent cash flows (e.g., changes in accounting methods) make the relation between market-to-book and *ROE* negative. Finally, negative earnings as a result of negative transitory items that do represent cash flows (e.g., adverse litigation outcome) have no significant effect on the relation between *ROE* and market-to-book.

where  $DNEG_t$  is the sign of earnings dummy variable,  $DUMMY_1$  (*FIRM*-dummy) is equal to 1 if a firm is a pioneering drug manufacturer, and 0 otherwise,  $DUMMY_2$  (*MS*-dummy) is equal to 1 if a firm has a market share in the upper half of the sample, and 0 otherwise,  $DUMMY_3$  (*PAT*-dummy) is equal to 1 if a firm has a number of patents per dollar of average past R&D that exceeds the median sample value, and 0 otherwise, and  $DUMMY_4$  (*RDG*-based) is equal to 1 if a firm has a growth in R&D that exceeds median sample growth, and 0 otherwise. I prefer to interact dichotomized versions of the value driver proxies with *ROE* and *RDBV* over the originally continuous value driver proxies, since the interaction effects are predicted to be nonlinear (see the table with partial derivatives on page 9). One derives the valuation coefficients for a specific firm by adding the estimated  $\alpha$ 's ( $\beta$  and  $\gamma$  respectively) multiplied by the values on  $DUMMY_{it}$  and  $DNEG_{it}$  that apply to that firm for year  $t$ . For example, the coefficient on *ROE* for a pioneering drug manufacturing with positive earnings, large market share, high relative patent output, and low growth is:  $\beta_0 + \beta_1 + \beta_2 + \beta_3$ . The coefficient on *ROE* for a generic drug manufacturer with an accounting loss, small market share, no patents and large growth is:  $\beta_5 + \beta_9$ . As indicated earlier, there are 32 possible summing schemes.

The specification of regression equation (11) with the five dummy variables and the interaction terms allows me to study the effect of each value driver proxy, discussed in section 5, on the valuation coefficients while controlling for all value driver proxies simultaneously. The  $\beta_0$  to  $\beta_4$  coefficients are the *ROE* valuation coefficients for the positive *ROE* firms, and the  $\beta_5$  to  $\beta_9$  coefficients are valuation coefficients on *ROE* for the negative earnings firms. The effect of *MS* on the *ROE* coefficient after controlling for the sign of earnings, type of drug firm, patent success, and growth, is reflected in  $\beta_2$  (for positive earnings firms) and  $\beta_7$  (for negative earnings firms). For example, a positive  $\beta_2$  indicates an upward effect on the *ROE* valuation coefficient by having a larger market share (see hypothesis 2). Based on the findings in earlier mentioned studies, the various possible sums of the *ROE*  $\beta$ 's for negative earnings firms, i.e. sums consisting of elements from  $\{\beta_5, \beta_6, \beta_7, \beta_8, \beta_9\}$ , is expected to be close to zero or negative. A sum close to zero is consistent with the abandonment option hypothesis discussed by Hayn (1995), Burgstahler and Dichev (1997) and Barth et al. (1998). A significantly negative sum would be consistent with the empirical findings of Burgstahler (1998) and Leibowitz (1999).

The empirical specification in eq.(11) also allows me to examine the difference in R&D valuation between positive and negative earnings firms, and the effects of the four value driver proxies on the R&D valuation coefficient. The  $\gamma$  coefficients of positive earnings firms ( $\gamma_0$  to  $\gamma_4$ ) only differ from the  $\gamma$  coefficients of negative earnings firms if the former type of firm has different values on  $\lambda$ ,  $\omega$ ,  $g$  and  $r$  than the latter. For example, I refer to the above studies that found lower earnings persistence for loss firms than for positive earnings firms. Lower earnings persistence reduces  $\omega$  in eq.(4) and therefore reduces (increases) the coefficient on  $RDBV$  if  $g < r$  ( $g > r$ ), ceteris paribus.

I estimate the coefficients in equation (11) on a pooled time series - cross-section dataset of pharmaceutical firms. I employ a multivariate outlier deletion rule to eliminate the 1% most extreme observations on both tails of the DFFITS distribution (see Belsley et al. (1980) for more details). The t-statistic on a coefficient indicates whether a value driver ( $FIRM$ ,  $MS$ ,  $PAT$  or  $RDG$ ) has a statistically significant effect on a valuation coefficient, i.e. the intercept,  $ROE$  and  $RDBV$  coefficient respectively. In order to compare the estimation results to prior studies, I also estimate the following four benchmark models on the pooled dataset:

$$\text{Model 1 : } MTB_t = \alpha + \beta ROE_t + \varepsilon_t$$

$$\text{Model 2 : } MTB_t = \alpha + \beta ROE_t + \gamma RDBV_t + \varepsilon_t$$

$$\text{Model 3 : } MTB_t = \alpha_0(1 - DNEG_t) + \alpha_1 DNEG_t + \beta_0(1 - DNEG_t) \times ROE_t + \beta_1 DNEG_t \times ROE_t + \varepsilon_t$$

$$\text{Model 4 : } MTB_t = \alpha_0(1 - DNEG_t) + \alpha_1 DNEG_t + \beta_0(1 - DNEG_t) \times ROE_t + \beta_1 DNEG_t \times ROE_t + \gamma_0(1 - DNEG_t) \times RDBV_t + \gamma_1 DNEG_t \times RDBV_t + \varepsilon_t$$

The first benchmark model simply specifies the market-to-book ratio as a linear function of  $ROE$ . The second model adds a separate term for R&D to the equation, and is identical to eq.(10). The third benchmark model is similar to the first model, but distinguishes between positive and negative earnings firms. I allow positive earnings firms to have a different intercept and slope coefficient on  $ROE$  from the coefficients of the negative earnings firms. Finally, the fourth benchmark is similar to the second model, but allows the intercept, coefficient on  $ROE$  and  $RDBV$  to differ between positive and negative earnings firms.

## 7 Sample and Descriptive Statistics.

### 7.1 Sample Selection and Data.

The sample consists of firms primarily active in the “Pharmaceutical Preparation” industry (SIC 2834), i.e. firms controlled by the FDA. I do not include firms from other pharmaceutical 4-digit industries: “Medicinal Chemicals And Botanical Products” (SIC 2833), “In Vitro And In Vivo Diagnostic Substances” (SIC 2834), “Biological Products, Except Diagnostic Substance” (SIC 2836), since these are not the typical pioneering and generic drug firms. The first year in the sample is 1975 since accounting rules for R&D were modified in 1975.<sup>16</sup> Similar to Leibowitz (1999), I eliminate firm-year observations with negative book values of equity (4.6% of the original sample), since *ROE* for these firm-years cannot be interpreted in economic terms, i.e., firms with losses and negative *BV* have a positive *ROE*. The final sample consists of 35 firms, listed in appendix C, and 620 firm-year observations.

The firm-year observations on *MTB*, *ROE* and *RDBV* are matched with the value driver proxies described in section 5. I use Moody’s 1998 Industrial Manual, OTC Unlisted Manual, OTC Industrial Manual, Bioscan (Febr 1999), and Securities Data Company’s Platinum M&A 1999 database to determine a sample firm’s subsidiaries, ownership and acquisition date. The Moody’s manual reports company name changes, important information for later name matches. I finally track the parent and subsidiary names in the FDA and patent databases.

The computation of the *FIRM* and *MS* variables is based on data provided by FDA. In particular, the Freedom of Information (FOI) Office of FDA provides comprehensive data on drugs approved by the FDA, such as tradename, generic name, firm, FDA approval date, and withdrawal date. However, the FOI CDROM has no information on therapeutic class. I collect

---

<sup>16</sup>In particular, in 1974 FASB published SFAS No.2 “Accounting for Research and Development Costs” to standardize and simplify accounting practice by requiring *all* R&D costs be charged to expense when incurred. Before SFAS No.2, the lack of explicit rules resulted in a broad range of different procedures being adopted for reporting purposes. Also, before 1975 some firms that were extensively involved in R&D did not disclose specific data regarding their R&D programs. R&D expenditures were often aggregated with other items on the income statement. The firms that did specify the amount of R&D in their financial statements, adopted one of the three following reporting alternatives. The majority included R&D expenditures as a separate line item in the income statement. Some included it in a footnote, while others reported it elsewhere in the unaudited sections of the financial statements (Dukes 1974, 18-21).

therapeutic class information from FDA’s National Drug Code (NDC) list, available at the FDA website. I match the FOI data with the therapeutic class based on tradename, generic name, firm and approval date. I use the Merck Index to assign a therapeutic class to drugs that could not be matched with the NDC list. I provide a list of therapeutic classes in appendix D. The final (merged) dataset contains 9,254 drugs approved between 1950 and 1998. That number includes all approved drugs, not only the number of drug approvals for the sample firms and their subsidiaries (3,168 drugs).

For the computation of *PAT*, I use information from the *U.S. Patents* database of Community of Science, a comprehensive bibliographic patent database that contains more than 1.7 million patents issued since 1975 (the first sample year). I search the database on parent and subsidiary name. I aggregate the patent application information at a firm-year level by adding all patents of a parent and its wholly owned subsidiaries in each year. The final patent dataset contains 41,834 patents for the 35 sample firms.

## 7.2 Descriptive Statistics.

### 7.2.1 Descriptive Statistics for the Pooled Sample.

Table 1 presents distributional features of the key Compustat variables in the study for the pooled sample: *MTB* (market-to-book), *ROE* (reported return on equity), *RDBV* (R&D to book value of equity ratio), *RDG* (shortterm growth in R&D), *LTRDG* (longterm growth in R&D), and *BVG* (yearly growth in book value of equity). The distribution of *ROE* is skewed to the left, since mean *ROE* is 4% lower than median *ROE*. The median market-to-book ratio is 3.6, much larger than the overall median ratio of 1.13 for all Compustat firms (Beaver and Ryan 2000). The distribution of *MTB* is skewed to the right. The median ratio of yearly R&D investments to book value is 11.7% with an interquartile range of 10.6%. Often in the literature on pharmaceuticals, advertisement is considered as an important intangible (Matraves 1999). For the current sample, the median ratio of advertisement expenses to book value of equity is 6% with an interquartile range of 9.4%. Thus, the magnitude of advertisement expenses is almost half that of R&D, with a larger variance.

The majority of firms in the sample are pioneering firms, so median  $FIRM$  is larger than 0.5. Median market share is 5.2% with an interquartile range of 6.2%. Patents ( $PAT$ ) are expressed in units of million dollars of R&D. Ninety percent of the observations are between 0 and 2.5 patents per million dollar spent on R&D, with a median of 0.25 patents. The short-term R&D growth measure,  $RDG$ , is much more volatile than the longterm measure,  $LTRDG$ . For example, the interquartile range of  $RDG$  is 20.5%, compared to 15.3% for  $LTRDG$ . This finding is not surprising, given that  $LTRDG$  is based on three year averages. By inspecting the difference between means and medians, the distribution of  $RDG$  is much more affected by extreme values than the distribution of  $LTRDG$ . Finally, the distribution of the growth rate of reported book value of equity,  $BVG$ , is very similar to the distribution of shortterm growth in R&D,  $RDG$ . That finding supports the assumption in the valuation model and  $ROE$  prediction model that  $k = g$ .

### 7.2.2 Descriptive Statistics by *High* and *Low* Value Driver Groups.

Table 2 contains medians of key variables calculated by “high” and “low” value driver groups. That is, I group the sample firms into two subsamples based on  $DFIRM$ ,  $MS$ ,  $PAT$  and  $RDG$  respectively. On the first line, I report the median for each high and low group, the difference between the group medians, and on the second line, I report the interquartile range (Q3-Q1) within each group and the Wilcoxon Z-score to test the significance of the difference in medians. The Wilcoxon Z-score is normally distributed.

The firms classified as generic drug manufacturers have a median proportion of 7.9% of pioneering drugs in their drug portfolio, whereas the pioneering firms have a median of 84.8% of pioneering drugs in their portfolio. Generic firms ( $FIRM \leq 0.5$ ) represent 36% of the total sample. Low  $MS$  firms have a median of 3.4% market share versus 7.6% in the high  $MS$  group. The patent output variable ( $PAT$ ) shows a median of 7 patents per 100 million R&D in the low  $PAT$  group versus 53 patents in the high  $PAT$  group. Finally, low growth firms have a median growth of 8.1% versus 28.3% for the high growth firms. Next, I discuss table 2 sequentially by each of the variables in the first column.

The market-to-book ratio is clearly different between generic and pioneering drug manufacturers, as expected: 3.29 for the former and 3.84 for the latter. Market share is also a variable that strongly differentiates between low and high market-to-book ratios. Both *PAT* and *RDG* are not able to discriminate between market-to-book ratios, since the difference in median *MTB* between the high and low *PAT* and *RDG* group are not statistically significant. The fact that low *RDG* and high *RDG* firms have almost identical market-to-book ratios is rather surprising. One explanation for the low association between R&D growth and *MTB* is that historical growth is a bad proxy for future growth opportunities, which are reflected in the market-to-book ratio. Another explanation is that the degree of R&D reporting bias in *BV* varies between the two growth portfolios, and thus differentially affects the denominator of the *MTB* ratio. The bias explanation is based on eq.(B.4): reported *BV* is less understated for high growth firms than for low growth firms. Since R&D reporting bias in *BV* is inversely related with growth, a high *MTB* ratio for the low-growth portfolio might result from an understated *BV*, whereas a high *MTB* ratio for the high growth portfolio might reflect the existence of future growth opportunities .

As expected, both  $\omega$  proxies, *DFIRM* and *MS*, are good discriminators for *ROE*. In particular, low values on *DFIRM* and *MS* correspond with lower values on *ROE*. Notice that even the median level of *ROE* in the generic group and low *MS* group is still high, i.e. 16.4%. The higher *ROE* level for pioneering drug firms compared to that of generic firms, is consistent with their higher market-to-book ratios. That could indicate either that pioneering firms are more profitable in an economic sense, or that the R&D reporting bias in *ROE* is positive for pioneering firms. Since I expect economic depreciation of R&D capital ( $1 - \lambda$ ) to be lower for pioneering drug firms than for generic firms (pioneering firms can profit more and longer from the results of their investments in R&D capital resulting from patent positions, superior knowledge, etc.), the latter explanation is highly plausible. Table 2 indeed shows a significantly higher *PAT* value for pioneering firms compared to the value for generic firms. Also, I show in eq.(B.6) that the sign and magnitude of R&D bias in *ROE* depends on the difference between R&D growth and reported *ROE*. *RDG* is 14.4% for pioneers versus 20.2% for generics (see below for a discussion), and the reported *ROE* level is 21.3% for pioneers versus 16.3% for generics, indicating an upward bias in *ROE* for pioneering firms, and downward bias in *ROE* for generic firms,

i.e., *ROE* reported by generic firms is lower than it would be under R&D capitalization. The observation on the bias effect for the pioneering firms is consistent with claims by Taylor (1999) and PhRMA (1999) that accounting profitability measures for pioneering drug firms are upward biased. Finally, the level of *ROE* is almost identical in the low and high *PAT* and *RDG* groups.

Pioneering drug firms have a median *RDBV* of 0.146 compared to 0.089 for generic firms. That is consistent with relatively higher R&D investments undertaken by pioneering firms in their long drug development process. Since firms with smaller market shares are typically generic firms, low *MS* firms also have lower *RDBV* values compared to high *MS* firms. Finally, *RDBV* is not different between high and low *PAT* and *RDG* groups.

Pioneering drug firms typically have high market shares, and generic firms typically have small market shares, indicating *MS* and *FIRM* are highly positively correlated. Proof of the high positive correlation is shown in the *MS* column of table 2: there is a significantly different value in *FIRM* between the low *MS* and the high *MS* groups (0.20 versus 0.75).

The patent output per million dollar invested in R&D is reflected in *PAT*. Since patents are key to pioneering drug firms to protect their developed drugs from competitors for a number of years, it is not surprising that pioneering firms have a significantly higher value on *PAT* than generic firms: 0.25 versus 0.20. Generic firms mainly have process patents, while pioneering firms have product patents. High *MS* firms also have significantly higher patent output than low *MS* firms: 0.27 versus 0.16 patents per million R&D. That is not surprising given the high correlation between *FIRM* and *MS*.

Finally, annual growth in R&D investments is reflected in *RDG*. As indicated above, pioneering firms have a significantly smaller growth in R&D investments than generic firms: *RDG* for pioneers is 14.4% versus 20.2% for generics. The high R&D growth rate of generic manufacturers is mainly attributed to the 1984 Hatch-Waxman Act that promoted the supply of generic drugs. At the same time of the Hatch-Waxman Act, changes in the demand for drugs - brought on by newer forms of health care delivery and financing - have influenced both the frequency with which generic and pioneering drugs are prescribed and the prices paid for them.

I refer to paper 2 of the Congressional Budget Office study on competition in the pharmaceutical industry for more details (CBO 1998). Not surprisingly, firms with lower market shares (typically generic firms) have a higher growth in R&D than firms with higher market shares: *RDG* is 17.2% for low *MS* firms versus 14.2% for high *MS* firms.

### 7.2.3 Correlation Analysis.

Table 3 presents descriptives on the bivariate correlation structure of the key variables. Assuming no spurious correlation due to the common deflator (*BV*), market-to-book ratios are highly correlated with *ROE* (42% Spearman corr.) and *RDBV* (30%). In other words, both *ROE* and *RDBV* are highly value-relevant. From the four value driver proxies, only *FIRM* and *MS* are correlated with *MTB*, *ROE* and *RDBV*. The high positive correlations between *FIRM* and *RDBV*, and between *MS* and *RDBV* confirm the findings in table 2. That is, pioneering firms are more research intensive and have larger market shares than generic firms.

*MS* is significantly positively correlated with *FIRM*, confirming the results in table 2. The high correlation between *FIRM* and *MS* strengthens the approach I take to formulate hypotheses 1 and 2. That is, both hypotheses focus on the effect of the competitive structure on  $\omega$ . I use *FIRM* and *MS* as proxies for the competitive structure, and therefore assume their main effect is on the  $\omega$  parameter in the valuation model, and not on  $g$ ,  $r$  or  $\lambda$ . *RDG* is significantly related with *LTRDG* (both are different proxies for R&D growth). *PAT* is weakly positively associated with *MS* and *FIRM*, since patenting firms are usually pioneering drug firms with larger market shares. Consistent with Scott Morton (1999), *FIRM* is negatively related with growth in R&D: the Hatch-Waxman Act of 1984 made it easier for generic firms to introduce their generic substitutes to the market place, and therefore resulted in faster growing R&D investments for generic firms. Since generic firms typically have smaller market shares, *MS* is negatively associated with R&D growth.

## 8 Regression results.

This section discusses the estimation results of the valuation equation (11). I also report and discuss the estimation of the four benchmark regressions discussed on page 21. All models are estimated on the pooled cross-sectional time-series dataset. The second to fifth column in table 4 contains the benchmark estimation results, while the last column in the table contains the results for the key valuation model.

The first benchmark regression (model 1) with *ROE* as the only explanatory variable in the valuation equation has very little explanatory power : the adjusted  $R^2$  is 0.1%. Once scaled R&D is added to the valuation model, the  $R^2$  increases from 0.1% in model 1 to 49% in model 2. The dramatic increase in  $R^2$  and the significance of the *RDBV* coefficient is evidence that  $\lambda$  is not zero, i.e., economic durability of R&D capital. The highly significant coefficient on *RDBV* supports the valuation specification that includes R&D as a separate term, and therefore suggests a non-zero  $\lambda$ . Lev and Sougiannis (1999), among others, also find significant valuation multiples on R&D expenses for R&D intensive firms. However, in the rest of this section, I extend previous studies by explaining the determinants of cross-sectional differences in the valuation multiples on both R&D and *ROE*.

Similar to model 1, the third benchmark model (model 3) only has *ROE* as explanatory variable, but allows a different intercept and slope between positive and negative earnings firms. Consistent with the findings of Leibowitz (1999) among others (see section 6), the *ROE* valuation coefficient is significantly positive for positive earnings firms, and significantly negative for negative earnings firms. The significantly negative *ROE* coefficient for negative earnings firms contradicts the predictions in the option-style valuation model of Burgstahler and Dichev (1997): the coefficient on negative *ROE* should be small and insignificant, consistent with the adaptation value hypothesis. However, if losses are generated by conservative accounting, such as R&D accounting, and if they do not reflect economic losses, then the market can indeed place a significantly negative valuation multiple on negative *ROE* (Leibowitz 1999).

The fourth benchmark model adds *RDBV* to model 3 and also distinguishes between positive and negative earnings firms. Interestingly, the slope coefficient on *ROE* is almost zero,

which is consistent with the abandonment option hypothesis. In contrast with model 3, model 4 controls for accounting conservatism by having  $RDBV$  as a separate term in the equation. The R&D coefficient is significantly positive for both negative and positive earnings firms, although much lower for negative earnings firms (7.53 versus 17.51 for positive earnings firms). The lower  $RDBV$  coefficient for negative earnings firms is most likely related to the difference in abnormal profitability persistence between positive and negative earnings firms. Consistent with prior literature, unreported results show that negative earnings firms have much lower abnormal profitability persistence ( $\omega$ ) than positive earnings firms. In section 3, I demonstrate that the partial derivative of the R&D valuation multiple with respect to  $\omega$  is positive when  $g > r$ . Thus, lower  $\omega$  values imply lower R&D valuation multiples. Finally, the change in explanatory power by adding  $RDBV$  to model 3 is again enormous: the adjusted  $R^2$  increases from 21% in model 3 to 56% in model 4.

The results for the key valuation model in this study are reported in the last column (“model 5”) of table 4. Overall, the explanatory power of the model that includes the value driver proxies is higher than that of the benchmark models: adjusted  $R^2$  is 66% compared to 56% of model 4. I discuss the effects of each of the four value driver proxies separately.

First, I focus on the effect of abnormal profitability persistence ( $\omega$ ) on the valuation coefficients. Both  $FIRM$  and  $MS$  measure  $\omega$ , and they affect the valuation coefficients through  $DUMMY_1$  and  $DUMMY_2$  respectively. Since both measures are highly correlated, I discuss their joint effect on equation (11). Both hypothesis 1 and 2 are strongly supported by the regression results. As expected, pioneering drug firms with large market shares have a significantly lower intercept (both coefficients on  $DUMMY_1$  and  $DUMMY_2$  are significantly negative) and higher coefficient on  $ROE$ . That is a result of pioneering firms having more persistent abnormal profitability than generic firms, since the latter experience more immediate competition when they introduce a drug (see section 2.1 for details). However, for the negative earnings firms (i.e.,  $NI < 0$ ) that represent 11% of the sample, both  $MS$  and  $FIRM$  effects on the intercept and  $ROE$  coefficient compensate each other, resulting in a zero effect on the two valuation coefficients. The effects of  $FIRM$  and  $MS$  on the R&D valuation coefficient depend on whether growth in R&D investments exceeds the cost of capital. The annual R&D growth rate for

pioneering firms is 14.4% (see table 2) and cost of capital estimates for these firms are reported by Taylor (1999) to vary between 15.6 (American Home Products) and 19.07 (Syntex) with an average of 16.7 in the period 1975-1991.<sup>17</sup> That would indicate that  $g < r$  for pioneering firms and, as a consequence, both *FIRM* and *MS* are expected to have a negative effect on *RDBV*. Table 4 shows a negative - but insignificant - effect of *FIRM* and *MS* on the R&D valuation coefficient.

Second, I discuss the effect of *PAT* on the *RDBV* valuation coefficient. The regression results strongly support hypothesis 3 that predicts a positive effect of patent success on the R&D valuation multiple. The coefficient on *DUMMY*<sub>3</sub> (*PAT*) is highly significant in the predicted direction. That is, firms with relatively more patents per dollar invested in R&D have a significantly higher valuation multiple on R&D. Depending on the rate of success of the R&D investments, firms will have higher or lower valuation multiples on R&D, after controlling for competition ( $\omega$ ) and growth ( $g$ ) effects. Thus, in addition to model 2 and model 4 in table 4 that include a separate R&D term, model 5 allows me to distinguish between the R&D coefficients of firms with different rates of R&D success. There is no significant effect of *PAT* on the R&D coefficient of negative earnings firms.

Third, the effect of growth in R&D investments on the valuation coefficients in equation (11) is reflected in the coefficients on *DUMMY*<sub>4</sub>. As predicted by hypothesis 4, *RDG* negatively affects the intercept and positively affects the coefficient on *ROE* in a statistically significant way. That is, high growth firms have a higher valuation multiple on the current level of *ROE*. Finally, the effect of growth on the R&D valuation multiple is predicted to be positive if the R&D growth rate is larger than  $\sqrt{\frac{R\lambda}{\omega}}$ , and negative otherwise. I refer to section 3 for more details. Table 4 shows a coefficient on *DUMMY*<sub>4</sub> (i.e., effect of *RDG*) for R&D that is not significantly different from zero. The effect is slightly positive for positive earnings firms, and negative for negative earnings firms. Prior literature (e.g., Hayn (1995)) documents lower persistence of losses compared to profits, i.e.  $\omega_{NI<0} < \omega_{NI>0}$ . As a result of a lower  $\omega$ , loss firms are expected to have a much higher pivot point  $\sqrt{\frac{R\lambda}{\omega}}$  than profit firms. For example, loss firms

---

<sup>17</sup>Ibbotson's Cost of Capital Quarterly 1999 Yearbook reports a median CAPM cost of capital of 15.5 for SIC 2834, with an interquartile range of [Q1,Q3]=[12.2,18.9]. When the small size premium is added to the CAPM estimate, the median becomes 17.6 and [Q1,Q3]=[13.5,20.8].

with  $r = 0.18$ ,  $\lambda = 0.7$  and  $\omega = 0.1$  have  $\sqrt{\frac{R\lambda}{\omega}} = 2.87$ , whereas profit firms with  $r = 0.16$ ,  $\lambda = 0.8$  and  $\omega = 0.7$  have  $\sqrt{\frac{R\lambda}{\omega}} = 1.15$ . Unreported analysis shows that the observed growth rate of loss firms is 0.19 compared to 0.15 for profit firms. Taking the above pivot points, the observed growth rate of 0.15 is exactly equal to the growth condition, whereas the observed growth rate of 0.19 is much below the pivot value of 2.87. The effect of growth on the valuation multiple of R&D is therefore predicted and found to be negative. The fact that the coefficient of -4.29 for the loss firms is not statistically significant is probably related to the high variance in annual growth for the sample loss firms ( $[Q1, Q3] = [0.01, 0.42]$  compared to  $[0.12, 0.18]$  for profit firms).

I illustrate the valuation specification for four different types of companies: (1) pioneering drug firms with a large market share, high patent output, low growth, and positive earnings, (2) generic drug firms with a small market share, low patent output, high growth, and positive earnings, (3) same as case 1 but with negative earnings, and (4) same as case 2 but with negative earnings. The following regressions are obtained by multiplying the  $DUMMY_i$  value with the estimated respective coefficients reported in table 4 (model 5):

$$\begin{aligned}
 1) \text{ pioneer } (NI > 0) : & \quad MTB_t = -1.96 + 14.2 ROE_t + 20.31 RDBV_t \\
 2) \text{ generic } (NI > 0) : & \quad MTB_t = 1.46 + 4.48 ROE_t + 18.52 RDBV_t \\
 3) \text{ pioneer } (NI < 0) : & \quad MTB_t = 2.31 - 3.47 ROE_t + 0.13 RDBV_t \\
 4) \text{ generic } (NI < 0) : & \quad MTB_t = 1.14 - 3.15 ROE_t - 1.03 RDBV_t
 \end{aligned}$$

For example, the intercept value of -1.96 for the pioneering drug firms with positive earnings is calculated from table 4 as  $2.52 - 1.72 - 1.63 - 1.13 = -1.96$ . The intercept value of 1.46 for the generic firms with positive earnings is calculated as  $2.52 - 1.06 = 1.46$ . Pioneering drug manufacturers with positive earnings have a much higher  $ROE$  valuation multiple than generic firms due to higher abnormal profitability persistence: the  $ROE$  coefficient is 14.2 for pioneers versus 4.48 for generics. The slightly higher R&D valuation multiple for pioneering firms is a result of a higher  $\lambda$  for these firms, measured by  $PAT$  (see table 2). The  $\lambda$  effect dominates the  $\omega$  effect (measured by  $FIRM$  and  $MS$ ) on the R&D multiple: the coefficient on  $DUMMY_3$  is 10.05 versus the sum of the coefficients of  $DUMMY_1$  and  $DUMMY_2$  of -7.47. The coefficient on  $ROE$  becomes slightly negative for loss firms: -3.47 for pioneers and -3.15

for generics. The abandonment option hypothesis would predict a coefficient close to zero on *ROE*. The negative valuation coefficient might be caused by conservative accounting methods used by negative earnings firms, that are not captured by the R&D term. Examples of such conservative accounting methods are the immediate write-off of purchased R&D (i.e., a negative special item in the year of the write-off), and accelerated depreciation of fixed assets. Again, I emphasize that the loss firm-year observations only represent 11% of the total sample. The R&D multiple is almost zero for both types of firms when they incur accounting losses.

To summarize, all four valuation hypotheses are strongly supported by the regression results. Not only do I find that R&D should be included as a separate term in the valuation model for research intensive firms, but the regression specification in equation (11) allows me to simultaneously examine the nonlinear effects of competitive structure, R&D success and growth in R&D investments on the valuation multiples of *ROE* and scaled R&D. Loss firms, representing 11% of the sample, have much lower valuation multiples on *ROE* and R&D than profit firms, indicating a smaller value relevance of *ROE* and R&D for negative earnings firms. Finally, I estimated an undeflated valuation specification as a sensitivity check, and the results remain unchanged.

## 9 Summary and Concluding Remarks.

This paper examines how accounting information is valued in a research-intensive industry, the pharmaceutical industry. I consider both economic context and R&D reporting bias as main sources of differences in the valuation of accounting information. Prior literature - and especially the intangibles literature - has taken a different approach as to how accounting and nonaccounting information is valued by the stock market. In particular, most prior studies include nonaccounting information (describing a firm's economic context) as separate linear terms in a valuation equation. I argue that economic context determines the valuation coefficients of the accounting variables. Bowen and Shores (2000), Hand (2001) and Shortridge (2001) are among the few studies that interact economic context variables with accounting variables for a sample of research-intensive firms, but none of these studies provide a theoretical valuation model to study these interaction effects. By using a theoretical model, I show that the interac-

tion effects are often highly nonlinear in nature.

The valuation model relates the market-to-book ratio to book return on equity (*ROE*) and R&D investments deflated by book value of equity (*RDBV*). Abnormal profitability persistence ( $\omega$ ), economic depreciation of R&D capital ( $\lambda$ ), and growth in R&D investments ( $g$ ) are the *value drivers* affecting the valuation coefficients of *ROE* and *RDBV*. I interact four proxies for these value drivers with the intercept, *ROE* and *RDBV* in the final regression equation. The four hypotheses on the effect of each of the value driver proxies are strongly supported by the regression results. First, differentiating between two types of drug firms (*FIRM*) shows that pioneering drug manufacturers have a higher coefficient on *ROE* and lower coefficient on *RDBV* due to higher persistence in abnormal profitability. For the same reason, firms with larger therapeutic market shares (*MS*) have a higher coefficient on *ROE* and lower coefficient on *RDBV*. Third, firms with more patent per dollar R&D investment (*PAT*) have a higher valuation multiple on scaled R&D due to a lower economic depreciation of R&D capital. Finally, higher growth in R&D investments positively affects the *ROE* valuation coefficient, and has no effect on the R&D multiple.

There are several limitations to this study. First, I derive a parsimonious valuation model in which R&D is the only source of reporting bias. Other sources of accounting conservatism, such as expensing of advertisement, are not considered. However, I argue that R&D is the key production factor for the sample companies, and the immediate expensing of the R&D investments most likely captures a large proportion of accounting conservatism sources for pharmaceutical firms. Second, the findings are based on a small sample of pharmaceutical firms, and might not be valid for a larger universe of Compustat firms. I suspect the conclusions of the study might be valid for other research intensive firms. Third, there are other sources of cross-sectional differences in the market-to-book ratio of pharmaceutical firms that are not used in the empirical analysis of the study. For example, I do not consider the valuation effects of strategic alliances between companies (Hagedoorn and Schakenraad 1994; Nixon 1996).<sup>18</sup> I

---

<sup>18</sup>For example, in November 1997 Merck signed a 145 million dollar research collaboration agreement with Biogen to develop VLA-4 for asthma and inflammation. Another common alliance is when pharmaceutical firms have marketing agreements, where one firm markets a drug that is developed by the second firm in a specific geographic region. For example, Glaxo Wellcome signed a marketing agreement with the Japanese drug manufacturer Fujisawa in 1994 to sell Imuran in the U.S. Finally, drug firms often have agreements with

also do not consider the outcome of clinical trials as a success indicator of R&D (I only use patents as an R&D output measure). Despite these limitations, I find strong support for the hypotheses on the effects of the four value drivers on the valuation coefficients of *ROE* and *RDBV*.

---

universities: drug firms finance research activities performed by academic laboratories to acquire familiarity with basic knowledge in a particular field. In addition, these agreements are important sources for recruiting qualified scientists and researchers (Gambardella 1995).

## References

- Allen, R.F. and A.S. Hagin. 1989. "Scale Related Efficiencies as a (Minor) Source of the Profits-Market Share Relationship." *The Review of Economics and Statistics* 71 (3): 523–526 (August).
- Amir, E. and B. Lev. 1996. "Value-Relevance of Nonfinancial Information: the Wireless Communications Industry." *Journal of Accounting and Economics* 22 (1): 3–30 (December).
- Barth, M.E., W.H. Beaver, and W. Landsman. 1998. "Relative Valuation Roles of Equity Book Value and Net Income as Function of Financial Health." *Journal of Accounting and Economics* 25 (1): 1–34.
- Beaver, W.H. and S.G. Ryan. 2000. "Biases and Lags in Book Value and Their Effects on the Ability of the Book-to-Market Ratio to Predict Book Return on Equity." *Journal of Accounting Research* 38 (1): 127–148 (Spring).
- Belsley, D., R. Kuh, and R. Welsch. 1980. *Regression Diagnostics*. New York: Wiley.
- Ben-Zion, U. 1984. Pages 299–312 in *The R&D and Investment Decision and its Relationship to the Firm's Market Value: Some Preliminary Results*. Chicago: The University Press of Chicago.
- Bernard, V.L. 1994, January. "Accounting-Based Valuation Methods, Determinants of Market-to-Book Ratios, and Implications for Financial Statement Analysis." Working paper, University of Michigan.
- Biddle, G., P. Chen, and G. Zhang. 2001, January. "When Capital Follows Profitability: Non-Linear Residual Income Dynamics." Working paper, Hong King University of Science and Technology.
- Bowen, R.M. and D. Shores. 2000, March. "Economic Context and the Value Relevance of Accounting Data." Working paper, University of Washington.
- Burgstahler, D. 1998. "Discussion of "Combining Earnings and Book Value in Equity Valuation"." *Contemporary Accounting Research* 15 (3): 325–341 (October).
- Burgstahler, D. and I. Dichev. 1997. "Earnings, Adaptation, and Equity Value." *Accounting Review* 72 (2): 187–215 (April).
- CBO. 1998. *How Increased Competition from Generic Drugs Has Affected Prices and Returns in the Pharmaceutical Industry*. Congressional Budget Office, Washington: CBO.
- Collins, D.W., M. Pincus, and H. Xie. 1999. "Equity Valuation and Negative Earnings: The Role of Book Value of Equity." *Accounting Review* 74 (1): 29–61 (January).

- Dechow, P.M., A.P. Hutton, and R.G. Sloan. 1999. "An Empirical Assessment of the Residual Income Valuation Model." *Journal of Accounting and Economics* 26:1–34.
- Dimasi, J.A., H.G. Grabowski, and J.M. Vernon. 1995. "R&D Costs, Innovative Output and Firm Size in the Pharmaceutical Industry." *International Journal of the Economics of Business* 2 (2): 201–219.
- Dukes, R.E. 1974, May. "Market Evaluation of Accounting Information: A Cross Sectional Test of Investor Response to Expensing Research and Development Expenditures." Ph.D. diss., Stanford University, Graduate School of Business.
- Egan, J.W., H.N. Higinbotham, and J.F. Weston. 1982. *Economics of the Pharmaceutical Industry*. New York: Praeger Publishers.
- Encaoua, D., B.H. Hall, F. Laisney, and J. Mairesse. 1998, June. "The Economics and Econometrics of Innovation: An Overview." Working paper, ENSAE.
- Gambardella, A. 1995. *Science and Innovation: the US Pharmaceutical Industry During the 1980s*. U.K.: Cambridge University Press.
- Geroski, P.A. 1990. In Mueller 1990, 15–34.
- Goto, A. and K. Suzuki. 1989. "R&D Capital, Rate of Return on R&D Investment and Spillover of R&D in Japanese Manufacturing Industries." *The Review of Economics and Statistics* 71 (4): 555–564 (November).
- Grabowski, H.G. and J.M. Vernon. 1992. "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act." *Journal of Law & Economics* 35 (October): 331–350.
- . 1994. "Returns on R&D on New Drug Introductions in the 1980s." *Journal of Health Economics* 13:383–406.
- Hagedoorn, J. and J. Schakenraad. 1994. "The Effect of Strategic Technology Alliances on Company Performances." *Strategic Management Journal* 15:291–309.
- Hall, B.H., A. Jaffe, and M. Trajtenberg. 1999, June. "Market Value and Patent Citations: A First Look." Working paper, NBER.
- Hand, J.R.M. 2001, December. "The Market Valuation of Biotechnology Firms and Biotechnology R&D." Working paper, UNC Chapel Hill.
- Hayn, C. 1995. "The information Content of Losses." *Journal of Accounting and Economics* 20:125–153.

- Hirschey, M. 1982. "Intangible Capital Aspects of Advertizing and R&D Expenditures." *The Journal of Industrial Economics* XXX (4): 375–390 (June).
- Hurwitz, M.A. and R.E. Caves. 1988. "Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmaceuticals." *Journal of Law & Economics* 31 (October): 299–320.
- Leibowitz, M.A. 1999. "Market-to-Book Ratios and Positive and Negative Returns on Equity." *The Journal of Financial Statement Analysis* Winter:21–30.
- Lev, B. and T. Sougiannis. 1996. "The Capitalization, Amortization, and Value-Relevance of R&D." *Journal of Accounting and Economics* 21:107–138.
- . 1999. "Penetrating the Book-to-Market Black Box: The R&D Effect." *Journal of Business Finance & Accounting* 26 (3): 419–449 (May).
- Lev, B., B. Sarath, and T. Sougiannis. 1999, May. "R&D-Related Reporting Biases and Their Consequences." Working paper, New York University.
- Malerba, F., L. Orsenigo, and P. Peretto. 1997. "Persistence of Innovative Activities, Sectoral Patterns of Innovation and International Technological Specialization." *International Journal of Industrial Organization* 15:801–826.
- Matraves, K. 1999. "Market Structure, R&D and Advertising in the Pharmaceutical Industry." *The Journal of Industrial Economics* 47 (2): 169–194 (June).
- Mueller, D.C. 1990. *The Dynamics of Company Profits: An International Comparison*. U.K.: Cambridge University Press.
- Nixon, B. 1996. "Research and Development Alliances and Accounting." *R&D Management* 26 (2): 169–175.
- O'Hanlon, J. 1997. "The Time Series Properties of the Components of Clean Surplus Earnings: UK Evidence." *Journal of Business Finance and Accounting* 23 (2): 159–189 (March).
- Ohlson, J.A. 1995. "Earnings, Book Values, and Dividends in Security Valuation." *Contemporary Accounting Research* 11 (2): 661–687.
- . 1999. "On Transitory Earnings." *Review of Accounting Studies* 4, no. 3 (forthcoming).
- Palepu, K.G., V.L. Bernard, and P.M. Healy. 2000. *Business Analysis and Valuation Using Financial Statements*. Cincinnati: South-Western College Publishing.
- Penman, S.H. 2001. *Financial Statement Analysis and Security Valuation*. New York: McGraw-Hill.
- PhRMA. 1999. *Pharmaceutical Industry Profile 1999*. Pharmaceutical Research and Manufacturers of America, Washington: PhRMA.

- Qi, D.D., Y.W. Wu, and B. Xiang. 2000. "Stationarity and Cointegration Tests of the Ohlson Model." *Journal of Accounting, Auditing and Finance* 15 (2): 141–160 (Spring).
- Scott Morton, F.M. 1998, September. "Barriers to Entry, Brand Advertising, and Generic Entry in the U.S. Pharmaceutical Industry." Working paper, University of Chicago.
- 1999. "Entry Decisions in the Generic Pharmaceutical Industry." *RAND Journal of Economics* 30 (3): 421–440 (Autumn).
- Shortridge, R.T. 2001, June. "Market Valuation of Successful Versus Non-Successful R&D Efforts in the Pharmaceutical Industry." Working paper, Ball State University.
- Taylor, C.T. 1999. "The Cash Recovery Method of Calculating Profitability: An Application to Pharmaceutical Firms." *Review of Industrial Organization* 14:135–146.
- Van Breda, M.F. 1981. *The Prediction of Corporate Earnings*. UMI Research Press.
- Weiss, L.W. 1969. "Advertising, Profits, and Corporate Taxes." *The Review of Economics and Statistics* 51 (November): 421–430.
- Zhang, X.-J. 2000. "Conservative Accounting and Equity Valuation." *Journal of Accounting and Economics* 29 (1): 125–149 (February).

## Appendix A

### Variable Definitions.

<i>Variables</i>	<i>Definition</i>
<i>ANDA</i>	abbreviated new drug application (for generic drug)
<i>BV</i>	reported book value of equity (Compustat item60)
<i>BVG</i>	yearly growth rate of book value of equity
<i>CBV</i>	book value of equity under R&D capitalization, see definition (B.1)
<i>CNI</i>	net income under R&D capitalization, see definition (B.1)
<i>CNI<sup>a</sup></i>	abnormal net income under R&D capitalization, $CNI_t - r CBV_{t-1}$
<i>CROE</i>	return on equity under R&D capitalization, see definition (B.5)
<i>CROE<sup>a</sup></i>	abnormal return on equity under R&D capitalization, $CROE - r$
<i>DEPRD</i>	economic depreciation of R&D capital as defined in (B.3)
<i>DFIRM</i>	dummy equal to one if $FIRM > 0.5$ and zero otherwise, see definition (5)
<i>DNEG</i>	dummy equal to one if a firm has negative $NI$ , and zero otherwise
<i>FIRM</i>	proportion of NDAs in the drug portfolio as defined in (5)
<i>LTRDG</i>	annualized three year growth rate of R&D as defined in (9)
<i>MS</i>	aggregate therapeutic market share as defined in (6)
<i>MTB</i>	market-to-book ratio, $\frac{MV_t}{BV_t}$
<i>NDA</i>	new drug application (for pioneering drug)
<i>NI</i>	reported net income before extraordinary items (Compustat item237)
<i>PAT</i>	number of patent scaled by average R&D expense of most recent three years
<i>RD</i>	research and development expense (Compustat item46)
<i>RDBV</i>	R&D intensity, R&D expense scaled by book value of equity
<i>RDG</i>	yearly growth rate of R&D as defined in (8)
<i>ROE</i>	reported book return on equity, $\frac{NI_t}{BV_{t-1}}$
<i>ROE<sup>a</sup></i>	abnormal book return on equity, $ROE - r$ , where $r$ is cost-of-capital
<hr/>	
<i>Parameters</i>	<i>Definition</i>
$\lambda$	economic durability of R&D capital
$\Phi_1$	parameter describing R&D capitalization bias in reported earnings, see (B.4)
$\Phi_2$	parameter describing R&D capitalization bias in reported book value, see (B.4)
$\Lambda_0$	intercept in valuation equation (4)
$\Lambda_1$	coefficient on current reported $ROE$ in valuation equation (4)
$\Lambda_2$	coefficient on current $RDBV$ in valuation equation (4)
$\omega$	AR(1) persistence parameter of $CROE^a$
$g$	firm-specific constant growth rate in R&D investments ( $G = 1 + g$ )
$k$	growth rate in book value of equity under R&D capitalization
$r$	cost-of-capital ( $R = 1 + r$ )

## Appendix B

### R&D Capitalization Bias in Earnings and Equity Book Value.

Weiss (1969) was the first to provide a framework to analyze the reporting bias in return on equity due to fully expensing of R&D. More recently, Lev et al. (1999) focus on this issue. Pharmaceutical firms (both pioneering and generic firms) invest heavily in intangible R&D assets (or R&D capital). As typically viewed by the literature, R&D capital represents the stock of *knowledge* a firm possesses at a certain point in time, and is the most important production factor for a pharmaceutical firm. The stock is accumulated through R&D efforts, and it depreciates and becomes obsolete due to such conditions as development of superior techniques and the decline in appropriability of knowledge as it diffuses (Goto and Suzuki 1989). Under current US GAAP, R&D investments are immediately expensed. As a result, two accounting effects occur. First, earnings under current GAAP differ from *capitalized* earnings (i.e. when intangible R&D capital would be capitalized and depreciated over its expected economic life) because current R&D outlays are immediately expensed. Second, book value of equity does not reflect investments in and depreciation of R&D capital.

I define the R&D capitalization effect on earnings and book value of equity as:

$$\begin{aligned}
 CNI_t &= NI_t + RD_t - DEPRD_t, \\
 CBV_t &= BV_t + \underbrace{\sum_{\tau=0}^{\infty} RD_{t-\tau} - \sum_{\tau=0}^{\infty} DEPRD_{t-\tau}}_{\text{R\&D capital}},
 \end{aligned}
 \tag{B.1}$$

where  $CNI$  is net income under R&D capitalization,  $NI$  is net income under current R&D GAAP,  $RD$  is R&D expense,  $DEPRD$  is depreciation charge on R&D capital,  $CBV$  is book value of equity under R&D capitalization, and  $BV$  is book value of equity under current GAAP. Eq.(B.1) holds for any capitalization scheme. In particular, under current US GAAP,  $RD_t$  is equal to  $DEPR_t$ , that is R&D investments are fully expensed. However, the R&D capitalization proposed in this study is one that measures economic depreciation.

To operationalize (B.1), I make two assumptions. First, I assume R&D expenses follow a non-stochastic exponential growth pattern with a constant growth rate  $g$ :

$$RD_{t+\tau} = (1 + g)^\tau RD_t \tag{B.2}$$

Second, similar to Hirschey (1982), I assume yearly R&D investments have a constant economic durability or retention rate  $\lambda$ . Assuming  $\lambda$  to be constant over time results in the following declining

balance depreciation formula:

$$DEPRD_t = (1 - \lambda) \frac{(1 + g)}{1 + g - \lambda} RD_t \quad (\text{B.3})$$

I implicitly assume  $\lambda < 1 + g$  and therefore do not allow negative growth rates of less than -1, since  $0 \leq \lambda < 1 + g$ . Firms conducting more successful R&D have a higher  $\lambda$  than firms with lower R&D success.

Finally, substituting (B.2) and (B.3) in (B.1), the bias in reported  $NI$  and  $BV$  results in expression (B.4):

$$\begin{aligned} \text{R\&D related } NI \text{ Bias} &= NI_t - CNI_t = -\frac{\lambda g}{1+g-\lambda} RD_t = -\Phi_1 RD_t, \\ \text{R\&D related } BV \text{ Bias} &= BV_t - CBV_t = -\frac{\lambda(1+g)}{1+g-\lambda} RD_t = -\Phi_2 RD_t. \end{aligned} \quad (\text{B.4})$$

where  $CNI$  is net income under R&D capitalization,  $NI$  is net income reported under current R&D GAAP,  $RD$  is R&D expense,  $CBV$  is book value of equity under R&D capitalization, and  $BV$  is book value of equity reported under current GAAP,  $g$  represents the time-constant R&D growth rate, and  $\lambda$  is the time-constant economic durability of R&D ( $1 - \lambda$  represents the R&D depreciation rate). If  $0 < \lambda \leq 1$ , then a portion of the R&D expenditures in the current period will carry over to subsequent periods and constitute an investment in intangible capital (or, R&D capital). On the other hand, if  $\lambda = 0$ , then R&D expenditures only have an influence in the current period, and should be expensed rather than capitalized. For example, unsuccessful R&D investments have little or no future benefits for a firm. The above equations implicitly assume the growth rate in R&D is smaller than the R&D depreciation rate, i.e.,  $g < 1 - \lambda$ . Consistent with Zhang (2000), higher growth ( $g$ ) increases the net income bias, while it lowers the book value bias. When there is no growth, net income equals earnings under R&D capitalization, but the equity book value bias remains different from zero (unless  $\lambda$  becomes zero).

Next, I derive the bias in return on equity,  $ROE$ , since  $ROE$  will be one of the key explanatory variables in the valuation model. Under R&D capitalization, the return on equity is defined as:

$$CROE_t = \frac{CNI_t}{CBV_{t-1}} = \frac{NI_t + \Phi_1 RD_t}{BV_{t-1} + \Phi_2 RD_{t-1}} = \frac{\frac{NI_t}{BV_{t-1}} + \Phi_1 \frac{RD_t}{BV_{t-1}}}{1 + \frac{\Phi_1}{g} \frac{RD_t}{BV_{t-1}}}. \quad (\text{B.5})$$

If the capitalization corrected  $CROE$  does not contain any reporting bias, then  $CROE$  reflects a firm's internal rate of return. In section 3, I assume a stochastic process for  $CROE$  with an expected value

equal to the cost of capital. The R&D bias in reported  $ROE$  is expressed as:

$$\text{R\&D related } ROE \text{ Bias} = ROE_t - CROE_t = \frac{\Phi_2 \frac{RD_{t-1}}{BV_{t-1}}}{1 + \Phi_2 \frac{RD_{t-1}}{BV_{t-1}}} [ROE_t - g], \quad (\text{B.6})$$

where  $ROE_t = \frac{NI_t}{BV_{t-1}}$  is book return on equity under current GAAP at time  $t$ . The above equation provides insight into the sign and magnitude of the R&D bias in reported  $ROE$ . The bias is negative (i.e.,  $CROE_t > ROE_t$  or reported  $ROE$  understates economic profitability) if  $g > ROE_t$ , and is increasing in absolute value with increasing R&D intensity ( $\frac{RD_{t-1}}{BV_{t-1}}$ ) and R&D economic durability ( $\lambda$ ). If  $ROE_t$  equals  $g$ , then the R&D bias in  $ROE$  disappears. Consistent with Zhang (2000), in the case of zero growth, reported  $ROE$  is always overstated (i.e.,  $ROE$  exceeds capitalized  $CROE$ ), and the difference between  $ROE$  and  $CROE$  is positively related with R&D intensity and R&D durability. For example, if R&D investments have no future benefits ( $\lambda = 0$ ), then  $ROE = CROE$  when R&D investments do not grow. Lev et al. (1999) provide a different approach in their proposition 1 to describe the sign and magnitude of the R&D reporting bias in  $ROE$ . In their model, the bias is positive if  $ROE \leq \frac{g}{1+g/2}$ . The latter  $ROE$  condition leads to a similar prediction on the sign of the reporting bias as in my eq.(B.6).

## Appendix C Sample Firms.

Obs	Cusip <sup>a</sup>	Permno <sup>b</sup>	Ticker <sup>a</sup>	Company Name <sup>a</sup>	Incorp. Date <sup>c</sup>	Type <sup>d</sup>
1	002824	20482	ABT	ABBOTT LABORATORIES	1900	P
2	020813	65832	ALO	ALPHARMA	1983	G
3	022615	64856	AZA	ALZA	1968	P
4	026609	15667	AHP	AMERICAN HOME PRODUCTS	1915	P
5	110122	19393	BMY	BRISTOL MYERS SQUIBB	1933	G
6	17253C	64178	RXC	CIRCA PHARMACEUTICALS	1978	G
7	266354	10685	DRMD	DURAMED PHARMACEUTICALS	1982	G
8	284131	31799	ELN	ELAN	1969	P
9	312024	54819	FAUL	FAULDING INC	1845	G
10	345838	45241	FRX	FOREST LABORATORIES	1956	G
11	368710	38280	GNE	GENENTECH	1976	P
12	37733W	75064	GLX	GLAXO WELLCOME	1873	P
13	406369	10116	HDG	HALSEY DRUG COMPANY	1935	G
14	448924	68340	ICN	ICN PHARMACEUTICALS	1960	G
15	478160	22111	JNJ	JOHNSON & JOHNSON	1887	P
16	482740	46950	KV.A	K V PHARMACEUTICAL	1941	G
17	532457	50876	LLY	LILLY (ELI) & CO	1876	P
18	589331	22752	MRK	MERCK & CO	1927	P
19	552880	54798	MOGN	MGI PHARMA	1979	P
20	628530	69550	MYL	MYLAN LABORATORIES	1970	G
21	631728	80004	NSTK	NASTECH PHARMACEUTICAL	1983	P
22	670100	63263	NVO	NOVO-NORDISK	1960	P
23	717081	21936	PFE	PFIZER	1900	P
24	716932	75429	3PHFR	PHARMACEUTICAL FORMULATIONS	1981	G
25	717125	61138	PRX	PHARMACEUTICAL RESOURCES	1978	G
26	716941	82647	PNU	PHARMACIA & UPJOHN	1886	P
27	76242T	39570	RPR	RHONE-POULENC RORER	1961	P
28	806528	77075	SHR	SCHERER (R P)	1972	G
29	806605	25013	SGP	SCHERING-PLOUGH	1935	P
30	871616	37102	SYN	SYNTEX CORP	1970	P
31	881624	75652	TEVIY	TEVA PHARMACEUTICAL INDUSTRIES	1985	G
32	904801	78116	UMED	UNIMED PHARMACEUTICALS	1962	P
33	910571	78749	UG	UNITED GUARDIAN	1942	P
34	934488	24678	WLA	WARNER-LAMBERT	1901	P
35	989365	69155	ZENL	ZENITH LABORATORIES	1968	G

<sup>a</sup> The cusip number, ticker, and company name are taken from the Compustat 1998 datafiles.

<sup>b</sup> The permanent number in the CRSP database of 1999.

<sup>c</sup> The incorporation date as mentioned in the history section of Moody's Industrial Manual or Bioscan February 1999 (American Health Consultants).

<sup>d</sup> This column classifies the sample firms into generic ("G") and pioneering ("P") drug firms. The classification is based on the level of *FIRM* (see definition 5): if the number of years  $DFIRM = 1$  greater than the number of years  $DFIRM = 0$  over the sample period 1975-1998, then P appears in the column of this table, and vice versa.

## Appendix D

### Definition of Therapeutic Classes.

Obs	FDA Class <sup>a</sup>	Therapeutic Class Name <sup>a</sup>	First <sup>b</sup>	Pop. <sup>c</sup>	Samp. <sup>c</sup>
1	0117	ANESTHETICS, LOCAL	1950	123	35
2	0118	ANESTHETICS, GENERAL	1951	27	11
3	0119	ANESTHESIA, ADJUNCTS TO/ANALEPTICS	1952	60	25
4	0281	ANTIDOTES, SPECIFIC	1953	54	14
5	0283	ANTIDOTES, GENERAL	1973	7	2
6	0346	PENICILLINS	1950	288	112
7	0347	CEPHALOSPORINS	1965	208	98
8	0348	LINCOSAMIDES/MACROLIDES	1964	178	85
9	0349	POLYMYXINS	1957	12	8
10	0350	TETRACYCLINES	1950	150	60
11	0351	CHLORAMPHENICOL/DERIVATIVES	1950	19	2
12	0352	AMINOGLYCOSIDES	1959	96	31
13	0353	SULFONAMIDES/RELATED COMPOUNDS	1950	113	35
14	0354	ANTISEPTICS,URINARY TRACT	1953	20	5
15	0355	ANTIBACTERIALS, MISCELLANEOUS	1950	173	82
16	0356	ANTIMYCOBACTERIALS (INCL ANTI LEPROSY)	1950	54	7
17	0357	QUINOLONES/DERIVATIVES	1987	11	4
18	0358	ANTIFUNGALS	1950	150	69
19	0388	ANTIVIRALS	1963	84	39
20	0408	DEFICIENCY ANEMIAS	1951	36	11
21	0409	ANTICOAGULANTS/THROMBOLYTICS	1950	78	27
22	0410	BLOOD COMPONENTS/SUBSTITUTES	1952	20	4
23	0411	HEMOSTATICS	1952	14	8
24	0501	CARDIAC GLYCOSIDES	1954	9	7
25	0502	ANTIARRHYTHMICS	1950	167	44
26	0503	ANTIANGINALS	1967	113	37
27	0504	VASCULAR DISORDERS, CEREBRAL/PERIPHERAL	1953	23	8
28	0505	HYPOTENSION/SHOCK	1950	57	27
29	0506	ANTIHYPERTENSIVES	1952	470	181
30	0507	DIURETICS	1953	268	70
31	0508	CORONARY VASODILATORS	1957	18	8
32	0509	RELAXANTS/STIMULANTS, URINARY TRACT	1974	19	4
33	0510	CALCIUM CHANNEL BLOCKERS	1985	61	19
34	0511	CARBONIC ANHYDRASE INHIBITORS	1957	3	1
35	0512	BETA BLOCKERS	1981	40	15
36	0513	ALPHA AGONISTS/ALPHA BLOCKERS	1974	11	4
37	0514	ACE INHIBITORS	1985	31	7
38	0600	CENTRAL NERVOUS SYSTEM	1979	2	1
39	0626	SEDATIVES/HYPNOTICS	1954	92	39
40	0627	ANTIAXIETY	1956	171	60
41	0628	ANTIPSYCHOTICS/ANTIMANICS	1954	162	45
42	0630	ANTIDEPRESSANTS	1959	162	63
43	0631	ANOREXIANTS/CNS STIMULANTS	1956	115	25
44	0632	CNS, MISCELLANEOUS	1953	36	17
45	0635	ANTIEMETICS	1954	19	8
46	0700	CONTRAST MEDIA/ RADIOPHARMACEUTICALS	1953	3	1
47	0789	DIAGNOSTICS, RADIOPAQUE AND NONRADIOACTIVE	1951	82	12
48	0790	DIAGNOSTICS - RADIOPHARMACEUTICALS	1957	78	2
49	0792	DIAGNOSTICS, MISCELLANEOUS	1954	11	2
50	0874	DISORDERS, ACID/PEPTIC	1962	71	28
51	0875	ANTIDIARRHEALS	1954	63	19
52	0876	LAXATIVES	1956	14	1
53	0877	GASTROINTESTINAL, MISCELLANEOUS	1953	99	27
54	0878	ANTISPASMODICS/ANTICHOLINERGICS	1950	22	9
55	0912	HYPERLIPIDEMIA	1964	34	17
56	0913	VITAMINS/MINERALS	1950	117	34
57	0914	NUTRITION, ENTERAL/PARENTERAL	1971	58	37
58	0915	REPL/REGS OF ELECTROLYTES/WATER BALANCE	1951	281	71
59	0916	CALCIUM METABOLISM	1953	20	7

## Continued

Obs	FDA Class <sup>a</sup>	Therapeutic Class Name <sup>a</sup>	First <sup>b</sup>	Pop. <sup>c</sup>	Samp. <sup>c</sup>
60	1032	ADRENAL CORTICOSTEROIDS	1951	308	85
61	1033	ANDROGENS/ANABOLIC STEROIDS	1958	35	11
62	1034	ESTROGENS/PROGESTINS	1951	131	53
63	1035	ANTERIOR PITUITARY/HYPOTHALMIC FUNCTION	1955	36	9
64	1036	BLOOD GLUCOSE REGULATORS	1957	121	73
65	1037	THYROID/ANTITHYROID	1956	13	5
66	1038	ANTIDIURETICS	1970	6	3
67	1039	RELAXANTS/STIMULANTS,UTERINE	1966	18	9
68	1040	CONTRACEPTIVES	1967	58	16
69	1041	INFERTILITY	1954	12	2
70	1042	DRUGS USED IN DISORDERS OF GROWTH HORMON	1960	8	2
71	1180	VACCINES/ANTISERA	1987	2	1
72	1181	IMMUNOMODULATORS	1968	22	11
73	1264	ANTISEPTICS/DISINFECTANTS	1950	28	8
74	1265	DERMATOLOGICS	1950	361	128
75	1268	TOPICAL STEROIDS	1961	75	27
76	1270	ACNE PRODUCTS	1973	15	5
77	1271	TOPICAL ANTI-INFECTIVES	1964	20	3
78	1272	ANORECTAL PRODUCTS	1957	3	2
79	1273	PERSONAL CARE PRODUCTS (VAGINAL)	1990	1	1
80	1274	DERMATITIS/ANTIPURETICS	1981	12	4
81	1371	EXTRAPYRAMIDAL MOVEMENT DISORDERS	1950	84	24
82	1372	MYASTHENIA GRAVIS	1951	11	4
83	1373	SKELETAL MUSCLE HYPERACTIVITY	1951	180	41
84	1374	ANTICONVULSANTS	1951	107	49
85	1400	ONCOLYTICS	1992	1	1
86	1479	ANTINEOPLASTICS	1950	219	106
87	1481	ANTIMETABOLITES	1968	3	1
88	1566	GLAUCOMA	1953	63	15
89	1567	CYCLOPLEGICS/MYDRIATICS	1961	28	1
90	1568	OCULAR ANTI-INFECTIVE/ANTI-INFLAMMATORY	1950	203	33
91	1569	OPHTHALMICS, MISCELLANEOUS	1953	35	1
92	1670	OTICS, TOPICAL	1953	28	7
93	1671	VERTIGO/MOTION SICKNESS/VOMITING	1951	261	77
94	1700	RELIEF OF PAIN	1996	2	1
95	1720	ANALGESICS, GENERAL	1954	208	73
96	1721	ANALGESICS-NARCOTIC	1950	204	62
97	1722	ANALGESICS-NON-NARCOTIC	1979	24	6
98	1723	ANTIMIGRAINE/OTHER HEADACHES	1953	25	6
99	1724	ANTIARTHRITICS	1952	225	83
100	1725	ANTIGOUT	1951	51	21
101	1726	CENTRAL PAIN SYNDROMES	1978	11	4
102	1727	NSAID	1959	109	35
103	1728	ANTIPTYRETICS	1989	4	3
104	1860	ANTIPROTOZOALS	1953	48	15
105	1862	ANTHELMINTICS	1953	21	11
106	1863	SCABICIDES/PEDICULICIDES	1957	16	3
107	1864	ANTIMALARIALS	1952	10	1
108	1940	ANTIASTHMATICS/BRONCODILATORS	1951	257	89
109	1941	NASAL DECONGESTANTS	1953	24	12
110	1943	ANTITUSSIVES/EXPECTORANTS/MUCOLYTICS	1951	90	36
111	1944	ANTIHISTAMINES	1950	148	50
112	1945	COLD REMEDIES	1951	104	40
113	1947	CORTICOSTEROIDS-INHALATION/NASAL	1982	8	6
114	2087	UNCLASSIFIED	1950	78	29
115	2095	PHARMACEUTICAL AIDS	1960	45	21
116	2096	SURGICAL AIDS	1950	27	5
117	2097	DENTAL PREPARATIONS	1954	30	7
		total		9254	3168

<sup>a</sup> FDA therapeutic class code and class name from National Drug Code Directory, available at: <http://www.fda.gov/cder/drug.htm>.

<sup>b</sup> First drug in this class since start of the database (Jan 1950).

<sup>c</sup> Pop. represents the total number of drugs in a particular class approved by the FDA between 1950 and 1998.

Samp. is the drug count for the 35 sample firms.

Table 1:  
Descriptive statistics for the pooled sample<sup>a</sup>

<i>Variable</i> <sup>b</sup>	<i>nobs</i>	<i>Mean</i>	<i>P5</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>P95</i>
<i>MTB</i>	620	4.161	1.206	2.310	3.600	5.635	8.956
<i>ROE</i>	615	0.159	-0.378	0.102	0.198	0.289	0.499
<i>RDBV</i>	608	0.149	0.039	0.078	0.117	0.184	0.356
<i>FIRM</i>	617	0.574	0.008	0.118	0.667	0.925	1.000
<i>MS</i>	583	0.061	0.018	0.035	0.052	0.080	0.136
<i>PAT</i>	536	0.568	0.000	0.078	0.250	0.618	2.458
<i>RDG</i>	600	0.295	-0.190	0.083	0.155	0.288	0.963
<i>LTRDG</i>	578	0.203	-0.058	0.116	0.163	0.269	0.580
<i>BVG</i>	616	0.285	-0.244	0.040	0.135	0.260	0.976

<sup>a</sup> The table provides summary statistics on the distribution of the key variables in the paper. The summary statistics are: *nobs* (number of sample observations), *Mean* (sample average), *P5* (5th percentile), *Q1* (1st quartile), *Median* (sample median), *Q3* (3rd quartile), and *P95* (95th percentile). The sample period covers 1975-1998.

<sup>b</sup> The variables are: *MTB* (market-to-book ratio), *ROE* (book return on equity), *RDBV* (the ratio of R&D to book value of equity), *FIRM* (proportion of NDAs in drug portfolio as defined in (5)), *MS* (market share as defined in (6)), *PAT* (number of patents to 3 year average R&D as defined in (7)), *RDG* (yearly growth rate of R&D investments as defined in (8)), *LTRDG* (annualized 3-year growth rate of R&D investments as defined in (9)), and *BVG* (yearly growth rate of book value of equity, defined similar to *RDG*).

Table 2:

Descriptive statistics by *high* and *low* value driver group <sup>a</sup>

Variable <sup>b</sup>	<b>High and low value driver groups</b>											
	<i>DFIRM</i>			<i>MS</i>			<i>PAT</i>			<i>RDG</i>		
	generic	pioneer	diff	low	high	diff	low	high	diff	low	high	diff
<i>MTB</i>	3.287 [3.341]	3.835 [3.948]	0.549 (3.50)	3.106 [3.875]	3.778 [3.122]	0.672 (3.72)	3.677 [3.512]	3.714 [3.688]	0.036 (0.41)	3.640 [3.484]	3.635 [3.543]	-0.005 (-0.11)
<i>ROE</i>	0.164 [0.217]	0.213 [0.177]	0.049 (3.67)	0.163 [0.217]	0.238 [0.174]	0.074 (7.41)	0.210 [0.188]	0.202 [0.192]	-0.008 (-0.28)	0.209 [0.190]	0.193 [0.212]	-0.016 (0.74)
<i>RDBV</i>	0.089 [0.087]	0.146 [0.116]	0.057 (8.09)	0.092 [0.091]	0.158 [0.107]	0.065 (8.19)	0.129 [0.133]	0.126 [0.110]	-0.003 (-0.36)	0.132 [0.128]	0.117 [0.105]	-0.015 (-1.33)
<i>FIRM</i>	<b>0.079</b> [0.173]	<b>0.848</b> [0.304]	0.769 (21.39)	0.200 [0.667]	0.748 [0.341]	0.548 (10.21)	0.752 [0.483]	0.694 [0.474]	-0.058 (-0.89)	0.697 [0.492]	0.609 [0.838]	-0.088 (-1.13)
<i>MS</i>	0.035 [0.018]	0.069 [0.054]	0.033 (14.18)	<b>0.034</b> [0.017]	<b>0.076</b> [0.052]	0.042 (18.91)	0.052 [0.049]	0.056 [0.044]	0.004 (-1.29)	0.059 [0.043]	0.045 [0.042]	-0.014 (-3.47)
<i>PAT</i>	0.209 [0.545]	0.253 [0.536]	0.044 (3.28)	0.155 [0.575]	0.265 [0.415]	0.110 (5.40)	<b>0.070</b> [0.193]	<b>0.529</b> [0.897]	0.459 (16.61)	0.246 [0.488]	0.241 [0.611]	-0.005 (0.17)
<i>RDG</i>	0.202 [0.490]	0.144 [0.130]	-0.058 (-2.89)	0.172 [0.325]	0.142 [0.121]	-0.030 (-2.08)	0.146 [0.194]	0.155 [0.163]	0.008 (1.09)	<b>0.081</b> [0.137]	<b>0.283</b> [0.394]	0.202 (20.33)

<sup>a</sup> The table provides descriptive statistics on key variables computed for each high and low value driver group. That is, I group the total number of sample firms into two subsamples based on *DFIRM*, *MS*, *PAT* and *RDG* respectively. The median per group together with the difference between the group medians are reported on the first line. The boldface numbers indicate the median values on the grouping variable. On the second line, the interquartile range (Q3-Q1) per group is reported between square brackets, and the Wilcoxon Z-score to test the significance of the difference in medians between the high and low group is shown between round brackets.

<sup>b</sup> See table 1 for the variable definitions.

Table 3:  
Correlation matrix for the pooled sample<sup>a b</sup>

	<i>MTB</i>	<i>ROE</i>	<i>RDBV</i>	<i>FIRM</i>	<i>MS</i>	<i>PAT</i>	<i>RDG</i>	<i>LTRDG</i>
<i>MTB</i>	<b>1</b>	0.175	0.331***	0.177	0.092	0.147	0.016	0.039
<i>ROE</i>	0.420**	<b>1</b>	-0.130	-0.012	0.177	-0.039	0.217	0.065
<i>RDBV</i>	0.301***	0.045	<b>1</b>	0.248**	0.165*	-0.060	-0.008	0.068
<i>FIRM</i>	0.208*	-0.010	0.286***	<b>1</b>	0.500***	0.143	-0.091	-0.134
<i>MS</i>	0.159	0.304*	0.284**	0.461***	<b>1</b>	0.009	-0.070	-0.026
<i>PAT</i>	0.096	0.026	-0.078	0.019	0.145	<b>1</b>	-0.023	0.034
<i>RDG</i>	0.033	0.080	-0.017	-0.061	-0.084	-0.016	<b>1</b>	0.243
<i>LTRDG</i>	-0.011	-0.016	0.036	-0.152	-0.052	-0.036	0.299*	<b>1</b>

<sup>a</sup> The table provides Pearson (upper triangle) and Spearman (lower triangle) cross-sectional correlation coefficients among the key variables in the paper. The correlations are averages of yearly cross-sectional correlations over the sample period 1975-1998.

<sup>b</sup> The variables in the table are: *MTB* (market-to-book ratio), *ROE* (book return on equity), *RDBV* (ratio of R&D to book value of equity), *FIRM* (proportion of NDAs in drug portfolio as defined in (5)), *MS* (market share as defined in (6)), *PAT* (number of patents to 3 year average R&D as defined in (7)), *RDG* (yearly growth rate in R&D as defined in (8)), and finally, *LTRDG* (annualized 3-year growth rate of R&D as defined in (9)).

<sup>c</sup> The asterisks represent the significance of a correlation coefficient. The significance level is calculated by counting the number of yearly correlations that have opposite signs than the majority of correlations. Three asterisks indicate that there are less than 5% yearly correlations with opposite signs. Two (one) asterisk(s) indicate there are between 5% (10%) and 10% (15%) yearly correlations with opposite signs.

Table 4:  
Estimation results for the *ROE*-based valuation model <sup>a</sup>

<i>Effect</i> <sup>b</sup>	Model 1	Model 2	Model 3		Model 4		Model 5	
			<i>NI</i> > 0	<i>NI</i> < 0	<i>NI</i> > 0	<i>NI</i> < 0	<i>NI</i> > 0	<i>NI</i> < 0
<b>Intercept</b>	4.67 (23.94)	1.62 (9.70)	2.77 (11.54)	3.66 (7.06)	0.18 (0.83)	2.69 (6.03)	2.52 (4.41)	0.97 (1.16)
<i>DUMMY</i> <sub>1</sub> ( <i>FIRM</i> ) –							–1.72 (–2.70)	4.70 (3.59)
<i>DUMMY</i> <sub>2</sub> ( <i>MS</i> ) –							–1.63 (–2.66)	–4.71 (–2.40)
<i>DUMMY</i> <sub>3</sub> ( <i>PAT</i> ) 0							–1.13 (–2.26)	1.35 (1.21)
<i>DUMMY</i> <sub>4</sub> ( <i>RDG</i> ) –							–1.06 (–2.10)	0.17 (0.14)
<b>ROE</b>	1.01 (1.88)	1.93 (7.24)	7.66 (11.53)	–5.62 (–5.85)	7.34 (11.02)	–0.04 (–0.11)	0.59 (0.37)	–1.07 (–0.77)
<i>DUMMY</i> <sub>1</sub> ( <i>FIRM</i> ) +							9.13 (3.88)	1.90 (0.59)
<i>DUMMY</i> <sub>2</sub> ( <i>MS</i> ) +							3.70 (1.53)	–4.88 (–1.35)
<i>DUMMY</i> <sub>3</sub> ( <i>PAT</i> ) 0							0.78 (0.47)	0.58 (0.21)
<i>DUMMY</i> <sub>4</sub> ( <i>RDG</i> ) +							3.89 (2.50)	–2.08 (–0.59)
<b>RDBV</b>		16.94 (22.94)			17.51 (23.14)	7.53 (5.00)	17.73 (4.89)	3.26 (8.89)
<i>DUMMY</i> <sub>1</sub> ( <i>FIRM</i> ) –							–4.78 (–1.29)	–12.46 (–3.00)
<i>DUMMY</i> <sub>2</sub> ( <i>MS</i> ) –							–2.69 (–0.76)	11.40 (1.47)
<i>DUMMY</i> <sub>3</sub> ( <i>PAT</i> ) +							10.05 (4.29)	–2.07 (–0.52)
<i>DUMMY</i> <sub>4</sub> ( <i>RDG</i> ) ?							0.79 (0.33)	–4.29 (–0.79)
adjusted R <sup>2</sup>	0.1%	49%	21%		56%		66%	
# observations	603	591	603		591		527	

<sup>a</sup> The *ROE* regression equations are estimated with OLS on the pooled sample, deleting the 1% most extreme observations at each side of the DFFIT distribution. The sample period covers 1975-1998. The dependent variable in all regressions is the market-to-book ratio (*MTB*). The independent variables are indicated in the first column “effect”. Models 1 to 4 are benchmark models (see page 21), while model 5 is the key valuation model in this study, i.e. eq.(11) on page 19.

<sup>b</sup> *DUMMY*<sub>*i*</sub> indicates a 0-1 variable and is based on one of the four value driver proxies: *DFIRM*, *MS*, *PAT* and *RDG*. *DUMMY*<sub>1</sub> (based on *FIRM*) is equal to 1 if a firm is a pioneering drug manufacturer, and 0 otherwise; *DUMMY*<sub>2</sub> (based on *MS*) is equal to 1 if a firm has a market share in the upper half of the sample (above the sample median), and 0 otherwise; *DUMMY*<sub>3</sub> (based on *PAT*) is equal to 1 if a firm has a number of patents per dollar of past R&D that exceeds the median sample number, and 0 otherwise; and *DUMMY*<sub>4</sub> (based on *RDG*) is equal to 1 if a firm has a growth in R&D that exceeds median sample growth, and 0 otherwise. The effect of the four dummy variables on the valuation coefficient of the intercept, *ROE* and *RDBV* is indicated by the + or – sign.