[Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012)](https://scholars.fhsu.edu/jbl)

Volume 2
Number 1 *Journal of Business & Leadership* Article 18

1-1-2006

In-Process R&D (IPRD) Write-off Misclassifications In Compustat: An Econometric Evaluation

Lolita Paff
Penn State Berks

Recommended Citation

Paff, Lolita (2006) "In-Process R&D (IPRD) Write-off Misclassifications In Compustat: An Econometric Evaluation," *Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012)*: Vol. 2: No. 1, Article 18.
DOI: 10.58809/AKCP5412
Available at: https://scholars.fhsu.edu/jbl/vol2/iss1/18

Journal of Business and Leadership: Research, Practice, and Teaching
2006, Vol. 2, No. 1, 158-169

IN-PROCESS R&D (IPRD) WRITE-OFF MISCLASSIFICATIONS IN COMPUSTAT: AN ECONOMETRIC EVALUATION

Lolita Paff, Penn State Berks

Over the period 1994-1999, for a sample of 57 firms, in-process research and development (IPRD) costs write-offs were reported in the 10-Ks of approximately 12% of the observations. The IPRD amounts ranged from $230,000 to over $167 million. In 38% of these cases, Compustat overstated R&D expense by including the IPRD write-off. Comparative econometric estimates obtained show larger parameter coefficients when the Compustat's R&D expense data was used. This suggests prior research on R&D tax credit effectiveness based on Compustat data may have been upwardly biased, overstating the tax credit's incentive effects. Policy implications and directions for future research are discussed.

INTRODUCTION

Standard and Poor's Compustat database is frequently used in accounting, finance and economics empirical research. Not surprisingly, a fairly extensive body of literature investigates this data source's integrity across variables and years. For example, San Miguel (1977) compares Compustat's reported 1972 R&D expense amounts to firms' 10-Ks, finding a discrepancy in 30% of the sample of firms tested. Rosenberg and Houglet (1974) compare stock prices in Compustat to the Center for Research in Security Prices (CRSP) quotations (during the 1960s) finding no material errors. Vasarhelyi and Yang (1983) examine total assets, current liabilities, net sales, net income before extraordinary items, inventories, gross plant, and depreciation for 200 firms in 1981. They find error rates of 0%, .5%, 1%, 3%, 4.5%, 6%, and 34.5%, respectively. Bean and Guerard (1989) compare research expenditures reported in Compustat to NSF and Census Bureau reports, finding the Compustat amounts tended to be higher due to differences in R&D definitions, particularly in 1978 and 1979. Kinney and Swanson (1993) examine the accuracy of Compustat's tax-related data fields over the period 1986-1988, finding a total error frequency of 11.65%. Similarly, Manzon (1994) finds Compustat incorrectly reported Net-Operating-Losses (NOLs) as zero or missing in 3.8% of the sample during the years 1982-1991, when a review of the firms' 10-Ks suggested otherwise. Kern and Morris (1994) consider the effect database choice (Compustat v. Value Line) has on empirical research. Based on data from 1977-1991, they find the choice of database "can affect the results of and inferences drawn from empirical research in ways more than anticipated by researchers" (p.284).

Ideally, tax-related empirical research should be performed with tax return data, since tax filings provide the most reliable information on items such as corporate tax liabilities, NOLs, expenses, and tax credits. However, since tax returns are confidential, most empirical research on R&D tax incentives on the firm-level has been based on a variety of public data sources. (Altshuler, 1989 is a rare exception.) In the case of R&D tax credit studies, data sources used include McGraw-Hill Surveys (Collins, 1983); National Science Foundation (NSF) reports (Bailey and Lawrence, 1992), 10-K filings with the Securities and Exchange Commission (SEC) (Paff, 2004 and 2005; McCutchen, 1993), and Compustat (Swenson, 1992; Berger, 1993; Eisner et al, 1993; Hall, 1993; Hines, 1993; Billings and Fried, 1999; Billings et al., 2001).

The literature evaluating the appropriateness and potential biases associated with the use of Compustat or other publicly available data in R&D tax credit related research has focused on the differences in definitions with respect to what constitutes "research" or "development" (Hall and Long, 1999), the effects of accounting reclassifications from non-R&D items to R&D expense (Hall and Wosinska, 1999), and the appropriateness of using R&D expense as a proxy for qualified research expenditures for R&D tax credit purposes (Hall and van Reenen, 2000). Along a related theme, Bean and Guerard (1989) compare R&D amounts reported in Compustat against NSF data. Although their results obtained with Compustat data were not materially different from those obtained from NSF data, they did observe significant differences for certain industries across certain years. "In general it appears that firms over-report R&D expenditures in Compustat relative to NSF/Census" (p.205). Similarly, Hall and Long (1999) evaluate the consistency and accuracy of R&D data from

two sources: firms' 10-K filings with the Securities and Exchange Commission (SEC) and the National Science Foundation (NSF). They write:

> Substantial effort appears to have gone into getting the R&D numbers "right" by the professional accounting world, by which we mean following the definitions and reporting requirements carefully and systematically, both within a single company and across companies... We have concluded, therefore, that the best practical course is to use the 10-K data as a benchmark... (p.27)

Interestingly, much less attention has been given to the potential for bias and erroneous findings that may result if the public data source suffers from discrepancies or errors. Of particular note, Kern and Morris's (1992 and 1994) effective tax rate studies focused on examining and reconciling the differences in empirical results obtained by testing the same theoretical model with two different databases, Compustat and Value-Line. This paper extends that vein of research in two ways. First, since we are unaware of any studies subsequent to San Miguel's comparison of Compustat v. 10-K reporting of 1972 R&D amounts, and given the fairly large number of R&D tax credit studies utilizing Compustat data, a more recent comparison of Compustat's R&D expense to the amounts reported in firms' 10-Ks during the period 1994-1999 is provided. Second, this paper provides comparative empirical results obtained by testing a model of R&D investment sensitivity to tax price changes with two alternate data sources, 10-K filings and Compustat, for the same sample of firms. In contrast to Morris and Kern, this research focuses on the differences between Compustat data and the original source data, each firm's 10-K filings.

Of particular relevance is the possibility of encoding and classification of errors in Compustat by including IPRD amounts in R&D expense. IPRD represents a portion of the acquired firm's past expenditure for research projects. Because these amounts do not represent current period expenditure for research activity they are not part of qualified research expenditure for R&D tax credit calculation purposes. Therefore, these amounts should be excluded from research and development expense when R&D expense is used as a proxy for qualified research expenditure in evaluating the effectiveness of R&D tax credit policy or testing models of research investment.

The accounting for and reporting of R&D is governed by two Financial Accounting Standards Board pronouncements, Statements of Financial Accounting Standards (SFAS) No. 2, Accounting for Research and Development Costs (1975) and No. 68, Research and Development Arrangements (1985). Despite the relative constancy of accounting policy, increases in merger and acquisition activity has complicated the accounting for a closely related item, acquisitions of in-process research and development (IPRD). This is because SFAS No. 2 "requires that R&D generally be expensed as incurred and that each year's total R&D be disclosed in the financial statements... SFAS 2 could lead to exaggerated results in reporting the acquisition of a company with in-process R&D" (Oliver 2003, 46). To illustrate the impact IPRD can have on a firm's financial statements, Stallworth and DiGregorio (2005) note that Network Associates paid $131 million for CyberMedia in 1998. They expensed 93% of the purchase price as IPRD. Upon completion of an SEC review, Network Associates reversed $214 of the IPRD write-off.

Although publicized examples of large write-offs did not occur until the mid-tolate-1990s, corporations have been allowed to write off IPRD since the late 1970s. One of the earliest and highly publicized cases occurred in 1990 when Lotus Development Corporation wrote off $53 million or over 81% of the purchase cost of Samna Corporation as in-process R&D. In 1993, Cisco Systems determined that 80 percent of its $120.5 million acquisition of Lightstream Corp. could be attributable to purchased R&D (McGoldrick, 1997). "By the end of 1996, the FASB was reportedly reviewing its rules governing R&D write-offs. The SEC was said to be investigating acquirers' valuations of purchased research and development" (Browning, 1997; 30). As San Miguel (1977) notes:

> Several factors may have contributed to the poor quality of R&D data in the 1972 Compustat tapes... One was probably the newly implemented 10-K R&D disclosure rules. The new rules required that different bits of R&D information be disclosed in a variety of locations... Unless these facts were carefully ferreted from the 10-K's numerous statements, footnotes, and schedules, errors in classifying the data were apt to result" (p.639).

Given the complexity of IPRD accounting and the reporting of R&D-related information in multiple financial statement locations, suggests Compustat may have some errors with respect to amounts reported as R&D expense. Indeed, our results show that in the cases

https://scholars.fhsu.edu/jbl/vol2/iss1/18
DOI: 10.58809/AKCP5412

of reported IPRD, Compustat erroneously included IPRD in R&D and thereby overstated R&D expense, in 38% of the observations. A comparison of the econometric results suggests Compustat data provides larger estimates of firms' R&D tax price sensitivity than those obtained with data derived directly from firms' 10-Ks. These findings suggest the tax price elasticity estimates obtained in prior R&D tax credit studies using Compustat data (including Swenson, 1992; Berger, 1993; Eisner et al, 1993; Hall, 1993; Hines, 1993; Billings and Fried, 1999; Billings et al., 2001 as noted previously) may have been upwardly biased.

The paper is organized as follows. Section two compares the 10-K and Compustat data, providing summary statistics, analysis of differences between R&D expense reported across the data sources, and potential explanations for the observed variations in reported amounts. Section three provides a description of the empirical model, tax price computation methodology, and the econometric results from the 10-K and Compustat data sets. The paper concludes with discussion of the implications and limitations of these findings, and suggestions for improved use of public source data in empirical R&D policy research.

10-K Data vs. COMPUSTAT

Data Gathering and Reporting from Compustat and 10-Ks: Each year, publicly owned corporations are required to file an annual report (10-K) of financial and operating information with the SEC. The SEC requires firms to submit their filings to the EDGAR system. EDGAR, the Electronic Gathering Analysis and Retrieval system, "performs automated collection, validation, indexing, acceptance, and forwarding of submissions by companies and others who are required by law to file forms with the U.S. Securities and Exchange Commission (SEC). Its primary purpose is to increase the efficiency and fairness of the securities market for the benefit of investors, corporations, and the economy by accelerating the receipt, acceptance, dissemination, and analysis of time sensitive corporate information filed with the agency" (www.sec.gov/edgar/aboutedgar.htm).

During the years 1994-1996, the SEC phased in the EDGAR filing requirement. For nearly all the firms in this study, filings prior to 1996 are not available in the EDGAR system. However, since firms are required to report prior periods' information for comparative purposes, the filings that are available usually include financial data extending back to 1992, assuming the firm was in operation.

The filings within EDGAR are in plain text format and HTML. This makes for quick online access, but does not allow for the downloading and saving of files in a format that is convenient for analysis. However, given the need for printable versions of this data by researchers, financial analysts, accountants, and investors, it is now possible to acquire the SEC filings in Rich Text Format (RTF). EDGAR-online provides RTF versions of EDGAR filings. The 10-Ks for all the firms in the sample were downloaded from this proprietary service. The filings downloaded from EDGAR-online are identical to the filings with the SEC in all respects except that they have been formatted to RTF. To verify the accuracy of the filings as provided by EDGAR-online, one income statement for every tenth firm in the sample was tested. The year of the income statement was rotated in reverse chronological order to avoid bias. There were no discrepancies between the information in EDGAR and the RTF version downloaded from EDGAR-online.

To verify that EDGAR-online had not omitted publicly available information from its database, the EDGAR database was searched for a sample of firms' filings and compared to the number of filings available for that firm through Edgar-online. Again, there were no cases where EDGAR-online supplied fewer filings than what was publicly available from the SEC. Based on these findings, we assume the data from the EDGAR-online files is equivalent to the data within the SEC's EDGAR database.

Some firms in the sample were not publicly owned or made their initial public offering (IPO) of stock during the study period (1994-1999). In these cases, 10-K filings may not be available. However, another valuable source of financial statement data is the S-1 registration, Section 11 of an S-1 registration, "General form for registration of securities under the Securities Act of 1933." In particular, the "Information with Respect to the Registrant" portion of the S-1 includes the same financial and operational data as a firm's 10-K. Therefore, in cases where 10-Ks were not available, the S-1 registration was used as an alternate source. The inclusion of S-1 registration data is an example of a difference between using firms' filings with the SEC and using Compustat for empirical research. Compustat data is limited to the 10-K as the source of its financial data. For example, if a firm went public in 1997, Compustat would not include prior periods' values, even if the data was reported in the firm's S-1 registration. Thus, financial results from periods before the firm's stock were publicly traded are not available in Compustat.

Another important difference between Compustat data and values obtained from SEC filings relates to restatements of financial statement amounts. Firms are required to report multiple periods' values in the financial statements. For example, a firm's 1996 fiscal year-end statements will include income statements for at least 1995, but will often include 1994 as well. If a significant event occurs during 1997, such as the sale of a business segment, or an acquisition of another firm, the 1997 year-end statements will segregate the material event and for comparability purposes, restate the financial results of prior periods. In this example, 1996 and 1995 will be restated as if the material event had occurred during those periods. The data in Compustat is intentionally adjusted for greater comparability across periods. However, for researchers using financial statement data as a proxy for qualified research expenditure in R&D tax credit policy evaluation, restated values are not appropriate. Therefore, for purposes of this study, the data derived directly from firms' 10-Ks were drawn from the earliest statement available first, working forward to the more recent filings. This ordering was necessary in order to obtain amounts closer to the values the firms likely used to compute R&D tax credits.

To summarize, Compustat's and 10-K's reported research and development expense amounts may differ for one or more reasons. Compustat may report fewer observations since it is limited to 10-K filings as a data source. The restatement of expense for comparative financial statement purposes, may affect analysis in time series analysis or when year-specific financial statement amounts are used as a proxy for a tax return item. Last, accounting and reporting complexity for items such as IPRD may result in coding and misclassification errors in Compustat.

Sample Selection and Descriptive Statistics: Compustat was used to generate a preliminary sample of firms based on the following screening criteria:

- Population Particulars: SIC Codes: 2834, 2835, 2836, 7372, 7373
- Fundamentals: R&D Expense > 0
- Company Specifics: State: California

The sample size was preliminary because some of the firms selected by the screening process could not be part of the study for one or more reasons. First, only firms with all research activity confined to the headquarter-state are included in the analysis. Limiting the sample to single-state R&D performers resulted in fairly large differences between the research intensities of the included and excluded pharmaceuticals. Note that testing single-state R&D performers exclusively is not the same as limiting the sample to firms operating entirely in one state. Indeed, over 47% of firms in the sample had sales and/or manufacturing offices located in two or more states. However, this selection criterion may have resulted in relatively young firms being included, and older, more established firms being excluded. Without detailed information regarding the spending by location of multi-site R&D performers, there is no way of estimating R&D tax prices those firms face.

Firms with insufficient or incomplete data were excluded. For example, some firms may not have publicly reported financial information available in prior years because they were privately held. Privately owned firms are not required to make public their financial statements. In other instances, a firm may not have been a stand-alone entity for the entire study period. Another reason for a firm to be excluded from the study related to its year of formation. In some cases, firms in operation in 1994 may have ceased to exist by the end of 1999. In other cases, firms operating in 1999 may not have existed in 1994. Therefore, this study's outcomes pertain to the research behavior of existing and surviving pharmaceutical and software firms.

In addition to the research intensity differences, excluded firms tended to be more profitable. It is unclear how this may have affected the results. If economies of scale or scope exist in the application of R&D results, it may be argued that larger firms with multiple-state research and more periods in which net income was reported stand to gain more from undertaking R&D and from increased research tax credit rates than smaller firms with consistent losses. This would mean the results obtained are slightly understated. Since Compustat provides complete data on publicly owned corporations exclusively, structural differences in R&D investment between public and private entities are not captured in these results. Himmelberg and Petersen (1994) find that a fairly large portion of investment can be explained by differences in internal finance. This suggests publicly owned companies may have an advantage in financing research investment than their privately owned counterparts. Although R&D spending for publicly owned small firms may be higher than for privately held entities, it is not clear whether a firms' ownership status affects incremental changes in R&D investment. Since both groups should be similarly motivated to reduce corporate income tax liability, it seems likely that firms in both categories would respond similarly. However, because publicly owned companies are influenced by

https://scholars.fhsu.edu/jbl/vol2/iss1/18
DOI: 10.58809/AKCP5412

investor ownership and public scrutiny, the investment patterns observed may reflect greater volatility than would be observed for privately owned firms (see Bushee, 1998).

Table 1: Descriptive Statistics

Research & Development Expense - 10-K & S-1 Filings					
	ALL	**Software**	**Biotech**	**94-'96**	**97-'99**
Mean	31,374	34,035	29,296	26,706	36,042
Std.Dev.	101,710	132,845	68,500	88,814	113,221
Minimum	36	36	287	85	36
Maximum	1,009,880	1,009,880	471,100	719,143	1,009,880
N	342	150	192	171	171
Research & Development Expense - Compustat					
	ALL	**Software**	**Biotech**	**94-'96**	**97-'99**
Mean	30,973	36,385	26,745	26,618	35,329
Std.Dev.	105,241	141,634	63,905	98,100	112,049
Minimum	36	36	287	36	83
Maximum	1,009,880	1,009,880	434,092	886,197	1,009,880
n	342	150	192	171	171
IPRD Observations- Derived from 10-K Filings					
	ALL	**Software**	**Biotech**	**94-'96**	**97-'99**
Mean	20,818	17,495	25,248	19,343	22,292
Std.Dev.	34,169	35,001	33,495	36,836	32,123
Minimum	230	255	230	230	992
Maximum	167,054	167,054	124,888	167,054	124,888
n	42	24	18	21	21

Note: Amounts are in $1000s, except sample sizes.

In summary, the firms in this study can be characterized as single-state R&D performers, publicly owned pharmaceuticals and software firms that have been in operation since at least the early 1990s. It is unclear whether any or all of the potential biases are present. Additional research is needed to gain perspective on whether the results obtained are representative of all firms in these industries and across other states.

Table 1 reports the descriptive statistics of the net sample in total and across various sub-samples across both sources of data. The sample of 342 represents 6 annual observations (1994-1999) for 57 firms (25 software and 32 pharmaceuticals/biotech). A review of the Compustat and 10-K research and development expense amounts shows a fair amount of consistency between the data sources, particularly with respect the observed minima and maxima. This is true across industries as well as time periods. With respect to the differences across the pharmaceutical and software industries, the software industry has a higher mean regardless of data source. In addition, the Compustat software R&D is larger and has greater deviation, reflecting in part the larger number of software firms that reported IPRD write-offs.

Interestingly, the mean R&D expense from the 10-K's is higher than the Compustat mean. However, the standard deviation from Compustat is greater. This suggests that Bean and Guerard's (1989) findings from the 1980s still holds: taken as a whole, Compustat is fairly consistent to other data sources, but there are instances where material differences occur. The variation appears to be attributable to the software firms during the years 1994-1996. This finding is consistent with software firms becoming more cautious in recording and writing down IPRD at the end of the decade in response to increased SEC scrutiny (Dowdell and Press, 2002).

Each IPRD observation reflects a year in which the firm wrote off some or all of its in-process research. These observations represent just over 12% of the sample. Consistent with the trends reported in the financial news, IPRD write-offs can involve large sums; the smallest observation was $230,000, the largest over $167 million. Also consistent with a prior expectation is the increase in the mean IPRD observed from the earlier to the later period. The slightly larger number of software firm IPRD observations also agrees with Dowdell & Press (2002) who observed that over half of all reported financial statement restatements occurred in the 737X SIC (Business Services - Computer Programming and Data Processing category) and concurrently, large numbers of firms in the industry made acquisitions that included IPRD.

Theoretical Model and Empirical Results

Theoretical Model of R&D Investment Based on Tax Price Estimates: The specification and underlying theoretical model follow Paff (2004 and 2005) and Hines (1991). The purpose of the model is to test for structural

shifts in the coefficients on the tax prices of R&D. The variables are:

It - Refers to firm i in time period t
δ,μ - Exponents referring to R&D input and non-R&D inputs respectively
ϕ_t - Parameter common to all firms in period t
ψ_i - Fixed effect for firm i
u_{it} - Normally distributed error term
S(R_C, R_H, I) - Firm sales are a function of in-house and contract research as well as non-R&D inputs
R_C - Contract research
R_H - In-house research
I - Non-R&D investments

Assuming, as did Hines, that firms' production follows a Cobb-Douglas form:

$$S_{it}(R_{Cit}, R_{Hit}, I_{it}) = (R_{Cit})^{\delta}(R_{Hit})^{\eta}(I_{it})^{\mu}\exp(\phi_t + \psi_i + u_{it}) \quad 1)$$

Separate specifications for each form of R&D require observations of each category of expenditure. However, only one variable, R&D expense, was observed for each firm. The single annual observation was divided into contract and in-house segments for the purpose of estimating the tax prices of each. Since there is only one observed independent variable, the econometric testing must be based on a single equation:

$$\ln(\underline{R}) = \lambda_t + \nu_i + \beta_1 \ln(PR_C) + \beta_2 \ln(PR_H) + \varepsilon_{it} \quad (2)$$

The left side represents the natural log of the annual total research expenditure observed for each firm. The specification indicates that R&D expenditure is a function of time and firm effects and the tax prices of both forms of R&D.

Tax Price Computations

In order to test firms' research investment responsiveness to changes in R&D tax price changes (PR_C and PR_H above), estimated tax credits and tax prices for each firm in each period were prepared. The credit rates used come directly from state tax statutes. The California credit is patterned after the federal legislation, with the purpose of rewarding only the R&D activity undertaken with the state. Unlike the federal credit, the California credit is permanent. For in-house R&D in excess of a threshold amount, the tax credit rates are:

- For tax years beginning prior to 1/1/97: 8%
- For tax years beginning prior to 1/1/99: 11%
- For tax years beginning prior to 1/1/00: 12%
- For tax years beginning on or after 1/1/00: 15%

For basic research payments, "contract research," above threshold levels, the tax credit rates are:

- For tax years beginning prior to 1/1/97: 12%
- For tax years beginning on or after 1/1/97: 24%

As Hall and Wosinska (1999) discuss, "The California law uses the federal definition of the base levels but with the important feature that although the R&D intensity used to determine the "Base Amount" is the same number used in the federal calculation, the sales figure by which it is multiplied is the California share of total sales. This has the strange, but possibly intended, effect that a firm with sales throughout the United States but which does all of its R&D in California can have a rather low base level of R&D spending relative to its current level, year after year, even though it is not increasing its R&D... The effect of the provision is to give firms a strong inventive to locate their R&D laboratories in California, even if the rest of the firm is nationwide. It is likely that this is one of the goals of the legislation" (p.8).

Several other parameters are also necessary for estimating the tax prices. First, note that because by design all firms in our sample do all their research activities in their home state, we set the in-state R&D fraction ω_i=1.0 as well as the U.S. R&D fraction υ_i=1.0 for all firms. Moreover, we assume no tax credits for inputs other than R&D. Another parameter is the marginal federal corporate tax rate, τ_{fed}. This varies by firm and by year according to the somewhat complex statutory levels and the 10K or S-1 reports by the firm on their level of before-tax profits, as follows:

- 15% if the year's profit before taxes > 0 & < \$50,000
- 25% if profit before taxes ≥ \$50,000 & < \$75,000
- 34% if profit before taxes ≥ \$75,000 & < \$100,000
- 39% if profit before taxes ≥ \$100,000 & < \$335,000
- 34% if profit before taxes ≥ \$335,000 & < \$10 million
- 35% if profit before taxes ≥ \$10 million & < \$15 million
- 38% if profit before taxes ≥ \$15 million & < \$18.333333 million
- 35% otherwise.

Finally, the federal R&D credit rates (ρ_C and ρ_H) were by statute 0.20 for all study periods, except 1995 and 1996. Congress let the credit lapse for the one year

https://scholars.fhsu.edu/jbl/vol2/iss1/18
DOI: 10.58809/AKCP5412

period between July 1995 - June 1996. Effectively, the credit was zero for half of each year. But our data is annual. Although not ideal because firms could have front or back loaded R&D expenditures, we used an average 0.10 federal credit rate for each of those full years.

For a typical example, one software firm in the sample faced the following characteristics relative to contract research in 1999:

- State R&D credit rate for contract research: $\gamma_C = 0.24$
- Federal R&D credit rate for contract research: $\rho_C = 0.20$
- Federal corporate tax bracket: $\tau_{fed} = 0.35$
- State corporate tax: $\tau_{in} = 0.0884$
- In-state sales apportionment fraction: $\alpha = 0.9575$
- In-state R&D fraction: $\omega_C = 1.0$ [True for all firms in sample]
- US R&D fraction: $\upsilon_C = 1.0$
- Average out of state tax: $\tau_{out} = 0.075$

So the effective tax price for contract research (P^t_C) for this firm in 1999 is:

- $P^t_C = [\alpha\ (1 - \tau_{in}) + (1 - \alpha)(1 - \tau_{out}) - \gamma_C\omega_C(1 - \tau_{in}) - \rho_C\upsilon_C](1 - \tau_{fed})$
- $P^t_C = [0.9575(1 - 0.0884) + (1 - 0.9575)(1 - 0.075) - 0.24(1)(1 - 0.0884) - 0.20(1)](1 - 0.35) =$
- $P^t_C = \$0.3207$

In other words, the added cost to the firm of one more dollar of R&D spending for this firm, after tax credits is only 32 cents. This reflects the combined 24 percent contract R&D credit in California, 20 percent Federal credit, the reductive effect of both the firm's 35 percent federal corporate tax bracket and the 8.84 percent California corporate tax rate, and accounts for a bit of out of state sales apportionment.

Note the effect of the change in California's tax laws. In 1996 and prior years, the state-level credit rate for contract research was only 12 percent, while the corporate tax rate was 9.3 percent. If the rates had not changed, a firm with otherwise similar characteristics would have faced an effective price $P^t_C = \$0.3893$, roughly 20 percent higher.

Empirical Results

Table 2 provides the pharmaceutical and software results for the entire period obtained from Compustat and firms' 10-Ks. The Contract ($LNPR_C$) and In-House ($LNPR_H$) variables are the natural logs of the estimated tax prices of basic research and in-house R&D respectively. Although there are many factors that affect a firm's R&D expenditure choice, the regression suggests a fair amount of the variation in research expenditure for the period 1994 through 1999 can be explained by the tax prices of R&D. Firms average sales and historical research intensity (R&D to sales ratio) were the only statistically significant fixed effects. Other potential right-hand-side variables were also tested, but were not found to have any significant explanatory power. These variables included the net income/loss position of the firm each year, SIC-specific dummy variables, a year-end dummy segregating the firms with a December 31 reporting date from firms with a fiscal year-end date, an initial public offering (IPO) variable and the age of the firm as of 1994.

Table 2: Comparative Regression Results

	Pharmaceuticals		Software	
	10-K	*Compustat*	*10-K*	*Compustat*
VARIABLE				
LNPRC	-1.7932***	-1.8095	-1.6491**	-5.7840**
	(.5344)	(.1495)	(.9529)	(2.6739)
LNPRH	-12.0254***	-27.3879**	-0.9758	25.7429***
	(4.4649)	(12.8986)	(3.3014)	(7.6263)
INTENSE	-0.0061**	0.0031	0.6241***	0.5981***
	(.0028)	(.0040)	(.0760)	(.1476)
LNSALES	0.0311***	0.0689***	0.7012***	0.5329***
	(.0115)	(.0268)	(.0347)	(.0425)
N	192	192	150	150
Adj-R^2	0.2620	0.1390	0.8310	0.5650

Note: ***, **, and * represent significance to .001, .05 and .10, respectively. White's adjusted standard errors are shown in parentheses.

Changes in the size of the coefficients or significance of the tax price variables provide evidence of structural shifts between the time periods. The most compelling evidence of the effect state-level tax credit changes on research expenditure is suggested by differences between the 1994-1996 and 1997-1999 reflecting firms' sensitivity

to the reduced tax prices of R&D that resulted from California's increases in research tax credit rates that took effect in 1997. The results for the combined 1994-1999 periods serve as a benchmark from which the robustness of the model in explaining research expenditure may be evaluated. With respect to the model's ability to explain R&D investment, there are dramatic differences in the coefficients across industries. Note the differences in coefficients obtained from the two sources of data.

The pharmaceuticals show consistently negative tax price coefficients, reflecting the inverse relationship between firms' investment in R&D and the tax price faced. The in-house R&D (LNPRH) variable shows consistently more negative coefficients, highlighting these firms' greater sensitivity to changes in the tax price of in-house research over basic research. This outcome may seem somewhat surprising. However, the 10-K disclosures of many of the pharmaceuticals included subcontracting relationships with larger, more mature pharmaceuticals. Firms hired to perform research for others are probably less likely to subcontract research to colleges and universities. If the firms are not investing in basic research, the tax price of research should not have significant explanatory power. Hence, the lower basic research (LNPRC) coefficient may be more a reflection of the characteristics of the pharmaceutical firms in the sample than the effect the tax price of contract R&D has of firms' research choice.

Table 3: Comparative Results by Time Period- Pharmaceuticals

	1994-1996 Pharmaceuticals		1997-1999 Pharmaceuticals	
VARIABLE	*10-K*	*Compustat*	*10-K*	*Compustat*
LNPRC	2.8072	-2.0500	-1.7642***	-1.5086**
	(3.4844)	(7.9349)	(.6954)	(.7602)
LNPRH	-148.8600***	-166.3730***	-30.8675***	-31.7852**
	(28.7007)	64.6168	(12.4532)	(16.0618)
INTENSE	-0.0024	0.0075	-0.0072**	-0.0012*
	(.0035)	(.0061)	(.0032)	(.0241)
LNSALES	0.0553***	0.0837**	0.0362**	0.0409*
	(.0149)	(.0393)	(.0177)	(.0241)
N	96	96	96	96
Adj-R^2	0.3960	0.1260	0.2880	0.2430

Note: ***, **, and * represent significance to .001, .05 and .10, respectively. White's adjusted standard errors are shown in parentheses.

The software in-house results are startling. Unlike the pharmaceuticals, the R&D intensity and sales variables are consistently significant and larger for the software firms. This most likely reflects the inherent research investment differences between the two industries. Many of the pharmaceuticals' 10-Ks indicated they are categorized as development stage companies; firms that have never had product sales. The software firms typically had multiple products and maintenance contract revenues. Thus, research intensity measured in terms of a ratio to sales and the logarithm of average sales have a greater relationship to software firms' R&D expenditure choice than the pharmaceuticals in this study. Of greater interest is the Compustat significance and positive sign on the tax price of in-house research. A positive sign seems to indicate a positive relationship between in-house R&D's tax price and firm investment. This seems unlikely.

Table 4: Comparative Results by Time Period- Software

	1994-1996 Software		1997-1999 Software	
VARIABLE	*10-K*	*Compustat*	*10-K*	*Compustat*
LNPRC	4.7864***	8.1383***	4.7864***	8.1383***
	(1.0531)	(1.1346)	(1.0531)	(1.1346)
LNPRH	9.9985*	-0.7377	9.9985*	-0.7377
	(5.6059)	(9.4517)	(5.6059)	(9.4517)
INTENSE	9.5698	38.3679***	9.5698	38.3679***
	(8.3013)	(12.2005)	(8.3013)	(12.2005)
LNSALES	0.6225***	0.6188***	0.6225***	0.6188***
	(.1073)	(.2435)	(.1073)	(.2435)
N	75	75	75	75
Adj-R^2	0.769	0.49	0.769	0.49

Note: ***, **, and * represent significance to .001, .05 and .10, respectively. White's adjusted standard errors are shown in parentheses.

DOI: 10.58809/AKCP5412

In order to better understand the unexpected sign of the Compustat software firms' in-house research coefficient, analysis across the two time periods needed. Tables 3 and 4 above report the comparative data source results across each time period for the pharmaceutical and software firms, respectively. The Compustat- and 10-K-based results for the pharmaceuticals are quite similar. The primary difference is from differences in the level of the coefficients. The tax prices, if significant, are consistently negative for the results based on both data sources. However, the coefficients from the Compustat sample tend to be larger. This is probably due to Compustat's inclusion of in-process R&D (IPRD) amortization as part of R&D expense.

Larger expense values used in the analysis would tend to bias the effect of an R&D credit upward. A likely explanation may reflect the limited amount of resources devoted by these firms toward basic research investment. Another explanation may be related to the inherent difficulties in appropriately labeling software R&D through the three-stage process required by the FASB. The investment choices of research managers may ultimately have little correlation to the amounts reported in the financial statements as research and development expense. More troubling from an empirical sense is the possibility that these accounting values may also bear little resemblance to the amounts the firms report as qualified research expenditure when applying for R&D tax credits. Thus, the use of public data for research on the software industry's sensitivity to changes in research tax prices may be inappropriate.

Further analysis of the 10-K and Compustat results reveals other important differences. There are inconsistencies in the parameter shifts from the early to later period, differences in signs, and differences in the levels of the coefficients. These results have several possible explanations. It may be that the model cannot detect the relationship between tax price and R&D choice for software firms because the tax prices were not that important to these firms in the 1990s. "Just to stay competitive, software developers have had to invest heavily in R&D at a pace that even exceeds their overall business growth patterns…A positive aspect of the rapid growth has been the lack of concern over any loss in government research funding due to federal budget cuts" (Research and Development, 1996: 4A). Thus, industry-specific factors such as the dot-com phenomenon may have been more influential in spurring the R&D decisions of software firms than changes in R&D tax credit policy.

Alternately, determining R&D expense precisely is more complex for software firms, thus introducing noise in the estimation. In particular, SFAS No. 86 Accounting for the Costs of Computer Software to be Sold, Leased or Otherwise Marketed (1985) requires firms to follow a three stage categorization of research costs related to software R&D. For example, software R&D costs incurred prior to product feasibility being established should be expensed. During software production, all relevant costs should be capitalized (reported as an asset on the balance sheet). Once the product is available to customers, amortization of the capitalized costs should begin. Determining when a product has progressed from stage one, where costs are expensed, to stage two where costs are capitalized is difficult and subjective. Thus, the amounts reported as research and development expense in firms' 10-Ks may not necessarily reflect the actual research expenditure of firms. Additionally, Compustat may inaccurately code or classify the capitalized and expensed portions of R&D. Further complicating the issue are the differences in definitions utilized by the IRS and Generally Accepted Accounting Principles (GAAP). Qualified research expenditure, as specified by Internal Revenue Code 174, is not defined the same as research expense in SFAS No. 2. For example, SFAS No. 2 excludes legal fees related to patent applications, but IRC 174 includes them (Oliver, 2003).

The inconsistencies in the software results are somewhat troubling because we cannot clearly determine if the inconsistencies are due to Compustat data gathering and reporting issues or troubles associated with using financial statement values in general to estimate confidential tax return information, or more generally to difficulties in accounting for software R&D. Additionally, analysis based on subsequent time periods would highlight whether subsequent changes in accounting and reporting requirements regarding mergers and acquisitions, including treatment of intangible assets and in-process research show persistent differences between Compustat and 10-K data.

CONCLUSION

Despite the uncertainties with respect to the software results, the regression results are consistent with prior federal-level studies that show R&D tax credit effects vary across industries (Mansfield 1986; Berger 1993; Mamuneas and Nadiri 1996). There are dramatic differences in the R&D spending, sales and tax price coefficient estimates obtained for the software and pharmaceutical firms in this study. Of particular note are the implications these results have for other studies based on Compustat data. Because Compustat frequently included the amortization of IPRD in the value of R&D

reported, R&D expense was often overstated. This means the tax price sensitivity estimates obtained from Compustat data may be biased upward. The tax price coefficients were consistently higher when the Compustat data was used than when the 10-K values were tested. For policymakers, this means that bigger changes in R&D tax credit rates may be needed for state policy setters to achieve a target amount of R&D expenditure change.

Although there may be sample selection effects associated with analysis limited to single-state R&D producing firms, we cannot predict with certainty the direction of those effects. Some may suggest that a study based on single-state, small R&D intensive firms may overstate the California's research tax credit policy effectiveness. However, larger, more profitable firms with multiple-state R&D may stand to benefit even more from a state-level incentive than the firms included in this paper. If the location of research activity is a strategic decision, then firms with multiple-state operations may react by reallocating the same aggregate spending to those locations with the lowest tax price in order to benefit from the increased credit rate. Given the incremental nature of the credit, in order for firms to benefit from higher rates of credit, their research intensity must increase over time. Thus, for multi-state R&D performers, the appearance of increased R&D intensity for state credit calculation purposes can be achieved by merely shifting the location of current levels of activity. In contrast, the single-state R&D performing firms can only continue to receive the credit by actually increasing research spending relative to their sales. Based on the available data, there is no way of knowing which group of firms responds to a greater degree to changes in state-level policy.

These findings are important to state- and federal-level R&D policymakers for several reasons. First, policy makers should consider the differences across industries when crafting research policy on the federal or state levels. The software and pharmaceutical firms exhibited variation in R&D investment behavior, and were divergent in their sensitivity to state-level R&D tax credits. A common policy concern focuses on the possibility the public sector is financing research activity that would have occurred even in the absence of the incentive. In this regard, the positive software tax price coefficient is particularly troubling. The notion of software firms exhibiting a positive relationship between the tax price of R&D and research investment suggests these firms may have been rewarded for R&D that would have been undertaken even if the credit were not available. At best, the incentive effects of state level research tax credits are modest with the results dependent on the data source used.

This paper provides further evidence of the limitations and care that must be employed in evaluating research results obtained with Compustat's R&D data. Future research should consider alternate time periods and firms in other industries in order to shed light on the robustness of the results herein. Ideally, evaluation of R&D tax credit policy should be performed using tax return data.

REFERENCES

Altshuler, R. 1988. A dynamic analysis of the research and experimentation credit. **National Tax Journal,** 41: 453-466.

Anonymous. 1996. Increasing software R&D strains supply of skilled programmers. **Research and Development,** 38: 4A-6A.

Bailey, M., & Lawrence, R. 1992. Tax incentives for R&D: What do the data tell Us? **Council on Research and Technology**, Washington, DC.

Bean, A., & Guerard, J. 1989. A comparison of census/NSF R&D data vs. Compustat R&D data in the financial decision-making model. **Research Policy,** 18: 193-209.

Berger, P. 1993. Explicit and implicit effects of the R&D tax credit. **Journal of Accounting Research**, 31: 131-171.

Billings, B., & Fried, Y. 1999. The effects of taxes and organizational variables on research and development intensity. **R&D Management,** 29: 289-301.

Billings, A., Glazunov, S., & Houston, M. 2001. The role of taxes in corporate research and development spending. **R&D Management,** 31: 465-477.

Browning, B. 1997. Maximizing R&D writeoffs to reduce goodwill. **Mergers and Acquisitions,** 32: 29-30.

Collins, E. 1983. **An early assessment of three R&D tax incentives provided by the Economic Recovery Tax Act of 1981.** Vol. PRA Report 83-7. National Science Foundation, Washington, DC.

Dowdell, T., & Press, E. 2002. **Restatement of in-process research and development write-offs: The impact of SEC Scrutiny.** (Working paper) Philadelphia: Temple University.

DOI: 10.58809/AKCP5412

Edgar-online, Inc. 2001.

Eisner, R., Albert, S, & Sullivan, M. 1986. The new incremental credit for R&D: Incentive or disincentive. **National Tax Journal,** 37: 171-183.

Financial Accounting Standards Board. 1974. **Accounting for research and development costs.** Statement of Financial Accounting Standards No. 2, Norwalk, CT.

Financial Accounting Standards Board. 1982. **Research and development arrangements.** Statement of Financial Accounting Standards No. 68, Norwalk, CT.

Financial Accounting Standards Board. 1985. **Accounting for the costs of computer software to be sold, leased, or otherwise marketed.** Statement of Financial Accounting Standards No. 86, Norwalk, CT.

Hall, B. 1993. R&D tax policy during the 1980s: Success or failure? **Tax Policy and the Economy,** 7: 1-36.

Hall, B., & Long, W. 1999. **Differences in R&D data on the NSF/census RD-1 form and the SEC 10-K form: A micro-data investigation.** Washington: Center for Economic Studies, U.S. Bureau of the Census.

Hall, B., & Van Reenen, J. 2000. How effective are fiscal incentives for R&D? A new review of the evidence. **Research Policy,** 29: 449-469.

Hall, B., & Wosinska, M. 1999. **The California R&D tax credit: Description, history and economic analysis.** A Report to the California Council on Science and Technology, Sacramento, CA.

Hines, J. 1991. On **the sensitivity of R&D to delicate tax changes: The behavior of U.S. multinationals in the 1980s.** Working Paper No. 3930. National Bureau of Economic Research, Cambridge, MA.

Hines, J. 1993. On the sensitivity of R&D to delicate tax changes: The behavior of U.S. multinationals in the 1980s. In A. Giovannini, R. Hubbard, & J. Slemrod, (Eds.), **Studies in International Taxation,** Chicago: University of Chicago Press: 149-94.

Kern, B., & Morris, M. 1992. Taxes and firm size: The effect of tax legislation during the 1980s. **Journal of the American Taxation Association,** 14: 80-96.

Kern, B., & Morris, M. 1994. Differences in the COMPUSTAT and expanded value line databases and the potential impact on empirical research. **Accounting Review,** 69: 274-284.

Kinney, M., & Swanson, E. 1993. The accuracy and adequacy of tax data in Compustat. **Journal of the American Tax Association,** 15: 121-135.

Mamuneas, T., & Nadiri, M. 1996. Public R&D policies and cost behaviour of the US manufacturing industries. **Journal of Public Economics,** 63: 57-81.

Mansfield, E. 1986. The R&D Tax Credit and Other Technology Policy Issues. **American Economic Review,** 76: 190-194.

Manzon, G. 1994. The role of taxes in early debt retirement. **Journal of the American Taxation Associati n,** 16: 87-100.

McCutchen, W. 1993. Estimating the impact of the R&D tax credit on strategic groups in the pharmaceutical industry. **Research Policy,** 22: 337-351.

McGoldrick, B. 1997. Goodwill games. **Institutional Investor**, 31: 145.

Oliver, J. 2003. Accounting and tax treatment of R&D: An update. **CPA Journal**, 73: 46.

Paff, L. 2004. The alternate incremental credit (AIC): An empirical analysis. **Technovation,** 24: 41-52.

Paff, L. 2005. State-level R&D tax credits: A firm-level analysis. **Topics in Economic Analysis & Policy,** 5: 17.

Rosenberg, B., & Houglet, M. 1974. Error rates in CRSP and Compustat databases and their implications. **Journal of Finance,** 29: 1303-10.

San Miguel, J. 1977. The reliability of R&D data in COMPUSTAT and 10-K reports. **Accounting Review,** 52: 638-641.

Securities and Exchange Commission. Washington, DC.

Stallworth, H., & DiGregorio, D. 2005. Earnings management and accounting for acquisition. **Internal**

Auditing, 20: 9-16.

Standard and Poor's COMPUSTAT. 2000.

Swenson, C. 1992. Some tests of the incentive effects of the research and experimentation tax credit. **Journal of Public Economics,** 49: 203-18.

Vasarhelyi, M., & Yang, D. 1988. **Financial accounting databases: Methodological implications of using COMPUSTAT and Value Line databases.** Working Paper, Rutgers University, NJ.

Lolita Paff is assistant professor of business and economics at Penn State Berks. She received her Ph.D. in economics from Lehigh University. Her research interests include the economics of technological innovation, tax credits, and state-level business development incentives. She provided expert testimony to Pennsylvania's Senate Finance Committee regarding R&D tax credit policy and has published articles in Technovation, Topics in Economic Analysis and Policy and Pennsylvania Economic Review (forthcoming).

https://scholars.fhsu.edu/jbl/vol2/iss1/18
DOI: 10.58809/AKCP5412