NBER WORKING PAPER SERIES

INDUSTRY EFFECTS AND APPROPRIABILITY MEASURES IN THE STOCK MARKET'S VALUATION OF R&D AND PATENTS

Iain Cockburn

Zvi Griliches

Working Paper No. 2465

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1987

We are indebted to the National Science Foundation (PRS85-12758 and SES82-08006), the Sloan Foundation, and the NBER Productivity Research Program for finanical support of this work. We are grateful to Richard Levin for providing us with access to the original Yale Survey data, and to Peter Reiss for helpful comments on an earlier version of this paper. The research reported here is part of the NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #2465
December 1987

Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents

ABSTRACT

This paper examines the stock market's valuation of a firm's innovative activity. We estimate the market's relative valuation of firms' tangible and intangible assets, focusing on knowledge capital in the form of accumulated "stocks" of R&D and patents. We tried to improve upon our estimates of the stock market's valuation of knowledge capital embodied in such "stocks" by bringing in measures of the appropriability environment facing a firm from the Yale Survey on *Industrial Research and Development*. The responses to Survey questions about the effectiveness of patents as a mechanism for protecting the returns from innovation turn out to be of some use: there is evidence of an interaction between industry level measures of the effectiveness of patents and the market's valuation of a firm's past R&D and patenting performance, as well as its current R&D moves. We find no evidence, however, that other appropriability mechanisms differ enough across industries to leave measurable traces in our data. The structure of the Yale Survey makes it possible to estimate the sampling error in the appropriability measures derived from it. This information was used by us in an errors-in-variables context, but with little success. In the absence of R&D variables, our estimates imply that a two standard deviation increase in our index of patent-effectiveness would raise the value of a patent held by our average firm from $0.4 million to $1.0 million. When R&D variables are introduced into the equations, the patents variables become insignificant — R&D expenditures are a better measure of input to the innovative function of firms than patents are of its output — but we estimate that the same experiment would induce changes in q of between 10 and 27 percent for the average firm, approximately doubling the market's valuation of this kind of capital.

Iain Cockburn
NBER
1050 Massachusetts Ave
Cambridge, MA 02138
(617) 868-3900

Zvi Griliches
Department of Economics
Littauer 125
Harvard University
Cambridge, MA 02138
(617) 495-2181

This paper examines the stock market's valuation of a firm's innovative activity. Following Griliches (1981), we estimate the market's relative valuation of firms' tangible and intangible assets, focusing on knowledge capital in the form of accumulated R&D efforts and patent rights, and ignoring other intangibles such as goodwill, advertising, and sector-specific human capital. We use accumulated stocks of R&D expenditures and the number of patents granted as proxies for knowledge capital or the output of innovations. One problem with using these numbers is that quality and not quantity is likely to be the crucial factor in the market's determination of the value of such assets. There is little we can do about this with the available data. An equally serious problem, however, is that even if all R&D and patents were of comparable quality, the market's valuation of a given amount of innovative activity will vary according to how successfully a firm can appropriate the returns from this investment, and this is the issue that the paper addresses.

Firms have a variety of possible mechanisms for preventing competitors from taking advantage of their investment in knowledge capital, and the availability and effectiveness of these mechanisms varies across firms and industries. In particular, the effectiveness of patents as a mechanism for appropriating the returns from R&D is not constant across firms and industries, and therefore the present value of returns to a firm from investing in patent protection should differ according to industry conditions and firm specific factors. Failure to take this into account may have been the cause of some puzzling results in the 1981 paper, where patent variables became insignificant (and in some cases wrong-signed) in the presence of R&D variables. (See Table 3.) Survey data collected by the Yale group (Levin, Klevorick, Nelson, and Winter) makes it possible, at least in principle, to construct measures of the appropriability of R&D at the industry level. Our primary interest here is in the patenting mechanism, and this paper presents results obtained from matching the Yale Survey to the NBER data on R&D and patenting intensity of large US manufacturing corporations in an attempt to control for inter-firm variability in the patenting environment.

The Equation to be Estimated

In a rational stock market, a firm's stock price should be the expected discounted value of the net income which will be derived from its assets. As a matter of definition, we can write:

$$V = g(\text{tangible capital, intangible capital}) \tag{1}$$

Under constant returns to scale, or as a local linear approximation (see Wildasin (1984)), this can be written as:

$$V = b[A + \delta K] = e^{(\lambda_t + \mu_i)}[A + \delta K] \tag{2}$$

where A is tangible capital, K is intangible capital and δ is its relative shadow price, while b, the average multiplier of market value relative to the replacement cost of total assets, consists of two multiplicative components: an overall market index λ_t and a firm-specific component μ_i. Re-arranging,

$$q = \frac{V}{A} = b(1 + \delta\frac{K}{A}) \tag{3}$$

taking logarithms, and exploiting the fact that $\log(1 + x) \simeq x$ when x is small,

$$\log(q) \simeq \lambda_t + \mu_i + \delta\frac{K}{A} \tag{4}$$

which is interpretable as a regression equation in which K is a vector of variables representing a firm's intangible assets. [1] To the extent that the valuation of various proxies for such intangible assets, e.g. patents, varies from industry to industry the estimated δ's need not be identical across firms or industries. We shall explore this possibility by allowing the estimated δ's to differ across industries, and/or by interacting the various measures of intangible capital with indices of the 'ease of appropriability' derived from the Yale Survey data.

This approach raises the question of the extent to which both the levels of past investment in R&D and the propensity to patent R&D results already reflect the appropriability situation directly. Both R&D and patenting are obviously endogenous with respect to appropriability indices. [2] Our limited success in detecting the effects of such indices may reflect this fact.

[1] The logarithmic transformation is used to allow easier comparisons over time. Since the average market q changes from year to year (and day to day), the multiplicative form can isolate it through the introduction of year dummies. To analyze one cross-section at a time, one could also have used a simple linearization of equation 1.

[2] The actual relation between the propensity to patent and our appropriability measures is very weak in our data: in regressions of the ratio of patents to R&D onto appropriability indices we obtain R^2's of less than 0.1 and contradictory coefficient estimates. If there were a strong selectivity bias in the propensity to patent vis-à-vis differences in appropriability, the original relation between market value and patents should have been much stronger than that which we have observed in the data.

FIG. 1.

Section I: Protecting the Returns from R&D

The questions in Section I ask about the effectiveness of alternative means of protecting the competitive advantages gained by successful research, development, and engineering.

*** Questions About New or Improved PROCESSES ***

I.A. In this line of business, how effective is each of the following means of capturing and protecting the competitive advantages of new or improved production processes?

[CIRCLE ONE NUMBER ON THIS SCALE FOR EACH]

	Not At All Effective			Moderately Effective			Very Effective
1. Patents to prevent competitors from duplicating the process	1	2	3	4	5	6	7
2. Patents to secure royalty income	1	2	3	4	5	6	7
3. Secrecy	1	2	3	4	5	6	7
4. Lead time (being first with a new process)	1	2	3	4	5	6	7
5. Moving quickly down the learning curve	1	2	3	4	5	6	7
6. Superior sales or service efforts	1	2	3	4	5	6	7

Comments:

*** Questions About New Or Improved PRODUCTS ***

I.B. In this line of business, how effective is each of the following means of capturing and protecting the competitive advantages of new or improved products?

[CIRCLE ONE NUMBER ON THIS SCALE FOR EACH]

	Not At All Effective			Moderately Effective			Very Effective
1. Patents to prevent competitors from duplicating the product	1	2	3	4	5	6	7
2. Patents to secure royalty income	1	2	3	4	5	6	7
3. Secrecy	1	2	3	4	5	6	7
4. Lead time (being first with a new product)	1	2	3	4	5	6	7
5. Moving quickly down the learning curve	1	2	3	4	5	6	7
6. Superior sales or service efforts	1	2	3	4	5	6	7

Comments:

The Data

We have combined two separate data sets in this study: the NBER RND-PANEL data set and the Yale Survey results.

The first is a large data set in panel format compiled from the Compustat and Patent Office tapes. Over 1800 firms are represented, with Compustat accounting data and patent figures from the late 1960's through 1984. The panel is not balanced, since complete data are not available for every firm in every year. The panel is widest around 1976, with significant attrition through time in both directions. (See Cummins, *et al.* (1985) for more detail.) We use a balanced, "cleaned" subset of these data, comprising 722 firms which met the joint requirements of:

- no missing observations on the variables of interest for the years 1973–1980.
- no 'jumps' in employment or capital stock, to avoid problems of mergers, takeovers, and the like, where a jump is defined as an absolute or relative change of more than 50 percent.

The Yale Survey took the form of a questionnaire on *Industrial Research and Development* mailed to R&D executives in 1562 business units in over 130 industries defined at the Line of Business (LB) level. In all 650 usable responses were obtained. The questionnaire posed over 120 questions about various mechanisms for appropriating returns from R&D and their effectiveness, the nature of technical progress, and the general relevance of science. Respondents were typically asked to answer on a 7-point scale from 1 = *not important* to 7 = *very important*. A sharp distinction was made between process and product innovation throughout. In Figure 1 we reproduce a page from the questionnaire to illustrate the type of question asked.

The Yale Survey contains a wealth of information, which has been reported elsewhere (Levin, *et al.* (1984), and Levin, *et al.* (1985)). However, the very richness of its detail makes it difficult to derive a single numeric measure of appropriability. Moreover, it is not immediately clear how responses about industries should be matched to our firms.

Measures of Appropriability

There are at least two ways of approaching the derivation of an 'appropriability conditions' index for an industry from the answers to the Yale Survey.

The first is to look for single questions which stand out as summaries of the information in the various sections of the questionnaire. The second is to use data reduction techniques to compress the information in related groups of questions into a single numeric measure. Since our interest is focused primarily on the role of patents, we have sought to draw a distinction between patent and non-patent mechanisms of appropriability in summarizing these data. In general we have not maintained the Yale group's distinction between products and processes: our dependent variable, the future net earnings of the firm capitalized in its market value, is affected by both cost reductions due to process innovation and revenue increases due to product innovation. Exploratory calculations indicate little gain from such a disaggregation.

Respondents' opinion on the effectiveness of a number of mechanisms for appropriating the returns from innovation (patents, secrecy, lead time, learning curve, sales and service effort) was asked in Questions IA1–IA6 (process) and IB1–IB6 (product).

The main variables we have constructed to measure the effectiveness of patenting as a mechanism for appropriating the returns from R&D are based on questions IA1 (*do process patents prevent duplication ?*) and IB1 (*do product patents prevent duplication ?*) A simple summary variable, *PPP* (**P**atents **P**rovide **P**rotection), was constructed as the sum of IA1 and IB1. However, since the scale of the responses is arbitrary, and may be quite non-linear (an answer scored as 6 may in some relevant sense be twice as far from 4 as 4 is from 5), we tried various transformations of these questions: calculating a "stretched" version of *PPP* where the distance between each score doubles going in either direction away from 4. As an alternative to algebraic transformations of scale, we also constructed two distribution-based variables: a "trichotomized" *PPP* where the variable takes on a value of -1 if the score is 2 or less, 0 if between 3 and 5, and 1 if 6 or more, and *FPP*, (**F**raction of respondents who say **P**atents **P**rovide **P**rotection) defined as the percentage of respondents within an industry answering 6 or 7 to question IB1 minus the fraction answering 1 or 2.

Sections IIC, IID, IIE, and IIF of the Questionnaire provide alternative measures of the effectiveness of patents in terms of imitation lag and imitation cost. The questions ask for an estimate of the imitation cost/lag in four cases: where an innovation is patented/non-patented and of a major/minor nature. Adding together the estimated imitation lag for major and minor patented innovations, and subtracting the estimated lag for major and minor non-patented innovations (and dividing by 4 to retain a natural scale) gives an estimate of the increment in imitation lag due to patents (*IML*). A similar calculation gives an estimate of the increment in imitation cost due to patents (*IMC*).

To measure respondents' opinion on the effectiveness of non-patent mechanisms (secrecy, lead time, learning curve, sales and service we used Questions IA3–IA6 (process) and IB3–IB6 (product). Rather than attach particular weight to any of these, following Levin (1986) we took each respondent's maximum score over all of these questions, as a measure of the availability of some non-patent appropriability mechanism to them (*MNP*), and also computed the mean of these eight questions (*NPP*).

An attempt was made to derive common factors from subsets of the Survey questions using factor analysis. There does appear to be some structure across the bulk of the questions at the individual respondent level which is consistent with our ***a priori*** beliefs: the pattern of factor loadings for the first two common factors supports our presumption that appropriability mechanisms fall into two largely orthogonal 'patent' and 'non-patent' classes. However the large amount of noise in the data at the individual respondent level prevented us from extracting any satisfactory simple factor structure: the standard χ^2 test rejected the hypothesis that four or fewer common factors can adequately explain the correlation in responses to the first 12 questions. Though we constructed factor scores for the first two 'patents' and 'non-patents' -like factors and carried them through the subsequent analysis, these new variables performed poorly in comparison to the simpler summaries of the data described above, presumably because of the signal-to-noise problem. Another more fundamental problem is that we have no real hypotheses about the causal structure underlying respondents answers to the Survey questions, and have just assumed the simplest possible orthogonal factor structure.

Matching Data Across Data Sets

Appropriability is presumably a property (or lack thereof) of a particular product or process in a particular legal and institutional setting. It may differ across industries if their products/processes differ enough in these dimensions. Where industry conditions (market structure, nature of technology, regulatory environment, etc) are such that there is a high degree of appropriability, firms are able to monopolize the returns from their innovative activity; where industry conditions are such that there is a low degree of appropriability, innovations will be quickly adopted by other firms in the industry, and the returns to an individual firm from doing R&D will be commensurately lower.

An industry in this sense is quite clearly defined at the conceptual level, but (as usual) is difficult to define in practice. Moreover, industry boundaries may not conform to firm boundaries, and hence may not fit too well with our basic data. Our approach has been to use the SIC product classification as a basis for defining industries, with some additional aggregation where sample sizes were small and the industries appeared to be similar in technological and other respects. Having defined some measure of appropriability, taking the mean response within each industry in the Yale data gave us a point estimate of appropriability levels for each such industry in their data set. Note that each industry has a different number of respondents, and hence appropriability is measured with different degrees of accuracy in different industries. Having estimated a set of industry appropriability levels, we then attributed them to each firm in the equivalent Compustat industry.

This raises a number of problems. The first is how to match industries. The institutional identities of the Survey respondents are not available to us, and we can only identify them by a line-of-business (LB) code assigned by the Yale group following the FTC's classification. There are 130 separate LB's so defined. For our matched firms we have only each firm's primary 4-digit SIC code assigned by Compustat. By using a correspondence between FTC LB codes and 4-digit SIC codes supplied by Levin it was possible to assign 2, 3, or 4 digit SIC codes to each respondent in the Yale Survey. In merging data sets we sought to match firms/industries at the lowest possible level of aggregation. However not all IDS-industries mapped into a single 4-digit or even 3-digit SIC-industry, and neither do all of our firms have a "genuine" 4-digit SIC code (3331 vs 3300). By forcing a 4-digit matching, we lost information by having to discard many firms, either because their product was not well-enough represented in the population of Survey respondents, or because we could only classify them as 2-digit firms. Conversely, by merging at a 2-digit level we stood to lose much of the richness of the Survey data, and ran the risk of introducing serious errors in assigning firms to industries in which the Survey responses were all based upon quite different products or processes (for example, a firm in SIC 3351, Copper Rolling and Drawing, would be assigned the appropriability level for SIC 33 when all the Survey respondents were from SIC 3312, Blast Furnaces and Steel Mills.)

The compromise solution was to do our analysis twice, using data sets formed first by matching firms and the Yale Survey respondents at a '3.5-

digit' level, in which IDS and SIC codes are mapped within each data set into a new 55 industry scheme called IND (devised by us from the SIC), then again at a '2.5-digit' level, according to a 24 industry scheme called NSF (which approximates the breakdown in the NSF R&D publications). In both cases, industries which had less than three respondents in the Survey were dropped on the grounds that at least three observations are needed to form reasonable estimates of the industry mean and its variance. By using the IND level matching we preserve some of the richness of detail available in the data at the expense of potentially mis-assigning firms to industries. Conversely, by using the NSF level matching we lose detail but are more confident that our firms are being assigned to the right industry.

Given that industries can be satisfactorily defined at some acceptable level of aggregation, we need to address the issue of the sampling error in our estimates of appropriability. Because each of the industries within the Yale Survey has a number of respondents, we can compute not only a different industry mean level for our appropriability measures, but also a within industry response variance and a corresponding standard error of the mean ($\sqrt{\sigma^2/n}$). In matching the data sets we can therefore take into account the fact that our appropriability measures are estimated with different degrees of accuracy for each industry.

Our main analysis will concentrate on a 1980 cross-section of manufacturing firms, using the log of q (the ratio of the market value of the firm to the replacement cost of its assets) as our dependent variable. 1980 is the latest year for which our data on patents granted by "date applied for" is reasonably complete. Our results have been checked against 1973 and 1979 cross-sections: essentially similar results were obtained, which we do not report here for the sake of brevity.

The major independent variables are: K the cumulated stock of past R&D expenditures (using a 15 percent depreciation rate); SP, the "stock" of cumulated past patents (using a 30 percent depreciation rate); and NR, an estimate of the current year's net investment in R&D, which is calculated as $NR = R\&D - 0.15K$, where $R\&D$ is the current year's R&D expenditure and K is the stock of R&D carried forward from the end of the previous year. These three variables have all been divided by the total fixed assets of the firm, and also interacted with the appropriability measures.

Table 1 below presents summary statistics for the main variables calculated across the whole of our sample. Industry means and standard deviations for some of the variables in our sample, together with some of the appropriability measures constructed from the Yale Survey are presented for both the IND and NSF levels of aggregation in Appendix C.

Table 1 somewhere here

An immediate question about these variables is the amount of useful information they contain about differences between industries. Table 2 presents the results of an analysis of variance for each of the variables constructed from the Survey, listing the F-ratios for testing the null hypothesis of no significant inter-industry variation in these measures. These F-ratios are calculated for the 55 IND classes and 24 NSF classes within the Yale Survey, and again for a smaller sample from which the drugs and computer industries have been excluded. The results for K/A (ratio of R&D stock to assets) and SP/A (ratio of Patents stock to assets) are presented for the same classes within our sample of firms to provide a benchmark (though note that the n's are different).

Table 2 somewhere here

Using the IND definition of industries, the F-statistic accepts the null hypothesis of no difference in industry means at the 5 percent level for all of our appropriability measures except *PPP*, and the transformed versions of it, *SRP* and *TRP*. For the NSF definition of industries there does appear to be a significant difference in industry means also for the *MNP* variable, but it disappears when we exclude the two "extreme" industries, drugs and computers, from this sample. Interpreting the F-statistic as the ratio of the between to the within variance, these results are less than encouraging. It is not obvious that there is much systematic between industries variance in these measures of appropriability, and hence their quality as indicators of differences in the inter-industry patenting environment may be rather low. [3]

[3]Notice also that the "stretching" and "trichotomizing" transformations of *PPP* have very little effect upon its F-statistic (nor did they have any effect upon the equivalent results for other variables, not reported here.) If our difficulties in deriving variables with a strong and systematic inter-industry variation are due to problems in finding the right scale with which to interpret the Survey responses, it is puzzling to find that these non-linear transformations have so little effect. On the other hand, these results may simply reflect the fact that in most industries the distribution of responses is tightly bunched around the 'neutral' score of 4, and these transformations may not be nonlinear enough in the right range. Another possibility is that our industry definitions may not reflect either

Table 1

Descriptive Statistics for 1980 Sample

VARIABLE	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE
log(V/A)	-0.273	0.697	-2.262	2.222
K/A	0.144	0.170	0.000	1.559
NR/A	0.008	0.023	-0.122	0.226
SP/A	0.105	0.206	0.000	3.817
PPP	7.632	1.307	4.400	11.412
SRP	28.828	4.719	17.000	43.176
TRP	-0.162	0.538	-1.400	1.235
IML	1.196	0.814	-0.201	3.729
IMC	0.254	0.322	-0.400	1.081
FPP	0.056	0.348	-0.667	1.000
NPP	4.720	0.378	3.333	5.375
TRN	0.285	0.131	-0.167	0.531
SRN	17.410	1.412	12.333	20.125
MNP	6.163	0.330	5.000	6.800
ASSETS	1137.696	3869.245	2.070	66498.563
EMP	14.926	42.864	0.050	746.000
SALES	1513.923	5725.416	0.653	103143.000

Full variable definitions are in Appendix B.
V = market value of the firm
A = Total net assets at replacement cost.
K = "Stock" of R&D using 15 percent depreciation rate.
NR = "News in R&D": current R&D less depreciation of the R&D stock.
SP = "Stock" of Patents using 30 percent depreciation rate.
PPP = Sum of scores on 'Patents Provide Protection' questions for products and processes (IA1 + IB1).
SRP = Stretched PPP.
TRP = Trichotomized PPP.
IML = Imitation Lag in years.
IMC = Imitation Cost, as fraction of innovator's.
FPP = Fraction in industry who answer 6 or 7 to question IB1 (Product Patents Provide Protection) – fraction answering 1 or 2.
NPP = Average effectiveness of Non-Patent appropriability mechanisms.
SRN = Stretched NPP.
TRN = Trichotomized NPP.
MNP = 'maximal' version of NPP.
SALES = Net Sales in $millions.
EMP = employment in thousands.

Table 2

ANOVA Results

Class:	IND	IND*	NSF	NSF*
PPP	3.23 (0.001)	2.61 (0.001)	5.73 0.0001	4.45 (0.001)
SRP	3.27 (0.001)	2.67 (0.001)	5.77 (0.001)	4.49 (0.001)
TRP	3.14 (0.001)	2.58 (0.001)	5.66 (0.001)	4.42 (0.001)
IML	1.18 (0.19)	0.98 (0.54)	2.19 (0.001)	1.80 (0.02)
IMC	1.36 (0.05)	1.25 (0.02)	2.25 (0.001)	2.08 (0.003)
MNP	1.07 (0.34)	0.97 (0.55)	1.56 (0.05)	1.37 (0.13)
NPP	1.08 (0.33)	0.97 (0.54)	1.39 (0.11)	1.16 (0.28)
SRN	0.97 (0.54)	0.84 (0.77)	1.20 (0.24)	0.94 (0.54)
TRN	0.98 (0.51)	0.87 (0.73)	1.25 (0.20)	1.02 (0.44)
K/A	9.78 (0.001)	9.42 (0.001)	19.47 (0.001)	18.52 (0.001)
SP/A	2.95 (0.001)	2.87 (0.001)	4.35 (0.001)	4.49 (0.001)

Table entries are the F-statistic for $H_0 : \mu_1 =, \ldots, = \mu_N = 0$. $Pr[> F]$ is in parentheses. The * indicates that the drugs and computers industries have been deleted from the sample.

For variable definitions see Table 1 and Appendix B.

A more exacting question is whether the data provide much more information than that contained by rather crude industry dummies. One way of asking it has already been done above by excluding the most obvious outlying industries: drugs and computers. Another way is to ask whether the continuous detail at the 55 industries level adds much above what might have been captured by just 10 2-digit industry dummy variables, and their interactions with our other explanatory variables. Of course these data provide an interpretation of such dummies and hence are of interest in their own right.

Estimation Results Controlling for Appropriability

In this section we report various attempts to improve upon our estimates of the market's valuation of patent outcomes and R&D investments by incorporating the different appropriability measures into the analysis of firm data. Tables 3a and 3b present the major results of our work using one of the better fitting appropriability indicators: *PPP* – "Patents Provide Protection". Table 3a shows the results of adding it to our standard valuation equation and interacting it with the various patent stock and R&D variables. Table 3b adds 10 2-digit level industry dummy variables to the same equations. In Table 4 we explore the use of alternative summary forms for the patents provide protection question. In Table 5 we look at the role of other "appropriability" mechanisms, such as secrecy and lead time, and compare the performance of measures based on these responses to the performance of the indices based on the responses to the effectiveness of patents questions. In Table 6 we compare the results of similar analyses for the 1979 cross- section and combine them with the 1980 data in a "seemingly unrelated" regression framework which allows, implicitly, for the presence of other unmeasured (but uncorrelated) individual firm effects.

Table 3a summarizes most of the major results. In the absence of R&D variables, past patenting does appear to capture some relevant aspects of "intangible" capital. Its coefficient is statistically significant and implies a valuation of approximately $0.5 million per patent granted. This is consistent with other evidence assembled in Griliches, Pakes, and Hall (1986).

---

the 'real' clustering of the respondents around particular technologies, or the industries that the respondents perceive themselves to be in, though limited experimentation with industry definitions had little effect upon these results. Nonetheless, the robustness of these ANOVA results leads us to believe that our inability to perceive of any useful inter-industry variation in much of the Survey responses is not simply a scaling problem, but a real effect, either in the nature of these data, or in the phenomena they seek to measure.

However, when measures of R&D are added to the equation this estimate either disappears (column 3) or is heavily attenuated (column 5). Adding a measure of the effectiveness of patent protection to these equations and interacting it with the patent stock and R&D variables improves the fit only marginally (by about .01) but does indicate the presence of an interaction. Without R&D variables the results imply a much higher valuation of patents in industries where patent protection is more effective. For example, column 2 in Table 3b could be read as indicating an average value of a patent of about $0.4 million which rises to about $1.0 million per patent in industries where the effectiveness of patents is two standard deviations higher than the average. When R&D variables are added in, the patent stock variables become less significant and the interaction is now attached to the R&D stock or the R&D "news" variable. The last columns of Tables 3a or 3b imply that the market values "news" in R&D much more highly than past investments or old patents and that such new R&D moves are valued about 50 percent higher in industries where patent protection is more likely to be effective. Adding separate industry intercepts to these equations attenuates these results somewhat, but does not eliminate them entirely.

Tables 3a and 3b somewhere here

Table 4 presents the results of trying to change the scale on which the responses to the patent effectiveness question were recorded. "Stretching" the scale, trichotomizing it, or measuring it by the excess of "high" responses over "low" makes surprisingly little difference to the results and neither did adding industry dummies. This was already visible in Table 2 where the ANOVA calculations yielded effectively identical results for the different versions of these variables, indicating that the lack of significant between industries variance in these measures is not an artifact of their scaling. The *FPP* measure does slightly better in terms of fit and the significance of some of the interaction terms, but not enough to change any of the conclusions significantly. We stick, therefore, with the simpler to interpret *PPP* measure in the rest of our analysis.

Table 4 somewhere here

Table 3a

The Stock Market's Relative Valuation of R&D and Patents

Dependent Variable: log(q)

	A0	A1	B0	B1	D0	D1
SP/A	0.493 (0.165)	0.785 (0.190)	0.111 (0.094)	0.192 (0.158)	0.246 (0.082)	0.309 (0.143)
PPP		-0.004 (0.019)		-0.012 (0.017)		-0.002 (0.017)
PPP*SP/A		0.333 (0.128)		0.076 (0.099)		0.094 (0.100)
K/A			1.374 (0.182)	1.442 (0.174)	0.741 (0.152)	0.694 (0.147)
PPP*K/A				0.303 (0.115)		
NR/A					11.99 (1.556)	12.82 (1.539)
PPP*NR/A						2.944 (1.249)
$\bar{R}^2$	0.027	0.037	0.125	0.133	0.258	0.265

N = 722. Mean of the dependent variable = -0.272, standard deviation = 0.697.
Heteroscedasticity-consistent standard errors in parentheses.
Matched by IND, 1980 Data.
All equations also contain an intercept term and the logarithm of Assets, whose coefficient was small but consistently significant, on the order of -0.03 (0.01).

For variable definitions see Table1 and Appendix B.

Table 3b

The Stock Market's Relative Valuation of R&D and Patents

Dependent Variable: log(q)

	A0	A1	B0	B1	D0	D1
SP/A	0.165 (0.100)	0.380 (0.171)	0.025 (0.097)	0.107 (0.167)	0.180 (0.093)	0.249 (0.155)
PPP		0.034 (0.024)		0.019 (0.024)		0.019 (0.023)
PPP*SP/A		0.236 (0.116)		0.075 (0.110)		0.098 (0.101)
K/A			0.837 (0.195)	0.932 (0.201)	0.385 (0.184)	0.335 (0.178)
PPP*K/A				0.365 (0.130)		
NR/A					11.18 (1.454)	11.96 (1.368)
PPP*NR/A						2.788 (1.231)
$\bar{R}^2$	0.166	0.172	0.191	0.200	0.304	0.310

Each equation has 10 2-digit industry dummies.

N = 722. Mean of the dependent variable = -0.272, standard deviation = 0.697.

Heteroscedasticity-consistent standard errors in parentheses.

Matched by IND, 1980 Data.

All equations also contain an intercept term and the logarithm of Assets, whose coefficient was small but consistently significant, on the order of -0.03 (0.01).

For variable definitions see Table1 and Appendix B.

Table 4

Comparision of Various Patent-based Appropriability Measures

	PAT=PPP	PAT=FPP	PAT=SRP	PAT=TRP
SP/A	0.192 (0.158)	0.254 (0.167)	0.192 (0.140)	0.181 (0.134)
PAT	-0.012 (0.017)	-0.018 (0.065)	-0.003 (0.005)	-0.001 (0.042)
PAT*SP/A	0.076 (0.098)	0.645 (0.516)	0.034 (0.034)	0.331 (0.352)
K/A	1.442 (0.174)	1.406 (0.175)	1.429 (0.172)	1.405 (0.180)
PAT*K/A	0.303 (0.115)	1.013 (0.500)	0.075 (0.032)	0.465 (0.322)
$\bar{R}^2$	0.132	0.135	0.132	0.129

N = 722. Mean of the dependent variable = -0.272, standard deviation = 0.697.
Heteroscedasticity-consistent standard errors in parentheses.
Matched by IND, 1980 Data.
All equations also contain an intercept term and the logarithm of Assets, whose coefficient was small but consistently significant, on the order of -0.03 (0.01).

For variable definitions see Table 1 and Appendix B.
PPP = Sum of scores on 'Patents Provide Protection' question for products and processes.
FPP = fraction in industry who answer 6 or 7 to question IB1 (Product Patents Provide Protection) – fraction answering 1 or 2.
SRP = "stretched" PPP
TRP = trichotomized PPP

Table 5 looks at the question of whether there is additional "power" in the responses to the other non-patent appropriability mechanisms questions. The basic answer is no. Neither by themselves nor in addition to the patent effectiveness measures do they add to the explained variance or result in significant interaction coefficients. In a way this was already foreshadowed in the ANOVA results. Given that there is little significant between industries variance in these measures it is not surprising that they cannot provide a sharp discrimination between the relevant environments that different firms find themselves in.

Table 5 somewhere here

Table 6 shows the results of combining 1979 data with the 1980 cross-section in a SUR framework. This procedure takes into account the serial correlation between the left-out individual firm components in such equations. The combined results are a bit stronger than the individual year ones but the conclusions remain the same: the R&D variables are "stronger" than the patent ones and the patent effectiveness measures improve the fit marginally and indicate the presence of some interaction between the "quality" of the appropriability environment and the market's valuation of a firm's R&D policy. Adding a 1973 cross-section to the SUR regression, gives similar results, not shown here.

Table 6 somewhere here

There remains the question how much do we gain by using such measures of the "effectiveness" of appropriability mechanisms relative to the use of a cruder interaction with 2-digit level industry dummies. If instead of interacting the SP/A and NR/A variables with *PPP* we interact them with our 10 industry dummies, we get adjusted R^2's of .179 and .315 versus the comparable values in Table 3b of .170 and .310 for columns 3 and 6 respectively. In this sense the *PPP* variable does quite well. It effectively accomplishes the same thing as 10 dummy variable cross-product terms and because it uses up only one degree of freedom it provides a more powerful test of the underlying hypothesis and a more useful interpretation of the data.

Table 5

The Additional Explanatory Power of Non-Patent Appropriability Measures

	1	2	3	4
SP/A	0.727 (0.172)	0.124 (0.188)	0.360 (0.170)	0.077 (0.183)
PPP	0.008 (0.020)	-0.006 (0.021)	0.035 (0.024)	0.023 (0.025)
PPP*SP/A	0.331 (0.137)	0.130 (0.147)	0.267 (0.133)	0.115 (0.142)
K/A		1.420 (0.202)		0.898 (0.224)
PPP*K/A		0.298 (0.161)		0.432 (0.172)
NPP	0.239* (0.071)	0.111* (0.083)	0.039* (0.079)	0.100* (0.089)
NPP*SP/A	-0.018* (0.402)	0.294* (0.441)	0.127* (0.388)	0.174* (0.432)
NPP*K/A		-0.179* (0.637)		0.263* (0.636)
$\bar{R}^2$	0.051	0.144	0.170	0.198
F-statistic	6.065	1.087	0.151	0.511
$Pr[> \quad F]$	0.002	0.354	0.860	0.679

The F-statistic tests the hypothesis that the starred coefficients in each equation are jointly equal to zero.
Equations 3 and 4 have 10 2-digit industry dummies.
N = 722. Mean of the dependent variable = -0.272, standard deviation = 0.697.
Heteroscedasticity-consistent standard errors in parentheses.
Matched by IND, 1980 Data.
All equations also contain an intercept term and the logarithm of Assets, whose coefficient was small but consistently significant, on the order of -0.03 (0.01).
For variable definitions see Table 1 and Appendix B.

Table 6

SUR Regressions using 1979 and 1980 Cross-Sections

Dependent Variable: log(q)

	Unrestricted				Restricted	
	79	80	79*	80*	79&80	79&80*
SP/A	0.189	0.293	0.146	0.234	0.114	0.106
	(0.103)	(0.139)	(0.103)	(0.138)	(0.101)	(0.101)
PPP	0.012	-0.003	0.040	0.021	0.009	0.034
	(0.016)	(0.017)	(0.020)	(0.022)	(0.015)	(0.020)
PPP*SP/A	0.132	0.134	0.123	0.144	0.116	0.109
	(0.068)	(0.094)	(0.068)	(0.094)	(0.066)	(0.067)
K/A	0.751	0.837	0.416	0.457	0.820	0.479
	(0.118)	(0.137)	(0.133)	(0.151)	(0.116)	(0.130)
NR/A	6.947	8.597	6.179	7.945	7.424	6.669
	(0.725)	(0.821)	(0.735)	(0.823)	(0.690)	(0.699)
PPP*NR/A	2.283	2.111	2.089	2.032	2.357	2.238
	(0.854)	(0.864)	(0.857)	(0.871)	(0.791)	(0.797)

The starred equations also have 10 2-digit industry dummies.
For variable definitions see Table 1 and Appendix B.

Estimation Results Correcting for Errors in Variables

The procedure by which we have matched the Survey data to our sample introduces two forms of error in our appropriability variables: sampling error in the estimation of industry means within the Survey, and a potential locational error in assigning firms/respondents to industries. While we do not address the locational error here, the sampling error in estimating industry means can be treated as a classical errors-in-variables problem. Recall that when some or all of the explanatory variables in a regression are measured with error, the cross-products matrix becomes "attenuated" by a multiple of the variance-covariance matrix of the measurement errors (see Appendix A for a fuller exposition.) If

$$y = X\beta + \epsilon \tag{5}$$

and

$$\tilde{X} = X + V \tag{6}$$

where X is the unobserved true design matrix, and V is a matrix of measurement errors with variance Σ then

$$E[\tilde{X}'\tilde{X}] = X'X + n\Sigma \tag{7}$$

which results in biased coefficients when equation 5 is estimated by OLS using $\tilde{X}$ as a substitute for X. In the case where Σ is known, or estimable, an consistent estimate of β can be obtained by applying OLS to the cross-products matrix after correcting for the attenuation by subtracting $n\Sigma$, with an appropriate adjustment of the standard errors.

Because we know the sampling error in our appropriability measures we can estimate Σ. The results of re-estimating the equations in Table 3a using a correction for the attenuation are presented in Table 7.

Table 7 somewhere here

The correction for attenuation appears to be quite successful for the simplest equation, doubling approximately the estimated patent stock coefficient and that on its interaction with *PPP*. However, for the more complex equations the patents variables again become insignificant, and the estimates become unstable, and implausible when the *NR/A* variable is included, and actually 'explode' when industry dummies are added. This reflects the fact that the estimated variance of the measurement error is almost equal to the total variance of our appropriability variables: the adjustment of the

Table 7

Regressions With "De-Attenuation" Adjustment for Errors-in-Variables

	A	B	D
SP/A	1.179 (0.262)	0.275 (0.327)	0.024 (0.360)
PPP	-0.014 (0.035)	-0.027 (0.034)	0.065 (0.031)
PPP*SP/A	0.783 (0.263)	0.157 (0.329)	-0.115 (0.328)
K/A		1.501 (0.202)	0.620 (0.155)
NR/A			16.201 (2.161)
PPP*NR/A			13.685 (5.800)
RMSE	0.678	0.646	0.586

Fuller's partial adjustment method is used. See Appendix A.

For variable definitions see Appendix B.

cross-products matrix brings it close to singularity, implying the absence of information in the appropriability variables above and beyond that which is already captured by the R&D and patent stock variables themselves together with a relatively small number of rather crude industry dummies.

An alternative solution to the errors-in-variables problem is the use of instrumental variables. Since measurement errors should be independent for variables constructed from disjoint sets of questions, *IML* is a valid instrument for *PPP*. The first three columns of Table 8 present the results of using *IML* and its interactions with the other variables as instruments for *PPP* and its interactions (we also use the industry dummies as instruments.) Again, this method produces sensible results for the simplest equation, but it makes little difference in the context of equations which contain R&D variables. The last two columns treat the patent stock variable as also being measured with error, using the R&D stock as an additional instrument. Now the patent variables have larger coefficients (perhaps unreasonably so) and are much more "significant", even in the presence of the R&D news variable. Since we do not have a good explanation for the shifting of these coefficients, we are not inclined to over-interpret these results.

Table 8 somewhere here

Conclusions

We tried to improve upon our estimates of the stock market's valuation of knowledge capital embodied in R&D and patents stocks by bringing in measures of the appropriability environment facing a firm from the Yale Survey. We found the responses to the questions about the effectiveness of patents as a mechanism for protecting the returns from innovation to be of some use. There is some evidence of an interaction between industry level measures of the effectiveness of patents and the market's valuation of a firm's past R&D and patenting performance, as well as its current R&D moves. There is no evidence, however, that other appropriability mechanisms differ enough across industries to leave measurable traces in such data.

Because the within industries variance of the Survey responses is so high, even for the somewhat better defined patents questions, our estimates are not very stable, and attempts to improve upon them using various errors-in-variables "de-attenuation" and instrumental variables methods were not particularly successful. Nevertheless, the estimated effects while not particularly precise, are not small. Table 9 presents the change in q implied by

Table 8

Estimation by Instrumental Variables

Dependent Variable: log(q)

	A	B	D	E	F
SP/A	0.901 (0.177)	0.272 (0.181)	0.400 (0.168)	3.934 (0.504)	2.328 (0.424)
PPP	-0.049 (0.030)	-0.042 (0.028)	-0.021 (0.026)	-0.069 (0.038)	-0.033 (0.029)
PPP*SP/A	0.466 (0.142)	0.167 (0.140)	0.197 (0.127)	0.929 (0.420)	0.658 (0.341)
K/A		1.416 (0.165)	0.681 (0.153)		
PPP*K/A		0.244 (0.174)			
NR/A			12.805 (1.176)		13.256 (1.340)
PPP*NR/A			3.170 (1.688)		1.496 (2.019)

Additonal Instruments :
Equation A: IML IML*SP/A
Equation B: IML IML*SP/A IML*K/A
Equation D: IML IML*SP/A IML*K/A IML*NR/A
Equation E: IML K/A IML*K/A
Equation F: IML K/A IML*K/A IML*NR/A
(all equations also have 10 2-digit industry dummies as instruments)
For variable defintions see Table 1 and Appendix B.

Table 9

Estimates of the Change in q implied by a 2 σ change in *PPP*

Equation	Δ q (%)
$\log(q) = f(PPP$, Patents stock, R&D stock, interactions) OLS, equation B1, Table 3a	20
$\log(q) = f(PPP$, Patents stock, R&D stock, news in R&D, interactions) OLS, equation D1, Table 3a	27
$\log(q) = f(PPP$, Patents stock, R&D stock, news in R&D, interactions) Restricted SUR, without dummies, Table 6	11
$\log(q) = f(PPP$, Patents stock, news in R&D, interactions) Instrumental variables, Table 8, equation F	13

Δ q is the percentage change in q implied by a 2 standard-deviation contrast in *PPP*.

some of our estimated equations for a two standard deviation contrast in the effectiveness of patents from their average level.

Table 9 somewhere here

The numbers range from 10 to over 25 percent, which is a rather large effect indeed. Given that R&D capital is about 14 percent of all other assets on average in our sample, this implies that such a change in the appropriability environment would come close to doubling its valuation.

The basic message of this paper is consistent with earlier work. There is some interesting information in patent counts, but it is subject to much error. Data on R&D expenditures, where available, are stronger measures of input to the process by which firms produce technical innovation than patents are of its 'output'. This difficulty with the patents numbers is not really eased by adding industry level information on the relative effectiveness of patents as a means of securing returns from innovation. But appropriability measures do appear to matter: we find significant interactions with either the patent stock or the R&D stock variables, implying that the market recognizes that similar R&D moves may have different payoffs in different appropriability environments. An alternative interpretation, which needs to be explored further, is that different appropriability environments imply different depreciation rates for R&D investment. These should have been incorporated in the construction of the R&D "capital" stock and the estimated interactions are our attempt to adjust for not having done so. We shall pursue some of these leads in our future work in this area.

Appendices

A. Sampling Error in the Appropriability Measures

Consider the following classical problem of errors in variables.

$$y = X\beta + \epsilon \tag{8}$$

$$\tilde{X} = X + V \tag{9}$$

where ϵ and the rows of V are i.i.d. with zero mean, and V follows some multivariate distribution with covariance matrix Σ

Substituting 9 into 8,

$$y = \tilde{X}\beta + (\epsilon - V\beta) \tag{10}$$

and multiplying through by $\tilde{X}$, we get

$$\tilde{X}'y = \tilde{X}'\tilde{X}\beta + \tilde{X}'(\epsilon - V\beta) \tag{11}$$

$$\tilde{X}'y = [\tilde{X}'\tilde{X} - \tilde{X}'V]\beta + \tilde{X}'e \tag{12}$$

Assuming that X, V, and ϵ are independent, that is

$$E[X'V] = E[X'\epsilon] = E[V'\epsilon] = 0 \tag{13}$$

it follows that

$$E[\tilde{X}'y] = X'y, \quad E[\tilde{X}'\tilde{X}] = X'X + V, \quad E[\tilde{X}'e] = 0 \tag{14}$$

and

$$E[\tilde{X}'V] = E[V'V] = n\Sigma \tag{15}$$

Which, using effectively the method of moments, gives the orthogonality condition

$$E[\tilde{X}'y - \tilde{X}'\tilde{X} - n\Sigma]\beta = 0 \tag{16}$$

If the covariance of the measurement errors, Σ, is known, or can be consistently estimated, then a consistent estimator for the parameters is just OLS applied to the cross-products matrix after correcting for the attenuation by subtracting $n\hat{\Sigma}$ or some fraction of it. (See Deaton (1985) and Fuller (1980)). Standard errors can be consistently estimated from the residuals, with an appropriate upwards adjustment in the case where Σ is not known with certainty.

The result of the data matching procedure described in the text is a design matrix of the form

$$X = [\, Z \,|\, X \,] \tag{17}$$

where the rows of Z are observations on the firm variables, all different, and the rows of X are made up of industry means of each variable from the Survey data, duplicated for all the firms in each industry. Our estimate of Σ is a weighted sum of the estimated Σ_i for each industry, where the weights depend on the number of Survey respondents for that industry within the Yale Survey and the number of firms within the industry in our sample.

Suppose we have a measure of appropriability, A, which has a different value for each industry A_i. The Survey data provides us with an estimate $\hat{A}_i$, the industry mean of the variable. $\hat{A}_i$ has sampling variance given by σ/n, which we can interpret as the variance of v in the equation

$$A_{ij} = A_i + v_{ij} \qquad j = 1, 2, \ldots, n_i \tag{18}$$

where j indexes individual respondents within industry i in the Yale Survey.

The error variance $\sigma_i^2 = E[v_{ij}^2]$ may be constant across industries, (homogeneous) or may vary across industries (heterogeneous). The calculation of the error variance of each appropriability measure is slightly different for each case.

Let n_i^S be the number of respondents in Yale Survey industry i, and n_i^F be the number of firms in our sample in industry i Consider the {k,k} element of Σ, the error variance of the k'th appropriability measure. Our estimate of its error variance within each industry in the Survey is $s_{ki}^2 = \hat{\sigma}_{ki}^2 / n_i^S$. After the data sets are matched, each industry mean appears n_i^F times in our sample. After forming the cross products matrix the {k,k} element of $V'V$ will be a weighted sum of the s_{ki}^2, where the weights reflect the size of the industry relative to the total number of observations in our sample.

$$\{V'V\}_{kk} = \sum_i \frac{n_i^F}{\sum_j n_j^F} \cdot \frac{\hat{\sigma}_i^2}{n_i^S} \tag{19}$$

Off-diagonal elements can be formed in a similar way using weighted sums of within-Survey-industry sampling covariances. [4]

[4] Using interaction variables in our regressions complicates the construction of $\hat{\Sigma}$. For example, suppose we have *PPP, SP/A, K/A, PPP*SP/A* and *PPP*K/A* as explanatory variables, where *PPP* is measured with error. $\hat{\Sigma}$ will have not only the diagonal term $\hat{\sigma}_{PPP}^2$ from *PPP*, but also off-diagonal terms of the form $\hat{\sigma}_{PPP}^2 \cdot \hat{\sigma}_{K/A}^2$, $\hat{\sigma}_{PPP}^2 \cdot \hat{\sigma}_{SP/A}^2$, and $\hat{\sigma}_{PPP}^2 \cdot \hat{\sigma}_{K/A,SP/A}$ from the cross-products among the variables.

The calculation is less difficult if we assume, alternatively, that the error variance is homogeneous:

$$\{V'V\}_{kk} = \hat{\sigma}_i^2 \sum_i \frac{n_i^F}{\sum_j n_j^F} \cdot \frac{1}{n_i^S} \tag{20}$$

where $\hat{\sigma}_i^2$ is estimated using a within-Survey total sum of squares after allowing for individual industry means.

Corresponding to the F-test for equality of means across industries, it is also possible to test for the equality of the within industry response variances (σ_i^2) using the statistic proposed by Bartlett (1937). This statistic was calculated for four of our appropriability variables, *PPP*, *IML*, *IMC*, and *MNP*, and except for *PPP*, the null hypothesis of equal within-industry variances was rejected at the one percent level. It is interesting to note that for variables which had significant variation in their mean across industries, we could not reject the hypothesis of homogeneous variance. On the other hand, where it is difficult to distinguish distinct industry means, variances appear to differ substantially, perhaps contributing to the difficulty in perceiving a consistent pattern of variation in the means.

Proceeding on the assumption of heterogeneous error variance, we estimated the diagonal elements of Σ with weighted sums of the within-Survey-industry sampling variances of our appropriability measures. Table A1 presents our estimates of $\hat{\Sigma}$.

For each matching level the estimated variance-covariance matrix of the errors in measuring each of the appropriability measures is given in correlation terms. The estimated error variance of each variable is presented in a separate column, together with the ratio of each variable's error variance to its total variance.

Table A1 somewhere here

The results of re-estimating our equations with an adjustment for attenuation are presented in Table 7. We also tried several other versions of these equations using our other appropriability measures, but had limited success: with a full adjustment for attenuation, the ratio of error variance to total variance is so high for our appropriability variables that for many of these regressions the adjusted cross-products matrix comes very close to being singular, introducing the potential for large numerical errors in the calculation of the estimates, and in some cases would not invert at all.

Table A1

Appropriability Measures: Estimated Error Variance

Matched by IND

	Correlation matrix: PPP	MNP	IMC	IML	Error Variance	as fraction of total (%)
PPP	1	0.0236	0.0369	0.4163	0.7102	41.5
MNP		1	-0.0455	-0.0941	0.0834	76.6
IMC			1	0.5744	0.5772	87.9
IML				1	0.0968	94.5

Matched by NSF

	Correlation matrix: PPP	MNP	IMC	IML	Error Variance	as fraction of total (%)
PPP	1	0.0503	0.1946	0.2842	0.2931	26.9
MNP		1	0.0016	0.1438	0.0376	65.3
IMC			1	0.5189	0.1672	42.1
IML				1	0.0298	46.7

We tried to avoid this problem by using Fuller's partial adjustment technique (see Fuller (1980), and Warren, White, and Fuller (1974)). This entails testing the smallest eigenvalue γ of the relevant submatrix against the critical value of $1+1/n$ and subtract either $(n-\alpha)\hat{\Sigma}$ or $((n\gamma-1)-\alpha)\hat{\Sigma}$ according to $\gamma < 1+1/n$ or $\gamma \geq 1+1/n$, which guarantees that the adjusted moment matrix will be positive definite. (Here α is an arbitrarily chosen constant introduced to lower the MSE, satisfying $k+1 \leq \alpha \leq k+4+2n/d$, where k is the number of regressors and d is the average number of degrees of freedom in estimating $\hat{\Sigma}$. We chose $\alpha = k+1$.)

In the case where Σ is unknown and unestimable, recourse is typically made to instrumental variable estimators, with well-known properties. Independently constructed appropriability measures may be valid instruments for each other if respondent errors are uncorrelated across questions, or if such correlations are attenuated enough by the within-industry averaging. The resulting IV estimates are presented in Table 8 in the text.

B. Definitions of Variables

There are two sources of data, the NBER RNDPANEL data set, which is fully documented in Cummins, *et al.* (1985), and the Yale Survey, which is documented in Levin, *et al.* (1984).

Variables Derived from the RNDPANEL:

V — Market value of the firm at the end of the year (EOY). This is calculated as the sum of the value of outstanding long term debt (converted to a standard maturity of 20 years using the *Survey of Current Business'* aggregate maturity structure for 1958 and each firm's history of net issues, and a matrix of bond prices from Moody's BAA Corporate Bond price series), plus the value of outstanding short term debt, plus the value of outstanding common stock (share price at EOY times number of shares outstanding at EOY), plus the value of outstanding preferred stock (reported preferred dividends paid divided by Moody's preferred dividend rate for medium risk companies), less the value of outstanding short term assets.

A — Total net tangible assets of the firm. It consists of the inflation-adjusted capital stock (net value of plant), plus inflation-adjusted inventories, plus investments in unconsolidated subsidiaries.

K — R&D stock, compiled from the series of annual R&D expenditures reported by Compustat, deflated by the RNDPANEL R&D Deflator series, using a 15 percent depreciation rate. Starting values were calculated on the assumption of an infinite history of previous growth in R&D expenditures at the same annual growth rate that obtained between 1972 and 1980. For up to two years of missing values between 1972 and 1980, the last non-missing year was carried forward until data became available again. For more than two years of missing data, annual R&D expenditures were set to zero, and the existing stock allowed to depreciate towards zero.

NR — The "news" in R&D, calculated as $NR = R\&D - 0.15K$. This is the current year's R&D expenditure less the depreciation on the stock of R&D carried forward from the end of the previous year, valued at current prices.

SP — The "stock" of patents held by the firm, compiled from annual patent count data supplied by OTAF using a 30 percent depreciation rate.

Starting values were calculated assuming an infinite history of previous growth at the annual rate which obtained for the period for which we have data. Missing data have been treated as zeros.

Variables From the Yale Survey:

PPP — "Patents Provide Protection" the sum of responses to questions IA1 and IB1, "do process/product patents prevent competitors from duplicating" the innovation.

NPP — The "average effectiveness of Non Patent Protection methods", the mean score on questions IA3–IA6 and IB3–IB6.

MNP — The "maximal NPP", the maximum score on questions IA3-IA6 and IB3-IB6.

FPP — The "Fraction of respondents in that industry who answer 6 or 7 to question IB1 (do product Patents provide Protection), less the fraction who answer 1 or 2.

IML — "IMitation Lag", the increment in imitation lag due to patents. This is calculated from questions IIE1–IIE4 and IIF1–IIF4: respondents were asked to estimate the time for a capable firm to effectively duplicate a competitor's innovation for the cases of "major"/"typical", patented/nonpatented and process or product innovations. Responses were on a discrete scale "Less than 6 months", "6 months to 1 year", "1 to 3 years" etc. We assigned a midpoint value to each of these ranges, (and 10 years for the extreme response of "Timely duplication not possible"), and calculated the increment in the reported imitation lag due to patents by subtracting "nonpatented" from "patented" scores in each of the eight cases (e.g. IIE1 (time to duplicate major patented process innovation) minus IIE3 (time to duplicate major nonpatented process innovation)), summing, and dividing by four to retain natural units. Thus $IML = ((IIE1 - IIE3) + (IIE2 - IIE4) + (IIF1 - IIF3) + (IIF2 - IIF4))/4$.

IMC — "IMitation Cost", the increment in imitation cost due to patents, as a fraction of the innovator's cost. This was calculated in the same way as *IML*, using instead questions IIC1–IIC4 and IID1–IID4. Here respondents were asked to report the cost for a capable firm to duplicate a competitor's innovation, on a scale running from "Less than

25 percent" to "Timely duplication not possible". Again we assigned a midpoint value to each of the ranges, and 300 percent to the extreme "not possible" response. Specifically, $IMC = ((IIE1 - IIE3) + (IIE2 - IIE4) + (IIF1 - IIF3) + (IIF2 - IIF4))/4$.

SRP, TRP, SRN, TRN are "stretched" and "trichotomized" versions of *PPP* and *NPP* respectively. "Stretching" is a non-linear transformation of the scale of responses where the original 1–7 scale is transformed into a 1–29 scale symmetric about the neutral response of 4, where each movement away from the center is double the previous move: 1 is twice as far from 2 as 2 is from 3, and 7 is twice as far from 6 as 6 is from 5. The actual stretching transformation is $1 \rightarrow 1$, $2 \rightarrow 9$, $3 \rightarrow 13$, $4 \rightarrow 15$, $5 \rightarrow 17$, $6 \rightarrow 21$, $7 \rightarrow 29$. "Trichotomizing" is a transformation where scores of 1 or 2 are assigned a value of -1, scores of 3, 4, or 5 are assigned a value of zero, scores of 6 or 7 are assigned a value of +1.

References

Bartlett, M. (1937), "Properties of sufficiency and statistical tests", *Proceedings of the Royal Society, Series A* 160, 268–282.

Cummins, C., B. Hall, S. Laderman, and J. Mundy (1985), "The R&D Masterfile: Documentation", NBER: unpublished.

Deaton, A. (1985), "Panel data from time series of cross-sections", *Journal of Econometrics* 30, 109–126.

Fuller, W. (1980), "Properties of some estimators for the errors-in-variables model", *Annals of Statistics* 8, 407–422.

Griliches, Z. (1981), "Market value, R&D and patents", *Economics Letters* 7, 183–187.

Griliches, Z., A. Pakes, and B. Hall (November 1986), "The value of patents as indicators of inventive activity", NBER Working Paper # 2083.

Levin, R., A. Klevorick, R. Nelson, and S. Winter (1984), "Survey research on R&D appropriability and technological opportunity", Research Paper, Yale University.

Levin, R., W. Cohen, and D. Mowery (1985), "R&D appropriability, opportunity, and market structure", *American Economic Review, Papers and Proceedings* 75(2), May, 20–25.

Schafer, D. (1986), "Combining information on measurement error in the errors-in-variables model", *Journal of the American Statistical Association* 81(393), March, 181–186.

Warren, R., J. White, and W. Fuller (1974), "An errors-in-variables analysis of managerial role performance", *Journal of the American Statistical Association* 69(348), December, 886–893.

Wildasin, D. (1984), "The q theory of investment with many capital goods", *American Economic Review* 74(1), March, 203–210.

Appendix C.

TABLE C1

Descriptive Statistics for firms by IND	N	Q		EMP		SALES $100m		ASSETS, $100m		STOCK OF RND / ASSETS		STOCK OF PATENTS / ASSETS	
		MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
1 MEAT PRODUCTS	9	.639	.231	3.71	2.38	4.72	3.85	1.31	1.21	.007	.013	.001	.004
14 SUGAR, CANDY, CHOCOLATE	7	.529	.135	4.95	4.14	5.73	5.30	3.53	3.22	.019	0.02	.007	.005
18 FATS & OILS	3	.746	.478	11.1	5.88	20.8	6.23	9.18	3.59	.032	0.02	.004	.003
21 BEER, LIQUOR, SOFT DRINKS	9	.838	.458	19.5	36.6	16.3	24.6	9.37	13.2	.004	0.01	.003	.004
24 LOGGING, SAWMILLS, LUMBER	21	.774	.501	7.58	13.5	7.15	13.9	7.80	16.8	.015	.028	.016	.038
26 PAPER, PAPER PRODUCTS	33	.583	.328	10.8	11.1	10.0	11.3	11.7	15.2	.037	.044	.028	0.04
45 DRUGS	19	1.55	.979	20.8	16.2	15.1	11.8	13.1	11.0	.334	.105	.181	.104
46 COSMETICS AND PERFUMES	13	1.18	.800	13.6	15.8	10.0	11.5	5.61	6.18	.156	.152	.098	.189
47 SOAPS, DETERGENTS, CLEANERS	9	1.44	1.40	10.1	19.4	14.6	35.0	9.26	23.5	.157	.143	.049	.071
48 PAINTS	7	.547	.207	7.97	10.2	6.01	6.98	3.18	3.94	.300	.136	.073	.092
49 INDUSTRIAL ORGANIC CHEMICALS	12	1.05	.992	29.1	38.5	30.2	37.7	33.2	50.7	.167	.083	.163	.167
53 PETROLEUM & REFINING	20	.982	.497	29.0	40.1	160	247	113	157	.018	.027	0.02	.034
54 TIRES, RUBBER, INNER TUBES	17	.665	.542	10.0	38.2	12.1	23.1	9.50	18.6	.141	.082	.096	.079
55 PLASTIC PRODUCTS	3	.515	.301	6.79	4.09	4.51	2.30	3.09	1.29	.149	.057	.163	.126
57 GLASS & GLASS PRODUCTS	10	.471	.191	17.9	18.6	11.8	13.2	11.6	14.3	.055	.076	.073	.071
58 CEMENT	6	0.49	.348	3.07	3.71	3.25	3.24	7.09	7.13	.003	.003	.000	.000
61 CONCRETE & GYPSUM	11	.434	.161	5.55	5.66	3.93	4.33	3.75	4.38	.016	.022	.021	.025
66 STEELWORKS, ROLLING & FINISH MILLS	24	.419	.256	10.7	19.0	9.56	15.4	11.7	20.4	.018	.036	0.01	.017
67 IRON & STEEL FOUNDRIES	5	.602	.363	5.96	9.65	4.86	8.29	4.53	8.14	.024	.037	.013	.017
84 SCREW MACHINE PRODS, BOLTS & NUTS	4	1.06	.852	2.77	2.97	1.78	1.76	1.32	1.25	.063	.056	.232	.282
90 VALVES PIPE FITTINGS, HEAT & PLUMB	7	1.48	1.26	5.23	6.38	4.02	5.25	2.96	4.32	.057	.035	.063	.043
93 FARM MACHINERY	5	.662	.144	31.1	40.3	24.9	31.2	21.2	29.2	.187	.049	.144	.127
97 OILFIELD MACHINERY & EQUIPMENT	5	2.55	1.11	14.9	7.54	11.8	7.17	10.8	6.72	.058	.027	.072	.045

(CONTINUED)

TABLE C1 (continued)

Descriptive Statistics for firms by IND	CU-SIP	Q		EMP		SALES $100m		ASSETS, $100m		STOCK OF RND / ASSETS		STOCK OF PATENTS / ASSETS	
	N	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
116 ELECTRONIC COMPUTING EQUIPMENT	23	1.70	1.40	42.2	74.7	26.3	55.8	22.1	49.5	.398	.351	.133	.116
131 COMMUNICATION EQUIPMENT	22	2.02	1.42	9.31	21.8	4.94	12.2	2.27	5.17	0.36	.254	.147	.233
143 SHIP & BOAT BUILDING & REP	5	.749	.229	20.1	36.1	11.3	20.3	5.34	8.78	.045	.061	.014	0.03
144 RAILROAD EQUIPMENT	3	.624	.182	23.9	20.3	15.4	9.80	12.9	6.95	0.04	.052	.103	.008
145 GUIDED MISSILES	2	1.06	.003	59.3	69.1	37.6	44.4	20.9	24.4	.184	.034	.343	.272
147 INSTRUMENTS	11	1.88	0.93	3.21	4.40	1.93	2.86	.972	1.28	.263	.139	.186	0.13
148 MEASURING & CONTROLLING DEV	17	1.41	.806	9.05	14.6	4.60	7.91	2.74	4.30	.337	.145	.164	.163
149 & 153 PHOTO, OPTICAL EQUIP, LENSES	12	1.25	.634	22.5	41.4	16.0	30.7	12.7	26.2	.395	.183	.581	1.04
150 & 151 DENT SURG & MED INSTR & APPL	14	1.00	.598	14.9	20.1	8.58	13.1	5.24	7.32	.219	.095	.136	.079
158 TOYS SPORT & ATHLETIC GOODS	8	0.46	.315	5.47	8.52	3.31	5.05	1.67	2.55	.223	.243	.107	.124
159 PENS, PENCILS, OFFICE MATERIALS	2	1.57	1.69	1.33	.099	.821	.266	.419	.039	.062	.057	.002	.002
200 PREPARED & PROCESSED FOODS	18	.718	.226	20.3	21.3	16.5	19.3	8.94	10.2	.043	.046	0.01	.015
204 GRAIN MILL PRODUCTS	5	.643	0.35	5.53	4.55	9.15	6.73	3.97	3.38	.028	.018	.007	.011
281 INDUSTRIAL INORGANIC CHEMICALS	15	1.02	.882	11.3	12.4	12.8	14.7	12.9	15.5	.116	.071	.093	.095
282 PLASTICS MATERIALS & SYNTHETICS	8	.943	.847	22.6	47.1	23.0	47.2	19.5	41.1	0.20	.143	.165	.297
287 AGRICULTURAL CHEMICALS	4	.699	.302	30.9	26.4	34.8	21.9	36.0	24.2	.093	.087	.065	.057
333 NON-FERROUS METALS	19	1.04	2.01	8.19	13.3	8.67	14.6	10.0	20.9	.046	.066	.077	.221
342 CUTLERY & HAND TOOLS	20	.808	.576	6.80	13.1	5.53	11.7	4.30	11.4	.069	.057	.058	.054
344 FABRICATED & STRUCT METAL PRODUCTS	43	.801	.855	3.32	3.86	2.33	2.68	1.32	1.80	.053	.071	.075	.118
351 ENGINES & TURBINES	6	.805	.625	20.8	14.3	14.4	9.69	8.28	5.65	.184	0.11	.148	.071
353 OTHER MACHINERY	12	.665	.167	11.2	22.8	10.4	24.0	8.77	21.1	.125	.072	.109	.099
354 MACHINE TOOLS & METALWKING MACH	11	.883	.467	5.35	5.11	3.15	3.01	2.41	2.42	.113	.084	.099	.101
355 SPECIAL INDUSTRIAL MACHINERY	26	1.01	1.03	4.04	6.75	2.61	4.21	1.79	3.31	.165	.189	.259	.382

(CONTINUED)

TABLE C1 (continued)

Descriptive Statistics for firms by IND	CU-SIP	Q		EMP		SALES $100m		ASSETS, $100m		STOCK OF RND / ASSETS		STOCK OF PATENTS / ASSETS	
	N	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
356 GENERAL INDUSTRIAL MACHINERY	34	1.32	1.19	5.17	9.03	3.17	5.56	2.11	4.02	.154	.118	.119	.115
361 ELECTRIC DISTRIB EQUIPMENT	5	.594	.225	88.6	176	55.3	109	39.5	78.6	.133	.076	.119	.042
362 MOTORS, GENERATORS, WELDING	6	1.08	.321	4.60	8.57	2.11	3.90	1.47	2.80	.289	.106	.197	.052
363 & 364 H-HOLD APPLIANCES & ELEC	18	.613	.357	10.6	17.6	5.99	8.65	3.10	4.25	.125	.099	.123	.118
365 RADIO & TV SETS, RECORDS & TAPES	6	.978	.459	26.7	53.0	15.8	31.8	8.19	17.0	.171	.188	.100	.155
367 ELECTRONIC COMPONENTS & ACC	35	1.63	1.33	6.62	15.7	2.88	7.06	1.48	3.19	.254	.187	.158	.127
369 STORAGE BATTERIES & AUTO ELEC	7	1.20	.926	4.53	4.97	2.57	2.62	1.83	2.16	.165	.169	.157	.148
371 MOTOR VEHICLES, PARTS & EQUIPMENT	32	0.50	.218	50.7	148	39.3	119	24.1	70.4	.124	.112	.051	.064
372 AIRCRAFT & PARTS	14	1.37	.716	27.2	35.5	19.5	29.0	10.3	15.2	.233	.184	0.17	.294
ALL	722	1.00	.937	14.9	42.9	15.1	57.3	11.4	38.7	.144	0.17	.105	.206

TABLE C2

Descriptive Statistics for Appropriability Variables by IND		PPP		IML (yrs)		IMC (pct)		NPP		MNP		FPP
	N	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	
1 MEAT PRODUCTS	3	7.33	3.21	.542	.938	.392	1.11	3.33	.938	5.67	.577	0.333
14 SUGAR, CANDY, CHOCOLATE	3	6.00	2.65	.333	1.16	.300	.829	4.58	.764	5.67	.577	-.667
18 FATS & OILS	7	7.71	2.21	1.09	3.33	.123	.705	5.12	.743	6.43	.535	.143
21 BEER, LIQUOR, SOFT DRINKS	3	4.67	1.15	.604	.355	.137	.606	4.79	.315	6.33	.577	-.667
24 LOGGING, SAWMILLS, LUMBER	7	6.71	2.69	1.12	1.69	-.24	.163	4.30	1.32	6.00	.577	-.429
26 PAPER, PAPER PRODUCTS	26	6.32	2.21	.965	2.23	.068	.577	4.89	1.02	6.16	.943	-.320
45 DRUGS	17	11.4	1.62	2.96	2.50	.819	1.03	4.38	.771	6.12	.857	1.000
46 COSMETICS AND PERFUMES	17	7.00	2.55	1.46	2.62	.374	1.01	5.11	1.08	6.29	.985	.059
47 SOAPS, DETERGENTS, CLEANERS	7	8.29	2.69	0.58	2.89	.046	.962	4.66	1.08	6.43	.787	.000
48 PAINTS	9	7.67	2.96	1.03	1.86	.295	.764	4.87	.933	6.44	.726	.111
49 INDUSTRIAL ORGANIC CHEMICALS	21	10.1	2.08	2.24	2.87	.906	1.21	4.57	1.05	5.89	1.00	.737
53 PETROLEUM & REFINING	10	9.33	2.00	-0.2	1.14	.217	1.14	4.03	.916	5.80	1.23	.222
54 TIRES, RUBBER, INNER TUBES	6	8.00	2.19	1.14	3.25	.354	.856	4.94	.552	6.33	.516	.167
55 PLASTIC PRODUCTS	29	8.17	2.51	.897	1.70	.159	.674	4.87	1.15	6.34	.897	.414
57 GLASS & GLASS PRODUCTS	3	9.67	1.53	1.44	.795	-.29	.636	4.79	.641	6.00	.000	.333
58 CEMENT	3	5.67	.577	1.81	1.23	-.06	.195	3.96	.887	6.67	.577	-.667
61 CONCRETE & GYPSUM	4	10.0	3.37	.484	1.34	.028	0.53	4.16	1.02	6.00	.816	.500
66 STEELWORKS, ROLLING & FINISH MILLS	10	8.60	3.50	1.29	1.91	.571	.812	4.79	.945	5.80	.789	.600
67 IRON & STEEL FOUNDRIES	7	6.86	2.97	1.00	1.25	.300	.637	4.02	1.27	5.43	1.27	-.286
84 SCREW MACHINE PRODS, BOLTS & NUTS	3	6.33	1.15	1.00	2.01	.096	.491	3.87	1.12	5.00	1.73	.000
90 VALVES PIPE FITTINGS, HEAT & PLUMB	10	7.11	1.62	1.95	1.90	.166	.325	4.21	1.48	5.90	1.60	.100
93 FARM MACHINERY	6	8.00	1.26	2.06	1.99	.396	0.77	4.54	.479	6.00	0.00	.167

(CONTINUED)

TABLE C2 (continued)

Descriptive Statistics for Appropriability Variables by IND		PPP		IML (yrs)		IMC (pct)		NPP		MNP		FPP
	N	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	
97 OILFIELD MACHINERY & EQUIPMENT	5	8.40	2.61	1.95	2.32	0.55	1.06	5.12	0.58	6.80	.447	.600
116 ELECTRONIC COMPUTING EQUIPMENT	22	6.68	3.15	.655	1.27	.049	.645	5.22	.783	6.68	.568	-.182
131 COMMUNICATION EQUIPMENT	17	6.75	2.11	.477	1.25	-.09	.367	4.90	1.42	6.29	1.16	-.118
143 SHIP & BOAT BUILDING & REP	4	5.00	3.16	3.13	4.78	.887	2.03	4.91	.717	6.00	.816	-.500
144 RAILROAD EQUIPMENT	4	6.50	1.73	1.47	2.19	.444	1.20	4.87	.577	6.25	.500	-.250
145 GUIDED MISSILES	5	4.40	1.14	.266	.896	.762	.688	4.90	.924	6.80	.447	-.400
147 INSTRUMENTS	9	7.37	2.07	.135	1.26	-.08	.293	4.56	1.08	6.44	.726	.000
148 MEASURING & CONTROLLING DEV	18	7.60	2.38	1.39	2.69	.262	.926	4.69	1.23	6.28	.958	.167
149 & 153 PHOTO, OPTICAL EQUIP, LENSES	4	6.00	.816	.328	.656	.344	1.04	5.37	.848	6.75	.500	-.250
150 & 151 DENT SURG & MED INSTR & APPL	12	8.27	2.15	1.39	1.77	.675	1.36	4.80	1.01	6.25	.754	.250
158 TOYS SPORT & ATHLETIC GOODS	5	8.20	2.17	.766	.704	0.13	.345	4.72	.271	6.20	.447	.200
159 PENS, PENCILS, OFFICE MATERIALS	4	7.50	1.91	.104	.662	-.36	.504	4.94	.298	6.50	.577	.000
200 PREPARED & PROCESSED FOODS	26	8.00	3.02	1.14	1.66	.214	.727	4.87	.884	6.23	.652	-.039
204 GRAIN MILL PRODUCTS	13	6.62	3.07	1.69	2.38	.300	.619	4.96	.969	6.31	.855	-.231
281 INDUSTRIAL INORGANIC CHEMICALS	20	9.60	2.37	2.04	2.67	.568	1.16	4.90	0.65	6.35	.587	.450
282 PLASTICS MATERIALS & SYNTHETICS	32	9.91	2.23	2.23	2.46	.452	1.01	4.88	.734	6.12	.833	.500
287 AGRICULTURAL CHEMICALS	11	10.4	2.84	1.55	3.38	1.08	1.16	4.36	1.42	6.00	1.00	.546
333 NON-FERROUS METALS	17	8.00	1.93	1.51	2.62	.361	.797	4.89	.914	6.41	.507	.063
342 CUTLERY & HAND TOOLS	3	8.33	2.89	3.73	4.31	-0.4	1.03	5.21	.832	6.33	.577	.333
344 FABRICATED & STRUCT METAL PRODUCTS	12	7.18	2.99	1.81	2.16	.623	.855	4.09	1.64	5.33	1.78	-.182
351 ENGINES & TURBINES	10	6.89	2.67	.607	2.16	-.23	.641	4.53	.971	6.10	.994	.100
353 OTHER MACHINERY	21	7.60	1.79	1.36	1.80	.046	.327	4.70	1.11	6.29	.784	.000

(CONTINUED)

TABLE C2 (continued)

Descriptive Statistics for Appropriability Variables by IND	N	PPP		IML (yrs)		IMC (pct)		NPP		MNP		FPP
		MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	
354 MACHINE TOOLS & METALWKING MACH	15	7.00	2.04	.264	1.24	-.04	.242	4.97	1.14	6.13	.834	.133
355 SPECIAL INDUSTRIAL MACHINERY	7	7.14	2.79	1.33	2.43	.462	1.14	4.66	.611	6.29	.756	-.286
356 GENERAL INDUSTRIAL MACHINERY	38	7.74	2.69	1.33	2.23	.137	0.74	4.74	1.18	6.11	1.29	.079
361 ELECTRIC DISTRIB EQUIPMENT	11	7.80	2.10	1.38	1.85	.292	.992	5.23	.831	6.73	.467	.000
362 MOTORS, GENERATORS, WELDING	13	6.50	2.78	2.11	2.34	.466	1.16	4.87	.724	6.15	.555	-.231
363 & 364 H-HOLD APPLIANCES & ELEC	3	5.00	2.00	.271	.469	.137	.413	4.83	.617	6.33	.577	-.333
365 RADIO & TV SETS, RECORDS & TAPES	5	6.20	2.39	-0.1	.863	-0.2	.275	4.70	2.20	6.00	1.73	-.600
367 ELECTRONIC COMPONENTS & ACC	12	7.42	1.93	.068	.235	-0.1	.221	5.12	.713	6.33	.778	.083
369 STORAGE BATTERIES & AUTO ELEC	3	7.33	1.15	.667	1.15	.325	0.46	5.33	1.08	6.67	.577	.667
371 MOTOR VEHICLES, PARTS & EQUIPMENT	24	8.37	3.52	1.42	2.16	.741	1.22	4.68	0.81	6.25	.897	.292
372 AIRCRAFT & PARTS	16	7.06	2.72	.746	1.39	.045	.575	5.12	.517	6.37	.619	-.125
ALL	632	7.85	2.73	1.28	2.15	.288	0.86	4.77	1.01	6.21	0.91	.056

Note: N may vary slightly across appropriability variables.

TABLE C3

Descriptive Statistics for firms by NSF		Q		EMP		SALES $100m		ASSETS, $100m		STOCK OF RND / ASSETS		STOCK OF PATENTS / ASSETS	
	N	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
NSF 1 FOOD & BEVERAGES PRODUCTS	51	.694	.299	13.1	20.8	12.4	16.3	6.45	8.76	.025	.034	.006	0.01
NSF 2 LUMBER, WOOD PRODUCTS, FURNITURE	21	.774	.501	7.58	13.5	7.15	13.9	7.80	16.8	.015	.028	.016	.038
NSF 3 PAPER AND ALLIED PRODUCTS	33	.583	.328	10.8	11.1	10.0	11.3	11.7	15.2	.037	.044	.028	0.04
NSF 4 INDUSTRIAL CHEMICALS	34	.968	.865	20.5	33.0	21.7	33.0	21.9	37.3	.152	.100	.134	.182
NSF 5 DRUGS	19	1.55	.979	20.8	16.2	15.1	11.8	13.1	11.0	.334	.105	.181	.104
NSF 6 OTHER CHEMICALS	34	1.10	.948	13.2	17.7	13.1	21.8	9.52	17.5	.179	.147	.076	.128
NSF 7 PETROLEUM REFINING	20	.982	.497	29.0	40.1	160	247	113	157	.018	.027	0.02	.034
NSF 8 RUBBER AND PLASTICS	20	.642	0.51	17.1	35.3	11.0	21.4	8.54	17.3	.142	.078	.106	.087
NSF 9 STONE, CLAY, GLASS	27	0.46	.215	9.59	13.4	6.68	9.26	7.40	10.0	.028	.052	.036	.054
NSF 10 FERROUS METALS & PRODUCTS	28	.438	.276	10.2	18.0	9.04	14.6	10.8	19.2	.019	.036	.011	.017
NSF 11 NON-FERROUS METALS & PRODUCTS	20	1.03	1.96	7.84	13.1	8.27	14.4	9.57	20.5	.044	.065	.074	.216
NSF 12 FABRICATED METAL PRODUCTS	74	0.88	.841	4.41	7.71	3.33	6.65	2.28	6.26	.058	.063	.078	.117
NSF 13 OFFICE COMPUTING ACCTING MACH	12	1.95	1.40	22.8	36.8	11.8	19.8	9.03	16.2	.434	.381	.114	.125
NSF 14 OTHER MACHINERY	110	1.14	1.05	13.9	37.2	9.71	28.3	7.82	25.1	.169	0.17	.153	.213
NSF 15 RADIO AND TV RECEIVERS	6	.978	.459	26.7	53.0	15.8	31.8	8.19	17.0	.171	.188	.100	.155
NSF 16 ELECTRONIC COMPONENTS	35	1.63	1.33	6.62	15.7	2.88	7.06	1.48	3.19	.254	.187	.158	.127
NSF 17 COMMUNICATION EQUIPMENT	22	2.02	1.42	9.31	21.8	4.94	12.2	2.27	5.17	0.36	.254	.147	.233
NSF 18 OTHER ELECTRICAL	36	.801	.546	19.3	67.1	11.5	41.4	7.64	29.8	.162	.125	.142	.109
NSF 19 MOTOR VEHICLES & PARTS	32	.500	.218	50.7	148	39.3	119	24.1	70.4	.124	.112	.051	.064
NSF 20 OTHER TRANSPORTATION EQUIPMENT	8	.702	.209	21.5	29.5	12.9	16.4	8.17	8.55	.043	.054	.047	.051
NSF 21 AIRCRAFT & MISSILES	16	1.33	.675	31.2	39.1	21.8	30.0	11.6	15.9	.226	.172	.192	.288
NSF 22 INSTRUMENTS & MEASURING DEVICES	28	1.60	.871	6.76	11.9	3.55	6.47	2.05	3.51	.308	.145	.173	.149
NSF 23 OPTICAL, SURGICAL MEDICAL EQUIP	14	1.00	.598	14.9	20.1	8.58	13.1	5.24	7.32	.219	.095	.136	.079
NSF 24 OTHER MANUFACTURING	22	.992	.746	14.4	31.7	10.0	23.4	7.58	19.8	.302	.224	.356	.795
ALL	722	1.00	.937	14.9	42.9	15.1	57.3	11.4	38.7	.144	0.17	.105	.206

Descriptive Statistics for Appropriability Variables by NSF	N	PPP MEAN	PPP STD	IML (yrs) MEAN	IML (yrs) STD	IMC (pct) MEAN	IMC (pct) STD	NPP MEAN	NPP STD	MNP MEAN	MNP STD	FPP
NSF 1 FOOD & BEVERAGES PRODUCTS	55	7.31	2.90	1.16	2.00	.232	0.69	4.82	.913	6.22	.686	-.109
NSF 2 LUMBER, WOOD PRODUCTS, FURNITURE	9	7.00	2.60	1.31	1.64	-.14	.327	4.44	1.20	6.11	.601	-.222
NSF 3 PAPER AND ALLIED PRODUCTS	26	6.32	2.21	.965	2.23	.068	.577	4.89	1.02	6.16	.943	-.320
NSF 4 INDUSTRIAL CHEMICALS	73	9.87	2.21	2.19	2.60	.611	1.11	4.80	.809	6.13	.861	.549
NSF 5 DRUGS	17	11.4	1.62	2.96	2.50	.819	1.03	4.38	.771	6.12	.857	1.000
NSF 6 OTHER CHEMICALS	45	8.18	2.93	1.35	2.72	.516	1.06	4.79	1.14	6.22	.951	.178
NSF 7 PETROLEUM REFINING	10	9.33	2.00	-0.2	1.14	.217	1.14	4.03	.916	5.80	1.23	.222
NSF 8 RUBBER AND PLASTICS	35	8.14	2.43	.939	1.99	.192	.698	4.89	1.07	6.34	.838	.371
NSF 9 STONE, CLAY, GLASS	20	8.95	2.68	.339	2.29	-.06	.777	4.66	.835	6.26	.562	.211
NSF 10 FERROUS METALS & PRODUCTS	17	7.88	3.31	1.18	1.65	0.47	.741	4.47	1.12	5.65	.996	.235
NSF 11 NON-FERROUS METALS & PRODUCTS	17	8.00	1.93	1.51	2.62	.361	.797	4.89	.914	6.41	.507	.063
NSF 12 FABRICATED METAL PRODUCTS	28	7.19	2.33	2.01	2.34	.264	.687	4.23	1.45	5.61	1.59	.000
NSF 13 OFFICE COMPUTING ACCTING MACH	22	6.68	3.15	.655	1.27	.049	.645	5.22	.783	6.68	.568	-.182
NSF 14 OTHER MACHINERY	102	7.54	2.34	1.21	2.03	.128	.691	4.75	1.05	6.19	.992	.078
NSF 15 RADIO AND TV RECEIVERS	5	6.20	2.39	-0.1	.863	-0.2	.275	4.70	2.20	6.00	1.73	-.600
NSF 16 ELECTRONIC COMPONENTS	12	7.42	1.93	.068	.235	-0.1	.221	5.12	.713	6.33	.778	.083
NSF 17 COMMUNICATION EQUIPMENT	17	6.75	2.11	.477	1.25	-.09	.367	4.90	1.42	6.29	1.16	-.118
NSF 18 OTHER ELECTRICAL	30	6.89	2.41	1.45	1.93	.346	.953	5.04	.776	6.43	.568	-.067
NSF 19 MOTOR VEHICLES & PARTS	24	8.37	3.52	1.42	2.16	.741	1.22	4.68	0.81	6.25	.897	.292
NSF 20 OTHER TRANSPORTATION EQUIPMENT	8	5.75	2.49	2.18	3.28	.634	1.47	4.89	.603	6.12	.641	-.375
NSF 21 AIRCRAFT & MISSILES	21	6.43	2.68	.645	1.30	.165	.636	5.07	.617	6.48	.602	-.190
NSF 22 INSTRUMENTS & MEASURING DEVICES	27	7.52	2.23	1.03	2.41	.146	.782	4.64	1.16	6.33	.877	.111
NSF 23 OPTICAL, SURGICAL MEDICAL EQUIP	16	7.67	2.13	1.09	1.59	0.58	1.25	4.95	.982	6.37	.719	.125
NSF 24 OTHER MANUFACTURING	9	7.89	1.96	.482	.721	-.01	.422	4.82	.287	6.33	.500	.111
ALL	645	7.87	2.73	1.27	2.18	.286	.866	4.77	1.01	6.21	.908	.056

Note: N may vary slightly across appropriability variables.