WORKING
PAPERS

USING LINKED PATENT AND R&D DATA TO MEASURE INTER-INDUSTRY TECHNOLOGY FLOWS

F. M. Scherer

WORKING PAPER NO. 57

March 1982

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

USING LINKED PATENT AND R&D DATA TO MEASURE INTER-INDUSTRY TECHNOLOGY FLOWS

F. M. Scherer

Northwestern University
Evanston, Illinois

Address until July 15, 1982:

Department of Economics
Northwestern University
Evanston, Illinois 60201

Running Head: Inter-Industry Technology Flows

This paper is based in whole or in part upon line of business data and has been authorized for release by the Commission. The paper has been reviewed to assure that it conforms to applicable statutes and confidentiality rules, and it has been certified as not identifying individual company line of business data.

USING LINKED PATENT AND R&D DATA TO MEASURE INTER-INDUSTRY TECHNOLOGY FLOWS

Revised Version

December 1981

F.M. Scherer*

1. Introduction

This paper discusses in some detail a methodology for linking patent and R&D data to construct a matrix of inter-industry technology flows through the U. S. economy. A somewhat aggregated (41 by 53) version of the matrix is presented, as are more detailed disaggregations of the row and column sums.

The motivation for developing these new data was straight-forward. During the 1970s the United States experienced a pronounced slump in the rate of productivity growth. One of many possible suggested causes was a slowdown in the emergence or absorption of new technology. New technology comes in significant measure from the research and development (R&D) activities of industrial enterprises. Beginning in the late 1960s, there was a deceleration in the growth of company-financed real (i.e., GNP deflator-adjusted) industrial R&D sufficiently large that, had growth trends continued, real 1979 outlays would have been roughly 50 percent higher than their measured values. The key questions remain, what quantitative links exist between R&D and productivity growth, and did the parameters of any such relationships shift between the 1960s and 1970s?

This is the sort of thing about which economists ought to know a considerable amount. Data on industrial R&D outlays have been collected under National Science Foundation auspices since 1953. However, serious obstacles have blocked the path to understanding.

For one, the NSF data leave a good deal to be desired. NSF's industry breakdowns are at a high level of aggregation. With one exception, R&D data are assigned to primary industries by the "whole company" method, which for multi-industry enterprises often leads to substantial misclassification of R&D in companies' secondary lines. NSF's newer and slightly more disaggregated "product field" statistics depart from the "whole company" approach, but the departures are unsystematic. The reporting instructions are confusing and virtually impossible to implement in decentralized companies, and it is evident from a 1975 survey that companies responded to the instructions inconsistently.[1]

Even more important is a fundamental conceptual problem. With the partial exception of the NSF product field data, all research and development spending surveys link R&D to an *industry of origin* -- usually the principal industry in which a surveyed enterprise operates. However, it has long been known (e.g., from McGraw-Hill surveys) that the bulk of industrial enterprises' R&D is oriented toward the creation and improvement of *new products* sold to others, as distinguished from R&D on new or improved production processes used internally within the performing company. The latter should in principle lead directly to productivity gains within the industry of R&D performance, assuming that the industry classification is correct. For product R&D, however, the linkage is much less clear. Both behavioral and measurement considerations lead us to believe that performing industries will secure at best only a modest fraction of the productivity benefits from their product R&D.[2]

On the behavioral side, an innovator will capture all the benefits from a productivity-enhancing new product only if it can engage in first-degree price discrimination. Under simple monopoly pricing, some of the benefits will necessarily be passed on to users. And when new product competition is vigorous, price competition may also break out, permitting innovators to retain only a small share of the superiority rents associated with their products.

There are also practical measurement problems. The first step in the compilation of productivity indices is estimating real output, usually by dividing some dollar measure of output by price deflators. If the price deflators were perfect hedonic price indices, innovators would be found to capture all or most of the productivity benefits resulting from their new products. Few price deflators meet this standard, however. More commonly, the actual prices of new products are linked into a price index at parity with the prices of older products, and the linking often occurs only after the new product has been on the market for a considerable time. One consequence is that price deflators severely underestimate product quality improvements, which in turn means that the measured output of product innovators is lower than it would be if hedonic price indices were used, so that measured productivity gains are not observed at the originating industry stage. (An exception may occur when, because of enhanced monopoly power, profit margins in the product-originating industry rise.) A further implication is that the productivity impact of new products is observed "downstream" at the buying and using industry stage, both because the prices measured for inputs used by buying industries do not reflect their superiority value and because (thanks to competition) the prices actually paid do not reflect that superiority value. Thus, to ascertain the productivity effects of new product R&D, one must trace the flow of technology from the industry in which a new product originated to the industry(ies) using the product.

The first to propose a solution to this set of problems was Jacob Schmookler (1966, Chapter VIII). He postulated a kind of input-ouput matrix of invention flows, in which the rows represented industries making an invention, the columns the sectors using inventions, and the diagonal elements process inventions. Row sums correspond more or less closely to the R&D data collected by NSF according to industries of origin. Column sums give the total amount of technology _used_ by an industry. With patent data, Schmookler was able to estimate column sums for a small sample of capital goods invention using industries, but his untimely death prevented him from progressing further toward the realization of a complete technology flows matrix.

Since then Nestor Terleckyj (1974, 1980) combined NSF survey data with conventional input-output statistics to estimate something like Schmooklerian matrix column elements as well as source data row sums. What is described in this paper is an effort to apply methods like those pioneered by Schmookler in estimating more disaggregated matrices at a higher level of precision.

Some substantive results, presented fully in other papers [Scherer (1981a) (1981b)], are summarized later. However, one finding deserves immediate attention. As expected, process R&D — that is, R&D devoted to improving a firm's own internal production processes -- was found to comprise only a small fraction of all company-financed industrial R&D: 26.2 percent when measured by a count of patents and 24.6 percent when measured by linked R&D expenditures data with adjustment for sample coverage. Most industrial R&D is indeed product-oriented. A fuller breakdown of the patent count by user category is as follows:

Process inventions		26.2%
Consumer good products only		7.4
Industrial capital good products		44.8
Subset used both as producer and consumer goods	7.8%	
Industrial material products		21.6
Subset used both as producer and consumer goods	8.7%	

2. The R&D Data

It seems clear that a full understanding of the impact of R&D on productivity growth requires one to go beyond mere industry of origin classifications and find out where the fruits of R&D are actually used. The starting point for such a venture should be R&D or other technological input data of good quality. "Good quality" here means at least three things: reasonable accuracy, recognizing the difficulties of measuring what is and what is not R&D; considerable disaggregation (especially since analyses of R&D-productivity links have proved sensitive to the degree of aggregation); and a correct matching of expenditures with industries. The third criterion, though obvious, deserves further attention. If, for example, as is standard National Science Foundation and McGraw-Hill survey practice, all R&D performed by Exxon is assigned to Exxon's primary industry category, substantial amounts of R&D occurring in the organic chemicals and resins, agricultural chemicals, synthetic rubber, office equipment, and communications equipment industries will be wrongly assigned to petroleum extraction and refining. This problem has grown increasingly severe as U.S. corporations have become more diversified. Already by 1972, before the most recent conglomerate merger wave, the average manufacturing corporation had 33 percent of its employment outside its primary (roughly three-digit) field (Scherer 1980a, pp. 76-77).

The data source that best satisfies these three criteria is the Federal Trade Commission's Line of Business survey. The first full survey, covering 443 large corporations,[3] was for the year 1974. It required reporting companies to break down their privately-supported and contract applied R&D outlays, among other financial items, into 262 manufacturing lines of business (LBs), usually defined at the three- or four-digit SIC level, and 14 nonmanufacturing categories. These 1974 Line of Business data were a principal basis for the work reported here.

They are not without problems. Perhaps most important, 1974 was the first year for which the survey was fully implemented. No survey can achieve perfect reporting, especially on the first iteration for an activity as difficult to measure as research and development. The data were therefore subjected before use to an extensive verification and correction process. Reported company R&D totals were compared to 10-K report R&D figures, individual 1974 line of business figures were compared against 1975 and 1976 reports, and a general check for significant omissions or peculiarities was made. Several classes of difficulties were discovered. First, R&D expenditure reporting was in some instances incomplete. The standard correction was to replace the 1974 figure with the comparable 1975 (or if need be, 1976) figure deflated by the ratio of 1974 to 1975 10-K R&D outlays. Second, some companies failed to distribute their R&D outlays over all relevant lines of business, instead lumping them together in a single (e.g., the largest) line or a few lines. Such problems were normally remedied by applying 1975 or 1976 distributional weights, although in a few cases, breakdowns were made on the basis of sales or (where some LBs were known to be more R&D-intensive than others) patents obtained by the various company LBs. Third, companies were asked to report basic research outlays in a single separate category for the whole corporation. There were quite clearly enormous qualitative differences among companies in what was categorized as basic research. Moreover, a few companies reported *all* their R&D outlays as basic research, and others reported all of their central corporate laboratory outlays, basic or applied, in the basic research category or in some other category such as "services." All outlays reported under basic research were spread by the author over the various lines of business in proportion to private applied R&D outlays. Special problems were resolved through other allocation methods, usually after consultation with company accountants. Fourth, companies were allowed to assign the costs and assets of LBs with sales of less than $10 million to a catch-all (99.99) reporting category. This option was exercised fairly frequently in connection with new endeavors that had high R&D outlays but low sales. The clearest cases were reclassified to their proper home industries, but less obvious or more complex cases had to be left as they were, and so the 99.99 category (included in "miscellaneous manufactures") entails misclassification and is unusually research-intensive. Fifth, companies were permitted to have a limited amount of "contamination" in their reports — i.e., activities that should ideally have been accounted for in a different LB, but whose segregation would have imposed appreciable accounting costs. The average level of contamination was on the order of four percent. (U.S. Federal Trade Commission, 1981, pp. 50-53.) For my analysis this poses problems whose solution will be described later. Finally, companies were to report only their domestic unregulated business activities; foreign operations were excluded. However, when domestic R&D expenditures supported manufacturing operations abroad, the R&D could be prorated between domestic and foreign branches, leading to some understatement of R&D relative to National Science Foundation definitions. Nothing could be done about this except to test its effect on the average number of patents received per million dollars of reported R&D expenditures. The elasticity of patenting with respect to the percent of total corporate sales occurring domestically was found to be -0.14, with a t-ratio of 2.34 (Scherer 1981a).

The total amount of company financed R&D reported by the 443 sample corporations after corrections was $10.64 billion, or 73 percent of the universe total in the National Science Foundation's 1974 R&D survey. Sample contract R&D outlays (mostly under federal government prime or subcontracts) amounted to $5.97 billion — also 73 percent of the NSF survey estimate.

These percentages are relatively high compared to the FTC sample's coverage of other financial variables such as sales (estimated to be roughly 54 percent of the total manufacturing sector universe) or assets, for which the coverage ratio was approximately 67 percent. (U.S. Federal Trade Commission, 1981, pp. 69-76.) Apparently, the FTC sampled relatively more heavily in R&D-intensive industries. Estimating exact coverage ratios is difficult because the FTC survey emphasized financial accounting variables whereas other universe figures are either heavily contaminated by the mixing of manufacturing and nonmanufacturing activities (e.g., the FTC's Quarterly Financial Report series and the

Internal Revenue Service Statistics of Income series) or have sales and assets (i.e., under Census reporting rules) defined on quite different bases.

Since individual industry coverage ratios were needed to implement my technology flows matrix concept, a different and rather unorthodox estimation technique was adopted. The basis of comparison was the set of four-digit industry value of shipments concentration ratios published in connection with the 1972 Census of Manufactures. For each industry, concentration curves were interpolated (or sometimes extrapolated) on log normal probability paper. Aggregations to the FTC Line of Business category level were carried out following the guidelines in Stigler (1963, pp. 206-211). One could then locate on the relevant concentration curve the maximum fraction of industry sales accounted for by the number of companies reporting in a given FTC line of business. This LB coverage estimate is biased upward to the extent that the company units reporting under the LB program are not uniformly the largest sellers in their industries. Downward biases intrude to the extent that the companies report as "contamination" with some other line of business sales that are reported to the Census Bureau in the correct industry category. The coverage ratios estimated in this way for manufacturing industries (which originated 95.1 percent of total sample R&D) ranged from .06 to .99. The value added-weighted mean coverage ratio was 0.61 — somewhat higher than the FTC's most closely comparable value of shipments coverage ratio estimate of 0.54.

The coverage ratios derived in this manner were used to inflate sample line of business R&D outlays and obtain whole-industry estimates. For company financed R&D, the sum of the inflated values across all lines of business is \$14.72 billion, which agrees quite closely with the 1974 NSF survey figure of \$14.65 billion. This suggests that measurement errors, sampling ratio estimation errors, sampling errors, and their intersection had on average no serious systematic bias. For contract R&D, there is an evident bias: the sum of the inflated estimates is \$6.77 billion, or 18 percent less than the NSF survey universe figure of \$8.22 billion.

3. Estimating Technology Flow Matrices

What has been described thus far is a procedure for getting R&D expenditure data organized by industry of origin. This is in principle what has been done in other surveys. The improvements consist mainly of considerably greater disaggregation and a more accurate match of expenditures to true origin industries.

Much more difficult and important steps were required to flow those originating industry outlays out to industries of use. The information needed to do so was obtained through a detailed analysis of invention patents. To begin with an overview, a sample of patents was drawn that matched as closely as possible the sample of companies on which R&D data by line of business were available. Each patent was inspected and coded as to industry (LB) of origin and industry(ies) in which use of the invention was anticipated. The industry of origin classifications were employed to link the patents to the lines of business in which corresponding R&D expenditures had been recorded. Each patent then became in effect a carrier of the average R&D expenditure per patent in its origin LB, transmitting by a fairly complicated algorithm those expenditures out to the coded using industries. Summed R&D outlays could then be collected for cells and columns of appropriately aggregated technology flow matrices.

The Patent Sample

The R&D data employed are for companies′ 1974 fiscal year, which is centered on the 1974 calendar year. U.S. and West German surveys suggest that the average lag between conception of an invention and filing a patent application is about nine months (Sanders 1962a and 1962b, p. 71; Grefermann et al., 1974, pp. 34-37). During the mid 1970s, the average period of patent pendency -- i.e., the lag between application and issue of a U.S. patent — was about 19 months. Thus, the total lag between invention, which is assumed to accompany R&D expenditure, and the issuance of a patent is estimated to be 9 + 19 = 28 months. The time span of the patent sample was therefore set for the ten-month period from June 1976 through March 1977, whose midpoint is lagged 28 months from June 30, 1974. Some timing error is inescapable here, since the distribution of patent application to issuance lags is skewed, with a few patents in the sample having been applied for as early as the 1940s. However, 92 percent of the sample patents had applications dated in the years 1974-76.

There is no simple consistent practice with respect to the names to which corporate patents are assigned. Some patents resulting from corporate R&D go only to the inventor, but this is now extremely rare in large corporations. Some patents issued to corporations are in fact acquired during their pendency from outside or spare-time inventors, but this too, our analysis suggested, also appears to be unusual. The principal company name matching problems come from mergers and the fact that many industrial patents are assigned not to the parent corporation but to some subsidiary. An extensive effort was made to identify patent-receiving subsidiaries. Mergers were identified through the Federal Trade Commission′s annual Statistical Report on Mergers and Acquisitions.

Several protocols were adopted to ensure that patents were in fact linked to the correct 1974 parent companies. See Scherer (1980b, p. 6). In cases where mergers following a parent company's 1974 fiscal year led to an undesirable scrambling of patents, company patent counsel were helpful in providing the needed unscrambling. Failure to have attended to these subsidiary and merger timing problems would have led to matching error rates on the order of 20 to 25 percent.

Some 75 companies were found to have obtained no patents during the 10 month sample period. For these a more extensive three-year sweep was made, yielding 69 additional patents accorded a weight of 10/36 each. This procedure imparts sampling bias, but a minor bias was considered acceptable in exchange for better coverage of low-patent industries. Unfortunately, there was no feasible means of identifying a universe and weights on the basis of which more efficient stratified sampling techniques could be applied.

Because the R&D expenditure data gathered were for U.S. operations only, patents whose inventor had a foreign address (or in the case of multiple inventors, all or most of whose addresses were foreign) were excluded from the sample.

Altogether, the final patent sample consisted of 15,112 patents counted with unit weights, or 15,062 patents when over-sampled company patents are fractionally weighted. After adjustment for foreign inventor exclusions, this was roughly 61 percent of all patents issued during the sample period to U.S. _industrial_ corporations (i.e., excluding universities, non-profit research institutes, patent management firms, retailers, public utility corporations, and the like). Of the 443 sample corporations, 397 were patent recipients. The most prolific assignee, General Electric, received 706 patents originally classified to 51 distinct lines of business.

Patent Classification

Once the patent sample was drawn, the printed specification of each patent was inspected individually by members of a team including an electrical engineering student, an organic chemistry major, a graduate management student with undergraduate honors in chemical engineering, and a "utility infielder" with a joint chemistry-economics major and a farming background. Mirroring the team's specialties, patents were pre-sorted into four groups: electrical inventions, organic chemical inventions, other chemical inventions, and everything else. The primary objective was to classify each patent according to industry of origin and industry(ies) of use. On the latter, up to three specific industries of use (including final consumption) could be identified, or the invention could be coded to either of two "general use" categories -- (1) use proportional to the origin industry's normal customer sales distribution (e.g., a machine tool invention); or (2) ubiquitous use throughout the industrial economy (e.g., a corporate jet aircraft invention).

Coding industries of use was for the most part the more straightforward and simpler task. U.S. law requires that inventions be useful to be patentable. Applicants therefore take some pains to point out in patent specifications what the actual or prospective uses of their inventions are. Instructions to classification team members emphasized the importance of coding uses to match as closely as possible the industrial locus where productivity impacts were most likely to occur. In cases of doubt, category (1) general use classification was favored. Of the 15,112 unit value patents in the final sample, 42 percent were classified to one specific non-consumer industry of use, 11 percent to two specific using industries, 6 percent to three specific using industries, 29 percent to category (1) general use, and 5 percent to category (2) ubiquitous use.

In coding industries of origin, it is not enough to say, e.g., "This is a petroleum refining invention, so that must be the relevant industry." See Scherer (1981c). A catalyst might come from an inorganic chemical maker, an anti-knock additive from the organic chemicals industry, or a process design from a company like UOP, whose home base is engineering services. Origin depends at least as much upon how the R&D-performing company is organized as upon function. Each classification team member was provided with a set of industry codes in which, according to published information, sample companies purported to operate, along with a qualitative description of the companies' product offerings. Even this, however, is not enough. The objective of the classification effort was not to identify industries of origin that were "correct" in some absolute sense, but to classify the patents in such a way that the origin industry codes corresponded with the LB codes in which enterprises chose to report the R&D expenditures that gave rise to the patents. Because of confidentiality restrictions, however, the structure of companies' LB reporting codes was not, and could not be, known in advance. This required in difficult cases a target-bracketing approach. As many as three industries of origin could be coded. In the original coding, 15.6 percent of the patents had two origin industries and 2.8 percent had three industries of origin. Uncertainty about company account organization was not, however, the only reason for multiple origin codes. Some inventions are genuinely joint: e.g., an aerospace company's metal fatigue testing system that can be used in either aircraft or missile assembly operations, or a fuel injection system microcircuit installed in either cars or trucks. Therefore, an additional set of codes was created to guide the ultimate patent - R&D dollar matching process. Inventions could be coded to be matched with a single preferred industry, and only if that match failed, with others; or a spread over multiple industries of origin could be specified to occur in equal parts; or the spread could be effected in proportion to matched LBs′ total R&D expenditures. Additional options existed to deal with problems of vertical integration, e.g., when it was expected that an electronic systems producer would report R&D concerning a semiconductor component production process under its systems LB code, even though the production (and hence the productivity impact) was likely to occur in a separate semiconductor plant.

Even after the classification team had acquired considerable expertise from on-the-job-training, 20 to 30 percent of the patents proved "hard" to classify. An important breakthrough in reducing that fraction was the discovery that from sources such as telephone books, annual reports, and a rich data base developed by Roger Schmenner at the Harvard-M.I.T. Urban Studies Center, one could tell what specific divisions or industry activities a company had at a geographic location. The company unit location in turn could often be inferred from the residence of the inventor, especially when there were multiple inventors with a similar patent specification address. All industry codings were double-checked by the author against abstracts in the Patent Office′s Official Gazette. In questionable cases, the entire specification was reviewed. Problem cases resistant to solution by these methods were resolved through telephone calls to company officials or the relevant inventors. In these and other ways, an attempt was made to enforce high standards of accuracy.

In addition to industries of origin and use, the individual patents were coded according to complexity (number of pages and claims), economic characterization (process vs. material vs. capital good vs. consumer good), technological characterization (system vs. device vs. circuit vs. composition of matter vs. chemical process), whether the invention originated under a federal government contract, and various other pieces of information. The federal contract invention coding proved to be unexpectedly difficult because, it was learned, contractors did not uniformly comply with the federal requirement that they include a notice of contract support in their patent specification, and the larger contracting agencies lacked complete records of their government-supported, contractor-owned inventions. Through an extensive effort, 325 contract inventions were identified, but it is believed that another 75 or so eluded the search. Since all were military-related, later adjustments could be made to minimize biases in estimating technology flow matrices.

A tape containing the original codings of the individual patents is available from the author on a cost reimbursement basis.

The Patent - R&D Link

With the main patent coding task completed, the patent tape was brought to the FTC's Line of Business program office in Washington, where the link to R&D data broken down into individual company lines of business commenced. At this point, the original list of 276 LB categories was condensed to 263, partly because it had proved impossible to make distinctions between certain origin industry categories (e.g., ethical and proprietary drugs, electric motors and motor controls, and storage vs. primary batteries) and partly to mirror industry consolidations made by the FTC for disclosure avoidance reasons. Following these consolidations, the number of individual company LBs to which a patent might be linked totalled 4,274. The average company broke its operations down into 9.65 LB categories.

After certain origin industry recodings to correct anticipated matching problems were made, the first link was executed. Among the 15,112 sample patents, there were matching problems on approximately 18 percent, including 1,101 patents on which no match at all was achieved and roughly 1,570 on which multiple origins had been coded, some but not all of which matched. Each patent with a partial or total matching problem was analyzed against company LB program submissions to determine the reason for the problem and to effect if appropriate a correct recoding. Extremely valuable in this effort were Schedule II of the FTC's LB reporting form, which broke down reporting LB sales to the five-digit product level of detail, and an appendix that gave the geographic location of every major establishment covered by a reporting LB. The principal reason for matching problems was that companies had not organized their LB reports according to our expectations. "Contaminated" reporting was one sub-reason. Another was that our salvo approach to questionable classifications had indeed both hit and missed the target. When all recodings were completed, there were 306 three-industry matches (as compared to 429 initial three-industry codings) and 1,619 two-industry matches (as compared to 2,359 codings originally). The remainder were single-industry matches. Altogether, 1,851 of the 4,274 reporting LBs had at least a finite fraction of a matched patent. Of the 3,003 individual company LBs that reported non-zero company financed R&D outlays, 1,691 had matched patents.

Because at low R&D expenditure levels (i.e., less than $1 million per year) the probability of patenting is finite but well below 1.0, the $732 million of private R&D in individual company LBs with no patents was spread proportionately over LBs <u>with</u> patents in the same industry before computing the average amount of private R&D associated with a patent. The average value of this inflation factor was 7.3 percent, although it ranged from zero to as much as 30 times the R&D of patent-receiving LBs within an industry.

For each individual company LB with patents, the average private R&D cost per patent, i.e., the quotient of inflated R&D divided by the weighted sum of matched patents, was computed. For each patent, the average cost of the patent was then tallied. When the patent had more than one matched industry of origin, the cost was a weighted average of the originating LBs' average costs, with the weights having either been prespecified to be equal for truly joint inventions or proportional to the matched originating LBs' total R&D outlays. Government contract invention patents were handled differently because it was known that not all such inventions had been identified. For them, the average contract R&D cost input was an industry-wide average, not an average within individual company LBs.

The final output of this matching effort consisted of two computer tapes, one organized by individual company line of business and one by individual patent. The patent tape contains for each patent all original input data plus matched LB codes, the weights assigned each matched LB, the average company-financed R&D expenditure underlying the patent (hereafter, ACP) and (when relevant) the average federal contract R&D expenditure underlying the patent (FACP). The company LB tape contains R&D expenditure totals, patent counts, average R&D costs per patent, and

12 weighted average values of characteristics (e.g., proportion of process patents, proportion of consumer goods patents, average patent length, etc.) of patents in that LB's portfolio. Since these tapes include individual company line of business information, they can only be accessed within the FTC Line of Business program office.

Technology Flow Matrix Estimation

The completed patent tape became a primary input into the computer programs creating technology flow matrices for the U.S. industrial economy. The essence of the problem was to take the R&D dollars (ACP or FACP) associated with a patent, inflate them by the reciprocal of the origin industry's sample coverage ratio, and then flow them through from industry(ies) of origin to industry(ies) of use, accumulating sums for each relevant cell.

The first substantive step was to retag patents by industry of origin. When the original coding procedure specified a preference for some single industry of origin, that preferred industry code was adopted, whether or not a match to LB reports had been achieved. In the absence of such a preference, multiple origin patents were divided among industries of origin in proportion to the weights determined through the earlier matching procedure. Patents originally coded as having probable vertical integration characteristics received special treatment. If the invention was a process and the vertical integration industry code differed from the industry code under which companies were expected to report their financial data, the invention was assigned to the industry of origin in which its actual use as a process was anticipated, whether or not an LB code match had been achieved. To have done otherwise would have generated process invention data inaccurate for purposes of analyzing the relationship between R&D and productivity growth.

After inflation to correct for differing origin industry coverage ratios, the inventions and their accompanying R&D dollars were flowed out to industries of use. For inventions coded as having a single industry of use or process inventions, this was quite simple. The R&D dollars went fully to the specific industry of use or, in the case of process inventions with multiple surviving origins, were divided among using industries in proportion to origin industry weights.

For inventions with multiple or general uses, the problem was more complex. Plainly, some using industries will use an invention more intensively than others. The question is, how does one determine the relative weights? The basic solution chosen was that inventions and their R&D would be flowed out to multiple using industries in proportion to the using industries' purchases from the origin industry. The natural basis for the needed "carrier matrix" $\underset{\sim}{A}$ was the 1972 input-output tables for the U.S. economy. However, substantial modifications had to be made before the input-output carrier matrices were consistent with our objective of tracing technology flows in such a way as to anlayze their productivity growth impact.

The starting point was the 496-order 1972 current transactions matrix recording the use of commodities by industries (U.S. Department of Commerce, 1979). This had to be aggregated down to the 263 x 286 industry level at which the most detailed technology flows estimates were to be prepared. Certain disaggregations also had to be made, usually on the basis of simple relative size weights. However, for the industrial gas, glass, and electron tube industries, new row vectors were estimated from primary data. Also, many input-output industries have large diagonal elements associated with inter-plant, intra-industry transfers. These might be viewed as surrogates for process inventions, but the correspondence is at best strained, and internal process inventions were in any event separately accounted for in our analysis. Therefore, the diagonals were "cleaned out" so that they did not exceed the row industry's proportionate share of aggregate output except in a few cases (such as organic chemicals) where productivity might plausibly have been affected by substantial intra-industry technologically advanced intermediate materials transfers. Corrections based upon primary data were also made to defense-oriented rows to reflect the fact that the sales pattern for products emerging from private R&D is different from the contract R&D pattern.

A more serious problem was posed by the input-output transactions matrix handling of capital goods-producing industries, which tend to be especially important R&D performers. Most of those industries' output is reported as sales to the "gross domestic fixed investment" column of final demand. This is obviously wrong in terms of identifying the industries in which capital goods technology is actually used. Basing technology use estimates on the small fraction of total output spread over using industries in the intermediate commodity output sector could be quite inaccurate. Therefore, the separately available capital flows input-output table for 1972[4] was integrated with the current transactions matrix -- something that, to the best of my knowledge, has not been done previously. The most detailed version of that table is available only in an 80 column (i.e., using sector) version, so the columns had to be disaggregated to 286 industries. For any capital flows matrix column spanning two or more of our industries, cell entries were split in proportion to the disaggregated industry's 1972-74 new capital investment as a fraction of the capital investment by all LB industries encompassed by the input-output column sector. Row aggregations to our level of detail were routine. Once a properly dimensioned capital flows matrix was available, the transactions and capital flows matrices were integrated. If T_{ic} is the capital formation element of the i^{th} row in the current transactions matrix and I_{ij} is a representative element of the capital flows matrix, a representative element T^*_{ij} of the revised transactions matrix was formed as $T^*_{ij} = T_{ij} + T_{ic}(I_{ij}/\sum_{i=1}^{285} I_{ij})$. This was done for 70 capital goods-producing industries with positive general use (category (1) or (2)) inventive activity.

Input-output conventions concerning the construction industry(ies) as a using industry posed similar problems. Substantial fractions of the output of the heating equipment, fabricated structural metal, office partitions, valves and pipe fittings, bathtub, and other industries are shown as used in the construction sector. It is true that construction is a large purchaser of such items, but it purchases them to install them for use by others. Inventions whose main utility lay in greater ease of installation by the building trades were specifically coded as having a construction industry use. Allowing the received input-output table structure to stand for general-use inventions would have inaccurately measured productivity-affecting technology flows. Consequently, output to construction industry sub-sectors was rerouted to "downstream" using sectors to the extent that the input-output table detail permitted. Where it did not, all or part of the remaining output originating from 29 industries and reported as used in construction was spread over all using industries in proportion to the industries' purchases of capital goods from the construction sector.

This problem has still another analogue. Consider the output of a technologically important component-producing industry such as semiconductors. According to the input-output tables, that industry's output flows to using industries like computer manufacturing, radio and television set production, and communications equipment. Yet who actually realizes the productivity-enhancing benefits of a more efficient large-scale integrated circuit: the computer maker who installs it in his newly-designed computer, or the university or bank or manufacturer who purchases (or leases) and uses the faster, higher-capacity computer? Some sharing of benefits may occur, but if the forces of competition are working with reasonable vigor or if deflators for new component-embodying systems are less than perfect hedonic price indices, one would expect much of the productivity-enhancing benefit from component product inventions to be passed on from the industry that assembles the components to the industries that buy and use the products embodying the improved components. To implement this notion, 22 industries that specialized in supplying components to some set of first-order using industries (usually assembly-type industries) were identified.[5] Relevant parts of component industries' output were subjected to a second-order (or for synthetic fibers, third-order) flow correction. Thus, let T^*_{ij} be the integrated first-order matrix sales of component origin industry i to assembly industry j. Then for any element k in industry j's output use row, the adjusted value is

$$T'_{jk} = T^*_{jk} + T^*_{ij} \; (T^*_{jk} \; / \sum_{k=1}^{285} T^*_{jk}),$$

with T^*_{ij} set equal to zero before second-order flows to row i (i]j) are computed. Because it was not clear a priori whether the benefits of component product inventions would in fact be passed on as measured productivity gains to second-order buyers, complete carrier matrices were calculated both with and without these component flow adjustments, and corresponding pairs of technology flow matrices were estimated. In fact, regression analyses of productivity growth revealed that technology flow variables <u>without</u> second-order component flow adjustments consistently had slightly greater explanatory power. See Scherer (1981b).

The result of these modifications was a set of input-output tables unlike any previously available, but suited as well as possible to performing the carrier matrix role in the estimation of technology flow matrices. All row elements were converted to ratios whose sum over all using sectors except end consumption equalled unity. After unneeded rows were purged and some further aggregation, a set of four carrier matrices $\underset{\sim}{A}$ -- one each with and without second-order component flows at the 263 x 285 and 263 x 56 levels of aggregation — was taken to the Federal Trade Commission to be linked with the patent data tape for the final technology flow matrix estimation stage.

For general-use inventions of category (1), the R&D cost of a patent ACP flowed through to using industry j (excluding the final consumption sector) from single origin industry i with coverage ratio c_i and carrier matrix coefficient a_{ij} was as a first approximation a_{ij} (ACP/c_i). For general-use inventions of category (2) (i.e., ubiquitous industrial use), a_{ij} ratios relating using industry value added to value added in all industries were applied. When there were three or fewer (e.g., M) specific industries of use, the coefficient for the k^{th} designated using industry was $a_{ik}/\sum_{k=1}^{M} a_{ik}$, except that when this value was less than 0.15, the coefficient was set equal to 0.15 and all the specifically designated large amounts of an origin industry's output should enjoy a larger technology flow than relatively small purchasers. One alternative considered and ultimately rejected was to multiply each ACP or FACP value by the numbers equivalent of the Herfindahl-Hirschman index for origin industry carrier matrix $\underset{\sim}{A}$ row elements before flowing out general-use invention values in proportion to the carrier matrix a_{ij} coefficients.[6] Instead, a suggestion made at an NBER workshop by Richard Levin was adopted. For any multiple industry of use invention, the using industry with the largest a_{ij} value was assigned a unit value and all other industries' a_{ij} coefficients were normalized to this value. That is, if the maximum a_{ij} is a_{im}, the coefficient for industry k would be a_{ik}/a_{im}, and so the R&D dollars distributed to that industry would be (a_{ik}/a_{im}) (ACP/c_i).

This convention, like the numbers equivalent approach, has the property of assigning greater weight to individual inventions, the larger the number of industries using the invention is and the more equal in size the using industries are. Ubiquitous use inventions in particular (with a numbers equivalent value of nearly 24) received far more weight in total than specific using industry inventions. Whether such weighting is appropriate cannot be determined on a priori grounds; the question is essentially an empirical one.

Another problem with the public goods approach is that, because R&D dollars are in effect double-counted, estimated R&D coefficients in regressions explaining productivity growth cannot be interpreted as steady-state returns on R&D investment. This is a significant disadvantage relative to the private goods approach, under which such rate of return inferences can (with appropriate caveats) be drawn.

using industry coefficients were renormalized to sum to unity. Although arbitrary, this convention ensured that specific-use industries received some of an invention's value even when input-output tables showed no relevant transactions between the origin and using industry pair.

The Public Goods Problem

Under the procedures described thus far, R&D dollars (or patents) were flowed out to using industries in such a way that the sum of the flows equalled the sum of the origin industry's R&D. An exception was made for final consumption goods uses, for which no productivity analysis could in any event have been contemplated. For any patent covering a consumer good, the final consumption sector column received the full R&D cost of that patent, whether or not there were also industrial uses. In effect, the consumer goods application of such inventions was treated as a public good not reducing the amount of R&D available for transmission to industrial sectors.

It can plausibly be argued that multi-use industrial inventions should also be handled as public goods, with use by industry k not reducing potential use by industry j. There are, however, both theoretical and practical difficulties in implementing such a public goods approach. It can be shown (Scherer 1981a) that as the number of using industries (i.e., the scope of the market) increases, firms will do more R&D and receive more patents, all else (such as the size of the average using industry) held equal. This increase in inventive activity may be channeled in either or both of two directions: perfecting a given narrow array of products, or increasing the variety of products geared to specialized demands of the diverse using markets. When product variety increases with rising market scope, particular inventions may be applicable in only a subset of the relevant using industries. This goes against the spirit of the public goods hypothesis. When R&D emphasizes perfecting a narrow array of products, other problems arise. If the same product is sold in many markets, it may pay to carry the product's development to a high state of refinement. For any single using market, considerable progression into the stage of diminishing marginal benefits is implied. In contrast, for single industry of use inventions, development is more apt to cease where the marginal benefit is high. This difference in marginal benefits per using industry is difficult to capture under a public goods approach.

If increased market scope led mainly to the perfection of a fixed range of products rather than increased product differentiation, one might also expect (because of increasing marginal invention costs) the R&D cost per patented invention to be greater, the broader the scope of the market. Crude tests of this increasing cost hypothesis failed to provide support. See Scherer (1981a). There was no significant evidence of systematically rising R&D cost per patent as the number of using industries increased from one to three and then from category (1) general use to category (2) ubiquitous use, all else equal.

Despite the possibility of increasing product differentiation and diminishing single-market marginal benefits as the number of using industries rises, an attempt was made to construct technology flow matrices under the assumption that multiple industry-of-use inventions were public goods. The conceptual problems were substantial. The very nature of public goods makes a certain amount of arbitrariness unavoidable. A basic guiding principle was that even though use by one industry should not detract from use by another, industries purchasing

Given the conceptual and practical difficulties faced in implementing a public goods approach to technology flow estimates, the question of which approach — public or private — to use in productivity analyses was left open. Some evidence will be presented be in a later section.

4. The Output

A principal end product of the effort described here is a set of technology flow matrices and vectors. Full matrices were constructed only at the 48 row by 57 column level of aggregation. These were estimated both for patents and company-financed R&D dollars under both the private and public goods assumptions, with and without adjustment for second-order component invention flows. For federal government contract R&D outlays, similar matrices were constructed only under the public goods assumption. Table 1 provides an example of a technology flows matrix for company-financed R&D expenditures. It is aggregated further to the 41 x 53 level, mainly to minimize confidential data problems. It is defined under the private goods assumption (except for final consumption) with adjustment for second-order component invention flows. The rows are industries of origin; the columns are industries of use; and the diagonal elements approximate internal process inventions (except for a few sectors like organic chemicals with extensive intra-industry intermediate product invention flows). All entries are in millions of dollars. Blank cells denote R&D flows of less than $50,000. Entries marked "d" had to be suppressed to comply with the FTC requirement that no underlying R&D data be disclosed for any group of fewer than four companies.

Examining row 3,4, we see that a majority of food and tobacco

Table 1. Technology Flows Matrix for Company Financed R&D

		ORIGIN R&D	Agriculture 1	Mining 2	Food products 3	Tobacco products 4	Textile products 5	Apparel products 6	Lumber & wood 7	Furniture 8	Paper products 9	Publishing 10	Inorganic chemicals 11	Organic chemicals 12	Synthetic resins etc. 13	Pharmaceuticals 14	Agri. chemicals 15	Other chem. products 16	Petroleum 17	Rubber & plastics 18	Leather products 19	Stone, clay, & glass 20	Ferrous metals 21	Nonferrous metals 22	ab. metal products 2
1	Agriculture & forestry	128.1	d																						
2	Mining, exc. petroleum	60.3		45.2											d			d		d					
3,4	Food & tobacco products	444.9	7.9		257.8	20.4	0.2				d	d		0.1	0.1	d		d	0.1	0.1		0.1	0.1		d
5	Textile mill products	179.3	0.8		0.2		128.4	8.1	0.3	0.8	0.5	0.2							0.1	21.6	0.9	0.1		0.1	d
6,19	Apparel and leather products	55.5	d	d	d	d	d	16.5	d	d	d	d		d				d	d	d		d	d	d	d
7	Lumber & wood products	72.6	0.4		0.2		0.2		64.2	0.6	0.1								0.1	0.1		0.3	0.2	0.2	0.
8	Furniture	51.1	0.1	0.1	0.3		0.1	0.1	0.3	7.8	0.2	0.2						0.1	0.1	0.1		0.1	0.2	0.1	0.
9	Paper mill products	202.3	6.0	0.1	25.8	0.1	1.2	0.3	0.4	0.4	86.4	3.1	0.1	0.1	1.0	0.7	0.1	1.0	1.5	1.2	0.2	0.9	0.1	0.1	0.
10	Printing & publishing	67.4	0.1		0.4		0.1				0.1	32.7			0.1	0.1		0.1	0.1	0.1		0.1	0.1		0.
11	Industrial inorganic chemicals	159.2	1.2	0.1	d	0.8	1.7	0.1	0.1		4.6	0.3	90.8	3.3	1.8	0.2	0.3	6.1	9.3	0.4		0.9	2.5	d	1.
12	Industrial organic chemicals	297.2	3.0	0.2	2.7	d	11.8	0.4	1.3	0.1	5.9	d	1.1	163.3	24.1	1.2	2.3	19.0	7.1	13.8	d	2.7	1.6	2.2	0.
13	Synthetic resins, fibers, rubber	601.6	1.6	0.5	2.8	d	32.0	11.8	1.1	0.8	13.9	0.5	d	0.8	70.8	d		23.9	2.3	169.1	1.7	6.4	0.9	10.0	2.
14	Pharmaceuticals	557.3	32.0		0.2						d					71.0		d							
15	Agricultural chemicals	186.7	142.8		d						d			0.1		d	34.2	d		d					
16	Paints, toiletries, explosives, and other chemical products	485.7	3.9	4.2	11.1	0.2	26.3	0.2	7.8	4.5	20.5	13.4	0.7	1.4	9.6	1.0		85.4	13.8	16.2	2.5	5.4	3.5	3.8	31.
17	Petroleum extraction & refining	380.3	4.0	0.7	1.1	0.1	0.3	0.4	0.9	0.1	1.0	0.3	0.2	0.3	0.3	0.1	0.1	0.8	312.7	0.7	0.1	0.6	0.6	0.4	1.
18	Rubber & plastic products	419.8	13.7	2.8	18.1	0.7	3.0	1.1	d	3.5	5.0	2.2	0.7	0.8	1.3	1.3	0.1	4.3	3.8	203.0	1.9	2.5	1.0	5.1	4.
20	Stone, clay, & glass products	265.0	0.9	0.1	d	0.1	d	0.1	2.4	1.0	d	0.3	0.1	0.2	0.1	d		1.3	d	d		155.3	11.3	0.3	2.
21	Ferrous metals	189.3	d	0.1	0.1				0.1	0.4				d		d	d		d	0.1		0.1	164.2	0.3	9.
22	Nonferrous metals	156.9	d	0.1	d		0.1	0.1	0.3	0.2	0.1	0.4	0.2	0.4		0.1		0.2	d	0.1	0.1	0.3	2.8	102.9	7.
23	Fabricated metal products	552.7	5.3	3.0	28.0	0.2	0.7	0.7	4.5	2.7	7.4	0.9	1.0	2.7	1.5	0.7	1.4	d	7.4	1.1	0.3	1.3	12.3	2.3	127.
24	Engines & turbines	282.2	10.4	3.2	2.2	0.3	0.6	0.6	1.1	0.3	2.5	0.9	0.4	1.0	0.5	0.2	0.3	0.5	6.5	0.6	0.1	1.1	2.0	1.1	1.
25	Farm machinery	199.3	165.4																						0.
26	Construction, mining, & materials handling equipment	351.2	2.0	65.9	4.6	0.3	1.2	0.3	7.3	0.3	1.3	0.6	0.3	0.8	0.7	0.2	0.2	0.3	17.7	0.4	0.2	5.8	3.4	1.3	1.
27	Metalworking machinery	121.5	5.4	0.2	0.4		0.1		2.8	0.5	d	0.1			0.1			d	0.1	1.4		0.9	14.5	7.5	15.
28	Other machinery	691.0	17.6	8.0	56.9	1.5	18.9	1.9	7.8	1.1	20.9	40.1	5.0	10.1	5.4	2.1	1.6	3.4	16.2	9.7	d	13.6	35.4	9.1	16.
29	Computers & office equipment	1153.0	2.9	2.4	15.9	0.6	9.4	3.4	1.6	1.7	10.1	13.4	1.8	5.0	4.8	6.2	1.3	8.6	20.6	8.1	1.6	8.5	19.0	5.3	10.
30	Industrial electrical equipment	205.9	0.3	1.6	1.3		1.1	0.2	0.5	0.3	2.5	0.3	0.7	1.6	0.6		0.5	0.3	4.0	0.5		d	8.2	4.9	5.
31	Household appliances	102.7	0.1		0.1		0.4	2.2	0.6										0.1						0.
32	Lamps, batteries, ignition, X-ray, and other electrical equipment	233.3	8.1	0.6	2.5	0.1	0.5	0.3	1.3	0.4	0.7	0.7	0.2	0.4	0.3	0.2	0.1	0.3	1.2	2.0	0.1	0.8	4.6	0.5	1.
33	Radio & communication equipment	1227.7	5.1	1.6	5.3	0.7	1.5	1.5	1.5	0.7	1.7	2.4	0.5	0.9	0.6	0.6	0.2	1.0	7.6	1.6	0.3	1.9	2.4	1.0	3.
34	Electronic components	594.9	0.5	0.2	0.7		0.6	0.2	0.2	0.1	0.4	0.7	0.1	0.3	0.2	0.2	0.1	0.3	0.9	0.4	0.1	0.6	0.7	0.2	0.
35	Motor vehicles & equipment	1518.0	78.0	10.6	25.4	1.1	5.0	3.6	17.1	3.9	5.4	10.0	0.9	2.2	0.9	1.7	0.6	2.3	16.7	2.6	1.2	11.4	4.2	2.6	14.
36	Aircraft	659.4	0.3		0.2				0.1		0.2	0.1		0.1					0.5	0.1		0.1	0.1	0.1	0.
37	Missiles, spacecraft, & ordnance	122.7			d									d		d			d						
38	Other transportation equipment	140.1		d										d			d		d				0.1		
39	Measuring & medical instruments, photo equipment, & timepieces	1036.4	6.3	2.1	8.1	0.8	3.3	1.5	2.0	0.9	7.2	20.7	2.4	5.3	2.8	4.2	1.3	4.3	13.5	4.0	0.4	5.1	6.6	3.1	5.
40	Miscellaneous manufactures	211.6	d	0.1	0.6	0.1	0.9	1.3	0.6	0.3	0.2	0.5	0.1	0.1	0.1	0.1		0.2	1.4	0.3	0.1	0.2	0.4	0.1	0.
41,42	Trade & finance & real estate	39.7																							
43	Transportation & public utilities	47.2																							
44,45	Construction & services, including R&D services	266.0	0.7	d	0.7	0.1	0.2	0.2	0.2	0.1	0.3	0.3	d	5.5	d	0.1	d		24.4	0.3		0.3	5.0	0.7	0.
	TOTAL R&D DOLLARS USED		561.8	157.3	493.4	29.8	250.8	57.2	131.1	33.7	206.0	147.7	108.8	207.6	332.1	95.3	45.7	180.3	496.5	470.0	16.6	232.0	307.8	166.1	270.

enditures

Engines & turbines 24	Farm machinery 25	Construction equip. 26	Metalworking mach'y. 27	Other machinery 28	Office equipment 29	Indust. electrical 30	Appliances 31	Other electrical 32	Communications equip. 33	Electronic components 34	Motor vehicles 35	Aircraft 36	Missiles & ordnance 37	Other trans. equip. 38	Instruments 39	Misc. manufactures 40	Trade 41	Finance, insurance, & real estate 42	Ground transportation 43A	Air transportation 43B	Water transportation 43C	Telecommunications 43D	Electric, gas, & sanitary utilities 43E	Construction 44	Medical services 45A	All other services 45B	General government 46	Defense & space 47	Final consumption 48	
																													d	1
																							d						0.0	2
											d						16.0	0.4	d	0.3		d	d	0.3	0.9	2.5	0.4	d	143.1	3,4
				0.2			0.1	0.1			0.6			d	0.2	0.3	0.8	1.0	0.2	0.3	0.1	0.1	0.1	0.8	0.7	d	0.4	0.2	18.2	5
		d		d				d	d		d				d	d	d	d	d	d		d	d	d	d	d	d	d	37.7	6,19
				0.1					0.1		0.1		0.1	0.2		0.3	0.6	0.4	0.2			0.1	0.4	0.3	0.2	1.5	0.4	0.2	2.8	7
		0.1		0.2	0.1		0.1	0.1	0.5	0.1	0.5	0.1		0.2	0.1	0.1	7.0	2.0	0.3	d		0.2	0.2	0.6	d	7.1	1.3	1.1	22.4	8
				0.3	0.1	0.1	0.2	d	0.2	0.1	0.3	0.1			0.9	0.6	14.7	1.0	0.5	0.3		0.2	0.3	1.4	2.4	4.0	1.6	0.4	73.3	9
				0.1					0.1		0.1	0.1			0.1		2.5	1.1	0.2	0.1		0.2	0.2	0.4	0.2	d	0.8	1.3	13.9	10
d	0.1	0.1	0.1	0.2	d			d	0.2	1.1	2.7	0.3	0.1	0.4	d	0.1	1.7	0.9	0.3	0.1		0.2	d	1.0	0.8	1.6	1.7	1.9	8.8	11
				0.4	d		0.1	0.6	d	2.0	0.3	0.1		0.1	d	0.3	1.6	0.7	0.3	0.1	d	0.2	0.8	0.6	2.9	2.3	1.5	9.9	14.1	12
				0.5	0.1	0.8	1.9	1.9	1.1	1.8	12.0	1.2	0.2	0.7	3.7	4.8	1.2	1.0	0.5	0.2	0.1	0.3	0.4	1.1	2.0	4.6	1.5	2.9	10.0	13
																									321.4	d	0.8	0.1	462.0	14
																	0.2	0.1	0.2				0.1	0.1	d	1.0	0.5	0.1	80.2	15
0.1	0.6	0.8	0.3	2.0	2.4	2.0	5.7	0.9	d	4.7	19.3	1.3	0.3	3.6	1.3	3.4	9.5	2.7	2.5	0.3	0.5	2.1	7.5	16.3	4.7	23.7	9.1	3.9	161.9	16
0.1	d	0.1	0.2	0.7	0.1	0.2		0.1	0.1	0.1	0.5	0.3	0.1	0.1	0.2	0.3	13.3	2.9	5.1	2.9	1.0	0.3	5.7	6.1	1.0	5.5	3.1	3.2	53.2	17
0.2	0.4	0.9	0.3	1.6	1.1	1.0	1.5	2.6	1.0	1.2	7.3	0.8	0.3	0.7	1.9	2.9	22.2	9.5	12.4	0.9	d	1.3	5.8	9.7	4.8	18.1	7.2	3.9	111.9	18
1.1	0.1	0.4	0.7	d	0.1	d	0.3	0.7	1.0	d	5.6	2.1	0.2	0.3	d	0.7	8.3	2.0	0.9	0.3	d	1.6	5.3	9.0	1.7	3.7	4.6	2.1	36.8	20
d	0.3	0.6	0.3	1.2	d	1.7	0.4	0.3	0.3	0.1	2.8	0.1	0.2	0.5	0.2	0.3	d		d		d		d	1.6		d		0.1	0.0	21
0.8	0.2	0.3	d	1.3	0.2	1.3	0.3	2.3	0.6	1.2	1.4	1.3	0.5	0.3	0.5	0.4	0.8	0.5	0.7			d	1.8	0.4	1.2	0.7	0.5	3.0	9.6	22
3.0	1.0	2.0	1.5	7.2	0.8	1.6	1.4	1.6	1.8	0.9	11.0	16.3	0.7	2.9	1.7	1.0	11.7	9.4	8.1	3.3	0.4	5.1	151.2	17.1	3.2	19.8	8.4	29.7	106.6	23
8.9	0.3	1.7	0.3	3.9	0.2	0.9	0.2	0.3	0.5	0.3	1.8	0.9	0.3	8.5	0.6	0.3	12.0	8.9	7.8	1.0	6.4	1.7	75.8	10.2	2.4	21.1	15.6	15.3	59.6	24
	d										0.1			0.1			d	5.9					0.1	2.4	0.3	4.6	0.5	0.1	33.1	25
1.0	0.5	16.0	0.4	2.0	0.7	0.4	0.4	0.5	1.0	0.9	5.7	0.2	0.1	2.4	0.4	0.3	10.2	1.6	3.9	2.3	0.2	0.4	1.2	154.4	4.7	3.5	8.1	10.5	4.0	26
1.0	1.3	2.2	d	7.4	0.8	1.8	0.7	2.0	1.3	1.1	12.2	3.2	0.8	0.7	3.9	1.1	2.3	0.1	0.3				d	6.6		6.6	1.6	0.9	12.2	27
1.1	0.8	1.6	2.7	64.9	1.5	1.3	8.0	2.5	1.8	3.3	16.2	1.5	0.9	2.7	2.4	2.1	64.0	11.5	8.4	3.5	1.2	2.3	33.0	20.5	7.8	64.3	12.1	26.8	111.6	28
5.5	3.5	7.3	2.9	12.9	110.5	4.2	1.4	3.3	16.5	6.6	17.1	20.5	3.8	3.5	16.8	3.7	97.5	176.4	20.9	4.3	1.5	25.8	26.1	2.3	19.8	224.2	75.6	75.2	45.0	29
0.7	0.6	2.9	2.2	15.6	2.6	34.4	3.8	1.7	0.7	0.8	1.4	0.4	0.2	2.4	1.6	0.8	2.1	1.8	2.5	0.3	0.4	0.5	52.0	1.8	13.6	7.9	5.2	4.7	25.4	30
							6.8	d	0.1		0.1			0.2			4.1	2.8	0.2			0.1	0.1	0.2	0.2	10.1	0.4	d	94.1	31
0.4	1.6	1.4	0.3	1.3	0.6	0.2	0.4	27.8	0.6	0.3	12.2	0.4	0.1	1.1	2.4	0.5	13.0	6.1	11.4	19.7	0.6	8.8	11.5	5.7	26.0	13.5	7.2	20.9	120.0	32
0.5	0.4	0.9	0.7	2.2	1.2	1.1	0.5	1.2	106.3	2.0	6.2	17.6	3.6	2.2	1.7	0.9	31.8	20.9	11.7	75.6	12.2	417.7	11.0	16.1	8.8	38.1	39.4	220.6	223.2	33
0.2	0.1	0.3	0.1	0.8	3.9	0.6	0.4	0.6	6.5	386.4	3.1	4.6	1.5	0.5	2.2	0.8	4.7	5.7	2.2	3.9	0.2	43.7	2.3	1.7	2.6	28.2	11.7	50.0	46.2	34
0.8	2.4	2.9	1.2	5.1	1.5	1.0	0.8	1.1	2.0	1.5	158.8	0.8	0.8	1.1	7.0	3.8	288.8	99.0	185.3	5.9	2.5	9.5	43.9	125.1	16.9	204.7	79.8	27.7	1345.9	35
				0.1							0.4	160.5	0.9	0.2			0.7	0.4	0.5	360.4		0.1	4.1	0.5	0.2	0.4	6.5	120.1	d	36
												d	22.0							0.3		d	d	d	d	d	d	93.2	d	37
											d			d			d	d	59.9	d	6.4	d			d	d	d	d	51.7	38
0.7	0.5	1.0	0.9	5.5	2.3	3.0	5.9	2.1	12.4	11.4	6.8	9.7	2.5	2.1	88.8	1.2	33.1	26.8	14.7	35.4	2.7	9.3	35.0	13.9	207.0	142.0	35.6	60.2	342.5	39
	0.1	0.1	0.1	0.3	0.1		0.1	0.1	0.1	0.1	0.5	0.1		1.2	0.4	77.5	3.9	2.3	0.7	d	d	d	0.6	1.6	1.7	29.0	4.3	1.8	86.9	40
																	(d)												0.0	41,42
																					d	d							0.[illegible]	43
		0.1	0.1	0.3	0.1			d	d	d	0.5	0.1		0.2	0.2	0.2	4.3	2.5	0.9	d	d	0.6	4.7	2.2	d	72.6	d	d	14.0	44,45
5.9	19.2	43.8	22.4	141.1	131.9	59.0	41.7	72.2	185.6	446.4	308.1	245.0	40.8	44.8	147.0	108.8	727.9	409.7	364.9	524.1	84.1	542.8	493.9	432.9	684.6	1000.3	378.1	842.2	4111.4	

products industry R&D is internal process-oriented, with most of the remainder flowing, not surprisingly, into final consumption or trade (i.e., restaurants and food stores). Reading down column 3, we see that the food products sector used appreciable amounts of R&D embodied in products purchased from the paper, miscellaneous chemicals (16), plastic products, fabricated metal products (e.g., containers), other machinery, office equipment, motor vehicle, and instruments industries. For food and tobacco products, the balance between R&D originated ($444.9 million) and R&D used ($493.4 + 29.8 = $523.2 million) is fairly even. This is not true for all sectors. At one extreme among manufacturing industries is the printing and publishing sector, which originated $67.4 million of R&D but used $147.7 million. At the other extreme is farm machinery, which originated $199.3 million but used only $19.2 million. Nonmanufacturing industries, as has been well known, originate very little R&D, but they use roughly half of the R&D originating in the manufacturing sector.

In Appendix A are presented more disaggregated industry R&D sums classified in three ways: by industry of origin, by industry of use with second-order component flows under the private goods assumption, and by industry of use with second-order component flows under the public goods assumption. The industry categories have been consolidated somewhat relative to the original source computations to avoid possible disclosure problems. Because nonmanufacturing industries perform so little R&D but are heavy users, a more detailed level of disaggregation is implemented on the use side of certain nonmanufacturing sectors.

Text Table 2 provides a matrix of the zero-order correlation coefficients between industry totals for some of the principal technology flow variables. Because of the asymmetry of origin vs. use disaggregation detail among nonmanufacturing industries, the correlations are for 247 manufacturing industries only. Note that the variables with and without second-order component flows are highly correlated: between USERD1 and USERD2, $r = 0.996$. There is more difference between the private and public goods measures; e.g., with component flows, $r = 0.877$.

Also included in Appendix A is a variable for each industry with origin data giving internal production process patents as a percent of total coverage ratio-inflated patents. Patents are the focus rather than R&D dollars because of disclosure limitations. The two, however, are fairly closely related. If PRD measures process R&D spending as a fraction (not percentage) of total origin industry spending and PP measures process patents as a fraction of total origin industry patents, the simple regression equation is:

$$\text{(R1)} \qquad PRD = .02 + \underset{(.026)}{.956}\ PP; \; r^2 = .855, \; SEE = .128.$$

Examining the individual data in Appendix A, one finds wide inter-industry differences in the degree of process patent orientation. But there are consistent and plausible similarities within like groups of industries. Thus, complex capital goods producers tend to be very product invention-oriented (i.e., with low process invention ratios), while producers of basic raw materials are process-oriented. It should be noted, however, that some of the process percentage values in the Appendix are computed from rather small numbers of patents, and so possibly substantial sampling errors may exist for the individual industry estimates.

Table 2

CORRELATION MATRIX: R&D SUMS, MANUFACTURING INDUSTRIES ONLY (N = 247)

	ORGPAT	ORGRD	USERD1	USERD2	USEPUB1	USEPUB2	USEPAT
ORGPAT	1.000	.724	.673	.690	.565	.591	.658
ORGRD		1.000	.608	.649	.681	.742	.422
USERD1			1.000	.996	.877	.847	.871
USERD2				1.000	.889	.871	.862
USEPUB1					1.000	.988	.754
USEPUB2						1.000	.715

Definitions:

ORGPAT: Coverage ratio-inflated count of patents by industry of origin.

ORGRD: Coverage ratio-inflated company-financed R&D outlays by industry of origin.

USERD1: R&D by industry of use, private goods assumption, with second-order component flows.

USERD2: R&D by industry of use, private goods assumption, without second-order component flows.

USEPUB1: R&D by industry of use, public goods assumption, with second-order component flows.

USEPUB2: R&D by industry of use, public goods assumption, without second-order component flows.

USEPAT: Coverage ratio-inflated patent count flowed to industries of use, private goods assumption, without second-order component flows.

There is another potential hazard in the process invention percentage estimates. They stem, as stated before, from detailed examination of 15,112 individual patents. It is generally believed that process inventions (used largely within the originating firm) are easier to keep secret than product inventions, and from this may follow a propensity for firms to patent relatively fewer process than product inventions, all else (such as the economic significance of the invention) held equal. See Scherer et al. (1959, pp. 153-154). If so, our process patent ratio estimates could have a systematic downward bias. When patents are linked to the privately-financed 1974 R&D dollars of the company LBs in which they originated, one finds that process inventions accounted for 24.6 percent of total coverage ratio-inflated sample R&D expenditures. There are two benchmarks against which this figure can be compared. Recent McGraw-Hill research and development expenditure surveys (1978, 1979) have asked inter alia what fraction of corporate respondents' R&D outlays involved process development and improvement. The universe estimates appear to be sensitive to survey response, varying since 1974 in the range of 17 to 24 percent. Second, the Strategic Planning Institute's PIMS data base contains among other things a breakdown of applied R&D expenditures between product and process categories. These estimates are made at the level of finely subdivided "businesses" within companies, and are likely therefore to be more accurate than the corporate aggregates estimated for McGraw-Hill surveys. The simple average process R&D share for some 948 businesses reporting in PIMS during the mid-1970s was 25.5 percent.[7] Thus, from comparison with available alternative benchmark data, there is no reason to believe that our process R&D share estimates are in fact seriously biased downward.

5. Productivity Relationships

Although the technology flow data described in this paper also provide new insight into a facet of American industry structure, the principal reason for compiling them was to permit a better-specified analysis of the links between R&D and productivity growth. The detailed results of that analysis are described elsewhere (Scherer 1981b). Here a brief overview must suffice.

Of three productivity data sets analyzed, we focus here on one following input-output industry definitions and published by the U.S. Bureau of Labor Statistics (March 1979) and supplemented by unpublished computer printouts. With 1974 R&D expenditures as the independent variable of central interest, the principal regression analyses examined annual labor productivity growth ΔLP (in percentage terms) over the peak-to-peak business cycle interval 1973-78. Productivity indices and corresponding gross capital stock change indices ΔK were available for a total of 87 industry groups, including nearly all of manufacturing plus agriculture, crude oil and gas, railroads, air transport, communications, and the electric-gas-sanitary utilities. Following a formulation developed by Terleckyj (1974, pp. 4-5), the industry R&D flow sums are divided by 1974 industry sales S.

As noted earlier, R&D outlays USERD2 linked to industries of use without second-order component flows had slightly greater explanatory power than the variable USERD1 <u>with</u> second-order flows. The simple correlation coefficients with ΔLP were 0.249 and 0.223 respectively. R&D flowed to industries of use under the public goods assumption had appreciably less explanatory power than under the private goods assumption; e.g., the zero-order productivity growth correlations were 0.160 for USEPUB1/S as compared to 0.223 for USERD1/S. A similar but even more pronounced disparity was found with other quite

60-67	Finance, insurance, real estate			409.8	3006.8
70-89	Services, including R&D services	238.0	4		
70	Hotels and motels			43.5	303.8
75	Auto repair services			116.5	1069.5
78	Motion picture production & exhibition			11.0	56.0
80	Medical, dental, and health services			686.8	1382.0
82	Educational services			147.8	693.6
	Other services			656.0	3557.2
	Government, except postal and defense			378.7	2244.7
	Defense and space operations			1206.3	2689.1
	Final consumption			4111.0	4111.0

---

*All R&D dollar figures are in millions of dollars.

differently measured industry productivity growth data sets. This implies either a lack of support for the public goods approach to technology flows measurement generally or deficiencies in the specific (and necessarily arbitrary) assumptions made to implement that approach.

A strong a priori hypothesis underlying this research was that R&D flowed through to industries of use would better "explain" productivity growth than R&D measured by industry of origin. Product R&D was expected to have especially little explanatory power. The support for this hypothesis with the BLS input-output data set was surprisingly equivocal. Where USERD1/S is the used R&D variable and PRODRD/S measures product R&D classified by industry of origin, the relevant full-sample multiple regression was:

$$\text{(R2)}\quad \Delta LP = -.14 + \underset{(.11)}{.35}\ \Delta K + \underset{(.144)}{.289}\ PRODRD/S + \underset{(.393)}{.742}\ USERD1/S;$$
$$R^2 = .193,\ N = 87;$$

with standard errors given in parentheses. Both R&D variables are significant at the .05 level, but product R&D has a slightly higher t-ratio (2.01 vs. 1.89).

The results were quite different when the industry sample was split into two mutually exclusive subsets, one for which the price deflators underlying the productivity indices were reasonably comprehensive in their industry product line coverage and another for which deflator coverage was skimpy. For the more comprehensive deflator subset, the hypothesis favoring used R&D is clearly supported:

$$\text{(R3)}\quad \Delta LP = -.16 + \underset{(.14)}{.40}\ \Delta K - \underset{(.337)}{.182}\ PRODRD/S + \underset{(.411)}{1.039}\ USERD1/S;$$
$$R^2 = .241,\ N = 51.$$

Used R&D is highly significant, product R&D negative but insignificant. For the subset based upon meager price deflators, the opposite pattern is observed:

$$(\text{R4})\quad \Delta LP = .08 + \underset{(.17)}{.31}\,\Delta K + \underset{(.205)}{.431}\ \text{PRODRD/S} + \underset{(.96)}{.096}\ \text{USERD1/S};$$

$$R^2 = .197,\ N = 36.$$

Since used R&D was also significant and product R&D insignificant in another quite different sample with well-measured productivity indices (see Scherer 1981b), it would appear that the superior performance of the product R&D variable in equations (R4) and (R2) is somehow associated with especially severe problems in measuring productivity growth. With somewhat less compelling support, one is inclined to conclude that the difficult task of tracing R&D flows to industries of use was indeed worthwhile.

6. Conclusions

I have described in some detail a methodology for estimating a technology flows matrix for the U.S. industrial economy. Many problems had to be overcome; there are undoubtedly appreciable errors of measurement; and the matrix is incomplete because it has no foreign, university, government laboratory, and individual inventor technology origin sectors. Yet from the standpoint of investigating the relationships between R&D and productivity growth, the data developed are surely much closer to what the relevant theory demands than anything previously available.

From regression equations (R2) and (R3) plus additional information, it can be ascertained that a two standard deviation increase in an industry's use of R&D was associated during the 1970s with an annual increase in labor productivity of 1.1 to 1.5 percentage points. Rates of return on investment in used R&D of from 74 to 104 percent are suggested. The magnitudes involved are important economically. I do not know how we can progress further toward understanding the impact of R&D on productivity growth without obtaining additional data similar to, but more accurate and comprehensive than, the R&D use data described here. Yet the thought of linking on an even larger scale patent to R&D data by the extremely labor-intensive methods used in my project is daunting, to say the least. A simpler and more accurate approach would be to have patent applicants provide the necessary information by filling out a form similar to the one used by my patent classification team. The marginal costs would be small, and the rewards in terms of improved information about the structure of technology flows and productivity growth could be substantial.

FOOTNOTES

[*] Professor of economics, Northwestern University. The underlying research was conducted under National Science Foundation grant PRA-7826526. This paper was drafted at the Max-Planck-Institut für ausländisches und internationales Patent-, Urheber- und Wettbewerbsrecht under a stipendium from the Max-Planck Gesellschaft. The author is also grateful to numerous research assistants, and especially to Chun-Yue Lai, Mary Gianos, Brett Spencer, and Pin Tai, who did most of the patent coding, and to Joe Cholka of the Federal Trade Commission, who provided indispensable computer systems assistance. Use is made of Line of Business data collected by the Federal Trade Commission. A review by FTC staff has determined that the figures presented here do not disclose individual company Line of Business data. The conclusions are the author's and not necessarily those of the FTC.

[1] U.S. Bureau of the Census (1975), question 30, reveals that 57 percent of the surveyed R&D expenditures were reported on an end product (i.e., line of business) basis, presumably contrary to instructions. Twenty-nine percent were reported (consistent with the instructions) in technological fields different from the end product fields. For the remaining 14 percent, the technological and end product fields were said to be identical.

[2] For surveys of the problems, see Griliches (1979) and Scherer (1979, pp. 200-204).

[3] The count of corporations and lines of business reported here does not agree exactly with official FTC figures because of slight differences in how both corporations and lines of business were consolidated.

[4] For a summary, see Coughlin et al. (1980). The detailed capital flows data were available on tape BEA IED 80-001. The detailed current transactions matrix is on tape BEA IED 79-005.

[5] The industries were weaving mills, fabric knitting mills, organic fibers, tires and tubes, rubber hose and belting, flat glass, pressed and blown glass, internal combustion engines, pumps, anti-friction bearings, compressors, speed changers and industrial drives, mechanical power transmission equipment, automotive carburetors etc., vehicular lighting equipment, electron tubes, cathode ray tubes, semiconductors, other electronic components, starter and traction batteries, aircraft engines, and buttons, zippers, etc. Not all elements in these industries' rows were subjected to second-order flows. Only those elements that were preponderantly of a "component sale to further assemblers" nature were so handled.

[6] Thus, with 285 industries of use, the numbers equivalent for origin industry i is $1/ \mathop{S}_{j=1}^{285} a_{ij}^{2}$, where a_{ij} is an element from the carrier matrix $\underline{A}_f$.

This numbers equivalent is in effect a purchasing industry dispersion index and, because there is a tradition of using such indices in industrial organization studies of pricing behavior, it has interest in its own right. For the 172 (out of 263) line of business categories on which complete capital flows, construction, and other corrections were made, the median numbers equivalent value without adjustment for second-order component flows was 8.8. The mean was 13.4. The highest values were for miscellaneous plastic products (71.3), paperboard containers (61.7), conveyors (48.2), industrial trucks (47.4), and metal-cutting machine tools (43.8). Twenty-one industries had values in the 1.00 - 1.99 range. The numbers equivalent for the total value added of all industries (i.e., the ubiquitous use carrier matrix row) was 23.9.

It should be noted that this analysis calls attention to what may be a serious problem in prior studies using purchasing industry dispersion indices. It is not clear what those studies do about the gross capital formation element in input-output transaction matrix rows. If it is included in the computation, there will usually be serious understatement of buyer dispersion for capital goods industries relative to what one would obtain integrating the transactions and capital flows matrices, as should be done. If it is excluded, actual sales patterns may be badly measured from intermediate output data alone.

[7] Because industries performing relatively little R&D tend to have relatively high process R&D ratios, the simple average of ratios for the 210 industries covered by Appendix A is 31.4. Relative to a weighted average, as our 24.6 percent figure is, the PIMS simple average could conceivably be similarly upward-biased.

REFERENCES

Coughlin, Peter E. et al. July 1980. New Structures and Equipment by Using Industries, 1972. Survey of Current Business 60: 45-54.

Grefermann, Klaus, Oppenländer, K.H., Peffgen, E., et al. 1974. Patentwesen and technischer Fortschritt. Göttingen: Schwartz.

Griliches, Zvi. Issues in Assessing the Contribution of Research and Development to Productivity Growth. Spring 1979. Bell Journal of Economics 10: 92-116.

McGraw-Hill Department of Economics. 23rd Annual McGraw-Hill Survey of Business′ Plans for Research and Development Expenditures, 1978-81 (May 1978).

---. 24th Annual McGraw-Hill Survey of Business′ Plans for Research and Development Expenditures, 1979-82 (May 1979).

Sanders, Barkev. Spring 1962. Speedy Entry of Patented Inventions into Commercial Use. Patent, Trademark, and Copyright Journal 6: 87-116.

---. 1962. Some Difficulties in Measuring Inventive Activity. In National Bureau of Economic Research Conference Report, The Rate and Direction of Inventive Activity. Princeton: Princeton University Press.

Scherer, F.M., et al. 1959. Patents and the Corporation. Revised ed. Boston: James Galvin and Associates.

Scherer, F.M. April 1979. The Causes and Consequences of Rising Industrial Concentration. Journal of Law and Economics 22: 191-208.

---. 1980. Industrial Market Structure and Economic Performance. Second ed. Chicago: Rand McNally.

---. 1980. Interim Progress Report: Research and Development, Patenting, and the Micro-Structure of Productivity Growth. Evanston: Northwestern University.

---. 1981. The Propensity To Patent. Working paper, Northwestern University.

---. 1981. Inter-Industry Technology Flows and Productivity Growth. Working paper, Northwestern University.

---. 1981. The Office of Technology Assessment and Forecast Concordance as a Means of Identifying Industry Technology Origins. Forthcoming in World Patent Information.

Schmookler, Jacob. 1966. Invention and Economic Growth. Cambridge: Harvard University Press.

Stigler, George J. 1963. Capital and Rates of Return in Manufacturing Industries. Princeton: Princeton University Press.

Terleckyj, Nestor. 1974. Effects of R&D on the Productivity Growth of Industries: An Exploratory Study. Washington: National Planning Association.

---. 1980. R&D and the U.S. Industrial Productivity in the 1970s. Paper presented at an International Institute of Management conference, Berlin, Germany.

U.S. Bureau of the Census. 1975. Findings on the Research and Development Response Analysis Project - 1975. Washington. D.C.: Mimeographed.

U.S. Bureau of Labor Statistics. March 1979. Time Series Data for Input-Output Industries: Output, Prices, and Employment. Bulletin 2018. Washington, D.C.: Government Printing Office.

U.S. Department of Commerce, Bureau of Economic Analysis. 1979. The Detailed Input-Output Structure of the U.S. Economy, 1972; Vol. I, The Use and Make of Commodities by Industries. Washington, D.C.: Government Printing Office.

U.S. Federal Trade Commission. September 1981. Statistical Report: Annual Line of Business Report, 1974. Washington, D.C.: Federal Trade Commission.

Appendix A

DETAILED INDUSTRY R&D DATA*

SIC Codes	Description	Origin R&D	Percent Process Patents	Used R&D: Private Goods	Used R&D: Public Goods
01-09	Agriculture, forestry, fisheries	128.2	25	556.2	1939.1
10	Metal mining	11.6	97	35.1	107.1
11,12	Coal mining	21.1	97	35.1	107.1
13	Oil and gas extraction	179.6	99	275.4	596.1
14	Nonmetallic mineral mining	27.6	77	49.0	177.4
15-17	Construction	28.0	0	435.8	2635.2
2011,13	Meat packing	31.5	86	57.8	173.4
2016-17	Poultry and egg processing	4.2	67	8.6	33.2
2026	Fluid milk	3.7	0	7.5	68.4
2021-24	Other dairy products	26.2	23	32.3	99.3
2032	Canned specialties	9.1	84	1919	71.0
2037	Frozen fruits and juices	8.5	100	12.7	43.7
2038	Frozen specialties	9.7	67	13.6	43.9
2033-35	Canned and dehydrated foods	40.6	83	51.3	135.1
2043	Cereal breakfast foods	12.5	62	11.8	26.0
2047	Pet foods	17.4	0	2.4	22.0

2048	Prepared feeds	39.3	60	33.9	63.7
2041,44,45	Flour and grain mill products	10.9	7	7.3	35.2
2046	Wet corn milling	16.6	54	13.3	30.2
2051	Bread and cakes	8.5	20	19.3	99.4
2052	Cookies and crackers	6.3	100	9.5	32.2
2061-63	Sugar	5.5	70	10.6	39.7
2065	Confectionary products	5.9	75	5.5	34.9
2066	Chocolate and cocoa products	2.0	87	2.0	10.4
2067	Chewing gum	6.8	13	0.8	5.9
2074-79	Fats and oils	20.0	65	15.8	58.9
2082-83	Beer and malt	12.1	83	36.1	128.3
2084-85	Wines and liquors	5.9	81	13.5	66.1
2086-87	Flavorings, syrups, and soft drinks	25.1	30	41.3	140.9
2095	Coffee	13.1	62	12.1	28.0
2091,92,97, 98,99	Miscellaneous food products	70.3	41	51.8	138.0
21	Tobacco products	31.5	43	29.8	91.4

221,222,223,226	Textile weaving and finishing#	179.3	65	250.8	457.7
2251,52	Hosiery#				
2253-59	Knitting mills#				
227	Carpets and rugs#				
228	Yarn and thread mills#				
229	Miscellaneous textiles#				
231,232	Men's clothing	15.4	90	16.4	74.5
233-238	Women's, children's, other clothing	13.0	50	17.3	95.2
239	Miscellaneous fabricated textiles	27.1	26	24.0	84.3
241,242	Logging and sawmills	56.8	100	92.0	296.4
243	Millwork and plywood	7.7	44	18.2	141.2
245	Wood buildings	0.0	0	12.5	80.3
244,249	Miscellaneous wood products	8.1	39	13.4	87.2
251	Household furniture	21.5	35	22.6	185.2
252	Office furniture	6.2	22	4.4	27.7
253,259	Miscellaneous furniture	12.3	6	3.3	30.0
254	Partitions and fixtures	11.0	0	3.3	37.8

#Provisionally combined because of Line of Business disclosure limitations.

261-263	Pulp, paper, and paperboard	45.8	73	108.7	340.8
2641-43	Bags, envelopes, and paper coating	39.5	24	30.9	120.1
2647	Sanitary paper products	74.1	26	35.6	65.0
2648	Stationery and tablets	0.7	100	1.3	8.5
2645,46,49	Converted paper products	5.1	39	4.6	24.4
265	Containers and boxes	32.6	18	18.9	109.0
266	Building paper and paperboard	4.5	48	4.5	16.7
271-273	Newspapers, periodicals, and books	27.4	19	56.8	260.3
274	Miscellaneous publishing	0.2	0	2.3	12.4
275	Commercial publishing	9.6	68	68.3	260.4
276	Business forms	4.5	39	5.7	24.3
277-279	Other printing and printing services##				
2813	Industrial gases	17.8	49	11.2	25.0
2816	Inorganic pigments	35.1	69	29.2	46.4
2812,19	Industrial inorganic chemicals	106.3	49	68.4	144.9
2821	Plastic materials and resins	289.6	35	132.7	258.0
2822	Synthetic rubber	97.8	50	50.2	78.9
2823-24	Organic fibers	215.9	59	151.2	255.7

##Provisionally suppressed because of Line of Business data limitations.

283	Drugs	557.3	12	96.6	244.0
2844	Toiletries	53.7	4	16.3	90.6
2841-43	Soap, detergents, polishes	96.5	13	51.0	181.6
2851	Paints	124.3	13	67.0	144.7
2861,65,69	Industrial organic chemicals	297.2	58	211.2	425.0
2873-75	Fertilizers	22.6	57	18.7	60.7
2879	Other agricultural chemicals	164.1	11	27.0	47.0
2892	Explosives	7.0	19	2.2	8.4
2891,93,95,99	Other miscellaneous chemicals	204.3	17	39.4	100.2
29	Petroleum products	201.3	64	202.8	489.2
301	Rubber tires and tubes	117.5	45	135.4	291.3
302-306	Other rubber products	33.7	35	47.2	132.1
3079	Miscellaneous plastic products	268.7	51	289.0	480.0
31	Leather goods	0.0	0	17.6	107.2
321	Flat glass	16.9	71	18.8	60.0
3221	Glass containers	24.9	87	28.7	70.4
3229,31	Other glass products	58.7	57	44.8	105.2

324	Hydraulic cement	7.3	33	16.9	69.2
325	Structural clay products	7.5	100	16.2	40.5
326	Porcelain and pottery products	42.3	59	28.7	53.9
3271-74,3281	Concrete and stone products	7.9	33	19.0	135.1
3275	Gypsum products	5.6	15	3.6	14.7
3291	Abrasive products	22.6	34	11.7	26.7
3292	Asbestos products	5.0	28	3.5	20.0
3296	Mineral wool	17.6	54	14.0	41.8
3293,95,97,99	Other mineral products	48.7	31	20.9	46.0
331	Steel mills	123.6	74	223.1	640.8
332	Iron and steel foundries	65.6	100	92.1	239.7
3331	Primary copper	14.4	95	17.6	34.6
3332	Primary lead	2.1	38	1.1	3.3
3333	Primary zinc	1.8	0	0.5	2.8
3334,53,54,55	Aluminum and aluminum products	56.9	64	51.3	137.5
3339	Other primary nonferrous metals	18.3	61	15.7	27.6
3341	Secondary nonferrous metals	6.1	70	8.3	29.6

3351,56	Nonferrous rolling and drawing	9.1	53	11.0	54.7
3357	Wire drawing and insulating	27.1	67	36.4	96.2
3361,62,69	Nonferrous foundries	0.3	67	8.9	54.8
3398-99	Miscellaneous primary metal products	20.8	65	14.8	33.3
341	Metal cans and barrels	54.9	39	46.1	158.7
3421	Cutlery	4.6	20	1.9	10.1
3423,25	Hand and edge tools	23.3	8	7.0	32.1
3429	Other hardware	13.5	0	7.1	69.7
343	Plumbing and heating ware	23.8	0	7.1	49.5
3441	Fabricated structural metal	2.7	33	8.8	74.4
3442	Metal doors, windows, etc.	3.7	0	3.8	36.0
3443	Nuclear reactors and fabricated plate	225.5	19	71.1	133.1
3444,46,48,49	Miscellaneous fab. metal work	52.1	11	21.3	148.2
345	Screw machine products	33.6	6	12.9	66.0
3462-63	Metal forgings	4.0	100	6.7	26.3

3465	Automotive stampings	8.2	12	13.8	101.3
3466	Crowns and closures	4.2	12	2.8	11.7
3469	Other metal stampings	2.7	17	8.5	66.5
3471,79	Metal plating and coating	44.4	24	24.2	117.6
348	Ordnance	22.6	10	6.6	44.1
3494	Valves and pipe fittings	40.2	0	7.8	66.8
349 x 3494	Other fabricated metal products	11.3	5	11.6	100.4
3511	Turbines and turbogenerators	117.2	6	15.2	56.9
3519	Internal combustion engines	165.0	10	41.1	106.1
3523	Farm machinery	172.8	2	15.3	97.0
3524	Lawn and garden equipment	26.4	0	3.9	31.2
3531	Construction machinery	242.4	0	26.3	129.2
3532	Mining machinery	22.9	0	1.9	14.7
3533	Oil field equipment	15.6	2	4.1	32.7
3534	Elevators and escalators	13.5	0	1.3	11.7
3535	Conveyors	12.3	0	2.1	16.6
3536	Hoists and cranes	4.5	0	1.5	10.4

3537	Industrial trucks	40.2	7	6.7	36.2
3541	Metal-cutting machine tools	30.9	0	2.6	22.1
3545	Machine tool accessories	8.2	0	2.9	21.6
3546	Power driven machine tools	20.6	0	2.2	18.5
3542,44,47,49	Other metal-working machinery	61.8	3	14.6	123.3
3551	Food products machinery	56.8	2	3.5	18.9
3552	Textile machinery	11.2	0	1.8	17.0
3553	Woodworking machinery	1.1	0	2.4	23.4
3554	Paper industries machinery	10.8	0	1.2	9.2
3555	Printing trades machinery	40.9	4	2.3	15.0
3559	Special industry machinery	40.1	1	7.4	56.5
3561	Pumps	38.3	0	7.4	54.0
3562	Anti-friction bearings	30.7	30	13.1	42.1
3563	Compressors	29.4	0	2.5	20.0
3564	Blowers and industrial fans	12.6	0	2.1	16.6
3566	Speed changers and drives	12.5	7	1.6	14.2
3567	Industrial furnaces and ovens	27.5	0	1.2	7.4

3568	Mechanical power transmission equipment	24.4	0	1.9	17.6
3569,69	General industrial machinery	70.4	10	9.9	34.7
3573	Computers and peripheral equipment	1027.3	6	115.9	307.6
3574	Calculators and accounting machines	51.1	4	11.8	50.4
3576	Scales and balances	5.3	0	0.5	5.3
3572,79	Other office machines (incl. typewriters)	69.3	0	5.1	39.5
3585	Refrigeration and heating equipment	110.1	10	31.9	188.2
3581,82,86,89	Service industry machinery	65.1	0	5.8	41.3
3592,99	Miscellaneous machinery	110.0	4	44.5	134.8
3612	Electrical transformers	23.6	11	7.8	34.8
3613	Switchgear	42.5	3	7.3	46.0
3621-22	Motors, generators, controls	64.4	15	21.7	101.6
3623	Electric welding apparatus	9.1	10	5.3	22.6
3624	Carbon and graphite products	13.5	48	7.7	19.3
3629	Other electrical industrial equip.	52.9	20	9.6	19.1
3631	Household cooking equipment	18.1	3	6.4	33.7
3632	Household refrigerators & freezers	16.8	11	13.1	69.7
3633	Household laundry equipment	13.1	5	8.7	48.1
3634,35,36,39	Other household appliances	54.8	3	16.1	82.9

364	Light bulbs and fixtures	67.4	12	23.6	118.7
3651	Radio, TV, and high fidelity sets	141.3	3	13.9	59.9
3652	Records and tapes	4.5	67	3.9	11.8
366	Telephone and communications equipment	1082.0	7	192.8	415.1
3671,73	Receiving and transmitting tubes	33.6	19	20.3	30.1
3672	Cathode ray picture tubes	34.7	47	19.6	29.3
3674	Semiconductors	436.3	50	329.2	410.6
3675-79	Other electronic components	90.3	43	71.0	159.4
3691,92	Batteries	86.5	22	13.9	35.6
3694	Engine electrical equipment	34.5	10	18.3	47.3
3693,99	Miscellaneous electrical equipment	44.9	2	2.1	14.1
3711	Passenger cars	1263.8	5	157.9	823.0
3713,95	Trucks, buses, and combat vehicles	139.1	2	49.1	377.1
3714	Motor vehicle parts	109.4	10	99.7	403.0
3715	Truck trailers	5.7	21	4.3	22.6
3721,28	Aircraft	306.6	22	137.1	361.1
3724	Aircraft engines	352.8	23	107.8	171.6

373	Ship and boat building	22.2	0	26.6	168.2
374	Railroad equipment	66.3	1	11.0	60.0
376	Guided missiles and spacecraft	100.1	16	34.2	97.5
3792	Travel trailers and campers	31.6	0	4.3	28.5
3751,99	Motorcycles, bicycles, and miscellaneous transportation equipment	20.1	0	3.1	27.7
3811	Scientific & engineering instruments	135.1	2	5.1	27.9
382	Measuring and controlling devices	159.6	3	19.2	111.8
383	Optical instruments	66.0	16	12.4	23.7
3843	Dental equipment	25.6	0	1.6	13.3
3841-42	Surgical equipment and supplies	117.5	7	16.8	78.2
3851	Opthalmic goods	14.4	22	6.4	17.1
3861 part	Photocopying equipment	177.9	9	26.9	79.9
3861 part	Other photo equipment and supplies	319.0	8	53.2	152.1
3873	Watches and clocks	21.4	2	3.4	17.3
3931	Musical instruments	3.7	0	1.4	10.2
3949	Sporting and athletic goods	18.0	19	14.9	52.1
3942,44	Dolls, games, and toys	45.5	8	9.9	69.8

395	Pens, pencils, office supplies	10.4	55	10.5	32.2
3911,14,15, 396	Jewelry and silverware	4.8	22	4.6	40.6
399,catchall	Miscellaneous manufactures (incl. FTC code 99.99)	129.2	35	67.6	157.0
40-49	Transportation and public utilities	47.2	97		
40	Railroads			101.5	579.3
41	Suburban transit			70.0	351.8
42	Trucking			177.8	1684.2
44	Water transportation			38.8	136.9
45	Air transportation			524.3	1253.6
46	Petroleum pipelines			9.4	44.8
49	Electric, gas, sanitary utilities			497.7	1638.5
43,47	Other transportation and utilities			15.5	117.7
50-67	Trade, finance, insurance, real estate	39.7	100		
54	Retail food stores			96.1	692.7
55	New car dealers and gas stations			102.1	859.8
58	Eating and drinking places			82.2	361.1
52,53,56, 57,59	Other retail trades			226.6	1874.1