THE MANUFACTURING SECTOR MASTER FILE: 1959-1987

Brownyn H. Hall

THE MANUFACTURING SECTOR MASTER FILE: 1959-1987

ABSTRACT

    This document describes the panel of publicly traded United
States manufacturing firms which was created and updated at the
National Bureau of Economic Research from 1978 through 1990 within
the Productivity Program. The panel consists of 2726 large
manufacturing firms with one to twenty-nine years of data each; the
period covered by the sampling frame was 1976 through 1987, with
data back to 1959 where possible. There are about 90 variables for
each firm-year of data: the variables give the complete income
statement, balance sheet, statement of changes, and data on the
market value of the common stock.

    The firms on the file are identified both by their CUSIP
number and by name, making it feasible to match this data to other
sources. A special feature of this data file is that all exits
from the file between 1976 and 1987 have been identified and the
reasons for exit have been tabulated in a diskette file. This file
is described in Appendix A of this document.

Brownyn H. Hall
National Bureau of Economic Research
204 Junipero Serra Blvd
Stanford, CA 94305

TABLE OF CONTENTS

1

1. Introduction

The Manufacturing Sector Master File is based on the Compustat Annual Industrial and Over-the-Counter Files for 1978 through 1987. It consists of 90 variables per year for each of approximately 2700 firms in the manufacturing sector (SIC codes 2000 through 3999) for as many years as data are available. The firms in question must have been on the files and in existence some time between 1976 and 1987; there are up to 30 years of data per firm. This document describes how the sample was constructed and defines the variables which are available for each firm-year observation (49,225 observations in all).

2. Dataset Description

The final dataset is on one file: RNDPAN87.SAS.DEC89, or RNDPAN87.MASTER.DEC89. Table 2.1 shows the number of firms and the distribution of these firms across beginning and ending years for the

dataset. The fixed format ASCII (or EBCDIC) version of this file has the following characteristics:

DSNAME=RNDPAN87.MASTER.DEC89

VOL=SER=volume ID

DCB=(RECFM=FB,LRECL=1316,BLKSIZE=26320)

The format is integer, floating point, and four character string variables. In Fortran, this would be

(2A4, 2A4, 7A4, 7A4, 7I5, 93F13.5)

and in SAS,

($8. $8. $28. $28. 7*5.0 93*13.5)

There also exists a SAS dataset on tape called RNDPAN87 (dataset name RNDPAN87.SAS.DEC89) containing these variables, which is preferred if you plan to use SAS to read the data. The missing value code on both the fixed format version of the file and on the SAS dataset is the usual SAS code . (period or dot).

In Table 2.2(a) I show an alphabetical list of the variables on the file with a short description of each one. The table also shows the Compustat item number if the variable came directly from that file, the type of the variable, and its byte location on the Fortran file. Character variables are indicated by 'CHAR.' Tables 2.2(b) and 2.2(c) show the same list sorted by the byte location on the tape and by the location of the variables on the original Compustat tape. Many of the variables come directly from the Compustat files and further information can be obtained by reference to that Codebook. Those which are computed for this dataset are described somewhat more fully in the section following the tables. All dollar values are in millions of current U.S. dollars.

TABLE 2.1

DATA AVAILABLITY BY YEAR

Ending Year

Beg. Year	75	76	77	78	79	80	81	82	83	84	85	86	87	Total	Total # of Firms
59	6	14	21	10	16	13	21	10	24	25	27	25	302	514	514
60	10	20	15	14	13	14	24	11	15	12	23	25	162	358	872
61		2	5	3	10	7	5	4	10	6	9	3	53	117	989
62	1	5	4	1	2	5	7	7	3	7	10	6	65	123	1111
63	2	2	1	1	6	2	5	1	7	7	3	7	45	89	1201
64	1	2	2	3	3	3	2	1	5	6	4	4	18	54	1253
65	2		1	5	3	2	6		6	3	8	7	40	83	1338
66	1	5	3	4	3	5	4	2	2	4	5	5	44	87	1422
67	2	3	2	1	1	2	4		7	3	7	9	40	81	1506
68	1	5	13	27	6	7	5	3	5	7	6	5	77	167	1666
69		4	2	2	3	4	4	1	1	5	5	5	39	75	1738
70		1	1	7	5	2	5	1	2	4	1	3	20	52	1793
71		3	3	4	1	3	9	4	4	5	6	8	45	95	1887
72	1	2	2	1	6	3	6	4	7	5	5	5	45	92	1977
73		1	1	1	2	2	1	2	3	7	2	13	45	80	2058
74			2		1	3	4	4	5	9	10	6	93	137	2193
75				1		2	2				1	2	12	20	2209
76			1		2	1		2			1	1	11	19	2212
77						2	1		1	2	1		8	15	2156
78								1	1	1	1		18	22	2102
79						2	2	2	3	4	3	5	31	52	2069
80							1		2	1	6	11	46	67	2052
81									1	2		9	25	37	2007
82									1	4	15	8	61	89	1978
83									1	2	1	6	47	57	1976
84											3	6	32	41	1900
85												4	49	53	1824
86												1	38	39	1700
87													11	11	1522
Total	27	69	79	85	83	84	118	60	116	131	163	189	1522	2726	49225

Each cell in the table shows the number of firms whose data begins in the row year and ends in the column year.

The last column shows the total number of firms whose data is available in the row year.

TABLE 2.2(a)

Variable List Sorted by Name

CS Loc	Name	Type	Loc	Description
	ADJ	NUM	1213	(Curr.Liab.-Debt)-(Curr.Assts-Invns.)
	ADJDEP	NUM	1252	Depreciation adjusted for inflation
27	ADJFACT	NUM	459	Adjustment factor for common stock
27	ADJFLAG	NUM	472	Adj. factor for common (fiscal year)
	ADJINV	NUM	1122	Inventories adjusted for inflation
	ADJTOT	NUM	1135	TOTAL Defl. using age structure
43	ADV	NUM	693	Advertising expense
	BKCAP	NUM	1226	Book value of capital stock
9	BKDEBT	NUM	225	Book value of long-term debt
3	BKINV	NUM	147	Book value of inventory
8	BKPLNT	NUM	213	Property, plant, and equip - total (net)
1	CASH	NUM	121	Cash and short-term investments
39	CDBTPSTK	NUM	615	Convertible debt and preferred stock.
hdr	CIC	NUM	78	CUSIP issue number and check digit
hdr	CNAME	CHAR	17	Company name
11	COMEQY	NUM	251	Common equity - tangible.
20	COMINC	NUM	368	Income available for common
41	COSTGOOD	NUM	641	Cost of goods sold
4	CURRASST	NUM	160	Current assets - total.
hdr	CUSIP	CHAR	1	CUSIP number for firm
44	DEBT1YR	NUM	680	Debt - due in one year.
N25	DEMP	NUM	1070	Employees include seasonal figures
14	DEPREC	NUM	290	Depreciation and amortization
35	DFRTAX	NUM	563	Deferred taxes and Inv tax credit (bal sht
74	DFRTAXBS	NUM	810	Deferred taxes (balance sheet).
26	DIV	NUM	446	Dividends per share
N22	DLCOMP	NUM	1057	LCOMP excludes employee benefits
N1	DMERGER	NUM	1083	Major merger dummy (CS footnote)
	DRND	NUM	1044	Code for R&D corrections
hdr	DUP	NUM	103	Compustat code for duplicate files
53	EARNSH	NUM	797	Earnings per share (primary).
29	EMPLY	NUM	485	Number of employees (1000s)
60	EQUITY	NUM	888	Common equity
hdr	FILE	NUM	83	Compustat file code
hdr	FYR	NUM	758	Month of fiscal year end
	GRATE	NUM	1187	Gross rate of return
	GROCAP	NUM	1174	Gross cap. stock adj. for inflation
7	GROPLT	NUM	199	Property, plant, and equip - total (gross)
hdr	INAME	CHAR	45	Name of Industry (from SIC)
18	INCOME	NUM	342	Income before xtry items
16	INCTAX	NUM	316	Income taxes - total.
71	INCTAXP	NUM	771	Income taxes payable.

Table 2.2 (a) (continued)

CS Loc	Name	Type	Loc	Description
33	INTANG	NUM	537	Intangibles.
15	INTRST	NUM	303	Interest expense.
37	INVCAP	NUM	589	Total invested capital
30	INVEST	NUM	498	Capital expenditures
59	INVMETH	NUM	875	Inventory valuation method.
42	LCOMP	NUM	654	Labor compensation
	LTDEBT	NUM	1200	Long-term debt adj. for age structure
38	MININT	NUM	602	Minority interest (balance sheet).
	NETCAP	NUM	1096	Net capital stock adj. for inflation
52	NOLCFU	NUM	784	Net operating loss carry forward
25	NOSHARES	NUM	433	Common shares outstanding (1000s)
	NPLANT	NUM	1109	Net plant value adj. for inflation
	NRATE	NUM	1239	Net rate of return
13	OPINC	NUM	277	Operating income before depreciation
24	PCLOSE	NUM	420	Price of common -close of calendar year.
43	PENSION	NUM	667	Pension and retirement expense
90	PENSPAST	NUM	862	Pension costs - unfunded past service
89	PENSVEST	NUM	849	Pension costs - unfunded vested benefits
22	PHIGH	NUM	394	Price of common - high during year.
23	PLOW	NUM	407	Price of common - low during year.
19	PREFD	NUM	355	Dividends on preferred stock.
10	PREFST	NUM	238	Preferred stock (liquidating value).
56	PREFSTR	NUM	836	Preferred stock (redemption value).
	RDINIT	NUM	1304	R&D stock in initial year.
2	RECEIV	NUM	134	Receivables - total.
47	RENTAL	NUM	719	Rental expense
36	RETEARN	NUM	576	Retained earnings.
46	RND	NUM	706	R&D expenditures in current $
	ROR	NUM	745	Calendar yr rate of return to common
	RSTOCK	NUM	1291	Stock of R&D capital
12	SALES	NUM	264	Sales in current $
111	SFDBT	NUM	953	Issuance of long term debt (stmt of ch).
108	SFEQU	NUM	914	Sale of common equity (stmt of changes).
109	SFINV	NUM	927	Sale of investments (stmt of changes).
110	SFOPI	NUM	940	Funds from operations (stmt of changes).
107	SFPPE	NUM	901	Sale of PPE (stmt of changes).
112	SFTOT	NUM	966	Sources of funds -total (stmt of ch).
hdr	SIC	NUM	73	Compustat 4-Digit Industry code
hdr	SMBL	CHAR	9	Stock ticker symbol
17	SPECIAL	NUM	329	Special items.
34	STDEBT	NUM	550	Debt portion of current liabilities
hdr	STK	NUM	98	Compustat stock ownership code
5	STLIAB	NUM	173	Current liabilities - total.
	TOTAL	NUM	1148	Invest in uncons subs & intang & other
31	TOTAL1	NUM	511	Investments and advances - equity method
32	TOTAL2	NUM	524	Investments and advances - other.

Table 2.2 (a) (continued)

CS Loc	Name	Type	Loc	Description
6	TOTASST	NUM	186	Assets - total/Liab plus equity - total
21	TOTDIV	NUM	381	Dividends on common - total.
129	UFACQ	NUM	1031	Acquisitions (from stmnt of changes)
128	UFCAP	NUM	823	Capital expense (stmt of changes).
114	UFDBT	NUM	992	Reduction in long term debt (stmt of ch)
127	UFDIV	NUM	628	Cash dividends (statement of changes).
115	UFEQU	NUM	1005	Repurchase of stock (stmt of changes).
113	UFINV	NUM	979	Increase in investments (stmt of ch).
116	UFTOT	NUM	1018	Uses of funds - total (stmt of changes).
	VAL	NUM	1161	Market value (D+E) in current $
	VCOMS	NUM	1265	Value of common stock in current $
	VCOMS2	NUM	1278	Alternate val of common in current $
hdr	XREL	NUM	93	Compustat code for industry group
48	XTRY	NUM	732	Extraordinary items and disc. operations
hdr	YEAR	NUM	108	Year of observation (data as of end yr)
hdr	ZLIST	NUM	88	Exchange listing and S & P Index code

TABLE 2.2(b)

Variable List Sorted by Location on Tape

CS Loc	Name	Type	Loc	Description
hdr	CUSIP	CHAR	1	CUSIP number for firm
hdr	SMBL	CHAR	9	Stock ticker symbol
hdr	CNAME	CHAR	17	Company name
hdr	INAME	CHAR	45	Name of Industry (from SIC)
hdr	SIC	NUM	73	Compustat 4-Digit Industry code
hdr	CIC	NUM	78	CUSIP issue number and check digit
hdr	FILE	NUM	83	Compustat file code
hdr	ZLIST	NUM	88	Exchange listing and S & P Index code
hdr	XREL	NUM	93	Compustat code for industry group
hdr	STK	NUM	98	Compustat stock ownership code
hdr	DUP	NUM	103	Compustat code for duplicate files
hdr	YEAR	NUM	108	Year of observation (data as of end yr)
1	CASH	NUM	121	Cash and short-term investments
2	RECEIV	NUM	134	Receivables - total.
3	BKINV	NUM	147	Book value of inventory
4	CURRASST	NUM	160	Current assets - total.
5	STLIAB	NUM	173	Current liabilities - total.
6	TOTASST	NUM	186	Assets - total/Liab plus equity - total
7	GROPLT	NUM	199	Property, plant, and equip - total (gross)
8	BKPLNT	NUM	213	Property, plant, and equip - total (net)
9	BKDEBT	NUM	225	Book value of long-term debt
10	PREFST	NUM	238	Preferred stock (liquidating value).
11	COMEQY	NUM	251	Common equity - tangible.
12	SALES	NUM	264	Sales in current $
13	OPINC	NUM	277	Operating income before depreciation
14	DEPREC	NUM	290	Depreciation and amortization
15	INTRST	NUM	303	Interest expense.
16	INCTAX	NUM	316	Income taxes - total.
17	SPECIAL	NUM	329	Special items.
18	INCOME	NUM	342	Income before xtry items
19	PREFD	NUM	355	Dividends on preferred stock.
20	COMINC	NUM	368	Income available for common
21	TOTDIV	NUM	381	Dividends on common - total.
22	PHIGH	NUM	394	Price of common - high during year.
23	PLOW	NUM	407	Price of common - low during year.
24	PCLOSE	NUM	420	Price of common -close of calendar year.
25	NOSHARES	NUM	433	Common shares outstanding (1000s)
26	DIV	NUM	446	Dividends per share
27	ADJFACT	NUM	459	Adjustment factor for common stock
27	ADJFLAG	NUM	472	Adj. factor for common (fiscal year)
29	EMPLY	NUM	485	Number of employees (1000s)
30	INVEST	NUM	498	Capital expenditures
31	TOTAL1	NUM	511	Investments and advances - equity method

Table 2.2(b) (continued)

CS Loc	Name	Type	Loc	Description
32	TOTAL2	NUM	524	Investments and advances - other.
33	INTANG	NUM	537	Intangibles.
34	STDEBT	NUM	550	Debt portion of current liabilities
35	DFRTAX	NUM	563	Deferred taxes and Inv tax credit (bal sht)
36	RETEARN	NUM	576	Retained earnings.
37	INVCAP	NUM	589	Total invested capital
38	MININT	NUM	602	Minority interest (balance sheet).
39	CDBTPSTK	NUM	615	Convertible debt and preferred stock.
127	UFDIV	NUM	628	Cash dividends (statement of changes).
41	COSTGOOD	NUM	641	Cost of goods sold
42	LCOMP	NUM	654	Labor compensation
43	PENSION	NUM	667	Pension and retirement expense
44	DEBT1YR	NUM	680	Debt - due in one year.
45	ADV	NUM	693	Advertising expense
46	RND	NUM	706	R&D expenditures in current $
47	RENTAL	NUM	719	Rental expense
48	XTRY	NUM	732	Extraordinary items and disc. operations
	ROR	NUM	745	Calendar yr rate of return to common
hdr	FYR	NUM	758	Month of fiscal year end
71	INCTAXP	NUM	771	Income taxes payable.
52	NOLCFU	NUM	784	Net operating loss carry forward
53	EARNSH	NUM	797	Earnings per share (primary).
74	DFRTAXBS	NUM	810	Deferred taxes (balance sheet).
128	UFCAP	NUM	823	Capital expense (stmt of changes).
56	PREFSTR	NUM	836	Preferred stock (redemption value).
89	PENSVEST	NUM	849	Pension costs - unfunded vested benefits
90	PENSPAST	NUM	862	Pension costs - unfunded past service
59	INVMETH	NUM	875	Inventory valuation method.
60	EQUITY	NUM	888	Common equity
107	SFPPE	NUM	901	Sale of PPE (stmt of changes).
108	SFEQU	NUM	914	Sale of common equity (stmt of changes).
109	SFINV	NUM	927	Sale of investments (stmt of changes).
110	SFOPI	NUM	940	Funds from operations (stmt of changes).
111	SFDBT	NUM	953	Issuance of long term debt (stmt of ch).
112	SFTOT	NUM	966	Sources of funds -total (stmt of ch).
113	UFINV	NUM	979	Increase in investments (stmt of ch).
114	UFDBT	NUM	992	Reduction in long term debt (stmt of ch)
115	UFEQU	NUM	1005	Repurchase of stock (stmt of changes).
116	UFTOT	NUM	1018	Uses of funds - total (stmt of changes).
129	UFACQ	NUM	1031	Acquisitions (from stmnt of changes)
	DRND	NUM	1044	Code for R&D corrections
N22	DLCOMP	NUM	1057	LCOMP excludes employee benefits
N25	DEMP	NUM	1070	Employees include seasonal figures
N1	DMERGER	NUM	1083	Major merger dummy (CS footnote)
	NETCAP	NUM	1096	Net capital stock adj. for inflation
	NPLANT	NUM	1109	Net plant value adj. for inflation
	ADJINV	NUM	1122	Inventories adjusted for inflation

Table 2.2(b) (continued)

CS Loc	Name	Type	Loc	Description
	ADJTOT	NUM	1135	TOTAL Defl. using age structure
	TOTAL	NUM	1148	Invest in uncons subs & intang & other
	VAL	NUM	1161	Market value (D+E) in current $
	GROCAP	NUM	1174	Gross cap. stock adj. for inflation
	GRATE	NUM	1187	Gross rate of return
	LTDEBT	NUM	1200	Long-term debt adj. for age structure
	ADJ	NUM	1213	(Curr.Liab.-Debt)-(Curr.Assts-Invns.)
	BKCAP	NUM	1226	Book value of capital stock
	NRATE	NUM	1239	Net rate of return
	ADJDEP	NUM	1252	Depreciation adjusted for inflation
	VCOMS	NUM	1265	Value of common stock in current $
	VCOMS2	NUM	1278	Alternate val of common in current $
	RSTOCK	NUM	1291	Stock of R&D capital
	RDINIT	NUM	1304	R&D stock in initial year.

TABLE 2.2(c)

Variables Sorted by Location on Compustat Tape

CS Loc	Name	Type	Loc	Description
1	CASH	NUM	121	Cash and short-term investments
2	RECEIV	NUM	134	Receivables - total.
3	BKINV	NUM	147	Book value of inventory
4	CURRASST	NUM	160	Current assets - total.
5	STLIAB	NUM	173	Current liabilities - total.
6	TOTASST	NUM	186	Assets - total/Liab plus equity - total
7	GROPLT	NUM	199	Property, plant, and equip - total (gross)
8	BKPLNT	NUM	213	Property, plant, and equip - total (net)
9	BKDEBT	NUM	225	Book value of long-term debt
10	PREFST	NUM	238	Preferred stock (liquidating value).
11	COMEQY	NUM	251	Common equity - tangible.
12	SALES	NUM	264	Sales in current $
13	OPINC	NUM	277	Operating income before depreciation
14	DEPREC	NUM	290	Depreciation and amortization
15	INTRST	NUM	303	Interest expense.
16	INCTAX	NUM	316	Income taxes - total.
17	SPECIAL	NUM	329	Special items.
18	INCOME	NUM	342	Income before xtry items
19	PREFD	NUM	355	Dividends on preferred stock.
20	COMINC	NUM	368	Income available for common
21	TOTDIV	NUM	381	Dividends on common - total.
22	PHIGH	NUM	394	Price of common - high during year.
23	PLOW	NUM	407	Price of common - low during year.
24	PCLOSE	NUM	420	Price of common -close of calendar year.
25	NOSHARES	NUM	433	Common shares outstanding (1000s)
26	DIV	NUM	446	Dividends per share
27	ADJFACT	NUM	459	Adjustment factor for common stock
27	ADJFLAG	NUM	472	Adj. factor for common (fiscal year)
29	EMPLY	NUM	485	Number of employees (1000s)
30	INVEST	NUM	498	Capital expenditures
31	TOTAL1	NUM	511	Investments and advances - equity method
32	TOTAL2	NUM	524	Investments and advances - other.
33	INTANG	NUM	537	Intangibles.
34	STDEBT	NUM	550	Debt portion of current liabilities
35	DFRTAX	NUM	563	Deferred taxes and Inv tax credit (bal sht)
36	RETEARN	NUM	576	Retained earnings.
37	INVCAP	NUM	589	Total invested capital
38	MININT	NUM	602	Minority interest (balance sheet).
39	CDBTPSTK	NUM	615	Convertible debt and preferred stock.
41	COSTGOOD	NUM	641	Cost of goods sold
42	LCOMP	NUM	654	Labor compensation

11

Table 2.2 (c) (continued)

CS Loc	Name	Type	Loc	Description
43	PENSION	NUM	667	Pension and retirement expense
44	DEBT1YR	NUM	680	Debt - due in one year.
45	ADV	NUM	693	Advertising expense
46	RND	NUM	706	R&D expenditures in current $
47	RENTAL	NUM	719	Rental expense
48	XTRY	NUM	732	Extraordinary items and disc. operations
52	NOLCFU	NUM	784	Net operating loss carry forward
53	EARNSH	NUM	797	Earnings per share (primary).
56	PREFSTR	NUM	836	Preferred stock (redemption value).
59	INVMETH	NUM	875	Inventory valuation method.
60	EQUITY	NUM	888	Common equity
71	INCTAXP	NUM	771	Income taxes payable.
74	DFRTAXBS	NUM	810	Deferred taxes (balance sheet).
89	PENSVEST	NUM	849	Pension costs - unfunded vested benefits
90	PENSPAST	NUM	862	Pension costs - unfunded past service
107	SFPPE	NUM	901	Sale of PPE (stmt of changes).
108	SFEQU	NUM	914	Sale of common equity (stmt of changes).
109	SFINV	NUM	927	Sale of investments (stmt of changes).
110	SFOPI	NUM	940	Funds from operations (stmt of changes).
111	SFDBT	NUM	953	Issuance of long term debt (stmt of ch).
112	SFTOT	NUM	966	Sources of funds -total (stmt of ch).
113	UFINV	NUM	979	Increase in investments (stmt of ch).
114	UFDBT	NUM	992	Reduction in long term debt (stmt of ch)
115	UFEQU	NUM	1005	Repurchase of stock (stmt of changes).
116	UFTOT	NUM	1018	Uses of funds - total (stmt of changes).
127	UFDIV	NUM	628	Cash dividends (statement of changes).
128	UFCAP	NUM	823	Capital expense (stmt of changes).
129	UFACQ	NUM	1031	Acquisitions (from stmnt of changes)
hdr	CIC	NUM	78	CUSIP issue number and check digit
hdr	CNAME	CHAR	17	Company name
hdr	CUSIP	CHAR	1	CUSIP number for firm
hdr	DUP	NUM	103	Compustat code for duplicate files
hdr	FILE	NUM	83	Compustat file code
hdr	FYR	NUM	758	Month of fiscal year end
hdr	INAME	CHAR	45	Name of Industry (from SIC)
hdr	SIC	NUM	73	Compustat 4-Digit Industry code
hdr	SMBL	CHAR	9	Stock ticker symbol
hdr	STK	NUM	98	Compustat stock ownership code
hdr	XREL	NUM	93	Compustat code for industry group
hdr	YEAR	NUM	108	Year of observation (data as of end yr)
hdr	ZLIST	NUM	88	Exchange listing and S & P Index code
N1	DMERGER	NUM	1083	Major merger dummy (CS footnote)
N22	DLCOMP	NUM	1057	LCOMP excludes employee benefits
N25	DEMP	NUM	1070	Employees include seasonal figures

3. Details of the Variable Construction

This section of the documentation describes the construction of each variable on the tape in somewhat more detail than the lists in Table 2.2. Users who wish to know more about the variables drawn directly from the Compustat files should consult that documentation for further information.

ADJ -- Value of net short-term assets, equal to current assets (CURRASST), less the value of inventories (BKINV), less current liabilities (STLIAB) plus the value of short-term debt (STDEBT). Short-term debt is added back in because it is treated elsewhere, and because it is assumed that interest is being paid on it, which is not the case for the other short-term liablities.

The short-term liabilities of a firm are composed of accounts payable, income taxes payable, accrued expenses and costs on contracts, dividends declared, employee benefits, customer deposits, and the current portion of long-term debt.[2] This variable does not include unfunded pension liabilities.

The short-term assets of a firm are cash and short-term investments, accounts receivable, inventories, prepaid taxes, estimated future tax benefits, security deposits, supplies and tools not in inventory, and prepaid expenses.

ADJDEP -- This year's depreciation adjusted for the effects of inflation. This variable is DEPREC deflated by the ratio of the GNP deflator for fixed nonresidential investment AA (see NPLANT for a definition of AA, average age) years ago to the current GNP deflator.

ADJFACT -- Compustat data item #27, the cumulative adjustment factor for common stock splits, stock dividends, etc. This factor, applied to the per-share data, converts such data into terms of the share units prevailing in the last year of data. Price data should be divided by ADJFACT to convert to current units, while shares data should be multiplied by ADJFACT.

ADJFLAG -- like ADJFACT, but lagged one year.

ADJINV -- the value of the firms' inventories adjusted for the effects of inflation. The inventories are valued at cost (book) each year unless the firm has specified LIFO as one of its methods of inventory valuation. The use of LIFO implies that the reported inventory valuation will be too low and an attempt is made to

_____

[2]This is included in short-term debt but also exist as a separate item on the Compustat tape. There is potentially some additional information here on the long-term debt issued 19 years ago.

13

adjust for this, using the change in inventories from year to year to obtain a measure of how old the inventories are and inflating them accordingly.

The process is started by setting ADJINV equal to the book value in the first year. Then the value of the inventories for each succeeding year is calculated as follows: if there is an increase in book inventory from last year to this, last year's adjusted inventories (ADJINV) are inflated by the ratio of the inventory price index this year to the price index last year and then the increase (assumed to be in current dollars) is added. If there was a decrease, last year's ADJINV is inflated in the same way and then written down by the ratio of this year's book inventories to last year's. The firm uses more than one method of inventory valuation, the book values are combined with the adjusted values using proportional weights derived from the reported ranking of the methods. Each firm uses up to three inventory methods and they are ranked in order of importance on the tape. I use the following rule of thumb to determine the weights:

Reported Number of Inventory Methods	Rank of LIFO	% LIFO as weight
1	1	100.0
2	1	66.7
2	2	33.3
3	1	50.0
3	2	33.3
3	3	16.7

ADJTOT -- Investments in unconsolidated subisdiaries and intangibles plus other investments (TOTAL) adjusted for inflation. In order to adjust the reported values of these items for inflation, we take the same approach as we did for inventories. In the first year, the value of ADJTOT is simply set equal to its book value. In each succeeding year, the previous year's ADJTOT is inflated by the ratio of the deflators for fixed residential investment in the two years and then the change in the book value of the investments is added. If there was a decrease in book, last year's inflated ADJTOT is multiplied by the ratio of this year's book value to last year's. As in the case of inventories, this approximation becomes more and more accurate as time goes by, but possibly at a slower rate, owing to the older age of the assets making up the TOTAL variable on average.

ADV -- Compustat data item #45, advertising expense.

BKCAP -- Net book value of capital stock. This variable is defined as

$$BKCAP = BKPLANT + BKINV + TOTAL$$

BKDEBT -- Compustat data item #9, book value of long-term debt.

BKINV -- Compustat data item #3, book value of inventories.

14

BKPLANT -- Compustat data item #8, net book value of the firm's physical plant.

CASH -- Compustat data item #1, cash and short-term investments.

CDBTPFSTK -- Compustat data item #39, outstanding convertible debt and carrying value of convertible preferred stock.

CIC -- the Compustat CUSIP issue number (two digits which identify whether security is a stock, bond, etc.) and check digit.

CNAME -- 28-character alphabetic name of the company (all caps).

COMEQY -- Compustat data item #11, Common equity including retained earnings, capital surplus, self-insurance reserves, and capital stock premium.

COMINC -- Compustat data item #20, income available for common.

COSTGOOD -- Compustat data item #41, cost of goods sold.

CURRASST -- Compustat data item #4, total current assets.

CUSIP -- CNUM, the Compustat identifying number for the firm (up to six digits).

DEBT1YR -- Compustat data item #44, debt due in one year.

DEMP -- A dummy which is one when the number of employees includes substantial (>10%) numbers of seasonal and/or part-time employees (Compustat footnote number 25).

DEPREC -- Compustat data item #14, depreciation and amortization.

DFRTAX -- Compustat data item #35, deferred taxes and investment tax credit.

DFRTAXBS -- Compustat data item #74, deferred taxes (balance sheet).

DIV -- Compustat data item #26, dividends per share. ADJFACT must be used to make this variable comparable from year to year.

DLCOMP -- A dummy which is one when labor expense (LCOMP) excludes employee benefits (Compustat footnote number 22).

DMERGER -- a dummy which is equal to one if Compustat reported that the data was affe6cted by a merger or acquisition (footnote 1 equal to AA) and a 2 if Compustat reported a major merger (footnote 1 equal to AB).

DRND -- A flag for the corrections to the R&D figures. If the value is zero, no corrections other than looking up a missing number have been made. The other values possible are the following, although

15

only 4 or 6 are included on this version of the tape (these come
from Compustat footnote number 3):

1   Customer-sponsored R&D removed
2   Engineering removed
3   Other redefinition of R&D, including interpolation
4   Engineering included
5   From the 157-firm sample (i.e., a Nadiri or other
    number)
6   Customer or Govt-sponsored R&D included.

DUP -- Duplicate company code, which identifies companies carried
      on more than one Compustat file. The codes possible are

00   No duplicate file
81   Primary Industrial file and Canadian file
83   Tertiary Industrial file and Canadian file
84   Supplementary Industrial file and Canadian file
86   Full Coverage file, OTC file, and Canadian file
91   Full Coverage file and Primary Industrial file
93   Full Coverage file and Tertiary Industrial file
94   Full Coverage file and Supplementary Industrial file
96   Full Coverage file and Over-the-Counter file
98   Full Coverage file and Canadian file

Since there are no Canadian firms on this file, all firms have
the code 00 in this edition.

EARNSH -- Compustat data item #53, earnings per share - including
      extraordinary items and discontinued operations.

EMPLY -- Compustat data item #29, number of employees. This is the
      number of company workers as reported to shareholders. It may be
      an average throughout the year or an end-of-year number; the
      latter is reported if both are given. It includes part-time
      employees and the employees of consolidated subsidiaries.

EQUITY -- Compustat data item #60, common equity (as reported). It inclues
      common stock outstanding, capital surplus, and retained earnings.

FILE -- the Compustat file code, with the following meaning:

1 Primary Industrial File
2 Bank File
3 Tertiary File
4 Supplementary Industrial File
5 Full Coverage File
6 Over-the-Counter File
11 Primary Industrial File and in the S&P Industrials Index
13 Full Coverage File

FYR -- the month in which the fiscal year ends for this firm and
      year of data. Fiscal years ending between January 1 and May
      31 are treated as though they ended in the prior calendar
      year.

GRATE -- the gross rate of return to capital, defined as the ratio of gross cash flows to GROCAP. Gross cash flows are the sum of the income before extraordinary items (INCOME), depreciation (DEPREC), and interest income (INTRST) less an inventory valuation adjustment and an imputed income from short-term assets (the prime rate times ADJ).

GROCAP -- the gross capital stock adjusted for inflation. This variable is computed as

GROCAP = GPLANT + ADJINV + ADJTOT.

GPLANT is gross plant revalued as for NPLANT but with the gross book value of the plant as input. GPLANT itself is not included on the tape.

GROPLT -- Compustat data item #7, gross book value of the firm's physical plant.

INAME -- the 28 character name of the 4-digit industry to which Compustat assigned this firm.

INCOME -- Compustat data item #18, income before extraordinary items and discontinued operations.

INCTAX -- Compustat data item #16, income taxes (total, from income statement).

INCTAXP -- Compustat data item #71, income taxes payable (balance sheet).

INTANG -- Compustat data item #33, intangibles.

INTRST -- Compuustat data item #15, interest expense.

INVCAP -- Compustat data item #37, total invested capital.

INVEST -- Compustat data item #30, capital expenditures (gross investment). The amount spent for the construction and/or acquisition of property, plant, and equipment, including that of purchased companies (acquisitions).

INVMETH -- Compustat data item #59, inventory valuation method.

LCOMP -- Compustat data item #42, labor and related expense including salaries, wages, profit sharing, payroll taxes, and employee benefits. (See DLCOMP also.)

LTDEBT -- The value of long-term debt adjusted for its age structure. Long-term debt is assumed to have been financed by 20-year bonds. Given a matrix of bond prices in year t for a bond due in year s and the distribution of the firm's long-term debt by the year incurred, it is possible to adjust the face value of the debt by the bond rates in any given year. We obtain such a matrix of prices from the Moody's corporate BAA bond price

series given previously. The difficulty is to determine the age structure of each firm's debt. In the absence of any history of the firm prior to 1958, the approach taken is to assume that all firms have an age structure of debt which is the same as the aggregate age structure from the 1958 Survey of Current Business. This distribution is shown in Table 2.4.

After 1958, we attempt to build up each firm's own long-term debt distribution in the following manner: starting with the 1958 distribution, in each successive year the amount of new long-term debt issued equals the change in long-term debt From the previous year to this plus the amount retired this year, which is assumed to be equal to that issued 20 years ago. There are two sources of error in this computation:

a) The firm's age structure of debt may not be the same as the aggregate in 1958.

b) The bonds may not always be for a twenty-year term.

Because of these sources of error, in any year the gross new debt issued may be negative. If this is the case, we set this contribution to zero and rescale the entire twenty year of long-term debt accordingly. Given the age distribution of book value of long-term debt for each of the twenty-three years in the and the matrix of bond prices we can now compute the market value of debt as

$$(1) \quad TL_t = \sum_{s=t-20}^{t-1} BV_s \cdot P_{st}$$

where BV is the book value distribution and P_{st} is the price in year t of a bond due twenty years from year s.

MININT -- Compustat data item #38, minority interest (balance sheet).

NETCAP -- the inflation adjusted net capital stock. This variable is computed as

NETCAP = NPLANT + ADJINV + ADJTOT

NOLCFU -- Compustat data item #52, net operating loss carry forward.

NOSHARES -- Compustat data item #25, the number of common shares outstanding. This variable should be multiplied by the adjustment factor ADJFACT to make it comparable across the years.

NPLANT -- the net value of the plant adjusted for inflation. This quantity is obtained by multiplying the book plant value by the ratio of the GNP deflator for fixed nonresidential investment in the current year to GNP deflator AA years ago.

AA is the average age of the plant and equipment for this firm which is deduced in the following manner: an average age series is obtained as the ratio of accumulated depreciation (gross plant minus net plant) to depreciation this year. This assumes straight-line depreciation.

Then a length of life of the current plant and equipment is also computed as the gross plant divided by this year's depreciation and a five-year moving average is taken of this series to smooth it. This year's average age is then adjusted by the ratio of this year's length of life to the moving average. This has the effect of smoothing the age series slightly. If the average age exceeds nineteen years, it is set to nineteen years, since the deflator is only available back to 1939.

NRATE -- the net rate of return to capital, defined as the ratio of net cash flows less the inflation adjusted value of depreciation (ADJDEP) divided by the net capital stock (NETCAP).

OPINC -- Compustat data item #13, operating income before depreciation.

PCLOSE -- Compustat data item #24, end of calendar year stock price. This should be divided by ADJFACT to make it comparable to data in other years.

PENSION -- Compustat data item #43, pension and retirement expense.

PENSPAST -- Compustat data item #90, unfunded pension costs, past or prior service.

PENSVEST -- Compustat data item #89, unfunded pension costs, vested benefits.

PHIGH -- Compustat data item #22, price of common - high during year. This variable should be divided by ADJFACT to make it comparable to data in other years.

PLOW -- Compustat data item #23, price of common - low during year. This variable should be divided by ADJFACT to make it comparable to data in other years.

PREFD -- Compustat data item #19, dividends on preferred stock.

PREFST -- Compustat data item #10, preferred stock (liquidating value).

PREFSTR -- Compustat data item #56, preferred stock (redemption value).

RDINIT -- R&D stock in initial year, calculated as described in the appendix.

RECEIV -- Compustat data item #2, receivables.

RENTAL -- Compustat data item #47, rental expense. All costs

charged to operations for the rental of space and/or equipment.

RETEARN -- Compustat data item #36, retained earnings.

RND -- Compustat data item #46, expenditures on research and development. The private (company-funded) expenditures on the development of new products and services including software but excluding customer of government-sponsored expenditures, exploration, engineering expense, inventor royalties, and market research and testing. Because continuous history of this variable is needed for the construction of RSTOCK, when it was randomly missing for one year in the middle of a continuous sequence of reported R&D, I interpolated a value using the model described in Appendix B. This model assumes that the logarithm of R&D expenditures evolves as a random walk, which is justified by the patterns of R&D spending actually observed. The interpolation affected less than ten percent of the firms.

Before use, this variable is generally deflated by the deflator shown in Table 2.3, which is a weighted average of an index of hourly labor compensation and the implicit price deflator in the non-financial corporate sector. The weights are 0.49 and 0.51 respectively, and the methodology is patterned on Jaffe (1972). The underlying data is from U.S. Department of Labor, Productivity Costs in Nonfinancial Corporations, various issues.

ROR -- This quantity is the one-period (year) rate of return to holding a share of the company's common stock over the calendar year. The definition of the one-period rate of return at time t is the following:

$$q_t = (p_t - p_{t-1} + d_t)/p_{t-1}$$

where p_t is the price of the common stock at the end of the year and d_t are the dividends received per share during the year. Note that for this definition it is important to be careful that p_{t-1} measures the same share of the company as p_t.

RSTOCK -- the stock of R&D capital, constructed from the history of R&D investment using a perpetual inventory model with declining balance depreciation. See Appendix B for details.

SALES -- Compustat data item #12, net sales. This is the amount of actual billings to customers for regular sales completed during the period, reduced by cash discounts, trade discounts, and returned sales for which credit is given to customers. Interest and equity income from unconsolidated subsidiaries, non-operating income, and income from discontinued operations are excluded.

SFDBT -- Compustat data item #111, issuance of long term debt (from statement of changes).

SFEQU -- Compustat data item #108, sale of common equity
(from statement of changes).

SFINV -- Compustat data item #109, sale of investments
(from statement of changes).

SFOPI -- Compustat data item #110, funds from operations
(from statement of changes).

SFPPE -- Compustat data item #107, sale of plant, property, and equipment
(from statement of changes).

SFTOT -- Compustat data item #112, source of funds - total
(from statement of changes).

SIC -- DNUM, the Compustat 4-digit Industry code. This variable should be
between 2000 and 3999. However, since some firms change industry
in some years (this most often happens in 1980-1982, when data
from the earlier Compustat tapes was spliced to data from the
1980 versions), there are a few firms with SIC codes outside of
manufacturing, at least in some years.

SMBL -- 8-character stock ticker symbol for this firm. Owing to
errors in the industrial file we received for 1959-1978, this
variable is garbage for some firms before 1979.

SPECIAL -- Compustat data item #17, special items (from income statement).

STDEBT -- Compustat data item #34, debt in current liabilities.

STK -- Stock ownership code from Compustat. The meaning is

    0 Publicly traded company
    1 Subsidiary of a public9ly traded company
    2 Subsidiary of a company that is not publicly traded
    3 Company that is publicly traded but is not on a major
       exchange

On this version of the file, there are no firms from the Full
Coverage tapes, and so the STK code is always equal to 0.

STLIAB -- Compustat data item #5, current liabilities - total.

TOTAL -- book value of investments in unconsolidated subsidiaries and
others and intangibles. Sum of TOTAL1, TOTAL2, and INTANG.

TOTAL1 - Compustat data item #31, investments in and advances to
unconsolidated subsidiaries where there is significant control,
all investments carried at equity.

TOTAL2 -- Compustat data item #32, investments in and advances to
unconsolidated subsidiaries where there is not control,
long-term receivables, and other investments.

TOTASST -- Compustat data item #6, total assets, equal to total liabilities

plus common equity.

TOTDIV -- Compustat data item #21, total dividends on common stock.

UFACQ -- Compustat data item #129, acquisitions
(from statement of changes)

UFCAP -- Compustat data item #128, capital expenditures
(from statement of changes)

UFDBT -- Compustat data item #114, reduction in long term debt
(from statement of changes)

UFDIV -- Compustat data item #127, cash dividends
(from statement of changes)

UFEQU -- Compustat data item #115, repurchase of common stock
(from statement of changes)

UFINV -- Compustat data item #113, increase in investments
(from statement of changes)

UFTOT -- Compustat data item #116, uses of funds - total
(from statement of changes)

VAL -- Market value of the firm, equal to the sum of the value of
the preferred stock (PREFST) the value of the common stock
(VCOMS), the long-term debt adjusted for its age structure
(LTDEBT), and the short-term debt (STDEBT), less the net short-
term assets (ADJ).

VCOMS -- Value of the common stock at the close of the calendar year.
The value is the price of the common stock times the number
of shares outstanding at the close of the year. This
computation is invariant to ADJFACT so long as the two
variables come from the same year of data.

VCOMS2 -- Average value of common stock throughout the calendar year.
The value is the average of the high and low price for the
year times the number of shares outstanding.

XREL -- a four-digit code that identifies the specific S&P industry
group in which each company in an S&P industry is contained.
See Compustat for details.

XTRY -- Compustat data item #48, extraordinary items and discontinued
operations.

YEAR -- two-digit year for this observation, 59 through 87.

ZLIST -- the Compustat Exchange listing and S&P Index code.

TABLE 3.1: DEFLATORS AND BOND YIELDS

Year	Preferred Dividend Rate Medium Risk Companies	Bond Yield	Fixed Investment Deflator	R&D Deflator
1939		0.0496	11.5	
1940		0.0475	11.9	
1941		0.0433	12.7	
1942		0.0428	13.3	
1943		0.0391	13.8	
1944		0.0361	14.0	
1945		0.0329	14.3	
1946		0.0305	16.4	
1947		0.0324	19.3	
1948		0.0347	21.0	
1949		0.0342	21.7	
1950		0.0324	22.4	
1951		0.0341	24.2	
1952		0.0352	24.2	
1953		0.0374	25.1	
1954		0.0351	25.2	
1955		0.0353	25.8	
1956		0.0388	27.7	
1957		0.0471	29.5	0.598
1958	.0514	0.0473	29.5	0.616
1959	.0499	0.0505	30.2	0.631
1960	.0518	0.0519	30.6	0.647
1961	.0482	0.0508	30.5	0.658
1962	.0481	0.0502	30.9	0.670
1963	.0469	0.0486	31.3	0.680
1964	.0467	0.0483	31.1	0.698
1965	.0460	0.0487	32.1	0.711
1966	.0503	0.0567	33.3	0.737
1967	.0534	0.0623	34.4	0.768
1968	.0583	0.0694	35.9	0.809
1969	.0662	0.0781	37.9	0.855
1970	.0770	0.0911	39.9	0.906
1971	.0711	0.0856	42.4	0.956
1972	.0703	0.0816	44.4	1.000
1973	.0729	0.0824	46.0	1.064
1974	.0837	0.0950	50.5	1.170
1975	.0847	0.1061	57.9	1.285
1976	.0792	0.0975	61.9	1.361
1977	.0779	0.0897	66.1	1.459
1978	.0824	0.0949	71.5	1.573
1979	.0911	0.1069	77.8	1.718
1980	.1060	0.1367	85.1	1.870

TABLE 3.1 (continued)

Year	Preferred Dividend Rate Medium Risk Companies	Bond Yield	Fixed Investment Deflator	R&D Deflator
1981	.1236	0.1604	93.4	2.010
1982	.1253	0.1611	100.0	2.100
1983	.1102	0.1355	98.8	2.175
1984	.1162	0.1419	97.9	2.279
1985	.1844	0.1272	97.7	2.381
1986	.0900	0.1039	100.2	2.468
1987		0.1058	100.4	2.560
1988		0.1083	100.7	2.645

TABLE 3.2

AGE STRUCTURE OF CORPORATE DEBT IN 1958

Age in Years	Fraction of Debt
19	.008
18	.020
17	.038
16	.038
15	.039
14	.047
13	.038
12	.038
11	.044
10	.060
9	.056
8	.058
7	.058
6	.062
5	.078
4	.075
3	.033
2	.059
1	.066
0	.085

4. Sample Construction

The construction of the sample proceeded in the following way (the peculiarities of datafile year selection are due to availability of the tapes): the sample was defined to be all firms in the R&D Master File constructed at the NBER in 1979-1982 and 1986-1987 (Hall et al. 1988), plus all manufacturing firms on the Compustat Annual Industrial and OTC files for 1986 and 1987, augmented by a few non-manufacturing firms which were formerly manufacturing firms or which acquired manufacturing firms during the period.

First, a list of firms to be selected was constructed by starting with the 1985 panel and adding all manufacturing firms which were on the Compustat Headers for the 1986, 1987 Industrial files, and the 1987 OTC file (the 1986 OTC file was unavailable). For all firms the latest appearance of the firm in the file was recorded; this produced a list of 2863 unique firms called PAN87HDA.JUN89, together with the name of the file from which their data was to be selected (IND, OTC, RES for the pre-1981 files, IND81-IND87, OTC81-OTC87, and RES84 for the post-1981 files). This list of headers was combined with the headers for the firms on the 1986 and 1987 Research Files to ensure that we had selected all manufacturing firms (only a small number of additional firms were taken from the Research files); the new list contained 2953 firms. All firms on this list (approximately 1200) which did not appear in one of the two current 1987 files were printed out.

This list of about 1200 firms was called (ultimately) EXITALL.JUL89, and it was used to investigate the reason for every exit from the panel between 1976 and 1987. The results of this investigation are documented in

a separate section (Appendix A). For the purposes of this project, the important fact is that it included approximately 150 firms which changed their name or reorganized, causing a change to the CUSIP number of the firm and thus a spurious exit. These cases needed eventually to have their data spliced between observations under the new and old CUSIPs. To prepare for this, the list of new CUSIPs was run against the current panel master header (PAN87HDA.JUN89), and a few firms were found which had been overlooked for some reason (usually a change of industry). The library work involved in identifying all the exits also identified a few problems (duplicates, etc.) in PAN87HDA, and the ultimate cleaned list of firms to be selected contained 2897 firms.

The next task was to obtain the actual data; the goal was to maximize both the number of years of data available for each firm and to maximize the number of non-missing values for each variable. To do this, at least two tapes (of different vintages) were used as sources for many of the firms. A series of jobs (SELIND.., SELOTC.., and SELRES..) were run to select those observations which were to come from the Annual Industrial 1981, 1983-1987, OTC 1981, 1983-1985, 1987, and Research Industrial 1984, 1987 files. The number of observations thus selected is shown in Table 4.1.

The output tapes from these jobs (ANNIND81-87, OTCINC81-87, and RESIND84-87) were merged and reformatted by MKCS8187 to a single file (CS8187.REFORMAT) which was organized one firm-year per record, approximately 40,000 observations in all. This was done so that the data could be updated more easily in the next step. In addition, to save space, all firm-year observations with no data were deleted and the last 100 of the 175 variables available for each firm-year were also deleted. A few of

27

these variables (locations 71, 74, 89, 90, 107-116, 127, 128, 129) were

moved to lower-valued locations (51, 54, 57, 58, 61-70, 40, 55, 71), since

they would be needed later. Also, the fiscal year month (FYR) was placed

in location 50 and footnotes 1 (mergers), 3 (R&D), 22 (labor compensation),

and 25 (seasonal and parttime employees) were recoded for dummies in

location 72-75. Finally, the lagged adjustment factor was placed in

location 28. This variable is necessary to provide continuity in the per

share data when Compustat tapes of different vintages are spliced.

A parallel selection and reformatting was constructed using the old

Compustat tapes CIND5981, CRES5978.EDIT, and COTC6182.FIX80 with the

following difference: all firms on these tapes (those in both the primary

and secondary columns of Table 4.1) were selected by data for long-lived

firms (more than 20 years old). This yielded a total of 1379 old

industrial firms, 166 old Research firms, and 495 old OTC firms. These

files (OLDCIND, OLDCOTC, and OLDCRES) were then reformatted by MKOLDIND,

MKOLDOTC, and MKOLDRES, yielding 45,927 firm-year observations in the same

format as CS8187.REFORMAT.

The same kind of selection and reformatting exercise was performed by

jobs SELRSIUP, SELR84UP, SELR86UP, SELR87UP, and MKCR8187 to obtain the 735

firms in the secondary column of Table 4.1, yielding about 14,700 firm-year

observations. This was done because occasionally firms exit from the main

Compustat files before their last one or two years of data are recorded,

so that it is necessary to obtain the additional data from the Research

file, where the firm is placed after it exits. This is a very important

step if we are interested in pre- and post-acquisition behavior of firms.

The five reformatted files OLD....REFORMAT, CS8187.REFORMAT, and

CSR8187.REFORMAT were merged by MRGCS87 to make CS87SLEP.INPUT in the

following way: the files were sorted by firm and year; the primary file was CS8187, which was merged with the OLD files so that only one firm-year was taken for each observation, but with the most recent data (in terms of vintage of the tape source). The file CSR8187 was then used to update this data in such a way that no missing values were allowed to replace existing data. This maximizes data availability for each firm. I ended up with about 73,000 observations, of which about 50,000 contained good data on most of the variables. This means that there are on average about 17 years of data per firm (out of 30 years possible at the maximum).

This file still needed to be updated for name and CUSIP changers; this was done (twice, since some firms changed names twice) by the jobs CS87NCR and CS87NCR2, which changed all the CUSIPs on the name changers to the most recent and then merged the data. The old names were saved on the early observations so such a splicing can still be detected. The end result was a file called CS87SLEP.INPUT with 49,225 good observations (since some duplicate observations were deleted after the merging). This file was reformatted back into a panel with 30 years of data per firm for use by the SLEPIAN program.[3] The final yield was 2726 firms (2897 less 171 name changers).

The file CS87SLEP.INPUT was processed by the SLEPIAN program in order to calculate the book values of the capital stock components adjusted for inflation, the market value of the common stock, preferred stock, and long term debt, and the stock of R&D capital. The list of the new variables created by SLEPIAN is shown in Table 4.2; the output file from this program (with the same 49,225 good records plus 32,795 phony records which

---

[3] This program owes its name to Arthur Slepian, a graduate student at Stanford University who wrote the first version for Brainard, Shoven, and Weiss in 1978-79.

had been added to create a balanced panel) was called SLEPIAN.OUT.

This file was reformatted slightly by deleting the 32,795 bad records and converting to fixed format in order to create the R&D Panel Master file described in this document. A SAS file containing SAS missing value codes was created at the same time.

TABLE 4.1

SOURCES OF THE MASTER FILE DATA

File	Year	Number of Firms (primary)	(secondary)
Annual Industrial	1978	566	813
	1981	25	
	1983	19	
	1984	34	
	1985	28	
	1986	84	
	1987	1134	
	TOTAL	1890	
Over-the-Counter	1980	218	277
	1981	10	
	1983	9	
	1984	11	
	1985	36	
	1987	490	
	TOTAL	774	
Research Industrial	1978	131	35
	1981	0	7
	1984	5	2
	1986	0	13
	1987	97	713
	TOTAL	233	
GRAND TOTAL		2897	

The number of firms shown in the primary column is the number for which the file is the original source of the firm data; this defines the sample. The number in the secondary column is the number of firms for which additional years of data were obtained from that file and spliced into the main observation.

The first file in each section (denoted 1978, 1980, and 1978 respectively) were actually merged versions of the older Compustat files constructed for the earlier master file. The ones for the Industrial and Research samples contained data for 1959 through 1981, and the one for OTC contained data for 1960 through 1982.

TABLE 4.2

VARIABLES ADDED BY THE SLEPIAN PROGRAM

Number	Name	Description
76	NETCAP	Inflation-adjusted value of the net capital stock, = NPLANT + ADJINV + ADJTOT
77	NPLANT	Net capital stock, adjusted for effects of inflation
78	ADJINV	Inventories, adjusted for effects of inflation and inventory method
79	ADJTOT	Investments in unconsolidated subsidiaries and others, and intangibles, adjusted for inflation
80	TOTAL	Book value of ADJTOT (= TOTAL1+ TOTAL2 + INTANG)
81	VAL	Market value of firm (= VCOMS + PREFDCAP + LTDEBT + STDEBT ----- ADJ)
82	GROCAP	Gross capital stock, adjusted for effects of inflation
83	GRATE	Gross rate of return [= (INCOME + DEPREC + INTRST ----- PRIME ----- ADJ ----- IVA)/GROCAP]
84	LTDEBT	Long-term debt (market value).
85	ADJ	Adjustments to market value (= STASST ----- BKINV ---- STLIAB + STDEBT)
86	BKCAP	Book value of the capital stock (= BKPLNT + BKINV + TOTAL)
87	NRATE	Net rate of return [= (INCOME + DEPREC + INTRST ----- PRIME ADJ ----- IVA ----- ADJDEP)/NETCAP]
88	ADJDEP	Depreciation, adjusted for inflation (as NPLANT adjustment)
89	VCOMS	Value of common stock at close of year
90	VCOMS2	Value of common stock (average of low and high during year)
91	RSTOCK	Stock of R&D investment (knowledge capital)
92	RDINIT	Startup value of R&D stock

REFERENCES

Blue Book of Canadian Business. Toronto: Canadian Newspaper Services
    International, 1988.

Brainard, W., J. Shoven, and L. Weiss. "The Financial Valuation of the
    Return to Capital." Brookings Papers on Economic Activity 2(1980).

Canadian Trade Index. Toronto: Canadian Manufacturers' Association,
    1989.

Capital Changes Reporter. 7 vols. Chicago: Commerce Clearing House, 1989.

Directory of Companies Required to File Annual Reports with the Securities and
    Exchange Commission. Washington, D.C.: U.S. Government Printing
    Office, 1988.

Directory of Corporate Affiliations. Wilmette, IL: National Register
    Publishing Co., 1972, 1976, 1989.

Directory of Obsolete Securities. 1983, 1987, 1989.

Hall, Bronwyn H., Clint Cummins, Elizabeth Laderman, and Joy Mundy. The
    R&D Master File: Documentation. NBER Technical Working Paper no. 72.
    1988.

Jaffe, S.A. "A Price Index for Deflation of Academic R&D Expenditures."
    NSF Bulletin 72-310, 1972.

Key British Enterprises. Multiple volumes. London: Dun & Bradstreet, 1988.

Lehn, Kenneth, and Annette Poulsen. 1989. "Free Cash Flow and Stockholder
    Gains in Going Private Transactions." Journal of Finance 44: 771-88.

Mergers and Acquisitions, The Journal of Corporate Venture. McLean,
    Virginia. Vol. 8-11.

National Science Foundation. Research and Development in Industry,
    NSF 80-307. Washington, D.C.: GPO, 1978.

National Science Foundation. Science Indicators. Washington, D.C.:
    GPO, 1985.

National Stock Summary. Jan. 1, 1989 edition. Jersey City, NJ: National
    Quotation Bureau, 1989.

San Miguel, Joseph. 1977. "The Reliability of R&D Data in Compustat
    and 10-K Reports." Accounting Review LII: 638-41.

Standard and Poor's Register: Corporations. Vol. 1. New York:
    McGraw-Hill, 1989.

U.S. Dept. of Labor. Productivity Costs in Nonfinancial Corporations,
    various issues. Washington, D.C.: GPO.

Wall Street Journal Index. Princeton, N.J.: Dow Jones Books.

Ward's Business Directory of U.S. Private Companies. Vol. 1. 1988.

Who Owns Whom: Continental Europe. London: Dun & Bradstreet, 1988.

Who Owns Whom: North America. London: Dun & Bradstreet, 1988.

Who Owns Whom: United Kingdom and Republic of Ireland. 2 vols. London: Dun
    & Bradstreet, 1988.

EXITLIST Documentation:

Exits from the Publicly Traded Manufacturing Sector 1977-1987

The EXITLIST dataset contains information on all the firms in the manufacturing sector which exited from the Compustat Annual Industrial and OTC Files between 1977 and 1987 (11 years of exits). The data which are supplied are the reason for exit, the actual year of the exit, and the name and CUSIP of the acquiring or other new firm if available. The data are contained in an ASCII file on floppy diskette, one line per exit, sorted by CUSIP number of the exiting firm. A complete listing of the variables on the file is given on page 10.

The file was constructed in the following way: A list of firms which were ever in manufacturing between 1976 and 1987 was constructed by choosing firms from the old R&D Panel (the 1985 edition; see Hall et al. 1988), and augmenting this with any other firms which were ever in SIC 2000 through 3999 and which existed on the Annual Industrial files for 1981 or 1983 through 1987. From this combined and merged list I selected any firms which had no data in 1987 and could thus be presumed to have exited for some reason. This yielded a list of approximately 1200 firms which existed on the files sometime between 1976 and 1986 but were not there in 1987, out of a total panel of approximately 3000 firms. I also retained the year in which each of the firms last had good data, and for some of them, the year and reason for exit which is given on the Compustat Research file.[1]

---

[1] Since this file contains only firms deleted from the Industrial file,

Using this list, I then investigated more thoroughly in the library the reasons for exit. The first source was the Directory of Obsolete Securities (1983, 1987, 1989), followed by the Capital Changes Reporter (various dates through 1989). Both sources are extremely good at identifying final acquirers (rather than just offers) as well as the price paid for common. They also record Chapter 7 and Chapter 11 filings (both of which I treat as bankruptcy, even though the firm may emerge from Chapter 11 a few years later and reappear on Compustat) and liquidations (the breaking up of the assets of the firm, with the shareholders receiving the proceeds). Finally, they give the new names when the firm changes names (leading to a change of CUSIP, and hence an "exit" from Compustat), and they describe any recapitalizations or reorganizations which may have occurred.

Neither of these sources clearly identifies the type of acquirer, whether publicly traded, private, or foreign. Nor do they state clearly whether the transaction was leveraged, although one can often tell by the type of deal involved. To obtain this information, I used the following strategies:

1) I identified publicly traded firms by finding them in the Compustat files and recording their CUSIPs (this enables one to match data for target and buyer later on).

2) I identified private acquisitions from the list published by Lehn and Poulsen (1989), and from various directories of public and private corporations (Directory of Corporate Affiliations, Key British Enterprises, etc.).

---

coverage after exit is incomplete for the OTC firms in the Compustat data.

3) I identified LBOs from the list supplied to me by Steven Kaplan (Kaplan 1989) and from the Merger and Acquisition Sourcebook, which identifies the type of transaction. This left me with quite a few acquirers unassigned; those with PLC (the U.K.), AG (West Germany), NV (Netherlands), AB (Sweden), SA (France or Switzerland), SpA (Italy), LTD (the U.K. or Canada), or obvious foreign names were identified as foreign, although if they had a CUSIP it was retained (some are traded directly or as ADR shares on the U.S. stock exchanges and thus have data available). I then looked up the remainder again in various Who Owns Whom directories; any subsidiaries were recoded to the name of the parent firm. This yielded positive identification of public and foreign firms; any remaining puzzles (very few) I coded as private. Public acquiring firms which are not on Compustat or CRSP have a CUSIP equal to 0 on my list.

4) Finally, I verified that all acquirers and new names which had CUSIPs were in my sample. Some were not, mostly because the firms had changed industries and become non-manufacturing. I added these firms to my sample so that data from them would be available for post-merger analysis, and I checked that these new firms had not also exited; for those which had, steps one through four were repeated until I had all possible acquirers.

The final list of exits contains 1224 firms and 6 variables. The data on the list are the following:

1. CUSIP -- The six-digit CUSIP of the firm which exited.

2. CNAME -- The name of the firm which exited.

3. ACQYR -- The year in which the actual acquisition, bankruptcy, liquidation, or name change took place. This is not necessarily the year in which the firm exited from the file; frequently it is later. There will be exits up until 89, since data may disappear much earlier from Compustat (and my sample is all firms which do not have data in

87 and possibly earlier years).

4. ACQRSN -- A code telling why the firm exited from the file:

M   The firm was acquired by another publicly traded firm.
P   The firm went private (or was purchased by a privately held firm).
PL The firm went private in a leveraged buyout.
F   The firm was acquired by a foreign entity.
C   The firm is controlled by another publicly traded firm and
    has been deleted from the file by Compustat.

B   The firm went bankrupt (no real distinction between Chapter 7
    and Chapter 11 in my data).
L   The firm was liquidated.

R   The firm underwent reorganization and may have later reappeared.
    (This includes some bankruptcies).
NC The firm changed its name (and hence its CUSIP). The new name
    and CUSIP are shown as the acquiring firm.

NO After much investigation, no reason for exit was found and the firm
    is still in existence.
D   The firm is still in existence and still (at least to a certain
    extent) publicly traded, but it has been delisted from an exchange
    and Compustat no longer carries its data.

5. ACUSIP -- The CUSIP of the acquiring or liquidating firm, or
new firm if a name changer or recapitalizer.

6. ANAME -- The name of the acquiring, liquidating, or
reorganizing firm. Lower case entries provide other information about
the exit (or give generic purchasers, such as "employees").

APPENDIX B

The Construction of the R&D Stock Variable:

Interpolating the Missing Values of R&D

To construct a variable that measures the stock of R&D capital owned by a firm, we use a method due to Griliches (1981; Griliches and Mairesse 1981, Griliches and Hall 1982). This method is based on a standard perpetual inventory equation with declining balance depreciation:

(1) $K_t = (1 - \delta) K_{t-1} + R_t$

where K_t is the end-of-period stock of R&D capital and R_t is the (real) expenditures during the year. The depreciation rate δ is chosen to be 15 percent per year; Griliches and Mairesse found that the exact choice of depreciation rate made little difference in production function estimates. This is not surprising since, if R&D expenditures are roughly constant in real terms, the stock of R&D capital is

$$K_t = \sum_{s=0}^{\infty} (1 - \delta)^s R_{t-s} = R/\delta$$

The variation across firms will then be approximately the same, regardless of the value of δ and the magnitude of the coefficient will just vary inversely with δ. This means that separate identification of δ and the coefficient of K_t in an equation will be difficult.

Two missing data problems must be confronted when making a stock out of a series of flow variables: first, the problem of initial conditions for the stock, and second, the fact that a single missing value for R&D in one year will cause all the associated stock variables to be missing. We solve the first problem by setting the initial stock to the R&D expenditures in the first year divided by the sum of the depreciation rate δ and a presample growth rate of new R&D of 8 percent per year. Thus the individual stock is approximately four times the level of R&D in the first year. The second problem is solved as described below, by interpolation where there are only one or two missing values in an R&D series. This procedure affects relatively few firms.

Assume that we observe R_t and R_{t+s} (with s>1) but not R_{t+1}, R_{t+2}, ..., R_{t+s-1}. $\{R_t\}$ is hypothesized to follow a random walk. How should we forecast R_{t+1}, R_{t+2}, ..., R_{t+s-1} given values of R_t and R_{t+s}? The unbiased estimator is

$$E(R_{t+i} \mid R_t, R_{t+s}) = R_t + E(\epsilon_{t+1} + \ldots + \epsilon_{t+i} \mid (R_{t+s} - R_t))$$

$$= R_t + E(\epsilon_{t+1} \mid \sum_{j=1}^{s} \epsilon_{t+j}) + E(\epsilon_{t+s} \mid \sum_{j=1}^{s} \epsilon_{t+j}) + \ldots + E(\epsilon_{t+i} \mid \sum_{j=1}^{s} \epsilon_{t+j})$$

$$= R_t + \sum_{k=1}^{i} E(\epsilon_{t+k} \mid \sum_{j=1}^{s} \epsilon_{t+j})$$

where $\epsilon_{t+k} \sim$ IIN $(0, \sigma^2)$.

So we need to compute the conditional expectation of the disturbance ϵ_{t+k}, conditioned on a sum of s such disturbances, where the sum includes the $\{\epsilon_{t+k}\}$ in which we are interested. (Note that

$E(\epsilon_{t+k} | \sum\limits_{j=1}^{s} \epsilon_{t+j}) = 0$, where $k \in [1,s])$.)

To perform this computation, I write the density of ϵ_{t+j} and $\Sigma\epsilon_{t+k}$ as a bivariate normal with zero means and covariance matrix equal to

$$\begin{bmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & s\sigma^2 \end{bmatrix}$$

so that the conditional density is

$$f(\epsilon_{t+k} | \sum\limits_{j=1}^{s} \epsilon_{t+j}) = N(\sum\limits_{j=1}^{s} \epsilon_{t+j}/s, \ \sigma^2/s)$$

by the formula for the conditional bivariate normal, and hence

$$E(\epsilon_{t+k} | \sum\limits_{j=1}^{s} \epsilon_{t+j}) = \sum\limits_{j=1}^{s} \epsilon_{t+j}/s.$$

This, in turn, implies that the optimal forecast of R_{t+i} is

$$R_{t+i}^{*} = R_{t} + (i/s)(R_{t+s} - R_{t})$$

or $\quad R_{t+i}^{*} = R_{t+s} + ((s-i)/s)(R_{t} - R_{t+s})$,

which is symmetric in the two endpoints as desired.

Now make the more realistic assumption that $r_{t} = \log R_{t}$, the natural logarithm of R&D expenditures follows a random walk instead, i.e.,

$$r_t = r_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma^2) \qquad t = \ldots -2, 1, 0, 1, 2, \ldots$$

From the preceeding, we know that

$$E(r_{t+i} \mid r_t, r_{t+s}) = r_t + (k/s)(r_{t+s} - r_t)$$

is the optimal forecast of r_{t+i}. To obtain the optimal forecast of R_{t+i} we must include the variance of e_t:

$$E(R_{t+i} \mid r_t, r_{t+s}) = \exp\left[r_t + i/s (r_{t+s} - r_t) + \sigma^2 \right]$$

To estimate σ^2, the variance of the shock, we use the method of moments and the observed growth rates in R_t:

$$\hat{\sigma}^2 = (T-1)^{-1} \left(\sum_{\tau=1}^{t-1} e_\tau^2 + \sum_{\tau=t+s}^{T} e_\tau^2 + \left(s^{-1} \sum_{\tau=t}^{t+s-1} e_\tau \right)^2 \right)$$

where the last term uses the information contained in the size of the jump over which we are trying to interpolate. e_t denotes $r_t - r_{t-1}$, the data estimate of ϵ_t. Given an estimate of σ^2, we can form an unbiased forecast of R_{t+i} using all the available data:

$$R_{t+i} = \exp\left[r_t + (i/s)(r_{t+s} - r_t + \hat{\sigma}^2) \right]$$

or

$$R_{t+i} = (R_t)^{1-(i/s)} \ (R_{t+s})^{i/s} \ \exp\left[(i/s) \ \hat{\sigma}^2 \right]$$

After interpolation to obtain a continuous R_t series and computation of an initial condition K_0, the stock series K_t is formed as in equation (1).