

Creating a Dataset of Tax Reconciliation Data

Historic Tax Information Extraction and Parsing

Dingjun Bian, Guangqi Li

Supervised by Anmol Bhandari and Ellen McGrattan



June 2, 2021

REPORT *APRIL 2021*



55 Corporations Paid \$0 in Federal Taxes on 2020 Profits

MATTHEW GARDNER AND STEVE WAMHOFF

Introduction

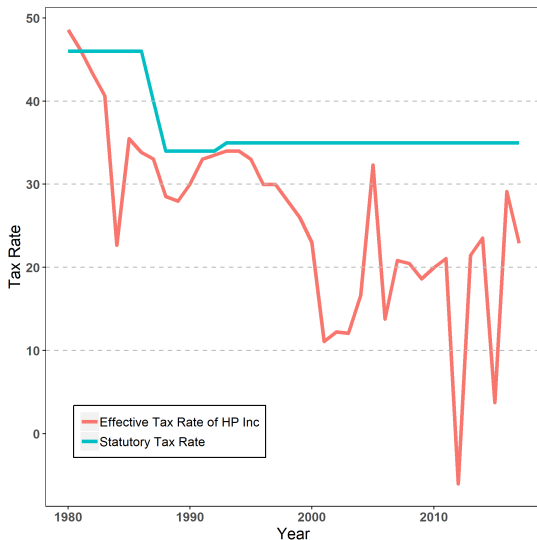


Figure: Statutory Tax Rate v.s. Effective Tax Rate of HP Inc.

Introduction

	2011	2010	2009
U.S. federal statutory income tax rate	35.0%	35.0%	35.0%
State income taxes, net of federal tax benefit	0.5	1.3	0.9
Lower rates in other jurisdictions, net	(24.0)	(18.3)	(12.2)
Research and development credit	(0.6)	(0.1)	(0.5)
Foreign net operating loss	—	—	(4.1)
Valuation allowance	5.2	0.8	(0.6)
Accrued taxes due to post-acquisition integration	—	—	0.6
Nondeductible goodwill	3.4	—	—
Other, net	1.7	1.5	(0.5)
	<u>21.2%</u>	<u>20.2%</u>	<u>18.6%</u>

Figure: Tax Reconciliation Table of HP Inc in 2011.

- Financial reports (10-K) in:
 - Mergent Archives: PDF, 1980-1995
 - SEC Edgar: HTML and TXT, 1994-2020

- Challenge:
 - How to locate tax reconciliation table.

Financial Reports in PDF

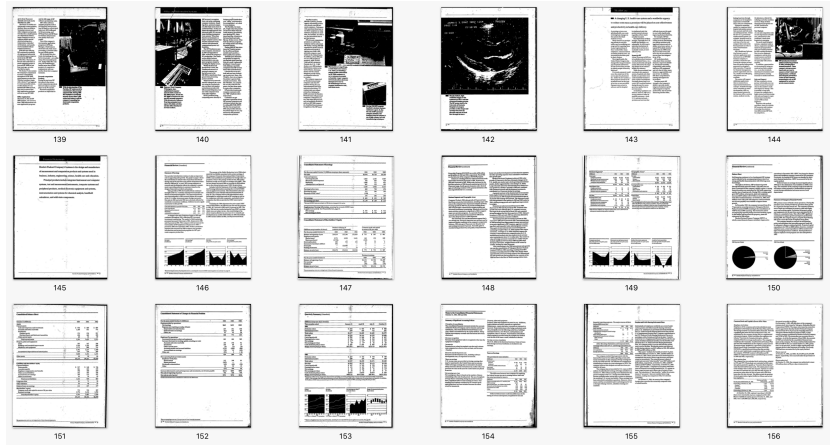


Figure: Part of the Financial Report of HP Inc in 1984.

The difference between taxes computed by applying the **federal income tax rate** to earnings before taxes and the actual provision for taxes is:

(Millions)	1984	1983	1982
Taxes on earnings at the U.S. statutory rate	\$ 395	\$335	\$311
DISC earnings	(145)	—	—
State income taxes, net of federal tax benefit	20	19	16
Research and development tax credits	(30)	(22)	(15)
Investment tax credits	(17)	(16)	(15)
Other	(28)	(20)	(4)
	\$ 195	\$296	\$293

Figure: Tax Reconciliation Tables of HP Inc in 1984.

- Challenge:
 - How to locate tax reconciliation table.
 - How to extract tax reconciliation table.

Notes to the Consolidated Financial Statements

October 31, 1984, 1983 and 1982

Summary of Significant Accounting Policies

Principles of consolidation

The consolidated financial statements include the accounts of Hewlett-Packard Company and its domestic and foreign subsidiaries, other than Hewlett-Packard Finance Company, which is accounted for by the equity method. All significant intercompany accounts and transactions have been eliminated.

Revenue recognition

Revenue from equipment sales is recognized at the time the equipment is shipped.

Inventories

Inventories are valued at standard costs that approximate costs computed on a first-in, first-out basis, not in excess of market.

Research and development costs

Research and development costs, including software development costs, are expensed as incurred.

Taxes on earnings

U.S. income taxes are provided on foreign earnings that may be repatriated to the United States and are not provided on foreign earnings that are intended to be indefinitely reinvested abroad. Investment tax credits reduce the provision for taxes in the year the related assets are placed in service.

Net earnings per share

Net earnings per share is based on the number of shares outstanding at the end of each period. The use of weighted-average shares outstanding during the period would not have a significant effect on net earnings per share. Outstanding stock options considered to be common stock equivalents have not been included because the effect would be immaterial.

Property, plant and equipment

Property, plant and equipment is stated at cost. Additions, improvements and major renewals are capitalized.

Maintenance, repairs and minor renewals are expensed as incurred. Depreciation is provided using accelerated methods, principally over the following useful lives: buildings and improvements, 15 to 40 years; and machinery and equipment, three to 10 years. Amortization of leasehold improvements is provided using the straight-line method over the life of the lease or asset, whichever is shorter.

Foreign currency translation

The U.S. dollar is the functional currency. Gains or losses from foreign currency translation are included in net earnings.

Taxes on Earnings

The provision for taxes includes:

(Millions)	1984	1983	1982
Federal taxes:			
Current	\$ 140	\$ 61	\$ 90
Deferred	37	105	84
Reversal of DISC deferred taxes	(118)	—	—
State taxes	37	35	30
Foreign taxes	99	95	89
	\$ 195	\$ 296	\$ 293

The difference between taxes computed by applying the federal income tax rate to earnings before taxes and the actual provision for taxes is:

(Millions)	1984	1983	1982
Taxes on earnings at the U.S. statutory rate	\$ 395	\$ 335	\$ 311
DISC earnings	(145)	—	—
State income taxes, net of federal tax benefit	20	19	16
Research and development tax credits	(30)	(22)	(15)
Investment tax credits	(17)	(16)	(15)
Other	(28)	(20)	(4)
	\$ 195	\$ 296	\$ 293

Deferred federal taxes result from differences in the timing of revenue and expense recognition for tax and

Financial Reports in PDF

(Millions)	1984	1983	1982
Taxes on earnings at the U.S. statutory rate	\$ 395	\$335	\$311
DISC earnings	(145)	—	—
State income taxes, net of federal tax benefit	20	19	16
Research and development tax credits	(30)	(22)	(15)
Investment tax credits	(17)	(16)	(15)
Other	(28)	(20)	(4)
	\$ 195	\$296	\$293

(Millions)	1984	1983	1982
Taxes on earnings at the U.S. statutory rate	\$395	\$335	\$311
DISC earnings	-145	-	-
State income taxes, net of federal tax benefit	20	19	16
Research and development tax credits	-30	-22	-15
Investment tax credits	-17	-16	-15
Other	-28	-20	-4
	\$195	\$296	\$293

Figure: Tax table in PDF and extracted tax table in CSV (Amazon Textract)

- Challenge:
 - How to locate and extract the tax reconciliation table (Fasttext)
 - How to parse the extracted tax reconciliation table into usable data

Financial Reports in HTML

[illegible]

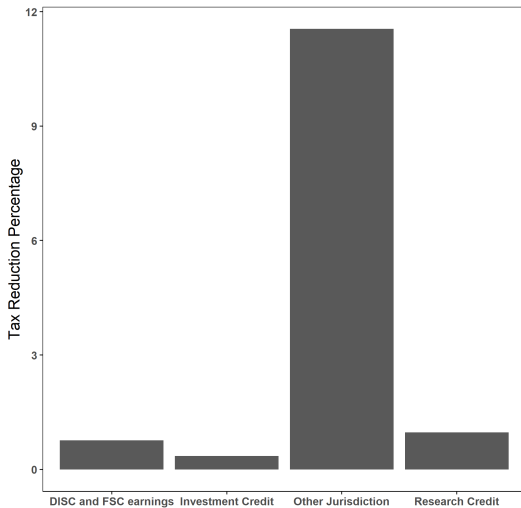
Figure: Raw HTML String from Fasttext

Financial Reports in HTML

	Unnamed: 0 cik	year_10k	Item	N_row	i_row	Years	value	Type
6643	0	47217	2008 u.s.federalsta	9	0	2008	35 %	
6644	1	47217	2008 stateincomef	9	1	2008	1.3	
6645	2	47217	2008 lowerratesin	9	2	2008	-16.9	
6646	3	47217	2008 researchhandi	9	3	2008	-0.4	
6647	4	47217	2008 valuationallo	9	4	2008	-1.7	
6648	5	47217	2008 u.s.federalta	9	5	2008	-7.9	
6649	6	47217	2008 accruedtaxes	9	6	2008	2	
6650	7	47217	2008 other,net	9	7	2008	-0.5	
6651	8	47217	2008 20.5####%	9	8	2008	20.8 %	
6652	9	47217	2008 u.s.federalsta	9	0	2007	35 %	
6653	10	47217	2008 stateincomef	9	1	2007	0.5	
6654	11	47217	2008 lowerratesin	9	2	2007	-13.2	
6655	12	47217	2008 researchhandi	9	3	2007	-0.6	
6656	13	47217	2008 valuationallo	9	4	2007	-1	
6657	16	47217	2008 other,net	9	7	2007	0.8	
6658	17	47217	2008 20.5####%	9	8	2007	13.8 %	
6659	18	47217	2008 u.s.federalsta	9	0	2006	35 %	
6660	19	47217	2008 stateincomef	9	1	2006	-0.1	
6661	20	47217	2008 lowerratesin	9	2	2006	-11.9	
6662	21	47217	2008 researchhandi	9	3	2006	-0.2	
6663	25	47217	2008 other,net	9	7	2006	-0.1	
6664	0	47217	2007 u.s.federalsta	9	0	2007	35 %	
6665	1	47217	2007 stateincomef	9	1	2007	0.5	
6666	2	47217	2007 lowerratesin	9	2	2007	-13.2	
6667	3	47217	2007 jobsactrepata	9	3	2007	22.4	
6668	4	47217	2007 researchhandi	9	4	2007	-0.6	
6669	5	47217	2007 valuationallo	9	5	2007	-1.7	
6670	6	47217	2007 u.s.federalta	9	6	2007	-7.9	
6671	7	47217	2007 other,net	9	7	2007	0.8	
6672	8	47217	2007 20.8####%	9	8	2007	13.8 %	
6673	9	47217	2007 u.s.federalsta	9	0	2006	35 %	
6674	10	47217	2007 stateincomef	9	1	2006	-0.1	
6675	11	47217	2007 lowerratesin	9	2	2006	-11.9	

Figure: Parsing Result

Statistics on HP Inc.



More Potentials

Possibilities:

- Look across the firms and observe their tax minimization strategies
- Look at a single tax deduction term through time and observe its trend based on historic events
-

Reference:

- Gardner, M.; Wamhoff, S. (2021, April). 55 corporations paid \$0 in federal taxes on 2020 PROFITS. Retrieved May 11, 2021, from <https://itep.org/55-profitable-corporations-zero-corporate-tax/>

Thank you for listening!

Dingjun Bian (bianx068@umn.edu)
Guangqi Li (li001122@umn.edu)