# WRDS Index Data Extraction Methodology 

J.A. Dobelman, H.B. Kang*, and S.W. Park*

January 2014

Technical Report TR-2014-01<br>Department of Statistics<br>Rice University


#### Abstract

This paper provides and validates an automatic procedure to generate accurate CRSP PERMNOs from Compustat GVKEYs for historical index constituents. We then validate the resulting PERMNO lists and examine some of the many pitfalls in other attempts to accomplish this, and provide cautionary guidance for WRDS index data researchers.


KEYWORDS: CRSP, Compustat, GVKEY, Historical Index Constituent, Indexes, WRDS Databases, PERMNO

[^0]THIS PAGE INTENTIONALLY BLANK

## Contents

1. Introduction ..... 5
2. Overview of WRDS Index Data ..... 6
3. Accurate Historical Components for Calibration ..... 6
4. WRDS database ..... 7
4.1 WRDS ..... 7
4.2. CRSP/Compustat Merged (CCM) Database. ..... 8
What is it? ..... 8
Linking Table [13] ..... 8
5. Ways to Obtain S\&P 500 from WRDS ..... 9
5.1 S\&P 500 historical constituents From WRDS Indexes ..... 9
5.2 Compustat Price, Dividend and Earnings (PDE) Monthly ..... 9
5.3 IHS Global Insight ..... 9
5.4 DSP, MSP lists from the CRSP using CRSPAccess Tool ..... 9
5.5 Direct Download of DSP500 ..... 10
6. Data Extraction-DSP500 for S\&P 500 ..... 10
6.1 What is the DSP ..... 10
6.2 Validation of DSP500 with True Constituents ..... 10
7. Data Crosswalks and the CRSP/Compustat Merged (CCM) Linking Table ..... 12
7.1 COMPUSTAT GVKEY $\rightarrow$ PERMNO ..... 12
7.2 COMPUSTAT GVKEY $\rightarrow$ PERMCO ..... 16
7.3 COMPUSTAT (GVKEY \& IID) $\rightarrow$ PERMNO ..... 16
7.4 Comparison among Conversion Methods ..... 17
8. Exploratory Enumeration of Compustat Historical Components ..... 17
8.1 Data Availability and Limitations ..... 17
8.2 Validation Result for Other Indices ..... 19
DJIA ..... 19
S\&P 100 ..... 19
NASDAQ 100. ..... 20
8.3 Summary ..... 20
9. Another Way to Obtain NASDAQ 100 ..... 20
10. Conclusions and Discussion ..... 21
11. References ..... 22
Appendix 1. Corporate Archeology Vignettes ..... 24
12. Burlington \& Berkshire ..... 24
13. Harrah's Entertainment \& Caesars Entertainment ..... 24
14. Viacom \& CBS Corp ..... 24
15. Clear Channel Communications \& CC Media Holdings ..... 24
16. SBC Communications \& AT\&T ..... 24
17. Boise Cascade \& Office Max ..... 24
18. Invitrogen \& Life Technologies Corp ..... 24
19. NTL Inc \& Virgin Mobile ..... 24
20. Bell Atlantic and Verizon ..... 25
21. Coastal Corporation ..... 25
22. K-mart ..... 25
23. Monsanto Company ..... 25
24. Northern Telecom ..... 25
Appendix 2. NASDAQ 100 Inclusion/Exclusion Table ..... 26

## 1. Introduction

The Dow Jones Industrial Average (DJIA), Standard \& Poor's (S\&P) 500 (SPX), S\&P 100 (OEX), or NASDAQ 100 (NDX) are well-known market indices which capture the market levels and dynamics and thus are used in many financial analyses. Many researchers are therefore motivated to conduct analyses on components of these market indices, including checking the levels, estimating model parameters or future prices based on the historical prices, or developing a new portfolio that gives optimal return. Regardless of what is believed about the efficacy of such activities, an accurate list of constituents is required. However, it is not an easy matter to obtain historically accurate index constituents at a low cost. One can typically arrange with an index provider to purchase historical constituent lists, but the cost for this is non-trivial, often in the thousands of dollars for multiple data extraction dates.

Consequently, much time is wasted by the naïve researcher searching for accurate historical constituent lists. There are a few lists on the Internet for certain dates in the past, but we generally need continuous lists from 1970s or 1980s, or earlier if possible. Many promising web links in fact point to obsolete university research sites which reference incorrect or inaccurate information, or resources which do not in fact provide historical constituents. There are many sources that give alternative methods to obtain the constituent lists; typical suggestions from these websites are to visit the index websites themselves, or to use proprietary software/data sources such as Datastream, Bloomberg, Factset, CapitalIQ, etc. For academic researchers, the Wharton Research Database Services (WRDS) is the resource of choice. It is beyond the scope of this paper to investigate the limitations of the other data sources; and since our institution subscribes to many of the WRDS databases, we limit our discussion to that resource.

Our motivation is to obtain pricing data on historical index constituents. Things get somewhat complicated due to the number of databases available on WRDS. However, there are two canonical databases which many researchers use for this purpose, i.e., Compustat and Center for Research in Security Prices (CRSP). The former is primarily a repository of accounting and metadata for individual stocks and indexes, and the latter is a pricing database for individual stocks and indexes. While each database is remarkable clean, unfortunately they do not use the same key variables; the key variable for stocks on Compustat is GVKEY, while that on CRSP is PERMNO. In order to perform financial and statistical analysis on levels and returns, we require the CRSP PERMNO for index constituents, which in turn requires a crosswalking method to map from Compustat GVKEY to CRSP PERMNO.

The WRDS site provides a frequently asked question (FAQ) list and other information on obtaining historical constituents for the S\&P 500 [1,2]. The FAQ suggests six ways to obtain various S\&P 500 information, including constituents. Each of these was investigated for purposes of obtaining historical constituents; the only resource that ultimately provided truly accurate historical constituents with PERMNO were the SAS files dsp500list. sas7bdata and msp500. sas7bdat files available by file transfer protocol (ftp) on the WRDS Unix system; and alas, there are not similar files for the DOW, OEX, NDX, etc.

Since we are interested in more than just the SPX, our main option is to use the WRDS webbased GUI interface and the Compustat Annual Updates - Index Constituents, but this approach brings with it several challenges outlined below. The remainder of the paper discusses an overview of the WRDS databases as they pertain to extracting historical constituents, how to clean and extract the constituents, as well as an account of some of the traps and difficulties in this process. We conclude with a procedure which generates the GVKEY crosswalk needed to capture the CRSP PERMNO for historical constituents for the desired stock index.

## 2. Overview of WRDS Index Data

Wharton Research Database Services (WRDS) provides access to many databases, which include Compustat North America from S\&P Capital IQ and CRSP from University of Chicago Booth. CRSP covers stock market pricing data on major stock exchanges (NYSE, AMEX, and NASDAQ) while Compustat covers basically accounting data for public, OTC and private companies. Since Compustat's main identifier is permanent company identifier GVKEY while CRSP's main identifiers are permanent company and security identifiers PERMNO and PERMCO, the need to link the two databases emerges. This need created CRSP/Compustat Merged Database (CCM) [3], which is included in CRSP.

However, CCM contains only Compustat data items, which can be searched by CRSP's PERMNO and PERMCO in addition to Compustat's GVKEY. What we need is the CRSP data items from Compustat's GVKEY. Thus, merging the data with the CRSP stock data requires additional steps. It may be possible to accomplish the task on WRDS through UNIX and SAS/FORTRAN programming, but our goal is to exploit the WRDS Web GUI for most tasks. Therefore, we provide our own method to accomplish this.

Compustat provides historical index constituents via their "Compustat Annual Updates - Index Constituents" database. This file provides much historical index constituent data on over 1,050 different indexes. Examples of some of the more popular indexes and their Compustat "ticker" symbols are in the table below. Note that the historical index constituents record is NOT complete for many indexes which the research workers might wish to use.

| Ticker | Index Name | Data Date |
| :--- | :--- | ---: |
| I0003 | S\&P 500 Comp-Ltd | $3 / 11 / 64$ |
| I0005 | Dow Jones Industrials-30 Stk | $3 / 17 / 97$ |
| I0006 | Dow Jones Transportation-20 | $1 / 1 / 1900$ |
| I0007 | Dow Jones Utilities-15 Stk | $1 / 1 / 1900$ |
| I0014 | S\&P 100-Ltd | $9 / 11 / 89$ |
| I0016 | S\&P Midcap 400 Index | $6 / 1 / 91$ |
| I0019 | S\&P Smallcap 600 Index | $10 / 1 / 94$ |
| I0020 | S\&P 1500 Super Composite | $12 / 30 / 94$ |
| I0028 | Nasdaq 100 | $11 / 5 / 04$ |

Further information about the WRDS database can be found in section 4, WRDS Database.

## 3. Accurate Historical Components for Calibration

We will discuss several ways to obtain the constituent lists. To confirm that this source of historical components is valid, we compare the lists obtained from the source with the true historical lists which we refer to as the calibration data on certain reference dates.

The true historical lists were also obtained in several ways, using various online sources ${ }^{1}$. During routine market monitoring in the past, we captured constituent data from primary source websites; these were obtained and stored for the S\&P 500, S\&P 100, DJIA, and NASDAQ 100 for several dates. Below are the indices and reference dates for which true historical constituents are available for calibration.

[^1]```
S&P 500 04/25/99, 09/03/02, 12/13/03, 12/29/04
S&P 100 11/01/02, 12/01/03, 01/04/05, 09/25/05, 02/06/06, 07/23/10
DJIA 07/03/02, 04/08/04, 02/06/06, 12/04/08
NASDAQ 100 05/16/02, 09/14/02, 11/02/02, 12/01/03, 09/01/05
```

The lists in hand were formatted with symbol and company name, and sometimes market capitalization.

Additionally, historical component lists of S\&P 500 for $08 / 23 / 2006$ and 12/19/2007 were obtained from a Standard and Poor's website [4]. The site used to give the historical lists from about 2000, and we downloaded data in late September and early October 2011, but the link was found to have been terminated when it was checked again on December 8, 2011. The lists from Standard \& Poor's contained symbol/ticker, company name, country, GICS, Sector, and Price. Constituents from these downloaded historical lists were manually checked and they matched those of the calibration lists.

One can obtain the historical DJIA list because the Wikipedia article[5], and other sites[6], provide the historical components of DJIA on every date on which the list has been changed ${ }^{2}$. The sites just give the company names and date of the change.

The daily NASDAQ 100 (NDX) list can be currently obtained from Yahoo Finance and the NASDAQ webpage, but historical constituents of NASDAQ 100 could not be found without negotiating with the company. WRDS only contains NDX data since Nov. 2004, although the index was launched in January 1985. It was necessary to find a source that contains the NASDAQ constituents before 2004.

We tried to obtain true historical components from the 1970s-80s. We looked at microfilms of Wall Street Journal, Financial Times, Barron's, and Chicago Tribune, but most past newspaper contain the alphabetical order of the whole NASDAQ and NYSE, and S\&P companies, but not for the specific indexes such as S\&P 500 or OEX. We did find however that the NASDAQ website has an inclusion/exclusion table for the NASDAQ 100 since 1995 (see section 9). In general, newspapers did not produce the needed historical constituents of stock indexes. It is possible that other sources such as the Value Line Investment Survey or Ibbotson's publications might list these stocks, but this was an ad hoc approach; the remainder of this paper seeks automatic and repeatable procedures.

## 4. WRDS database

### 4.1 WRDS

In our research, the Wharton Research Data Service (WRDS) database collection was used. Since 1993, University of Pennsylvania developed the database that provides financial, economic and marketing data available on the internet. WRDS provides access to Compustat from Capital IQ, CRSP from University of Chicago Booth, IBES ${ }^{3}$, NYSE-TAQ and many others, but in this report only Compustat and CRSP, which contain daily historical constituents, were used [12].

Because CRSP and Compustat are constructed and managed by different companies[16], both databases have different permanent unique identifiers; CRSP identifies individual companies with PERMNO, while COMPSTAT identifies companies with GVKEY. It is necessary to mutually compare both historical constituents. PERMCO is CRSP's permanent company identifier and PERMNO is CRSP's permanent issue identifier; each company has a unique PERMCO, but can have multiple PERMNOs.

[^2]
### 4.2. CRSP/Compustat Merged (CCM) Database

## What is it?

The CRSP/Compustat Merged (CCM) Database is composed of CRSP and Compustat data, together with the link and link-history references between these two databases. It includes Standard \& Poor's Compustat data, reformatted into CRSP's proprietary CRSPAccess database format. The CRSP Link provides a matching of CRSP historical price, distribution, and total return data with Compustat fundamental data by associating identifiers that are unique to each database. CRSP Link - an array of data, linking permanent unique identifiers: CRSP's PERMNO and PERMCO and Compustat's GVKEY.

## Linking Table [13]

WRDS offers a linking table in CCM. Linking option needs to be fulfilled to get the data. The linking tables provide the basis for creating a crosswalk between GVKEY and PERMNO.

The LINKDT and LINKENDDT signify when the link begins and when the link ends. Currently trading stocks show ' $E$ ' for LINKENDDT. We have to regard this as "Last Updated On" date under "Variable Descriptions." The crosswalk we use most in this paper is last updated on 08/10/2011.

Moreover, LIID is the linking IID. IID is Compustat's permanent issue identifier. There is an identifying relationship between IID and GVKEY. One GVKEY can have multiple IIDs, and both should be used as a pair to properly identify a security.
There is also LINKPRIM, the primary issue marker for the link, defined as follows.
$\mathrm{P}=$ Primary, identified by Compustat in monthly security data.
$J=$ Joiner secondary issue of a company, identified by Compustat in monthly security data.
C = Primary, assigned by CRSP to resolve ranges of overlapping or missing primary markers from Compustat in order to produce one primary security throughout the company history.
$\mathrm{N}=$ Secondary, assigned by CRSP to override Compustat. Compustat allows the U.S and Canadian security to be both marked as Primary at the same time. For Purposes of the link, CRSP allows only one primary at a time and marks the others as N .
There are basically eight options for LINKTYPE. LC and LU are normally used because they are typically the most accurate.

LC - Link research complete. Standard connection between databases.
LU - Unresearched link to issue by CUSIP
LX - Link to a security that trades on another exchange system not included in CRSP data.
LD - Duplicate Link to a security. Another GVKEY/IID is a better link to that CRSP record.
LN - Primary link exists but Compustat does not have prices.
LS - Link valid for this security only. Other CRSP PERMNOs with the same PERMCO will link to other GVKEYs.
NR - No link available, confirmed by research
NU - No link available, not yet confirmed

## 5. Ways to Obtain S\&P 500 from WRDS

As mentioned in the introduction, WRDS provides various guidance documents and FAQ's on obtaining historical constituent data. We initially focus on the S\&P 500, in "General Information FAQs: S\&P 500 Data and Constituents" [1, 2]. In these FAQ's, recommendations are made for the researcher in obtaining historical data on the SPX, and these are summarized below, with a disposition for our purpose of obtaining historical constituent data (i.e., not pricing/returns data).

### 5.1 S\&P 500 historical constituents From WRDS Indexes

This appears to be a workhorse database for historical constituents. We first obtained data from WRDS - Compustat- North America - Index Constituents from Dec 2003 to Dec 2003 with TIC of i0003 (S\&P 500 Comp-Ltd). The query returned 500 GVKEYs.
Index constituents offer the entire stock information that existed on the specific duration or a month. The index name (CONM) shows what index the companies belong to. In addition, similar to the DSP500 list described in section 6, the index constituents give the date that the companies were included in the index and the date that the companies were excluded. For example, Abbott Laboratories has existed in the S\&P 500 list since March 31, 1964 to the last day of the month that we searched for. Unfortunately, this report does NOT include PERMNO's needed for pricing data available from CRSP.

| GVKEY | GVKEYX | from | thru | Conm | TIC | co_conm | co_tic |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 1078 | 3 | 19640331 |  | S\&P 500 Comp-Ltd | I0003 | ABBOTT LABORATORIES | ABT |
| 1300 | 3 | 19640331 |  | S\&P 500 Comp-Ltd | I0003 | HONEYWELL INTERNATIONAL INC | HON |

### 5.2 Compustat Price, Dividend and Earnings (PDE) Monthly

In Compustat's annual updates, Price, Dividend and Earnings' monthly format gives the S\&P 500 lists for the end date of each month, and all 500 lists were exactly the same as the S\&P 500 data from the first method. The lists from Dec 31, 2004, Nov 30, 2003, and Dec 31, 2006 were selected to validate the method. This second method only gives a limited data for the last dates that we can also obtain from the first method; therefore, this second method is impractical.

### 5.3 IHS Global Insight

This database is another source that contains the historical S\&P 500. IHS Global Insight is a database for economic forecasts and industry analyses. At present, our institution does not subscribe to this database. However, because this database is different from the historical list from Compustat and CRSP, this method has a potential to be valid and useful; most likely this could be CUSIP- or GVKEYbased, and PERMNO would not be immediately extractable.

### 5.4 DSP, MSP lists from the CRSP using CRSPAccess Tool

We could generate the DSP500 file (see section 6 below), which shows each company's inclusion and exclusion dates in the S\&P 500 lists. To obtain the list, it is necessary to use the CRSPAccess tool, which provides CRSP stock information and CCM databases. This, and other database tools, are found in The CRSP Utilities and Program Libraries Guide (CUPL)[14]. Here, the stkprint utility generates the S\&P 500 historical constituent list and gives Daily or Monthly Stock Product and the Index Product. At one time this utility had other groups besides the S\&P 500 available, but this has not been the case
since at least 2011 [11, p.13]. Example queries are provided in the CUPL, but the lack of ease with which this tool generates a DSP500 list bespeaks the efficacy of direct downloading of the DSP500 itself.

### 5.5 Direct Download of DSP500

Direct download of the DSP500 and MSP500 list is recommended in the FAQ's and is described in section 6 below.

## 6. Data Extraction-DSP500 for S\&P 500

### 6.1 What is the DSP

dsp500list.sas7bdata is a SAS file obtainable via ftp/ssh from WRDS. For the S\&P 500, it provides a list of all stocks by PERMNO which have been in the S\&P 500, along with their inclusion and exclusion dates. It is thus eminently useable for historic backtesting purposes. In obtaining the historical constituent list, we used a 2007-vintage dsp500list, herein referred to as DSP500. This was the file originally used in J.R. Thompson's groundbreaking research on the MaxMedian portfolio [8, 9], and was checked to be valid against the true S\&P 500 data for several reference dates. The DSP500 is formatted as below.

| $\underline{\text { PERMNO }}$ | $\underline{\text { Start }}$ | $\underline{\text { Ending }}$ |
| :--- | :---: | :---: |
| 22787 | 19570301 | 19700318 |
| 29962 | 19680801 | 19700318 |
| 23579 | 19631017 | 19700422 |
| 41515 | 19650215 | 19700422 |
| 28089 | 19680404 | 19700422 |

The second column "Start" signifies that the company of such PERMNO entered S\&P 500, and "Ending" signifies when the company left the S\&P 500. The DSP500 we used had an end date of 20071231.

One can obtain this file by ftp in: wrds.wharton. upenn.edu/wrds/crsp/sasdata/a_index/
A simple login script would be ssh -l username wrds.wharton.upenn.edu
This file is updated once per year, after the most recent calendar year is updated in CRSP (usually around February-March of the subsequent year). For example, the CY2012 became available February 18, 2013, and included data through December 31, 2012. Care must be exercised when augmenting a previously-downloaded dataset. As mentioned below, PERMNOs can change retroactively, requiring the acquisition of a new DSP500 list and new data download.

### 6.2 Validation of DSP500 with True Constituents

The 500 companies on each date from DSP500 were obtained with R code ${ }^{4}$. Then, with the PERMNOs from DSP500, the company names were obtained from CRSP on Wharton Research Data Services (WRDS). Then, the company names of the calibration data and the company names from CRSP

[^3]of the DSP 500 PERMNO list were manually compared. The results of comparisons are summarized as below. In short, in this sample of calibration dates, the DSP500 has been found to be valid with the exceptions of Ford Motor and R.J. Reynolds Tobacco.

In the table, "Date" signifies the date of lists that were compared, and "\# in constituent" signifies the number of companies in the real data, and "\# in dsp500" signifies the number of companies obtained from CRSP with the PERMNOs from DSP500. "Problem" signifies the companies which were missing, and are shown with their tickers. First, the lists of $9 / 3 / 02$ were first compared with the calibration data, and also checked against the list from the now-defunct Standard \& Poor's website.

| date | $\#$ in constituent | $\#$ in dsp500 | problem |
| :--- | :--- | :--- | :--- |
| 19990425 | 500 | 500 |  |
| 20020903 | 500 | 498 | F, RJR |
| 20031213 | 500 | 499 | F |
| 20041229 | 500 | 499 | F |
| 20060823 | 500 | 499 | F |
| 20071219 | 500 | 499 | F |

The "problem" column shows which companies were problematic ${ }^{5}$. There were two companies, Ford Motor and RJ Reynolds Tobacco. CRSP had retroactively changed the PERMNO for Ford Motor from 88394 to 25785 . All data with 88394 has been removed, according to the 2008/09 CRSP release note[10]. According to the release note, CRSP issued a new PERMNO, 25785, for Ford Motor Company at the time of its reorganization, which is why the complete time-series for Ford is split into two PERMNOs. CRSP researchers extensively reviewed the handling of Ford and as a result, decided to combine the two historical data series into one. PERMNO 88394 has been removed from CRSP database and all history for Ford may be found under PERMNO 25875. However, the 2007-vintage DSP500 file had Ford as 88394, so when the PERMNOs were put in CRSP, the output does not include Ford.

Also, R.J. Reynolds Tobacco had PERMNO of 14218 before 4/28/89 and of 86946 between 6/15/99 and $8 / 1 / 04[11]$. According to the DSP500, the "start" of 86946 is 20020904 and "ending" is 20071231. However, RJR existed in the actual S\&P 500 in 9/3/02 according to the validation lists and from Standard and Poor's. We see that, as the case of Ford signifies, it is necessary to update the PERMNO list periodically if one wants to do ongoing backtesting. If one needed to re-download all of the data, then a new PERMNO list must be obtained. Otherwise, one must keep the archived data since new downloads can and will arbitrarily have retroactive changes of companies' identifiers.

[^4]
## 7. Data Crosswalks and the CRSP/Compustat Merged (CCM) Linking Table

Compustat does offer the historical constituent lists for various indices as discussed above, but it does not give return data, providing instead various accounting figures such as ROE, market capitalization, number of shares outstanding, asset turnover rate; and meta-information about companies such as address, telephone, etc. Therefore, to use these index components, we have to convert the data keys from Compustat to CRSP because CRSP does provide returns; In order to link Compustat with CRSP, we use the CRSP/Compustat Merged (CCM) Database under CRSP of WRDS and use the linking table ${ }^{6}$. There are various ways to exploit the linking table.

### 7.1 COMPUSTAT GVKEY $\rightarrow$ PERMNO

We used the crosswalk file downloaded from CRSP/Compustat Merged Database - Linking Table and conducted crosswalking by pulling out PERMNOs that match GVKEY of historical constituent list we obtained from Compustat.

There are many discrepancies with the GVKEY $\rightarrow$ PERMNO conversion method, resulting from multiple PERMNOs being associated with a single GVKEY. We take 12/13/2003 as an example and explain why these discrepancies occur. All the other dates have similar stories.
2003.12.13.

| GVKEY | Ipermno | Conm | in real constituent list |
| :---: | :---: | :---: | :---: |
| 2435 | 29938 | BROWN-FORMAN -CL B | TRUE |
| 2435 | 29946 | BROWN-FORMAN -CL B | FALSE |
| 3226 | 89565 | COMCAST CORP | TRUE |
| 3226 | 89525 | COMCAST CORP | FALSE |
| 7146 | 89155 | MCCORMICK \& CO INC | TRUE |
| 7146 | 52090 | MCCORMICK \& CO INC | FALSE |
| 7506 | 76234 | MOLEX INC | TRUE |
| 7506 | 54827 | MOLEX INC | FALSE |
| 13714 | 76226 | CBS CORP | FALSE |
| 13714 | 75104 | CBS CORP | TRUE |
| 62689 | 89495 | TRAVELERS COS INC | FALSE |
| 62689 | 89346 | TRAVELERS COS INC | TRUE |

In most cases, there are two affiliates with different PERMNOs. Brown Forman (distillers of Jack Daniels whiskey) with PERMNO 29938 is that of the thinly-traded and closely held Class A shares, and Brown Forman with PERMNO 29946 is of Class B. Note that the company name in COMPUSTAT is Class B for both!

COMCAST with PERMNO 89525 is of Class A, and 89565 of Class K.

[^5]McCormick \& Co's PERMNO 52090 exists from 19721214 to the current download date and PERMNO 89155 exists from 20010917 to the current download date.

For Molex, PERMNO 54827 exists from 19721214 to the current download date, and PERMNO 76234 exists from 19900726 to the current download date.

CBS Inc., PERMNO 75104 exists from 19870610 to 19901231 and from 19910101 to the current download date, and PERMNO 76226 exists from 19900614 to 19901231 and 19910101 to the current download date.

Traveler's Property, PERMNO 89346 existed from 20020322 to 20020731, from 20020801 to 20020820, and from 20020821 to 20040401. PERMNO 89495 of Travelers Cos Inc. existed from 20020821 to 20040401.
1999.04.25.

| GVKEY | lpermno conm | in real <br> constituent <br> list |  |
| :--- | :--- | :--- | :--- |
| $\mathbf{1 2 3 9}$ | 69163 | ALBERTO-CULVER CO | TRUE |
| $\mathbf{1 2 3 9}$ | 42083 | ALBERTO-CULVER CO | FALSE |
| $\mathbf{2 4 3 5}$ | 29946 | BROWN-FORMAN -CL B | FALSE |
| $\mathbf{2 4 3 5}$ | 29938 | BROWN-FORMAN -CL B | TRUE |
| $\mathbf{5 9 0 5}$ | 83824 | INCO LTD | TRUE |
| $\mathbf{5 9 0 5}$ | 12546 | INCO LTD | FALSE |
| $\mathbf{8 9 7 2}$ | 85658 | RAYTHEON CO | TRUE |
| $\mathbf{8 9 7 2}$ | 24942 | RAYTHEON CO | FALSE |
| $\mathbf{1 3 7 1 4}$ | 76226 | CBS CORP | FALSE |
| $\mathbf{1 3 7 1 4}$ | 75104 | CBS CORP | TRUE |
| $\mathbf{1 4 5 9 0}$ | 75294 | FREEPORT-MCMORAN COP\&GOLD | TRUE |
| $\mathbf{1 4 5 9 0}$ | 81774 | FREEPORT-MCMORAN COP\&GOLD | FALSE |

2002.09.03.

| GVKEY | lpermno conm | in real <br> constituent <br> list |  |
| :--- | :--- | :--- | :--- |
| $\mathbf{1 2 3 9}$ | 42083 | ALBERTO-CULVER CO | FALSE |
| $\mathbf{1 2 3 9}$ | 69163 | ALBERTO-CULVER CO | TRUE |
| $\mathbf{2 4 3 5}$ | 29938 | BROWN-FORMAN -CL B | TRUE |
| $\mathbf{2 4 3 5}$ | 29946 | BROWN-FORMAN -CL B | FALSE |
| $\mathbf{7 5 0 6}$ | 76234 | MOLEX INC | TRUE |
| $\mathbf{7 5 0 6}$ | 54827 | MOLEX INC | FALSE |
| $\mathbf{1 3 7 1 4}$ | 75104 | CBS CORP | TRUE |
| $\mathbf{1 3 7 1 4}$ | 76226 | CBS CORP | FALSE |
| $\mathbf{6 2 6 8 9}$ | 89346 | TRAVELERS COS INC | TRUE |
| $\mathbf{6 2 6 8 9}$ | 89495 | TRAVELERS COS INC | FALSE |

2004.12.29.

2006.08.23.

| GVKEY | lpermno | Conm | in real <br> constituent <br> list |
| :--- | :--- | :--- | :--- |
| $\mathbf{2 4 3 5}$ | 29938 | BROWN-FORMAN -CL B | TRUE |
| $\mathbf{2 4 3 5}$ | 29946 | BROWN-FORMAN -CL B | FALSE |
| $\mathbf{2 7 1 0}$ | 64899 | CONSTELLATION BRANDS | TRUE |
| $\mathbf{2 7 1 0}$ | 69796 | CONSTELLATION BRANDS | FALSE |
| $\mathbf{3 2 2 6}$ | 89525 | COMCAST CORP | FALSE |
| $\mathbf{3 2 2 6}$ | 89565 | COMCAST CORP | TRUE |
| $\mathbf{3 5 0 5}$ | 59248 | MOLSON COORS BREWING CO | FALSE |
| $\mathbf{3 5 0 5}$ | 90562 | MOLSON COORS BREWING CO | TRUE |
| $\mathbf{6 6 6 9}$ | 52708 | LENNAR CORP | FALSE |
| $\mathbf{6 6 6 9}$ | 89731 | LENNAR CORP | TRUE |
| $\mathbf{7 1 4 6}$ | 52090 | MCCORMICK \& CO INC | FALSE |
| $\mathbf{7 1 4 6}$ | 89155 | MCCORMICK \& CO INC | TRUE |
| $\mathbf{7 5 0 6}$ | 54827 | MOLEX INC | FALSE |
| $\mathbf{7 5 0 6}$ | 76234 | MOLEX INC | TRUE |
| $\mathbf{1 2 8 8 6}$ | 90441 | NEWS CORP | FALSE |
| $\mathbf{1 2 8 8 6}$ | 90442 | NEWS CORP | TRUE |
| $\mathbf{1 3 7 1 4}$ | 75104 | CBS CORP | TRUE |


| 13714 | 76226 | CBS CORP | FALSE |
| :--- | :--- | :--- | :--- |
| 157858 | 90251 | FREESCALE SEMICONDUCTOR INC | TRUE |
| 157858 | 90435 | FREESCALE SEMICONDUCTOR INC | FALSE |
| 165675 | 91063 | VIACOM INC | FALSE |
| 165675 | 91066 | VIACOM INC | TRUE |

2007.12.19.

| GVKEY | lpermno | cOnm | in real <br> constituent <br> list |
| :--- | :--- | :--- | :--- |
| $\mathbf{2 4 3 5}$ | 29946 | BROWN-FORMAN -CL B | FALSE |
| $\mathbf{2 4 3 5}$ | 29938 | BROWN-FORMAN -CL B | TRUE |
| $\mathbf{2 7 1 0}$ | 64899 | CONSTELLATION BRANDS | TRUE |
| $\mathbf{2 7 1 0}$ | 69796 | CONSTELLATION BRANDS | FALSE |
| $\mathbf{3 2 2 6}$ | 89525 | COMCAST CORP | FALSE |
| $\mathbf{3 2 2 6}$ | 89565 | COMCAST CORP | TRUE |
| $\mathbf{3 5 0 5}$ | 90562 | MOLSON COORS BREWING CO | TRUE |
| $\mathbf{3 5 0 5}$ | 59248 | MOLSON COORS BREWING CO | FALSE |
| $\mathbf{6 6 6 9}$ | 89731 | LENNAR CORP | TRUE |
| $\mathbf{6 6 6 9}$ | 52708 | LENNAR CORP | FALSE |
| $\mathbf{7 1 4 6}$ | 52090 | MCCORMICK \& CO INC | FALSE |
| $\mathbf{7 1 4 6}$ | 89155 | MCCORMICK \& CO INC | TRUE |
| $\mathbf{7 5 0 6}$ | 76234 | MOLEX INC | TRUE |
| $\mathbf{7 5 0 6}$ | 54827 | MOLEX INC | FALSE |
| $\mathbf{1 2 8 8 6}$ | 90442 | NEWS CORP | TRUE |
| $\mathbf{1 2 8 8 6}$ | 90441 | NEWS CORP | FALSE |
| 13714 | 76226 | CBS CORP | FALSE |
| 13714 | 75104 | CBS CORP | TRUE |
| $\mathbf{1 6 5 6 7 5 ~}$ | 91066 | VIACOM INC | TRUE |
| $\mathbf{1 6 5 6 7 5 ~}$ | 91063 | VIACOM INC | FALSE |
|  |  |  |  |

Summary
There are two PERMNOs associated with one security because some companies have two affiliates with different classes. They have the same GVKEYs and are trading on the same date, so they cannot be filtered with linking dates. Therefore, we usually obtain more than 500 companies with this conversion method. This may cause distortions in financial analysis with too large a portfolio.

### 7.2 COMPUSTAT GVKEY $\rightarrow$ PERMCO

Because discrepancies using PERMNO cannot be ignored, it was necessary to find other possible linkages between the constituent lists from the Compustat data and the CRSP data. PERMNO is permanent issue identifier and PERMCO is permanent company issue identifier on CRSP. One PERMNO belongs to only one PERMCO but one PERMCO can have one or more PERMNO.[16]

### 1999.04.25.

The resulting PERMCOs are 498 in total. SML Corp, which exists in the real data, is not returned when we convert them into PERMCO. Also, there are two Sprint securities in the real data, but PERMCO does not differentiate between the two, Sprint FON ("fiber optic network") and Spring PCS ("personal communications services"). [17]
2002.09.03.

RJR is excluded in the Compustat historical list. A total of 499 securities are converted into PERMCO, and 498 are returned because here also since PERMCO does not differentiate the Sprint stocks as given above.
2003.12.13.

The number of PERMCOs returned is 499. Sprint causes problem here again.
2004.12.29., 2006.08.23., 2007.12.19.

The total number of PERMCOs for the three dates is 500 , and everything is congruent with the calibration data.

## Summary

Unlike PERMNO, PERMCO seems to give fewer distortions because the number of stocks that are off from the real data is much less than in the case of PERMNOs. However, with PERMCO, we get fewer than 500 companies in the first three dates because it does not differentiate stocks from the same company but with different classes, so analysis would cause problems such as underestimated portfolio return. Also, since there are two different securities with different prices and returns available with one PERMCO, we do not know which of the two would be given as output with the PERMCO.

### 7.3 COMPUSTAT (GVKEY \& IID) $\rightarrow$ PERMNO

As PERMCO seems to give less than 500, we again look to locate PERMNO. Indeed, with IID we can link one GVKEY to one PERMNO (see section 4.2).
1999.04.25.

Total number of components obtained by Compustat is 500 . However, in the process of converting these 500 components into PERMNO using GVKEY and IID, SML Holdings is excluded.
2002.09.03.

Total number of components obtained is 499 . RJR is not obtained by Compustat.
2003.12.13., 2004.12.29., 2006.08.23., 2007.12.19.

For all four dates, total number of components obtained is 500 . Everything matches the real list.

This linking method using both GVKEY and IID works very well. It is certainly far better than the method that links only GVKEY to PERMNO. There will be fewer distortions in financial research using this method than with the other two.

### 7.4 Comparison among Conversion Methods

The following is a summary of the conversion methods for the calibration dates. The first number in each column is the number of matches to each target coding; the number in parentheses is the number of GVKEYs corresponding to the unique PERMNO/PERMCO.

|  |  | GVKEY+IID <br>  GVKEY | $\rightarrow$ PERMNO | GVKEY $\rightarrow$ |
| :--- | :--- | :--- | :--- | :--- |
| PERMNO | GVKEY $\rightarrow$ <br> PERMCO |  |  |  |
| 19990425 | 500 | $499(499)$ | $505(499)$ | $498(498)$ |
| 20020903 | 499 | $499(499)$ | $504(498)$ | $498(498)$ |
| 20031213 | 500 | $500(500)$ | $506(500)$ | $499(499)$ |
| 20041229 | 500 | $500(500)$ | $507(500)$ | $500(500)$ |
| 20060823 | 500 | $500(500)$ | $511(500)$ | $500(500)$ |
| 20071219 | 500 | $500(500)$ | $510(500)$ | $500(500)$ |

It is seen therefore that the GVKEY $\rightarrow$ PERMNO method returns more than the number of stocks necessary, and GVKEY $\rightarrow$ PERMCO returns less than the necessary number of stocks. On the other hand, GVKEY \& IID $\rightarrow$ PERMNO does not have such problems. Therefore, we confirm that GVKEY \& IID $\rightarrow$ PERMNO is the best way to link Compustat GVKEY with CRSP PERMNO.

## 8. Exploratory Enumeration of Compustat Historical Components

### 8.1 Data Availability and Limitations

The Compustat Constituent list not only gives the S\&P 500 constituent list, it also gives DJIA, S\&P 100, NASDAQ 100, and many other S\&P indices. Therefore, we can use this method in obtaining historical constituent lists for other indices. The most frequently used market indices are DJIA, S\&P 500, S\&P 100, and NASDAQ 100, so we provide the table below, which was introduced earlier in section 2.

| Name | GVKEYX | TIC | Start | Indicated <br> Date |
| :--- | :--- | :--- | :---: | :---: |
| S\&P 500 Comp-Ltd | 3 | I0003 | 19640331 | 19640331 |
| Dow Jones Industrials-30 Stk | 5 | I0005 | 19970317 | 19071110 |
| Nasdaq 100 | 208 | I0028 | 20041105 | 20041105 |
| S\&P 100-Ltd | 664 | I0014 | 19890911 | 19890911 |

There is a cautionary item in this table. The indicated date shows how far back the index is supposed to have data for. The data start date is different from the indicated date for DJIA because although its first indicated date is $11 / 10 / 1907$, on that date, we only obtain only one stock (which is General Electric.) I0005 does not return 30 stocks until 03/17/1997. Going further back from 3/17/97 we see that the constituent list fails to accumulate the stocks from previous reconstituation dates. What seems to be happening with the DJIA is that it fails to include the stocks which have been dropped from the index on each reconstituation date prior to $3 / 17 / 1997$. For that date is picks up the new stocks Johnson \& Johnson, Walmart, Citigroup and Hewlett Packard, but it fails to keep in the record the four stocks dropped on the new date - Westinghouse, Bethlehem Steel, Texaco and Woolworth are no longer in the historical record. This happens with each reconstituation date in the past until in 1907 we only have one stock in the "historical record".

From 3/17/1997 on, the index returns thirty stocks, and constantly has thirty stocks for all subsequently available dates. After concernedly checking, we can say that the other indexes here and in the table in section 2 do not suffer from this problem. We have been in coordination with WRDS to resolve this issue and hope to report its resolution soon.

It is even true that the other indices show fluctuations in the number of total historical components on a daily basis when one searches after the indicated data start date. For example, the S\&P 500 shows the total number as low as 453 in June 1976 and 483 in September 1968 and June 1970. On other days, the total numbers are from 488 to 502 , but mostly 500 , so long one is pulling from dates no earlier than the data start date. Prior to that, I0003 suffers from failing to pick up constituent changes as well. The table below shows how many dates have how many components for S\&P 500 since $03 / 31 / 1964$.

| total number of <br> components <br> how many dates <br> with the number | 29 | 553 | 483 | 488 | 490 | 491 | 492 | 493 | 494 | 495 | 496 | 497 | 498 | 499 | 500 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |

The table below is obtained by DSP500 with the last date of $12 / 30 / 2007$. This shows how many dates have how many components for S\&P 500 since 12/31/1925. We know that CRSP's S\&P 500 index data shows 90 stocks until March 1957; indeed, [19] tells us that 500 stocks constituted the SPX on $3 / 4 / 1957$ and onwards. The fact that we never obtain 90 stocks and the high frequency of non-500 stock counts tells us that DSP500 might not be reliable before 3/31/1964 as well.


| total number of components | 432 | 433 | 438 | 440 | 442 | 445 | 446 | 447 | 448 | 449 | 450 | 451 | 452 | 454 | 455 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| how many dates with the number | 14 | 22 | 42 | 15 | 7 | 7 | 35 | 28 | 8 | 35 | 21 | 48 | 2 | 2 | 14 |
| total number of components | 457 | 460 | 461 | 463 | 464 | 465 | 467 | 469 | 471 | 473 | 475 | 480 | 483 | 484 | 485 |
| how many dates with the number | 65 | 36 | 14 | 21 | 26 | 10 | 7 | 7 | 7 | 14 | 21 | 57 | 36 | 28 | 7 |
| total number of components | 486 | 487 | 489 | 490 | 491 | 492 | 493 | 494 | 495 | 496 | 497 | 498 | 499 | 500 |  |
| how many dates with the number | 28 | 25 | 21 | 50 | 1 | 7 | 12 | 40 | 22 | 43 | 257 | 110 | 758 | 13116 |  |

However, when we take the dates from 01/01/1970 for DSP500, we obtain a table as below. Therefore, we can conclude that the S\&P 500 composite list of early dates from 1920s to 1960s with DSP500 is mostly not complete.

| total number of components | 454 | 487 | 491 | 493 | 494 | 495 | 496 | 497 | 499 | 500 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| how many dates with the number | 2 | 1 | 1 | 5 | 4 | 9 | 21 | 108 | 758 | 13116 |

The table below is for S\&P 100 since the data start date, 09/11/1989.

| total number of components | 98 | 99 | 100 |
| :--- | :--- | :--- | :--- |
| how many dates with the number | 1 | 13 | 7912 |

The table below is for NASDAQ 100 since the data start date, 11/05/2004.

| total number of components | 91 | 98 | 99 | 100 | 101 |
| :--- | :--- | :--- | :--- | :--- | :--- |
| how many dates with the number | 8 | 52 | 307 | 1907 | 16 |

### 8.2 Validation Result for Other Indices

DJIA
We had true historical components for the Dow Jones Industrial Average for 07/03/2002, $04 / 08 / 2004,02 / 06 / 2006$, and $12 / 04 / 2008$. We validated that the lists of these dates match with constituents obtained by Compustat.

## S\&P 100

We had true historical components for S\&P 100 for $11 / 01 / 2002,12 / 01 / 2003,01 / 04 / 2005$, $09 / 25 / 2005,02 / 06 / 2006$, and $07 / 23 / 2010$. Except for $01 / 04 / 2005$, lists for all the other dates
are validated. Therefore, we assume that the calibration list of 01/04/2005 may be labeled with an incorrect date because none of the its components match the Compustat list of 1/4/05.

## NASDAQ 100

We had true historicals for NASDAQ 100 for 05/16/2002, 09/14/2002, 11/02/2002, 12/01/2003, and $09 / 01 / 2005$. We were able to validate only $09 / 01 / 2005$ because NASDAQ 100 obtained by Compustat is available from 11/03/2004, the "start date" in the table in section 8.1.
On 09/01/2005, Compustat gives 99 components for NASDAQ 100, and it does not include Celgene Corp (Ticker: CELG).

### 8.3 Summary

Compustat seems to provide all major indices, DJIA, S\&P 500, S\&P 400, S\&P 100, and NASDAQ 100 with constituents. It does give sufficiently far-reaching constituent lists for S\&P 500 although it also shows fluctuations in the total number of stocks for the index. The other indices do not show as much fluctuation, but they are not available as far back in time as the S\&P 500. Especially, the NASDAQ 100 starts only from 11/05/2004, the OEX begins only as far back as $9 / 11 / 1989$, and the DJIA to $3 / 17 / 1997$, making these impractical for use in thorough, long-horizon financial analysis and backtesting.

## 9. Another Way to Obtain NASDAQ 100

The NASDAQ 100[20] is composed of 100 of the largest domestic and international nonfinancial securities listed on the NASDAQ Stock Market, based on market capitalization. The index mainly contains major industries except financial companies. The NASDAQ 100 launched on January 31, 1985, but as Compustat only gives historical constituents after 11/05/2004, we felt it is necessary to attempt to obtain NASDAQ 100 from a different source. We looked at newspapers to obtain the data, and found that Barron's has the NASDAQ 100 list. We are unable to verify how back this list may be found.

The NASDAQ constituents have not varied as much as S\&P 500 historical constituents. We also note that NASDAQ shares their constituents' change on their website[4]. The table contains included companies and excluded companies; for example, Monster Worldwide, Inc was replaced by Seagate Technology on 11/10/2008.

|  | Additions |  | Deletions |  |
| :--- | :--- | :--- | :--- | :--- |
| Date | Symbol | Company | Symbol | Company |
| $11 / 10 / 2008$ | STX | Seagate Technology | MNST | Monster <br> Worldwide, <br> Inc. |

Full change information (through 11/10/2008) is included in the NASDAQ 100 Inclusion/ Exclusion Table in appendix 2. Unfortunately, this website seems to have ceased being updated in 2008, and we know there have been several component changes since then, with the subsequent changes found for example in the NASDAQ 100 Wikipedia article [20].

## 10. Conclusions and Discussion

Backtesting usually requires performing historical tests on constituents of major market indices. A technique for obtaining unique CRSP PERMNOs from Compustat's Indices' constituents' GVKEYs had been described, as well as other index constituent resources available using the graphical user interface (GUI) at WRDS.

One first downloads from WRDS - Compustat- North America - Index Constituents the list of constituents for the desired index. Then, from WRDS - CRSP/Compustat Merged Database - Linking Table one obtains a crosswalk of GVKEYs, IIDs and linking PERMNOs from which the appropriate filtering can generate a list of unique PERMNOs for constituents on a given date. It should be noted that the linking table uses company names as of the date of the linking table extraction, so that careful cross-checking and some corporate archeology is required to match company names with calibration data. For example (see appendix 1), when validating the OEX constituent list for $4 / 25 / 1999$, eight percent of the companies had name changes occurring sometime "in the future." These had to be reconciled. We found however that all the companies were accounted for using our technique.

If S\&P 500 constituents are desired, it is most efficient to simply use the DSP500 list obtainable by ftp from the WRDS Unix system. This provides a direct inclusion/exclusion table for all PERMNOs that were ever in the S\&P 500, at least since 3/31/1964, the earliest date that we have been able to verify historical constituents. Care should be taken when using DSP500 for PERMNO constituent pulls for dates prior to $3 / 31 / 1964$ until further research can explain the frequency of stock counts significantly less than 500.

Using the crosswalk technique and with little additional programming effort, we can generate a list of validated CRSP PERMNOs for any date from among the most popular indices that Compustat provides. The PERMNO list is of similar accuracy to that obtained from the DSP500, except now we would perform backtesting using the OEX, S\&P 400, S\&P 1500, etc. Further research would demonstrate the efficacy of using Compustat's sector indices or global index constituents.

In Compustat, NASDAQ 100 has constituents only back to 2004, and, pending resolution of the trouble ticket, the DJIA only goes back to 1997. If we wanted to restrict ourselves to these start dates it may be possible to manually merge the inclusion/exclusion tables found on the various websites with the narrative versions in the Wikipedia articles. This would result in about 20 years for NASDAQ 100 constituent data, and 100 years for the DOW industrials. GVKEYs or PERMNOs would then have to be obtained. Unfortunately, we would again have the issue of retroactively changing PERMNOs that was encountered in the SPX case, arguing against such a labor-intensive process.

## 11. References

[1] General Information FAQ's; Accessing SEP 500 Data, WRDS, http://wrds-
web.wharton.upenn.edu/wrds/support/Additional\ Support/WRDS\ Knowledge\ Base\ w ith\%20FAQs.cfm?folder_id=744\&article_id=1061 or http://tinyurl.com/787rgoh (login required).
[2] SEP 500 Datasets and Constituents, WRDS, http://wrds-
web.wharton.upenn.edu/wrds/support/Data/_004Research\ Applications/_015Other\ Topics/S \&P\%20500\%20Datasets\%20and\%20Constituents.cfm or http://tinyurl.com/lk5r7qa (login required).
[3] Overview of CRSP/Compustat Merged Database in WRDS, WRDS, http:/ / wrdsweb.wharton.upenn.edu/wrds//support/Data/_001Manuals\ and\%_20Overviews/_002CRSP/_001 General/_002WRDS\%20Overview \% 20of\%20CCM\%20\%28CRSP-
Compustat\%20Merged\%29\%20Database.cfm or http://tinyurl.com/mtjyf8m (login required).
[4] Standard \& Poor's webpage, link no longer operational, last visited in October 2011, http://wwweq.standardandpoors.com/portal/site/sp/en/us/page.topic/indices_500/2,3,2,2,0,0,0,0,0,2,3,0,0,0,0,0. html
[5] Historical components of the Dow Jones Industrial Average, Wikipedia, http://en.wikipedia.org/wiki/Historical_components_of_the_Dow_Jones_Industrial_Average
[6] Dow Jones Industrial Stocks, http://www.dow-jones-djia.com/2007/03/03/djia-first-published-july-3-1984/, also available http://tinyurl.com/c2242yg
[7] Historical Data - NASDAQ-100 Index, Nasdaq Corporate webpage, http://www.nasdaq.com/indexshares/historical_data.stm
[8] Thompson, J.R. and Baggett, L.S. "The MaxMedian Rule for Portfolio Design," The Proceedings of the U.S. Army Conference on Applied Statistics (2007)
[9] Thompson, J.R. Empirical Model Building: Data, Models, and Reality. Wiley, 2nd Ed. (2011) p. 356
[10] Stock E Indices Release Notes, July 2008 Monthly Update, Center for Research in Security Prices, http://www.google.com/url?sa=t\&rct=j\&q=\&esrc=s\&frm=1\&source=web\&cd=1\&ved=0CCcQFjAA\& url=http\%3A\%2F\%2Fwww.crsp.chicagobooth.edu\%2Fdocumentation\%2Fpdfs\%2Frelnotes\%2F2008\%2 Fmdaz_200807.pdf\&ei=ROftTui4EojosQLX2oXTCQ\&usg=AFQjCNFWoMp31CGEsiJHH31y6su1Sn7lz Q\&sig2=p9yRs0DZNib4whCmmQCpmA (defunct link); also available http://tinyurl.com/omkwkyt.
[11] CRSP ts_print PERMNO Search Tips, JMSB Tech Note, last modified September 19, 2007, http://johnmolson.concordia.ca/addon_components/JMSBSearchTool/TechNotes/CRSP_tsprint_PERMNO_Search_Tips.pdf
[12] Wharton Research Data Services main page, http://wrds-web.wharton.upenn.edu/wrds/
[13] CRSP/COMPUSTAT Merged Database Guide, Chicago Booth Center for Research in Security Prices, February 2009, http://wrds-
web.wharton.upenn.edu/wrds/support/document_show.cfm?key=/CRSP\ -
\%20Compustat\%20Merged \%20Database\%20(CCM).pdf, or http:// tinyurl.com/n7tkq74
[14] CRSP Software Guide: CRSP US Stock \& US Indices Databases, Chicago Booth Center for Research in Security Prices, June 2013, http://www.crsp.com/products/documentation/cupl-guide-crsp-us-stock-us-index-databases-and-crspcompustat-merged-database
[15] How do I access SEP 500 constituents?, Chicago Booth Center for Research in Security Prices, http://www.crsp.com/documentation/kb/crspaccess/ca-0004.html (defunct link)
[16] CRSP/COMPUSTAT Merged Database Guide, Chicago Booth Center for Research in Security Prices, January 2006, http://www.crsp.com/products/research-products/crspcompustat-mergeddatabase?activetab=docs\#tabhash
[17] Sprint Corporation History, FundingUniverse webpage,
http://www.fundinguniverse.com/company-histories/Sprint-Corporation-Company-History.html
[18] NASDAQ 100, Nasdaq Corporate webpage, http://www.nasdaq.com/markets/indices/nasdaq100.aspx
[19] SEP 500, Wikipedia, https://en.wikipedia.org/wiki/S\%26P_500
[20] NASDAQ 100, Wikipedia, https://en.wikipedia.org/wiki/NASDAQ-100.
[21] "Berkshire Buys Burlington in Buffett's Biggest Deal (Update5)", Bloomberg, November 3, 2009, http://www.bloomberg.com/apps/news?pid=newsarchive\&sid=aseXHPBeRtRE
[22] "Harrah's Entertainment Inc. changes name to Caesars Entertainment Corp.", Las Vegas Sun, November 23, 2010, http://www.lasvegassun.com/news/2010/nov/23/us-harrahs-name-change/
[23] Viacom (1971-2005), Wikipedia, http://en.wikipedia.org/wiki/Viacom_(1971\�\�\�2005)
[24] "CBS, Viacom Formally Split", CBS Money Watch, February 11, 2009, http://www.cbsnews.com/2100-500395_162-1176111.html
[25] Clear Channel Communications, Wikipedia, http://en.wikipedia.org/wiki/Clear_Channel_Communications
[26] "SBC Communications to Adopt AT\&T Name", AT\&T, October 27, 2005, http://www.att.com/gen/press-room?pid=4800\&cdvn=news\&newsarticleid=21850
[27] OfficeMax, Wikipedia, http://en.wikipedia.org/wiki/OfficeMax
[28] "Invitrogen and Applied Biosystems Complete Merger", Life Technologies Corporation Press Release, 2008, http://www.lifetechnologies.com/us/en/home/about-us/news-gallery/press-releases/2008/invitrogen-and-applied-biosystems-complete-merger.html
[29] Virgin Media, Wikipedia, http://en.wikipedia.org/wiki/Virgin_Media
[30] Bell Atlantic, Wikipedia, https://en.wikipedia.org/wiki/Bell_Atlantic
[31] Coastal Corporation, Wikipedia, https://en.wikipedia.org/wiki/Coastal_Corporation
[32] Kmart, Wikipedia, https://en.wikipedia.org/wiki/K-mart
[33] Monsanto, Wikipedia, https://en.wikipedia.org/wiki/Monsanto
[34] Northern Telecom, Wikipedia, https://en.wikipedia.org/wiki/Northern_Telecom

## Appendix 1. Corporate Archeology Vignettes

In matching the true historical components with the lists provided by other methods, there are some apparent discrepancies that actually turn out can be reconciled, provided proper corporate matching effort is made. These include:

1. Burlington \& Berkshire

Berkshire Hathaway acquired Burlington in 2009[21].

## 2. Harrah's Entertainment \& Caesars Entertainment

Harrah's Entertainment Inc. changes name to Caesars Entertainment Corp. on November 23, 2010[22].

## 3. Viacom \& CBS Corp

Effective on December 31, 2005, Viacom changed its name to CBS Corporation[23] and later CBS and Viacom are formally split[24].

## 4. Clear Channel Communications \& CC Media Holdings

They are actually the same company with different names[25].

## 5. SBC Communications \& AT\&T

SBC Communications Inc. (NYSE:SBC) announced that it will change its name to AT\&T, Inc. following completion of its acquisition of AT\&T (NYSE:T), completed in late 2005[26].

## 6. Boise Cascade \& Office Max

On December 9, 2003, Boise Cascade Corporation acquired 100 percent of the voting securities of OfficeMax, Inc. The company's name changed from Boise Cascade Corporation to OfficeMax Incorporated, and the names of the office product segments changed from Boise Office Solutions, Contract and Boise Office Solutions, Retail to OfficeMax, Contract and OfficeMax, Retail.[27]

## 7. Invitrogen \& Life Technologies Corp

Invitrogen and Applied Biosystems Complete Merger: "CARLSBAD, Calif.--(BUSINESS WIRE) - Invitrogen Corporation (NASDAQ:IVGN) and Applied Biosystems Inc. (NYSE:ABI) announced the successful completion of their merger transaction. The new company was named Life Technologies Corporation." [28]

## 8. NTL Inc \& Virgin Mobile

The company was formed in March 2006 through the merger of NTL and Telewest, which created NTL:Telewest. A further merger with Virgin Mobile UK in July 2006 created the first "quadruple-play" media company in the United Kingdom, offering television, internet, mobile phone and fixed-line telephone services. All of the company's consumer services were rebranded under the name of Virgin Media in February 2007. And a reorganization split NTL itself into NTL Inc.[29]
9. Bell Atlantic and Verizon

In 2000, Bell Atlantic merged with former independent phone company GTE, and adopted the name "Verizon"[30].
10. Coastal Corporation

Coastal Corporation was a major energy producer when it merged with the El Paso corporation in 2001 [31].
11. K-mart

The chain purchased Sears in 2005, forming a new corporation under the name Sears Holdings Corporation. [32].
12. Monsanto Company

In December, Monsanto merged with Pharmacia \& Upjohn, and the agricultural division became a wholly owned subsidiary of the "new" Pharmacia [33].
13. Northern Telecom

Nortel is simply another name for Northern Telecom [34].

## Appendix 2. NASDAQ 100 Inclusion/Exclusion Table

| Date | Additions |  | Deletions |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Symbol | Company | Symbol | Company |
| 11/10/2008 | STX | Seagate Technology | MNST | Monster Worldwide, Inc. |
| 07/21/2008 | FLIR | FLIR Systems, Inc. | UAUA | UAL Corporation |
| 05/19/2008 | CA | CA, Inc. | TLAB | Tellabs, Inc. |
| 04/30/2008 | DTV | The DIRECTV Group, Inc. | BEAS | BEA Systems, Inc. |
| 12/24/2007 | FMCN | Focus Media Holding Limited | ERIC | LM Ericsson Telephone Company |
| 12/24/2007 | HANS | Hansen Natural Corporation | PTEN | Patterson-UTI Energy, Inc. |
| 12/24/2007 | HOLX | Hologic, Inc. | ROST | Ross Stores, Inc. |
| 12/24/2007 | SRCL | Stericycle, Inc. | SEPR | Sepracor Inc. |
| 12/24/2007 | STLD | Steel Dynamics, Inc. | XMSR | XM Satellite Radio Holdings Inc. |
| 12/04/2007 | BIDU | Baidu.com, Inc. | CKFR | CheckFree Corporation |
| 10/08/2007 | LEAP | Leap Wireless International, Inc. | CDWC | CDW Corporation |
| 10/02/2007 | HSIC | Henry Schein, Inc. | MXIM | Maxim Integrated Products, Inc |
| 07/12/2007 | FWLT | Foster Wheeler, Ltd. | BMET | Biomet, Inc. |
| 06/01/2007 | CEPH | Cephalon, Inc. | MEDI | MedImmune, Inc. |
| 03/08/2007 | UAUA | UAL Corporation | AEOS | American Eagle Outfitters, Inc. |
| 02/14/2007 | RYAAY | Ryanair Holdings plc | APCC | American Power Conversion Corporation |
| 02/01/2007 | LOGI | Logitech Technology SA | CMVT | Comverse Technology, Inc. |
| 12/18/2006 | INFY | Infosys Technologies Limited ADS | JDSU | JDS Uniphase Corporation |
| 12/18/2006 | LVLT | Level 3 Communications, Inc. | LNCR | Lincare Holdings Inc. |
| 12/18/2006 | VRTX | Vertex Pharmaceuticals Incorporated | URBN | Urban Outfitters, Inc.. |
| 12/12/2006 | AEOS | American Eagle Outfitters, Inc. | RHAT | Red Hat, Inc. |
| 10/24/2006 | LINTA | Liberty Media Corporation, Liberty Interactive Series A | ATYT | ATI Technologies Inc. |
| 05/08/2006 | MICC | Millicom International Cellular S.A. | PIXR | PIXAR |
| 04/20/2006 | AKAM | Akamai Technologies, Inc. | CHIR | Chiron Corporation |
| 02/01/2006 | ISRG | Intuitive Surgical, Inc. | SEBL | Siebel Systems, Inc |
| 01/09/2006 | AMLN | Amylin Pharmaceuticals, Inc. | MCIP | MCI, Inc. |
| 01/04/2006 | JOYG | Joy Global Inc. | MERQE | Mercury Interactive Corporation |
| 12/19/2005 | ATVI | Activision Inc. | CECO | Career Education Corp. |
| TR2014-01.20140110.doc |  |  |  | 1/10/2014 |


| 12/19/2005 | CDNS | Cadence Design Systems, Inc. | DLTR | Dollar Tree Stores Inc. |
| :---: | :---: | :---: | :---: | :---: |
| 12/19/2005 | CKFR | Checkfree Corporation | ISIL | Intersil Corporation |
| 12/19/2005 | DISCA | Discovery Holding Company | IVGN | Invitrogen Corproation |
| 12/19/2005 | EXPE | Expedia Inc. | LVLT | Level 3 Communications Inc. |
| 12/19/2005 | GOOG | Google Inc. | MLNM | Millennium Pharmaceuticals Inc. |
| 12/19/2005 | MNST | Monster Worldwide, Inc. | MOLX | Molex Inc. |
| 12/19/2005 | NIHD | NII Holdings, Inc. | NVLS | Novellus Systems Inc. |
| 12/19/2005 | NVDA | NVIDIA Corporation | QLGC | QLogic Corp. |
| 12/19/2005 | PTEN | Patterson-UTI Energy Inc. | SANM | Sanmina-SCI Corp |
| 12/19/2005 | RHAT | Red Hat, Inc. | SNPS | Synopsys Inc. |
| 12/19/2005 | URBN | Urban Outfitters, Inc. | SSCC | Smurfit-Stone Container Corp. |
| 08/15/2005 | SEPR | Sepracor Inc. | NXTL | Nextel Communications, Inc. |
| 07/01/2005 | CELG | Celgene Corporation | VRTS | VERITAS Software Corporation |
| 12/29/2004 | CTSH | Cognizant Technology Solutions Corporation | PSFT | PeopleSoft, Inc. |
| 12/20/2004 | ADSK | Autodesk, Inc. | CEPH | Cephalon, Inc. |
| 12/20/2004 | ERICY | LM Ericsson Telephone Company | CPWR | Compuware Corporation |
| 12/20/2004 | LBTYA | Liberty Media International, Inc. | FHCC | First Health Group Corp. |
| 12/20/2004 | MCIP | MCI, Inc. | GNTX | Gentex Corporation |
| 12/20/2004 | NTLI | NTL Incorporated | HSIC | Henry Schein, Inc. |
| 12/20/2004 | SIRI | Sirius Satellite Radio Inc. | NVDA | NVIDIA Corporation |
| 12/20/2004 | WYNN | Wynn Resorts, Limited | PTEN | Patterson-UTI Energy, Inc. |
| 12/20/2004 | XMSR | XM Satellite Radio Holdings Inc. | RYAAY | Ryanair Holdings plc |
| 08/19/2004 | KMRT | KMart Holding Corporation | SPOT | PanAmSat Corporation |
| 11/13/2003 | SNDK | SanDisk Corporation | BGEN | Biogen, Inc. |
| 12/22/2003 | ATYT | ATI Technologies Inc. | ADCT | ADC Telecommunications, Inc. |
| 12/22/2003 | CECO | Career Education Corporation | BRCD | Brocade Communications Systems, Inc. |
| 12/22/2003 | GRMN | Garmin Ltd. | CIEN | CIENA Corporation |
| 12/22/2003 | ISIL | Intersil Corporation | ERICY | LM Ericsson Telephone Company |
| 12/22/2003 | LRCX | Lam Research Corporation | HGSI | Human Genome Sciences, Inc. |
| 12/22/2003 | LVLT | Level 3 Communications, Inc. | ICOS | ICOS Corporation |
| 12/22/2003 | MRVL | Marvell Technology Group, Inc. | MNST | Monster Worldwide Inc. |
| 12/22/2003 | RIMM | Research in Motion Limited | RFMD | RF Micro Devices, Inc. |
| 06/03/2002 | DLTR | Dollar Tree Stores, Inc. | ADLAE | Adelphia Communications Corporation |
| 07/15/2002 | SIAL | Sigma Aldrich Corporation | IMNX | Immunex Corporation |


| 07/24/2002 | LNCR | Lincare Holdings Inc. | WCOEQ | WorldCom, Inc. |
| :---: | :---: | :---: | :---: | :---: |
| 11/07/2002 | TEVA | Teva Pharmaceutical Industries Limited | CEFT | Concord EFS, Inc. |
| 11/19/2002 | CMCSA | Comcast Corporation Class A | CMCSK | Comcast Corporation - |
|  |  |  |  | Special Class A |
| 12/16/2002 | PDCO | Patterson Dental Company | GMSTE | Gemstar-TV Guide International Inc. |
| 12/23/2002 | APCC | American Power Conversion Corporation | ABGX | Abgenix, Inc. |
| 12/23/2002 | CHRW | C.H. Robinson Worldwide, Inc. | ADRX | Andrx Group |
| 12/23/2002 | EXPD | Expeditors International of Washington, Inc. | AMCC | Applied Micro Circuits Corporation |
| 12/23/2002 | FAST | Fastenal Company | ATML | Atmel Corporation |
| 12/23/2002 | FHCC | First Health Group Corp. | CHTR | Charter Communications, Inc. |
| 12/23/2002 | GNTX | Gentex Corporation | CNXT | Conexant Systems, Inc. |
| 12/23/2002 | HSIC | Henry Schein, Inc. | CYTC | Cytyc Corporation |
| 12/23/2002 | LAMR | Lamar Advertising Company | IDTI | Integrated Device Technology, Inc. |
| 12/23/2002 | PETM | PETsMART, Inc. | IMCL | ImClone Systems Incorporated |
| 12/23/2002 | PIXR | Pixar | ITWO | i2 Technologies, Inc. |
| 12/23/2002 | PTEN | Patterson-UTI Energy, Inc. | PDLI | Protein Design Labs, Inc. |
| 12/23/2002 | ROST | Ross Stores, Inc. | PMCS | PMC - Sierra, Inc. |
| 12/23/2002 | RYAAY | Ryanair Holdings plc | RATL | Rational Software Corporation |
| 12/23/2002 | WFMI | Whole Foods Market, Inc. | SEPR | Sepracor Inc. |
| 12/23/2002 | XRAY | DENTSPLY International Inc. | VTSS | Vitesse Semiconductor Corporation |
| 02/13/2001 | BRCD | Brocade Communication Systems, Inc. | SDLI | SDL, Inc. |
| 03/13/2001 | NVLS | Novellus Systems, Inc. | BMCS | BMC Software, Inc. |
| 05/31/2001 | NVDA | NVIDIA Corporation | VSTR | VoiceStream Wireless Corporation |
| 10/04/2001 | ADRX | Andrx Group | ATHMQ | At Home Corporation |
| 10/05/2001 | GILD | Gilead Sciences, Inc. | EXDSQ | Exodus Communications, Inc. |
| 12/17/2001 | APOL | Apollo Group, Inc. | XOXO | XO Communications, Inc. |
| 12/24/2001 | CDWC | CDW Computer Centers, Inc. | ARBA | Ariba, Inc. |
| 12/24/2001 | CEPH | Cephalon, Inc. | BVSN | BroadVision, Inc. |
| 12/24/2001 | CHTR | Charter Communications, Inc. | CMGI | CMGI, Inc. |
| 12/24/2001 | CYTC | CYTYC Corporation | CNET | CNET Networks, Inc. |
| 12/24/2001 | ESRX | Express Scripts, Inc. | COMS | 3Com Corporation |
| 12/24/2001 | ICOS | ICOS Corporation | INKT | Inktomi Corporation |
| 12/24/2001 | IDTI | Integrated Device Technology, | LVLT | Level 3 Communications, Inc |


| Inc. |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| 12/24/2001 | IMCL | ImClone Systems Incorporated | MCLD | McLeodUSA Incorporated |
| 12/24/2001 | IVGN | Invitrogen Corporation | MFNX | Metromedia Fiber Network, Inc. |
| 12/24/2001 | PDLI | Protein Design Labs, Inc. | NOVL | Novell, Inc. |
| 12/24/2001 | SEPR | Sepracor Inc. | PALM | Palm, Inc. |
| 12/24/2001 | SNPS | Synopsys, Inc. | PMTC | Parametric Technology Corporation |
| 12/24/2001 | SYMC | Symantec Corporation | RNWK | RealNetworks, Inc. |
| 06/09/2000 | VRSN | VeriSign, Inc. | NSOL | Network Solutions, Inc. |
| 09/07/2000 | JNPR | Juniper Networks, Inc. | VISX | VISX, Incorporated |
| 10/27/2000 | ARBA | Ariba, Inc. | NTLI | NTL Incorporated |
| 10/30/2000 | BRCM | Broadcom Corporation - Class A | LCOS | Lycos, Inc. |
| 11/06/2000 | PALM | Palm, Inc. | GBLX | Global Crossing Ltd. |
| 12/18/2000 | ABGX | Abgenix, Inc. | ADPT | Adaptec, Inc. |
| 12/18/2000 | BEAS | BEA Systems, Inc. | APCC | American Power Conversion Corporation |
| 12/18/2000 | СНКР | Check Point Software Technologies Ltd. | APOL | Apollo Group, Inc. |
| 12/18/2000 | EXDS | Exodus Communications, Inc. | DLTR | Dollar Tree Stores, Inc. |
| 12/18/2000 | FLEX | Flextronics International Ltd. | LGTO | Legato Systems, Inc. |
| 12/18/2000 | HGSI | Human Genome Sciences, Inc. | MLHR | Herman Miller, Inc. |
| 12/18/2000 | IDPH | IDEC Pharmaceuticals Corporation | NETA | Network Associates, Inc. |
| 12/18/2000 | INKT | Inktomi Corporation | NWAC | Northwest Airlines Corporation |
| 12/18/2000 | MERQ | Mercury Interactive Corporation | PHSY | PacifiCare Health Systems, Inc |
| 12/18/2000 | MLNM | Millennium Pharmaceuticals, Inc. | QTRN | Quintiles Transnational Corp. |
| 12/18/2000 | RATL | Rational Software Corporation | SIAL | Sigma-Aldrich Corporation |
| 12/18/2000 | TMPW | TMP Worldwide Inc. | SNPS | Synopsys, Inc. |
| 01/13/1999 | CMVT | Comverse Technology, Inc. | HBOC | HBO \& Company |
| 03/10/1999 | ATHM | At Home Corporation | TCOMA | Tele-Communications, Inc. |
| 03/18/1999 | CMGI | CMGI, Inc. | NSCP | Netscape Communications Corporation |
| 05/05/1999 | UNPH | Uniphase Corporation | JCOR | Jacor Communications Inc. |
| 05/25/1999 | CNET | CNET, Inc. | MCCRK | McCormick \& Company, Incorporated |
| 05/28/1999 | LCOS | Lycos, Inc. | FORE | FORE Systems, Inc. |
| 06/10/1999 | VISX | VISX, Incorporated | NOBE | Nordstrom, Inc. |
| 06/25/1999 | SEBL | Siebel Systems, Inc. | ASND | Ascend Communications, Inc. |
| 07/14/1999 | CNXT | Conexant Systems, Inc. | AMFM | Chancellor Media Corporation |


| 08/04/1999 | CIEN | CIENA Corporation | QNTM | Quantum Corporation |
| :---: | :---: | :---: | :---: | :---: |
| 09/09/1999 | GBLX | Global Crossing Ltd | FDLNB | Food Lion, Inc. |
| 10/06/1999 | EBAY | eBay Inc. | CNTO | Centocor, Inc. |
| 10/27/1999 | VSTR | VoiceStream Wireless Corp. | CEXP | Corporate Express, Inc. |
| 11/22/1999 | RNWK | RealNetworks, Inc. | COMR | Comair Holdings, Inc. |
| 12/20/1999 | ADLAC | Adelphia Communications Corporation | WTHG | Worthington Industries, Inc. |
| 12/20/1999 | AMCC | Applied Micro Circuits Corporation | CBRL | CBRL Group Inc. |
| 12/20/1999 | BVSN | BroadVision, Inc. | FHCC | First Health Group Corp. |
| 12/20/1999 | DISH | EchoStar Communications Corporation | ROST | Ross Stores, Inc. |
| 12/20/1999 | ITWO | i2 Technologies, Inc. | FAST | Fastenal Company |
| 12/20/1999 | LGTO | Legato Systems, Inc. | ANDW | Andrew Corporation |
| 12/20/1999 | MEDI | MedImmune, Inc. | EFII | Electronics for Imaging, Inc. |
| 12/20/1999 | MFNX | Metromedia Fiber Network, Inc. | RTRSY | Reuters Group PLC |
| 12/20/1999 | NSOL | Network Solutions, Inc. | CATP | Cambridge Technology Partners, Inc. |
| 12/20/1999 | NTAP | Network Appliance, Inc. | ADSK | Autodesk, Inc. |
| 12/20/1999 | NXLK | NEXTLINK Communications, Inc. | MUEI | Micron Electronics, Inc. |
| 12/20/1999 | PMCS | PMC - Sierra, Inc. | LNCR | Lincare Holdings, Inc. |
| 12/20/1999 | QLGC | QLogic Corporation | STEI | Stewart Enterprises, Inc. |
| 12/20/1999 | RFMD | RF Micro Devices, Inc. | RXSD | Rexall Sundown, Inc. |
| 12/20/1999 | SDLI | SDL, Inc. | TECD | Tech Data Corporation |
| 12/30/1999 | GMST | Gemstar International Group, Limited | QWST | Qwest Communications International, Inc. |
| 01/07/1998 | AMFM | Chancellor Media Corporation | OSSI | Outback Steakhouse, Inc. |
| 01/07/1998 | APOL | Apollo Group, Inc. ClA | IFMX | Informix Corporation |
| 01/07/1998 | AWIN | Allied Waste Industries, Inc | PETM | PETsMART, Inc. |
| 01/07/1998 | CATP | Cambridge Technology Partners, Inc. | BOST | Boston Chicken, Inc. |
| 01/07/1998 | CTXS | Citrix Systems, Inc. | GEMS | Glenayre Technologies, Inc. |
| 01/07/1998 | DURA | Dura Pharmaceuticals, Inc. | KMAG | Komag, Incorporated |
| 01/07/1998 | ERICY | LM Ericsson Telephone Co ADR | INTU | Intuit Inc. |
| 01/07/1998 | IMNX | Immunex Corporation | PAGE | Paging Network, Inc. |
| 01/07/1998 | JCOR | Jacor Communications Inc. | IDXX | IDEXX Laboratories, Inc. |
| 01/07/1998 | RTRSY | Reuters Holdings Plc ADR | RPOW | RPM, Inc. |
| 01/07/1998 | TECD | Tech Data Corporation | CRUS | Cirrus Logic, Inc. |


| 02/11/1998 | RXSD | Rexall Sundown, Inc. | AGREA | American Greetings Corporation |
| :---: | :---: | :---: | :---: | :---: |
| 08/27/1998 | LVLT | Level 3 Communications, Inc. | VKNG | Viking Office Products, Inc. |
| 08/28/1998 | QWST | Qwest Communications International, Inc. | DIGI | DSC Communications Corporation |
| 09/15/1998 | USAI | USA Networks, Inc. | GART | Gartner Group, Inc. |
| 09/16/1998 | YHOO | Yahoo! Inc. | MCIC | MCI Communications Corporation |
| 12/21/1998 | AMZN | Amazon.com, Inc. | GNCI | General Nutrition Companies, Inc. |
| 12/21/1998 | COMR | Comair Holdings, Inc. | OFIS | U.S. Office Products Company |
| 12/21/1998 | DLTR | Dollar Tree Stores, Inc. | SYBS | Sybase, Inc. |
| 12/21/1998 | INTU | Intuit Inc. | ADTN | ADTRAN, Inc. |
| 12/21/1998 | LNCR | Lincare Holdings, Inc. | PHYC | PhyCor, Inc. |
| 12/21/1998 | MCLD | McLeodUSA Incorporated | DURA | Dura Pharmaceuticals, Inc. |
| 12/21/1998 | NTLI | NTL Incorporated | OXHP | Oxford Health Plans, Inc. |
| 12/21/1998 | VRTS | VERITAS Software Corporation | WCLX | Wisconsin Central Transportation |
| 12/21/1998 | VTSS | Vitesse Semiconductor Corporation | PAIR | PairGain Technologies, Inc. |
| 12/30/1998 | SANM | Sanmina Corporation | AWIN | Allied Waste Industries, Inc. |
| 05/22/1997 | NSCP | Netscape Communications Corporation | GATE | Gateway 2000, Inc. |
| 06/04/1997 | MCHP | Microchip Technology, Incorporated | SNDT | SunGard Data Systems, Inc. |
| 06/12/1997 | MLHR | Herman Miller Incorporated | USRX | U.S. Robotics Corporation |
| 06/12/1997 | ROST | Ross Stores, Inc. | RWIN | Republic Industries, Inc. |
| 07/01/1997 | OFIS | U.S. Office Products Company | CSCC | Cascade Communications Corp. |
| 07/24/1997 | JJSC | Jefferson Smurfit Corporation | STRY | Stryker Corporation |
| 08/13/1997 | STEI | Stewart Enterprises, Inc. Cls A | FORT | Fort Howard Corporation |
| 10/15/1997 | SPOT | PanAmSat Corporation | TYSNA | Tyson Foods, Inc. Class A |
| 02/29/1996 | OXHP | Oxford Health Plans, Inc. | BNET | Bay Networks, Inc. |
| 07/10/1996 | ASND | Ascend Communications, Inc. | STRM | StrataCom, Inc. |
| 07/17/1996 | CSCC | Cascade Communications Corp. | USHC | U.S. Healthcare, Inc. |
| 08/02/1996 | RWIN | Republic Industries, Inc. | AMER | America Online, Inc. |
| 09/30/1996 | FORE | FORE Systems, Inc. | AESC | AES Corporation (The) |
| 11/25/1996 | SBUX | Starbucks Corporation | STJM | St. Jude Medical, Inc. |
| 12/23/1996 | ADTN | ADTRAN, Inc. | AKLM | Acclaim Entertainment, Inc. |
| 12/23/1996 | BBBY | Bed Bath \& Beyond Inc. | ASAI | Atlantic Southeast Airlines, Inc. |
| 12/23/1996 | CEXP | Corporate Express, Inc. | BOBE | Bob Evans Farms, Inc. |


| 12/23/1996 | CEFT | Concord EFS, Inc. | GIDL | Giddings \& Lewis, Inc. |
| :---: | :---: | :---: | :---: | :---: |
| 12/23/1996 | EFII | Electronics for Imaging, Inc. | JBHT | J.B. Hunt Transport Services, Inc. |
| 12/23/1996 | FISV | Fiserv, Inc. | KELYA | Kelly Services, Inc. |
| 12/23/1996 | FORT | Fort Howard Corporation | LRCX | Lam Research Corporation |
| 12/23/1996 | KMAG | Komag, Incorporated | MLHR | Herman Miller, Inc. |
| 12/23/1996 | MCAF | McAfee Associates, Inc. | MTEL | Mobile Telecommunication Technologies Corp. |
| 12/23/1996 | PAIR | PairGain Technologies, Inc. | PRGO | Perrigo Company |
| 12/23/1996 | PHYC | PhyCor, Inc. | RDRT | Read-Rite Corporation |
| 12/23/1996 | QTRN | Quintiles Transnational Corp. | SHLM | A. Schulman, Inc. |
| 12/23/1996 | SNDT | SunGard Data Systems, Inc. | SSSS | Stewart \& Stevenson Services, Inc. |
| 12/23/1996 | SNPS | Synopsys, Inc. | VCELA | Vanguard Cellular Systems, Inc. |
| 12/23/1996 | WCLX | Wisconsin Central Transportation | WMTT | Willamette Industries, Inc. |
| 12/31/1996 | FAST | Fastenal Company | MFST | MFS Communications Company, Inc. |
| 01/27/1995 | RPOW | RPM, Inc. | DOLR | Dollar General Corporation |
| 01/27/1995 | STRM | StrataCom, Inc. | MGMA | Magma Power Company |
| 02/13/1995 | SHLM | A. Schulman, Inc. | QVCN | QVC, Inc. |
| 02/27/1995 | KLAC | KLA Instruments Corporation | SMLS | SciMed Life Systems, Inc |
| 03/06/1995 | HBOC | HBO \& Company | SONO | Sonoco Products Company |
| 07/05/1995 | MFST | MFS Communications Company, Inc. | LOTC | Lotus Development Corporation |
| 07/19/1995 | ALTR | Altera Corporation | LGNT | LEGENT Corporation |
| 08/18/1995 | NWAC | Northwest Airlines Corporation | BRNO | Bruno's Inc. |
| 09/27/1995 | USRX | U.S. Robotics Corporation | PTCM | Pacific Telecom, Inc. |
| 10/03/1995 | GEMS | Glenayre Technologies, Inc. | LINB | LIN Broadcasting Corporation |
| 11/06/1995 | AMER | America Online, Inc. | PHYB | Pioneer Hi-Bred International, Inc. |
| 11/27/1995 | GATE | Gateway 2000, Inc. | ROAD | Roadway Services, Inc. |
| 11/27/1995 | INTU | Intuit Inc. | MMEDC | MultiMedia, Inc. |
| 12/18/1995 | BOST | Boston Chicken, Inc. | ACCOB | Adolph Coors Company |
| 12/18/1995 | GART | Gartner Group, Inc. Class A | ALEX | Alexander \& Baldwin, Inc. |
| 12/18/1995 | GNCI | General Nutrition Companies, Inc. | ASTA | AST Research, Inc. |
| 12/18/1995 | IDXX | IDEXX Laboratories, Inc. | CHRS | Charming Shoppes, Inc. |
| 12/18/1995 | MUEI | Micron Electronics, Inc. | HONI | Hon Industries Inc. |
| 12/18/1995 | MXIM | Maxim Integrated Products, Inc. | INEL | Intelligent Electronics, Inc. |


| $12 / 18 / 1995$ | PETM | PETsMART, Inc. | NDSN | Nordson Corporation |
| :--- | :--- | :--- | :--- | :--- |
| $12 / 18 / 1995$ | PSFT | PeopleSoft, Inc. | TECUA | Tecumseh Products Company Class A |
| $12 / 18 / 1995$ | RDRT | Read-Rite Corporation | YELL | Yellow Corporation |


[^0]:    * Hyun Bin Kang and Sung Woo Park were students at the Department of Statistics when this work was done. Ms. Kang is currently a statistics graduate student at Pennsylvania State University, and Mr. Park is a corporal in the South Korean army.

[^1]:    ${ }^{1}$ These include vendor websites and Yahoo Finance when the latter provided downloadable current constituents; it does not do so for most indexes now.

[^2]:    ${ }^{2}$ In fact we have found minor errors on the Dow Jones constituent history website.
    ${ }^{3}$ Institutional Brokers' Estimate System (I/B/E/S)

[^3]:    ${ }^{4}$ permnolist <- subset(dsp500, start <= date \& ending > date)

[^4]:    ${ }^{5}$ We had mis-labeled one date in the calibration data, which purported to be from 12/29/04; when compared with the DSP500 list for $12 / 4 / 04$, there were five discrepancies including Ford Motor, but there was only one problem with Ford when it was compared with the list for $12 / 29 / 04$. It turned out that when the list dated $12 / 4 / 04$ was compared with the list dated $12 / 29 / 04$ from Standard \& Poor's, they all matched. Therefore it was concluded that the list was labeled incorrectly, and it actually pertained to the date $12 / 29 / 04$. Therefore, it was cited as 20041229 in the result table above

[^5]:    ${ }^{6}$ This is found on WRDS in the CRSP as CRSP/Compustat Merged Database - Linking Table

