

# Working with International Financial Data Sets: Potential Issues and Solutions

November 8, 2018

Todd Hines, Stanford University  
Shikha Sharma, MIT

# International Company Research

- Historically research involving companies has focused on US publicly traded companies.
  - See the Todd Hines and Kathleen Dreyer presentation at 2014 Beyond the Numbers conference, “Linking Financial Data Sets: Possible Problems and Solutions”
- More and more interest in non-US research. More interest in private non-US company research

# International Company Research (cont.)

- More issues than working with only US data
  - More companies
  - More countries with different accounting and disclosure rules
- Not many guides/overviews have been attempted
- Not possible to do comprehensive overview in allotted time period
  - Presentation will concentrate on some of the major issues

# This presentation

- Discussion of important issues
  - Todd
- Discussion of heavily-used sources
  - Todd
- Primary company ids used by international companies
  - Shikha
- Examples of linking between databases using international company data
  - Shikha

# Important Issues with Intl. Company Research Data

- Backfiles not as deep as for US public companies
- Company IDs
  - More complex
  - More issues with linking
- More data providers
  - For US – Compustat and CRSP
  - For International – many more competitors

# International Fundamentals (Financials) Databases

- Backfile dates
- Three primary fundamentals databases
  - Worldscope (Factset)
  - Compustat Global
  - Osiris
- Comparison of coverage prepared by WRDS
  - <https://bit.ly/2CSN6ii>
    - International Accounting Databases on WRDS: Comparative Analysis. Rui Dai. Wharton Research Data Services Working Paper (2012).
    - Note that in this document Worldscope = Factset

# International company – Daily Stock Price Data

- Backfile dates
- Most heavily used sources
  - Datastream
  - Compustat Global
    - Advantage - stock price links easily to Compustat Global fundamentals
  - Bloomberg
    - Has download limits

# International Private Company

- Coverage dependent upon disclosure requirements of each country
  - United States vs. France
- Orbis
  - Most requested
  - Many weaknesses
  - Paper with good overview of the coverage for many of the countries
    - Kalemli-Ozcan, Sebnem, et al. How to construct nationally representative firm level data from the ORBIS global database. No. w21558. NBER, (2015).
- Capital IQ
  - Coverage not documented
- WorldBase
  - Only scholar who has ever gotten access is Laura Alfaro at Harvard Business School



# Identifiers

Used to:

- Track entities and securities through time
- Merge datasets
- Filter data

# Major Identifiers

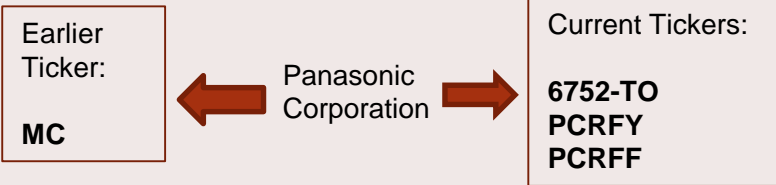
## Commonly used identifiers:

- Company Name
- Ticker
- CUSIP
- ISIN
- SEDOL

## Proprietary identifiers

- GVKEY
- BvD ID
- Worldscope ID

# Identifiers: Ticker & Company Name

Identifier	Example	Issues
Company Name	Toyota Motor Corp Toyota Motor Corporation Toyota Mtr Cr Corp	Names may be abbreviated differently making them unusable for constructing large datasets.
Ticker	<p>ZEN - was ticker for Zeneca before it became AstraZeneca. Currently, assigned to ZenDesk</p>  <p>The diagram illustrates a company, Panasonic Corporation, with two different tickers. On the left, a box labeled 'Earlier Ticker:' contains the ticker 'MC'. On the right, a box labeled 'Current Tickers:' contains three tickers: '6752-TO', 'PCRFY', and 'PCRFF'. A double-headed arrow connects the two boxes, with 'Panasonic Corporation' written above it.</p>	<p>Tickers can change. They can also be reassigned.</p> <p>A company can have multiple tickers</p>

# Identifiers: CUSIP

Identifier	Properties	Issues
CUSIP (Committee on Uniform Securities Identification Procedures)	Used to identify securities trading in US and Canadian exchanges.  6, 8, or 9 characters  CUSIPs for Toyota Motor:  <b>892331</b> – first 6 digits identify the issuer 892331 <b>30</b> – next 2 digits identify type of security 89233130 <b>7</b> – 9 <sup>th</sup> character is a check digit	Reliable but can change with change in company name, merger or other developments  Never reassigned  May have to convert 6 digits into 8 or 9 digits for linking datasets

# Identifiers: ISIN & SEDOL

Identifier	Properties	Issues
ISIN (International Securities Identification Number)	Consists of 12 alphanumeric characters  Toyota Motor Corp : <b>JP36334400001</b>	Reliable but can change due to merger, change in company's name and other events.  Never reassigned
SEDOL (Stock Exchange Daily Official List)	Consists of 7 characters  Toyota Motor Corp: <b>6900643</b>	SEDOLs can change overtime due to change in company name, mergers, change in company headquarters and other events

# Proprietary Identifiers

Identifier	Example	Issues
Compustat GVKEY	<p>Represents a company</p> <p>Usually a 6 digit number</p> <p>GVKEY for: Panasonic Corporation: <b>007114</b></p>	<p>Most reliable identifier if working within Compustat only</p> <p>GVKEY does not change through the life of a company.</p> <p>Panasonic's history goes back to 1918 but the GVKEY for the company remains the same</p>

# Identifiers Available in Major International Financial Datasets

Database	Ticker	CUSIP	ISIN	SEDOL	GVKEY	Other
Compustat Global	Yes	No	Yes (in output)	Yes	Yes	SIC & NAICS
Datastream	Yes	Yes	No	No	No	DS Code
ORBIS	Yes	No	Yes	No (as of Oct. 2015)	No	BvD ID
SDC (Global New Issues)	Yes	Yes (6 digit; in output)	Yes	Yes	No	
SDC (M&A)	Yes	Yes (6 digit; in output)	No	Yes		
Worldscope	Yes	Yes (9 digit)	Yes	Yes	No	WS Perm ID IBES Ticker
Zephyr	Yes	No	Yes	No	No	BvD ID

# Other Issues with Identifiers

- **Header vs. Historical Identifiers**

- Headers present the most current data. In most datasets, variables such as company name, ticker, CUSIP, ISIN, address, and industry codes reflect only the most current information. Earlier name changes and identifiers are not retained.
- Historical identifiers have date ranges in which they are valid. Compustat Global includes both header and historical industry codes but other identifiers are header data.
- Understand which type of identifier your data uses.

- **Issue vs. Entity**

- Some databases such as SDC provide 6 digit CUSIPs that represent the issuer. Other datasets with 8 and 9 digit CUSIPs correspond to the security.



# Let's try an example!

1990s saw big mergers in the pharmaceutical industry

Is pharmaceutical **merger** and acquisition the key to strengthening the pipeline?

467 words

21 August 2000

Chemical Marketing Reporter

CMKRA

English

(c) 2000 Royal Society of Chemistry

One of the most significant issues ongoing in the pharmaceutical industry is the pressure to build critical mass to finance research and development. But the key question is whether mergers and acquisitions really strengthen the product pipeline. The pharmaceutical industry is forecast to invest \$26.4 bn in research and development in 2000, an all-time high. This is a 10% rise on 1999's level of \$24.0 bn. Worldwide pharmaceutical **merger** and acquisition transactions reached an estimated \$71.8 bn in 1999, three times the 1998 level of \$24.3 bn. among the top 20 pharmaceutical companies in 1999 and 2000, five key mergers took place. These include the **merger** of **Astra** and **Zeneca** to create **AstraZeneca**; the **merger** of Rhone-Poulenc and Hoechst Marion Roussel to create Aventis; the **merger** of Pharmacia and Upjohn with Monsanto to create Pharmacia; the **merger** of Warner-Lambert and Pfizer; and the **merger** of **Glaxo Wellcome** and **SmithKline Beecham** which is scheduled to close in Sep 2000.

# Major Sources for M&A transactions, Financial Statements and Security Prices

**SDC Platinum:** Provides M&A transactions and for companies worldwide.

**Worldscope:** Provides financial statements and annual security prices for US and non-US companies. To pull data:

- Use SEDOLs if you have them OR convert 6 digit CUSIPs from SDC into 9 digits to make them work in Worldscope

**Compustat Global:** Provides financial statements and daily security prices for companies outside North America. To pull data:

- Use SEDOLs if available OR conduct an industry search around relevant dates. Filter out companies that don't match your results from SDC.

Source: Thomson Reuters SDC Platinum														
Date Announced	Date Effective	Target Name	Target CUSIP	Target's Sedol	Target Ticker Symbol	Target Primary SIC Code	Target Industry Sector	Target Nation	Acquiror Name	Acquiror CUSIP	Acquiror's Sedol	Acquiror Ticker Symbol	Acquiror Primary SIC Code	Value of Transaction (\$mil)
01/20/95	05/01/95	Wellcome PLC	949478	0947293	WCM	2834	Drugs	United Kingdom	Glaxo Holdings PLC	377327	0371784	GLX	2834	14,284.837
09/13/95	09/13/95	Cascan GmbH(Merck)	14748C			2834	Drugs	Germany	Glaxo Wellcome PLC	37733W	0371784	GLX	2834	
12/05/97	01/28/98	Pofa Poznan(Poland)	73115Q			2834	Drugs	Poland	Glaxo Wellcome PLC	37733W	0371784	GLX	2834	220.001
01/14/98	01/14/98	Zeneca Abdi Ibrahim Ilac	98930Z			2834	Drugs	Turkey	ZENECA Group PLC	98934D	0989529	ZEN	2834	
06/09/98	06/09/98	Mundogen	62613W			2834	Drugs	Spain	Glaxo Wellcome PLC	37733W	0371784	GLX	2834	
12/09/98	04/19/99	Astra AB	046298		ASTZ	2834	Drugs	Sweden	ZENECA Group PLC	98934D	0989529	ZEN	2834	34,636.776
01/17/00	12/27/00	SmithKline Beecham PLC	832378	0135656	SBH SBH	2833	Drugs	United Kingdom	Glaxo Wellcome PLC	37733W	0371784	GLX	2834	75,960.847
05/15/06	08/22/06	Cambridge Antibody Tech Grp	132148	0166225	CATGY	2834	Drugs	United Kingdom	AstraZeneca PLC	046353	0989529	AZN	2834	1,035.062
02/01/07	03/02/07	Arrow Therapeutics Ltd	04282P			2834	Drugs	United Kingdom	AstraZeneca PLC	046353	0989529	AZN	2834	150.001
04/22/08	06/02/08	Sirtis Pharmaceuticals Inc	82968A	B1XN4P5	SIRT	2834	Drugs	United States	GlaxoSmithKline PLC	37733W	0925288	GSK	2834	715.584
10/29/08	01/07/09	Genelabs Technologies Inc	368706	B0V2509	GNLB	2834	Drugs	United States	GlaxoSmithKline PLC	37733W	0925288	GSK	2834	56.789




Ticker, CUSIP, SEDOL, ISIN, Company Name and SIC Codes are all header data!

Source: Worldscope

year_	Ticker	CUSIP	SEDOL	ISIN	Employees	Total Assets	Company Name	SIC 1	SIC 2
1995	AZN	46353108	2989044	US0463531089	32000	7841640000	ASTRAZENECA PLC	2834	5122
1996	AZN	46353108	2989044	US0463531089	30400	8510280100	ASTRAZENECA PLC	2834	5122
1997	AZN	46353108	2989044	US0463531089	32100	8202995070	ASTRAZENECA PLC	2834	5122
1998	AZN	46353108	2989044	US0463531089	34600	8953539630	ASTRAZENECA PLC	2834	5122
1999	AZN	46353108	2989044	US0463531089	58000	19345319746	ASTRAZENECA PLC	2834	5122
2000	AZN	46353108	2989044	US0463531089	52300	18245032841	ASTRAZENECA PLC	2834	5122
2001	AZN	46353108	2989044	US0463531089	54000	17838950051	ASTRAZENECA PLC	2834	5122
2002	AZN	46353108	2989044	US0463531089	58700	21350053589	ASTRAZENECA PLC	2834	5122
1995	GSK	37733W105	2374112	US37733W1053	52419	13257806400	GLAXOSMITHKLINE PLC	2834	8731
1996	GSK	37733W105	2374112	US37733W1053	53808	13907544260	GLAXOSMITHKLINE PLC	2834	8731
1997	GSK	37733W105	2374112	US37733W1053	53068	13536344510	GLAXOSMITHKLINE PLC	2834	8731
1998	GSK	37733W105	2374112	US37733W1053	54350	15114384180	GLAXOSMITHKLINE PLC	2834	8731
1999	GSK	37733W105	2374112	US37733W1053	109041	26715688870	GLAXOSMITHKLINE PLC	2834	8731
2000	GSK	37733W105	2374112	US37733W1053	107517	27423929960	GLAXOSMITHKLINE PLC	2834	8731
2001	GSK	37733W105	2374112	US37733W1053	107470	26360916000	GLAXOSMITHKLINE PLC	2834	8731
2002	GSK	37733W105	2374112	US37733W1053	104499	29191699680	GLAXOSMITHKLINE PLC	2834	8731
2003	GSK	37733W105	2374112	US37733W1053	100919	35283843890	GLAXOSMITHKLINE PLC	2834	8731
2004	GSK	37733W105	2374112	US37733W1053	100019	40385884990	GLAXOSMITHKLINE PLC	2834	8731

ISINs can be used to pull annual security prices from Worldscope

Data on Target Companies from Worldscope											
year_	freq	Ticker	CUSIP	SEDOL	ISIN	Employees	Total Assets	Company Name	NATION	SIC 1	SIC 2
1995			46298105	2037954	US0462981053	16962	6535392500	ASTRA AB- ADR	SWEDEN	2835	2834
1996			46298105	2037954	US0462981053	19851	7578746880	ASTRA AB- ADR	SWEDEN	2835	2834
1997			46298105	2037954	US0462981053	22206	7843543200	ASTRA AB- ADR	SWEDEN	2835	2834
1998			46298105	2037954	US0462981053	24958	9387815400	ASTRA AB- ADR	SWEDEN	2835	2834
1999			46298105	2037954	US0462981053			ASTRA AB- ADR	SWEDEN	2835	2834
1997		CATG	132148107	2765909	US1321481079	70	76807302.04	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
1998		CATG	132148107	2765909	US1321481079	123	77654749.48	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
1999		CATG	132148107	2765909	US1321481079	150	57949053.75	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2000		CATG	132148107	2765909	US1321481079	161	249906239.5	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2001		CATG	132148107	2765909	US1321481079	247	254005192.5	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2002		CATG	132148107	2765909	US1321481079	293	245937599.2	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2003		CATG	132148107	2765909	US1321481079	279	230533623.4	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2004		CATG	132148107	2765909	US1321481079	281	219809272.7	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2005		CATG	132148107	2765909	US1321481079	284	366534717.8	CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2006		CATG	132148107	2765909	US1321481079			CAMBRIDGE ANTIBODY TECHNOLOGY GROUP PLC	UNITED KINGDOM	8731	
2004		SIRT	82968A105	B1XN4P5	US82968A1051			SIRTRIS PHARMACEUTICALS, INC.	UNITED STATES	2834	
2005		SIRT	82968A105	B1XN4P5	US82968A1051		34720000	SIRTRIS PHARMACEUTICALS, INC.	UNITED STATES	2834	
2006		SIRT	82968A105	B1XN4P5	US82968A1051	42	52086000	SIRTRIS PHARMACEUTICALS, INC.	UNITED STATES	2834	
2007		SIRT	82968A105	B1XN4P5	US82968A1051	57	125215000	SIRTRIS PHARMACEUTICALS, INC.	UNITED STATES	2834	
2008		SIRT	82968A105	B1XN4P5	US82968A1051			SIRTRIS PHARMACEUTICALS, INC.	UNITED STATES	2834	

SEDOL Search Results from Compustat Global														
	gvkey	indfmt	consol	fyear	fyr	datadate	at	emp	exchg	isin	sedol	conm	fic	sic
 GVKEY for each Company remains the same through all the years.	5180	INDL	C	1995	12	19951231	8541	52.419	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	1996	12	19961231	8314	53.808	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	1997	12	19971231	8437	53.068	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	1998	12	19981231	9346	55.273	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	1999	12	19991231	10427	61.83	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2000	12	20001231	21590	108.201	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2001	12	20011231	21917	107.899	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2002	12	20021231	22327	106.166	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2003	12	20031231	23975	103.166	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2004	12	20041231	22578	99.837	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
Company name, ISIN, SEDOL, SIC are header data.	5180	INDL	C	2005	12	20051231	27198	100.728	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2006	12	20061231	25553	102.695	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2007	12	20071231	31003	103.483	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2008	12	20081231	39393	99.003	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	5180	INDL	C	2009	12	20091231	42862	99.913	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
	28272	INDL	C	1995	12	19951231	5093	31.4	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
	28272	INDL	C	1996	12	19961231	5000	31.1	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
	28272	INDL	C	1997	12	19971231	5010	31.4	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
	28272	INDL	C	1998	12	19981231	5410	34	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
	28272	INDL	C	1999	12	19991231	19816	58	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
28272	INDL	C	2000	12	20001231	18434	57	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2001	12	20011231	17985	52.6	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2002	12	20021231	21576	58.7	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2003	12	20031231	23573	61.9	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2004	12	20041231	25616	64.2	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2005	12	20051231	24840	65.3	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2006	12	20061231	29932	66	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2007	12	20071231	47957	67.4	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	
28272	INDL	C	2008	12	20081231	46784	65	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834	

## Compustat Global SIC Code Search Results

gvkey	indfmt	datafmt	consol	fyear	datadate	at	emp	exchg	isin	sedol	conm	fic	sic
5180	INDL	HIST_STD	C	1995	19951231	8541	52.419	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
5180	INDL	HIST_STD	C	1996	19961231	8314	53.808	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
5180	INDL	HIST_STD	C	1997	19971231	8437	53.068	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
5180	INDL	HIST_STD	C	1998	19981231	9346	55.273	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
5180	INDL	HIST_STD	C	1999	19991231	10427	61.83	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
5180	INDL	HIST_STD	C	2000	20001231	21590	108.201	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
5180	INDL	HIST_STD	C	2001	20011231	21917	107.899	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
5180	INDL	HIST_STD	C	2002	20021231	22327	106.166	194	GB0009252882	925288	GLAXOSMITHKLINE PLC	GBR	2834
28272	INDL	HIST_STD	C	1995	19951231	5093	31.4	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
28272	INDL	HIST_STD	C	1996	19961231	5000	31.1	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
28272	INDL	HIST_STD	C	1997	19971231	5010	31.4	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
28272	INDL	HIST_STD	C	1998	19981231	5410	34	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
28272	INDL	HIST_STD	C	1999	19991231	19816	58	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
28272	INDL	HIST_STD	C	2000	20001231	18434	57	194	GB0009895292	989529	ASTRAZENECA PLC	GBR	2834
30938	INDL	HIST_STD	C	1995	19951231	43715	16.962	256	SE0000101008	5241503	ASTRA AB	SWE	2834
30938	INDL	HIST_STD	C	1996	19961231	52224	19.851	256	SE0000101008	5241503	ASTRA AB	SWE	2834
30938	INDL	HIST_STD	C	1997	19971231	62280	22.206	256	SE0000101008	5241503	ASTRA AB	SWE	2834
30938	INDL	HIST_STD	C	1998	19981231	76138	24.958	256	SE0000101008	5241503	ASTRA AB	SWE	2834

Compustat Daily Prices for Astra and Zeneca around Deal Announcement Date							
gvkey	datadate	conm	prccd	exchg	isin	sedol	dldte
28272	19981207	ASTRAZENECA PLC	24.2	194	GB0009895292	989529	
28272	19981208	ASTRAZENECA PLC	25.2	194	GB0009895292	989529	
28272	19981209	ASTRAZENECA PLC	27.11	194	GB0009895292	989529	
28272	19981210	ASTRAZENECA PLC	25.9	194	GB0009895292	989529	
28272	19981211	ASTRAZENECA PLC	25.75	194	GB0009895292	989529	
28272	19981214	ASTRAZENECA PLC	25.14	194	GB0009895292	989529	
28272	19981215	ASTRAZENECA PLC	25.87	194	GB0009895292	989529	
28272	19981216	ASTRAZENECA PLC	26.36	194	GB0009895292	989529	
28272	19981217	ASTRAZENECA PLC	26.3	194	GB0009895292	989529	
28272	19981218	ASTRAZENECA PLC	26.14	194	GB0009895292	989529	
30938	19981207	ASTRA AB	<b>147</b>	256	SE0000101008	5241503	19990426
30938	19981208	ASTRA AB	<b>151</b>	256	SE0000101008	5241503	19990426
30938	19981209	ASTRA AB	<b>170.5</b>	256	SE0000101008	5241503	19990426
30938	19981210	ASTRA AB	<b>160.5</b>	256	SE0000101008	5241503	19990426
30938	19981211	ASTRA AB	<b>162</b>	256	SE0000101008	5241503	19990426
30938	19981214	ASTRA AB	<b>161</b>	256	SE0000101008	5241503	19990426
30938	19981215	ASTRA AB	<b>157.5</b>	256	SE0000101008	5241503	19990426
30938	19981216	ASTRA AB	<b>164.5</b>	256	SE0000101008	5241503	19990426
30938	19981217	ASTRA AB	<b>163.5</b>	256	SE0000101008	5241503	19990426
30938	19981218	ASTRA AB	<b>161.5</b>	256	SE0000101008	5241503	19990426
30938	19981207	ASTRA AB	142	256	SE0000101016	5205974	19990426
30938	19981208	ASTRA AB	146	256	SE0000101016	5205974	19990426
30938	19981209	ASTRA AB	169	256	SE0000101016	5205974	19990426
30938	19981210	ASTRA AB	158	256	SE0000101016	5205974	19990426
30938	19981211	ASTRA AB	161.5	256	SE0000101016	5205974	19990426
30938	19981214	ASTRA AB	159.5	256	SE0000101016	5205974	19990426
30938	19981215	ASTRA AB	156.5	256	SE0000101016	5205974	19990426
30938	19981216	ASTRA AB	162.5	256	SE0000101016	5205974	19990426
30938	19981217	ASTRA AB	162.5	256	SE0000101016	5205974	19990426
30938	19981218	ASTRA AB	160.5	256	SE0000101016	5205974	19990426

AstraZeneca  
data here is  
for  
Zeneca  
before  
the merger



# Summing Up

- More choices available for intl financial data. Understand what your institution owns.
- Backfile dates are shorter than for US financial data.
- Which identifier to use? Will depend on your starting point.
- If starting from scratch, begin in a dataset that can yield most number of identifiers.
- Use tickers and company names sparingly as they are least reliable.
- Remember that most datasets only provide the most current company name, address, ticker, CUSIPs and ISINs – also known as header data.
- Only file we know with previous company identifiers
  - WRDS Capital IQ Identifiers file
    - Only goes back to late 90s or so