International Accounting Databases on WRDS: Comparative Analysis

Rui Dai[†]
Wharton Research Data Services
University of Pennsylvania

Abstract

While multiple data vendors have claimed to offer a comprehensive international coverage of accounting and financial items for firms worldwide, there is a growing demand for better understanding of the differences across international databases. This paper compares several international accounting databases on WRDS to address this traditional obstacle confronting empirical researchers. I document important issues on the data coverage at the firm and country level, the sample overlap, and the discrepancies in coverage of various data items for FactSet Fundamentals, Compustat Global, and Bureau Van Dijk's international databases. The main conclusions are threefold: 1) Compustat Global features greater coverage of large companies in more developed countries and provides a wider range of accounting data items than any other databases; 2) BvD Osiris offers lesser variety of accounting data items, but it also contains a higher number of small firms from developing countries; 3) finally, FactSet Fundamentals Database provides a balance in the firm size and quantity of accounting items with a reasonable geographical coverage.

This version: March 16th, 2012

[†] I would like to gratefully acknowledge the comments and suggestions from Denys Glushkov, Mark Keintz, Rabih Moussawi and Luis Palacios. The views expressed in this study are personal of the author and do not necessarily reflect those of WRDS. All errors are my own. Address correspondence to Rui Dai, Wharton Research Data Services, The Wharton School, University of Pennsylvania 3819 Chestnut Street, Suite 300, Philadelphia, PA 19104, or e-mail: rdai@wharton.upenn.edu.

Electronic copy available at: https://ssrn.com/abstract=2938675

1. Introduction

Empirical research in international finance and accounting is a field that has gained increased interest from academics and analysts around the world in the past decade. The reliability of those studies is ultimately determined by the sample coverage and data quality of the international accounting data. A normal practice to choose a source of data among many available alternatives is to follow the convention established in previous literature, which, in turn, may also be affected by the access to different data sources. Such practice has been the standard approach for the US financial and accounting studies. However, the existing international studies diverge widely in the use of accounting data, presumably, due to the availability and suitability of international data as well as the lack of tradition in literature.

Currently, three well-known vendors of global financial statement data on WRDS platform are FactSet Fundamentals (a product built on a copy of Worldscope), Compustat Global (Global Vantage), and the Bureau Van Dijk (Osiris and Amadeus). These databases provide a wide coverage at both firm and country level, and have their proprietary procedures to ensure the consistency among the accounting items from different countries and across accounting standards. In addition to accounting items, they sometimes provide other closely related market items, such as market price, and foreign exchange rate. This paper is the first formal examination of the data properties of those databases to provide researchers with some insights for their choice on these data sources. Even though several studies have examined and compared accounting databases (e.g. Lara, Osma, and Nogue 2006), these previous papers usually focus on the data properties within a single country or region. Additionally, this paper also offers a guide on data issues and collection procedures to help researchers choose an international accounting database that is adequate for their research needs.

Electronic copy available at: https://ssrn.com/abstract=2938675

The results show that BvD Osiris provides the broadest coverage at both the firm and the country level, while Compustat Global tends to include the largest companies around world. FactSet fundamental is more likely to include a large number of companies in the more developed markets. The empirical evidence also suggests that FactSet Fundamentals has larger number of observations and longer covered time periods for some frequently used accounting variables. Osiris may be the best candidate for the researchers who want to investigate the companies in emerging markets, while Compustat Global may be a database of choice for the researchers in global ETF and Indices.

The paper proceeds as follows. Section 2 deals with some background of three international accounting databases. Section 3 explains the general structure of those databases, discussing in particular some potential problems in each database. Section 4 presents some related literature for database comparisons. Section 5 provides the data preparation, which is quite detailed and intended as a guide for data collection for each database. Section 6 compares the coverage of each database, and Section 7 illustrates a qualitative comparison for some common data items from different databases. Section 8 concludes.

2. International Accounting Databases

The accounting database is one of the most crucial components in the fields of empirical finance and accounting research. Compustat North America has long been recognized as the primary source for accounting data for the empirical studies of US public companies, largely due to the data availability and the common practice of empirical research. However, there is no such agreement among researchers in international finance. A keyword search among the finance and accounting journals used by *the Financial Times* in compiling the Business School research rank indicates that finance studies tend to prefer the Worldscope, whereas accounting research

Electronic copy available at: https://ssrn.com/abstract=2938675

appears to slightly favor Compustat Global. Table 1 offers a brief summary of the databases cited in the papers published in the top finance and accounting journals. Below is a brief review of international databases currently available on WRDS.

FactSet Fundamentals:

In 2008, FactSet acquired a copy of Worldscope and a forwarding right of reuse to develop and brand it as FactSet Fundamentals (FactSet).[1] Due to this twin feature, the FactSet Fundamentals database shares a great deal of similarity with the Thomson Reuters' Worldscope. For example, 1) FactSet contains general and segment financial information for a board selection of firms and countries since 1980 at both annual and quarterly basis; 2) FactSet focuses on a relatively small number of large companies in early years; and 3) The financial statement data in FactSet are reported in a standardized format according to a proprietary global template to provide a dataset with international comparability.[2] However, FactSet available at WRDS seemingly reports relatively fewer variables than the Worldscope. For example, its annual data contains 665 time-series accounting items, including ratios and derived values, while Ulbricht, and Weiner (2005) documents that Worldscope includes 1,284 time-series annual variables.

S&P's Compustat Global:

[1] In April 2008 FactSet purchased a copy of the Worldscope Database from Thomson Reuters which has been providing daily updates to the database until April 2010. As a part of contract, FactSet hired some previous Worldscope team members responsible for managing, maintaining, and collecting the database, and obtained all software for collecting and maintaining the database. Beginning May 2010, FactSet Fundamentals has been solely collected by FactSet. In additional to securities available in Worldscope, FactSet indicated that they added 2,000 new dually listed securities, and has also added nearly 6,000 new securities to the database on a net basis prior to the period in which the paper is written.

[2] Standardization of accounting data means that the value of accounting variable is not presented as reported value in financial statement when a known disparity of definitions exists among the accounting standards. Usually the accounting data are adjusted with the intention of making data more comparable to US GAAP (Garcia Lara et al. 2006). Osma and Pope (2001) offers a comparison of WorldScope with as-reported accounting information from Extel Financials for all EU firms between 1995 and 2006, revealing that WorldScope provides adjusted net income (stockholders' equity) in 47% (59%) of all cases. These differences in net income (stockholders' equity) were over 5% of the original value in over 15% (19%) of cases.

Compustat Global, previously known as Global Vantage, is provided by Standard & Poor's. Similar to Compustat North America, it provides quarterly and annual financial statement information at both consolidated and unconsolidated level as well as the segment data. In particular, the annual accounting items are available since 1987. An important benefit of Compustat to researchers is that data are collected and normalized according to the country accounting principles, disclosure methods and specific data item definitions. Even though the detailed information is proprietary, the standardization of Compustat is widely considered in line with the regulations and standards of the SEC, U.S. GAAP and the IFRS. The Compustat is also well known to provide a wide range of accounting variables. For example, the Compustat Global annual contains 964 time-series of various accounting items with few derived and ratio variables.

BvD's Osiris:

Osiris is a product offered by Bureau van Dijk (BvD), and contains financial information on globally listed public companies, including banks and insurance firms. Osiris strives to cover all publicly listed companies worldwide, as well as other major non-listed firms that are primary subsidiaries of publicly listed firms. Osiris also claims to provide both standardized and "as reported" accounts to reflect the different configurations of their accounts.[3] However, Osiris only collects financial information on annual basis and provides a much smaller selection of variables. In particular, the accounting data for industrial companies only contain 207 time-series variables. Though Osiris provides financial information for four different levels of consolidation with the unique ownership information, it does not provide a ready-to-use segment data. Furthermore, Osiris provide annual accounting items since 1982 and dramatically increased the number of

[3] The response from BvD New York regarding the details of its financial reporting format indicates that Osiris only provides WRDS the standardized accounting items. The evidence in the section 7 of this paper, however, suggests that the standardization procedure of Osiris may largely diverge from others, if the accounting items in Osiris data feed are standardized in a consistent manner.

covered companies in recent years, and, therefore, often has a limited number of years for annual accounting items.

BvD's Amadeus:

Amadeus from BvD is a European financial database containing annual accounting information for over ten million firms from Pan-European countries with a focus on private companies. Due to its specific focus, the number of accounting variables is limited. The annual data only includes 24 balance sheet items, 25 profit and loss account items, and 26 ratios. Similar to Osiris in many aspects, Amadeus provides no explicit segment data, and different consolidation levels, and only offers the most recent 10 year annual report.

3. Database Structure

Data items provided in FactSet, Compustat Global, BvD Osiris and Amadeus can be broadly categorized into two types: descriptive and time series items. The descriptive items are header (most recent) information such as company name, country, and permanent id (which is assumed to be constant across time). The time series items are accounting variables collected at a certain frequency, i.e., annually or quarterly. In most cases, the time-series accounting variables are maintained historically through time, while the descriptive variables are valid as of the date of the latest data vintage.

All four databases provide identification records at both company and security level. In Compustat Global and BvD Osiris and Amadeus company level codes serve as the primary identifier for accounting data record. However, FactSet's primary identifier is at the security level, therefore, WRDS provides a unique one-to-one link between security level and entity

level.[4] Interestingly, FactSet entity not only includes industrial or financial companies, but it also features Venture Capitals, Hedge Funds, and etc.[5]

In addition to database specific identifiers, each database also contains international standard security identifier: ISIN and SEDOL. The ISIN is a 12-digit unique issue identifier, while the SEDOL is 7-digit market level identifier. A cross-market security can have only one ISIN, but it may have multiple SEDOLs. For example, the common shares of PetroChina Company Limited (PetroChina) listed in Hong Kong stock exchange are assigned an ISIN CNE1000003W8. The SEDOL number for this security is 6226576, when it traded in Hong Kong, and it is 5939507, when it is traded in Germany.

ISIN is not a unique identifier for issuers with multiple securities. The NYSE ADR of PetroChina's Hong Kong shares is assigned an ISIN US71646E1001, while its mainland China shares are assigned an ISIN CNE1000007Q1. Indeed, PetroChina has 2,569 active and 694 inactive market level identifiers according to the London Stock Exchange as of this writing (including bonds, non-common stock, preferreds, etc). Additionally, all databases provide ISIN and SEDOL as descriptive (most recent) variables and do not maintain history of ISIN and SEDOL, which complicates the use of ISIN numbers as a common link among databases.

This paper relies primarily on the annual data updated to the end of 2010 to investigate the firm and country coverage and the availability of accounting data. However, as of this writing (Sep-Oct 2011), the 2010 annual data from Amadeus have not been updated to include 2010 data. All databases collect their information from multiple data sources, and different countries have different financial reporting conventions. Therefore, it is not surprising to observe

[4] The detail methodology of this one-to-one link between the security level and the firm level is documented in WRDS FactSet Fundamental Overview, available at http://wrds.wharton.upenn.edu.

[5] When FactSet entity variable is labeled as Extinct, the information about the organizational structure is sometimes more limited.

that the annual data are not available or incomplete for some countries for fiscal year 2010 (2009 for Amadeus). For those reasons and for the sake of consistency with the most recent descriptive information, such as ISIN and country code in 2010, most cross-sectional analyses in this paper are performed using the cross-sectional data in fiscal year 2008.

Since this paper aims to investigate the accounting information at the firm level, careful treatment of different consolidation levels in financial reporting is an important issue. For Osiris and Amadeus, four consolidation codes are used: consolidated and unconsolidated statements with or without unconsolidated companions. Compustat Global indicates whether a financial statement is consolidated or not. Even though FactSet provides certain information about consolidation such as consolidated net income, no explicit information about consolidation level of accounting data is provided. To remove the replications at different consolidation levels in each database, the only records with the highest consolidation are kept in cases when there are multiple records for a firm in a fiscal year in Compustat Global and BvD Osiris and Amadeus.

Each database provides certain geographic information about the issuers. Osiris and Amadeus offer a 2-digit country ISO code, but without further specifications. Compustat Global, on the other hand, provides two 3-digit country ISO codes to indicate where the issuer is incorporated and headquartered. In additional to those two location codes about incorporation and headquarters location, FactSet offers another six 2-digit country codes, such as the country where the issues are listed.

BvD offers IMF exchange rate from local currency to US dollars at the closing date of the statement. Compustat Global also provides an exchange rate record between the local currency and reported currency within the annual dataset, however, does not distinguish between the conversion rates used for balance sheet and income statement as FactSet does.

Compustat and FactSet usually keep the time series history of accounting items since the moment a company is included into the database. In fact, Compustat is well known to have certain tendency to back-fill the accounting data for some small successful firms (see, Fama and French, 1993). BvD databases, on the other hand, have either an explicit rule to keep time series items for certain number of years or a relatively shorter history for accounting items. For example, Amadeus only keeps up to 10 years of time series accounting data, and only 10 percent companies from Osiris *hybrid insurance* have more than 9 years of history.[6] The numbers of time series items also vary among different databases. Compustat offers the largest number of most time series variables (more than 900 variables), while BvD databases, especially Amadeus, provide the least number of accounting variables.

Table 2 demonstrates the direct availability for some frequently-used accounting variables in each database. Both FactSet and Compustat offer all balance sheet and income statement variables (see Table 3), however, some variables, such as Capital Expenditures and Accounts Payable, are not provided in BvD. In addition to basic fundamental items, FactSet provides some most frequently used derived items and ratios, while Compustat offers virtually no ratio items.

Besides accounting data, stock price data are also available in all three databases. In WRDS, monthly security trading data are available for FactSet since 1991, while BvD Osiris offers both monthly and weekly trading data with a short horizon. Compustat provides security trading records at the daily level since 1984. Similar to its accounting time series data items, BvD often keeps security trading records up to a limited number of years. For example, Amadeus only keeps the most recent 12 month trading records, while Osiris keeps no more than

[6] There five different datasets in Osiris including Banks, Industrials, Insurance (Composites), Life Insurance, and Non-Life Insurance. The variables in each dataset are usually defined differently. For the purpose of this study, only the variables in Osiris Industrials database are retained for the final Osiris sample.

5- (8-) years weekly (monthly) trading records. Daily exchange rate information is also available in both FactSet and Compustat. For each day, FactSet provides two variables for each currency exchange rate: Currency in USD (e.g. $1.1/EUR) and Currency per USD (e.g. €0.907/USD), while Compustat provides one daily exchange rate for each currency: Currency in GB. Except the exchange rate for the currency in US dollar at the closing day of fiscal year within the annual accounting data, Osiris provides no additional exchange rate information.

4. Related Literature

To the best of our knowledge, there are few empirical studies to investigate and compare the international accounting databases. One exception is Lara, Osma, and Noguer (2006), which examine whether the choice of database has an effect on the results of empirical studies. They mainly focus on fourteen member states of the European Union and using seven international databases: Datastream, Global Vantage, Company Analysis, Worldscope, Thomson Financials, Extel Financials and BvD Osiris. The authors find that the results of the Ohlson (1995) model change significantly depending on the database chosen. They conclude that company size and differential coverage among databases explains the difference.

The other empirical research that compares multiple databases is more country specific studies. Kern and Morris (1994) present differences and similarities between the Compustat and ValueLine databases based on total assets and sales for US companies. The authors document that the mean differences of these two variables increase significantly from 1971 to 1990. Ulbricht and Weiner (2005) conducted an investigation for U.S. and partly Canadian data from Compustat and Worldscope from 1985 to 2003. They show that the use of two databases should lead to comparable results, but also find that the size bias is a crucial factor to the quality of results. Alves, Beekes and Young (2007) compare the coverage and content of the Datastream,

Worldscope, Extel, Company Analysis, and Thomson Research for UK companies. Their results suggest that these products are not perfect substitutes in terms of 1) coverage of firms and accounting items and 2) the values of accounting items. Their replication of four empirical tests indicates that the results are sensitive to the data source.

There are also some papers focusing on the quality of equity trading data across different data sources. For example, Ince and Porter (2006) compare individual US equity return data from Datastream with similar data from the CRSP to evaluate Datastream for use in studies involving large numbers of individual equities in markets outside the US. They demonstrate that after careful screening of the Datastream data, inferences drawn from Datastream are similar to those drawn from CRSP. Additionally, Schmidt, Arx, Schrimpf, Wagner and Ziegler (2011) document that appropriately screened data from Datastream and Worldscope can be used to replicate closely not only U.S. market returns and the corresponding momentum risk factor, but also the widely-used U.S. size and value risk factors. The authors then build pan-European and country-specific momentum, size, and value risk factors by using the same data screen.

5. Data Preparation for Database Comparison

Due to the fact that BvD databases contain only annual data, and Compusat Global only provides non-North American companies, we compare non-NA (i.e., excluding US and Canada) annual dataset across each database. As mentioned above, FactSet and Compustat primarily focus on public companies, while Amadeus is generally designed for European private companies, and Osiris also includes a number of private subsidiaries for public firms. To have a meaningful comparison among databases, I first construct a data set containing only public firm for each database.

To be consistent with the commonly used research conventions, fiscal year is shifted one year forward if fiscal year end falls between Jun 1 and May 31 of the next year. Since each database structure is different in many perspectives, the database-specific filters applied to screen for public firms are briefly described below. I also discuss the limitations associated with using those filters.

Public Firms:

For FactSet and Compustat, I first consider an entity to be a public company in a fiscal year if there is non-missing closing price for the entity in the last calendar month of the fiscal year. This is to control for companies that became private within that year. However, in cases where companies do not have trading records (e.g. price) at all but are given a valid ISIN or SEDOL, they are also considered as public companies to align with the treatment for BvD databases below. Additionally, I keep only the last observation in a given fiscal year to control for multinational companies and the changes in fiscal year ends. For FactSet, I further remove the observations when the two descriptive organizational type variables indicate that the entity belongs to non-corporate categories, such as Hedge Funds, Asset-Backed, and Sovereign Wealth Managers. For Compustat, I further remove the unconsolidated observations when both consolidated and unconsolidated records are available (for a given GVKEY).

For BvD databases, the definition of public companies is slightly different due to the fact of the BvD practice of data keeping. As mentioned above, the number of years available for the security trading data is limited in BvD databases. Additionally, Osiris does not provide any trading record for more than 22% of its public companies, while less than 2% of public companies miss trading records in other databases.[7] An entity in BvD databases is considered a

[7] Osiris provides security trading records for 77.64% of its public companies, while these figures for other databases are 98.76% (FactSet), 98.16% (Amadus), and 98.15% (Compustat).

public company if a) the company trading records are available in the month of the fiscal year end or 2) the company has a valid ISIN, SEDOL or ticker. Additionally, a company is also considered public in Osiris if the company is specified as listed.[8] Therefore, I keep all historical accounting information for the public companies in Osris and Amadeus unless there is a year in which its security ceased trading.[9] Finally, I removed the unconsolidated observations when there is consolidated company with same ISIN (and market capitalization) in the same fiscal year for Osiris (Amadeus).[10] There are approximately 92% public companies that are *industrials* in Osiris, and all public companies in Amadeus are identified to belong to *large company* datasets.[11]

There are several limitations associated with above filters. When FactSet or Compustat incorporates a firm's accounting records into its database before its securit(ies)y trading records for a public company, my algorithm would consider it as private for a few years. This problem is possible when a data vendor obtains accounting and security trading data from different data sources and at different time periods, especially in the case, FactSet, via WRDS, provides its security data since 1991 while provides its accounting items since the 80's. I am also unable to identify when a company became public or return to be private for the companies without trading records. This is likely to result an upward bias for the number of public companies especially for

[8] Osiris offer an additional variable called LIST, which indicates whether some securities of the company are listed in exchanges.
[9] BvD databases keep the latest security trading record for some companies, even if those companies are delisted. For example, a company is delisted in Jan 1, 2006, Amadeus keeps the 12 month trading records in 2005.
[10] A large number of observations with a consolidation code C* are found in Osiris. For this type of observations, usually all accounting items are empty. Based on the discussion with a technical representative from BvD New York, those observations are dummy header items to help online query at Osiris official web site. Those observations are removed from the final sample data. Additionally, some ISIN items in Osiris are specified as "delisted" or "unlisted" instead of a 12-digit code. The observations with "unlisted" ISIN are removed when there is no security trading information available for the company.
[11] The companies in Amadeus are categorized into three groups: large, medium and small. The variables in the data sets for each group are defined in similar way.

Osiris at early years. Finally, a public company could be considered private, if the assumption that data vendor updates its security data more frequently than its accounting data is not valid.

6. Coverage Analysis

This section will address the coverage of the reviewed international accounting databases along the following dimensions. The first part compares the historical coverage of three global databases in terms of numbers of countries and public companies over time. Second part presents the analysis for the number of companies covered by all databases in any given year (2008). Finally, an investigation for the coverage of each database in some selected countries will be provided. Since the main purpose of this paper is to examine the financial and accounting databases in a global perspective, I will primarily focus on FactSet, Compustat, and Osiris from now on.

6.1 Time Series Comparison

Panel A and B in Figure 1 represent time series coverage of countries and public companies by all databases from 1987 to 2010. In panel C, the median of medians of total assets represents the median of country total asset medians for a year in million USDs. An upward trend in the numbers of both companies and countries is observed for all three international databases. Panel C indicates that those databases increase the number of companies by adding more relatively smaller companies. Osiris dominates other databases in term of the number of countries, FactSet tends to cover fewer countries in nearly all the years, and Compustat contains the relatively larger companies.

The number of companies covered by Compustat stays below other databases' coverage, while the median of medians in total assets for companies in Compustat is always above those of the others, suggesting that Compustat database includes relatively larger companies. On the other

hand, the Osiris covers more countries and the companies of smaller size, suggesting Osiris has the largest coverage in terms of pure companies count. An upward size bias may introduce some survivorship bias into the sample, as larger firms tend to survive longer than their smaller counterparts; while the downward bias in size could reduce the numbers of observations available for time series and panel data. Researchers should consider the consequences from those potential biases in their research designs.

Table 3 provides the numbers of companies from 1991 to 2010 for a select group of countries. A country is selected if 1) there are more than 100 companies at least in one year for this country in any one of the database, and 2) companies must have more than 5 years of accounting data for the country across the databases. Fifty-three (53) countries that passed those criteria are reported alphabetically in Table 3.

It is noteworthy that Osiris tends to include more companies in developing countries in recent years. For example, since 2002 Osiris consistently contains 1,000 Indian companies more than Compustat or FactSet. Additionally, Figure 2 provides a geographic representation of the coverage for both countries and firms for each database in year 2008. The number of companies is measured by databases' primary identification, and companies must have positive total assets or sales items. Interestingly, Osiris is the only database providing accounting information for the some emerging markets, such as Iraq, and Mongolia.

6.2 Overlap in coverage across databases

The ISIN and SEDOL are the only common identifiers among different databases in question. As discussed above, there are some potential problems associated with ISIN and SEDOL. For example, there are no historical records in all databases (i.e., only the most recent identification of whether the company is public or private is available), and ISIN and SEDOL are

not unique identifiers at the company level. Those problems may result in inconsistent use of ISIN and SEDOL among different databases.

PetroChina, for instance, is one of largest Chinese public firms in Asia. The ISIN for PetroChina is CNE1000007Q1 in Compustat, while it is CNE1000003W8 in FactSet and Orisis. The Compustat's ISIN is for PetroChina's shares listed in mainland China, whereas the ISIN in the other two databases are for its Hong Kong shares. The similar inconsistency is observed even between Osiris and Amadeus. Therefore, the "overlap" analysis that uses ISIN and SEDOL for matching may understate the actual numbers of firms shared among databases. It is noteworthy that the ISINs used by each database in this paper are the ISINs that are considered by each data vendor as the primary equity ISIN for a company. Even though there are certain degree of disagreement among data vendors about the primary equity ISIN, FactSet and Compustat often keep track the ISINs for different issues for same issuers in a separated dataset.[12]

In Figure 3, instead of database's primary identifiers, I use ISIN to measure the number of companies in each database in 2008. Panel A provides an idea of the overlap among all the databases for non-North American (i.e., excluding US and Canada) companies. Panel A shows that over 70% of companies can be found in all three databases. Compustat has the least number of companies (slightly more than 3%) that are not shared by any other databases, but Osiris has most companies not covered by others. Panel B reports a similar Van Diagram for European countries among four databases. It is interesting to notice that the Amadeus, which claims to focus on European private sector, covers the least number of public companies in Europe.

Osiris has relatively smaller number of public firms not covered by other databases, which, given the largest number of unmatched firms in Osiris (see Panel A), suggests that Osiris

[12] In particular, FactSet keeps the three ISINs for PetroChina in a dataset called h_security_isin, while Compustat offers those three ISINs in its G_secnamesd dataset.

focuses more covering the companies in developing regions. A considerable proportion of public companies do not share common coverage between Osiris and Amadeus, indicating that the different data sources are used by two databases from BvD. FactSet includes the most public companies in Europe, which is consistent to the conventional belief that Worldscope (FactSet) covers most companies in developed countries.

Figure 4 indicates that the distribution of natural log of sales and total asset in 2008 for each database (the data items are not winsorized, unlike in Table 6). Though the overall histogram and kernel density distribution are similar, the left tail is fattest in Osiris, less so in FactSet and the least in Compustat for both total asset and sales, also indicating that Compustat tilts towards covering the relatively larger international companies, Osiris includes lots of relatively smaller firms with Factset being in the middle.

6.3 Coverage in Selected Countries

Table 4 serves two purposes. The first is to investigate the overall company coverage for selected countries, and the second is to understand what the country code provided by Osiris stands for. I have already shown the number of companies that are covered by each database for some countries in Table 3. The next question to answer is the how well the databases cover companies within those countries. Additionally, unlike Compustat and FactSet, the Osiris/Amadeus does not specify whether its country code is for the country of incorporation or the country of listing.[13] It is important to answer this question for several analyses provided in this paper.

[13] In this paper, the number of companies in one country is calculated according to the country code where the companies incorporated for FactSet (FF_COUNTRY_ISO) and Compustat (FIC), however, Osiris does not provide any specification about its country code (CNTRYCDE).

I first collect the number of companies listed in 42 countries (excluding US and Canada) from World Federation of Exchanges (WFE) in its 2008 annual book.[14] The first 3 columns in Table 4 present the numbers of stocks for the total, domestic and foreign companies listed in a given WFE member country. It has been well documented that many companies choose to list their common equity in the countries other than its country of incorporation for various reasons, such as regulation, market liquidity, etc. On the other hands, many companies are attracted to incorporate in some countries, so called tax havens, which offers foreign businesses little or no tax liability in a politically and economically stable environment. The number of companies incorporated in their domestic countries should be expected to exceed the number of stocks listed in their domestic exchanges at least for some developing countries, (e.g. China), and tax havens (e.g. Bermuda), when the accounting database covers a large proportion of the companies listed in their domestic exchanges.

Table 4 indicates that the FactSet and Osiris seem to have similar coverage for the world major economies (G20), while Compustat only covers more companies than reported by WFE in 3 out of 13 G20 countries in year 2008. Furthermore, the number of companies covered by Osiris exceeds the domestic (total) companies reported by WFE in year 2008 for 7(6) out 13 G20 countries, suggesting that the country code used by Osiris is unlikely to represent the listing country.

Table 5 provides the coverage statistics for the 11 major international indices with fixed number of components. Since FactSet and BvD currently do not provide their index constituent data to WRDS, the lists of ISIN for the indices are obtained from the Compustat database.

[14] The information regarding World Federation of Exchanges is available at http://www.world-exchanges.org.

Therefore, it is expected that Compustat should have a better coverage.[15] FactSet provides similar coverage to Compustat's for those indices, but Osiris contains, on average, over 10% less constituents in Australia, India, Europe, France, United Kingdom, and Germany. The difference between Osiris and Compustat appears to be driven mainly by the fact that Osiris includes much fewer constituents in early years.

7. Common Data Item Comparison

To compare some qualitative issues of the accounting data items across different databases, Table 6 provides some summary statics for a select group of frequently used items from Compustat, FactSet, and Osiris. Panel A and B present the balanced sheet, income statement, cash flow, and derived values respectively. The accounting items from a database are first converted to millions in local currency and then translated to US dollars in millions by using the exchange rate from the database at the end of fiscal years. The data items are further winsorized at 1% and 99% in each fiscal year for all databases. The first six columns in Table 6 present some summary statistics for each variable by database.

Among the databases, both FactSet and Osiris have 6 accounting items with the largest number of observations out of total 14 balance sheet items. FactSet has 5 accounting items, including sales, with the largest number of observations out of 7 income statement and cash flow items. The mean and median for total assets, sales, and many other items monotonically decrease from Compustat to FactSet to Osiris, which is consistent to the previous results in this paper. Interestingly, the standard deviation of total assets, sales and some other items also decreases in the same order. A further check indicates that there are 2,549 observations in Osiris with

[15] As mentioned before, different data vendors may disagree on the primary stocks for some large international companies, and therefore may use different constituent ISINs for those companies in each index. A minor discrepancy from Compustat does not necessarily indicate an inferior coverage for the constituent stocks, especially in the international large cap indices.

repeated positive sale or total asset items in different years for same companies, while only 21 such repeated sale or total asset items in both Compustat and FactSet. Both finding suggests there may be relatively larger amount of stale or interpolated observations in Osiris.

Another important aspect comparing accounting data items is the number of non-missing records available for a company through time which has a direct impact on the accuracy of empirical time series and panel data regression models. The 7th and 8th column in Table 6 report the average and median number of years available for each database. FactSet has 11 out 14 balance sheet items and 6 out of 7 income statement and cash flow items with longest time span for both mean and median. Compustat has longest time span among databases for the 4 remaining accounting items: Net Goodwill, Accounts Payable, Accounts Receivable, and Depreciation and Amortization.

Results in columns 9 to 12 provide some statistics for paired comparisons among three international databases. In the first step, the companies from different databases are paired by ISIN and fiscal year. Other than inconsistent ISIN and SEDOL, there are some potential problems associated with this direct matching for comparing the availability and similarity of accounting data. Even though the consolidation level has been controlled for in each database, it is still possible to have different level of consolidation across databases and, therefore, the different reporting entity and period. It has also been observed that different databases report annual accounting data in different currencies for the same companies and in the same fiscal year, suggesting the databases may collect financial statements for the same companies following different accounting standards. For example, the financial reporting currency for BP plc is GBP in FactSet and Amadeus, while is USD in Compustat and Osiris. The foreign exchange rates provided by different databases are collected from different data sources, which

may also contribute to the differences in item value. Additionally, the FactSet and Compustat have their own proprietary procedures to standardize the accounting items, while Osiris claims to report both as-reported and standardized accounting items. Finally, each database may have differential treatment of the restated financial statements, which is often not specified by the data vendors. All those reasons may result in a large difference in accounting item values for the pairs with same ISIN and fiscal year. Therefore, a further fuzzy matching filter is applied for the paired observations to compare variable values across databases. The filter is that the absolute difference between the total assets scaled by the average of the total assets must be less than 5%, i.e.

$$\frac{Abs\left(Total\ Asset_{i,t}^{database_1} - Total\ Asset_{i,t}^{database_2}\right)}{Abs\left(Avg\left(Total\ Asset_{i,t}^{database_1}, Total\ Asset_{i,t}^{database_2}\right)\right)} \leq 5\%,$$

when both total asset items are available; otherwise the absolute difference between the sales scaled by the average of the sales must be less than 5% i.e.

$$\frac{Abs\left(Sales_{i,t}^{database_1} - Sales_{i,t}^{database_2}\right)}{Abs\left(Avg\left(Sales_{i,t}^{database_1}, Sales_{i,t}^{database_2}\right)\right)} \leq 5\%,$$

when the both sale items are available.

Columns 9&10 in Table 6 report the number of paired observations that have same ISIN and fiscal year in different databases and pass the 5% fuzzy matching filter and also have the absolute difference between the values of paired accounting items scaled by their average value less than 1%, i.e.

$$\frac{Abs\left(item_{i,t}^{database_1} - item_{i,t}^{database_2}\right)}{Abs\left(Avg\left(item_{i,t}^{database_1}, item_{i,t}^{database_2}\right)\right)} \leq 1\%.$$[16]

Table 6 (columns 9 to 10) reports the paired correlation for the accounting items passing 5% fuzzy matching filters mentioned above. Consistent with the previous findings in this paper, the number of matching and the correlation between Compustat and Osiris are usually inferior to those between FactSet and others, especially for the accounting items more likely to be subject to standardization, such as EBIT, and Interest Expense.

In summary, the number of items matched between FactSet and Osiris tends to be less than those matched between FactSet and Compustat almost all the time, which indicates a probable difference in methodologies used by Osiris to record accounting items. Furthermore, the number of paired observations and paired correlations for interest expenses between Osiris and other databases are not comparable to the other paired results, indicating a large disparity in the measures of some items between Osiris and other databases. Additional check for the data across different databases shows that 2,008 (1,967) fiscal year end date in company-year matched observations are different between Osiris and FactSet (Compustat), while only 121 observations are different between FactSet and Compustat.

8. Conclusions and Implications

This paper has demonstrated various differences among the international accounting databases on WRDS (FactSet Fundamentald, Compustat Global, BvD Osiris and Amadeus) and the potential effect of those differences on results and findings of international empirical research. The results reveal that the available databases may not be perfect substitutes for one

[16] Osiris reports negative value for Cost of Goods Sold and Depreciation and Amortization data items. The numbers of records for those two items are calculated by using absolution value of valid records in Osiris.

another in some cases. The differences are mainly attributable to the variation in firm and country coverage and the availability of accounting data items.

Compustat Global traditionally tends to focus on the larger companies in more developed countries and provides the most comprehensive accounting data items. BvD Osiris covers more companies in emerging countries and offers a limited variety of accounting data items in a mixture of standardized and as-reported formats. Amadeus actually provides the least number of public European companies with the limited number of accounting items available. FactSet has a balance between the company size and coverage, with a reasonable selection of accounting items.

Among international databases, Osiris and FactSet often have the largest number of observations with accounting items, while Compustat and FactSet tend to have longer time span of those items at company level. In addition, Compustat and FactSet have most similarity in accounting items matched by ISIN in terms of values, comovement, and the fiscal year ends.

References

Alves, P., W. Beekes, and S. Young. "A Comparison of Uk Firms' Financial Statement Data from Six Sources." Working paper, University of Lancaster, 2007.

Fama, Eugene F., and Kenneth R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33, no. 1 (1993): 3-56.

Ince, Ozgur S., and R. Burt Porter. "Individual Equity Return Data from Thomson Datastream: Handle with Care!". *Journal of Financial Research* 29, no. 4 (2006): 463-79.

Kern, Beth B., and Michael H. Morris. "Differences in the Compustat and Expanded Value Line Databases and the Potential Impact on Empirical Research." *The Accounting Review* 69, no. 1 (1994): 274-84.

Lara, Juan Manuel García, Beatriz García Osma, and Belén Gill de Albornoz Noguer. "Effects of Database Choice on International Accounting Research." *Abacus* 42, no. 3-4 (2006): 426-54.

Ohlson, James A. "Earnings, Book Values, and Dividends in Equity Valuation*." *Contemporary Accounting Research* 11, no. 2 (1995): 661-87.

Schmidt, Peter S., Urs Von Arx, Andreas Schrimpf, Alexander F. Wagner, and Andreas Ziegler. "On the Construction of Common Size, Value and Momentum Factors in International Stock Markets: A Guide with Applications." *SSRN eLibrary* (2011).

Ulbricht, Niels, and Christian Weiner. "Worldscope Meets Compustat: A Comparison of Financial Databases." *SSRN eLibrary* (2005).

Table 1: Database Cited for Papers on International Finance and Accounting Journals

The numbers of papers are collected by using Google Scholar with keyword: database name and "International" or "Global" or "Emerging" or "Developing" in each key journal. The database name keyword is "Compustat Global" or "Global Vantage" for Compustat Global, is "FactSet" or "Worldscope" for FactSet, and is "Bureau van Dijk" or "AMADEUS" or 'OSIRIS' for Bureau van Dijk. The following abbreviations for journal names have been used: JF, Journal of Finance; JFE, Journal of Financial Economics; RFS, Review of Financial Study; JFQA, Journal of Financial and Quantitative Analysis; TAR, The Accounting Review; CAR, Contemporary Accounting Research; JAE, Journal of Accounting and Economics; JAR, Journal of Accounting Research; and RAS, Review of Accounting Studies. The number does not necessarily mean that a paper conducts research by using the database. For example, Charia, Henry (2008) only mention some characteristics of Worldscope and Global Vantage. Additionaly the number is not mutual exclusive, i.e. one paper could also conduct research based on several databases, such as Pope and Walker (1999).

Panel A: Finance Journals

	JF	JFE	RFS	JFQA	Total
Compustat Global (Global Vantage)	10	11	5	5	31
Bureau Van Dijk (Osiris or Amadeus)	9	14	6	2	31
FactSet (WorldScope)	44	63	25	21	153

Panel B: Accounting Journals

	TAR	CAR	JAE	JAR	RAS	Total
Compustat Global (Global Vantage)	10	9	10	19	7	55
Bureau Van Dijk (Osiris or Amadeus)	2	1	2	1	0	6
FactSet (WorldScope)	10	10	8	10	7	45

Table 2: The Availability of Most Used Accounting Variables in International Databases

This table demonstrates the direct availability of a selection of accounting variables in Compustat Global, FactSet, and BvD Osiris and Amadeus. The direct availability means the accounting variables are directly available in a given database, so it does not necessarily indicate the variable may not be derived from other available variables. For example, the ROA can be derived from net income and total asset.

Balance Sheet Items	FactSet	Compustat	Osiris	Amadeus
Accounts Payable	×	×		
Accounts Receivable, Gross	×	×	×	
Cash And Short-Term Investments	×	×	×	
Common Equity	×	×	×	
Current Assets	×	×	×	×
Current Liabilities	×	×	×	×
Goodwill, Net	×	×	×	
Intangible Assets, Net	×	×	×	×
Inventories	×	×	×	
Long Term Debt, Total	×	×		×
Property, Plant And Equipment, Gross	×	×	×	
Retained Earnings	×	×	×	
Total Asset	×	×	×	×
Total Liabilities	×	×	×	
Income Statement and Cash Flow Items				
Cost Of Goods Sold	×	×	×	×
Capital Expenditures	×	×		
Depreciation (And Amortization)	×	×	×	×
Interest Expense, Total	×	×	×	×
Net Cash Flow From Operation	×	×	×	
Net Income[18]	×	×	×	×
Sales	×	×	×	×
Derived Values and Ratios				
EBIT	×	×	×	×
EBITDA	×	×	×	×
Gross Profit Margin	×		×	×
Net Profit Margin	×		×	
Operating Margin	×			×
Quick Ratio	×			
Current Ratio	×		×	
ROA	×		×	×
ROE	×		×	
Inventory Turnover	×			
Receivables Turnover	×			
Pretax Interest Coverage Ratio	×		×	×
Fixed Charge Coverage Ratio	×			

[18] The Net Income variable (NI) in Compustat Global contains only missing values when this paper is written, However, Compustat provides other related net income variables that are not missing, such as Consolidated Net Income (NICON), and Unconsolidated Net Income (NINC).

Table 3: A Longitude Demonstration for Selected Countries

This table demonstrates the number of companies available for selected countries over period 1991-2010. A country is selected if this country has at least 100 public companies available in one year and has at least 5 years observations in FactSet, Compustat, or Osiris. The country is the country where the company is incorporated for FactSet and Compustat. All companies in the table must have either no missing total asset or no missing net sale in a given year. The numbers represent public firms only and no replications due to different level of consolidations.

		2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993	1992	1991
Australia	**Compustat**	1,721	1,771	1,805	1,708	1,567	1,453	1,273	1,169	847	624	512	478	472	438	307	250	228	152	148	141
	FactSet	1,799	1,808	1,851	1,829	1,700	1,555	1,469	1,363	1,279	1,258	825	541	396	344	321	278	227	206	203	61
	Osiris	1,728	1,798	1,853	1,874	1,767	1,652	1,512	1,372	1,296	1,218	835	497	437	387	361	329	292	256	232	214
Austria	**Compustat**	82	91	95	99	94	85	84	85	88	102	105	102	103	100	94	60	48	43	36	36
	FactSet	84	92	98	103	105	106	110	107	116	125	129	130	127	126	115	85	90	90	85	75
	Osiris	85	98	105	106	107	109	106	98	103	107	110	101	98	86	73	66	55	47	40	32
Bangladesh	**Compustat**	44	50	42	28	23	12	8	7	3	0	0	0	0	0	0	0	0	0	0	0
	FactSet	13	14	11	9	6	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Osiris	92	124	107	92	86	79	70	59	50	81	76	59	49	48	47	45	21	8	9	9
Belgium	**Compustat**	126	138	141	141	127	128	124	128	130	139	129	119	109	104	104	70	62	42	37	35
	FactSet	129	143	147	158	157	162	159	148	146	153	161	166	164	158	139	118	119	120	115	111
	Osiris	133	161	162	168	166	173	173	168	157	152	165	168	144	128	103	89	83	76	63	55
Bermuda	**Compustat**	502	513	515	511	495	483	466	453	432	355	303	278	264	241	158	92	83	61	41	36
	FactSet	84	93	94	102	97	83	77	68	58	48	31	16	10	11	11	10	8	7	4	4
	Osiris	473	674	685	695	688	682	661	606	579	567	524	449	373	340	300	242	207	142	100	75
Brazil	**Compustat**	320	325	331	333	275	255	216	207	195	184	153	147	144	127	69	62	51	41	41	37
	FactSet	368	386	395	399	400	398	411	374	355	362	385	381	184	154	146	137	107	88	81	76
	Osiris	407	443	473	467	470	443	455	438	461	397	392	236	200	131	110	94	0	0	0	0
Cayman Islands	**Compustat**	401	386	377	359	311	293	260	226	173	74	59	50	46	41	18	11	9	5	1	1
	FactSet	32	34	35	31	20	17	14	11	8	5	4	1	1	1	1	1	1	0	1	1
	Osiris	602	716	701	653	581	529	490	402	327	254	181	114	76	57	48	36	29	20	17	12
Chile	**Compustat**	150	154	150	149	143	144	143	138	136	136	134	126	126	112	25	25	20	13	12	9
	FactSet	206	218	219	214	219	221	218	216	215	188	190	192	128	92	87	76	70	61	57	45
	Osiris	196	228	233	232	236	241	231	226	225	168	221	221	206	209	195	175	157	145	127	106
China	**Compustat**	2,310	1,979	1,913	1,725	1,496	1,448	973	472	400	303	167	151	143	132	81	65	56	34	10	1
	FactSet	2,903	2,942	2,893	2,621	2,085	1,902	1,888	1,750	1,545	1,310	1,280	1,215	154	142	121	98	48	39	24	14
	Osiris	2,269	2,248	2,268	2,235	1,911	1,647	1,571	1,413	1,302	1,144	929	699	287	41	35	19	9	4	2	1

		2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993	1992	1991
Colombia	**Compustat**	36	34	32	26	25	23	23	23	23	22	23	21	20	19	7	6	6	4	2	2
	FactSet	39	44	43	43	39	36	41	40	40	34	35	31	33	30	31	30	29	27	26	21
	Osiris	75	85	86	91	109	98	113	116	103	89	93	51	54	63	30	24	9	8	3	0
Cyprus	**Compustat**	35	36	36	35	27	24	20	13	15	15	10	6	4	4	2	0	.	.	.	.
	FactSet	63	63	63	48	4	3	0	0	0	0	0	0	0	0	0	0	.	.	.	.
	Osiris	46	134	141	142	126	31	19	11	6	5	5	6	6	6	3	3	.	.	.	.
Denmark	**Compustat**	113	151	177	179	176	169	168	175	185	159	142	143	137	127	90	66	60	57	30	31
	FactSet	184	202	210	208	202	192	188	185	194	202	213	223	227	218	202	158	158	152	149	141
	Osiris	154	186	197	210	213	207	199	180	189	193	197	201	193	190	184	176	139	124	99	81
Egypt	**Compustat**	27	31	36	36	35	26	26	23	19	18	19	14	11	7	0	0	0	0	0	0
	FactSet	108	133	129	109	83	77	52	46	32	23	20	13	14	6	0	0	0	0	0	0
	Osiris	201	236	314	408	514	623	708	707	418	120	116	85	64	35	34	30	20	8	7	6
Finland	**Compustat**	124	127	129	134	130	129	124	127	129	132	117	98	86	77	72	51	49	49	28	28
	FactSet	124	131	134	141	140	148	151	154	155	152	158	159	158	147	121	102	97	100	93	93
	Osiris	105	122	124	128	131	135	138	134	139	144	143	135	102	71	59	47	45	45	36	34
France	**Compustat**	620	678	686	720	685	647	651	668	693	718	712	552	542	548	531	361	298	195	177	174
	FactSet	692	746	793	820	858	866	887	904	913	922	986	1,027	1,036	982	890	625	636	654	655	634
	Osiris	656	867	932	926	908	847	860	870	852	837	867	860	705	586	523	477	436	374	359	328
Germany	**Compustat**	646	742	784	807	735	657	639	670	713	799	749	592	491	432	399	296	273	246	203	196
	FactSet	778	880	934	976	984	980	962	947	944	997	1,034	1,030	1,038	976	824	633	621	622	559	502
	Osiris	683	831	878	922	898	828	787	745	753	780	819	822	710	534	433	357	323	303	274	247
Greece	**Compustat**	190	206	203	199	191	187	170	145	146	137	99	93	82	62	34	14	9	2	2	2
	FactSet	279	291	305	315	317	329	328	317	319	336	326	265	202	193	191	127	127	123	104	68
	Osiris	248	266	277	291	298	261	243	132	134	187	175	152	125	123	111	90	77	54	45	26
Hong Kong	**Compustat**	426	383	364	349	298	271	251	245	240	220	212	200	192	188	123	93	88	71	55	54
	FactSet	1,048	1,079	1,075	1,078	1,050	1,049	1,046	1,018	977	917	810	500	425	416	399	345	212	169	147	118
	Osiris	182	225	220	228	226	227	222	212	206	204	201	191	185	180	166	149	140	108	95	85
India	**Compustat**	1,563	1,752	1,671	1,390	993	859	756	684	590	547	483	396	355	317	135	102	82	61	51	38
	FactSet	2,289	2,551	2,555	2,378	2,269	2,080	874	762	581	470	457	407	348	324	330	328	297	205	210	173
	Osiris	720	3,438	3,574	3,567	3,510	3,152	2,672	1,839	1,649	1,244	791	643	295	242	164	137	74	56	55	51
Indonesia	**Compustat**	289	296	283	262	271	262	267	272	261	250	229	234	232	187	117	83	55	41	35	33
	FactSet	419	425	428	418	385	368	358	340	337	328	313	221	171	169	160	146	108	93	88	79
	Osiris	362	335	360	352	332	325	310	311	300	131	61	71	80	87	89	87	68	46	28	17

		2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993	1992	1991
Israel	**Compustat**	212	221	191	204	168	141	110	102	88	89	64	53	53	42	31	21	1	0	0	0
	FactSet	273	281	259	244	215	220	205	195	155	136	120	86	69	61	41	37	39	37	25	0
	Osiris	405	596	629	654	652	373	338	178	173	168	155	136	104	75	53	44	32	24	22	22
Italy	**Compustat**	265	278	286	298	274	262	254	254	265	273	237	213	198	184	178	120	121	101	96	92
	FactSet	278	288	301	312	320	327	321	307	307	303	305	301	282	263	239	214	217	216	223	217
	Osiris	264	292	308	324	339	348	321	259	253	262	252	214	179	172	148	127	100	100	85	90
Japan	**Compustat**	3,659	3,775	3,895	3,973	3,956	3,821	3,724	3,624	3,602	3,538	3,399	3,295	3,203	3,055	2,910	2,772	2,505	2,007	1,705	1,667
	FactSet	3,679	3,792	3,913	4,028	4,107	4,124	4,133	4,028	3,919	3,809	3,742	3,458	3,413	3,227	2,514	2,505	2,464	2,431	2,375	2,251
	Osiris	3,025	3,970	4,169	4,400	4,518	4,542	4,551	4,514	4,467	4,317	2,622	1,787	1,413	1,001	947	860	610	333	233	171
Jordan	**Compustat**	10	169	163	157	125	75	55	52	47	35	12	8	7	5	2	0	0	0	.	.
	FactSet	94	93	68	56	46	41	36	34	24	13	9	6	6	4	0	0	0	0	.	.
	Osiris	195	241	247	219	195	167	153	146	131	124	107	20	17	16	13	13	7	1	.	.
Korea, South	**Compustat**	1,570	1,664	1,719	1,678	1,333	1,204	1,061	1,032	911	895	579	537	507	500	462	421	131	116	106	88
	FactSet	1,111	1,206	1,256	1,157	1,128	1,120	1,069	995	954	837	786	738	473	367	316	270	253	221	146	114
	Osiris	1,728	1,769	1,773	1,692	1,690	1,665	1,657	1,643	1,610	1,572	1,205	1,178	1,110	1,079	1,014	961	845	787	755	745
Kuwait	**Compustat**	93	153	148	133	112	68	58	46	31	25	0	0	0	0	0	0	0	0	.	.
	FactSet	196	200	192	182	167	148	112	61	0	0	0	0	0	0	0	0	0	0	.	.
	Osiris	174	207	212	210	186	165	150	112	64	39	25	20	12	12	11	10	4	1	.	.
Malaysia	**Compustat**	916	933	954	971	983	970	889	811	761	672	627	622	586	553	378	246	220	157	119	105
	FactSet	968	999	1,031	1,063	1,075	1,086	1,063	1,026	950	848	767	530	450	436	417	362	268	226	224	156
	Osiris	740	955	961	998	1,019	1,034	1,013	948	882	800	771	730	631	567	605	537	392	156	178	172
Mexico	**Compustat**	109	109	108	104	101	96	94	90	86	82	79	81	90	84	60	51	38	36	39	25
	FactSet	130	131	130	129	135	135	142	143	146	152	150	153	129	90	88	76	75	61	56	39
	Osiris	117	125	127	129	126	131	138	137	142	146	144	136	101	82	76	100	64	50	45	17
Mongolia	**Compustat**	.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.	.	.	.
	FactSet	.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.	.	.	.
	Osiris	.	2	61	38	105	119	136	134	138	106	68	42	34	20	14	4	.	.	.	.
Netherlands	**Compustat**	141	150	160	175	168	170	174	180	185	199	208	213	199	188	176	125	102	84	80	75
	FactSet	151	168	185	197	196	203	212	212	212	216	231	252	259	259	236	187	189	186	187	182
	Osiris	138	155	167	181	182	185	199	204	210	218	236	262	248	221	191	158	134	120	113	110
New Zealand	**Compustat**	114	118	118	124	112	110	104	102	85	90	84	77	75	58	37	24	19	22	20	18
	FactSet	129	132	139	149	144	138	138	126	111	110	100	79	68	65	57	52	46	32	28	16
	Osiris	111	131	138	142	147	148	143	136	125	117	92	78	65	53	42	40	35	31	26	24

		2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993	1992	1991
Nigeria	**Compustat**	38	58	65	66	47	35	33	30	23	14	6	0	2	2	1	2	2	1	0	0
	FactSet	42	93	106	88	58	39	25	17	13	3	0	0	0	0	0	0	0	0	0	0
	Osiris	20	78	120	134	133	131	133	136	129	122	96	91	97	87	73	61	49	17	6	1
Norway	**Compustat**	198	216	237	240	206	180	161	149	164	158	132	121	127	102	85	55	52	45	43	43
	FactSet	205	227	247	263	254	243	221	198	202	200	198	210	216	206	189	112	112	109	102	105
	Osiris	152	162	181	207	219	222	212	182	180	173	165	157	156	143	117	99	87	73	48	38
Oman	**Compustat**	62	72	64	60	58	54	42	40	35	29	9	0	0	0	0	0	.	.	.	.
	FactSet	115	118	118	117	86	37	22	2	2	2	0	0	0	0	0	0	.	.	.	.
	Osiris	101	117	123	124	125	123	118	111	103	79	60	7	6	6	6	6	.	.	.	.
Pakistan	**Compustat**	222	236	224	210	191	169	128	105	88	92	102	96	89	76	33	13	5	0	0	0
	FactSet	173	198	198	199	175	144	129	124	122	121	117	103	102	100	81	75	78	71	52	16
	Osiris	343	502	506	408	341	285	268	264	236	211	246	248	206	82	46	26	15	8	5	1
Peru	**Compustat**	79	79	75	69	58	50	48	47	43	38	34	33	32	28	7	4	3	1	1	1
	FactSet	101	105	107	104	97	94	79	82	84	88	89	70	47	37	36	19	24	26	17	5
	Osiris	127	151	164	161	157	165	164	151	152	146	141	134	128	112	20	17	3	3	3	0
Philippines	**Compustat**	198	202	201	196	189	187	183	179	176	160	162	159	163	160	51	32	15	11	10	8
	FactSet	231	242	245	242	243	244	235	233	234	227	220	168	130	123	122	105	72	69	58	41
	Osiris	215	231	233	235	240	241	232	232	232	224	186	84	63	50	61	57	45	33	26	16
Poland	**Compustat**	353	358	333	278	206	171	130	114	99	90	86	76	69	49	16	2	0	0	0	0
	FactSet	335	358	360	363	348	322	257	202	126	94	92	81	70	53	54	49	16	14	9	1
	Osiris	100	137	134	107	91	95	87	81	68	67	68	60	51	33	22	15	11	10	8	4
Portugal	**Compustat**	43	48	50	49	47	49	50	52	53	55	51	60	62	61	57	23	20	8	8	8
	FactSet	53	57	60	59	61	63	64	69	71	78	87	90	95	105	110	75	72	62	61	55
	Osiris	51	59	61	63	63	61	64	68	70	66	64	66	52	47	39	37	33	23	16	8
Russia	**Compustat**	149	166	158	143	112	72	63	52	41	37	37	31	24	14	8	0	0	.	.	.
	FactSet	198	219	217	224	224	212	131	83	52	39	33	26	24	30	19	1	0	.	.	.
	Osiris	259	342	362	391	409	360	262	176	125	92	77	46	30	18	13	15	8	.	.	.
Saudi Arabia	**Compustat**	128	125	120	96	50	33	22	21	18	14	0	0	0	0	0	0	0	0	0	.
	FactSet	145	138	127	102	95	89	69	58	5	0	0	0	0	0	0	0	0	0	0	.
	Osiris	115	120	120	111	108	96	90	85	75	56	53	28	9	9	9	9	1	1	1	.
Singapore	**Compustat**	637	638	635	635	592	552	502	432	407	376	292	283	255	244	181	162	145	102	65	61
	FactSet	643	666	676	690	665	635	628	623	579	486	433	290	247	235	220	213	162	126	110	85
	Osiris	533	620	638	652	657	635	611	527	478	443	403	348	329	288	238	171	167	149	139	140

		2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993	1992	1991
South Africa	**Compustat**	296	311	317	309	289	280	282	282	285	276	257	218	213	216	160	107	86	73	72	72
	FactSet	335	350	364	381	390	376	377	372	383	446	515	563	552	309	189	176	177	172	156	88
	Osiris	269	344	361	364	358	337	332	323	332	351	339	326	298	225	165	100	78	65	65	56
Spain	**Compustat**	139	145	147	147	143	133	138	139	142	152	150	149	138	141	135	113	106	83	67	67
	FactSet	143	149	153	156	157	162	173	172	179	192	195	202	210	209	203	158	158	161	160	153
	Osiris	131	157	167	172	202	206	192	148	159	164	164	128	126	110	103	92	83	88	89	90
Sri Lanka	**Compustat**	183	183	177	166	138	83	46	33	27	22	12	9	7	7	2	0	0	0	0	.
	FactSet	190	193	191	186	106	47	35	33	30	34	51	47	21	20	18	18	18	17	6	.
	Osiris	69	215	219	217	220	173	95	54	61	142	131	25	18	16	11	9	1	1	1	.
Sweden	**Compustat**	381	389	376	380	324	293	275	279	288	292	270	215	201	149	117	81	79	72	67	69
	FactSet	473	488	500	501	466	444	385	350	345	350	359	361	323	301	250	177	171	156	152	143
	Osiris	148	422	448	476	444	429	358	292	276	277	268	293	300	252	226	177	137	104	82	64
Switzerland	**Compustat**	245	258	264	271	271	252	248	253	264	268	225	215	209	201	187	143	127	114	66	65
	FactSet	253	264	272	282	282	282	285	277	282	288	289	260	251	243	226	191	192	187	176	176
	Osiris	239	267	274	281	274	267	259	249	261	262	257	245	231	201	189	171	126	112	104	95
Taiwan	**Compustat**	1,504	1,491	1,454	1,264	1,234	968	878	827	604	282	226	202	192	184	61	32	24	14	10	7
	FactSet	1,654	1,673	1,686	1,665	1,608	1,604	1,579	1,516	1,375	1,255	564	434	251	242	227	210	125	55	26	22
	Osiris	1,309	1,643	1,595	1,646	1,628	1,486	1,431	1,092	1,034	888	818	739	626	373	96	72	42	17	15	12
Thailand	**Compustat**	502	507	501	500	485	451	420	375	351	340	342	346	347	336	212	181	163	127	122	103
	FactSet	540	547	548	554	558	546	539	517	453	412	380	280	274	280	272	248	235	225	162	98
	Osiris	491	515	537	546	537	478	478	476	432	398	344	196	189	200	214	191	132	108	106	92
Turkey	**Compustat**	156	151	157	154	140	129	108	102	86	74	65	60	62	50	37	15	9	4	3	3
	FactSet	284	285	278	261	236	240	239	225	209	177	160	144	112	95	70	55	54	54	42	30
	Osiris	273	318	324	299	278	180	148	138	98	63	46	34	28	78	77	73	44	39	29	6
United Arab Emirates	**Compustat**	90	91	89	84	62	36	19	16	10	6	1	1	1	1	0	0	0	0	.	.
	FactSet	95	99	95	90	77	66	55	29	0	0	0	0	0	0	0	0	0	0	.	.
	Osiris	92	102	104	100	99	95	83	52	38	31	28	24	18	17	17	16	2	2	.	.
United Kingdom	**Compustat**	1,582	1,808	1,976	2,140	2,121	2,008	1,828	1,743	1,763	1,766	1,724	1,612	1,642	1,640	1,483	1,065	942	876	850	799
	FactSet	1,519	1,733	1,921	2,125	2,199	2,217	2,188	2,139	2,043	1,995	1,867	1,764	1,784	1,876	1,868	1,557	1,485	1,442	1,412	1,292
	Osiris	1,492	2,067	2,293	2,478	2,506	2,437	2,264	2,115	2,124	2,141	1,989	1,884	1,891	1,774	1,513	1,406	1,188	936	755	644
Vietnam	**Compustat**	49	110	49	27	3	0	0	0	0	0	0	0	0	0	0	0	0	.	.	.
	FactSet	174	160	118	45	8	3	2	0	0	0	0	0	0	0	0	0	0	.	.	.
	Osiris	428	441	433	361	306	155	81	46	35	16	8	8	8	7	7	6	3	.	.	.

Table 4: Country Coverage Benchmarked by World Federation of Exchanges Yearbook 2008

This table demonstrates the number of companies available for the non-north American member countries of World Federation of Exchanges as well as for those countries in FactSet, Compustat, BvD Osiris, and Amadeus. The country is where the company is incorporated for FactSet and Compustat. All companies in the table must have either no missing total asset or no missing net sale in a given year.

Country	WFE			FactSet	Compustat	Osiris	Amadeus
	Total	Domestic	Foreign				
Latin American							
Bermuda	52	16	36	94	515	685	
Mexico	373	125	248	130	108	127	
Argentina	112	107	5	92	72	97	
Brazil	392	383	9	395	331	473	
Chile	238	235	3	219	150	233	
Colombia	89	89	0	43	32	86	
Peru	244	201	43	107	75	164	
Europe							
Austria	118	101	17	98	95	105	80
Cyprus	119	119	0	63		141	4
Germany	832	742	90	934	784	878	796
Greece	285	282	3	305	203	277	236
Hungary	43	40	3	40	20	20	28
Ireland	68	58	10	72	65	84	59
Italy	300	294	6	301	286	308	189
Luxembourg	262	34	228	54	42	60	20
Malta	19	19	0	17	11	16	7
Norway	259	209	50	247	237	181	166
Poland	458	432	26	360	333	134	286
Slovenia	84	84	0	17	19	25	41
Spain	3,576	3,536	40	153	147	167	365
Switzerland	323	253	70	272	264	274	210
United Kingdom	3,096	2,415	681	1,921	1,976	2,293	1,642
Africa							
Egypt	373	372	1	129	36	314	
South Africa	411	367	44	364	317	361	
Asia Pacific							
China	1,604	1,604	0	2,893	1,913	2,268	
Hong Kong	1,261	1,251	10	1,075	364	220	
India	6,327	6,327	0	2,555	1,671	3,574	
Indonesia	396	396	0	428	283	360	
Iran	356	356	0			4	
Israel	642	630	12	259	191	629	
Japan	3,786	3,769	17	3,913	3,895	4,169	
Jordan	262	262	0	68	163	247	
Korea, South	1,793	1,789	4	1,256	1,719	1,773	
Malaysia	976	972	4	1,031	954	961	
Philippines	246	244	2	245	201	233	
Singapore	767	455	312	676	635	638	
Sri Lanka	235	235	0	191	177	219	
Taiwan	722	718	4	1,686	1,454	1,595	
Thailand	525	525	0	548	501	537	
Turkey	317	317	0	278	157	324	
Australia	2,009	1,924	85	1,851	1,805	1,853	
New Zealand	172	147	25	139	118	138	

Table 5: Major Index Coverage

This panel A, B, and C represent the number of companies of the major large cap indices for worldwide, Asian, and European countries respectively. The indices are from Compustat index data and the index constituents are matched by ISIN and the years of match are implied by the closing date of the fiscal year for the constituent companies. The empty records indicate that either the years are prior to the inception days of indices or the data is not available in Compustat index data. The S&P 700 is designed to be a highly liquid and tradable index with the total market capitalization of which is large enough to approximate the appropriate market segment with a fixed number of 700 components. The STOXX Europe 600 represents large, mid and small capitalization companies across 18 countries of the European region with a fixed number of 600 components. The S&P Asia 50 is an equity index from four major Asian markets: Hong Kong, Singapore, South Korea, and Taiwan. The S&P Latin America 40 is an equity index from five major Latin American markets: Brazil, Chile, Colombia, Mexico, and Perú. The S&P Japan 500 is designed to represent the Japanese investable market and the constituents are eligible companies listed on the Tokyo, Osaka, or JASDAQ exchanges. ASX 100 represents the large and mid-cap universe for Australia. The Hang Seng contents 45 biggest companies of Hong Kong stock market is recorded and monitored by this index. BSE100 was has 100 Indian companies with varying weightages. The FTSE 100 is a share index of the 100 most highly capitalized UK companies listed on the London Stock Exchange. The SBF 120 is based on the 120 most actively traded stocks listed in Paris. And DAX30 is a blue chip stock market index consisting of the 30 major German companies trading on the Frankfurt Stock Exchange.

Panel A: A Selection of World and Continental Indices

	S&P 700			Dow Jones STOXX Europe 600				S&P Asia 50			S&P Latin America 40		
Year	**FactSet**	**Compustat**	**Osiris**	**FactSet**	**Compustat**	**Osiris**	**Amadeus**	**FactSet**	**Compustat**	**Osiris**	**FactSet**	**Compustat**	**Osiris**
2000	3	3	3	476	531	351	237				32	37	25
2001	5	5	4	528	578	396	305				34	38	28
2002	5	5	5	549	591	411	312				36	38	29
2003	385	402	313	567	592	439	335	45	46	44	37	39	30
2004	604	625	512	578	594	466	354	49	50	48	37	39	31
2005	614	635	541	591	607	502	392	52	53	51	36	38	30
2006	618	633	563	587	598	525	404	49	50	48	36	38	31
2007	622	635	579	587	595	544	423	50	51	48	36	38	31
2008	623	636	598	592	601	574	445	49	50	48	36	38	31
2009	625	638	617	582	593	576	406	49	50	48	37	39	37
2010	624	633	614	588	596	583	43	48	50	47	37	39	36
Avg. % 01-10	**67.5%**	**69.2%**	**62.1%**	**82.1%**	**84.9%**	**71.7%**	**60.2%**[19]	**97.8%**	**100.0%**	**95.5%**	**90.5%**	**96.0%**	**78.5%**

[19] Since the annual report for Amadeus is not available in year 2010 when the paper is written, the average percentage is calculated by using average percentage from 2000 to 2009 for Dow Jones STOXX Europe 600.

Panel B: A Selection of Asian Indices

	S&P Japan 500			S&P ASX 100			Hang Seng			BSE 100		
Year	**FactSet**	**Compustat**	**Osiris**	**FactSet**	**Compustat**	**Osiris**	**FactSet**	**Compustat**	**Osiris**	**FactSet**	**Compustat**	**Osiris**
2000				74	89	49	32	32	32	82	87	59
2001				78	88	53	31	31	31	78	84	56
2002				84	91	55	33	33	33	78	81	57
2003	30	30	25	83	89	59	33	33	33	94	95	71
2004	483	500	438	85	90	65	32	33	32	97	98	82
2005	485	499	442	85	96	69	33	34	33	99	99	84
2006	485	496	451	87	95	71	38	38	36	100	100	88
2007	491	501	461	90	98	78	41	43	39	97	97	91
2008	491	499	473	89	97	85	40	42	39	92	92	88
2009	495	499	475	88	94	88	40	42	39	99	99	97
2010	496	499	480	89	94	90	43	45	42	99	100	97
Avg. % 01-10	**86.4%**	**88.1%**	**81.1%**	**85.8%**	**93.2%**	**71.3%**	**80.9%**	**83.1%**	**79.3%**	**93.3%**	**94.5%**	**81.1%**

Panel C: A Selection of European Indices

	FTSE 100				SBF 120				DAX 30			
Year	**FactSet**	**Compustat**	**Osiris**	**Amadeus**	**FactSet**	**Compustat**	**Osiris**	**Amadeus**	**FactSet**	**Compustat**	**Osiris**	**Amadeus**
2000					94	113	76	65	28	30	25	18
2001					108	119	84	78	29	30	24	19
2002					110	117	88	81	31	31	25	22
2003					113	116	92	87	30	30	26	22
2004					114	116	97	90	30	30	26	22
2005					117	118	104	94	30	30	26	22
2006					119	120	109	98	30	30	27	23
2007					118	120	114	103	30	30	28	23
2008	68	70	65	45	119	120	119	108	29	29	30	24
2009	95	98	94	68	120	121	120	106	28	28	29	23
2010	100	102	100	21	121	122	122	5	29	29	29	0
Avg. % 00-09	**81.5%**	**84.0%**	**79.5%**	**56.5%**	**94.3%**	**98.3%**	**83.6%**	**75.8%**	**98.3%**	**99.3%**	**88.7%**	**72.7%**

Table 6: Financial Statement Items Comparison

This table demonstrates the availability, quantity, and overleaping properties of some frequently used accounting data items. Panel A and B present the balanced sheet, and income statement, cash flow and derived value respectively. The first six columns of the table provide summary statistics for the selected variables in each database. The empty values represent that the given variable is not directly available in the database. The next two columns indicate the average and median years available for each variable. The following two columns demonstrate the number of observations matched and the last two columns indicate the correlation between the matched observations. The records from different databases in this table are firstly matched by ISIN and year. The matched data are further restricted by the condition that the difference in values of total assets from matched pairs scaled by the average of two total assets is no greater than 5%. When the total assets are not available in either matched database, the criterion is applied by using net sales. All values have been translated into US dollar in millions by using the exchange rates directly from the original database at the closing day of the fiscal year end. The variables are winsorized all at the upper and lower one percentiles for each database. All firms are public and no replications due to different level of consolidations.

Panel A: Balance Sheet Items

		Summary Statistics								Matched		Correlation	
	Database	No. of Obs.	Avg	Std.	Med.	P01	P99	Avg. Year	Med. Year	FactSet	Osris	FactSet	Osris
	Compustat	317,042	2,326	8,612	197	2	54,197	9	7	214,403	168,622	100.0%	100.0%
Total Asset	FactSet	384,874	1,921	7,415	146	1	47,629	9	8	-	180,608	-	100.0%
	Osiris	407,708	1,729	6,897	114	1	44,681	8	7	-	-	-	-
	Compustat	316,669	1,757	7,251	92	0	45,866	8	6	198,496	106,847	99.9%	93.6%
Total Liabilities	FactSet	384,158	1,443	6,143	67	0	38,709	8	7	-	111,146	-	91.6%
	Osiris	368,886	481	1,673	43	0	9,960	6	5	-	-	-	-
	Compustat	316,283	472	1,327	82	-7	7,956	8	7	212,980	148,393	99.6%	98.2%
Common Equity	FactSet	384,764	382	1,140	60	-20	6,761	8	7	-	164,502	-	98.9%
	Osiris	373,233	292	872	44	-15	5,221	7	6	-	-	-	-
	Compustat	263,929	446	1,270	80	0	7,595	7	5	177,216	138,868	99.4%	98.9%
Current Assets	FactSet	329,118	352	1,027	59	0	6,065	7	6	-	149,731	-	99.5%
	Osiris	376,219	305	950	42	0	5,338	7	6	-	-	-	-
	Compustat	263,918	358	1,090	53	0	6,649	6	5	180,067	130,824	99.5%	98.5%
Current Liabilities	FactSet	328,994	285	897	39	0	5,431	7	6	-	146,448	-	99.3%
	Osiris	372,430	240	802	27	0	4,577	6	5	-	-	-	-
	Compustat	291,650	45	244	-	0	1,424	3	1	62,125	41,321	98.1%	93.4%
Goodwill, Net	FactSet	248,070	60	298	-	0	1,917	2	1	-	44,544	-	95.1%
	Osiris	108,868	129	537	4	0	3,165	1	0	-	-	-	-
	Compustat	293,333	75	378	1	0	2,153	4	2	100,781	68,357	97.9%	97.7%
Intangible Assets, Net	FactSet	366,405	61	312	1	0	1,684	4	3	-	90,548	-	98.0%
	Osiris	239,986	80	365	2	0	2,121	3	1	-	-	-	-

Panel A: Balance Sheet Items, cont'd

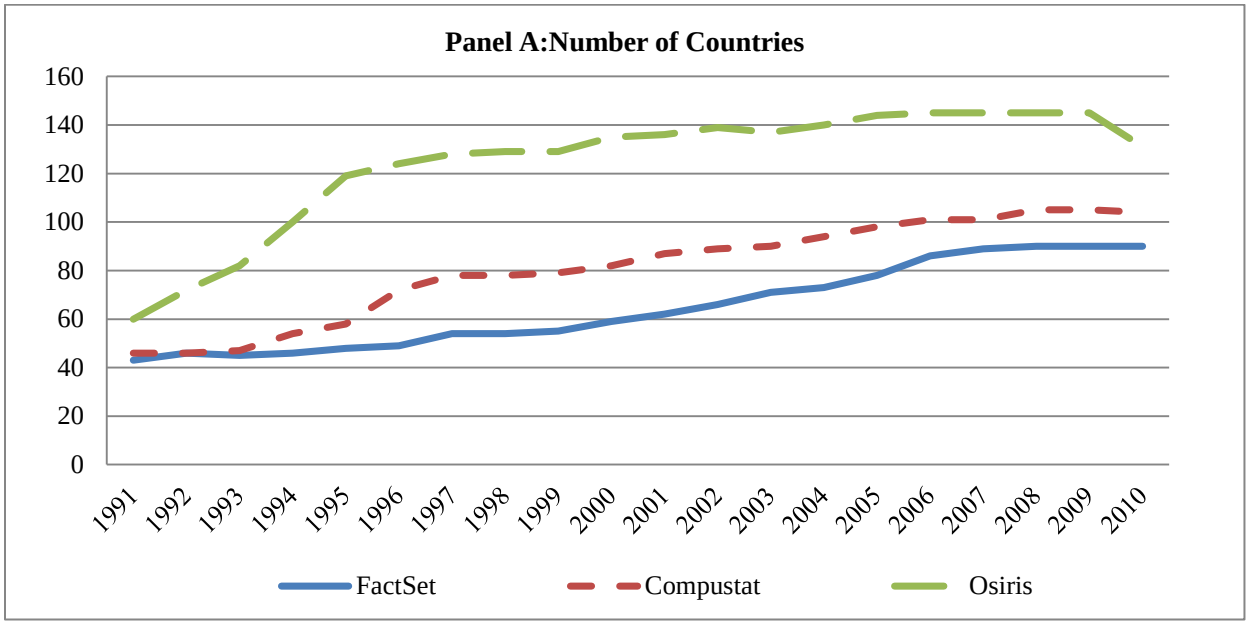
		Summary Statistics								Matched		Correlation	
	Database	No. of Obs.	Avg	Std.	Med.	P01	P99	Avg. Year	Med. Year	FactSet	Osiris	FactSet	Osiris
	Compustat	315,112	97	313	11	0	1,801	6	4	160,628	121,690	98.5%	94.2%
Inventories	FactSet	351,494	87	277	10	0	1,565	6	4	-	127,533	-	94.7%
	Osiris	313,205	82	269	10	0	1,488	4	3	-	-	-	-
	Compustat	307,023	329	1,284	11	0	7,992	6	4	140,225	-	96.9%	-
Long Term Debt, Total	FactSet	383,182	259	1,058	6	0	6,813	6	4	-	-	-	-
	Osiris	-	-	-	-	-	-	-	-	-	-	-	-
	Compustat	298,310	606	2,110	52	0	14,352	6	5	158,408	4,974	96.9%	73.5%
Prop. Plant and Equip.	FactSet	321,917	628	2,136	63	0	13,912	7	6	-	6,655	-	74.5%
	Osiris	295,069	79	308	4	0	1,927	4	2	-	-	-	-
	Compustat	304,639	161	623	12	-245	3,363	7	6	143,528	103,850	95.1%	92.7%
Retained Earnings	FactSet	335,173	147	591	9	-275	3,148	7	6	-	124,881	-	95.7%
	Osiris	363,448	109	456	6	-250	2,552	6	5	-	-	-	-
	Compustat	303,445	117	377	13	0	2,305	6	4	152,464	-	98.1%	-
Accounts Payable	FactSet	302,633	96	296	11	0	1,821	5	4	-	-	-	-
	Osiris	-	-	-	-	-	-	-	-	-	-	-	-
	Compustat	263,378	173	494	28	0	3,037	6	4	31,279	41,539	96.2%	96.1%
Accounts Receivable, Gross	FactSet	147,889	206	568	37	0	3,618	3	1	-	45,084	-	95.0%
	Osiris	338,047	101	323	12	0	1,919	5	4	-	-	-	-
	Compustat	263,608	121	361	18	0	2,131	6	4	167,193	116,734	98.9%	97.6%
Cash and Short-term Investments	FactSet	362,037	97	306	12	0	1,784	7	6	-	150,177	-	97.9%
	Osiris	369,391	81	267	8	0	1,492	5	4	-	-	-	-

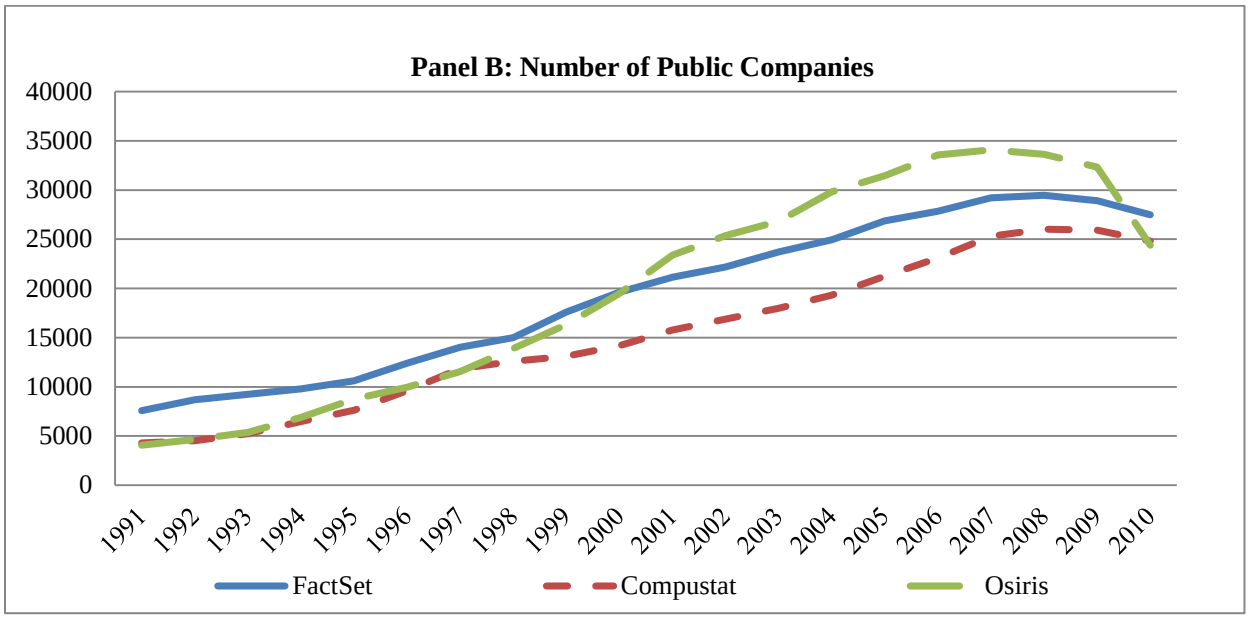
Panel B: Income Statement, Cash Flow and Derived Items

		Summary Statistics								Matched		Correlation	
	Database	No. of Obs.	Avg	Std.	Med.	P01	P99	Avg. Year	Med. Year	FactSet	Osris	FactSet	Osris
	Compustat	264,244	916	2,658	146	0	16,180	7	5	173,163	119,423	99.7%	98.9%
Sales	FactSet	386,017	788	2,458	100	0	15,365	9	8	-	139,900	-	99.2%
	Osiris	351,996	644	2,056	76	0	12,443	6	5	-	-	-	-
	Compustat	264,196	42	170	4	-128	928	6	5	173,112	128,702	98.1%	97.1%
Net Income	FactSet	385,950	35	143	3	-116	744	7	6	-	151,699	-	97.6%
	Osiris	373,839	26	110	2	-95	616	6	5	-	-	-	-
	Compustat	248,559	568	1,766	74	0	10,464	6	5	30,224	59,907	94.3%	-92.8%
Cost of Goods Sold	FactSet	330,843	557	1,690	77	0	10,497	7	6	-	28,178	-	-97.8%
	Osiris	314,722	-453	1,455	-52	-8,589	0	6	5			-	-
	Compustat	219,417	86	627	5	0	1,517	4	3	109,233	-	97.4%	-
Capital Expenditures	FactSet	353,785	77	3,415	4	0	1,424	6	4	-	-	-	-
	Osiris	-	-	-	-	-	-	-	-	-	-	-	-
	Compustat	288,325	40	137	4	0	927	5	3	136,336	95,949	98.2%	-94.8%
Depreciation and Amortization	FactSet	328,555	36	125	3	0	860	4	3	-	99,792	-	-95.6%
	Osiris	352,166	-28	104	-2	-659	0	4	2	-	-	-	-
	Compustat	302,134	49	268	2	0	977	4	3	120,642	4,999	72.4%	27.4%
Interest Expense, Total	FactSet	372,766	19	70	1	0	407	4	3	-	1,047	-	-29.8%
	Osiris	378,796	8	177	-1	-263	581	4	2	-	-	-	-
	Compustat	219,050	83	299	7	-51	1,878	5	4	119,748	92,747	98.5%	98.0%
Net Cash Flow from Operation	FactSet	339,229	81	310	5	-104	2,036	7	6	-	84,686	-	96.8%
	Osiris	202,661	52	206	3	-63	1,317	3	3	-	-	-	-
	Compustat	311,239	100	364	9	-47	2,238	7	6	117,585	53,919	94.0%	91.4%
EBIT	FactSet	328,941	48	173	5	-50	1,008	6	5	-	49,963	-	94.2%
	Osiris	372,514	47	173	4	-56	1,039	6	5	-	-	-	-
	Compustat	308,868	138	483	14	-30	3,045	7	6	103,903	69,154	96.3%	94.7%
EBITDA	FactSet	321,442	84	288	9	-24	1,753	6	5	-	68,086	-	97.3%
	Osiris	351,119	81	280	7	-19	1,814	5	4	-	-	-	-
	Compustat	-	-	-	-	-	-	-	-	-	-	-	-
Current Ratio	FactSet	328,748	2	3	1	0	21	8	7	-	8,304	-	96.5%
	Osiris	398,637	3	5	1	0	28	3	2	-	-	-	-
	Compustat	-	-	-	-	-	-	-	-	-	-	-	-
ROA	FactSet	349,204	1	17	3	-86	31	9	8	-	9,490	-	95.1%
	Osiris	366,922	1	14	3	-61	31	5	5	-	-	-	-
	Compustat	-	-	-	-	-	-	-	-	-	-	-	-
ROE	FactSet	339,947	2	37	7	-193	73	8	7	-	9,842	-	90.3%
	Osiris	359,528	1	40	7	-211	68	6	6	-	-	-	-

Figure 1: Numbers of Countries and Companies Covered by Global Accounting Databases

Panel A represents the number of country measured by the number of different country codes in each database. The country codes for FactSet and Compustat are the country codes for the country where the companies incorporated in each year. The number of companies in Panel B is calculated by the different database identifiers for the companies, such as gvkey in Compustat in each year. The Panel B represents public firms only and no replications due to different level of consolidations. The Panel C demonstrates the median of the country medians in every year for the total asset measured by million USDs. The median of medians for Compustat in year 1987 is 6,030 million USDs.





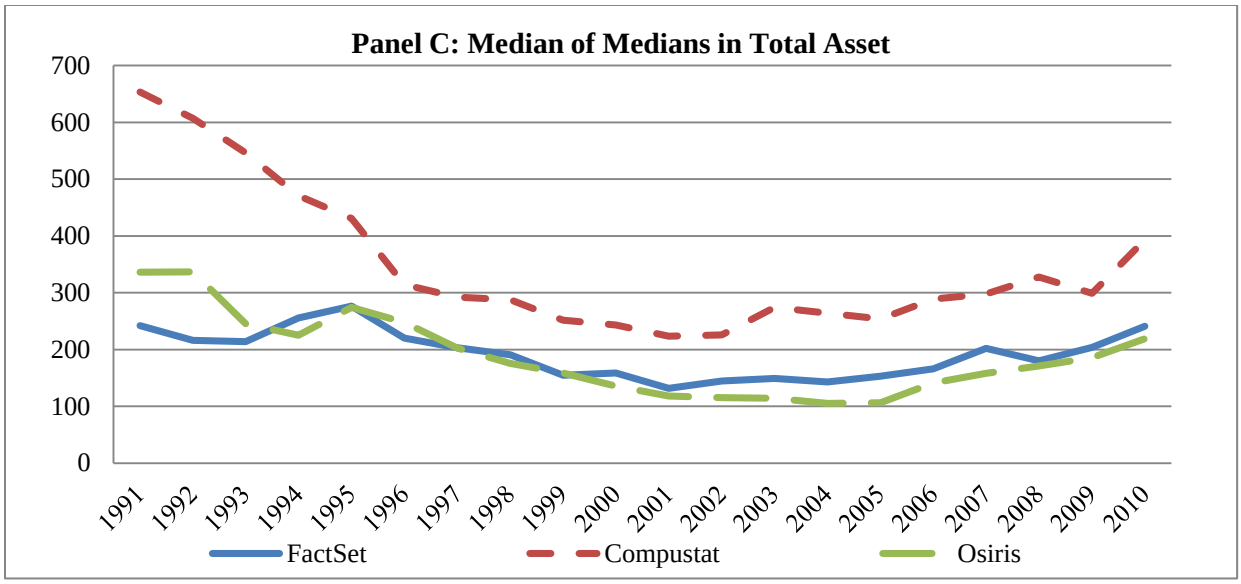
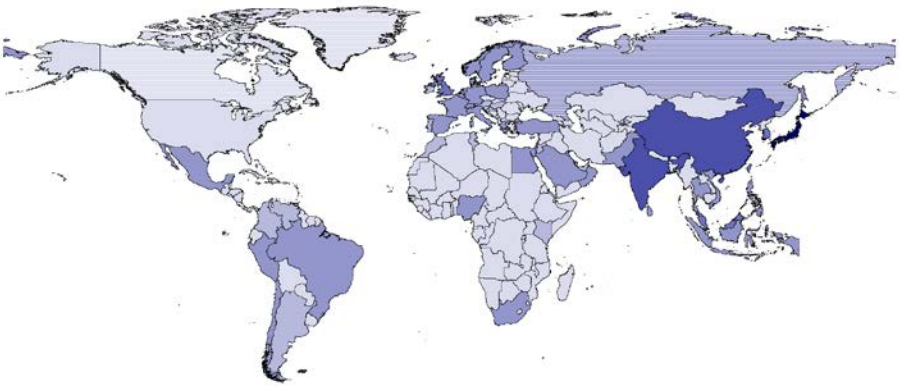


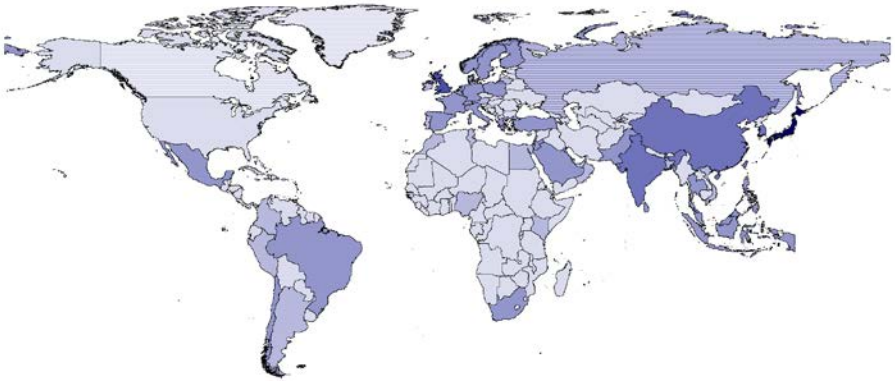
Figure 2: 2008 Geographic Demonstrations of Countries Available in Global Accounting Databases

The geographic demonstration in year 2008 for FactSet, Compustat, and Osiris is provided in Panel A, B, and C respectively. The number of companies is calculated by the different database identifiers for the companies, such as gvkey in Compustat. The countries of FactSet and Compustat are identified by the country where the companies incorporated in year 2008. The number of companies presented in Panel A, B & C are public firms only and no replications due to different level of consolidations. The Global Maps are generated by SAS GMAP procedure. Since GMAP does not consider Hong Kong as dependent political region, the number of companies in Hong Kong is not reflected in this demonstration.

Panel A: FactSet Non-North American Public Companies



Panel B: Compustat Non-North American Public Companies



Panel C: Osiris Non-North American Public Companies

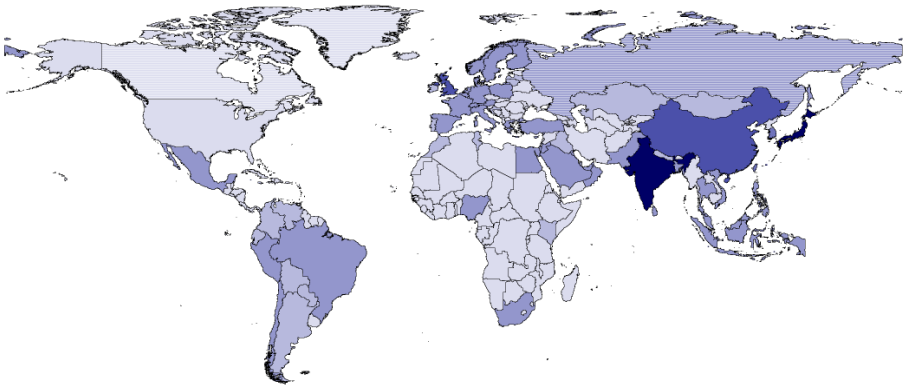
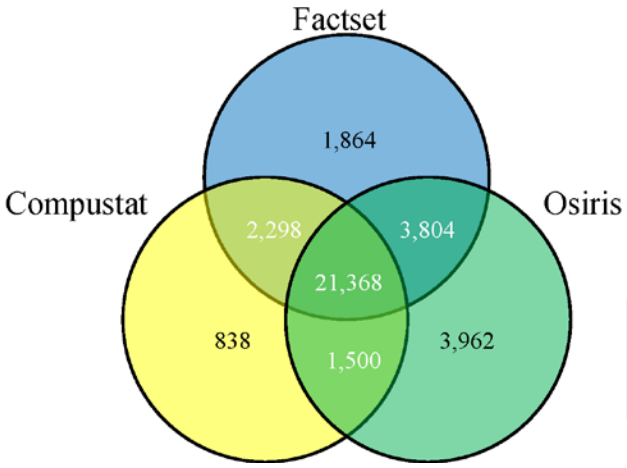


Figure 3: 2008 Venn Diagram of Companies in Global Accounting Database

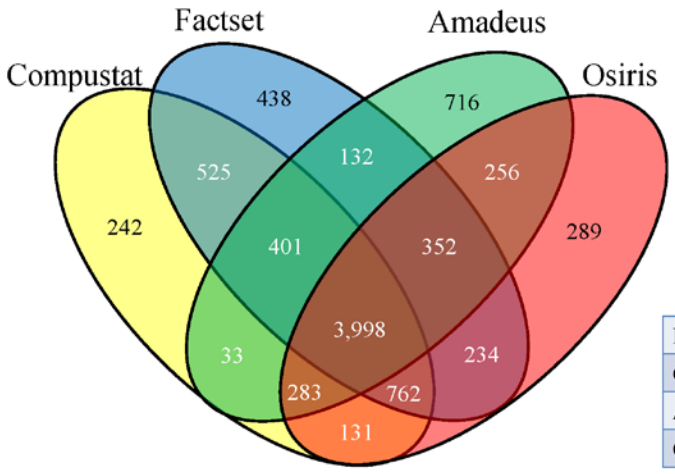
The Venn Diagram for companies outside North American is provided in Panel A. The Venn Diagram for European companies is provided in Panel B. The number of companies is calculated by ISIN of the companies in year 2008. The country ISO codes for FactSet and Compustat are the country codes for the country where the companies incorporated in year 2008. The all data presented in Panel A & B are public firms only and no replications due to different level of consolidations.

Panel A: Companies outside North American



FactSet:	29,334
Compustat:	26,004
Osiris:	30,634

Panel B: European Companies

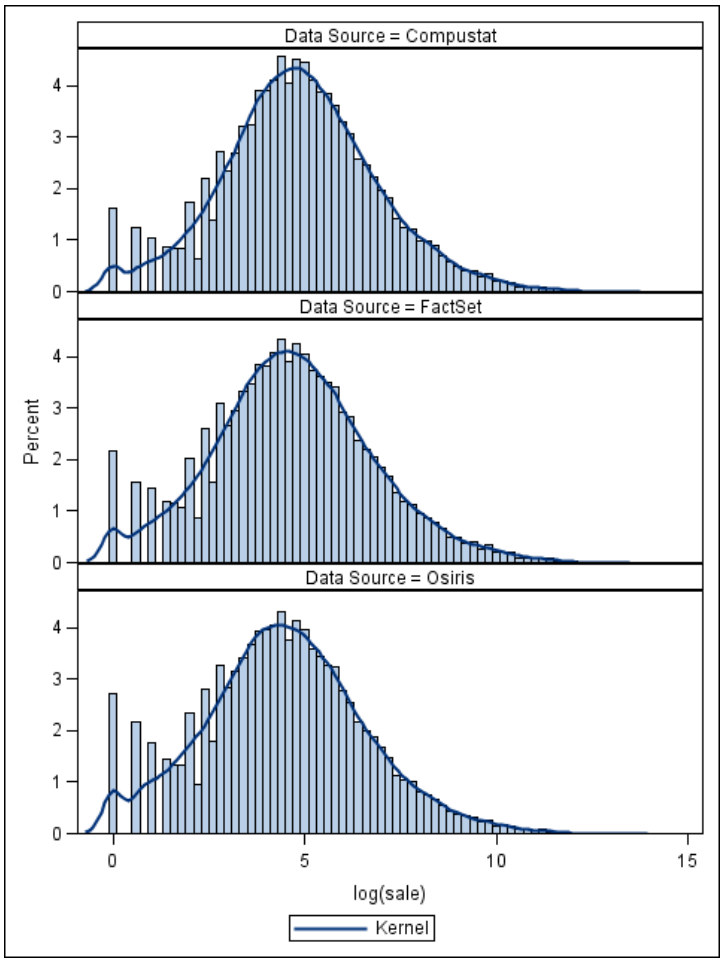


FactSet:	6,944
Compustat:	6,375
Amadeus:	6,171
Osiris:	6,305

Figure 4: The Distribution of the Natural Logarithm of Sales and Total Asset

This Panel A and B represent the histogram and kernel density estimate for nature logarithm of sale and total assets in 2008 for Compustat, FactSet, and Osiris respectively. The net sale and total assets are restricted to be positive and the value presented is not winsorized. The histogram and kernel density estimate of Compustat in top cell and those of FactSet or Osiris are in the middle and bottom cells respectively.

Panel A: Logged Sales



Panel B: Logged Total Asset