วารสาร คณะเศรษฐศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ปีที่ 7 ฉบับที่ 2 ก.ค. - ธ.ค. 2543

เศรษฐศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

KASETSART UNIVERSITY JOURNAL OF ECONOMICS

ISSN 0858-9291

50.-

Business Research with the I-SIMS Database

Pattaragit Netiniyom[1/]

Data is one of the most important components of business research. However, under the circumstances of new technology, the development of electronic databases encourages the researchers to spend more time on their theory and model developments, rather than in data collection. Two major research data criteria, reliability and validity, always require the researcher's attention. The I-SIMS is the electronic database of the Thai listed company that is designed and prepared by the Stock Exchange of Thailand. Nevertheless, an electronic database has some points that the users need to consider when operating. This paper presents the definition and basic concepts of business research, impacts of the data to business research ; validity and reliability notion, the definition of electronic database, the I-SIMS database profiles and the effect of the I-SIMS database on reliability and validity concepts.

Business research with the I-SIMS database

There are many ways to obtain information. For example, some people may obtain information from their hearings while others depend upon personal beliefs. These ways vary from person to person. To ensure that people have the same information, experiment and research methods are developed to provide information with support evidence that no one can deny. However, the experiments are more suitable for scientific studies rather than studies in the social sciences, of which the business field belongs.

Research method is "the diligent inquiry or examination in seeking facts or principles ; laborious or continued search after truth ; as, researches of human wisdom" (Webster, 1996). By this definition, research is a process of problem solving. It is commenced by various

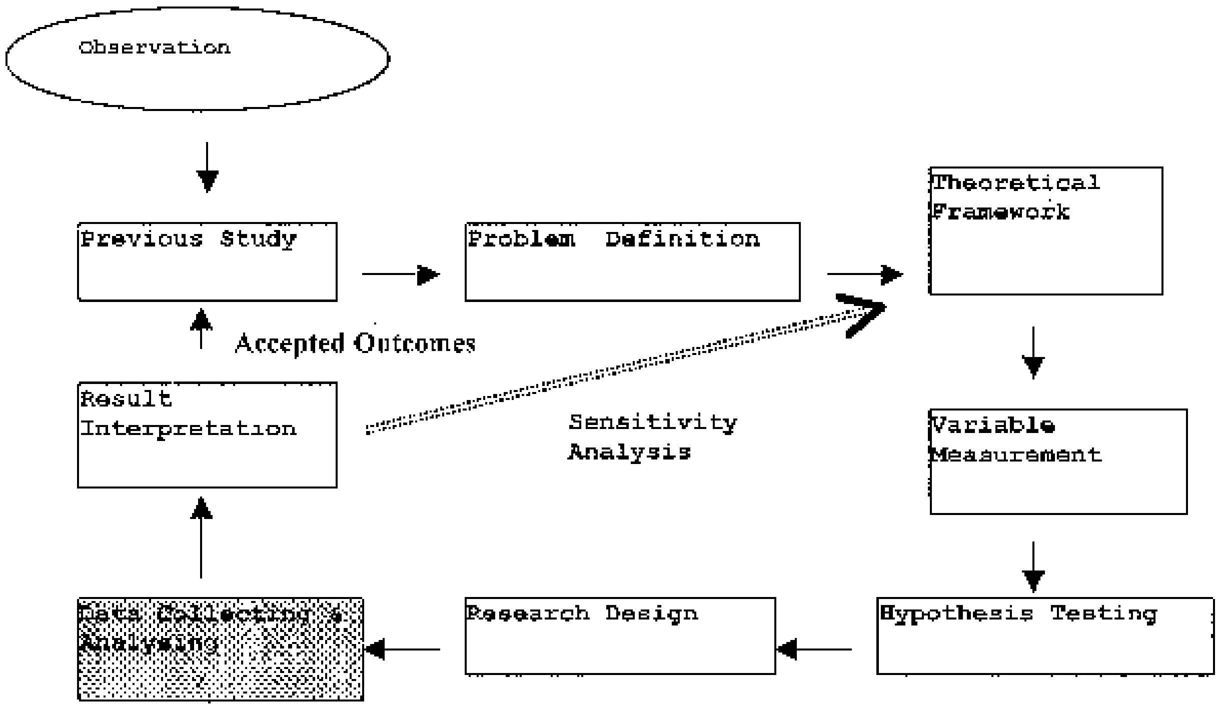
[1/] A Lecturer, Department of Finance, Faculty of Business Administration, Kasetsart University (At present, the author is doing a Ph.D degree at the Commerce Department, University of Queensland)

I am very grateful to my supervisor, Professor Frank Finn, for his support in my study. Thanks are also extended to the staff of the Faculty of Business Administration who have encouraged me in scholarship, and also to the Stock Exchange of Thailand for the I-SIMS database support.

observations or measurements that lead to interest identified questions being raised. Basing on various assumptions, hypotheses are the expected answers to the questions. The variables are selected to test the hypotheses. Sequentially, data are collected and analysed to get the results. Finally, if the researchers repeat their tests in different scenarios and the results confirm their previous assumptions they will be developed along with the previous studies into a theory. However, if the results do not properly explain the questions, the assumptions need to be revised. In figure 1, the diagram shows the global picture of business research process.

Figure 1

Research process



There are two basic types of research. The first is a pure research that is mostly inspired by curiosity. Another type is an applied research that aims to answer the specific questions. Both of them are important to develop knowledge. Most business researches are of the second type.

In the real world, there are many questions that could instigate research. Nevertheless, not all questions will inspire researchers to investigate the questions in detail. The objectives of research have to be clearly identifiable because the research process consumes plenty of time and requires a significant budget. The expected results should be objective so that the conclusion would be useful for the scope of applicability. Moreover, the research topics should be relevant to theory development, useful for policy implementation and research-able in terms of data or data availability. The ethical standard and cost of the research also are needed to be considered.

In business research their limitations still exist, which includes "precision" and "confidence interval". The precision refers to the degree of findings that are close to the "reality" while the confidence interval is a measure of the correct level of the researcher's estimation. If the degree of precision is high and the onfidence interval is wide, the research has complied to replicability, generalizability and objectivity, which are the qualifications of good research. In contrast with scientific research, business research generally involves human behaviors that are unreliable and different from person to person. This problem distorts the accessibility to the above mentioned quality of good research.

Business research is defined as "the systematic and objective process of gathering, recording and analysing data to aid business decision making. It reduces the uncertainty of a firm's environment" (Sekaran, 1992). To apply the concept of research process in business areas, the managers must identify either the threats that may trouble the firm or the opportunities that may enhance the firm's value. They have to select and implement a suitable course of action. The implementation always needs evaluation to monitor the results. The process of identifying the problems, implementing the suitable actions and continuously monitoring the outcomes are major activities of applied research in the business field.

There are many criteria for classifying the type of business research being conducted. Business research can be classed according to its business function, as in the areas of accounting, finance, marketing and management. However, some economic research is very close to business research, making it difficult to identify one from the other. Business research can be grouped by the techniques utilised in the research such as experiment, survey and observation. The category of the purpose is also a factor, which can be separated in to three objectives. The first group is exploratory research that is conducted to clarify the studied problem. The second group is descriptive research that describes either the characteristics of population or sample of research. The last category is casual research that conducts to identify the cause and effect of the relationship among variables which are specified in the research problem. Although the research in the business fields are varied in their objectives, a common purpose is to define and to reduce any threats to businesses and to enhance the opportunities of firms. These results support the long term existence of any firm's establishments.

Most business researches begin with observations. For example, a study of capital structure theory can be started by studying the differences between struggling and successful firms. The other example is about a well known corporate finance theory of the 1990s, an agency cost theory, was started by noticing the management behaviors among businesses. To date, there are numerous business research studies that have been done around the world. Because the research results generally bring fellow researchers to extend or to repeat former studies. Many people believe that these practices are simply the nature of research. The widely accepted research conclusions are predominantly based upon their assumptions or their circumstances. As time progresses, the conclusions are sometimes not true any more. Environmental changes generally inspire fellow researchers to conduct further research to provide their proposal of theory changes.

Even though all of the steps of a research process are important, this paper focuses on the part of "researchable in term of data" that relates to the electronic database characteristics. Data collecting is the most critical stage of research. Any mistakes in this step may bring the wrong conclusions and make the research useless. It is important that researchers organise the data once it is collected because the raw data rarely provides direct answers. After organizing the information the data pattern can be achieved. It will give more meaningful answers, especially when a typical patterns significantly exist.

Although the impact of input data could be considered in many different ways, such as in econometric view points i.e. homoscedasticity or stationary process, this paper concentrates only on the criteria for good research measurements, the reliability and validity concepts. These concepts are the elementary descriptions of how data do affect on research works.

The reliability of data is the extent to which data is consistent or the degree to which research measures are free from errors. The measurement results should be the same over time and across situations. This situation likes when we measure the length, the length we got should not vary by the instrument types. The reliability concept is very important in repeating of the research (Huck and Cormier, 1996, p.75). The test-retest method makes confirmation of the outcomes of the same issue at two separate points of time. To the test-retest of reliability, the statistical methods such as Spearman Brown or Kurder Richardson # 20 (KR-20) are applied to the measurements of the reliability coefficient. To examine the reliability of data sometimes a researcher only changes the statistical instruments or divides data into different groups. Each group is then tested and the results are compared. This means that data must be consistent, otherwise the conclusions from the study can not result in knowledge development. However, reliability measures only the homogeneity of the data, it does not scale the data stability.

Validity, is the degree of data accuracy with in the specified measurement process (Huck and Cormier, 1996, p.88). This means it is the ability to measure what a researcher intends to scale. The validity considerations are directly related to the purpose of the study. The identified measurement process can be designed at the time of data collection (concurrent validity), or it may be designed before the data is collected (predictive validity). Validity is a characteristic of the data and an instrument measuring. Although the validity should relate to other measures of the same issue, it does not follow this criterion sometimes. The reason is that the validity is quite subjective to the researchers because its degree is only an estimation.

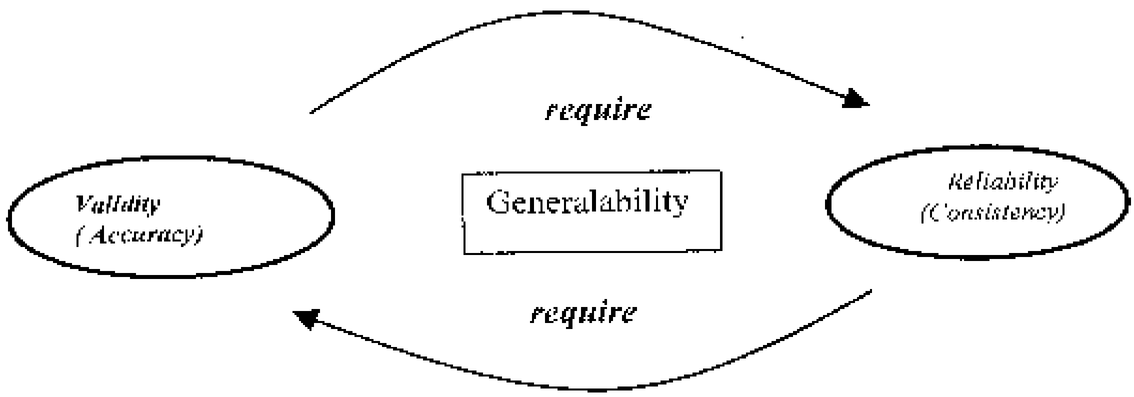
Overall, both reliability and validity are necessary in research, and the quality of data and statistical measurement selection are the main factors that directly affect them. However, not all research studies can automatically obtain both reliability and validity from their data at the same time. This means that when a high reliability coefficient exists, the validity may or may not be high as well, it depends upon the data. For example, when we conduct a preliminary survey for candidate A, we may find that about 20% of the sample intends to vote for A. However, this figure represents merely a fraction of the population, the result may change under different situations, such as if different people are present during the survey, or it is given at a different time. When the real vote comes out, it could not confirm how much the voters indeed chose the candidate A. This example shows that a reliable instrument (preliminary survey) are not always valid (real vote). Yet, they do have an effect on each other as shown in figure 2, but it is necessary for researchers to test each one of them separately. The reason is that the robust combination of validity and reliability, which refer to the same outcomes across firms and time would lead to the generalability that is the significant purpose of business research.

An electronic database is a database that can be accessed by computer (WorldNet-Princeton University, 1997). The electronic database is occasionally known as the on-line database, the computer database or the electronic information service. They have great influence on the world, especially in research works. For researchers, collecting data is the most tedious job. However it is necessary as most researches require a huge pile of data to prove their theories. The evolution of electronic databases has allowed researchers more time to develop their theories instead of handling the data collection themselves.

In general, favorable characteristics of elec-

Figure 2

Reliability and Validity relationship



tronic database for research purpose include : reliability, accuracy, comparability and consistency. Customer supporting service is also important. The feature details of a good electronic database are as follows ;

- **Reliability and accuracy :** Researchers expect high accuracy of electronic database but the electronic database deals with an enormous amount of data. Sometimes manual data collection is still required. Entering and repeating check the data takes time and money. However, precise information from databases is highly valuable, especially in terms of implications of the research results.
- **Standardized information :** In business research, the different practices of data recording of each firm needs to be recognized. For example, financial data from the financial statements vary due to an individual firm's definition and its disclosure. To form a useful electronic database, the transformation of all data to the same scale is desirable because the data must be used to compare with other data. For this reason, the comparable "one by one" concept is required.
- **Comprehensive coverage :** At present, there are many vendors who sell electronic databases. However, the economy of scale concept plays the big role, it is accepted as the guide line. A seller that covers more varieties of data, in terms of types and time periods, will be more utilizable. Market competition in turn leads to cheaper prices for the customer. The main reason is that it is very risky for anyone to use different sources of data without comparing the data profiles. Yet, comparing of different sources of data often demands that researchers spend more time in the completing of their papers.
- **Up to date information :** When research has been commenced, it takes time to complete the paper, especially in pure research where new ideas are investigated In contrast, applied research is generally adjusted from pure research or is the repetition of former studies in a different time period. In this case, up to date data is required. For example, many researches in finance area of share price based on risks and returns concept are accepted as direction of an investor's investment decision. Without any newly significant theory challenges, an accepted former study model is used to forecast the future with the latest data series.
- **Customer support :** Although the electronic databases that are provided by the various vendors are generally in the same format, the sellers' utilization of the whole database is indis-

pensable. It lessens the total cost of the computer database. A company's account representative must inform any potential customers of their availability of electronic database services for each customer. These services encourage more research to be done in the expanding business field.

At present, the most well known electronic database is provided by the Standard & Poor's Compustat, located in Denver, Colorado. It is a division of the McGraw-Hill companies. Established in 1962, the Standard & Poor's Compustat now provides different kinds of data such as S & P Equity Investor Services, S & P Comstock, S & P Securities Inc. and S & P Securities Inc.

The S & P Compustat database usually supports data of prices, dividends and earnings (PDE) of companies in the North American marketplace. Special electronic databases are also available. At present, the notable special types are the Global Vantage Database, the Executive Compensation Database (Execu-Comp) and the EDGAR. For academic purposes, the Compustat BackData that covers the quarterly data from 1950 to 1962 is also very useful.

Although the S & P Compustat is widely accepted by both scholars and general users, there are some concerns of utilizing the S & P Compustat. It should be acknowledged that all electronic databases may suffer from these concerns. The paper by Robert B. McElreath and C.Donald Wiggins in 1984 described the problems of the S & P Compustat database used in financial research. There are three meaningful errors of the S & P Compustat database that users need to be aware of when handling the research data. There are as follows :

1. Incorrect data : In collecting piles of data, data entering mistakes may occur at anytime. The effects of data errors sometimes change the conclusion of a study. Users can apply the statistical methods to monitor data errors. For example, the first order serial correlation methods such as Godfrey's LM test can check the error of a data series. However, at present, errors of data are curtailed by technological advances. Financial data is generally submitted to marketplaces through an on-line channel by a company itself but it is safer for users to consider and check with anomalous figures when they are visible.

2. Definition of data : Data of ny firms have a higher chance of facing this problem when the periods of study are longer. Although a firm favors to keep its businesses going, there are many things to change when environmental variations occur, particularly in merger and take over covenants. These transactions normally lead to significant change of fundamental business practices that a researcher must reckon case by case. Moreover, the regulation changes by government agencies such as tax systems must be viewed with additional attention. The definition of data directly affects the consistency of research and the comparability of results in the research methods.

3. Survivorship bias : Survivorship means the ex-post selection. It happens when the companies that exist at the beginning of the study. periods are not present in the latter periods. Its effects can be clearly explained via the example of share return. When researchers calculate the share return in some industries, they must select the companies that fulfil all studied variables through the periods of interest. Any companies that de-list from the market or declare bankruptcy are not included in the data set. It is obvious that the results of calculations do not demonstrate the whole set of truths and thus the researches do not accomplish their purposes.

For the S & P Compustat database, the limitations of size bias and timing must be recognized as well. Size bias results from the data vendors' profitability objective. They prepare data to sell and the companies that are big in size in terms of total assets or market capitalization are more highly preference by buyers than smaller ones. Nevertheless, the S & P compustat now is covering different size sectors of listed companies in the North American marketplace. There are S & P 500, S & P MidCap 400 and S & P SmallCap 600. Researchers can typically choose to do their research using these data sources and the results from each one generally represent the results of the others. Timing is normally influenced by lag effects. For example, the P/E ratio of December 1999 does not mean that the price on 30 December 1999 is directly cited to the earnings of the year 1999. The announcement of the financial statement of year 1999 must be a few months after the end of the year.

Moreover, a researcher must realize that there are more than two acceptable accounting standards. The items in financial statements do not explain themselves as shown by a single digit, however they need to be re-checked with regards to their assumptions and accounting methods.

In Thailand, there are on-line databases of Thai business surroundings such as www.worldbank.org or http://msn.byu.edu and www.corporateinformation.com/thcorp.html. Unfortunately, the information at those sites is superficial and not up to date. The Stock Exchange of Thailand or SET has developed its electronic database for a decade, distinctively beginning when the SET implemented a computerized trade system. In the past few years, the development of the Integrated SET Information Management System (I-SIMS) has significantly benefited both local and foreign users. The I-SIMS is composed of listed company information on the SET Information Management System database and trading information from the Automated Trading System for the SET (ASSET) database.

The I-SIMS database is one of the SETINFO electronic database groups. The SET further provides other forms of on-line databases. These are as follows ;

- **Price Reporting System (PRS) :** In the PRS database, users can retrieve the information of market summaries, real time and historical price trading, the listed companies' news, and listed company information.

- **On line ASSET Quotations (OAQ) :** The OAQ database is a real time market data because it is directly linked to the Automated Trading System for the Stock Exchange of Thailand (ASSET). It gives data of trading information.
- **Internet SET Daily (ISD) :** The Internet SET daily service is an on-line serviced data. It covers daily news of the SET or any listed companies, full financial statements and daily market trading information.
- **Data Request (DRQ) :** When users need a special type of data, they can contact the SET and be informed whether the SET electronic database contains that information or not.
- **Listed Company Info CD-ROM :** The information in the Listed Company Info is collected from the above mentioned database types. However, the data in this form is not in present time. They regularly lag behind current time about one or two quarters.

In this paper, electronic information of the Listed Company Info has been selected to concentrate on. This type of database covers all other types, and it is useful for all scholars and researchers to complete their researches. The I-SIMS program design is friendly to user. The design includes pull down menus, drop down lists and radio buttons, and check boxes for the user's selection. The main screen of the program is shown in figure 3.

Figure 3

The main screen of the I-SIMS in CD-ROM database

In Figure 3, the users can choose to collect the data that they are interested in just by opening the pull down menu in the company label. The label has all the listed companies' abbreviation symbols. By selecting the sector label, the users may choose to compare the data among industries. The data of past trading is also provided in this program. All data can automatically be exported to the Excel program.

The selection function label is very useful for both scholars and researchers. The selection menu as in figure 4, shows all listed companies in alphabetical order. The users can select all data lists and set the conditions to get a result by screening command.

The fundamental analysis of financial statement data can be supplied by the I-SIMS database. As shown in figure 5, the selection from the pull down menu and check boxes allow

Figure 4

The selection function pull down menu

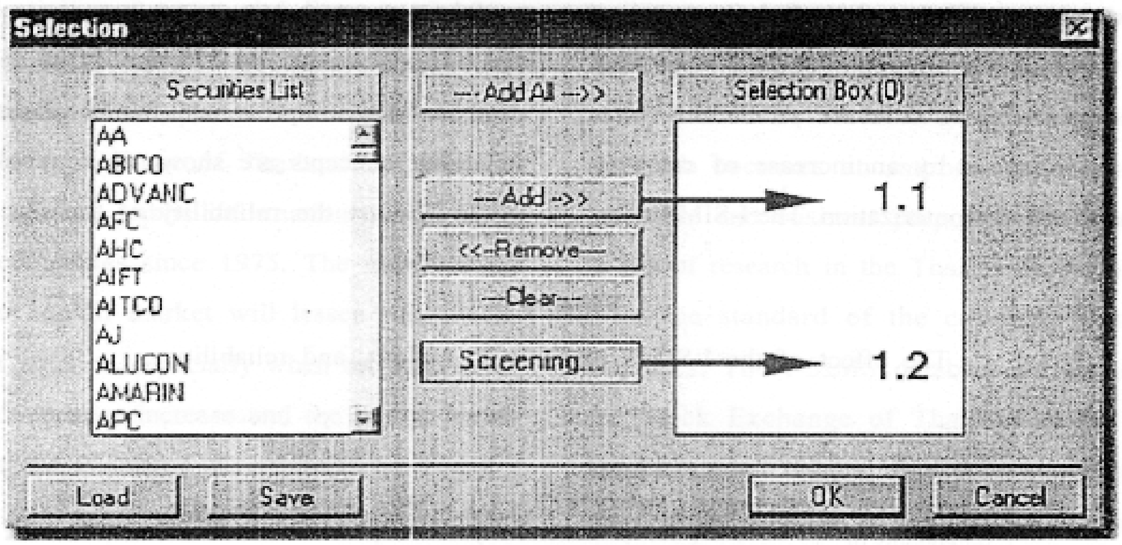
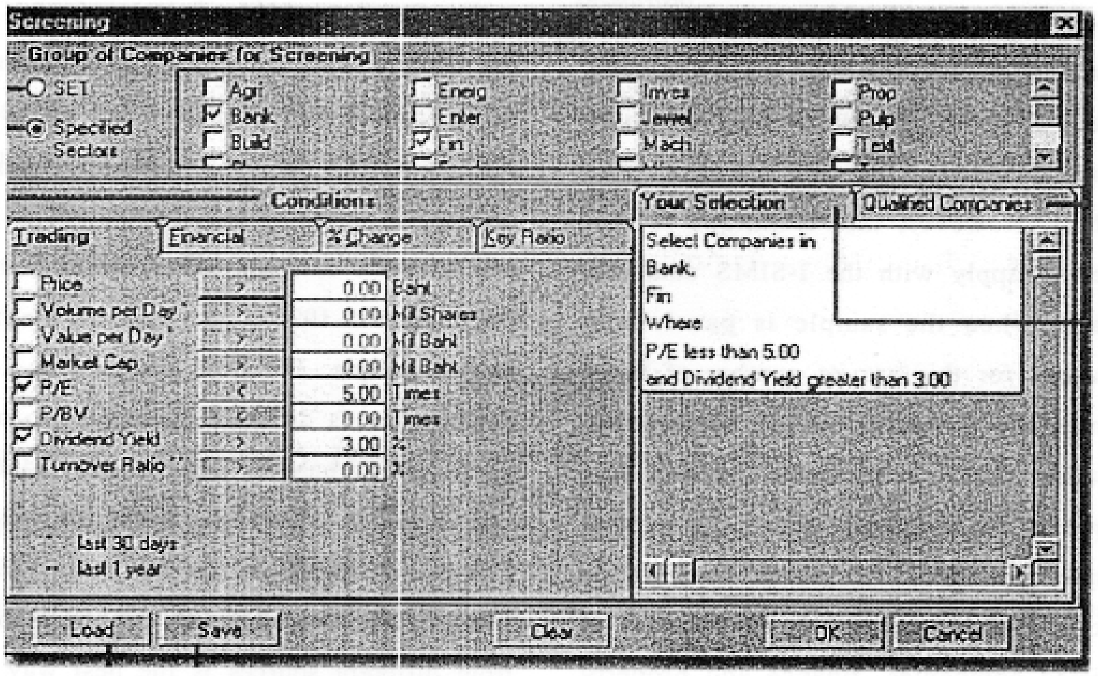


Figure 5

The fundamental financial statement analyses



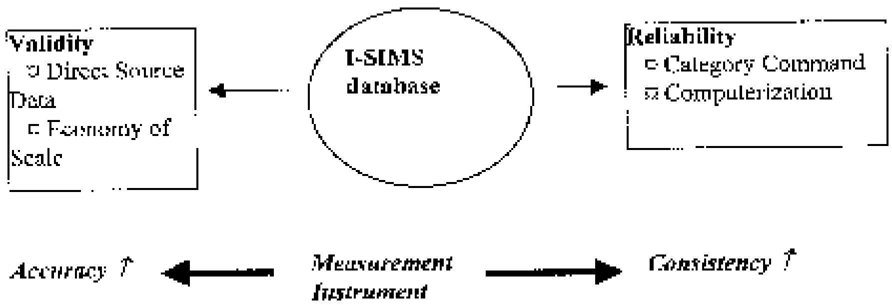
for quick research, providing researchers more time to develop their theories and models.

In overall the I-SIMS database encourages an improvement of validity and reliability concepts for business research in Thailand. The research validity is improved by an accuracy of the source of data because the I-SIMS database has transferred the financial statement data files which is submitted after an audited firm has verified its client. The economy of scale in maintaining the I-SIMS database would encourage the development of the database. It supports the new categories of data providing in the future and makes possible new researches as well as reducing database price. The reliability is enhanced by an increase of category command and computerization. The I-SIMS database makes comparison among groups efficiencies by a computerized database, which is very useful for research robust test (sensitivity analysis). The summarized effects of the I-SIMS database to the improvement of the validity and reliability concepts are shown in figure 6.

To test the reliability of the electronic database, the researchers may select the split-half method that measures the consistency by checking one half of the data set against the other half. However, this technique is still confined to apply with the I-SIMS database, especially when the sample is panel data. The reasons for this are its number of listed companies and the period length of database. Ramanathan (1998) implements the split half technique as the diagnostic tests. The number of sample data should be greater than 40 observations. Unfortunately, only two sectors of Thai listed companies, finance and property sector, are qualified. Moreover, the Thailand economic crisis of 1997 has increased the number of members of the REHABCO sector (the companies under the rehabilitation). The REHABCO sector refers to a company that does not comply with the minimum criteria of the Stock Exchange of Thailand, such as the requirements of the financial statement information disclosure or the financial status. The period length of the database that provides completed information in 1995 is also impeded by the split half technique. For both reasons the I-SIMS database needs different techniques to test for up to date reliability and validity. McElreath and Wiggings (1984) have tested the validity of Compustat database for 20 years (1960 to 1980). They recommended that the rechecking from different sources is the best way to get the highest validity. When the researchers get the results from their observations, they need to identify some outputs that highly differ from the others. They need to verify the input data from different sources such as other data ven-

Figure 6

The effect of the I-SIMS database to validity and reliability



dors. Although the database of Thai listed companies can be obtained from many data providers, the cheapest and the most precise way for local researchers is to obtain these information directly from the original source. The Stock Exchange of Thailand has supplied all the original public information papers of its listed companies since 1975. The development of Thai capital market will lessen this inconvenient practice, especially when the number of listed companies increase and the period length of database is longer.

In conclusion, the business research that consists of a problem identifying process and a problem solving process could be widely developed through the supports of electronic database. The existence of reliable and convenient electronic databases helps researchers to do more studies in shorter periods of time. The researchers are able to access databases that are up to date and to revise their papers as often as they need. This is important for the present globalization environment. More users lead to a more efficient database while the cost in maintaining databases is lower than before. After the Thailand economic crisis of 1997, the role of business research is growing, as the research becomes an important material for academic and business purposes. The developments of research in the Thai environment will raise the standard of the country's business activities. The I-SIMS database developed by the Stock Exchange of Thailand is the Thai listed company database and it is beneficial for researchers working in the Thai environment. However, it is the users' responsibility to check and re-check under the electronic database awareness concepts, the reliability and validity, including the uses of the I-SIMS database. While the program is continually being developed and improved, the progress of Thai electronic databases such as the I-SIMS database is a good sign for Thai business research.

References

Easterby, M., R. Thrope and A. Lowe, *Management Research*. 1995. Saga Publications

Huck, S. and W. Cormier. 1996. *Reading Statistics and Research*. Addison Wesley

McElreath & Wiggings. 1984. "Using the Compustat Tapes in Financial Research : Problems and Solutions". *Financial Analysts Journal*, Jan.-Feb., 1984

Ramanathan R. 1998. *Introductory econometrics with application 4th Edition*, Fort Worth : Dryden : Harcourt Brace College Publishers

Rosenberg & Houghlet. 1974. "Error Rates in CRSP and COMPUSTAT Data Bases and Their Implications", *Journal of Finance*, Sept. 1974

Sekaran, U. 1992. *Research Methods for Business.* John Wiley & Sons

The Standard & Poor Compustat. 1999. www.compustat.com

The Stock Exchange of Thailand, I-SIMS in CD-ROM and www.set.or.th

Zikmund W.G. 2000. *Business Research Methods 6th edition.* Dryden Press