

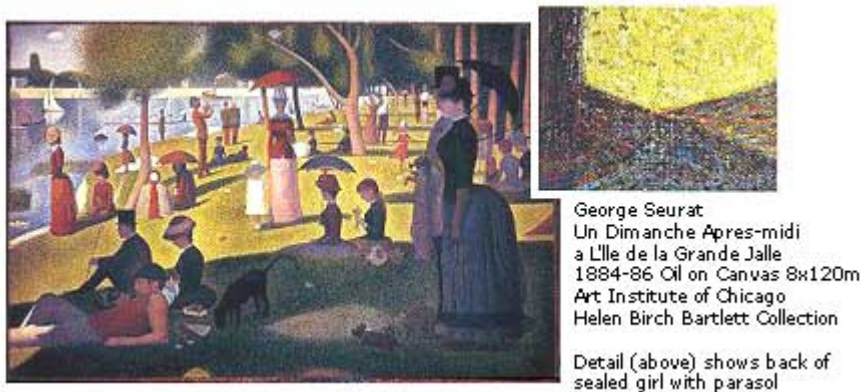
A Work of Art: CRSP's Database

By Marilyn Kun
GSB Chicago
March 1995



Great works of art take many shapes and forms. As GSB alumni, you may have visited the Art Institute of Chicago to view a special exhibit or a favorite collection. Perhaps you stood in awe before Georges Seurat's *A Sunday on La Grande Jatte*. As you walked closer to get a more detailed view, you may have marveled at the painter's technique—pointillism—and wondered how someone could have had the patience and precision to cover a canvas with millions of tiny dots.

Now, to appreciate the achievement that is CRSP, imagine that each of those dots is just one piece of stock information from the New York Stock Exchange. Individually, they seem overwhelming, don't they? But when you step back to take a look at the whole picture, the result is the original CRSP Master File.



George Seurat
Un Dimanche Apres-midi
a L'Ile de la Grande Jatte
1884-86 Oil on Canvas 8x120m
Art Institute of Chicago
Helen Birch Bartlett Collection

Detail (above) shows back of
seated girl with parasol

In 1959, Louis Engel, a vice president at the firm then known as Merrill Lynch, Pierce, Fenner & Smith, telephoned James H. Lorie, '47 Ph.D., (then professor of business administration, now professor emeritus since 1991) and asked whether anyone knew how well people were doing in the stock market relative to other investments. Although he was unable to answer the question, Lorie knew he could find the answer.

He proposed that Merrill Lynch fund a project to gather, clean, and make complete the prices, dividends, and rates of return of all stocks listed on the NYSE. With the new capabilities offered by computers, the timing was right for Lorie to develop a database that would maintain accurate securities over time. "There was a bit of foresight because that was the time in which large scale computers were first coming into play," explains Roberts. Hamada, dean and Edward Eagle Brown Distinguished Service Professor of Finance, director of CRSP from 1980 to 1985. "You could store, analyze, and create large databases. So Jim saw this as the time in which you could really find out the right answer to that very basic question."

Such a complete and accurate database would be invaluable to empirical research. No longer would researchers need to compile their own data. But, however worthwhile the project, collecting the data had to be done by hand, a painstaking task as time consuming as Seurat's.

In 1960, the Center for Research in Security Prices (CRSP) was established with a grant of \$300,000 from Merrill Lynch & Co., Inc., (from 1964 to 1986 their gifts reached in excess of \$1 million) and Lorie became the center's first director, a position he would hold until 1975. He and Lawrence Fisher, then associate professor of finance and associate director of CRSP, collaborated to gather the data. The two were faced with the awesome responsibility of checking the accuracy of each piece of stock information, and filling in the blanks for missing stock prices. In the end, Lorie estimated that between two and three million pieces of information were entered onto magnetic tape. In 1964, their database was complete, and they successfully demonstrated the capabilities of computers by analyzing total return—dividends received as well as changes in capital as a result of price changes—of all common stocks listed on the NYSE from January 30, 1926, to the present.

A seminal article by Lorie and Fisher in the *Journal of Business* reported the results. The article proclaimed that the average of the rates of return on common stocks listed on the NYSE was 9 percent. The front page of the *New York Times* financial section heralded the pair's findings. "If I had to rank events, I would say this one (the original CRSP Master File) is probably slightly more significant than the creation of the universe," said Rex A. Sinquefield, '72, chief investment officer and co-chair of Dimensional Fund Advisors, Inc. "The entire field of finance has been changed and developed through that database. And how appropriate it is that it happened at the University of Chicago, a university that was set up as a research university and known since day one in 1892 as one committed to empiricism."

The center currently offers subscriptions to data files in the following areas: common stocks on the NYSE, AMEX, and NASDAQ; government bonds; selected financial and agricultural futures; indices based on NYSE and AMEX stocks, US Treasury bonds and bills, and the consumer price index; and files of daily returns for stocks listed on both the NYSE and AMEX. Current subscribers include not only universities but also brokerage houses, corporations, banks, and government agencies, which use the CRSP files for commercial forecasting; fees for the services are modest.

In addition to providing grants for faculty and student research, since 1965 the center has continued to hold its semiannual Seminar on the Analysis of Security Prices, a forum to disseminate new work in the field of investment analysis. The seminar was the original idea of Lorie to present research studies early, often before they are published, for practitioners. Most of these studies come from the CRSP database.

According to Hamada, not only has the center been the basis for everything we have learned empirically about the US equity markets but it also has been emulated now in almost every country. "I think CRSP is the epitome of research centers in any field in terms of the kind of impact it has had in its field," he said.

Methodology

CRSP first sorts all stocks on the NYSE by market cap and breaks the universe into ten equal groups by number of names. These are called "deciles". Decile 1 is the group of the largest stocks on the NYSE and decile 10 is the group of the smallest stocks on the NYSE. CRSP then includes all equivalently sized AMEX and NASDAQ (OTC) stocks into the NYSE size decile in which they fit by market cap. All Small Cap Indexes are rebalanced quarterly.

CRSP 9-10 Index: The smallest fifth of NYSE stocks by name and all equivalents from other exchanges. Sometimes referred to as "micro-cap" stocks.

CRSP 6-10 Index: The smallest half of NYSE stocks by name and all equivalents from other exchanges. Sometimes referred to as "low-cap" or "small-cap" stocks. Similar in size to Russell 2000 Index.

CRSP 6-7-8 Index: Deciles 6, 7 and 8 of NYSE stocks and all equivalents from other exchanges.

This article contains the opinions of the author(s) and those interviewed by the author(s) but not necessarily Dimensional Fund Advisors Inc. or DFA Securities Inc., and does not represent a recommendation of any particular security, strategy or investment product. The opinions of the author(s) are subject to change without notice. Information contained herein has been obtained from sources believed to be reliable, but is not guaranteed. This article is distributed for educational purposes and should not be considered investment advice or an offer of any security for sale. Past performance is not indicative of future results and no representation is made that the stated results will be replicated.

Dimensional Fund Advisors Inc. is an investment advisor registered with the Securities and Exchange Commission. Consider the investment objectives, risks, and charges and expenses of the Dimensional funds carefully before investing. For this and other information about the Dimensional funds, please read the [prospectus](#) carefully before investing. Prospectuses are available by calling Dimensional Fund Advisors Inc. collect at (310) 395-8005; on the Internet at www.dimensionalfund.com; or, by mail, DFA Securities Inc., c/o Dimensional Fund Advisors Inc., 1299 Ocean Avenue, 11th Floor, Santa Monica, CA 90401.

Mutual funds distributed by DFA Securities Inc.