

Professor Lubos Pastor



Professor John Heaton



CRSP Launches Investable Indexes

How Booth faculty helped create
new groundbreaking products

By Patricia Houlihan and Diane Treuthart | Photos by Matthew Gilson

From academics to asset managers, anyone who wants the “gold standard” of stock market databases turns to the Center for Research in Security Prices (CRSP). For more than 50 years, CRSP has been the leader in providing research-quality market and index data. (See “CRSP Marks 50 Years,” page 32.) In November, CRSP leveraged its academic roots and introduced the first of 26 real-time investable indexes that reflect the way asset managers invest today.

To help construct the new equity indexes, CRSP turned to **Lubos Pastor**, Charles P. McQuaid Professor of Finance, and **John Heaton**, Joseph L. Gidwitz Professor of Finance, who shared with *Chicago Booth Magazine* the story behind how the groundbreaking products were created. The indexes blend advancements in academic research with current commercial practice. Offering a new approach to security migration, the indexes minimize turnover while retaining style purity.

Feature CRSP

How did you get involved with the CRSP index project?

Heaton: I'm on the CRSP board, and [chief operating officer] **David Barclay, '70**, and I have talked extensively about asset pricing and the need for new indexes from CRSP. The CRSP index project is related to my research not only in asset pricing, but also portfolio allocation.

Pastor: I teach a course in portfolio management, and I welcomed the opportunity to work on a project related to what I teach. I often debate with my students the merits of active versus passive management. The CRSP indexes should be of interest to both active and passive money managers.

How did you approach the task of developing investable indexes?

Pastor: We started at the end. We wanted to create indexes that would be of interest to both academics and practitioners. And, of course, we wanted them to be better than the existing practitioner indexes.

Heaton: We spent a lot of time designing the indexes to match the way we think about style from an academic perspective. We then worked hard to make them investable and practical.

How do the new CRSP indexes differ from CRSP's existing index products?

Pastor: They are high-frequency indexes that will be updated every second, whereas the highest frequency data that has been available at CRSP is daily. CRSP also will offer value and growth style indexes for the first time.

Heaton: And they're investable. The current ones are not. The current CRSP indexes are constructed in deciles and there's no worry about migration or trading costs. In the new indexes we move stocks more smoothly across indexes, making them much more practical for trading.

Pastor: We made these indexes easy to track by money managers, meaning a money manager can create a fund that mimics the performance of the new CRSP indexes at a low cost.

How are the new CRSP indexes different from the indexes of other index providers? What makes CRSP's new indexes unique?

Heaton: With the new CRSP indexes, we're sorting the eligible universe by market capitalization and assigning securities to the appropriate indexes according to their market cap. In a way, that's different and more robust than some of the current index providers. Some providers assign a fixed number of securities to an index, others selectively choose securities

for an index based on what they are trying to represent. But it's important that the new CRSP indexes don't deviate too much from common practice, because we are trying to make something that is practical.

The way we think about migrating securities across indexes is different. We tried to create an index where there's not a lot of churning, and there's not necessarily a lot of trading. At Booth, when we think about mutual fund performance, we think that indexing is great and you ought to minimize cost. That's what these indexes are meant to do. That's a really important feature.

Pastor: We've come up with a different way to do migration. Our approach relies on the concept of "packetting." When a security crosses a band separating two indexes, we migrate 50 percent of the company's holdings—"packet"—to the adjacent index. If a security just barely crosses into the neighboring index territory so that it's still within the band, we don't migrate anything. This "packetting" approach reduces churning and trading costs.

Packetting will be used with the market cap indexes as well as with the value/growth style indexes. The value/growth style-based indexes will use insights from academic research, some of it done here at Booth, to design value and growth dimensions that are different from other providers. The value and growth dimensions will be defined by multiple value and growth factors. Compared to other indexes, we will allocate stocks differently, and we will migrate them differently from one index to another.

When the new indexes were announced at the CRSP Forum in November, what questions did you hear from the audience?

Heaton: We were asked if we were trying to design the best set of indexes to track exposure to factors. Factors are variables that explain differences in average returns across asset classes or track risk characteristics. We thought about various ways to delineate stocks within different styles. But we're not saying we know how to construct indexes perfectly in order to track the best set of factors, risk exposures, or portfolios that explain distinct differences in average returns. The new CRSP indexes do achieve some of that, but are not meant to be the final statement on risk and return in the stock market. The new indexes will serve as new benchmarks for further research in this area.

Pastor: There was also a question about front-running.

Heaton: When a stock moves into an index, it can get a bump



“I often debate with my students the merits of active versus passive management. The CRSP indexes should be of interest to both active and passive money managers.”—Lubos Pastor

because institutional money follows it. This creates the possibility of market manipulation in advance of index reconstitution. Active managers could try to push a stock by buying. For an illiquid stock this buying pressure could really drive the price up and bump it over the market capitalization boundary for an index. That trading then becomes self-fulfilling because institutional demand is created.

CRSP has a randomization procedure that is meant to mitigate this type of front-running.

Pastor: We go further than the existing index providers in protecting index investors against front-running.

How will the new CRSP indexes change what researchers will be able to do? And what new type of research products will be available?

Pastor: It will be easier than ever before to work with high-frequency return data at the index level. So, I expect it will be easier for us to analyze issues related to market liquidity. Liquidity has become an important research topic in light of the recent financial crisis and the “flash crash” last May. People want to understand how prices move at very high frequencies as opposed to analyzing monthly or annual price data.

Heaton: Research in performance evaluation is very impor-

Feature CRSP

tant. The world of investing has evolved. People now ask, “Why are we paying such large fees when we don’t really get diversification or improved returns?” I think there’s going to be more emphasis on index investment as a way to minimize fees so the new investable indexes will be useful from a research perspective because they provide strategies that investors can actually do.

Also as more questions are asked about performance, we’ll have a set of indexes that provide a good set of benchmarks. For example, we’ll have off-the-shelf indexes with style delineated on value and growth dimensions in ways not available in other easily accessible academic indexes.

We’ll be exploiting CRSP’s advantage in the academic market, so we’ll have a set of indexes to go along with that advantage. CRSP is running these indexes internally, so we have access to all the data behind them in order to produce new and interesting research products. Data products may include things like corporate actions, quarterly performance, and some of the high-frequency data I mentioned like price, number of trades, and volume at the security level.

What was the greatest challenge in putting the indexes together?

Heaton: To address some of the practical questions in a way that is not too theoretical. In research, you come up with a model and make it theoretically sound, but sometimes you can lose sight of the practicality.

Pastor: Another challenge was to realize that, 10 years from now, people can look back and ask, why did CRSP do it this way? We wanted to make sure that the decisions we make will be good for years to come.

Heaton: We’d come up with an idea and the staff at CRSP would run hundreds of permutations to help us understand the boundaries of what’s feasible, what’s interesting, what is useful. Once we had all the results, we had to determine which approach would be the most robust, most useful, and have long-lasting applicability.

David Barclay insisted that the CRSP staff be very careful and methodical in terms of exploring different approaches to banding and migration, including even slight nuances. Between February and July, we explored more than 40 different approaches to banding and migration.



How did you settle on the final approach, and how would you describe it?

Pastor: CRSP did a very thorough analysis of our top 15 approaches to banding and migration. They ranked all approaches on total turnover, “bad” turnover, and style impurity. Bad turnover is turnover where shares traded for a specific stock by a fund exceed 10 percent of the average daily trading volume of that stock. CRSP tabulated and charted the results, and we selected our final approach based on how well it maintained style purity while minimizing turnover. That’s the key tradeoff—style purity versus turnover.

Heaton: I’d describe our final approach as practical. It accurately reflects the investment process of asset managers. It emphasizes cost efficiency and successfully balances style purity and turnover.



“The new indexes will serve as new benchmarks for further research in this area.”

—John Heaton

What did you enjoy most while working on these indexes?

Pastor: The practical angle. I initially had some reservations about how interesting this project would be. I was thinking, how hard can it be to capture the performance of a market segment? It's not the most difficult problem out there, but it has been more interesting than I expected. The practical challenges, especially regarding the migration of securities from one index to another and doing it in a cost effective way, were very interesting.

Heaton: With academic finance, almost all the research is immediately applicable. But a lot of it views the world from a pretty high level, like a helicopter hovering over the terrain. In this work, we were down in the trenches. We both thought, how hard was that trench work, really? It turns out it's difficult, and it has been more interesting than I expected; I learned a lot doing it. It gave both of us a new appreciation for what's done in practice and why things are done in practice.

It seems like this is a unique opportunity—to have access to new CRSP data and to be able to see your theory in practice. Is this different from what faculty do at other schools?

Heaton: CRSP itself is unique, so we have a unique opportunity because we have CRSP, an academic data provider—the first, the biggest, the most well known, the most respected. CRSP data has a long history and is checked and validated by researchers in a way that a lot of commercial data is not. The unique opportunity is here because of CRSP and we're building on it. So, that's what's different about this place.

Pastor: There's nothing like CRSP.

Heaton: It's the gold standard.



To learn more about CRSP and the investable indexes, visit crsp.com/index.html.

CRSP Marks 50 Years

A Blue-Chip Idea Pays Dividends **By Patricia Houlihan**

In 1960, Chicago Booth took a look at stock market data in such a transformational way that it allowed the rates of return on common stock to be measured for the first time. For 50 years, the Center for Research in Security Prices (CRSP) has provided high-quality market and index data to academic researchers and asset managers.

In 1959, associate dean **James Lorie** fielded an intriguing call from Louis Engel, a vice president at Merrill Lynch, Pierce, Fenner & Smith. The firm wanted to advertise how well people had done investing in common stocks, but Engel needed solid data and wanted Booth faculty to help collect it. With an initial grant of \$50,000 from Merrill Lynch, CRSP was born.

Building the first database of information from 1926 to 1960 took three-and-a-half years, a herculean effort by associate professor of finance **Lawrence Fisher**, whose attention to detail set the standard for producing clean data, said **Eugene Fama**, **MBA '63, PhD '64**, Robert R. McCormick Distinguished Service Professor of Finance. “The fact that it took so long to build was mostly Larry getting everything right. It was accurate data on dividends, shares outstanding, splits, all kinds of capital changes—and all on paper,” he said. “And it was pushing the limits of computing because computers were primitive at the time.”

In January 1964, using the CRSP data, Fisher and Lorie published “Rates of Return on Investments in Common

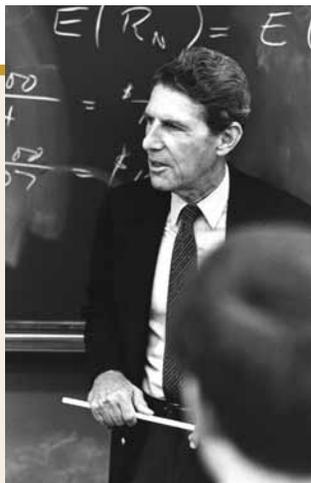
Stocks” in the *Journal of Business*, breaking new ground. The *New York Times* called the paper “the most comprehensive and accurate measurement so far of the rates of return on common stock investment.” But Lorie was concerned about whether others would embrace CRSP data and encouraged Fama to use it for a study of stock splits. The resulting paper, “The Adjustment of Stock Prices to New Information,” published by the *International Economic Review* in 1969, was the first to use CRSP data in an academic study. Subsequently, more than 400 universities in 30 countries have become subscribers of CRSP data, proving that Lorie’s concerns were unfounded.

Lorie as Visionary

Lorie’s other idea for promoting CRSP was to host seminars twice yearly for finance professionals where researchers would present their new empirical research on the market. Among those who attended was **John “Mac” McQuown**, who had heard Fisher and Lorie present their preliminary findings at Merrill Lynch’s New York office in 1963. When McQuown took a job with Wells Fargo, that institution became one of the first to sign up for the CRSP seminars, he said. Fama recalled the events were “very popular among business people. It was a way to introduce all the new research in finance at that time to the business community.”

“A big job.”

Professor **James Lorie** (right) fielded a call in 1959 from Louis Engel, vice president at Merrill Lynch. Engel’s request for data on how well people had done investing in common stocks led to the research that set the foundation for CRSP.



1926–60

Lorie proposed constructing a stock database that would involve gathering, cleaning, and providing the prices, dividends, and rate of returns for all stocks listed on the NYSE from 1926 to 1960.

Dawn of the computer age

New computer technology allowed for the maintenance of accurate securities information over time—once the original data had been collected manually. Lorie collaborated with associate professor of finance **Lawrence Fisher** to collect and research the accuracy of each piece of stock data.

The seminars, combined with the new databases, played an even more important role. Lorie, working with then-dean **W. Allen Wallis**, realized that business education in the post-World War II environment needed to become more professional and discipline-based. Moving Chicago Booth to the forefront would take a two-pronged approach: embracing empirical research, and building a community of top thinkers who would draw others to the school. The CRSP database, initially available only to Booth faculty, enabled the faculty to produce quantitative work in finance, and helped create the momentum that ultimately turned the school into a finance powerhouse.

Wallis and Lorie had played critical roles. “Jim Lorie was one of the key people in bringing in those who changed business education,” **David Booth, '71**, said. “And he recognized the importance of empirical data. Now, everybody takes empirical research for granted, but back in those days, it was a struggle. Computers weren’t big enough. Databases weren’t available. But he saw it was an important part of what business schools needed to do and the kind of research faculty do, and he made it happen. Since then, every other business school has moved our way.”

CRSP continued to develop new products throughout the 1970s and 1980s, when Booth and **Rex Sinquefeld, '72**, co-funded the new NASDAQ database with the National

Association of Securities Dealers to produce data they used at Dimensional Fund Advisors. CRSP added more products through the 1990s and 2000s. In 2009, after CRSP expanded yet again to offer daily data from before 1962, it became clear that the market was saturated. “It was time to introduce a new product that takes CRSP in a new direction, extending CRSP’s competencies from research into commercial investable products,” said chief operating officer **David Barclay, '70**. CRSP introduced the CRSP U.S. Total Market Index at its November forum, the first of seven investable indexes that will ultimately be part of a group of 26 indexes.

Behind it are principles and properties developed by **John Heaton**, Joseph L. Gidwitz Professor of Finance, and **Lubos Pastor**, Charles P. McQuaid Professor of Finance; they developed high-level objectives as well as the idea that indexes should reflect how money managers invest, a key consideration being cost. (See “CRSP Launches Investable Indexes,” page 26.)

In addition to creating the investable indexes, CRSP will develop research products from the indexes, providing new and different data for academics. Barclay said, “These new indexes are logical extensions of our core products and a way to continue CRSP’s long-term growth.” ■

ON THE WEB To learn more about the history of CRSP, visit crsp.com/50/index.html.

Building the database

Completed in 1964, the original stock market database was estimated to contain between two and three million pieces of information.

Results of the research

The new data set allowed for the average rate of return to be measured for the first time. Lorie and Fisher’s article in the *Journal of Business* stated the average rate of return on common stocks listed on the NYSE was 9 percent.



CRSP in the new millennium

In 2006, CRSP released the Pre62 database, which contains daily data from 1926 to 1962. Previously, only monthly data was available for the time period.