KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Spring 2014

OnAnalytics

diagnose. predict. optimize.

DISCOVER WHAT'S INSIDE
Converting Data to Insights

Welcome

As the Founding Corporate Partner of the Kelley School of Business Institute for Business Analytics (IBA), I am very privileged to open this spring's *OnAnalytics* magazine. In this issue you'll hear from two of Deloitte Consulting LLP's analytics thought leaders: Matt Schwenderman, who discusses the changing role of the CFO; and Oliver Page, who writes about building the right environment for analytics in the supply chain. Through these articles runs a common thread: the importance of integrating analytics into everything we do.

It's been a great year seeing the IBA develop into a mature organization and the interest it has generated among students, corporate partners, and other universities. I am regularly asked by my colleagues at Deloitte to share how we have established such a strong relationship with the IBA. Of course it starts with the talented faculty involved with the IBA and with Frank and Vijay's leadership, but it goes much deeper than that to a culture of innovation and flexibility that truly makes it successful.

As the sponsor for our campus recruiting at IU, I am also extremely impressed by the number of both graduate and undergraduate students we saw this year that had a relationship with the IBA or related business analytics curriculum. Clearly, it's a hot topic in the market—and IU is taking a leadership position among universities.

Evidence of the interest in learning more about business analytics was a presentation on analytics essentials in January by another one of my Deloitte Consulting LLP partners and analytics leaders, John Lucker. John's session at the newly renovated Social Science Research Commons at Indiana University was attended by more than 100 students and alumni, both on campus and via live webcast.

The term and discipline of business analytics has many meanings. At our first IBA board meeting, we had a lively discussion on this topic. From a strategic perspective, Deloitte recognizes business analytics as a "disruptive" technology in the digital enterprise and sees a convergence in the future of three major trends—big data, the Internet of Things, and machine learning.

Our vision for companies achieving competitive advantage in this area now revolves around new organizational, transformational, and technology approaches for enabling the analytics enterprise. From my vantage point in technology consulting, this includes identifying high-value business use cases and employing data visualization and discovery techniques; integrating big data with traditional data in data warehouses; optimizing core business intelligence and reporting environments; and architecting purpose-built, high-performance analytic technology ecosystems.

From an organizational and transformational perspective, there is no one-size-fits-all approach. Many companies have created new functions or repurposed existing ones to lead the charge on operationalizing analytics within their organization. Many questions exist, however—particularly concerning which use cases drive the most value (or cost take-out), who owns which components, knowing which technologies to use where, and what techniques exist to deliver data quickly while still maintaining control. It is our contention that without a relentless focus on the analytics enterprise, funding and long-term interest will waver and early initiatives will fail to deliver value, merely becoming science experiments in many organizations.

The IBA is a very important vehicle to foster broad adoption in the marketplace around business analytics concepts. I look forward to seeing you at the annual Kelley Forum on Business Analytics, April 3–4.

With best regards,

Mark Zozulia
Principal, Deloitte Consulting LLP
National Business Intelligence and Data Warehousing Capability Leader
Indiana University Kelley School of Business Alumnus

From the Editors

This issue marks the third year of *OnAnalytics* and features articles by Kelley School faculty, Kelley graduate students, and executives from Deloitte Consulting LLP. Deloitte Consulting was instrumental in supporting the launch of Kelley's analytics initiative and is the Founding Partner of the Institute for Business Analytics. We greatly appreciate the leadership and strong support of Mark Zozulia.

In this issue, we again focus on four key areas: accounting/finance, healthcare, marketing, and supply chain analytics. There are two reasons for these foci. First, they represent significant opportunities for analytics. Second, these areas cover many of the activities and interests of Kelley School faculty, students, and corporate partners. We have been pleasantly surprised by the variety of strategic and operational problems brought to our attention in these areas.

Accounting Professor Messod D. Beneish uses archival financial data to assess the likelihood that a firm is engaging in earnings manipulation, a practice in which reported earnings are more favorable than the reality. He finds that enhanced short-term results are frequently followed by negative longer-term returns, thus providing another tool for investors. Deloitte principal Matt Schwenderman suggests that the role of chief financial officer (CFO) is poised to expand beyond traditional activities. He argues that the CFO is well positioned to become a champion of advanced analytics throughout the organization.

Kelley Professor of Operations and Decision Technologies Jonathan E. Helm reports on a sophisticated optimization model for patient scheduling developed in collaboration with four hospitals. MS in Information Systems graduate student Mischella Felix-Jayachandran reports on her internship at a large health benefits firm. There, a major project to consolidate multiple data sources into a usable single source will ultimately support decision making.

Marketing Professor Shibo Li describes research using advanced statistical techniques to model a customer's sequence of views on a retail website. By tracking and predicting the path that online shoppers take, retailers can use the model to ultimately increase conversion rates. MSIS student Sean Yang spent his internship on a project to strengthen participation in a company's online community. He led a project to install a gamification system, which more than doubled the level of user engagement. MBA student Miguel Becerra worked with a not-for-profit university foundation to develop a measure of alumni engagement. This measure, once validated and implemented, will be used to enhance and direct fundraising efforts.

Supply chain analytics is concerned with the entire flow of goods and services. Operations Management Professor Gilvan "Gil" C. Souza reports a case study using advanced analytics to improve the efficiency of remanufacturing processes, an effort which ended up saving the company in question an estimated $5 million per year. Kelley MBA student Christopher Stafford spent his internship working on the distribution link of the supply chain. By defining and implementing a revised metric for transportation decisions, the company was able to substantially reduce shipping costs. Oliver Page, a senior manager at Deloitte, discusses unrealized opportunities for analytics in the supply chain. He cites the example of emerging technologies that track consumer data and demand in order to support supply chain operations, but notes that few manufacturers are using them to augment their internal data.

The breadth of applications described in this issue of *OnAnalytics* helps one to understand why analytics has become such an important activity and why it continues to receive so much attention in both the popular and professional media. The variety of ways in which analytics can be used is truly amazing and seemingly without limit. We hope you share our excitement for the ideas presented here and find this issue enlightening.

Frank Acito
Professor of Marketing
Max Barney Faculty Fellow
acito@indiana.edu

Vijay Khatri
Associate Professor of Information Systems
Arthur M. Weimer Faculty Fellow
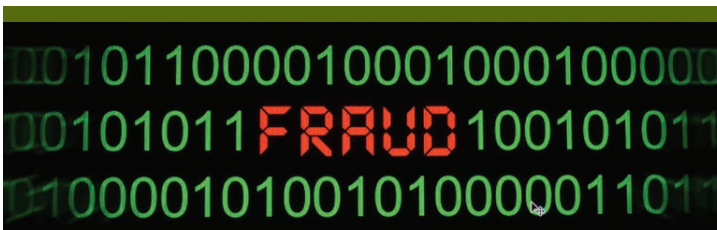vkhatri@indiana.edu

Predicting Firms that Manipulate Disclosed Earnings

Messod D. Beneish
Sam Frumer Professor
of Accounting

dbeneish@indiana.edu

> When I started this work, I intended simply to try and detect earnings manipulation. However, one of the validity tests I conducted early on involved determining whether the model was useful in creating a trading strategy. It was one way of persuading people that there was something predictive about the model.

Fraudulent financial practices have long been an issue in business, affecting not only a company's bottom line but also its investors. While negative press about high-profile companies such as Enron is familiar to anyone who watches the evening news, the actual incidence of earnings manipulation is likely much higher than is reported.

Forensic accounting and other financial analysis techniques can be used to detect the probability of earnings manipulation. The Beneish model, developed in the late 1990s by the author, applies financial ratios to data from a company's financial statements in order to determine its *M*-score—the likelihood that it has *manipulated* its disclosed earnings to appear more favorable than they really are.

In this study, the researchers take the *M*-score a step further, using it to predict stock market results. They propose that companies with traits similar to proven earnings manipulators are potentially dangerous investments. The researchers apply *M*-score analysis to company earnings and compare their computed likelihood of earnings manipulation to their actual performance in the stock market. Although these companies are more likely to appear as high performers in the market, the probability of them maintaining that momentum is slim.

Statement of Problem

A company that is likely to manipulate its earnings fits into a typical profile: its sales are increasing quickly; it shows deterioration in the quality of its assets and in its gross margins; and it uses aggressive accounting practices. Companies with these traits also tend to look like high performers in the stock market—but time after time they are unable to maintain those results.

Investors must look at many factors to determine whether a company is a good investment. The *M*-score is not the sole tool available to distinguish between a true high performer and one that uses tricks to

disguise its true earnings. It is, however, a good diagnostic tool when combined with critical thinking and with other proven predictors of earnings manipulation discussed below. Cornell University students used the *M*-score to flag Enron as an earnings manipulator several years before its downfall.

Data Sources Used

The sample was drawn to begin in 1993 and end in 2010. Because the *M*-score was estimated before 1993, the tests in this paper are conducted entirely out-of-sample. The authors required financial statement data to be available in the Compustat Industrial, Research, and Full Coverage files, and stock return data to be available through the Center for Research in Security Prices (CRSP) for the studied time period.

Analytic Techniques

The researchers eliminated financial services companies, any company with less than $100,000 in sales or in total assets, any company with a market capitalization of less than $50 million during the period prior to the portfolio start date, as well as any company with insufficient data available to estimate the *M*-score.

The researchers then applied the Beneish model to financial information for all companies in the sample, flagging the companies whose earnings had a high likelihood of having been manipulated. Next, the researchers compared returns for flagged and not-flagged companies.

The researchers continued their evaluation of the *M*-score's ability to predict returns by separating the companies into high- and low-*M*-score groups and examining their performance in those groupings. The researchers also compared the *M*-score to five proven predictors of future returns to determine whether the *M*-score could offer any information above and beyond those well-documented predictors.

Results

It is clear that *M*-score plays an important role in predicting future returns above and beyond the information provided by other commonly used predictive measures such as accruals. Large accruals are widely documented to indicate poor future performance, while small accruals indicate good future performance. By splitting small-accrual companies into those with high and low *M*-scores, the researchers found that they were able to find poorly performing firms even among those with small accruals.

Applying the Beneish model to the full sample resulted in flagging 17.4 percent of the sample observations as potentially fraudulent. Although this number is high relative to the number of firms that are caught or that admit manipulating earnings, the return results indicate that looking like a manipulator means poorer future return performance.

For the entire sample period, flagged companies showed statistically significant negative returns, while not-flagged companies showed similarly noteworthy positive returns. Not-flagged companies appreciably outperformed flagged companies in 13 of the 18 years studied. The researchers demonstrated that those companies with *M*-scores in the top 10 to 20 percent had a high rate of failure. Conversely, those with *M*-scores in the lowest 10 to 20 percent tended to fare well over time.

When the companies were separated into groups of 10 based on five proven predictors of future returns—market value equity (MVE), BTM (book to market), momentum, short-interest ratio (SIR), and accruals—not-flagged companies again outperformed the flagged companies. Additionally, in every instance, trading strategy results could be improved by taking into account the *M*-score of each company in the portfolio.

Business Implications

Fraudulent reporting costs shareholders billions of dollars when transgressions are uncovered. Not only are shareholders affected; misreporting also requires regulative and legislative bodies to waste resources dealing with the issues that arise, and negatively affects investor confidence. Reducing these incidences of fraud will in turn positively affect financial markets.

Both private and public initiatives must be used to fight accounting fraud. Legislation and enforcement are important, but perhaps the most effective action comes from sophisticated investors who are able to identify and hold accountable in the market those companies that are likely to manipulate their earnings. While it provides important information, using the *M*-score to predict whether a company's earnings will persist is only one component in uncovering the potential for fraud and applying that knowledge to making savvy investments.

Messod D. Beneish, Charles M. C. Lee, and D. Craig Nichols, "Earnings Manipulation and Expected Return," *Financial Analysts Journal*, 69 (2), pages 57–82, 2013.

Analytics and the CFO

Matt Schwenderman

Matt Schwenderman, principal, Deloitte Consulting LLP, provides consulting services targeted at helping companies improve their enterprise performance management and enhance the value of information. He currently has national responsibility for Deloitte's Performance Management Technology (PMT) practice.

mschwenderman@deloitte.com

Once, the CFO was the keeper of the company's finances: an overseer but not necessarily a facilitator. That role—and in fact the role of the entire finance team—is evolving. No longer only a steward, today's CFO is also a catalyst and a strategist who influences the direction of the entire company. Finance is increasingly becoming a strong business partner, particularly when it comes to evaluating trends. As a cross-functional area, it has the unique ability to examine all the factors in play and put them into a context on which the organization can base its decisions.

The traditional activities of finance—for example, audits, reporting, and planning—make it a natural choice to champion increased use of analytics through a more balanced approach. Rather than relying on finance to separate the good data from the bad after the fact, or provide context to historic results, a more sensible approach is to involve finance throughout the entire lifecycle of

business decision making. Using financial analytics, companies can better assess risk, cut costs and control expenses, manage working capital, enhance revenue, and improve operational efficiencies.

Championing Analytics throughout the Organization

Why is finance so well positioned to champion analytics throughout an entire organization? First, it has the history and experience of being data driven. It's traditionally been concerned with return on investment, something that analytics projects must ultimately demonstrate. Its workforce has the training and skills to analyze and bring econometric and advanced techniques to bear on business decisions. And finally, it has the core competency of being able to evaluate and maintain data quality.

For example, a company's marketing department would traditionally put together a

revenue plan and then submit it to finance and the executive team, who would then allocate resources. But this might require finance to question the assumptions the marketing team made. Rather than being in a position of questioning others' assumptions and activities, why not have finance work with the marketing team to use advanced analytics from the start? If marketing is responsible for putting together strategies, then they should be using analytics, quality data sets, and robust assumptions to run models and analyze trends.

Emerging Role of CFO

This is where the emerging new role of the CFO comes in. Instead of producing reports, handling financial governance, cutting costs, and distributing information, the CFO has the chance to advance predictive analytics, develop strategic cost analysis, define growth strategies, and drive innovation. These activities provide a more integrative view of the company and allow it to adjust more quickly to changes in the market. Putting finance at the forefront of planning allows a company to be more nimble.

CFOs are being driven to deliver more value-added services more quickly. The current push for real-time planning comes primarily from competitive pressures influencing the speed at which decisions must be made, increased volatility and velocity of change, and the availability of enormous amounts of supporting data, all tied into the latest available technology. But real-time planning is not without its own set of challenges. To be effectively involved in strategic planning, finance must shift from examining historical data to applying knowledge to predictive modeling. Finance and accounting teams can no longer be siloed; they should be integrated into the organization. Perhaps most importantly, management must realize that development programs that

emphasize talent rotation and short-term assignments lead to loss of continuity in understanding and missed opportunities to maximize the use of analytics.

And it's imperative that companies understand that a solution that works today might not always be valid. Companies tend to track the same metrics over time without necessarily ensuring that those metrics are still relevant. It's vital that a company pay attention to the right key performance indicators (KPIs). We have found that companies with formal financial management processes performed better in terms of shareholder value than their peer group.

Increasing Shareholder Value through Analytics

Finance has a real opportunity to increase shareholder value through the use of analytics in real-time planning and forecasting. For example, media companies have developed social media sentiment analysis techniques to forecast the potential revenue curve for an upcoming movie. While this can certainly influence marketing techniques as that particular movie's release date draws near, the information gathered will be much more powerful if the studio applies that information to a broader financial forecasting model. By applying what they learn from each upcoming release to movies that are coming next year or the year after, the studio is increasing profits because they're more appropriately allocating resources (capital) to the marketing and promotion initiatives with the highest likely return. This is the crux of real-time forecasting—continually learning and feeding data back into the model, and seeing how that translates over time. This model is a logical step forward, putting finance at the forefront of company strategy.

Optimizing Efficiency through Patient Scheduling

Jonathan E. Helm
Assistant Professor of Operations & Decision Technologies

helmj@indiana.edu

With a limited number of beds to go around, hospitals must schedule procedures as efficiently as possible. Patient scheduling is a foundation on which both quality of care and financial stability are built. From the need to find balance was born the Hospital Admission Scheduling and Control (HASC) problem, which researchers in the 1970s addressed through simulations. In 1980, full implementation of the Hill-Burton Act resulted in hospitals adding beds to their facilities, and the need to address efficiency decreased sharply. An aging population and changes in health care in the 30 years since have found hospitals once again facing scheduling issues.

Statement of Problem

The variable patient load—*census variability*—on any given day is an administrative problem all hospitals face. This shifting load leads to decreases in patient safety, quality of care, and access to services, as well as to increases in costs, patient deaths, nurses' job dissatisfaction, and staff turnover. Overcrowding of wards can lead to patients being shifted to a ward not best suited to their needs—or worse, to procedure cancellations or patients being diverted to a different hospital. Strategically scheduling patients for procedures to avoid overcrowding—*census smoothing*—is vital.

While a simulation model was effective in the 1970s, changes in technology have opened the door to new methods of solving the HASC problem. The researchers propose new models for elective patient scheduling that generate optimal schedules by taking into account the entire hospital. Additionally, these models address the conflicting objectives of admitting as many elective patients as possible and not admitting so many that they must turn away or relocate them.

Data Sources Used

The researchers spent more than four years collaborating with four small-, medium-, and large-sized teaching and non-teaching hospitals in the United States, Netherlands, Singapore, and Canada. While they based their analysis of optimal scheduling practices on one hospital in particular, the researchers were able to validate their model across all four. The researchers extrapolated significant prior research and simulation models to develop an analytical model, which was verified through a year's worth of historical raw patient data provided by the four collaborating hospitals.
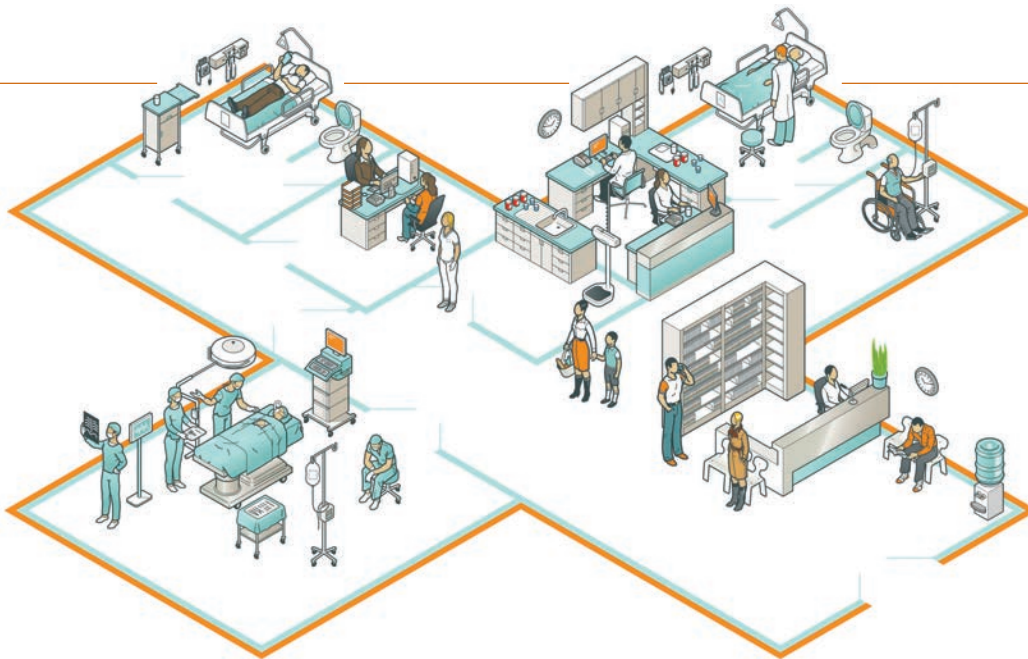
Analytic Techniques

Unlike much of the prior work on the HASC problem, the researchers create a stochastic model that looks at individual hospital wards as an interconnected whole and maps the patient visit from admission to discharge. Additionally, they include two elements not previously considered: in-hospital transfers due to changes in patient condition and patients being served by a ward other than their preferred one due to overscheduling. This unique model was based in part on research done on cellular communication in the mid-1990s. Cellular phone users transfer from one cell tower to another as they move, and studies were done to predict traffic between towers: how long would a cell user stay on one tower? What tower would they move to next? This pattern of movement is not dissimilar to the path a patient takes when moving through the hospital.

The researchers first determined the typical path for each type of patient. To align with typical scheduling structures, patient type was chosen based on the admitting service (e.g., cardiology). The patient path, then, comprised the probable type, timing, and duration of resource usage—that is, beds in various wards—during the patient's hospital stay. Using this path, the researchers were able to model the total patient population (both elective and emergency) in each ward for a given schedule on a given day of the week. They then linked this process to variables typical of the elective admission procedure to determine the ideal mix and volume of patients to admit at any given time.

Results

Past attempts to address HASC have involved simplifying the model of patient care so it does not take into account the fact that patients follow a path through the hospital. It's not as simple as "check in—occupy a bed—check out." The researchers' use of a model that reflects the reality of a typical patient stay in the hospital yielded promising results. They have created a decision tool that provides hospital administration with the flexibility to determine their optimal schedule, rather than a mandated approach to scheduling that does not take into account all possible factors. Once a hospital has used this tool to optimize its schedule, it need only implement simple decision rules provided by the tool to increase its efficiency and decrease its costs.

These rules take the form of caps on the number of elective admissions to schedule on each day. That is, the hospital limits procedures to a maximum of x procedures of type y on any given day, which is similar to the common practice of block scheduling. The decision tool the researchers have developed makes it simple for the administration to determine optimal numbers for x.

Business Implications

While the researchers' decision tool has the ability to increase efficiency for any hospital that implements it, the problem lies in the power the administration has to address the problem. The health care industry is rooted in well-established historical and cultural norms. For example, many surgeons prefer to operate on Mondays and Tuesdays, while they traditionally avoid doing elective procedures on Fridays. Consequently, Mondays and Tuesdays too often see overfilled beds, but by the weekend the patient population has dropped significantly. This cycle repeats endlessly.

If a hospital wants to run more efficiently, it must change its culture before a decision tool like that developed by the researchers can be effective. This depends primarily on hospital leadership— are they willing to undertake the necessary changes? The bottom line is that changes will come if the financial incentives are significant enough. In the current health care climate, the researchers' tool may be just what the doctor ordered.

Jonathan E. Helm and Mark P. Van Oyen, "Design and Optimization Methods for Elective Hospital Admissions," 2014. Working paper.

Data Consolidation in Health Care

Mischella Felix-Jayachandran

MS '14, majoring in information systems

Interned at a large health benefits company

mfelixja@indiana.edu

Health care is a hot topic these days, and companies providing health benefits are often in the spotlight. It is essential that these companies develop and use tools that allow them to provide affordable health care while still maintaining a strong bottom line.

Having grown significantly through recent acquisitions, a leading health benefits company needed to build new financial analytics and reporting systems to meet these goals. As an intern, Mischella Felix-Jayachandran found that the basis for developing such analytical and reporting capability was data consolidation.

How did data consolidation benefit your internship company?

As a major health insurance player, the company was very interested in developing analytics on financial risk and uncertainty. Consolidating data ensured that the extracted information was reliable and valid for effective decision making. Data consolidation also drove automation and improvements in other processes such as financial forecasting and project reporting, saving the company time, effort, and money.

What were the steps taken to complete the project?

I worked on consolidating and streamlining data to build an analytical and reporting capability for the company's portfolio planning team. I first identified available data sources

and what targeted information was needed from the data, and then created a data map to identify redundancies and inefficiencies in the current system. Next, my team spoke with all business and technology stakeholders to determine the best database for the project based on size and cost constraints. Once a database was chosen, we modeled the data to find the best consolidation method to fit the business need. We then migrated the data to the target database and built the analytical and reporting capabilities.

What tools can be used for managing data consolidation?

Toad for Data Analysts is a user-friendly tool that can be used to extract data from different systems, transform them into a standard format, and load them into a target system. Microsoft Excel and Access and the PowerPivot add-in for Excel are also valuable tools for smaller, decentralized databases.

How is data consolidation important for analytics?

Consolidating data helps an organization get a single view of the information required for effective decision making. Most companies deal with multiple data sources, but building analytics on such disparate sources is expensive and yields inaccurate results.

How did your courses at Kelley prepare you for this project?

The MSIS core classes gave me the practice and experience to analyze and solve business problems. The Kelley data analytics competition introduced me to common data challenges, giving me hands-on experience with analytical tools. Case-based learning and knowledge-first learning sharpened my problem-solving skills.

What did you learn from this experience?

I learned the importance of determining the big picture—how the scope of your work fits into the overall scheme of a project. I also learned that establishing a standard data management process is essential for successful integration between teams and applications. I gained valuable data management skills that I will carry with me as I start full-time employment.

Tracking Customer Behavior for More Effective Marketing
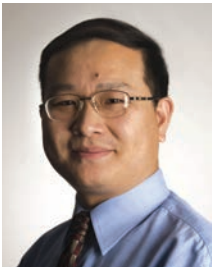
Shibo Li

Associate Professor of Marketing; Arthur M. Weimer Faculty Fellow

shili@indiana.edu

Online stores track demographics, purchase history, and the means by which a customer accesses the store in order to provide information targeted to that particular customer. While these are important pieces of information, there is a much richer data stream to be mined—*clickstream data*. These data deliver information about the full path a user takes both across websites and within a website, from the first page viewed to the user's exit.

The researchers have developed a method for capturing and mining these data, allowing online retailers access to a much wider scope of information. Not only can retailers use this information to precisely target individual consumers, but they can also predict from these data whether a user will purchase an item. This predictive ability allows the retailer to potentially influence browsing (i.e., non-purchasing) consumers to make a purchase.

Statement of Problem

Although there is considerable customer information at their disposal, online stores have been limited in their ability to precisely target individual customers. Without a methodology to analyze the paths customers follow through the retailer's website, it is difficult to determine what steps to take to most effectively market to each consumer. Looking at a customer's purchase history in isolation is informative, but there is much more to be learned.

The researchers' methodology allows managers to analyze a customer's path through a website and predict whether or not the customer will complete a transaction. Looking at the customer's path and whether or not it led to a purchase is critical in terms of designing a website or deciding on the optimal marketing mix variables (product, place/distribution, promotion, and price).

Data Sources Used

The researchers examined data for popular online bookseller Barnes & Noble for the month of April 2002. The data were obtained from Comscore Media Metrix (CMM), an Internet marketing research firm that tracks online consumer behavior. CMM provided software to a panel of volunteers, who installed it on their home computers. This software tracked detailed data about their online usage, but it was missing a key element: the type of content on each page a user visited.

The researchers wrote a computer program to recreate browsing information from these data—most importantly, the actual content of each web page. This enabled them to categorize each page into a certain type (e.g., home, account, category), creating a more complete picture of customer behavior.

Analytic Techniques

The researchers developed a *dynamic multinomial probit model with a hidden*

Markov process, which takes into account the effect of different variables to predict likely outcomes of a user's browsing session. This model tracks two important elements: the category into which each individual web page falls, and the ability to track the sequence of events that make up a browsing session. Additionally, the model allows the researchers to consider consumer demographics and other characteristics and their influence on the user's web page category choices, including purchases.

The researchers' methodology addresses an additional level of complexity that had not previously been examined in this context: the idea that consumers can switch from browsing to purchasing (or vice versa) during the course of a single visit to the website. They implemented a hidden Markov chain to take into account the unobserved goals of the user (i.e., whether or not the user came to the site with the goal of making a purchase).

To test their model, the researchers applied it to the data and compared their results to those obtained from other models, such as zero-order and first-order Markov models, latent-class models, and other types of multinomial probit models. In every case, models that took into account the sequence of events performed an order of magnitude better than those that did not. And the researchers' model outperformed them all.

Results

The path a consumer takes through a retail website reflects that particular consumer's goals. It's not enough to track the number of times a consumer views a particular page; the sequence of pages viewed also contains critical information. Using the information provided by their model, the researchers were able to calculate the probability that any given user would make a purchase.

Purchase conversion, or the percentage of a site's customers that complete a transaction, generally remains lower than 7 percent

for the focal online retailer. By using their model to observe customers' paths, the researchers were able to correctly predict a purchase up to 42 percent of the time. Retailers could use these predictions to dynamically tailor the site to the individual consumer, potentially increasing the purchase conversion rate from 7 percent to more than 9 percent.

Business Implications

The researchers determined that if Barnes & Noble's website changed dynamically depending on whether a customer's goal was predicted to be purchasing or just browsing, the retailer could increase purchase rates enough to substantially increase profitability. This is potentially applicable across all manner of online retail sites.

Path analysis has broad business implications well beyond online marketing. Examination of a path can be online or offline, and gathered from a wide range of tracking technologies such as security cameras, RFID, GPS systems in mobile phones, and other devices. When combined with marketing and customer purchase behavior, such as that gathered through loyalty cards, a much more complete picture of customer behavior and the effects of marketing emerges.

Currently there are myriad companies using targeted marketing— for example, Facebook and Google both take a user's browsing history into account, delivering messages that their algorithms predict that particular user would like to see. Incorporating path analysis marketing decisions such as these would increase the likelihood that the message reaches its target market and therefore potentially increases sales.

Alan L. Montgomery, Shibo Li, Kannan Srinivasan, and John C. Liechty, "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23 (4), pages 579–595, 2004.

Enhancing a Social Network through Gamification

Sean Yang
MS '14, majoring in management information systems

Interned for an enterprise software company's social network team

yuebyang@indiana.edu

Regular participation in an online professional community increases both personal knowledge and satisfaction and engagement with the community itself. One enterprise software company was interested in increasing user participation on its social network. This is where master's candidate Sean Yang stepped in. Sean interned as the technical lead on the project team that introduced gamification, which uses game-related thinking to make non-gaming situations more engaging.

Members of the company's social network can now complete missions and earn badges demonstrating their involvement in the community. From the simple act of logging in to giving a presentation at an industry event, members have myriad ways to build a good reputation in the community and to demonstrate their professional achievements.

How did gamification benefit your internship company?

The introduction of gamification to the company's online community significantly increased user engagement. The number of actions taken by our users has more than doubled since gamification was implemented, and we were able to drive appropriate community behaviors by setting up missions that encouraged these particular behaviors. Given that the online community is a critical platform where customers both learn about our products and share their own knowledge and experience, the increased activity and improved experience is likely to help the company retain and enlarge its customer base.

What were the steps taken to complete the project?

I joined the project team in the development and implementation stages. We followed an agile methodology of software development, going through several rounds of testing concurrent with fixing bugs and introducing features. Once the team had addressed all critical issues, we migrated user data from the legacy system to the new, gamified system. We had a silent "go-live" for a week, during which the gamified user interface wasn't enabled, but the backend was attached to the production system. This allowed us to ensure everything worked as expected without disrupting regular user activities. Once the system was deemed successful, all we had to do was push the button to reveal the new gamification elements to our users.

How can data from gamification be used for analytics?

First, a gamified system tracks real-time user interaction, providing a comprehensive data set for performance evaluation. Second, gamification data can be used to analyze and moderate users' actions, keeping the game fair for all users. Finally, gamification data can be used to improve the gamification itself—we can redesign the game to encourage more desired behavior.

What tools can be used for reporting in this type of application?

Given the volume of data we were getting, we started with Excel. Over time, we turned to data visualization software and then to MySQL to handle the increase in data.

How did your courses at Kelley prepare you for this project?

The MSIS core classes informed many aspects of my experience with the project. Project management classes helped me to coordinate team resources, and courses in the software development lifecycle allowed me to communicate more effectively with developers and engineers.

What did you learn from this experience?

I became a superuser of a leading gamification platform. I learned a couple of analytical tools from scratch and put them into practical use. I experienced how different teams work together to complete a project, and how a good decision-making process allows individual perspectives and concerns to be addressed, resulting in a robust project launch.

Data Analysis for Effective Fundraising

Miguel Becerra

MBA '14, majoring in business analytics and marketing

Interned at the fundraising organization for a public university

mibecerr@indiana.edu

As state appropriations for higher education decrease, fundraising grows more important for state universities. Private gifts are key in keeping high-quality education affordable. To be most effective, fundraising organizations must be able to efficiently collect and analyze data about giving.

This is where intern Miguel Becerra stepped in for the fundraising organization for a state university. They wanted to analyze the level of engagement of the university alumni population, because studies indicate that highly engaged alumni give more and more often. Although the organization had collected copious data, they had no metric for analyzing them. Miguel clarified the problem, deciphered the data, and developed a methodology for analysis. He then used data visualization to present complex results in an easy-to-understand chart.

What were the steps taken to complete the project?

Identifying the right question—"How can we define alumni engagement, and which factors should we evaluate for determining if alumni are engaged or not?"—was essential to designing the analysis strategy. To find an answer, I surveyed the organization staff to understand their criteria for assessing alumni engagement. This helped me understand that the organization did not have a standard way to measure or classify the data they wanted to analyze.

After defining a methodology for measuring engagement data, which resided in several databases, the next step was to prepare it for analysis. This involved cleaning the data set and finding additional sources to complete it. Current analytics tools can facilitate the process of integrating data from different sources, so the main challenge was identifying and keeping track of those sources.

When the data were prepared, I analyzed them and classified alumni by levels of engagement. I then created a visualization highlighting the main U.S. cities with a significant number of highly or moderately engaged alumni. This information allowed the organization to determine where alumni engagement activities and fundraising campaigns would make the most difference.

It is important to note that the analytics tools I had available simplified the analysis stage by allowing me to test and compare models in a very short time.

How would you address challenges with numerically expressing business imperatives (metrics) in future projects?

The most challenging aspect of this project was that different departments had different metrics for analyzing the data according to their own needs. In order to overcome this obstacle in future projects, I would work from the beginning with the users of those metrics to understand their approaches and to develop a standard methodology for analysis. I would also work with them to define the best way to visualize the data—a good representation of the data may increase adoption of the metric across the board.

How would you address challenges with data in future projects?

I would focus first on having a good understanding of the problem, then on evaluating whether I have the required data resources. I also think it is very important to help stakeholders set realistic expectations about the results the analysis will provide.

How did your courses at Kelley prepare you for this project?

The courses offered in the business analytics major allowed me to get hands-on experience with analytics tools that are widely used in the corporate world. Additionally, I learned how to apply statistical and visual analysis techniques to business problems. The problem-solving skills I learned at Kelley were also very valuable for clarifying the business problem and finding the correct question to answer.

What did you learn from this experience?

A data analytics project involves much more than just working with data—without knowing the right question to answer, analysis is difficult. Although organizations may think that having large amounts of data is sufficient, without a good collection strategy the data they have may be useless. But unstructured data can still be extremely valuable for gathering insights.

Finding the Most Profitable Alternative in Remanufacturing

Gilvan "Gil" C. Souza

Associate Professor of Operations Management

gsouza@indiana.edu

> The most important thing is the difference between the highest cost to manufacture the worst quality return and [the cost to remanufacture] the best quality return. How the cost changes according to the quality of the returned product is not that important. This is a significant insight, because it s a difficult thing to obtain in practice.

As equipment leasing becomes more prevalent, companies are increasingly turning to remanufacturing returned items. When functional products are returned at the end of a lease term, it makes sense for a company to recover value from those items rather than scrapping them wholesale. For original equipment manufacturers (OEMs), refurbishing and reselling those items makes good sense—not just from a financial perspective but also from a "green" perspective. However, remanufacturing items is a complex process that involves acquiring the used products, determining whether they're more suitable for remanufacture or salvage and distributing them accordingly, remanufacturing products, and remarketing them.

In this study, the researchers examine how grading the quality of returned items can help companies determine the best course of action for remanufacturing operations. Is it worth it to add an additional level of complexity to the steps that must already be taken in the remanufacturing process?

Statement of Problem

A company that remanufactures returned products must have a way to determine the most profitable alternative for each product. When a customer returns a product at the end of a lease, the company must decide whether it should be remanufactured or if it must be salvaged (i.e., dismantled for parts or recycled for materials recovery). Many companies currently remanufacture on demand—in other words, returned products are held in inventory indefinitely until demand arises for that product to be remanufactured. For a large portion of these companies, the volume of returns is higher than the demand for remanufactured items, so they must pay overhead to store these items for a long time.

It is in a company's best interest to use a formal capacity planning system; otherwise, the company may be spending too much on storage. The researchers developed an analytical model that takes into account demand and returns forecasts and the quality level of returned products, and uses those

numbers to efficiently determine which products should be remanufactured and which should be salvaged.

Data Sources Used

The researchers gathered data from several OEMs including Bosch, Cisco, and IBM. The primary source for developing their model was Pitney-Bowes, which provides hardware, software, and services for mail management. Because the devices they offer are primarily mechanical (e.g., postage meters, scales, and sorters), Pitney-Bowes was considered to be generally representative of many remanufacturing companies. Pitney-Bowes managers provided data for approximately 10 individual products, including monthly demand forecasts and observations about costs to remanufacture the different items. Once the model was developed, the researchers generalized it further using data gathered from the other OEMs.

Analytic Techniques

The goal of this research was to determine how having a precise number of quality grades affects a company's remanufacturing operations, and what the optimal number of grades is. The researchers used three techniques in developing their analytical model. First, they used *dynamic programming*, an optimization method that takes into account the future in order to make decisions about the present. It examines questions such as what the impact is if the company remanufactures *x* number of items in the current period when future demand is uncertain. The researchers used this method primarily to create an algorithm for finding the optimal solution when there are few capacity constraints in the remanufacturing shop. This algorithm states that a company should start from the highest quality grade and remanufacture until certain properties are met in order to maximize profit.

The researchers next looked at *linear programming*, which is much the same as dynamic programming but takes the demand forecast and the returns forecast as givens and considers capacity constraints. The linear programming model then solves for the best production plan that meets all capacity constraints.

The researchers used their algorithm and linear programming to solve the problem for 245,000 different instances, where each instance was a combination of parameter values possible in industry. This created a dataset of problems and optimal solutions. After determining a company's optimal profit in an instance, the researchers compared the results to a case where the firm doesn't grade products into different qualities. The incremental profit that arises from quality grading is the value of grading. Regression analysis was then used to summarize the relationship between the value of grading and the independent variables of cost, demand, and returns.

Results

The researchers' model provides an optimal quality grading policy for cases where returns occur on a known schedule and demand

rates are also known with reasonable certainty. A company can use this model to compute the absolute value of the grading process and compare it to the fundamental costs of implementing a quality grading process.

The true benefit of a grading system arises from the lower cost of remanufacturing the higher-quality returns. The researchers determined that a company with a policy of grading returns derives approximately 4 percent better results than a company with no such grading policy, on average. Additionally, they calculated that the most benefits are realized with five quality categories—the law of diminishing returns takes effect past that point, and benefits decrease as the categories become more complex.

Benefits of a grading system increase when returns are high relative to demand. In this case, a company derives more value from remanufacturing only the best returns and salvaging all others rather than paying to store those lower quality items. If the number of returns is similar to the amount of demand, however, having a grading system in place yields very little benefit. Quality grading is virtually useless when a company must remanufacture every returned product to meet demand. Additionally, when salvage values are high, quality grading is unimportant because any given item has a relatively high value regardless of whether a company remanufactures or salvages it.

Business Implications

As a result of this study, Pitney-Bowes has established a quality grading system in line with the optimal system proposed by the researchers. This system is estimated to save them $5 million per year in inventory costs.

The researchers' model is relevant for any remanufacturer who leases products or is otherwise able to predict returns and demand for remanufactured products. A quality grading system increases complexity of operations, especially in terms of establishing rigorous guidelines for sorting items into various quality categories. Additionally, instead of maintaining one stockpile of items, a company must be able to maintain a stockpile for each category of item. While there is a slight learning curve, any company implementing a quality grading system could likely recoup startup costs within a few months.

Although this model was developed for companies that could forecast demand for returns and remanufactured items with reasonable accuracy, the model can be extended to provide guidance to other companies. While doing so is heuristic and cannot be guaranteed to provide an optimal solution, it can provide insight into making better decisions about remanufacturing.

Mark Ferguson, V. Daniel Guide Jr., Eylem Koca, and Gilvan C. Souza, "The Value of Quality Grading in Remanufacturing," *Production and Operations Management*, 18 (3), pages 300–314, 2009.

Building a Metric for Efficiency

Christopher Stafford
MBA '14, majoring in supply chain and business analytics

Interned at an international building products company

ctstaffo@indiana.edu

Transportation experts excel at using visual maps to judge the efficiency of even the most complex freight shipments. But when sending out more than 1,700 trucks per week, a large building products company needed a faster, more practical way to identify shipping inefficiencies.

When Christopher Stafford joined this company as a summer intern in 2013, he focused on defining what constitutes an "efficient" shipment, assessing the ability of the company's current metric to measure efficiency, developing new metrics or modifying the existing one, and designing and implementing a sustainable reporting system.

Christopher's work is already paying off for the company. His automated reporting system immediately began saving them several hours of tedious work each week. And with the PivotTable he built, transportation managers at the company have more control over their data than ever before.

What steps were taken to complete the project?

I concluded that there are two main categories of efficiency the company should track: distance efficiency, or how straight the route is; and load efficiency, or how the truck is used along the route. The company's existing metric accurately measured load efficiency, but didn't account for distance efficiency—"efficient" routes often doubled back or zigzagged.

To calculate distance efficiency, I used historical data to measure the distance from each plant to each destination (thousands of combinations). With these data I could easily calculate the mileage increase that resulted from additional stops.

Finally, I worked with the company's transportation lead analyst to design and implement an automated reporting system that would collect each week's freight data, calculate the efficiencies using the new

metrics, and more accurately find inefficient shipments. Managers have used this information to track and fix patterns of inefficiency.

How would you address challenges with numerically expressing business imperatives (metrics) in future projects?

First, it's important to know the goal of the metric. Business imperatives are frequently vague—for example does "efficiency" mean load, distance, or cost efficiency, or something else entirely? Also, metrics must be meaningful to the people who will be affected by them. By putting the metric in simple terms—how many extra miles did the truck drive out of its way?—it's easy for shipment schedulers and planners to understand.

How would you address challenges in applying analytic techniques in future projects?

Understanding the data set is vital. If the data are nebulous or include information irrelevant to the metric at hand, including those data may corrupt the results. Knowing the implications of those data up front will prevent potential missteps in developing a metric.

How did your courses at Kelley prepare you for this project?

My courses in data mining, supply chain, and cost accounting were all crucial to this project. The most beneficial course, however, was Wayne Winston's "Decision Analysis through Spreadsheet Modeling." In this course, I learned how to use Excel to conquer business problems—an area where Kelley students are unmatched.

What did you learn from this experience?

I learned how to gain insights by working backward through multiple huge data sets. I also learned the value of building personal credibility. Veteran employees were initially reluctant to change their metrics and key performance targets, which determined a significant portion of their salary. In the end, however, my results won them over.

Building Supply Chain Competency in Analytics-Driven Organizations

Oliver Page

Oliver Page is a senior manager in Deloitte Consulting LLP's Technology Strategy practice, concentrating for the last ten years on the consumer products and hospitality industries. In early 2013, he completed a major initiative for the Grocery Manufactures Association (GMA) focused on how to capitalize on innovation by harnessing big data and analytics to drive top- and bottom-line growth.

opage@deloitte.com

Today's supply chains are not only global and more complex, but also face a stiffer and more demanding regulatory environment and highly volatile commodity prices. However, manufacturers also have access to more information and improved technology to enable decision making. These pressures, information, and technology together are driving transformation of a manufacturer's analytical capability. Improved use of analytics can enhance supply chain and manufacturing operations performance by enabling increased quality, lowered costs, improved delivery rates, and reduced risk.

While POS and inventory data sharing is still a primary focus, a few leading manufacturers are augmenting internal data with new external data assets that will provide better insights into what is actually happening on retailers' shelves. Emerging technology and digital innovations help provide a granular, real-time, and end-to-end view of the supply chain. Manufacturers have the opportunity to leverage these data assets to generate insights that drive value across the supply chain.

Emerging Data Assets

Emerging technology and digitization affect all aspects of the supply chain: risk, plan, source, make, and deliver. For example, manufacturers are increasingly turning to the "crowd" to gather real-time data and perform micro tasks at retail locations. Innovators in this space are developing tools that securely connect manufacturers to trusted eyes, ears, and arms at specific locations around the world. In exchange for a small payment, manufacturers are able to gather nearly real-time in-store visibility to improve retail compliance and execution.

In addition to tapping into the crowd, supply chain visibility is being transformed by tiny economical sensors and "smart" packaging. This emerging technology that combines condition monitoring into sensors, RFID, or even directly into packaging can enable real-

time visibility into the supply chain. Leading manufacturers are optimizing distribution, reducing spoilage, and better managing inventory and operational costs using these technologies.

The supply chain is ripe for emerging technology innovation. This field is evolving rapidly and leading organizations have developed a strategy to capitalize on emerging innovation (e.g., tapping into the crowd, using environmental sensors and open collaboration platforms, social listening) and how data from these innovations affect their business.

Building the Right Environment

Traditional and emerging capabilities in analytics create tremendous opportunities. To take full advantage, however, manufacturers will need to evaluate whether traditional approaches to delivery will meet tomorrow's needs. Manufacturers should consider how they organize their analytical talent to optimize this critical workforce segment. Traditionally, analysts have been embedded into individual departments. The shortcoming of this model is that it creates little incentive to find cross-functional value that would result in an integrated information management capability. Organizations would need to determine whether talent should remain functionally embedded, become more consultative, or be centralized to best address the needs of the organization.

Regardless of the model, analytics should require a leader with the business acumen and the seniority to be an "analytics champion"— to inject analytics into company strategies and critical decisions. A key role for this leader will also be to drive cultural change—one of the hardest elements to transform—within the organization. Without a focus on changing culture and using analytics to change process, new analytical capabilities won't drive true business and operational change.

Technology

Most manufacturers have spent years, and millions of dollars, developing broad information architecture that does not effectively answer the types of questions senior leaders are now asking. While massive investments in enterprise data warehouses and enterprise resource planning systems have enabled precision, the systems often lack the agility to answer questions such as, "What customer and SKU combinations will be the most profitable; then how will regional or local economic conditions affect those predictions?" Developing predictions that take into account internal, customer, and external information is challenging the use of traditional technology environments.

Leading manufacturers are enabling internal, external, and cross-functional capability through more flexible architectures. This helps integrate data assets more rapidly so analysts spend more time on analysis versus data manipulation in offline repositories. And finally, it is often the last mile in which the analyst communicates the insight where the chain breaks. Transforming the user experience through analytical visualizations enables organizations to rapidly derive and communicate new insights from both existing and emerging information sources. This has proved to be a "quick-win" opportunity for many manufacturers.

Start Small, Prove Value, Scale Fast

Enterprise analytics is in the spotlight. Organizations are no longer questioning the importance of fact-based decisions; rather, they are asking where to start and how to enhance their existing capabilities.

Improving an organization's enterprise analytical capability is a journey. Successful analytical transformation balances capability building with the delivery of near-term "wins" to build momentum and excitement for the program. One approach that has been successful for many leading organizations is to think about the journey in three steps: Focusing on illustrating capability and demonstrating value, evolving those capabilities and scaling to the needs of the internal customers, and embedding more advanced analytical capabilities and transforming "business as usual."

Building supply chain and enterprise analytics competency requires sustained leadership commitment. Value can be delivered quickly; however, longer-term capabilities must be thoughtfully developed over time to deliver against the promise of analytics.

IBA Affiliated Faculty

Frank Acito
Professor of Marketing; Max Barney Faculty Fellow; Co-Director, Institute for Business Analytics

Herman Aguinis
Professor of Organizational Behavior & Human Resources; John F. Mee Chair of Management; Director, Institute for Global Organizational Effectiveness

Goker Aydin
Associate Professor of Operations & Decision Technologies; ArcelorMittal USA Undergraduate Faculty Fellow
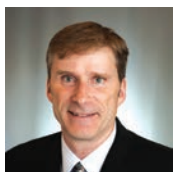
Hillol Bala
Assistant Professor of Information Systems; Whirlpool Corporation Faculty Fellow

J. Doug Blocher
Chairperson and Associate Professor of Operations & Decision Technologies; Arthur M. Weimer Faculty Fellow
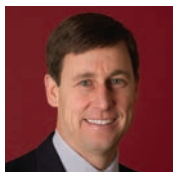
Kurt M. Bretthauer
Professor of Operations & Decision Technologies; W. Michael & William D. Wells Life Sciences Faculty Fellow

Raymond R. Burke
Professor of Marketing; E. W. Kelley Chair of Business Administration; Director, Customer Interface Laboratory

Kyle Cattani
Associate Professor of Operations Management; W. W. Grainger, Inc. Faculty Fellow

Sandeep Chandukala
Assistant Professor of Marketing

Hai Che
Associate Professor of Marketing

Jonathan E. Helm
Assistant Professor of Operations & Decision Technologies

Randy Heron
Professor of Finance; Roger & Barbara Schmenner Faculty Fellow

F. Robert Jacobs
Professor of Operations Management; Chase Faculty Fellow; Honors I-Core Faculty Lead

Vijay Khatri
Associate Professor of Information Systems; Arthur M. Weimer Faculty Fellow; Co-Director, Institute for Business Analytics

Shibo Li
Associate Professor of Marketing; Arthur M. Weimer Faculty Fellow

Haizhen Lin
Assistant Professor of Business Economics

Girish Mallapragada
Assistant Professor of Marketing

Anne P. Massey
Professor of Information Systems; Dean's Research Professor; Chairperson, Doctoral Programs; Associate Vice President for University Academic Planning and Policy

Alex Mills
Assistant Professor of Operations &
Decision Technologies

Alfonso J. Pedraza-Martinez
Assistant Professor of Operations &
Decision Technologies

Philip T. Powell
Clinical Associate Professor of Business Economics
and Public Policy; Chairperson of Kelley Direct

Jeff Prince
Associate Professor of Business Economics; Cathie
and Jerry Anderson Faculty Fellow

Babur De los Santos
Assistant Professor of Business Economics

Kim Saxton
Clinical Associate Professor of Marketing

Scott B. Smart
Clinical Associate Professor of Finance;
Whirlpool Finance Faculty Fellow; Associate
Chair of the Full-Time MBA Program; Director of
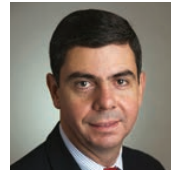Strategic Finance Academy

Daniel C. Smith
Professor of Marketing; Former Dean of the Kelley
School of Business; President and CEO,
IU Foundation

Ash Soni
Professor of Operations & Decision Technologies;
ArcelorMittal Faculty Fellow; Executive Associate
Dean for Academic Programs;

Gilvan "Gil" C. Souza
Associate Professor of Operations Management

Noah Stoffman
Assistant Professor of Finance

Ramesh Venkataraman
Associate Professor of Information Systems;
Lawrence D. Glaubinger Professor of
Business Administration; Chair of the
Undergraduate Program

Munirpallam Venkataramanan
Professor of Decision Sciences; Jack R. Wentworth
Professor; Vice Provost for Strategic Initiatives

James M. Wahlen
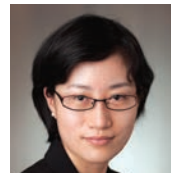Professor of Accounting;
James R. Hodge Chair of Excellence

Rockney G. Walters
Professor of Marketing

Matthijs R. Wildenbeest
Associate Professor of Business Economics

Lucy Yan
Assistant Professor of Operations &
Decision Technologies

Jingjing Zhang
Assistant Professor of Information Systems;
3M Faculty Fellow

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Indiana University Kelley School of Business
1275 E. Tenth Street
Bloomington, IN 47405

IBA
Institute for Business Analytics

About Us

The Kelley Institute for Business Analytics uses the resources of the prestigious Kelley School of Business to produce insightful research and train professionals who can excel in this exciting new field.

What Is Business Analytics?

Simply put, it's using data to make better business decisions. And it's becoming big business.

For years, companies have collected data about their practices and consumers. Now, thanks to inexpensive computing, more companies are putting their data to work—using techniques such as predictive analytics, optimization, and simulation to make fact-based decisions that improve productivity, increase profits, and create a competitive advantage.

Kelley: Leading the Way

To make the most of business analytics, companies need innovative ideas and well-trained professionals. That's where Kelley comes in.

One of just a few business analytics programs nationwide, Kelley's IBA supports:

- An academic program that prepares students to solve business problems using analytics
- Corporate partnerships that shape Kelley's understanding of analytics and help companies tap into Kelley's talent
- Cross-disciplinary research by Kelley's expert faculty
- Seminars, conferences, a speaker series, and case competitions that bring together professionals, faculty, and students

Contact Us

http://kelley.iu.edu/iba
kiba@indiana.edu
Linkedin: http://tinyurl.com/linkedin-kiba
YouTube: http://tinyurl.com/youtube-kiba