KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

OnAnalytics

Fall 2016

diagnose. predict. optimize.

FIVE YEARS
OF DISCOVERING
THE CONNECTIONS

Converting Data to Insights

Welcome

As the Founding Corporate Partner of the Kelley School of Business Institute for Business Analytics (IBA), I am thrilled to open this issue of *OnAnalytics* magazine. Inside, you'll hear from two of my colleagues at Deloitte Consulting LLP: Greg Szwartz, who shows how analysis of complaints data is being used to improve quality and safety; and Lorraine Cohen, who writes about using analytics to better manage business travel-related tax concerns.

Looming large over any discussion of business analytics today, of course, is the topic of cognitive computing—self-learning systems that employ data mining, pattern recognition, and natural language processing to simulate human thought processes in a computerized model. Because of all the buzz about cognitive, your instincts probably tell you to ignore it as much as possible—to let everything cool down a bit so that we can recognize it for what it really is.

There's just one problem with that approach.

Cognitive computing is already huge, and it's expected to only get bigger. I don't say that lightly. I've seen plenty of next-big-thing flameouts, just like anyone who's been engaged in the analytics world for years. But when it comes to cognitive, in my work with businesses around the world today, I'm seeing signals that remind me not of the flameouts but of the monumental advances made in technology. Remember that moment when we all realized that there were as many mobile devices in the world as there were people? Or, going way back, when it became clear that the Internet was not solely the domain of government researchers, and had serious commercial applications? That's the kind of moment we're experiencing right now with cognitive.

Think about it. Computing capabilities are unbelievably strong today. There's a greater discipline in algorithms than we've ever seen. Then there's the constantly decreasing cost of data storage—it costs around three cents to store a gigabyte of data today, which was inconceivable not so long ago. Put it all together, and you realize that whatever we've done in cognitive computing today will soon be considered quaint early indicators of the seismic changes that follow. We are heading down an exponential change curve.

Cognitive computing is one of several key topics that we'll explore together at the Kelley Forum on Health Care Analytics and the Kelley Forum on Finance Analytics, planned for September 16 and October 28 respectively. I hope to see you there—this is one more way IBA is working to bring together students, faculty, and practitioners with a shared interest in business analytics concepts.

Sincerely,

Paul Roma
Principal, Deloitte Consulting LLP

From the Editors

Celebrating a Milestone

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

With Kelley's Institute for Business Analytics (IBA) completing five years, this issue of OnAnalytics commemorates an important landmark. OnAnalytics now reflects the perspective of our 50+ faculty, 350+ undergraduate, graduate, and online students, and 30+ corporate partners around the theme of analytics.

Kelley faculty contributed four articles to the current issue. Ruomeng Cui, assistant professor of operations and decision technologies, reports on the results of tests on the accuracy of crowd-sourced forecasts. Such forecasts were shown to be accurate, viable tools, especially where historical data is not available or is of limited relevance.

By paying for specific keywords used in online searches, advertisers can generate traffic on their sites and potentially increase sales. Alice Li, assistant professor of marketing, provides recommendations for allocating keyword advertising budgets.

Evaluating the economic benefits of investments in renewable energy involves considerations of multiple factors. Gilvan Souza, professor of operations and decision technologies, examines the effects of the granularity of measurements on the economic feasibility of investments in renewable sources.

Brian Miller, assistant professor of accounting, conducted a study of the complexity, length, and readability of financial reports. Among the findings from the study were that complexity affected total trading activity and that large investors had a greater degree of consensus in trades than small investors.

Three Kelley MBA students report on internship experiences involving applications of analytics. Vikram Swamy worked with a company that provides customer leads to automobile dealerships. By analyzing effectiveness of the leads, he was able to make strategic recommendations. Swetha Prasad analyzed data from a technology firm's customer relationship management tool to better understand and align the relationship between customers and the company's development team. Matthew Davie used a large e-commerce firm's database to develop and recommend improvements in the after-sales experiences of customers. Rahul Singh, an MS in Information Systems student, reports on his experience optimizing a client's portfolio of keywords being purchased from online search engine firms.

This issue includes articles from the 2015 Analytics Award Competition. The winning entry from Humana used data to predict seniors' likelihood of falling. Their model was successfully used to identify high-risk individuals, and a subsequent outreach program effectively reduced falls. Runner-up C. H. Robinson developed a model to more efficiently track employee workloads to optimize staffing. This model was instrumental in increasing operating margins for the firm.

Lorraine Cohen, a partner at Deloitte Tax LLP, reports on an analytics-based system to track employee travel patterns. This system is used by Deloitte to assess the tax implications of interstate and international work assignments. Finally, a team from Deloitte Consulting LLP— Greg Szwartz, Steve Ellis, and Geetanjali Chakraborty—has developed a platform that analyzes patient complaints data in a life sciences company, finding problems earlier and fixing issues before they become problems.

As the field of business analytics matures, there has been an increase in industry-specific applications. To take advantage of this trend, the Institute for Business Analytics has revised the structure of its annual forums. In place of the general forums, two special-issue forums—one on digital marketing and the other on sustainability of resource usage and efficiency—were held in fall 2015. Summaries of each are included in this issue. The program of industry-specific forums will continue for 2016, with events planned on analytics in health care and finance.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

OnAnalytics

diagnose. predict. optimize.

How Investors React to Complex Reports

Brian P. Miller
Associate Professor of Accounting; Arthur M. Weimer Faculty Fellow

bpm@indiana.edu

Efforts to make financial data more easily consumable date back more than 80 years to the Securities Act of 1933, commonly known as the truth in securities law. Although a 1998 Securities and Exchange Commission (SEC) rule set forth requirements for the use of plain English in particular sections of financial reports, complexity of these documents has continued to increase to the point where certain investors may not be fully capable of processing them.

In this paper, the author examines the effect of more complex (longer, less readable) filings on trading behaviors of small and large investor classes.

Statement of Problem

Despite regulation and ongoing concerns about the length of financial reports, firms have continued to increase the amount of information they disclose. Whether this information is beneficial to investors is unclear. Some accounting practitioners argue that useful disclosures are now hidden among boilerplate, redundant, immaterial, or irrelevant data. As a result, the cost to interpret these reports may be too high for some small investors and so they will not initiate trades in response to a report filing.

Although previous research has provided insights into the effect of report length on market responses, there has not before been a study of the effect on small and large investors' specific trading behavior relative to the complexity of the filings. The author therefore examines two hypotheses:

- Abnormal trading volume during the time period around a filing is lower when reports are more complex, and the effects of complexity on abnormal trading are most obvious among small investors.

- Trading consensus within a class of investors is lower when reports are more complex.

Data Sources Used

The author examined a set of 10-K reports filed between 1995 and 2006, downloaded from the EDGAR company filings database maintained by the SEC. To compare across classes of investors, the author retained only firms that had both small and large trade activity. After discarding observations that met certain other criteria (e.g., stock price less than a dollar, prior year filing date unavailable), the sample of observations available for primary analysis consisted of 3,809 unique firms with 12,771 firm year observations. This reduced to a sample set of 4,724 observations for which data were available to calculate the total number of words plus the number of table cells in each document.

"Increased availability of information is of no benefit to investors if reports are too long and complex to process."

Analytic Techniques

This study focused on the effects of report length and readability, examined separately and together. In addition to considering the word length of the document, the author uses two distinct measures of readability. The first measure, the Fog Index, is a widely accepted readability formula developed in 1952. The second measure uses output from a proprietary software application called "StyleWriter–Plain English Editor" to develop a unique, multidimensional measure of readability based on the factors set forth in the SEC's plain English guidelines.

Prior evidence indicates that the most visible response to public disclosures is trading volume occurring around a 10-K filing. The author focused on abnormal trading activity during this time period to determine investors' responses to variations in complexity. He examined the sample set of reports in search of evidence showing that more complex reports would negatively affect trade behavior during days around the report filing.

The author performed the research in stages. First, he examined the impact of reporting complexity on the aggregate trading behavior of small, medium, and large investors, as well as specific trading behavior of small and large investors. Next, he examined the effect of reporting complexity on trading consensus within each investor class. Finally, he examined whether the association between complexity and lower abnormal trading is driven by variations in report complexity over time and/or variations in firms' disclosure attributes.

Results

The author's overall results show evidence that more complex reports cause a decrease in total trading activity. The author found that this decrease appears to be associated with both variations in the readability of the disclosures in a given year as well as variations in the report length over time. The reduction in total trade activity due to reporting complexity appears to be driven by small investors making fewer trades. In fact, the author found that increases in report complexity over time significantly affect small investors, but have only a limited impact on large investors. The evidence is consistent with it being too costly for some small investors to process more complex filings.

More complex reports also appear to lead to decreased consensus among the small investors who do elect to trade around a 10-K filing. Conversely, there is some evidence that large investors are more likely to have increased consensus when a report is less readable and more data are provided. This is likely due to the fact that large investors are better equipped than small investors to interpret more complex filings.

Additional analysis reveals that the effects of readability and length appear to act as substitutes, with length of the reports appearing to overshadow the readability of the reports when both measures of complexity are analyzed simultaneously. This evidence suggests that length of the document is potentially a greater concern for small investors.

Business Implications

The results of this study indicate that regulations requiring more disclosure may not level the playing field between small and large investors as previously thought. Additionally, more disclosure does not necessarily aid investors in their trading decisions. Although SEC regulation has increased small investors' ability to access financial information, this increased availability is of no benefit if reports are too long and complex to process. Improvements in readability, however, can lead to an increase in small investor trading activity.

The results highlight the importance of considering the costs of processing information in reports when developing further regulation. As processing costs increase, small investors will tend to delay trading around filings, either to have the time to process the reports or to trade instead around events where easier-to-access information is made available. As such, introducing initiatives that reduce processing costs may serve to ensure that information is quickly impounded into market prices.

Brian P. Miller, "The Effects of Reporting Complexity on Small and Large Investor Trading," *The Accounting Review,* 85 (6), pages 2107–2143, 2010.

Managing Business Travelers Leveraging Analytics

Lorraine Cohen

Lorraine Cohen is a partner at Deloitte Tax LLP with over 25 years of experience in expatriate tax program management consulting, tax planning, and human resource policy consulting.

If you travel to other states or countries for business, your employer is required to pay income taxes in the locales where you're performing your work. And even if your employer fails to withhold and pay taxes properly, you are required to do so. Taxing authorities around the globe are becoming increasingly vigilant and implementing tools that help them track compliance. It's in everyone's best interest for your company to confirm that it is withholding and remitting taxes appropriately, because failing to do so could result in hefty fines and penalties.

But what happens when a company's payroll team doesn't know where the employees are actually working? How can the team determine where to allocate an employee's time or detect when a given employee is about to cross a taxable threshold? For a large international company with thousands of employees, many of whom travel to various locations at various times, keeping track of everyone can be a logistical nightmare. This is where workforce analytics can help a company mitigate risks, comply with tax laws, and manage its overall costs.

Deloitte Tax LLP has developed an innovative approach that leverages analytics to manage tax concerns with respect to business travelers. Deloitte uses data analytics to create dynamic reports that:

- Identify travel patterns and risk assessment profiles by employee, business unit, position, level, country, region, and other company-specific data

- Identify jurisdictions in which employees have triggered tax filing obligations

- Identify individuals who are approaching tax thresholds in order to address potential filing issues

- Illustrate income allocation to enable payroll reporting and appropriate tax remittance

- Provide actionable information that allows a company to manage risk

Traditionally, employees have generally been required to self-report their work locations. This task increases administrative overhead and takes employee time away from other critical business needs. And in some cases, employees completely fail to report, resulting in incomplete data. Rather than relying on self-reporting, Deloitte examines and cross-checks volumes of data including travel records, expense reports, security badge scans, virtual private network logins, corporate jet records, and other sources to accurately determine where employees are traveling and how long they've been there. Once this data is extracted and transformed, it is analyzed for employee taxability in 78 countries and all 50 states. The resulting information can then be tailored according to the company's business rules, visualized in Qlikview, and imported into the company's payroll system to facilitate required withholding and reporting.

For a single company, Deloitte analyzed a sample year that included 80,000 travelers and discovered that more than 10,000 of them were above a threshold set by the company. Tracking these employees and implementing alternatives to business travel in advance of a liability being triggered, such as videoconferencing, could save the company a significant amount of money.

Whether a company's goal is to comply with tax laws, to prevent employees from triggering taxation, to keep track of where employees are, or some combination of the three depends upon the individual company. But no matter the goal, having the ability to track employee travel will benefit the company. Not only can Deloitte's approach help companies manage withholding and reporting, but it can also help them obtain information needed to negotiate better rates for travel services, locate their employees in the event of an emergency such as a natural disaster, and apply for refunds for value-added and other taxes that may not apply to traveling employees. Deloitte's approach enhances a company's ability to make informed decisions, institute policies to control costs, increase employee satisfaction, and manage risks.

Analyzing Profit, Loss, and Strategic Workflow

2015 KELLEY ANALYTICS LEADERSHIP AWARD RUNNER-UP

Mike Duffey

Mike Duffey is a business analyst with a fine arts background. He extracts narrative from information to better explain the underlying story.

duffmik@chrobinson.com

The runner-up for the 2015 analytics leadership award was a tool called "The Wiz Reporting Suite 1.0: Activity-Based P&L and Strategic Workflow," developed by third-party logistics provider C. H. Robinson and led by business analyst Mike Duffey. The Wiz provides oversight and management through an activity-based approach to tracking employee work and overall office and team structure. It was created to help management better understand personnel costs and account profitability in an environment where several hundred employees could work on any of thousands of customer accounts during a single month.

The Wiz offers guidance in maximizing both staffing efficiency and account profitability through two modules. The employee workflow module shows what tasks each employee performs for which account at what time of day, allowing insight into the balance of work within teams. Before The Wiz, the company knew that employees had uneven daily workflows but had no data to prove it. Using The Wiz, they track a set of core tasks, applying a weight to each task based on the time it took to perform and allocating the work and the cost to perform that work across customer accounts based on what the employee was doing for a given customer. The company uses the resulting information to make staffing decisions and to structure groups, resulting in more efficient use of resources—and more employee satisfaction.

The account level profitability module rolls up monthly employee and commission costs at various levels and pulls in load volumes and revenues. Creating a single report with all of these data allows the company to analyze key efficiency and profitability metrics by account, sales team, branch, office, and more. The company is able to examine their pricing and determine whether or not they are turning a profit on a given customer. Together, the employee workflow and account level profitability modules allowed C. H. Robinson to increase operating margins and improve productivity in 2014.

Identifying Seniors at Risk for Falling

2015 KELLEY ANALYTICS LEADERSHIP AWARD WINNER

Harpreet Singh

Harpreet Singh, PhD, is manager of predictive modeling for Humana Inc.

hsingh3@humana.com

Vipin Gopal

Vipin Gopal, PhD, is enterprise vice president of clinical analytics for Humana Inc.

vgopal@humana.com

A project called "Predictive Model-Based Identification and Outreach for Seniors at High Risk of Falling" won the 2015 analytics leadership award. A team led by Vipin Gopal and Harpreet Singh at Humana Inc. developed this project.

In the United States, according to the Centers for Disease Control and Prevention, an adult age 65 or older dies every 20 minutes due to injuries sustained from a fall. Every year, one in three older adults falls. As one of the country's largest Medicare Advantage providers, Humana was looking for a way to identify and assist seniors at a high risk of falling. In 2012, they developed a predictive model to address this need.

To create the predictive model, the team identified more than 200 relevant predictors in Humana databases and determined the characteristics and circumstances of people who have fallen in the past. Using that information, rigorous predictive modeling techniques, and a wide variety of data sources including medical and pharmaceutical, the team developed a statistically sound and robust predictive model. The model was transitioned over to Humana's clinical organization for a proof-of-concept pilot study. This pilot showed that outreach was effective in leading participants to take steps to reduce their fall risk. Success of the pilot resulted in a large-scale deployment of the model by Humana to assist seniors.

A current initiative, Personal Emergency Response Service (PERS), uses the predictive model to identify high-risk seniors who would benefit from falls prevention efforts, including a wearable remote monitoring device. Participants in the program wear a device, around the neck or on the belt, that signals emergency response for immediate assistance in case a fall is detected: automatically by device sensor, or if the senior pushes a help button on the device. In case of fall detection, a 24 x 7 service talks to the member through the device and arranges for emergency medical services if needed.

Improving Quality and Safety through Complaints Data Analysis

Greg Szwartz
Greg Szwartz is practice lead for Life Sciences Advanced Analytics at Deloitte Consulting LLP.

gszwartz@deloitte.com

Steve Ellis
Steve Ellis is a manager within Advanced Analytics and Modeling at Deloitte Consulting LLP.

sellis@deloitte.com

Geetanjali Chakraborty
Geetanjali Chakraborty is a senior consultant within Advanced Analytics and Modeling at Deloitte Consulting LLP.

gechakraborty@deloitte.com

The medical device industry is highly regulated, with rigorous reporting and compliance requirements. Device companies are intent on keeping their products safe in order to manage expected returns from shareholders, while also managing reputational and strategic risk. To comply with regulations and support the business's drive for high-quality products requires sifting through massive volumes of complaint data to determine which complaints are worth investigating—which ones might be reportable to the FDA and/or which ones represent opportunities for safety and quality improvements. There are numerous examples of product complaints leading to quality improvements that benefit patients (consumers) as well as the bottom line. This article explores a solution designed to separate important signals from a large and deep "lake" of complaints.

By nature, complaints data are messy. At one end of the spectrum, a patient may complain about something that is not a safety issue (e.g., a broken shipping box), and at the other end are legitimate complaints involving potentially serious health and safety issues the company needs to know about. The analytical challenge lies in distinguishing between the two—separating the signal from the noise. This is made even more difficult by the variety and volume of complaints a large device company receives—from anywhere in the world, in any language, and from any customer or third party.

A team from Deloitte developed a platform that allows its client, a life sciences company, to analyze these complaints data. The solution is dubbed the "Signal Detection Management (SDM)" application. It assembles data from disparate sources into a unified relational database and applies statistical and visualization techniques to identify potential safety signals within the data. It also includes a sophisticated graphical user interface (GUI) designed specifically to be able to access, filter, and apply rigorous statistical tests on large complaints data

sets. The data asset, visualizations, advanced statistical tests, and GUI are all part of a GxP validated system—one that complies with quality guidelines required by the FDA in life sciences.

One potential signal that SDM might identify is a packaging complaint, which seems mundane but is in fact more nuanced. A dent in a box *could* be a serious issue, because it means the product inside is damaged. If that product will be going into a patient, this could become a matter of life or death. In other cases, a dent in a box is just a dented box. How does the company determine which is the true signal? If every dented box is a signal, the company must expend resources investigating the cause of every dent and not much else. SDM helps the company focus on investigating the *true* signals and finding the real problems rather than examining every box. Ultimately, the patients benefit, because true signals are identified and corrected before injuries occur. Speed is important in anything related to safety.

Data are captured for SDM through the company's complaints handling system, where complaints come from patients, doctors, hospitals, literature, and social media. The company spent millions streamlining its complaints handling process, but still needed a way to analyze the data. This is where SDM's advanced analytics come into play. The platform includes a standard analysis toolset that is run on every product each month, and the GUI allows an analyst to create alternate views of the data.

The statistical and advanced analytical components of SDM are impressive, but statistical models often overlook the importance of the human element—people who know the products and who understand how products operate and interact with a patient. It can be difficult to connect domain experts with the data, because they may not have the statistical knowledge to understand the analytical process. The SDM GUI allows these experts to interact with the models in a way that pulls their knowledge into the system naturally and easily. This in turn allows the analyst to continually tune the statistical models to filter out false signals, bringing true ones to the forefront more effectively than an automated statistical process.

When a true signal is detected, an analyst flags it and assigns the signal to an investigator using a workflow tracker. The data set is then frozen, the issue is reported to the FDA as necessary, and the company investigates until the case is closed. Finding problems earlier and faster means saving lives—fixing issues before they become true problems.

In some cases, signals are closed because they are known problems, are not significant quality or safety issues, or are caused by misuse of the product rather than by the product itself. In other cases, there are product deficiencies or manufacturing issues that need to be corrected. And in all cases, the company must determine whether an adverse event has occurred, because those must be reported to the FDA. Adverse events are, however, much less common than signals that are closed for other reasons.

SDM incorporates more than a dozen advanced data mining and statistical techniques including time series algorithms (e.g., EWMA, CUSUM, Shewhart charts), disproportionality analysis (e.g., PRR, Multi-item Gamma Poisson Shrinker), Poisson regression models, and rare event detection to analyze the entire product lifecycle and myriad product types. For mature products with known defect rates, the platform primarily uses time series algorithms. For products at the end of their lifecycle, disproportionality analysis is the method of choice. For products new to the market, EWMA allows analysts to develop a trend and cycle up to a steady state fairly quickly to determine if an issue is becoming a concern right away. SDM also contains two sample tests and other less sophisticated methods when data do not allow for deeper analysis.

The application was built using HTML5, JavaScript, Node.JS, Rabbit.MQ, R, Oracle, and Tableau and is distributed across four independent production servers comprising 64 CPUs with 32GB of RAM each.

Decreasing Bad Buying Experiences

Matthew Davie

MBA '16, majoring in finance and business analytics

Interned at an e-commerce company

mdavie@indiana.edu

An e-commerce company must ensure that a customer has a good buying experience if they hope to retain that customer. Part of the buying experience includes maintaining an effective search algorithm so customers can find the items they're looking for. When one large e-commerce company wanted to decrease its customers' bad buying experiences, they charged intern Matthew Davie with examining the relevant data and making recommendations.

What were the steps taken to complete the project?

First, I had to understand the e-commerce site's database and the business problem. I spent the first few weeks of the project reading internal documentation and speaking with stakeholders to understand the site's search algorithm, internal attitudes toward buyers, and what constituted a "bad buying experience."

After writing down what I felt were the important questions and the types of data I would need to answer those questions—as well as my own personal hypotheses—I extracted and cleaned data from the site's internal database. I then used the data to identify relevant patterns, as well as areas where the site can improve its after-sales buying experience by more aggressively differentiating between sellers.

Finally, my hypotheses—which seemed reasonable given past transaction data—were to be A/B tested on the site in the following months.

What did you find out about analytics roles in organizations?

I came to realize that many analytics professionals act as internal consultants within organizations. You must be able to synthesize and understand techniques used by data scientists to come up with business solutions. You must then be able to present those solutions to a business audience in a way that convinces them to pursue your ideas.

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

One of my biggest issues was the vast data set I was working with, which led to initial confusion on my part. While I felt I did a reasonable job attempting to first understand the data, in the future I would make it a priority to speak with the company's data experts.

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

S523, a data visualization course, gave me a better idea of how to present my data and also helped me refine my SQL skills. The statistical techniques employed in K513, a data mining course, will serve me well working in analytics-related roles in the future.

What did you learn from this experience?

The earliest steps in the process are the most important. In our rush to a solution, we often overlook the need to understand the business and have a clean data set—particularly in large, data-driven organizations. Focusing on the early stages will make the actual analysis less taxing.

What's the best way to prepare for this type of role?

Learn as many statistical analysis techniques as possible. Learn how to use different software packages to run your analysis—many organizations typically use SAS, but newer organizations will often employ R.

Managing Keyword Portfolios

Rahul Singh

MS '16, majoring in information systems

Worked at a data analytics company

rahusing@indiana.edu

When a South American airline wanted to optimize its keyword portfolio, it turned to a data analytics company. Kelley MSIS student Rahul Singh was tasked with identifying primary keywords and determining how much the client should bid for them on search engines. The client had spent a significant amount of money and wanted to maximize profits by bidding on only the most effective keywords.

Because this was the first time Rahul's company had worked with the client, meeting expectations was key to building a lasting relationship. Due to the nature of the project, the team faced unprecedented roadblocks and had to undertake a very different process to suit the client's needs.

What were the steps taken to complete the project?

After speaking with the client about their expectations and the business problem, we developed a list of variables and data on which we would run our analytics. We then downloaded data for the previous four years from the client's data vendor. The data included keyword information, clicks, time period, cost, and value obtained across different dimensions such as account name, region, and branch. We also obtained information on their promotions, discounts, and other marketing strategies.

Cleaning the data took two weeks. The data set was nearly 10 gigabytes and contained issues such as outliers, errors, and special characters in numerical columns. We used descriptive statistics and our own macros to identify these discrepancies and clean the data. Next, we split the data set into a validation set and a development set, then ran a time-series ARIMA model in SAS to identify the primary keywords in the portfolio. We studied the pricing the client had done on these keywords and matched it against a competitor's pricing strategy. We used linear regression in SAS and Solver in Excel to identify the optimum price for the keywords. Finally, we prescribed changes to the bidding on the keywords.

What challenges did you face in the initial stages of the project?

We spent about 80 percent more time cleaning the data than originally intended. In retrospect, we realized that we should allot more time for data cleansing when working on a project with a new client or getting data from unfamiliar sources.

What were the apt variables used for the model?

We used different variables for each of the two models we built. For the ARIMA model, we used account name, branch name, keyword, timestamp, customer information, click information, and ticket amount. For the linear regression model, we used keyword, price, competitor pricing, and other keywords influencing the keyword.

What were the challenges with applying analytic techniques?

We had to use tools that could handle such a large data set, such as SAS and R, and we had to build macros from scratch. Obtaining the data was a challenge due to security and access restrictions. Additionally, the data were in Spanish, and several keywords were unknown. Finally, we spent a significant amount of time fulfilling client changes to the final requirements.

What did you learn from this experience?

We should realize that certain requirements can be fulfilled only when you are given adequate information and time. Due to the assumptions we had to make, it was challenging to meet every request for a change halfway through the project. It is important to stick to timelines and to the scope of the results that are needed.

What was central to providing value for your client, and how did you do it?

The most important thing was to build a business case out of our recommendations. The client's interest in which statistical method we used was only in knowing the accuracy of the model. The client wanted answers to "How should I spend my money on these keywords?" and "How much revenue will I gain if I do that?"

In order to provide extra value to the client, we created a user-friendly tool in Excel using VBA macros. The tool was designed to show the different revenue and results obtained by changing bids and prices of keywords in the entire portfolio. We then trained the client so they could maximize their use of the tool.

The Role of Attribution Metrics on Keyword ROI

Alice Li
Assistant Professor
of Marketing

aliceli@indiana.edu

When potential customers use a search engine to find a product, their keywords can range from general ("jewelry") to specific ("silver necklace with ruby"). Search engines help customers find useful information about a product, whether they are just beginning to browse or are ready to commit to a purchase. Retailers can compare the keywords customers use to the resulting purchases, or conversions, to determine how best to apply their advertising budget to buying sponsored search results.

The researchers examine the influence of attribution metrics—that is, whether a retailer should attribute a sale to the first or last search term a customer uses—on the return on investment a retailer will get when purchasing sponsored search results. In this study, the first of its kind, the researchers give insights into the relationship between the attribution metric and the type of keywords—general or specific—that a customer uses.

Statement of Problem

Customers tend to use more general keywords when they are in the early, or upper, part of the purchase funnel. Conversely, customers in the lower part of the purchase funnel—those who are near their decision—tend to use more specific keywords. To ensure the best use of its

advertising budget, a retailer must be sure to bid for the correct mix of upper funnel and lower funnel keywords in its advertising to avoid missing out on large segments of potential customers. Additionally, when a customer clicks on multiple paid search results before making a purchase, the retailer must determine how to allocate the conversion credit to the keywords. A new allocation results in new performance of the keywords. What will happen to the ROI of search campaigns when a retailer changes its attribution metrics?

Data Sources Used

The researchers used six months of keyword data provided by an online jewelry retailer. The data set consists of 505 unique search keywords on Google and Bing that were clicked at least once between January 21 and July 18, 2012. The data includes daily information about each keyword: the number of impressions, the number of clicks, the average cost per click, average ad position, quality score, and disguised revenue.

Analytic Techniques

The researchers used the keyword data provided by the retailer to analyze the ROI of paid search campaigns. The retailer had changed attribution

metrics halfway through the time period captured by the data, allowing the researchers to estimate the impact of both first-click and last-click attribution. To highlight the importance of accounting for the attribution metric when determining an advertising strategy, the researchers developed a quantitative model to capture the simultaneous changes on the revenue and cost of keywords to the retailer, how the search engines position the retailer's keywords, and the customer's response to the keyword ads, such as click-through rates and conversion rates. The researchers also illustrated a counter-factual scenario, proposing a third attribution metric to account for a keyword's contributions under both attribution methods rather than one or the other. This can help the retailer improve the ROI of its search ads by directing better use of its budget.

Results

The research yielded five equations: the advertiser's revenue model, the bidding model, the search engine's ad position decision, and the aggregation of customer click-through rates and conversion rates. Their results show:

- A keyword portfolio with more upper funnel keywords will generate more revenue when using the first-click attribution method, while a portfolio with more lower funnel keywords will generate more revenue using the last-click method.

- The presence of a brand name in a keyword can improve the position of an ad and increase the conversion rate, but does not significantly affect the click-through rate.

- The click-through rate decreases when paid search results are ranked lower.

- The click-through rate demonstrates an inverted-U shape with respect to how specific a keyword is. That is, the click-through rates of very broad and very specific keywords are lower.

- Conversion rates do not change significantly depending on the position of the ad in search results.

The researchers' results indicate that using a first-click attribution method negatively affects this particular retailer's revenue. Using a fractional attribution metric takes into account a keyword's potential contribution under both the first-click and the last-click attribution strategy. By using a fractional attribution strategy rather than a first-click attribution strategy, the retailer could improve its revenue by at least seven percent without increasing its advertising budget.

Business Implications

The researchers propose that their modeling framework will help retailers better understand how attribution metrics affect their ROI and, as a result, better allocate their keyword marketing budget. Retailers' advertisers can experiment with and change the metrics they use, then use the researchers' model to discover the effect each attribution metric has on revenue.

> "To ensure the best use of its advertising budget, a retailer must understand the impact of attribution on their search campaigns."

Because the performance of keywords will differ depending on attribution metric, advertisers will be able to identify important upper funnel and lower funnel keywords. This will allow advertisers to determine the best metric for each keyword and adjust their bidding accordingly. Additionally, the keywords may provide insights into which product categories benefit more from which keywords.

The researchers note that if a particular upper funnel keyword performs very well under the last-click metric, this may indicate that a retailer's website is well designed, because once the customer arrives at the retailer's website, the customer will remain engaged until conversion rather than going back to the search engine and searching for another retailer. By analyzing keywords and the roles they play in purchases, retailers can judge how they might improve their websites to more effectively increase the conversion rate.

Hongshuang (Alice) Li, P. K. Kannan, Siva Viswanathan, and Abhishek Pani, "Attribution Metrics and Return on Keyword Investment in Paid Search Advertising," presented at INFORMS Marketing Science Conference, Baltimore, MD, June 2015.

The Promise of Digital Marketing and Metrics that Matter

KEYNOTE:

Michael Wilhite
Senior Vice President, 84.51º

PANEL MODERATOR:

Jeff Kavanaugh
Vice President-Managing Partner, Infosys

PANELISTS:

Byron Hardie
Senior Marketing Director, Angie's List

Mike Kaplan
Vice President, Salesforce

Michael Lancor
Director, Procter & Gamble

Tracey Moon
Chief Marketing Officer, Brillio

Michael Wilhite opened the 2015 Kelley Forum on Marketing Analytics with a keynote address about the promise of digital marketing. He described two groups: those who think that traditional metrics are more important, and those who think they have digital marketing figured out. Wilhite believes that the truth is in the middle. "That's why it's the promise of digital marketing. You have to understand what will drive major trends and how analytics will play a strong role." Digital marketing provides new data sources and the opportunity to operationalize analytics.

Wilhite pointed out that the need for immediate responsiveness is vital because competition comes from all facets. "You always have to be able to adapt," he said. "Digital marketing is already a big deal."

While offline advertising is barely growing, he said, online advertising is increasing at 14 percent. At the same time, consumers are opting out. There's no recipe for success other than programmatic research to determine how your customer base wants to interact with you. Personalization helps—the consumer learns that you're going to provide relevant content through targeted email. And analytics are the key to understanding what works.

The opportunity exists to target the right ads to the right people; however, consumers think of online ads as a nuisance. Wilhite cited Facebook as an example of a business model focused on advertising revenue. "Most consumers aren't looking to Facebook as that channel," he said. "They're trying to understand what's going on with their family and friends, and ads are intrusive. As ad blocking gets more effective, consumers will opt out. In each case, it's about using analytics to drive more effectiveness."

Wilhite then talked about how consumer feedback has been one of the most effective marketing tools in the last decade. "The power of online reviews and user-generated content relies on the principles of transparency and trust." Business models are being developed based on a foundation of user-generated content.

Finally, Wilhite mentioned the golden triangle of disruption: mobile, social, and real time. He recommended that companies assess technology changes to learn what's important and where they should look for future investments.

Moderator Jeff Kavanaugh then asked the digital marketing metrics panelists to talk about something they found interesting about marketing analytics.

Michael Lancor stated that he isn't sure that there's anything nondigital in marketing. "I think when we're talking about digital, what we're probably talking about is effective marketing."

Byron Hardie finds that many of the same tenets of product development are coming into digital marketing. "It's about being able to identify user intent, align it with the marketing message, and deliver the right message at the right time."

Mike Kaplan said that he has the opportunity to discover what customers are doing with their marketing efforts. "Over the years, we've seen technology play such a big role."

Tracey Moon opined that buyers tend to be very cynical. "We're using analytics and data to figure out who influences them, and how we can use marketing or user-generated content to surround them with good news about our company."

A question and answer session followed these opening statements.

Q: Is digital marketing everything?

MOON: We use analytics to share content with influencers in the industry, and to try to improve all the way through the funnel. There's still an element of old-school marketing in what I do, including meeting people. We find if you can combine the digital and physical, it works.

KAPLAN: We're seeing technology embedded in all of the events we participate in. Everything is tied together. Customer service is as much marketing as it is selling or servicing today.

LANCOR: Ten years ago, talking about digital marketing was helpful because it was new. From a marketing analytics and measurement standpoint, almost everything is digital now. Digital analytics give us the ability to measure what's real and target the right messages to the right people at the right time.

Q: What are some of the challenges of measurement? How do you get metrics you can make decisions on?

MOON: For us, marketing is measured by impact on revenue—did marketing bring in the lead, and what happened? There are so many tools, and I'm trying almost every one of them.

HARDIE: Ideally, you would have a single view of the user: first-party data including channel engagement, demographic information, and intent mapping blended with third-party data from other marketing channels. Start by connecting a few key points such as email engagement, time-on-site, return visits, and conversion, then build on that.

Q: How do you apply marketing techniques and metrics collaboratively in an ecosystem?

LANCOR: That ecosystem of information is becoming critical, but nobody wants to share data. As a manufacturer, you're trying to bring together as much information as possible to make better products and advertising. What's helping is the ability to bring data streams into a single dashboard to provide greater access.

Q: What about advertising connected to telematics?

HARDIE: The biggest thing is to identify user intent. As you guide a user through a process, you're acquiring data and providing a service. We've gone far in the science of analytics, and I think the pendulum will swing back to the artistry of it. You can draw incorrect conclusions from data, and that's where analysis comes in. Truly great analysis requires artistry and humanization. Remember, segments don't evangelize your brand, share your content, or buy your product. People do. We should strive for the segment of one.

Q: What works when it comes to data ownership?

LANCOR: No one person or group really owns information anymore. We've started putting media and analytics into a single group that crosses different organization lines, making it clear that it's company data. Data analytics have to be a top strategy.

KAPLAN: We find many companies where email and social are not connected. How is anyone running effective multichannel marketing without connectivity and collaboration? So we're seeing more and more where the marketing organization has its own data function.

Q: What are you seeing with technology?

MOON: We use HubSpot with Salesforce, and that gives us a better view. But we need a lead forensics tool. And then you have intelligence tools like RainKing and LinkedIn Sales Navigator. There's no shortage of tools out there. I wanted to set a strong foundation of marketing automation and sales and then start pinpointing where we needed to improve.

KAPLAN: We're seeing more connecting technologies. You should look at the ability of technologies to send and receive data in real time and to have the openness to connect. The app ecosystem is very rich.

HARDIE: When you look at artificial intelligence, you don't have a single marketing message. You have a customized message that changes as new data comes in. That really is the future of marketing—pinpointing the individual user and what they want, and servicing their need in real time.

LANCOR: Programmatic marketing is still in its infancy. I think you'll see algorithms and machine learning transform marketing because they enable precision and real-time A/B testing. Also, technology is transforming how people shop and will have a strong impact on how we think about marketing in the future.

Analyzing Data to Develop a Marketing Strategy

Swetha Prasad

MBA '16, majoring in marketing and business analytics

Interned at a global technology corporation

swprasad@indiana.edu

The marketing team at a large technology corporation uses a customer relationship management (CRM) tool to aggregate, analyze, and monetize information collected from various systems and programs through which the team engages with developers and customers.

As part of her internship, Swetha Prasad performed data analytics on this CRM tool. The useful insights she gained allowed her to devise recommendations for a more focused marketing strategy.

What were the steps taken to complete the project?

I first needed to understand the nature of the business, the customers, and current marketing efforts. Reaching out to relevant stakeholders helped me narrow my analysis to answer three main questions and develop a specific marketing strategy focus for each. The questions were:

- How does the performance of different industry segments affect developer activity?

- How are marketing efforts aligned to—and increasing—sales?

- In what phase does a given developer fall in the company journey: awareness, interest, purchase, or advocacy?

My next step was to evaluate the data preparation efforts that were required before analytics could be performed. Preparation efforts included cleansing, normalization, and conversion of qualitative data to quantitative wherever possible for better analysis.

I then used Excel to conduct a trend analysis and created reports illustrating developer activity related to marketing efforts in different industry segments. I used Excel to conduct a correlation analysis of developer and sales data to understand the effectiveness of marketing efforts on sales. Finally, I conducted a cluster analysis using IBM SPSS Modeler to sort developers into the four phases in order to identify target customers.

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

Data preparation consumes more than 90 percent of the time spent on analytics. Having an efficient collection and/or storage mechanism can help minimize this effort. It is also important to have standardized metrics across different functional groups within an organization when collecting and storing this information.

No recommendation is solid if it is not backed by relevant data. For continual improvement and innovation, it is important to analyze data and develop insights to solve problems or provide direction. Translating qualitative data to quantitative information can help better visualize the story the data tells.

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

A core course in quantitative methods gave me a good foundation in understanding types of data and an introduction to regression analysis. Introduction to Spreadsheet Modeling gave me a great overview of Excel's capabilities. Finally, Data Mining gave me a solid foundation in using IBM SPSS Modeler.

What did you learn from this experience?

I learned the importance of collecting the right data in the right format to assist with analytics, as well as the importance of translating the results into a business story that paints a clear picture for business leaders. Visualizing the results using Excel and Tableau helped me translate these results effectively.

Given what you've learned from the project and the coursework, how are you better prepared for this type of role?

I better understand the process of data analytics, and have a good structure and methodology to follow in future projects. I understand the power of data analytics in evaluating the market or customer behavior and using the results to formulate strategies. I have a better understanding of various tools and how to apply them. And I have developed a much deeper understanding of how critical translating data into business results is in communicating the value of analytics.

Updating a Value Proposition

Vikram Swamy

MBA '16, majoring in business analytics and finance

Interned at an online automotive listing company

vikswamy@indiana.edu

Due to increasing competitive pressure in the industry, one online automotive listing company decided that it needed to update its value proposition for the first time in more than a decade. Intern Vikram Swamy helped the company discover how to redefine its communication methods and strategically adjust pricing for its core products in order to better serve consumers.

What were the steps taken to complete the project?

Though I worked on more than a dozen projects during my internship, the one I contributed the most to was the company's value definition initiative. The main objective was to listen to the voice of the customer—car dealerships—and to update the value proposition for the company's main product while ensuring that pricing objectives matched accordingly.

My main task for this project was a lead attribution study. I worked with the company's business intelligence (BI) team and with an outside vendor to see whether leads were still providing customers with the value that they assumed. I first had to talk with the vendor to determine what information they needed in order to give us the best data on vehicle ownership. Once these data were acquired, I worked with BI staff to figure out how the data could be analyzed to give the most accurate picture of the value proposition of the company leads.

I then communicated the findings to the vice president sponsoring the project and provided a strategic recommendation based on the findings. I also had a meeting at the end of the project where managers asked questions, requiring me to defend my conclusions.

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

One of the biggest lessons I learned from this project is never to jump into a complex and vast data set without a strict action plan. It's easy to get lost in the data, and having such a plan provides an anchor to a strategic mindset. This was also instrumental in the question and answer session, because the mental heuristics I undertook to come up with the plan provided me with the expertise I needed to answer the questions.

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

I was able to succeed mainly due to my Kelley coursework. Some beneficial classes included a market-based analytics course, a spreadsheet modeling course, and a database warehousing course.

What did you learn from this experience?

This experience gave me an appreciation of project management and methodologies. It had me constantly aware of deadlines, resources, and potential issues on a project level—something I had never been exposed to before. One of my biggest lessons is to never go unprepared into any meeting. Always have questions to ask and think of potential responses to questions you may be asked.

Given what you learned from the project and the coursework, how are you better prepared for this type of role?

One of the biggest challenges for me was strategic exposure. My previous roles have been operational, with little experience regarding corporate strategy or analytics. Because of Kelley academics and my internship, as well as my pre-MBA work, I've become much more adept at bridging the operational and strategic sides of a company. Thanks to Kelley and to my internship, I now always have a plan before I began analyzing large data sets.

Factors Influencing Investment in Renewable Technologies

Gilvan C. Souza
Professor of Operations
Management; Rifkin Family
Faculty Fellowship

gsouza@indiana.edu

Before an organization chooses to invest in renewable energy technology such as solar panels, it should first determine the most cost-effective energy-generating capacity to purchase. By examining current energy bills and rate schedules, the amount of solar radiation in the organization's geographic area, the amount of power each panel can put out, and the cost and expected useful life of each panel, it is possible to find the optimal number of panels to purchase.

In this paper, the researchers set forth the idea that the level of granularity of the solar radiation and electricity demand data—minutes, hours, days, and so on—significantly affects the determination of this optimal capacity. If an organization uses only data aggregated into days, months, or years to make their capacity decision, renewable energy sources may appear to be economically infeasible when in fact a closer look at the data would paint a different picture.

Statement of Problem

Organizations have an unceasing need for energy. Constantly rising energy prices and an organization's desire to reduce its carbon footprint may encourage the investigation of alternative, renewable technologies such as solar or wind power. Whether such a system is feasible for any given organization—and if so, which technology is best—depends on several factors: expected energy prices, availability of government incentives, and the organization's demand for energy versus the expected supply. An organization must be able to determine its optimal capacity investment in order to obtain a renewable technology that best meets its needs while at the same time providing a cost-effective alternative to the standard electrical grid.

Data Sources Used

The researchers examined two cases: a national bank's investment in a solar photovoltaic system at branches in Los Angeles and in South Carolina, and a hotel's investment in a combination of solar and natural gas to heat water in San Francisco, California.

For the first case, the bank provided its Los Angeles branch's electricity meter readings for 2013 in 15-minute intervals. The researchers obtained solar radiation data in minute

increments for the same location and time period, monthly electricity rates, and governmental incentive rates. For comparison, they also obtained the same types of data from a branch in Bluffton, South Carolina.

For the second case, the researchers used the hotel's water heating demand observations from 2004, along with solar radiation data from the same time period.

"The current approach to capacity investment decisions usually examines average efficiency, which can lead to an overinvestment in renewable technology."

Analytic Techniques

The researchers examined a situation in which an organization planned a one-time investment in a renewable energy technology. They developed a model that takes into account physical constraints, energy and operating costs, variability of renewable energy sources, and government incentives for returning power to the grid.

In the case of the Los Angeles bank branch, the researchers used the most granular data available—15-minute intervals—for both demand and solar yield to calculate the true optimal solution. They next aggregated adjacent intervals into hours, calculated the optimal solution for various levels of granularity, and compared the results to the true optimal solution. The researchers also compared their solution to a heuristic that uses average yield efficiency. They then did similar calculations for the South Carolina branch.

For the San Francisco hotel, the researchers used granularity to analyze the value of energy storage, since water heated by a solar thermal system is not necessarily used immediately. In this application, government incentives do not come into play because no energy is returned to the grid. To find the optimal capacity of the renewable system, they added units of capacity up to the point where the natural gas technology provided the same marginal cost.

Results

The researchers have shown that examining renewable yield and energy demands at a very granular level can significantly affect capacity investment decisions. Using more granular data better takes into account the interaction between random demand and solar yield.

In the case of the Los Angeles bank branch, the optimal solution remains close to the same at the 15-minute and one-hour levels. As granularity decreases further, the spread in possible solutions increases and introduces more uncertainty into the investment decision. When the researchers compared their solution to the heuristic, the result of the heuristic was a solution with a capacity more than three times greater than the optimal solution.

At the bank's South Carolina branch, optimal capacity is smaller due to a more negative correlation in solar yield and demand and the fact that the state does not offer government incentives for power returned to the grid. The difference between the two branches demonstrates that rather than being standard, investment decisions should consider factors that vary significantly across geographic locations.

In the case of the San Francisco hotel, the researchers determined that a renewable technology could meet a significantly high portion of demand, particularly due to the water heater's ability to store energy. Energy storage makes renewable technologies more attractive.

Business Implications

Organizations should consider re-evaluating previous investments in renewable energy technology if those decisions were made using non-granular data. The current approach to capacity investment decisions usually examines average efficiency, or the average random yield over a year. This can lead to an overinvestment in renewable technology, because this average fails to incorporate the intermittency that occurs between supply and demand. Using the decision tool set forth by the researchers, organizations can better determine their optimal portfolio of energy-generating technologies. This will lead to a more cost-effective investment in renewable energy technology and should prevent both overinvestment in capacity and the potential for complete abandonment of the idea of installing renewable technology due to a perceived lack of benefit.

S. Hu, G. Souza, M. Ferguson, and W. Wang, "Capacity Investment in Renewable Energy Technology with Supply Intermittency: Data Granularity Matters!" *Manufacturing & Service Operations Management*, 17 (4), pages 480–494, 2015.

Using Prediction Markets for Accurate Forecasts
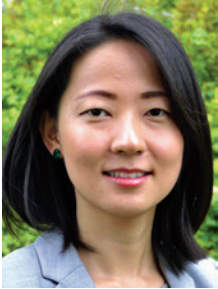
Ruomeng Cui

Assistant Professor of Operations
and Decision Technologies

cuir@indiana.edu

The ability to create accurate forecasts is essential in many business processes, including demand planning, procurement planning, and new product development. Sales forecasts are particularly important in operations management, since they facilitate decisions such as production planning. But in the absence of historical data, how can a company create an accurate forecast?

One method is the prediction market, which is used to collect and quantify—or crowdsource—knowledge from a group of diverse individuals. Prediction markets run much like the stock market, with a group of individuals trading on outcomes they believe to be most likely. Even if most people in the group are not well informed, through a phenomenon known as the wisdom of crowds, the group as a whole can reach a wise decision.

The researchers investigate using prediction markets in operations management, particularly regarding sales forecasts, price commodity forecasts, and predictions of popular product features. They also determine the effect of crowd size on forecast accuracy.

Statement of Problem

When historical data is available, companies can analyze it to generate future predictions. But in some cases, such as a new product release, data may not exist and traditional forecasting methods cannot be used. Consequently, companies must use other methods and information sources to generate their forecasts.

To determine the efficacy of prediction markets in instances where historical data is unavailable, the researchers seek to answer two questions:

- How accurate are forecasts generated via prediction markets in corporate and public settings?

- What are the drivers of forecast accuracy and how can they be controlled to achieve a better forecast?

Data Sources Used

The researchers obtained data from Cultivate Labs, which facilitates prediction markets for a wide range of organizations. For this study, they chose to use data from Cultivate Labs' public prediction markets and three private corporations (chemical, retail, and automotive). Each data set contains the market information, stock information (i.e., outcome choices), trading history, and price history of each stock.

The public prediction markets were primarily in the areas of finance and economy, sports, entertainment, and politics. The public data set includes 9,558 prediction markets with 735,455 trades placed by 28,754 users from 2006 to 2009.

The chemical company focused on commodity price forecasts, launching 27 prediction markets in which 31 employees participated. The retail company focused on predicting regional sales, launching 27 markets in which 319 employees participated. The automotive company focused on forecasting weekly car sales, launching 84 markets in which 301 employees participated.

The researchers also gathered data from their own prediction market, in which 52 second-year MBA students participated.

> "Through the wisdom of crowds, a group of diverse individuals can reach a wise decision."

Analytic Techniques

The researchers chose to focus their efforts on continuous-variable markets, which use data such as sales and prices, because most variables that support operations decisions are continuous.

The researchers applied a quantile assessment approach to test their hypothesis that prediction market distribution forecasts equal the actual distribution for the chemical company, the retail company, and the public market. For the automotive company, they used mean absolute percentage error to measure forecast accuracy and compare performance relative to both industry average and a focus group conducted by the company.

Next, they used a Brier score to measure the accuracy of the predictions in each public market, followed by a regression analysis to examine the impact of group size on each market. Because easier prediction markets—those with less complex questions—most likely have a more accurate outcome, the researchers designed an experiment that would allow them to manipulate the size of the group and control for other factors. The experiment subjects were MBA students enrolled in an elective supply chain management course. The students were randomly divided into two groups of 8 and two groups of 18. The students used a prediction market to determine the demand for 38 summer trips. The researchers offered incentives to help ensure good forecasts.

Results

The researchers determined that prediction markets produce accurate forecasting results. Additionally, the size of the group does have an impact on forecast accuracy. The researchers determined that their groups of 18 students created forecasts that were significantly more accurate than those of the smaller groups.

The researchers observed increased accuracy in trades occurring later during the time a prediction market was open. This indicates that as time passes, participants get more information. Therefore, prediction markets will successfully incorporate new information as it becomes available.

In a corporate context, the researchers' results show that a crowd of employees can accurately forecast operations variables such as commodity prices and sales. For the automotive company, crowd wisdom produced forecasts that outperformed both industry average and the company's internal focus groups.

The researchers have shown that with proper information aggregation methods, prediction markets can successfully translate crowd wisdom into accurate distribution forecasts suitable for use in the different stages of a supply chain.

Business Implications

The researchers' results suggest that forecast accuracy can be improved by increasing the number of participants, which will in turn increase diversity in background and opinions. They indicate that companies should create groups with at least 8 members to elicit good predictions. However, larger groups are better—18 participants create more accurate forecasts than 8.

Properly implemented crowd forecasting could help companies earn (or save) millions. For certain areas such as consumer goods sales forecasts, crowd wisdom from the general public can be helpful. Areas in which the public has good knowledge could potentially facilitate company operations and marketing decisions.

The researchers' results indicate that companies can confidently make strategic decisions based on the outcomes of prediction markets. Prediction markets can be used not just for sales forecasts, but also for research and development, project management, and a variety of other business problems.

Practical Sustainability and Sustainability Metrics

KEYNOTE:

Mitch Jackson
Vice President, FedEx Corporation

PANEL MODERATOR:

Patti Murdock
President, Clean Logistics

PANELISTS:

Steve Raetz
Director, C. H. Robinson

Karen Albertson
Vice President, 3M

Ron Mims
National Sales Manager,
AcornVac, Inc.

Mitch Jackson's keynote address for the 2015 Kelley Forum on Supply Chain Analytics focused on practical sustainability. Jackson's company, FedEx, delivers millions of shipments every day for customers around the world. At a broader level, the company connects people and possibilities. "We're built on the belief that local economic growth requires connectivity with the rest of the world. But we've got to do it responsibly, resourcefully, and sustainably," Jackson said. "It's something that permeates our thinking and translates into the concept of practical sustainability."

Practical sustainability is the application of sustainability through strategic and transformational stewardship that adds value. Jackson talked about the four building blocks: performance, transparency, innovation, and leadership. "We wanted to integrate sustainability into the value proposition of the company—to make it part of the business," he said. "Performance is first and foremost. If you're not performing each and every day, you will not be around in the future."

The second building block, transparency, concerns sharing information about what a company is doing and how, so that others may learn from it.

The third building block is innovation. "I define innovation as 'applied inspiration,'" Jackson said. He pointed out that a company's competitors must take action in response to successful innovation.

The fourth building block is leadership. Jackson cited FedEx's push for fuel efficiency greenhouse gas standards in commercial vehicles. "Leadership is sometimes uncomfortable and uncertain," he said. "That's why not everybody does it. We didn't know how it would ultimately take shape, but we needed to be a part of that process to improve transportation. What was already occurring didn't work."

Jackson cautioned that anyone implementing sustainability initiatives should avoid what he calls pinball leadership. "If you don't have a direction in mind and are bouncing around and changing direction from external factors, you're not leading. You're simply reacting. Lacking data or not analyzing data contributes to that." He went on to say that companies should enable a license to optimize—give teams the information they need, the general direction they should go, and the freedom to do it. "They'll often exceed what you expected," he said.

Next he talked about enlightened serendipity. "It's being enlightened enough to know that change is necessary and to have a direction in mind. The information and data you get will show you some of that. And you're sometimes going to succeed—ultimately, to make a difference," he said. "There's an old adage in sustainability that what gets measured gets managed. That's leaving out something. What gets measured *and matters* gets managed. Data matters. It tells us where we are, where we're going, and, more importantly, where we should be going," Jackson concluded.

After the keynote, moderator Patti Murdock introduced the panelists and asked for opening statements.

"We use technologies, products, and processes to make products and solutions for our customers," said Albertson. "Manufacturing is a core strength. That's important because a lot of people think of sustainability as a manufacturer's responsibility. But it's much bigger than that. We consider every employee in charge of sustainability," she added. "It's foundational, whether you're a scientist in the lab, an engineer in the factory, a production employee, or a marketer." Sustainability at 3M isn't only an internal priority—the company also helps its customers with their own sustainability efforts. Albertson cited an example of microscopic glass bubbles that are being incorporated into vehicles to make them more lightweight and efficient, yet improve strength.

Mims spoke next about vacuum waste systems, which can conserve water and allow for more flexibility in building design. He discussed a correctional facility in which AcornVac had installed a vacuum waste system, saving the facility three million gallons in water usage yearly compared to a conventional gravity system. Mims explained that not only does vacuum waste save water and money, but it can also save space, time, and energy. "We use the combined energies of gravity, air, and vacuum to move waste. That means saving material costs inside a building."

Next, Raetz explained the role of less than truckload (LTL) in sustainability. The challenge of moving a partial load lies in understanding how much of a truck's emissions can be attributed to each shipment on the truck. The two most recognized standards for estimating carbon emissions of a full truckload are the Greenhouse Gas Protocol and the Environmental Protection Agency's (EPA's) SmartWay. Raetz explained that applying either standard to LTL ignores the complexities of freight movement. To map this movement, C. H. Robinson commissioned a project at MIT's Center for Transportation Logistics. There, faculty and students developed an accurate representation of a shipment. "If you want to truly get the emissions of a shipment, you have to use the detailed model," said Raetz. "The problem is, the only people who know the data are the LTL carriers. So we created a simple model to try to replicate the detail and got within a three percent variance." C. H. Robinson has presented their findings to the EPA and has proposed this model as the new national standard.

A question and answer session followed.

Q: How are you helping customers with their sustainability goals?

ALBERTSON: We have a sizeable business with the aerospace market. Our customers are constantly trying to improve fuel efficiency by reducing weight and maintaining or improving strength. As a core material science company, we are constantly innovating to help them achieve these goals.

Q: Do you find sustainability to differ in focus in different parts of the world?

ALBERTSON: With a global customer, sustainability goals will be global goals. Their execution at the local level may have to meet local metrics or outcomes, but they're going to have a global target and global capabilities. We are all aware of those local expectations and laws, as 3M is a global company with sales in more than 200 countries worldwide.

Q: Do you have specific targets for parts of the world where some things are more important than others?

ALBERTSON: We have very targeted goals for our water scarcity locations. Factories have a very different ecosystem because we have to do things differently—not just because it's legislated, but because it's the right thing to do.

RAETZ: I talked about the EPA's SmartWay program. That didn't exist in Europe. We were a founding member of Green Freight Europe. It's meeting needs similar to SmartWay, but is pan-Europe and member-based versus government-based. It seems reasonable that each economic or governmental region may have some customized solution working toward a common objective.

Q: Steve, what behavior and decisions would a shipper make differently based on your model?

RAETZ: I think most of our customers today focus on their environmental impact as it relates to things that occur within buildings and assets that they own. We experience little engagement today from our customers around transportation emissions. We wanted to be ahead of the curve and try to set some of the standards.

Q: How do firms deal with incompatibility between goals of environmental sustainability and cost reduction?

RAETZ: It seems that most supply chains have a lot of opportunity to reduce waste from a cost standpoint. That said, transportation-related emissions reduction typically follows the waste/cost reductions. I think if regulation changed in a material way and something drastic was imposed on freight movement, cost and emissions tradeoffs may not be the same and the decisions could become more difficult.

ALBERTSON: Where we have challenges is usually in product performance. For example, we would really like to be solvent-free. It's difficult to design solvent out of some formulations and achieve customer expectations in performance. Our goal is to use science to achieve the desired outcomes and provide upside benefits to the customer and market.

2015 Year in Reflection

2015 Corporate-Student Roundtable

Participating companies: Deloitte, Accenture, Humana, Eli Lilly & Company, C. H. Robinson, Roche Diagnostics, Protiviti, Whirlpool, Indiana University, ArcelorMittal, Tableau, SAS-JMP, Angie's List, Essendant, Cummins

Thomas Andreesen
Protiviti

Rob Lowden
Indiana University

Kylie Cherco
Deloitte

2015 Analytics Leadership Award

Judging panel: Curt Hinrichs (chair), SAS-JMP; Dion Rudnicki, IBM; Bill Russell, Allegient

Dion Rudnicki, IBM, **Curt Hinrichs,** SAS-JMP, **Jerry Oglesby,** SAS

Runner-up: Mike Duffey, The Wiz Reporting Suite 1.0: Activity Based P&L and Strategic Workflow, C. H. Robinson

Winner: Harpreet Singh, Vipin Gopal, Predictive Model-Based Identification and Outreach for Seniors at High Risk of Falling, Humana

Keynote speaker: Ansa Sekharan, Informatica

2015 Forum on Marketing Analytics–Digital Marketing

Co-chairs: Professor Shibo Li and Michael Wilhite, 84.51°

Michael Wilhite
84.51°

Panel on "Digital Marketing Metrics that Matter: The Good, the Bad, the Ugly"

- Panel Moderator: Jeff Kavanaugh, Infosys
- Panelists: Byron Hardie, Angie's List; Mike Kaplan, Salesforce; Michael Lancor, Procter & Gamble; Tracey Moon, Brillio

2015 Forum on Supply Chain Analytics–Sustainability

Co-chairs: Professor Gil Souza and Glenn Wegryn, Analytic Impact LLC

Mitch Jackson
FedEx Corporation

Panel on "Sustainability Metrics"

- Panel Moderator: Patti Murdock, Clean Logistics
- Panelists: Steve Raetz, C. H. Robinson; Karen Albertson, 3M; Ron Mims, AcornVac, Inc.

2015 Predictive Analytics Challenge

Judging participants: Deloitte, Accenture, Humana, Eli Lilly & Company, C. H. Robinson, Roche Diagnostics, Protiviti, Whirlpool, Indiana University, ArcelorMittal, Tableau, Cummins, SAS

Winning MBA students Nate Buyon, Brad Kleinschmidt, Eric Kuehl with the judges

Winning MSIS students Ben Hartman, Rahul Singh, Peter Buschbacher, Siddharth Pasumarthy with the judges

Winning MBA students Jayanti Ghate, Swetha Prasad, Yang Luo, Sho Majina with the judges

IBA Affiliated Faculty

Frank Acito

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

Goker Aydin

Associate Professor of Operations & Decision Technologies;
ArcelorMittal Faculty Fellow of Supply Chain

Hillol Bala

Associate Professor of Information Systems;
Whirlpool Corporation Faculty Fellow

George Ball

Assistant Professor of Operations & Decision Technologies

Azi Ben-Rephael

Assistant Professor of Finance

Ruth Beer

Assistant Professor of Operations & Decision Technologies

J. Doug Blocher

Chairperson and Professor of Operations & Decision
Technologies; Arthur M. Weimer Faculty Fellow

Kurt M. Bretthauer

Professor of Operations & Decision Technologies;
John & Esther Reece Professor

Dan Carroll

Lecturer, Operations & Decision Technologies;
Co-Director, MBA Supply Chain Academy

Kyle Cattani

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

Hai Che

Associate Professor of Marketing

Ruomeng Cui

Assistant Professor of Operations & Decision Technologies

Isaac Hacamo

Assistant Professor of Finance

Jonathan Helm

Assistant Professor of Operations & Decision Technologies

Randy Heron

Professor of Finance; OneAmerica Foundation
Endowed Chair

John Hill

Clinical Assistant Professor of Operations & Decision
Technologies

Craig Holden

Professor of Finance; Boquist-Meyer Faculty Fellow

Ryan Israelsen

Assistant Professor of Finance

F. Robert Jacobs

Professor Emeritus of Operations

Niket Jindal

Assistant Professor of Marketing

Vijay Khatri

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

Kristoph Kleiner

Assistant Professor of Finance

Alice Li

Assistant Professor of Marketing

Shibo Li

Professor of Marketing; Arthur M. Weimer Faculty Fellow

Alex Lopes

Clinical Associate Professor of Information Systems

Girish Mallapragada

Assistant Professor of Marketing

Alfonso J. Pedraza-Martinez

Assistant Professor of Operations & Decision Technologies

Brian P. Miller

Associate Professor of Accounting; Arthur M. Weimer
Faculty Fellow

Alex Mills

Assistant Professor of Operations & Decision Technologies

Neil Morgan

Professor of Marketing; PetSmart Inc. Distinguished
Professor of Marketing Chair

Veronika Krepely Pool

Associate Professor of Finance; Gregg T. and Judith A.
Summerville Chair of Finance

Philip T. Powell

Clinical Associate Professor of Business Economics and Public Policy; Daniel C. Smith Faculty Fellow; Associate Dean of Academic Programs, Indianapolis

Jeff Prince

Associate Professor of Business Economics; Harold A. Poling Chair of Strategic Management

Anjanette Raymond

Assistant Professor of Business Law and Ethics

Lopo L. Rego

Associate Professor of Marketing

Daniel Sacks

Assistant Professor of Business Economics and Public Policy

Kim Saxton

Clinical Associate Professor of Marketing

Richard Shockley

Associate Professor of Finance; CenterPoint Energy Faculty Fellow

Rebecca Slotegraaf

Professor of Marketing; Whirlpool Faculty Fellow; Chair, Kelley School of Business Doctoral Programs

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

Professor of Operations Management; Rifkin Family Faculty Fellowship

Noah Stoffman

Associate Professor of Finance

Ramesh Venkataraman

Professor of Information Systems; Lawrence D. Glaubinger Professor of Business Administration; Chair of the Undergraduate Program; Chairperson of Kelley Direct

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

Wayne Winston

Emeritus Professor of Operations & Decision Technologies

Owen Wu

Associate Professor of Operations & Decision Technologies

Lucy Yan

Assistant Professor of Operations & Decision Technologies

Qiuping Yu

Assistant Professor of Operations & Decision Technologies

Jingjing Zhang

Assistant Professor of Operations & Decision Technologies; 3M Faculty Fellow

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

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

Founding Partner and Advisory Board Member

Deloitte Consulting LLP

Supporting Partners and Advisory Board Members

3M
Allegient LLC
C. H. Robinson Worldwide, Inc.
Indiana University
Informatica
Protiviti
SAS

Advisory Board Members

84.51°
Accenture
Angie's List
ArcelorMittal
Booz Allen Hamilton
Brillio
Citibank
Cummins
Discover
Dow AgroSciences
Eli Lilly and Company
Ernst & Young
FedEx Services
Hanapin Marketing
Hill-Rom
IBM Corporation
Infosys
McKinsey and Company
Methodist Le Bonheur Healthcare
Qlik
Roche Diagnostics Corporation
SAS-JMP
Tableau
Whirlpool