FINANCE | HEALTH CARE | MARKETING | SUPPLY CHAIN

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Spring 2013

On Analytics

diagnose. predict. optimize.

GET TO THE HEART OF DATA

Four Applications of Converting Information to Insights

Welcome

Dear Colleagues,

Companies that seek a competitive edge are turning to business analytics in exponential numbers. As this exciting field continues to grow, the Kelley School is working to share the latest innovations in the classroom and in the boardroom, in the United States and around the globe.

For our MBA program, we have developed a major and minor in business analytics. We have created and delivered business analytics certificate programs for Deloitte Consulting and Booz Allen Hamilton. We have recently partnered with the Indian Institute of Management at Lucknow to develop a graduate program in business analytics for students in India.

As part of our commitment to remain in the forefront of this emerging field, our Institute for Business Analytics is organizing the *Forum on Business Analytics,* which we plan to hold annually. The April event will bring together Kelley stakeholders—students, faculty and corporate partners—to discuss opportunities and challenges in applying business analytics.

This edition of Kelley *OnAnalytics* focuses on just some of the wide-reaching research by our faculty and the experiences of our students who put their skills to work during internships with top companies. We hope you find this research and the field applications interesting, and we invite you to join the conversation in this exciting discipline by connecting with Kelley's Institute for Business Analytics.

We look forward to seeing you at the forum in April.

Idalene Kesner

Interim Dean, Kelley School of Business
Associate Dean of Faculty and Research
Frank P. Popoff Chair of Strategic Management
Professor of Management

From the Editors

With the turn of 2013, we are now celebrating the second year of *OnAnalytics* magazine. This issue of *OnAnalytics* features articles by six Kelley School faculty from our Departments of Finance, Business Economics, Marketing, and Operations and Decision Technologies.

Professor of Finance Noah Stoffman studied the performance of nonprofessional investors and offered the sobering conclusion that many new investors do not seem to improve their decisions over time and would actually be better off if they quit. Haizhen Lin and Matthijs Wildenbeest, professors in the Department of Business Economics, found that the premiums for Medigap insurance plans for seniors showed considerable differences, even controlling for the risk levels of a company's customers, services provided, and other factors. The conclusion offered is that seniors could financially benefit from careful shopping prior to selecting a Medigap plan. Sandeep Chandukala, a professor in the Kelley Marketing Department, tackled the issue of identifying unmet demand for consumer products. Conjoint analysis was used to disentangle the effects of product features and attributes from price, promotion, and other factors.

For firms considering investments in sustainable technologies, evaluating the financial implications can be complicated by the uncertainty in energy costs. Shanshan Hu and Gil Souza, professors in the Operations and Decisions Technologies department, developed a model that incorporated this uncertainty as well as seasonal energy price fluctuations for a decision on fleet vehicles for Coca-Cola Enterprises (CCE). Their model demonstrated that for CCE, the best decision was to purchase a combination of diesel and hybrid diesel-electric delivery trucks rather than all of one or the other.

Students at the Kelley School are finding that more and more career opportunities are becoming available in which business analytics is playing an important role. To highlight some recent examples, we asked four of our MBA and MS in Information Systems students to report on their summer 2012 internships to discuss their experiences, specifically focusing on how analytics was used and how their coursework at Kelley contributed to their efforts. Their internship assignments were in very different industries, but all involved using data to drive understanding and decision making. Ryan Melnikas, for example, analyzed data on sales and inventory costs at a large direct marketing firm to rationalize the assortment of products carried by the firm, thereby reducing inventory costs while still meeting customer needs for options. Natalie Basich worked with a major insurance company to develop a performance model for its recent centers. Her model identified the profit and cost drivers that will enable the company to manage the centers more effectively. Jorge Alvarez developed a forecasting model for a specialty printing supplier, which enabled him and his team to more efficiently target the most promising customers and sales territories. Finally, Mayuri Pardeshi worked with a Fortune 500 client of a major consulting company to analyze and map data flows in their firm's information systems. The result of this effort was improved understanding of and confidence in analytics reports, thus enabling these reports to inform and improve decisions.

While we use this space in each issue of *OnAnalytics* to tell our readers about the articles in the pages ahead, we also want to acknowledge the professional support, creativity, and hard work of those involved in the production of each issue. Specifically, we thank Anne Auer and Jeni Donlon at Marketing and Communications in Kelley; Caroline Gilley, Dennis Hill, and Michael Nelson at IU Communications; Sarah Mote; Elisabeth Andrews; and Jenny El-Shamy.

We hope that you enjoy this issue of *OnAnalytics* that focuses on our faculty research and analytics-savvy student projects!

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

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

OnAnalytics

diagnose. predict. optimize.

14

Identifying Unmet Demand

Sandeep R. Chandukala

What do consumers want? The question can’t only be answered by observing purchasing behavior, as the most pertinent information for marketers and product developers is not which needs are already being met by the marketplace but rather which demands are as yet unmet. With this study, the researchers propose a model for identifying unmet demand that separates out different segments of consumers, enabling decision makers to address the preferences of consumers with an interest in the product category while minimizing the influence of those who would be unlikely to make a purchase.

16

Forecasting for Hot Prospects

Jorge Alvarez

MBA student Jorge Alvarez helped a speciality printing supplies company make its nationwide sales force better by developing a forecasting model that allowed its sales reps to call on the “right” customers and employ the “right” marketing actions. By ranking and segmenting customers by their potential value to the company, the model allowed their territory-based sales team to share best practices and increase marketing efficiency.

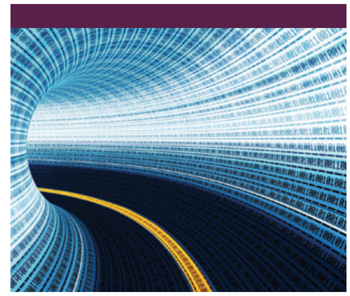
18

One Hybrid Doesn’t Fit All

Shanshan Hu and Gil Souza

Lower operational costs versus higher initial price tag—deciding whether to invest in sustainable technologies for your businesses doesn’t have to be a gamble. A new computational model weighs all the factors of cost and uncertainty—fuel costs, seasonal demand, and maintenance costs—to give businesses an optimal solution to energy efficiency. Taking Coca-Cola Enterprises’ recent investment in hybrid electric trucks as an example, the researchers show how the best investment in sustainable technology may be a hybrid itself.

20



Getting Your Data Warehouse in Order

Mayuri Pardeshi

The ability to make good decisions depends on a strong lineage—data lineage, that is. During her internship at a major consulting company, Master of Science in Information Systems graduate Mayuri Pardeshi helped a health care firm clean up its data warehouse from its disparate origins in various IT systems and through all its transformations on its way to analytics reports. As a result, decision makers gained a renewed confidence in the data’s ability to guide strategic decisions.

Learning by Trading and Learning to Quit

Noah Stoffman
Assistant Professor of Finance
nstoffma@indiana.edu

We know that nonprofessional investors tend to suffer from certain behavioral biases that hurt them financially. Our question was whether these individual investors can learn from their trading experiences to overcome these biases and improve their trading skills. We realized that the existing literature failed to take into account the high rates of attrition among new traders, so we wanted to separate out whether new traders were learning to make better trades or simply learning that they weren t very good at trading.

In the absence of professional training, what causes some individual investors to succeed while others falter? Is trading prowess a matter of inherent ability, or do new traders improve over time? With this study, the researchers track which new investors stay in the game and which ones drop out, offering a clearer picture of the lessons investors are learning. Their findings indicate that while consistently active traders do improve their abilities, the primary insight gained by many new traders is that they would be better off if they quit.

Statement of Problem

The performance of individual investors is a relatively new area of study, with the past decade offering intriguing insights into patterns of behavior. For example, new investors often suffer from a "disposition effect," in which they tend to sell assets on which they have experienced gains and hold assets on which they have experienced losses (appearing to regard the stock only with reference to the value at which they purchased it rather than a broader picture of the stock's history). Some investors, however, perform better than others, perhaps because they are inherently more gifted at trading or alternatively because they eventually acquire the skills to overcome challenges like the disposition effect. This nature-versus-nurture question motivated the researchers to investigate whether individual investors were learning to make better trades or learning about their own existing abilities.

Data Sources Used

The researchers used data from the Nordic Central Securities Depository containing information on all trading of Finnish stocks between 1995 and 2003. More than 22 million trades by individual investors were included. The total number of accounts for which the researchers were able to estimate the disposition effect was 11,979.

Analytic Techniques

The panel structure of the data enabled the researchers to track individual investors over time in order to measure performance, attrition, improvements in portfolio returns, and reductions in the disposition effect.

The researchers ascertained performance persistence among individuals in three ways: first through a regression, then a Spearman rank correlation, and finally by sorting investors into quartiles and plotting the average performance of each of the quartiles over several years. They used a hazard model to measure the disposition effect, running the regression for each investor-year. The researchers also ascertained correlation between the disposition effect and performance by examining investor returns over disposition effect quintiles, and by implementing a difference-in-difference specification using regressions for low- and high-disposition groups to compare the returns of stocks sold at a gain with those that could have been sold at a loss but were not. Lastly, the researchers tested the stability of disposition as an attribute by estimating the rank correlation of account-level disposition coefficients over adjacent time periods.

For their learning model, the researchers began with a simple model to assess whether experience is related to returns, testing experience either in years or in cumulative trades. They estimated the regressions with weighted least squares to account for the estimation error in disposition effect. They added individual fixed effects to the simple learning model, then used it to measure whether attrition was random. Using a two-stage modified Heckman selection model to account for survivorship bias and individual heterogeneity, the researchers employed cross-sectional probit regressions to predict whether or not an individual ceases to trade in a given period. The second stage of the model adjusts for year fixed effects and uses regressions to test whether investors are improving their returns over time.

Robustness checks included substituting the market return for each stock's return to disentangle learning to select stocks from learning to time the market, clustering the standard errors at the individual level, performing tests with 30-day returns instead of actual time of sale, and conducting a bootstrap experiment to keep constant the number of observations used to estimate the disposition effect.

Results

Performance persistence was found to be statistically and economically significant, providing evidence that the most successful investors continue to outperform the least successful investors over time. The disposition effect was found to be statistically and economically quite large, with the median investor 2.8 times more likely to sell a stock whose price is above its cost at purchase than one that has fallen in value since the time of purchase. Investors with the highest disposition effect had substantially lower returns than those in the lowest disposition quintile, particularly when viewed at longer horizons. The difference-in-difference specification also indicated that the disposition effect is costly to investors by showing that high-disposition investors sell stocks that subsequently outperform the stocks they could have sold, but low-disposition investors do not. The persistence of the disposition effect at the individual level was found to be statistically significant across adjacent time periods. Taken together, these results indicate that the disposition effect is an economically important behavioral bias.

Using only the simple learning model revealed positive and significant relationships between trading returns and experience whether measured in years or number of trades completed. An additional year of experience increases average 30-day post-purchase returns by approximately 3 percent, while each 100 additional trades increases returns by approximately 0.8 percent. The disposition effect also appears to decline with experience.

The selection model, however, reveals that poor performance is related to attrition: an investor whose performance is one standard deviation worse than the mean is 15 percent less likely to continue trading. When the characteristics of this attrition are taken into account in the statistical analyses, learning estimates are reduced by as much as 75 percent, with learning as a function of duration shrinking to the point of becoming negligible. These findings imply that learning occurs as a function of the number of trades conducted rather than the passage of time. Moreover, the bulk of observed learning can actually be attributed to performance-related attrition; that is, "learning about ability" is considerably more significant than "learning by doing."

The robustness tests left the results unaffected.

Business Implications

The prevalence of the disposition effect and its correlation with poor returns indicates that behavioral biases are an important feature of financial markets. From an individual standpoint, the inclination of low-ability investors to cease trading may point to a need for screening mechanisms that would enable prospective traders to test their abilities before they begin investing. A broader perspective also calls for more opportunities for training, as the presence of these biases among investors will likely affect asset prices. Individual investors should also be informed that learning correlates with the number of trades rather than the number of years an individual spends investing, which may counter an intuitive inclination to start slowly when beginning to trade.

This study shows how analytics can be used to disentangle the effects of various business factors. What may seem on the surface to be a straightforward relationship—such as individual investors improving over time—can turn out to result from a very different set of behaviors. Analytics permits a type of systematic investigation that can reveal hidden influences and surprising solutions.

Amit Seru, Tyler Shumway, and Noah Stoffman, "Learning by Trading," *Review of Financial Studies,* 23 (2), pages 705–739, 2010.

Taking Stock of SKUs

Ryan Melnikas

MBA '13, majoring in marketing and business analytics

Interned at a leading direct marketing company

rymelnik@indiana.edu

To provide a wide variety of choices to their customers, direct marketers typically hold a large number of SKUs (stock keeping units). Such a lush inventory of choices is good for customers, until it becomes excessive inventory, costing the company money, human resources, and space to track and house the overflow.

During his internship, Ryan Melnikas was charged with curbing a leading direct marketer's growing number of SKUs while still giving customers options and variety. Ryan combined the insights he found in his data analysis with the company's qualitative market research and designed recommendations that addressed how its customers shop, what they value, and what their expectations are.

What were the steps taken to complete the project?

First, I talked to people in different departments (merchandising, inventory, catalog editors, forecasting) and read the company's management reports to define the problem and determine whether the SKU growth could continue without costing the company more to produce and house the products. The company tracked SKU growth and excessive inventory at a high level (total SKUs/excess inventory) across every product line, so I had to dig deeper into which products' SKUs were growing the fastest and which were responsible for the most excess inventory. Interviewing the departments and looking at more detailed information allowed me to address smaller pieces of the overall problem and figure out where making changes would have the biggest impact.

How would you address challenges with data in future projects?

The direct marketing firm had detailed information on its customers. I found it tempting to start my project by poring over spreadsheets and data, but I was able to resist doing that until I had a better sense of what exactly the company's problems were and what information I really needed to be able to make useful recommendations.

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

I found that I spent a lot of time explaining my regression model rather than the implications of what I'd found. I should have put together a very clear cheat sheet to help explain what a regression model can show you and how it can be used to help you figure out what's going on with your business.

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

The courses I took at Kelley helped me identify which data analysis methods would work best with the information I had and kept me focused on the "so what?" of the analysis. Many people can run the models and do the statistical analysis, but the more important piece of data analysis is the "so what?" What do the results mean for your business and how can you use them to make a more informed decision. My professors always stressed that analysis without business insight is useless.

What did you learn from this experience?

Model and data analysis is only useful if it helps you clearly explain something about your business and allows you to make a more informed decision. While analysis helps you reduce uncertainty about a decision, it doesn't make the decision for you. You must be able to sell your recommendations as well as explain how the model works.

Accounting for Price Differences for Identical Medigap Coverage

Haizhen Lin
Assistant Professor of Business Economics
hzlin@indiana.edu

Matthijs R. Wildenbeest
Assistant Professor of Business Economics
mwildenb@indiana.edu

Medigap plans, sold by private insurers, fill the considerable financial gap between what seniors aged 65 and older receive from Medicare and what its high deductibles and co-payments leave on their tabs. Coming in a dozen varieties labeled A though N, Medigap plans are standardized by the Centers for Medicare and Medicaid Services.

However, even though Medigap is a standardized product, the monthly premiums for identical coverage can vary by hundreds of dollars from one insurer to the next, even within the same state.

With this study, the researchers measured price dispersions in Medigap plans and evaluated how search costs and insurer differentiation affected premiums. Finally, the researchers revealed that eliminating search costs by making pricing information readily available could lower premiums and improve consumer welfare.

Statement of Problem

Choosing a Medigap plan is tricky. Seniors not only have to weigh which of a dozen plans will fit their needs, but they also have to shop insurance companies and premiums, which can vary by hundreds of dollars for identical coverage.

Take a Medigap Plan F policy, the most popular choice, as an example: the average coefficient of variation—a measure of the extent of variability in relation to the average of the population—across states for Plan F is 0.27, which is substantial. To translate that into dollars: in 2009, a 65-year-old woman could pay between $1,223 and $3,670 for a Plan F policy in Indiana, depending on which insurer she chose.

With this study, researchers propose a model to explain why price dispersion exists and what could happen if seniors had easy access to both policy and pricing information.

Data Sources Used

The researchers pulled data from two sources. The National Association of Insurance Commissioners' (NAIC) Medigap Experience Files provided data on market shares, allowing researchers to focus on active individual policies, total premiums, and claim volumes. The Weiss Ratings provided detailed pricing information for Medigap plans, including insurers, location, plan type, gender, age related to a plan, rating method, and smoking status.

Merging these datasets gave researchers information on prices and sales at the state level, for a total of 4,704 observations, accounting for more than 80 percent of the total policies sold between 2007 and 2009.

Analytic Techniques

The researchers used this data on Medigap plans to explain price dispersion and to estimate the effect of search frictions on demand parameters.

Measuring price dispersion across insurers, researchers established a coefficient of variation across insurers' rating methods. Because high-priced plans could attract fewer seniors, the researchers corrected for market share by weighting the coefficient of variation by the number of policies sold.

Researchers then used the data to determine why these variations in premiums occur. First, they ran a regression model to see if high-priced insurers served more high-risk seniors, which would suggest that cost differentials determine pricing.

Next, they considered other dimensions of product differentiation—the insurers' branding, billing services, reputation, financial ratings—as a possible explanation for price dispersion. They evaluated each insurer's price against the overall price distribution in a state and then further compared the R-squared of a regression of premiums on plan dummies, rating methods,

What's interesting about the Medigap market is that all the plans, regardless of which insurer administers it, are the same. Government regulation has made things easier for seniors by standardizing these plans, so that whichever of the 12 plans they choose, seniors know exactly what they're getting. However, there's a huge price variation in the same market across different private insurers. This means that seniors are paying a different price for exactly the same apple.

and state dummies to the R-squared of a similar specification that also includes insurer fixed effects.

Finally, they evaluated whether high search costs might dissuade seniors from robust comparison shopping—and give an incentive for insurers to increase premiums. They charted price versus market share to see if it resulted in a negative and monotonic relationship.

From those results, they took both product differentiation and search costs for their theoretical and empirical framework, estimating the equation using constrained two-stage least squares.

The researchers then used their search model, where seniors would compare anywhere from one to five different insurers, against a model that assumes seniors have full information. To do that, they estimated a standard conditional logit model of demand.

Results

Researchers found substantial price differences across plan types and states. They sorted those variations in premiums by the insurers' own rating methods: attained-age, where premiums increase with the age of the policy holder; issue-age, where premiums are charged based on the age at enrollment; and community-based, where premiums are uniform across all subscribers in the community.

For all three rating methods, the coefficient of variation is substantial, with attained-age (the most popular) and issue-age rating methods showing slightly higher variation than the community-based rating. In a few states, plan types show hardly any price variation; however, the coefficient of variation can be as high as 0.65 for the issue-age rating method. Most plan types have an average coefficient of variation that exceeds 0.20.

So why is there such price variation for a standardized product? Of the three determinants, researchers found that search costs (the time, money, and effort it takes to conduct a robust search of insurers, plans, and premiums) and product differentiation (the firm's age, service, and financial rating) each play a role in explaining the price differences.

The researchers' model also studied what happens to prices and market shares when search costs are lowered. While researchers estimated that search costs fall between $22 and $49, they found that decreasing those costs resulted in up to a five percent decrease of premiums, or an average savings of up to $87. With that savings on search costs and a resulting reduction in premiums and expansion in market share, they found that overall consumer welfare could increase by up to $321, a magnitude of up to four times the average price decrease.

Business Implications

The hard work of shopping around for a Medigap provider is a good investment for seniors. For insurers, the research model shows that providing those consumers with free and ready access to information about plans and premiums can be a competitive advantage. Once you remove the search frictions and provide transparency about premiums, insurers will likely adjust pricing structures, market shares will increase, seniors will save money, and consumer welfare will rise.

Beyond the world of supplemental insurance, the research model applies to any environment where price dispersion exists and search frictions and product differentiation play a role in purchasing decisions.

Haizhen Lin and Matthijs R. Wildenbeest, "Price and Search in the Medigap Insurance Market," 2012. Working paper.

Big Data and Health Care Insurance

Natalie Basich

MBA '13, majoring in finance and business analytics

Interned at a major health care insurer in the Midwest

nbasich@indiana.edu

Many industries face the challenges of big data and an ever-expanding number of applications, databases, and data warehouses. But the health care industry and health care insurers grapple with a unique set of tests. Once collectors of information, health care insurance companies now use data to make informed decisions, and they do so in a rapidly changing regulatory environment.

During her internship, Natalie Basich worked with a major health care insurance company to analyze how its retail centers were performing. The company wanted a more robust financial model that pulled multiple years of data and allowed it to view sales projections as well as assess past, current, and potential performance.

What were the steps taken to complete the project?

I spent time with the current financial model to understand, line by line, what was used to develop the forecasting numbers. Then I talked to the stakeholders—the developer of the model, my manager, different department heads—to understand what their goals were, what the model would be used for, who would be using it, and what mix of products I needed to consider along with the associated contribution margins, for each market.

Once I built the model, I reviewed it with my manager and the finance team. My analysis not only looked at the retail center program as a whole but also included market level analysis. Through that financial model, the company could evaluate the drivers of value for each market and see what was making the most impact on profitability.

How would you address challenges with data in future projects?

You have to know who to ask and be persistent in getting what you need. I discovered that when you develop relationships, you get information faster and easier. You also need to fully understand the end goal of the data and how it will be used so you can best present your results.

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

Explicitly list all assumptions you've made, so they are clear to both you and senior management and you can easily update information without having to question the model. I would also add, to always think about the "so what?" Why does the information and analysis matter? That answer will help you present your findings in the best way.

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

I knew the right analytical tools to use and what you can do with data once you have it. It's not just about going into the analytics; it's about asking the right questions and not being afraid to ask them over and over again.

What did you learn from this experience?

From my internship, I learned that it takes longer to understand what is needed and gather it from different parts of the organization than to analyze the data. Relationships and the right approach are key to speeding up the process. Data is a big buzzword in health care and this internship helped me develop competitive skills.

Since I'll be working in the health care field after graduation, I'll be able to take those lessons directly into my job and my career. And I'm excited about this field. There are huge changes coming in regulation, and I like the idea of being a part of solving the challenges behind the changes.

Identifying Unmet Demand

"My main research interest is in the advertising of new products. If you re a marketer, you need information not just about what people buy but also about the primary motivations that drive their preferences. A problem that I saw with research in this area was that the relative importance of different product beliefs were being measured across one collective sample, when in fact not every product attribute was relevant to every consumer. The main contribution of this study was to demonstrate a means of combining multiple statistical methods to more accurately identify demand for product attributes across a heterogeneous group of consumers.

Sandeep R. Chandukala
Assistant Professor of Marketing
sarchand@indiana.edu

Unmet demand refers to customers' existing needs or preferences that are not being satisfied by the products on the market. Information about unmet demand provides crucial direction for marketers and product developers, but is difficult to identify through market research. Low sales in a category, for example, could be related to the perceived absence of desired product attributes, or could be due to any number of other factors, including price and promotion strategy.

Conjoint analysis, which enables researchers to obtain information about the value consumers place on hypothetical attributes and benefits, can offer a more meaningful picture of unmet demand. Results can be skewed, however, by respondents who have little interest in the product category or whose preferences vary greatly across the sample.

With this study, the researchers propose a model for separating out such effects within the context of conjoint analysis. They test the model using data collected in a national survey of preferences and beliefs regarding toothpaste. The model and empirical application demonstrate a strategy for identifying the unmet demands of heterogeneous groups of consumers.

Statement of Problem

Marketers and product developers require information about consumers' unmet needs and preferences. To be meaningful, such information must relate to consumers who have an interest in purchasing the product and an opinion about its attributes. Moreover, the data must accurately capture heterogeneity in the population so that disengaged respondents do not distort the data and groups of consumers with opposing preferences do not effectively cancel each other out. A useful model for identifying unmet demand must separate out groups of consumers according to their product beliefs and values in order to provide a more nuanced picture of potential demand.

Data Sources Used

Working with the marketing research firm Yankelovich Partners, the researchers collected data on toothpaste preferences from a national sample of 757 respondents. Through a conjoint study, respondents indicated their preferences for a set of 30 product attributes and benefits (e.g., "Helps prevent cavities"; "Fights bad breath"). They also indicated their level of agreement with each of six product belief statements (e.g., "Toothpastes are too strong tasting"; "Toothpastes don't really work to prevent dental problems").

Analytic Techniques

To construct their model for unmet demand, the researchers began with a random-effects specification for heterogeneity describing variation in product beliefs. They added a variable selection matrix in order to capture instances where a respondent's behavior is affected by the variables and diminish the noise of those for whom those variables play no role. Finally, the researchers employ a multivariate normal distribution for the random effects and introduce a mixture component model to allow for distinct response segments. Bayesian estimation of the model introduced a latent variable pointing to respondent segment membership.

The ranking conjoint task of the empirical study led to an exploded logit form of complete preference ordering. Markov chain Monte Carlo methods were used to estimate the model parameters.

The researchers assessed fit statistics for eight model variations, beginning with a standard random-effects model and adding specifications with each iteration. Models 7 and 8 contained all the components of interest to the researchers, combining covariates, normal components, and variable selection into one model. Comparisons of log-marginal densities were used to identify the best in-sample fit.

Applying the toothpaste data, the researchers calculated covariate relationships between respondents' product beliefs and their preferences toward product attribute components.

Results

Model 7, a variation of the full model proposed by the researchers, offered the best fit to the data. Compared to the baseline model (model 2, which added product beliefs to random effects), model 7 resulted in a 17 percent improvement in fit.

Though the researchers were prepared to account for multiple response segments with distinct needs (model 8), the data revealed that only two components were required for the model. In essence, one component reflected those respondents who indicated their preferences and beliefs, and the other component captured those who were disengaged from the study. By separating out respondents who did not appear to hold meaningful views on toothpaste, the researchers' heterogeneity model offered a clearer picture of the consumers of interest to marketers.

Applying the model to the toothpaste data yielded a number of insights into consumer preferences, such as the association of unmet demand for medical benefits with beliefs that toothpastes are too strong tasting, irritate one's mouth, cost too much, and generally do not work. As another example, respondents who concurred with the statement that breath freshening does not last long enough were more likely to value heightened taste attributes (e.g., fresh tasting, gives your mouth a tingle). This association suggests a potential demand for products offering longer-lasting taste sensations.

Business Implications

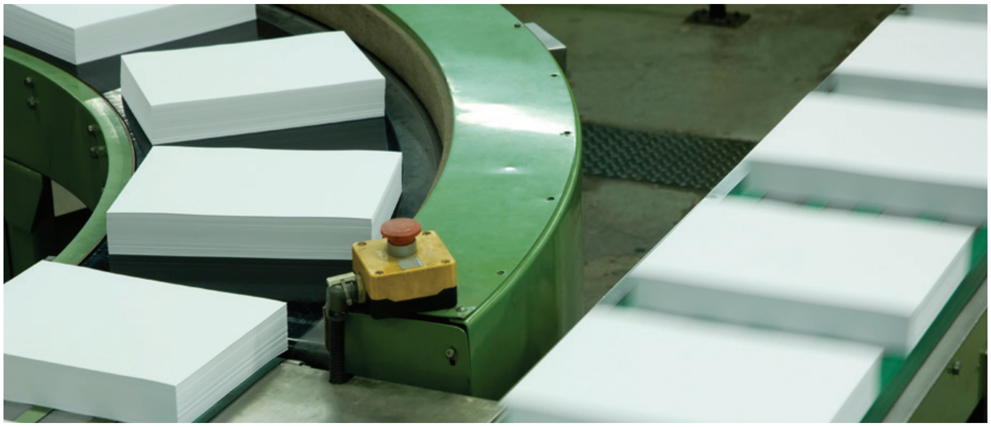
Identifying unmet demand enables businesses to adjust their marketing strategies to highlight the attributes consumers are seeking, or, if the product does not exhibit those attributes, to reformulate or diversify to meet those needs.

The model of heterogeneity proposed in this study can be used to better understand market segments and their unmet product demands. In particular, the model offers a means of removing the distortion caused by disengaged respondents, enabling marketers and product developers to concentrate on the consumers who would actually have an interest in buying their product.

This study demonstrates the capability of business analytics to disentangle multifaceted data to deliver straightforward and applicable information about consumers.

Sandeep R. Chandukala, Yancy D. Edwards, and Greg M. Allenby, "Identifying Unmet Demand," *Marketing Science,* 30 (1), pages 61–73, 2011.

Forecasting for Hot Prospects



Jorge Alvarez

JD/MBA '13, majoring in business analytics and marketing

Interned at a large specialty printing supplies company

joralvar@indiana.edu

For a large speciality printing supplies company, territory-based sales initiatives rely on reps' industry expertise, customer relationships, and entrepreneurial drive to entice and educate customers about its innovative substrate applications and how they can benefit their printing businesses.

Jorge Alvarez's summer internship with a large speciality printing supplies company focused on bringing data and analytics to that nationwide field to help its sales force identify the most valuable customers, determine effective terms of engagement, and drive new marketing efficiencies.

How did the forecasting model benefit the speciality printing supplies company? The company had already implemented a new Enterprise Resource Planning (ERP) system to unify its business operations. Building a forecasting model would help it rank and segment customers based on their predicted value and match those segments with the most effective marketing actions.

And it's working. With that forecasting model, it deployed the customer ranking and segmentation to 23 sales territories to help improve targeting and marketing plans—and identified a $2.8 million opportunity in net revenue among active customers.

What were the steps taken to complete the project?

Although data from legacy systems had already been transferred to the new system, other data relevant to the project existed in various internal and external reports that had to be collected and consolidated.

Once we had that initial data set, we then worked with the company's leadership, managers, and sales reps to identify key customer characteristics and marketing actions for ongoing analysis. Then, we developed a causal map that linked key characteristics and actions to their primary outcome measure, the Margin Before Freight (MBF).

Using active customers, we ran regressions on the causal maps and built MBF forecasting models that ranked and segmented customers by opportunity—growth, core, and slow. We then built a messaging matrix that we could share across the sales force to help the company target the most impactful marketing actions. For example, "growth" customers responded more strongly to sales calls, while "core" customers responded best to sampling.

To help increase marketing efficiency, we used that data to modify routing schedules to reflect the appropriate allocation of time for each customer segment. For example, "growth" customers should receive 60 percent of total sales calls whereas "core" and "slow" customers should receive 30 and 10 percent, respectively.

How would you address challenges with data in future projects?

Gathering the data was the most time-consuming phase of the project. It's important to take the time to consider what information you need now and even in the future—and then solicit that data from customers and sales representatives up front.

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

Preparing data for analysis is crucial for analytic techniques to work. In addition, there are many different types of transformations possible. Time must be allocated to try different combinations of transformations and analytic techniques.

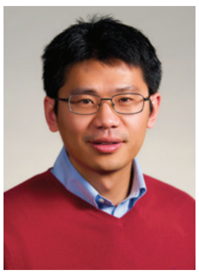
How did your courses at Kelley prepare you for this project?

My business analytics courses gave me a framework for analyzing business problems—starting with understanding the business before moving on to the data and analysis. The problem-based teaching approach provided me with valuable experience applying various analytic techniques to solve real-world problems.

What did you learn from this experience?

Spending time up front with the data—gathering, cleaning, exploring, transforming—saves time on the back end. Applying analytic tools on successfully prepared data is a pleasure and takes relatively little time. Instead, spend time on the back end focusing on what the analytics mean.

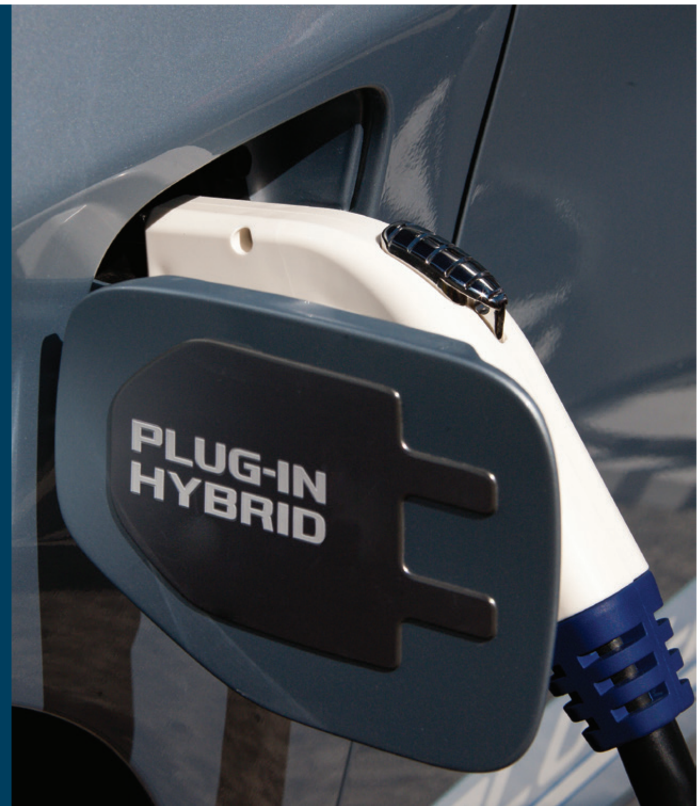
One Hybrid Doesn't Fit All



Shanshan Hu
Assistant Professor of Operations and Decision Technologies
hush@indiana.edu

Gil Souza
Associate Professor of Operations Management
gsouza@indiana.edu

"Firms struggle to justify their investment in sustainable technology. They know, for example, that hybrid electric vehicles (HEV) are more fuel efficient, but HEVs also cost 50 percent more than regular diesel trucks. Most decision models stop there, without accounting for volatile fuel costs, regular maintenance costs, how often these vehicles are used, or seasonal demand. Our model takes all of these factors into account and gives firms an optimal solution, whether they re investing in delivery trucks or any other energy efficient technology.



When it comes to investing in sustainable technologies, it's not all or nothing. For firms looking to shift their production and distribution into energy efficient technologies, a hybrid approach is often best.

When businesses make decisions about whether to invest in traditional or sustainable technologies, they typically use a net present value comparison, which ignores the volatility of fuel costs and uncertainty factors and essentially gives those firms an either/or solution—you're either all in with one technology, traditional or sustainable—or not.

With this study, researchers weigh acquisition costs against energy efficiency, but also take into account the uncertainty of fuel costs, the fluctuations in seasonal demand, and the impact of heavy use and deterioration—and show firms how to ease into energy efficiency with a mixed portfolio of sustainable and traditional technologies.

Statement of Problem

Firms shopping for sustainable technologies for the production and distribution of their products face a challenge. While these sustainable technologies may save energy and money over the long term, firms must stare down steep acquisition costs.

Deciding whether to invest in sustainable technologies, however, is a little more complicated than just weighing the immediate acquisition price tag versus long-term operating costs, and a little more complex than previous decision-making models admit. A more useful methodology factors in uncertainty and offers firms a dynamic solution to their investments in energy efficiency.

Take for example the choice Coca-Cola Enterprises had to make when they decided to start replacing their fleet of diesel delivery trucks with diesel-electric hybrid vehicle (HEV) trucks. How they choose to invest those dollars also depends on volatile fuel costs, usage-based deterioration, and seasonal demand.

Data Sources Used

To build their observations for Coca-Cola Enterprises (CCE), researchers used CCE's historical maintenance costs and detailed demand data for a particular market to estimate CCE's seasonal parameters. Along with CCE's data, researchers then simulated diesel prices using historical data provided by the U.S. Department of Energy. They used those costs to calibrate a standard stochastic model that simulates diesel prices dynamically.

Analytic Techniques

This study provides a tractable decision support model to help firms make better choices about how much to invest in new, sustainable technologies that factor in the uncertainties of fuel costs and demand, as well as fixed maintenance costs and deterioration. Using an infinite-horizon stochastic dynamic optimization—or dynamic programming—researchers can define the optimal policy, at any time, based on a reasonable evolution of diesel prices, and under a long-term horizon.

Using regression techniques, researchers also created models to predict seasonal demand and dynamic diesel prices.

Results

The researchers lay out a methodology that finds the optimal portfolio for capacity investments that not only weighs acquisition, fuel, and maintenance costs but also considers how often vehicles are used and how to best maximize fuel efficiency.

In the case of CCE's delivery trucks, they found that, optimally, the company should include both hybrid electric vehicles (HEV) (54 percent) and conventional diesel trucks (46 percent) in their capacity portfolio. CCE can use HEV trucks to meet their average, baseline demand, and then deploy diesel trucks to supplement the delivery fleet during their peak demand seasons.

The model or CCE's observations can easily transfer to other companies with seasonal fluctuations. For those rare companies with a steady stream of demand all year long, researchers suggest a total investment in HEV, or the technology with lower energy consumption, over those technologies that may cost less but require more energy, and therefore incur more cost, to operate.

For companies that have made a commitment to sustainable technology across the board and regardless of acquisition costs, the researchers found that the total discounted cost (acquisition and operating costs) of an HEV-only fleet is 1.24 percent below optimal cost. Environmentally friendly technologies are also economically attractive.

Business Implications

As energy efficient technologies enter the marketplace, firms are faced with a choice. Do they invest? Do they stay put? And if they do invest, how much of their capacity portfolio should they hand over to these new technologies? This study gives those firms a framework with which to make these decisions, more dynamically, more easily, and more accurately than in the past.

With this study, the researchers provide a more nuanced ISD type (Invest/Stay Put/Disinvest) policy that meets each realization of demand and fuel costs in a period. Basically, the researchers looked at the capacity of conventional technology versus the capacity of sustainable technology and show that the optimal solution is defined by four curves. This model allows firms to quickly determine their curves and their optimal solution, reducing computational time from days to minutes.

Using CCE's data on seasonal demand and maintenance costs alongside historical data on diesel prices, they show what an optimal, real-world capacity portfolio looks like. Companies that have a similar profile can then draw on those insights to inform their own decision making.

In addition to this computational model, the researchers offer eight observations that they drew from the CCE example. These observations cover what happens and how firms should act when fleet sizes are reduced or substituted, when diesel prices or demand spikes, or when fleets are highly utilized.

For policy makers, these observations also offer insights into the efficacy of environmentally friendly government incentives. Neither tax credits nor fixed fuel taxes are cost-effective policy measures to reduce carbon footprint, especially for firms with high seasonal demand.

Dynamic Capacity Investment with Two Competing Technologies, Wenbin Wang, Mark Ferguson, Shanshan Hu, and Gilvan C. Souza, *Manufacturing & Service Operations Management,* forthcoming.

Getting Your Data Warehouse in Order

Mayuri Pardeshi
MSIS '12

Interned at a major consulting company

pardeshi@indiana.edu

During her summer internship at a major consulting company, Mayuri Pardeshi worked on a project with one of its Fortune 500 healthcare clients. Years of mergers and acquisitions had left it with a motley collection of IT systems—and a data warehouse in disarray.

In fact, many Fortune 500 companies struggle to develop sophisticated and reliable data warehouses. These companies have come to understand the power of analytics, but realize that before they can reap the benefits, they first have to get their data warehouses in order.

Working with the consulting company's information management service line that dealt with business intelligence, Mayuri and her team established a data lineage to give the client end-to-end understanding of their disparate information systems and how their data flowed and transformed, from data collection to target analytics reports.

How did data lineage benefit your client?

The client had experienced that, more often than not, the users of analytics reports questioned the reliability of the data. Data lineage helped those decision makers, which included C-level executives, understand and visualize data flow and transformations, giving them the confidence that the data and analytics reports could inform business and administrative decisions.

What were the steps taken to complete the project?

We had extensive conversations with the client's key stakeholders and the consulting company's subject matter experts to gather requirements and understand the targeted outcome. We then mapped the client's current technical landscape and compared it to industry standards. From there, we recommended establishing data lineage to improve data tracking and then

conducted impact analysis. We implemented that recommendation first as a pilot to generate proof-of-concept, and then later, we scaled it into a full-blown project solution and implementation plan for the client.

What tools can be used for managing data lineage?

Data lineage is established through metadata management using metadata repository tools, which are widely available in the marketplace. When we shopped for the right tool, we looked for one that would fit with client's existing systems and the vendor support provided.

How is data lineage important for analytics?

Data analytics allows organizations to analyze past performance and forecast future outcomes. However, businesses can only draw conclusions from that data if they start with accurate, consistent, reliable data. Data lineage helps businesses understand how their data was collected and how it has been transformed.

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

Kelley gave me confidence and competency to jump in and provide value, right away. Because of my MSIS courses, I already knew the language—the core concepts, terminology, and processes—of data warehousing and analytics. This helped me quickly understand the client's problems and gave me the confidence to provide valuable input throughout the project. With the skills I learned from the project management class, I created a project plan for the client without any training and with minimal supervision.

What did you learn from this experience?

A key takeaway for me was the importance of humility and continuous learning. I came into my internship with the latest concepts and industry-valued tools in analytics and data warehousing. However, I realized that even with the latest knowledge, it's not possible to be prepared 100 percent for real-world challenges.

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; Dean's Research Professor; Director, Institute for Global Organizational Effectiveness

Goker Aydin

Associate Professor of Operations & Decision Technologies

Hillol Bala

Assistant Professor of Information Systems

J. Doug Blocher

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

Kurt M. Bretthauer

Professor of Operations & Decision Technologies; Kimball Faculty Fellow

Raymond R. Burke

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

Kyle Cattani

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

Sandeep Chandukala

Assistant Professor of Marketing

H. Sebastian "Seb" Heese

Associate Professor of Operations Management

Randy Heron

Professor of Finance; Roger & Barbara Schmenner Faculty Fellow

F. Robert Jacobs

Professor of Operations Management; Chase Faculty Fellow

Vijay Khatri

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

Haizhen Lin

Assistant Professor of Business Economics

Anne P. Massey

Chairperson, Doctoral Programs; Professor of Information Systems; Dean's Research Professor

Philip T. Powell

Chairperson of Kelley Direct; Clinical Associate Professor of Business Economics and Public Policy

Alfonso J. Pedraza-Martinez

Assistant Professor of Operations & Decision Technologies

Jeff Prince

Associate Professor of Business Economics

Babur De los Santos

Assistant Professor of Business Economics

Kim Saxton

Clinical Assistant Professor of Marketing

Scott B. Smart

Associate Chair of the Full-Time MBA Program; Clinical Associate Professor of Finance; Whirlpool Finance Faculty Fellow and Director of Corporate Finance Academy

Daniel C. Smith

Former Dean; Professor of Marketing; Clare W. Barker Chair of Marketing

Ash Soni

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

Gilvan "Gil" C. Souza

Associate Professor of Operations Management

Noah Stoffman

Assistant Professor of Finance

Ramesh Venkataraman

Chair of the MSIS Program; Associate Professor of Information Systems; Whirlpool Corporation Faculty Fellow

Munirpallam Venkataramanan

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

James M. Wahlen

Professor of Accounting; James R. Hodge Chair

Rockney G. Walters

Professor of Marketing

Matthijs R. Wildenbeest

Assistant Professor of Business Economics

Wayne L. Winston

Professor of Operations & Decision Technologies; John & Esther Reese Professorship

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

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

Nonprofit Organization
U.S. Postage
PAID
Bloomington, IN
Permit No. 351

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