

How Non-Scientific Granulation Can Improve Scientific Accountics

Bob Jensen at **Trinity University**

Accountics is the mathematical science of values.

Charles Sprague [1887] as quoted by McMillan [1998, p. 1][\[NH1\]](#)

For me, she says, "this really showed the beauty of science, that you can have this personal experience that isn't reflected in big data."

Jennifer Jacquet as quoted by Robin Wilson, *Inside Higher Ed*, October 22, 2012 ---

http://chronicle.com/article/The-Hard-Numbers-Behind/135236/?cid=at&utm_source=at&utm_medium=en

In quantitative finance and accountics science, we call important factors not reflected in big data, or otherwise that cannot be scientifically quantified, "black swans" or "causal factors."

The purpose of this paper is to show how non-scientific research cannot only add value to quantitative studies, non-scientific research can find granules of causation that cannot be discovered in most quantitative studies.

Introduction

Example 1: An Accountics Science Illustration

Suggested Interview Research Granulation Searches for Causality

What granules of causation might be discovered in the interviews of clients who changed auditors after the ChuoAoyama audit firm scandal became revealed to the public?

Example 2: granular Non-Science Research and Database Variables

Qualitative Research

Mechanical Turk and the Limits of Big Data: The Internet is transforming how researchers perform experiments across the social sciences

Introduction

A simple operating definition of accountics science research is any accounting research focused primarily upon analysis with mathematical equations and/or statistical inference tables. Since the 1980s the leading academic accounting research journals are almost exclusively accountics science journals that seldom publish non-accountics articles or even commentaries on published research.

“An Analysis of the Evolution of Research Contributions by *The Accounting Review*: 1926-2005,” (with Jean Heck), *Accounting Historians Journal*, Volume 34, No. 2, December 2007, pp. 109-142.

<http://www.trinity.edu/rjensen/395wpTAR/Web/TAR395wp.htm>

Nearly all of the empirical accountics science articles fall into two types:

1. Multivariate models (especially regression) models using purchased databases such as Compustat, Audit Analytics, CRSP and other very large commercial databases.
2. Behavioral experiments that usually use students as surrogates for real-world decision makers.

As a result, accountics scientists seldom leave the campus to obtain databases used in multivariate models and statistical inference analysis.

Typical Absence of Causal Analysis in Accountics Science

Accountics scientists rarely do causation analysis. Their multivariate regression and other data mining outcomes rely upon correlation and/or tenuous causation inferences rather than direct search for causation. Findings dependent upon experiments student behavior must generally be extrapolated tenuously to untested hypotheses concerning real world decision making and risk-taking behavior. There are of course some exceptions, but causal discoveries are generally not found in accountics science. The old saying that correlation is not causation is nagging limitation of so many, many accountics science findings.

Typical Presence of Causal Analysis in Granular Non-Science (protocol analysis, interviews, surveys, cases, anecdotes, and field studies)

Whereas accountics scientists have to indirectly infer/assume causality, granular studies focus more directly upon uncovering causes of outcomes. It's obvious that it's possible to add a great deal to confirm or disconfirm accountics science findings with granular non-science findings. But therein lies the problem, because granular findings are often subjective, anecdotal, and possibly cherry picked. This creates doubt about the reliability and robustness of their findings. More importantly, granular studies may be so expensive that only small samples are practical, including samples of just one company or one executive.

This does not mean that, when doing accountings science research on large samples, the researchers did not do granular research supplements such as preliminary field research or small sample (e.g., case) studies. However, due to the limitations of the non-scientific nature of granular research, accounting research journal referees may require deleting any mention of the granular research outcomes that can be very misleading due to their non-scientific nature and potential biases.

Example 1: An Accountings Science Illustration

Granularity --- <http://en.wikipedia.org/wiki/Granularity>

Granularity is the extent to which a [system](#) is broken down into small parts, either the system itself or its description or observation. It is the extent to which a larger entity is subdivided. For example, a yard broken into inches has finer granularity than a yard broken into feet.

Coarse-grained systems consist of fewer, larger components than **fine-grained** systems; a **coarse-grained** description of a system regards large subcomponents while a **fine-grained** description regards smaller components of which the larger ones are composed.

The terms **granularity**, **coarse**, and **fine** are relative, used when comparing systems or descriptions of systems. An example of increasingly fine granularity: a list of nations in the [United Nations](#), a list of all states/provinces in those nations, a list of all counties in those states, etc.

The terms *fine* and *coarse* are used consistently across fields, but the term *granularity* itself is not. For example, in [investing](#), *more granularity* refers to more [positions](#) of smaller size, while [photographic film](#) that is *more granular* has fewer and larger chemical "grains".

For example, in the granular [protocol analysis](#) the researcher might record the verbalized thoughts of a single real-world decision maker while making a real-world decision such as how many equity shares of a company to add to a client's portfolio or a loan officer's decision as to the maximum amount of credit to extend to a borrower. Interviews and case studies typically do not entail protocol recordings of actual decision making thought processes, but interviews, surveys, and case studies often rely upon self-reporting by decision makers on how decisions are reached ---

<http://www.trinity.edu/rjensen/000aaa/thetools.htm#Cases>

Purchased databases such as Compustat do not contain the level of granular detail found in protocol and case analyses. At the same time, protocol and case studies do not contain the sample sizes of purchased databases of much coarser variables. This is why case studies are sometimes called non-scientific "small-sample studies." There is no basis for scientific inference from small sample studies except in the rare instance where even one anomaly will destroy a hypothesis.

A recent accountings science study suggests that audit firm scandal with respect to someone else's audit may be a reason for changing auditors.

"Audit Quality and Auditor Reputation: Evidence from Japan," by Douglas J. Skinner and Suraj Srinivasan, *The Accounting Review*, September 2012, Vol. 87, No. 5, pp. 1737-1765.

We study events surrounding ChuoAoyama's failed audit of Kanebo, a large Japanese cosmetics company whose management engaged in a massive accounting fraud. ChuoAoyama was PwC's Japanese affiliate and one of Japan's largest audit firms. In May 2006, the Japanese Financial Services Agency (FSA) suspended ChuoAoyama for two months for its role in the Kanebo fraud. This unprecedented action followed a series of events that seriously damaged ChuoAoyama's reputation. We use these events to provide evidence on the importance of auditors' reputation for quality in a setting where litigation plays essentially no role. Around one quarter of ChuoAoyama's clients defected from the firm after its suspension, consistent with the importance of reputation. Larger firms and those with greater growth options were more likely to leave, also consistent with the reputation argument.

• • •

To test whether the F2006 auditor switches away from ChuoAoyama are unusually frequent, we estimate a logit model of factors that explain auditor changes. The control variables are drawn from previous research on auditor switches and include firm size (log of total assets), growth (percentage change in total assets), leverage, change in leverage, profitability (ROA), a loss dummy, U.S. listing, keiretsu inclination, auditor industry expertise, earnings quality as measured by accruals, whether the firm completed an M&A transaction in the preceding two years, and industry fixed effects.²² We provide details of data sources and variable definitions in Appendix B. The keiretsu inclination variable measures whether and to what extent these firms are part of the large corporate groups common in Japan (e.g., Aoki et al. 1994; Hoshi and Kashyap 2001).

We include dummy variables for whether the client is a ChuoAoyama client (CA), for fiscal year 2006 (F2006), and for the interaction of these two dummies (CA_F2006). The interaction variable is our primary interest because it measures the extent to which client firms switch away from ChuoAoyama in fiscal 2006, the period in which we argue that auditor reputation drives switching.

• • •

Our results are largely consistent with the importance of reputation effects. We find evidence that a relatively large number of ChuoAoyama's clients left the firm for other auditors as the seriousness of ChuoAoyama's quality problems became evident. The rate of client turnover at ChuoAoyama in fiscal year 2006, before it became apparent that the firm would be shut down but after audit-quality questions had been raised, was substantially higher than would otherwise be expected, consistent with clients leaving once the firm's reputation for quality was seriously diminished. Moreover, we find that the likelihood of switching is higher for larger clients and clients with higher market-to-book ratios, characteristics associated with a demand for higher-audit quality, and lower for firms with greater managerial ownership, indicating a lower demand for audit quality in such firms. Clients that moved to Aarata were also larger, with higher market-to-book ratios, a greater extent of cross-listing, and higher foreign ownership. These switches are not the result of clients following their audit teams to new auditors. Our event study results weakly support the auditor-quality argument, but are likely to lack power because questions about ChuoAoyama's audit quality were revealed over an extended period.

Our conclusions are subject to two caveats. First, we find that clients switched away from ChuoAoyama in large numbers in Spring 2006, just after Japanese regulators announced the two-month suspension and PwC formed Aarata. While we interpret these events as being a clear and undeniable signal of audit-quality problems at ChuoAoyama, **we cannot know for sure what drove these switches** (emphasis added). It is possible that the suspension caused firms to switch auditors for reasons unrelated to audit quality. Second, our analysis presumes that audit quality is important to Japanese companies. While we believe this to be the case, especially over the past two decades as Japanese capital markets have evolved to be more like their Western counterparts, **it is possible that audit quality is, in general, less important in Japan** (emphasis added) .

These are very honest admissions that extend to the entire history of most accountings science studies. The Skinner and Srinivasan inference that the audit firm's loss of reputation caused a third of the clients to switch is very tenuous and superficial since two thirds of the clients remained loyal and did not switch. This suggests at a minimum that reasons for switching are far more complicated than assumed by Skinner and Srinivasan.

In other words, like most accountics science papers causality that is inferred could be slightly off base or largely off base. There's no way of knowing because the accountics models cannot see the granules of causation. This is where non-science granular research might be of some help.

Non-science [protocol analysis](#) is not of much use as a follow up to the Skinner and Srinivasan study since changing auditor decisions in this study are one-time past historical events for the PwC-affiliated ChuoAoyama auditor client switching and are not frequently repeated observable decision events such as portfolio decisions of a trust investor or a bank's decision to set a credit limits of borrowers.

Non-science mail survey research where the clients of the ChuoAoyama audit firm at the time are surveyed are not likely to be of much use since there's no incentive for those clients to respond at all, and if some of them respond the results will be questionable since the respondents quite likely to provide answers they think the researchers and public want to hear.

Suggested Interview Research Granulation Searches for Causality

So what could be the reasons for switching away from the ChuoAoyama audit firm other than the firm's auditing scandal and resulting loss of reputation?

What comes to mind are those clients who may have used the audit firm scandal as an excuse rather than a reason to switch to another audit firm.

The ChuoAoyama audit firm is a Japan Big Three audit firm (PwC) that is likely to be at or near the top in terms of Japan's audit fees. Perhaps clients initially chose this audit firm affiliated with PwC to enhance their own appeal to investors in Japan's fledgling equity markets. After the fact, it's very difficult to change such an audit firm without possibly having a huge negative impact on the client's stock price and credit standing.

But some clients in retrospect may be very unhappy about the high audit fees relative to what it views as the quality of the audit and the importance of a ChuoAoyama audit for equity prices relative to less expensive audit firm alternatives.

Thus we cannot rule out that some proportion of the clients that changed audit firms did so because the highly publicized scandal concerning the ChuoAoyama audit firm gave them an excuse to switch with positive publicity rather than negative publicity.

Switching audit firms as an excuse rather than a reason is entirely consistent with the accountics science findings of the Skinner and Srinivasan article published in *The Accounting Review*.

So how could non-scientific granulation studies provide added value to the Skinner and Srinivasan scientific findings?

I would look toward personal face-to-face interviews on causality. Case research could also be used, but the number of clients that can reasonably be expected to participate is probably larger for focused interviews vis-à-vis more extensive cases. Interviews and case research has some advantages over mail survey research if the interview/case researchers meet on-site and face to face with respondents. There's a better chance of getting at the real causes, although there still might be reluctance to have those real causes publicized. Interviewers should probably assure respondents that their responses will remain anonymous.

Ideally the clients included in the interview studies would be randomly picked if the entire population is not interviewed. One reason this type of interview research is non-scientific is that there's no accounting for clients that absolutely refuse to participate.

The interviewers ideally should be highly respected in the Japanese business community and be fluent in the Japanese language. This is why they should probably be current or former Japanese citizens. However, there might be exceptions such for well known case researchers like Robin Cooper who is highly respected in Japan for his case research and writing focused on Japanese companies.

Interview research experts should decide how best to phrase the key questions and where to couch them in the entire interview. There are many nuances in interview research to be considered when trying to get potentially sensitive answers. This is where promises of anonymity may be extremely important.

There are the usual scientific arguments against interview and case research, including the possibility of cherry picking the clients to be studied. The clients might not be entirely truthful about sensitive causal factors. And the number of clients studied is miniscule relative to the number of clients included in the accountants science study. However, this may be less of a problem in the Skinner and Srinivasan since there is a relatively small population of clients who switched audit firms.

What granules of causation might be discovered in the interviews of clients who changed auditors after the ChuoAoyama audit firm scandal became revealed to the public?

Possible Answer 1

Skinner and Srinivasan suggest (but could not conclude) that nearly all the clients that changed audit firms did so because of the possible adverse effect keeping the scandalous audit firm would have on cost of capital increases for clients who used a scandal-ridden audit firm. But this suggestion is weak because it cannot explain why a majority of the ChuoAoyama audit firm's clients did not switch auditors.

Possible Answer 2

Skinner and Srinivasan did not consider the possibility that some clients switched auditors because the scandal gave them an excuse to dump an expensive and possibly over-priced auditor while at the same time appearing to be more noble when switching from a scandal-ridden auditor. For example, the client may strongly suspect the audit firm is padding the work hours for no good reason. *If at least one interview found that the scandal was an excuse rather than the reason for switching auditors we have slightly more evidence of causality than we had with just the accountants science study that can say zero about causality.*

Possible Answer 3

Skinner and Srinivasan did not consider the possibility that some clients switched auditors because the scandal gave them an excuse to change to an auditor having a local office nearby that promised better service due to response times and at lower cost due to such things as lower travel expense billings. Auditors having nearby offices also improve relationship building at civic meetings, golf outings, etc. This may not be ideal from the standpoint of independence considerations, but clients are generally less concerned about independence than investors.

Possible Answer 4

Skinner and Srinivasan did not consider the possibility that some clients switched auditors because the scandal gave them an excuse to change from an audit firm that communicated poorly with some clients. Reasons in general that companies give for changing auditors are that their auditors communicated poorly with management and audit committees.

Possible Answer 5

Skinner and Srinivasan did not consider the possibility that some clients switched auditors because the scandal gave them an excuse to change from an audit firm that was inefficient and superficial in the audit. For example, the audit teams might be comprised of novice auditors having little or no experience with the industry and/or the types of accounts being audited. For example, auditors being assigned to audit interest rate swaps might keep asking naïve questions about derivative instruments contracts and hedging.

Possible Answer 6

Skinner and Srinivasan did not consider the possibility that some clients switched auditors because the scandal gave them an excuse to change from a newly assigned partner in charge that the client really disliked relative to previous partners in charge. Audit firms change partners in charge of audits for various reasons, and client experiences with a new partner and charge may greatly sweeten or sour the audit experience.

There are of course many other possible reasons for switching and/or retaining audit firms. We won't really know until we ask.

Conclusion

The point here is that non-scientific research methods have chances of finding granules of causation that are impossible to find when the granules of causation cannot possibly be uncovered in the accountics science studies that do not drill down to granules of causation.

Example 2: Granular Non-Science Research and Database Variables

On occasion, databases have granularity that's ignored in scientific study because the granularity is not easily placed in mathematical models. Sometimes there's just too much granularity. And the granularity data may be too subjective and/or immeasurable. An example of bank stress testing is shown below.

Banks must also submit much more granular information, including dozens of details about individual loans.

"**Stress for Banks, as Tests Loom**," by Victoria McGrane and Dan Fitzpatrick, *The Wall Street Journal*, October 8, 2012 ---

http://professional.wsj.com/article/SB10000872396390444024204578044591482524484.html?mod=WSJ_hp_LEFTWhatsNewsCollection.

U.S. banks and the Federal Reserve are battling over a new round of "stress tests" even before the annual exams get going later this fall.

The clash centers on the math regulators are using to produce the results. Bankers want more detail on how the calculations are made, and the Fed thus far has resisted disclosing more than it has already.

A senior Fed supervision official, Timothy Clark, irked some bankers last month when he said at a private conference they wouldn't get additional information about the methodology, according to people who attended the event in Boston. Wells Fargo WFC -0.78% & Co. Treasurer Paul Ackerman said at the same conference that he still doesn't understand why the Fed's estimates are so different from Wells's. His remarks drew applause from bankers in the audience, said the people who attended.

The annual examinations in their fourth year have become a cornerstone of the revamped regulatory rule book—and a continuing source of tension between the nation's biggest banks and their overseers.

Smaller banks will soon have to grapple with similar requirements. On Tuesday, the three U.S. banking regulators—the Fed, the Comptroller of the Currency and the Federal Deposit Insurance Corp.—plan to complete rules requiring smaller banks with more than \$10 billion in assets to also run an internal stress test each year. That would widen the pool of test participants beyond the Fed's current requirement of \$50 billion in assets, a group comprised of 30 banks.

The stress tests, which started in 2009 as a way to convince investors that the largest banks could survive the financial crisis, now are an annual rite of passage that determines banks' ability to return cash to shareholders.

The financial crisis taught regulators that they need to be able "to look around the corner more often than in the past," said Sabeth Siddique, a director at consulting firm Deloitte & Touche, who was part of the Fed team that ran the inaugural stress test in 2009.

The Fed asks the big banks to submit reams of data and then publishes each bank's potential loan losses and how much capital each institution would need to absorb them. Banks also submit plans of how they would deploy capital, including any plans to raise dividends or buy back stock.

After several institutions failed last year's tests and had their capital plans denied, executives at many of the big banks began challenging the Fed to explain why there were such large gaps between their numbers and the Fed's, according to people close to the banks.

Fed officials say they have worked hard to help bankers better understand the math, convening the Boston symposium and multiple conference calls. But they don't want to hand over their models to the banks, in part because they don't want the banks to game the numbers, officials say.

It isn't clear if smaller banks will have to start running their tests immediately, as regulators have issued guidance indicating that midsize banks will have at least another year until they have to run the tests.

One new frustration for big banks is that the information requested by the Fed is changing. This year the Fed began requiring banks to submit data on a monthly and quarterly basis, in addition to the annual submission. **Banks must also submit much more granular information, including dozens of details about individual loans.**

Fed officials say the new data gives them the information they need to build their stress-test models and to see banks' risk-taking over time. Banks say the Fed has asked them for too much, too fast. Some bankers, for instance, have complained the Fed now is demanding they include the physical address of properties backing loans on their books, not just the billing address for the borrower. Not all banks, it turns out, have that information readily available.

Daryl Bible, the chief risk officer at [BB&T Corp.](#), [BBT -0.77%](#) a Winston-Salem, N.C.-based bank with \$179 billion in assets, challenged the Fed's need for all of the data it is collecting, saying in a Sept. 4 comment letter to the regulator that "the reporting requirements appear to have advanced beyond the linkage of risk to capital and an organization's viability," burdening banks without adding any value to the stress test exercise. BB&T declined further comment.

The Fed has backed off some of its original requests after banks protested. For example, the Fed announced Sept. 28 that it wouldn't require chief financial officers to attest to the accuracy of the data submitted after banks and their trade groups argued that the still-evolving process was too fresh and confusing for any CFO to be able to be sure his bank had gotten it right.

Banks needed more time to build up the systems and controls to report data reliably, the Fed said. But the regulator also warned that it may require CFO sign-off in the future.

Accounting scientists and financial analysts typically ignore the granular data when they build mathematical models of bankruptcy risk of a bank. For example, typical mathematical models are the [Value at Risk](#) (VaR) model and the [Altman Z-Score model](#). Neither model analyzes the granular detail of a bank's loans to specific individuals.

Much more subjectivity in valuation becomes necessary for "granular factors" that take uniqueness of each loan into consideration. The typical valuation model is discounted cash flow (DCF economic value) adjusted by granular factors. In 1932, Bill Paton (in his *Accountants Handbook*), Bill Paton outlines granular "appraisal factors" in the following categories:

1. Length of time the account has run.
2. Customer's practice with respect to discounts.
3. General character of dealings with the customer.
4. Credit ratings and similar data.
5. Special investigations and reports.

This highlights how ostensibly scientific databases can contain non-scientific elements of data. For example, when the Fed's stress test database contains a data element for working capital, there's little concern over the accuracy and interpretation of this data point. But when the database contains a description of the creditor's general character there's a much more subjective aspect to this data even if the data is a single point on a [Likert Scale](#).

Whereas bank managers and bank auditors may examine Paton's granular detail on some type of sampling basis, the VaR and Altman Z-Score scientists do not build such granular detail into their models even on a sampled basis.

Qualitative Research

Qualitative Research --- http://en.wikipedia.org/wiki/Qualitative_research

Qualitative research is a method of inquiry employed in many different academic disciplines, traditionally in the [social sciences](#), but also in [market research](#) and further contexts.^[1] Qualitative researchers aim to gather an in-depth understanding of [human behavior](#) and the [reasons](#) that govern such behavior. The qualitative method investigates the *why* and *how* of [decision making](#), not just *what*, *where*, *when*. Hence, smaller but focused [samples](#) are more often needed than large [samples](#).

In the conventional view, qualitative methods produce information only on the particular cases studied, and any more general conclusions are only propositions (informed assertions). [Quantitative methods](#) can then be used to seek empirical support for such research hypotheses. This view has been disputed by Oxford University professor [Bent Flyvbjerg](#), who argues that qualitative methods and [case study research](#) may be used both for hypotheses-testing and for generalizing beyond the particular cases studied

"How to qualitatively assess indefinite-lived intangibles for impairment," Ernst & Young, October 18, 2012 --- [Click Here](#)

[http://www.ey.com/Publication/vwLUAssets/AL/TechnicalLine_BB2420_Intangibles_18October2012/\\$FILE/TechnicalLine_BB2420_Intangibles_18October2012.pdf](http://www.ey.com/Publication/vwLUAssets/AL/TechnicalLine_BB2420_Intangibles_18October2012/$FILE/TechnicalLine_BB2420_Intangibles_18October2012.pdf)

What you need to know

Page | 12

- Companies that use the optional qualitative assessment and achieve a positive result can avoid the cost and effort of determining an indefinite-lived intangible asset's fair value.
- Using the new qualitative assessment will require significant judgment.
- Companies that use the qualitative assessment will have to consider positive and negative evidence that could affect the significant inputs used to determine fair value.
- Companies that have indefinite-lived intangible assets with fair values that recently exceeded their carrying amounts by significant margins are likely to benefit from the qualitative assessment.
- Using the qualitative assessment does not affect the timing or measurement of impairments.

Overview

The Financial Accounting Standards Board (FASB or Board) introduced an optional qualitative assessment for testing indefinite-lived intangible assets for impairment that may allow companies to avoid calculating the assets' fair value each year.

Accounting Standards Update (ASU) 2012-021 allows companies to use a qualitative assessment similar to the optional assessment introduced last year for testing goodwill for impairment.² The goal of both standards is to reduce the cost and complexity of performing the annual impairment test.

ASC 3503 requires companies to test indefinite-lived intangible assets for impairment annually, and more frequently if indicators of impairment exist. Before ASU 2012-02, the impairment test required a company to determine the fair value of

Continued in article

Bob Jensen's threads on intangibles and contingencies ---

<http://www.trinity.edu/rjensen/theory01.htm#TheoryDisputes>

The purpose of this paper was to show how non-scientific qualitative research cannot only add value to quantitative studies, qualitative research can find granules of causation that cannot be discovered in most quantitative studies.

Bob Jensen's threads on case method research ---

<http://www.trinity.edu/rjensen/000aaa/thetools.htm#Cases>

"Mechanical Turk and the Limits of Big Data: The Internet is transforming how researchers perform experiments across the social sciences," by Walter Frick, MIT's *Technology Review*, November 1, 2012 --- [Click Here](#)

http://www.technologyreview.com/view/506731/mechanical-turk-and-the-limits-of-big-data/?utm_campaign=newsletters&utm_source=newsletter-daily-all&utm_medium=email&utm_content=20121102

It's telling that the most interesting presenter during *MIT Technology Review's* [EmTech session on big data](#) last week was not really about big data at all. It was about [Amazon's Mechanical Turk](#), and the experiments it makes possible.

Like many [other researchers](#), sociologist and Microsoft researcher [Duncan Watts](#) performs experiments using Mechanical Turk, an online marketplace that allows users to pay others to complete tasks. Used largely to fill in gaps in applications where human intelligence is required, social scientists are increasingly turning to the platform to test their hypotheses.

The point Watts made at EmTech was that, from his perspective, the data revolution has less to do with the amount of data available and more to do with the newly lowered cost of running online experiments.

Compare that to Facebook data scientists Eytan Bakshy and Andrew Fiore, who presented right before Watts. Facebook, of course, generates a massive amount of data, and the two [spoke of the experiments they perform](#) to inform the design of its products.

But what might have looked like two competing visions for the future of data and hypothesis testing are really two sides of the big data coin. That's because data on its own isn't enough. Even the kind of experiment Bakshy and Fiore discussed—essentially an elaborate A/B test—has its limits.

This is a point political forecaster and author Nate Silver discusses in his recent book *The Signal and the Noise*. After discussing economic forecasters who simply gather as much data as possible and then make inferences without respect for theory, he writes:

This kind of statement is becoming more common in the age of Big Data. Who needs theory when you have so much information? But this is categorically the wrong attitude to take toward forecasting, especially in a field like economics, where the data is so noisy. Statistical inferences are much stronger when backed up by theory or at least some deeper thinking about their root causes.

Bakshy and Fiore no doubt understand this, as they cited plenty of theory in their presentation. But Silver's point is an important one. Data on its own won't spit out answers; theory needs to progress as well. That's where Watts's work comes in.

Continued in article