Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory

Authors: Arch Woodside, Roberta MacDonald

Post-print version of an article published in Journal of Business Research 66(4): 463-472. doi:10.1016/j.jbusres.2012.12.021.

Post-print version of an article published in Journal of Business Research 66(4): 463-472. doi: 10.1016/j.jbusres.2012.12.021

MOVING BEYOND MULTIPLE REGRESSION ANALYSIS TO ALGORITHMS:

Calling for Adoption of a Paradigm Shift from Symmetric to Asymmetric Thinking in Data Analysis and Crafting Theory

Arch G. Woodside, Boston College

Submission: December 2012

Revision: December 2012

Acceptance: December 2012

Published: *Journal of Business Research*, 66 (4), 463-472.

The author appreciates the insights and suggestions for revising by Carol M. Megehee, Coastal Carolina University, to earlier drafts of this version of the paper. Send correspondence to Arch G. Woodside, Boston College, Carroll School of Management, Department of Marketing, 140 Commonwealth Avenue, Chestnut Hill, MA 02467; telephone/fax: 1 617 552 3069 / 6677 (arch.woodside@bc.edu).

Post-print version of an article published in Journal of Business Research 66(4): 463-472. doi: 10.1016/j.jbusres.2012.12.021

MOVING BEYOND MULTIPLE REGRESSION ANALYSIS TO ALGORITHMS

Calling for Adoption of a Paradigm Shift from Symmetric to Asymmetric Thinking in Data Analysis and Crafting Theory

ABSTRACT

This editorial suggests moving beyond relying on the dominant logic of multiple regression analysis (MRA) toward thinking and using algorithms in advancing and testing theory in accounting, consumer research, finance, management, and marketing. The editorial includes an example of testing an MRA model for fit and predictive validity. The same data used for the MRA is used to conduct a fuzzy-set qualitative comparative analysis (fsQCA). The editorial reviews a number of insights by prominent scholars including Gerd Gigerenzer's treatise that "Scientists' tools are not neutral." Tools impact thinking and theory crafting as well theory testing. The discussion may be helpful for early career scholars unfamiliar with David C. McClelland's brilliance in data analysis and in introducing business research scholars to fsQCA as an alternative tool for theory development and data analysis.

Keywords: algorithm; causal recipe; configuration; consistency; coverage; fit validity; fuzzy set qualitative comparative analysis; multiple regression analysis; predictive validity

INTRODUCTION: TOOLS-TO-THEORY PERSPECTIVE

MRA is more than just a statistical tool—the method shapes thinking and theory crafting. "Scientists' tools are not neutral" (Gigerenzer, 1991, p. 19). This editorial is an echo and an application of Gigerenzer's (1991) general thesis that scientific tools (both methods and instruments) suggest new theoretical metaphors and theoretical concepts once they are entrenched in scientific practice; familiarity with the tools within a scientific community also lays the foundation for the general acceptance of the theoretical concepts and metaphors inspired by the tools. This editorial is not to suggest that researchers should always avoid using MRA.

The editorial does suggest that most MRA applications in business research and JBR submissions are done badly and that researchers should think and craft algorithms for building and testing theory much more often they do now. The comments and recommendations concerning MRA apply to structural equation modeling (SEM) as well.

Additional comments on the severe limitations of MRA and SEM research using fixed-point five- and seven-point self-report scales to learn cognitive processes appear elsewhere (Woodside, 2011). The limitations of using one-shot, one-person-per-firm, or one-person-per household, self-reports as valid indicators of causal relationships of actual processes are so severe that academics should do more than think twice before using such surveys as the main method for collecting data – if scholars seek to understand and describe actual thinking processes additional methods are necessary for data collection. The relevant literature includes several gems of exceptionally high quality, validity, and usefulness in the study of actual processes; reading these studies are a useful step toward reducing the reliance on one-shot self-report surveys (Woodside, 2011, describes some of these exceptionally high-quality studies).

A CALL TO MOVE BEYOND MRA

Several tenets support this call to move beyond MRA to crafting and testing theory using algorithms. First, researchers using MRA focus on estimating whether or not the influence (i.e., the effect size) of each hypothesized independent variable associates significantly with a dependent variable after separating out the influence of other independent variables in an equation involving two or more independent variables—a "net effects" estimation approach to research. Frequently, such research reports include comparisons of models with specific independent variables having significant versus insignificant net effects depending on the presence or absence of other independent variables in the models.

Given that multi-collinearity (i.e., significant correlations among the independent variables) always occurs with a high number of variables in a model (e.g., ten variables), a researcher may show that none of the independent variables has a significant net effect while at the same time the model explains a substantial share of the variance in the dependent variable or that a given variable of high interest (e.g., private-equity ownership) shifts from significant to nonsignificant status in influencing a dependent variable (e.g., loan default) depending upon what other variables the researcher includes in the models (Hotchkiss, Smith, & Strömberg, 2013, Tables 6-9; Mauro, 2001, "Table VI").

The focus on net effects is misleading for several reasons and more useful perspectives on theory and method are available. Reasons not to rely on MRA exclusively include the point that cases counter to the observed net effects nearly always occur—not all the cases in the data support a negative or positive relationship between the independent and dependent variable. Rather than showing a limited number of models in which X has a positive (or negative) net influence on Y, the researcher can increase the contribution of the study by showing the combinatory conditions for which X is a positive influence on Y as well as the combinatory conditions when X is a negative influence on Y. For example, in an award-winning paper on adoption of industry certification standards in the cut flower industry in Colombia and Ecuador, Prado (2012) shows that a dummy country variable (with Colombia equal 1 and Ecuador equal to zero) results in a consistent negative net-effect influence on adoption. Yet, many firms in Colombia adopt the industry standards. Prado (2012) does not address the issue of how the seemingly negative country influence is overcome to achieve the outcome of adoption—what combination of influences of antecedent conditions in Colombia leads to industry certification adoption?

Second, reality usually includes more than one combination of conditions that lead to high values in an outcome condition (i.e., high values in a dependent variable); thus, reality usually indicates any insightful combination of conditions has an asymmetrical relationship with an outcome condition and not a symmetrical relationship. MRA tests the extent to which the relationship between a causal statement (i.e., statement X) involving one and more weighted variables and an outcome variable Y is symmetrical. In symmetrical relationships, low values of (a single or complex statement of) X associate with low values of Y and high values of X associate with high values of Y.

Figure 1a shows a symmetric relationship for a causal statement and a dependent variable. Figure 1b shows an asymmetric relationship for a causal statement and a dependent variable. A symmetric relationship indicates that high values of X are both necessary and sufficient for high values of Y to occur and that low values of Y occur with low values of X. The asymmetric relationship in Figure 1b indicates that high values of X are sufficient for high values of Y to occur but high values of X are not necessary for high values of Y to occur; high values of Y occur when values of X are low, indicating that additional "causal recipes" (i.e. simply and complex X statements) associate with high values of Y. "Causal recipes" (Ragin, 2008) are combinatory statements of two or more simple antecedent conditions, for example the combination statement of "old, wealthy, and divorced male" is a conjunctive statement of four antecedent conditions—a possible causal recipe with high values on this statement associating with a high score on buying a Lexus convertible automobile.

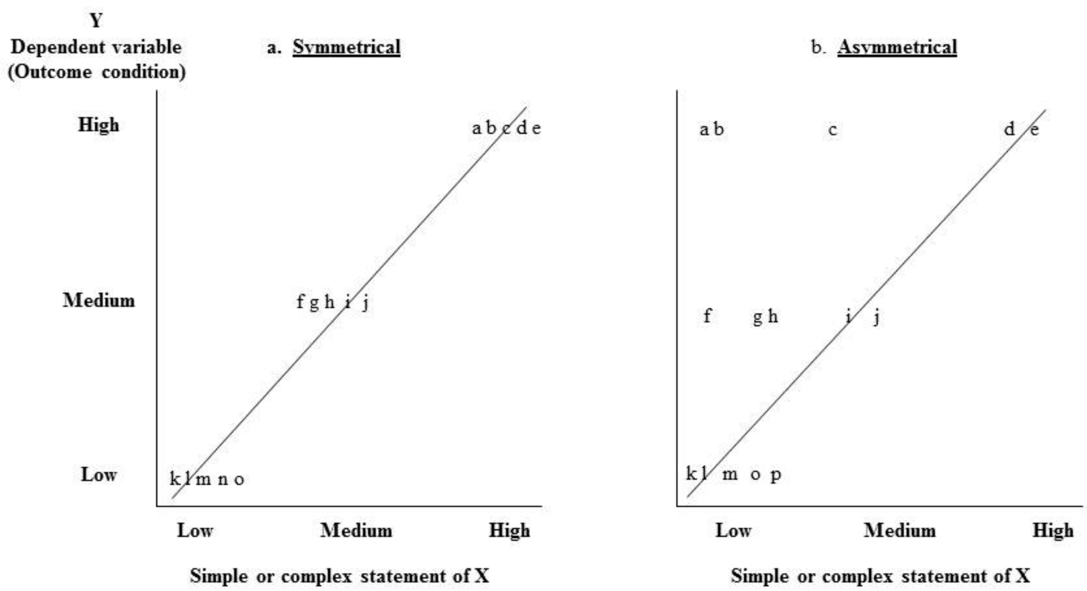


Figure 1
Symmetrical and Asymmetrical Relationships
between X and Y for 15 Cases of Synthetic Data

Significant correlations above .80 indicate symmetric relationships; significant correlations in the range of .30 to .70 indicate asymmetric relationships. Except for findings of tests for reliability of items in a measurement scale, significant correlations between unique variables usually fall below .70 because different combinations of independent variables associate with high values of Y, and any given X statement that relates substantially with Y has both low as well as high values that relate to high values for Y.

Table 1 includes data that matches with Figure 1. In Table 1 the correlation for the data in Figure 1a equals 0.98—indicating a symmetric relationship. In the second data set in Table 1 the correlation for the data for Figure 1b equals 0.49—indicating an asymmetric relationship. Using the software program for fuzzy set qualitative comparative analysis (available at fsQCA.com), the first two data sets are transformed to “calibrated” scores in the third and fourth parts of Table 1. For the calibrated scores, Table 1 reports “consistency” and “coverage” indices.

Table 1
Correlation and QCA Tests of Symmetric and Asymmetric Relationships

Symmetrical Data				Asymmetric	Data	Calibrated	Symmetric	Data	Calibated	Asymmetric
Case	X	Y		xx	yy	X	Y		xx	yy
a	2.6	3		1	3	0.9	0.98		0.02	0.98
b	2.7	3		1.1	3	0.93	0.98		0.03	0.98
c	2.8	3		1.7	3	0.95	0.98		0.25	0.98
d	2.9	3		2.9	3	0.97	0.98		0.97	0.98
e	3	3		3	3	0.98	0.98		0.98	0.98
f	1.7	2		1	2	0.25	0.5		0.02	0.5
g	1.8	2		1.3	2	0.32	0.5		0.07	0.5
h	1.9	2		1.4	2	0.41	0.5		0.1	0.5
i	2	2		1.8	2	0.5	0.5		0.32	0.5
j	2.1	2		1.9	2	0.59	0.5		0.41	0.5
k	0.8	1		1	1	0.01	0.02		0.02	0.02
l	0.9	1		1.1	1	0.02	0.02		0.03	0.02
m	1	1		1.2	1	0.02	0.02		0.05	0.02
n	1.1	1		1.3	1	0.03	0.02		0.07	0.02
o	1.2	1		1.4	1	0.05	0.02		0.1	0.02
	r = .98			r = .49		consistency	0.98		consistency	0.95
						coverage	0.91		coverage	0.44

The consistency index is analogous to a correlation and the coverage index is analogous to the “coefficient of determination” (i.e., r^2). Details appear below on calculating consistency and coverage; the point for now is that whether or not a relationship between X and Y is symmetrical or asymmetrical has little impact on consistency scores. Note that the consistency scores equal 0.98 and 0.95 for the symmetric and asymmetric data sets in Figure 1. Unlike correlation analysis, consistency is a test for sufficiency and not a test for sufficiency and necessity. While correlation and multiple regression analysis are matrix algebra applications, consistency and coverage are Boolean algebra applications. Appendix 1 shows the formula with example calculations for consistency and coverage.

For most contexts in reality no one simple or one complex statement of an independent variable (X) is necessary for high values of a dependent variable (Y). The dominant-logic approach to theory proposals of one given model that leads eventually to a principal dependent variable needs replacing to account for the reality of multiple combinations (i.e.. causal recipes, alternative routes) resulting in high values in the dependent variable.

Consider the findings of Cooper and Kleinschmidt (2007)—authors of a series of highly cited studies on the effects of key success factors (KSFs) and profitability (numbers in parentheses are correlations of the KSFs with profitability). The study uses five- point Likert scales to measure each item. In one of their studies, the correlations below of 161 firms engaging in new product development indicate that the presence versus absence of a factor is not sufficient for high profitability:

- A high-quality new product process (.416)
- A defined new product strategy for the business unit (.228)
- Adequate resources-people and money-for new products (.244)
- R&D spending on new products (as % of the business's sales) (ns=not significant)
- High-quality new product development teams (.196)
- Senior management commitment to new products (.268)
- An innovative climate and culture in the business unit (.243)
- The use of cross-functional teams for product development (.230)
- Senior management accountability for new product results (.228).

The use of the expressions, "key success factors" and "critical success factors," are misleading in that none of the correlations indicate necessary or sufficiency for high profitability. The effect size of these correlations indicate that while some of these actions may be useful in combinations with other actions, none alone are sufficient for high profitability. None of the factors are necessary or sufficient for a highly successful product development.

If these 9 dimensions represent somewhat unique KSFs, what combinations of high versus low values among the 9 KSFs lead to high profitability? Any one firm among firms with a highly profitable new product is unlikely to achieve level 5 (highest) evaluations for all 9 dimensions. Using a property-space approach (Lazarsfeld 1937), considering three levels for each dimension—low, moderate, high—a total of 19,683 combinations are possible (p. 39). A few of these paths are likely to result in highly profitable new product outcomes—possibly 10% of the paths or about 200 paths. About 30% of the paths are likely to result in substantial losses—about 600 paths. The remaining paths are likely to be untried and most may represent non-implementable decision recipes. Multiple "key success paths" (KSPs) relate to high scores for product innovation success rather than KSFs.

Third, referring to success and competencies, McClelland (1998) stresses that many relationships among a dependent variable and independent variables are not linear and not well described by correlation coefficients. Instead, such relationships are describable as "tipping points" (Gladwell, 1996). What sociologists often observe in changes in a societal variable making little difference until the changes reach a certain level is likely to occur in business research contexts as well.

McClelland (1998) illustrates this tenet for the relationship of competency frequencies and levels to success as an executive. For example, Figure 2 shows that for "Impact and

Influence", the T (i.e., typical executives) group is more likely than the O (i.e., outstanding executives) group to have a frequency score anywhere from 0 to 7; the O group is more numerous than the T group only when the frequency score reaches 8 to 10; further, this O versus T difference does not change at higher frequency scores, above 10. So it would be a misrepresentation of the relationship to describe it in terms of a linear correlation coefficient. For the data graphed in Figure 1, for example, the biserial r is .22, $p < .10$, between O versus T status and frequency of the competency "Impact and Influence", but this statistic understates the significance of the relationship (55% of the O executives vs. 20% of the T executives had frequencies of 8 or more, $p < .001$) and misrepresents its nature for frequencies below 8.

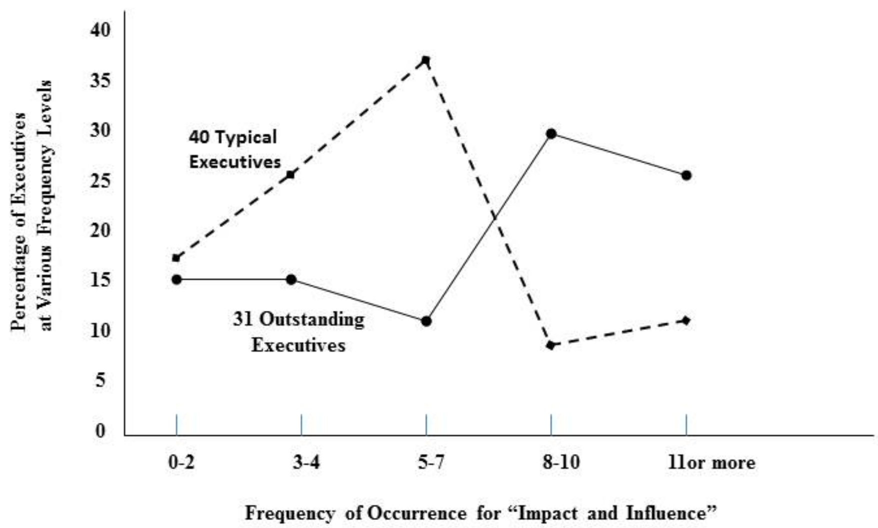


Figure 2
Percentages of Outstanding and Typical Executives Showing Different Frequencies of the Competency "Impact and Influence"
Source: McClelland (1998, Figure 1, p. 334).

Thirteen studies of managers were examined to see whether the O group satisfied the following algorithm: mean frequency or maximum-levels core significantly higher than that of the T (T = typical) managers (a) on at least one of the initiative and one of the organizational competencies and (b) on a total of five competencies drawn from the list in McClelland's Table 1

(McClelland's Table 1 includes12 competencies from "Achievement Orientation" to "Team Leadership"). The O groups in 11 (85%) of the studies satisfied this algorithm, compared with only 1 out of 8 (13%) studies of individual contributors, that is, technical and professional personnel such as geologists, consultants, and insurance raters($p < .01$ for the difference in proportions).

Thus, in McClelland's (1998) study competency algorithms that associate with success in various types of executive positions are observable by using the principle of substitutability; that is, a variety of different but functionally equivalent alternative predictor variables may relate to an outcome criterion. To some extent, therefore, different competencies can substitute for each other.

In a seminal paper, Mauro (1995) makes the same point about substitutability in his research on the impact of country-level corruption, red-tape, and institution inefficiency on total investment as well as GDP growth. Mauro's (1995) data set consists of the Business International (BI) indices on corruption, red tape, and the inefficiency of the judicial systems for 1980-1983. The indices are based on standard questionnaires filled in by BI's correspondents stationed in about 70 countries. He restricts his analysis to nine indicators; each averaged over four years—"a less noisy indicator of institutional variables, which we may expect to change only slowly" (Mauro, 1995, p. 684).

The BI indices are integers between 0 and 10 and a high value of the index means that the country in question has "good" institutions. In his "Section III" and the first five columns after "nation" in Appendix 2 to this editorial, each indicator is the simple average for the country in question for the period 1980-1983. Mauro grouped together each of the nine indicators into one of five summary indicators based on "closely related on a prior grounds, the indices that I choose to group together are more strongly correlated with each other" (Mauro, 1995, p. 686).

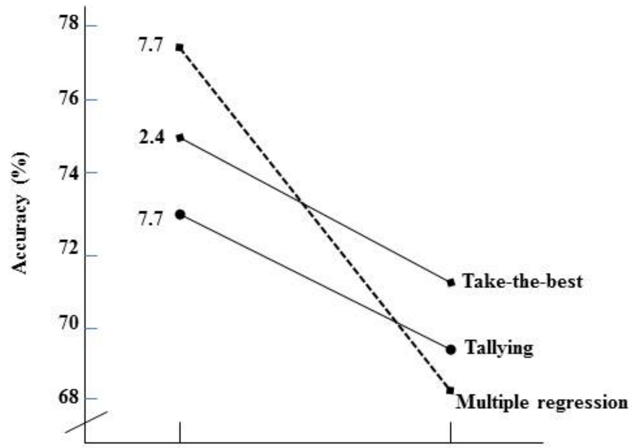
Mauro (1995) observes that all BI indices are positively and significantly correlated before and after controlling for gross domestic product (GDP) per capita. "A number of mechanisms may contribute to explaining the positive correlation[s] among all categories of institutional efficiency. Corruption may be expected to be more widespread in counties where red tape slows down bureaucratic procedures… At the same time this multicollinearity makes it difficult to tell which of the several institutional factors examined is crucial for investment and growth. As a consequence, it may be desirable to combine groups of variables into composite indices" (Mauro 1995, pp. 685-686).

The difficulty is overcome if the researcher moves beyond thinking in terms of which of the several institutional factors are crucial; none are crucial but a few combinations of these variables are likely to associate with high levels of investment and high levels of growth. Rather than developing theory and thinking in terms of relative impacts of independent variables, thinking in terms of alternative mechanisms (i.e., algorithms) indicates that several causal recipes relate to high economic growth.

The following additional point has profound theoretically and practical importance. For one or more cases a **low** score on anyone antecedent condition (such as "Achievement Orientation" in McClelland's study or corruption in Mauro's study) may combine with other antecedents to result in a **high** score on the outcome condition. With medium-to-large sample sizes, cases occur with seemingly unusual scores on any one simple antecedent condition ("independent variable") that are counter to the primary influence (the "main effect") of the simple condition and the outcome.

While in Mauro's (1995) study the BI indices all correlate positively, at the case level combinations occur that run counter to this finding—as Mauro reports in "Table II" in his paper. While some countries have relatively high or low scores in all five indices relating to corruption, red tape, and efficiency, other countries have surprising combination of low, medium, and high scores. For example, consider Zimbabwe's scores in Appendix 2 for 1980-1983; the scores include high calibrated values for judicial efficiency, red tape (indicating low red tape), corruption (indicating low corruption), and bureaucracy efficiency—and a low calibrated score for political stability (indicating high instability). Does such a country causal recipe associate with high economic growth? In the case of Zimbabwe, the answer is no—calibrated growth is zero.

As McClelland (1998) and others (Gigerenzer and Brighton, 2009) stress, the critical question is whether or not a model (e.g., an empirical multiple regression model or an algorithm) predicts a dependent variable in additional samples—holdout samples that are separate data sets from the data sets used to test the fit of data to a theory. Gigerenzer and Brighton (2009, p. 118) confirm "that achieving a good fit to observations does not necessarily mean we have found a good model, and choosing the model with the best fit is likely to result in poor predictions. Despite this, Roberts and Pashler (2000) estimated that, in psychology alone, the number of articles relying on a good fit as the only indication of a good model runs into the thousands." Currently, this bad practice occurs for most submissions to the JBR and likely for most submissions and published articles in all business-related journals.



"Fig. 1. Less is more effects. Both tallying and take-the-best predict more accurately than multiple regression, desipte using less information and computation. Note that multiple regression excels in data fitting ("hindsight"), that is, fitting its parameters to data that are already known, but performs relatively poorly in prediction ("foresight," as in cross-validation. Take-the-best is the most frugal, that is, it looks up on average, only 2.4 cues when making inferences. In contrast, both multiple regression analysis and tallying look up 7.7 cues on average. The results shown are averaged across 20 studies, including psychological, biological, sociological, and economic inference tasks (Czerlinkski, Gigerenzer, & Goldstein, 1999). For each of the 20 studies and each of the three strategies, the 95% confidence intervals were <= 4 percentage points."

Figure 3
Moving Beyond Fit to Prediction Validity
Source: Gigerenzer and Brighton (2009, Figure 1, p. 112).

Gigerenzer and Brighton's (2009) study explains in-depth why high model fit results in low predictive validity. Their observation and conclusions are central to the purpose of this editorial. Their Figure 1 (Figure 3 here) is profound in illustrating Armstrong's (2012) observations about MRA.

> Analysts assume that models with a better fit provide more accurate forecasts. This ignores the research showing that fit bears little relationship to ex ante forecast accuracy, especially for time series. Typically, fit improves as complexity increases, while ex ante forecast accuracy decreases – a conclusion that Zellner (2001) traced back to Sir Harold Jeffreys in the 1930s. In addition, analysts use statistics to improve the fit of the model to the data. In one of my Tom Swift studies, Tom used standard procedures when starting with 31 observations and 30 potential variables. He used stepwise regression and included only variables where t was greater than 2.0. Along the way, he dropped three

> outliers. The final regression had eight variables and an R-square (adjusted for degrees of freedom) of 0.85. Not bad, considering that the data were from Rand's book of random numbers (Armstrong 1970). I traced studies on this illusion back to at least 1956 in an early review of the research on fit and accuracy (Armstrong 1985). Studies have continued to find the fit is not a good way to assess predictive ability (e.g., Pant and Starbuck 1990). The obvious solution is to avoid use of t, p, F, R-squared and the like when using regression. (Armstrong, 2012, p. 690)

Armstrong's (2012) observations are valuable for referencing in particular when looking at an MRA table with six-to-twenty independent terms in the attempt to control for influences beyond the focal independent variables. "Users of regression assume that by putting variables into the equation they are somehow controlling for these variables. This only occurs for experimental data. Adding variables does not mean controlling for variables in non-experimental data because many variables typically co-vary with other predictor variables. The problem becomes worse as variables are added to the regression. Large sample sizes cannot resolve this problem, so statistics on the number of degrees of freedom are misleading" (Armstrong, 2012, p. 691).

Armstrong (2012) recommends against estimating relationships for more than three variables in a regression—findings from Goldstein and Gigerenzer (2009) are consistent with this rule-of-thumb. A complementary recommendation is not to report MRA findings without also reporting findings from using simple algorithms and never report findings for fit validity only—always report predictive validity findings from tests of models with holdout samples.

Post-print version of an article published in Journal of Business Research 66(4): 463-472. doi: 10.1016/j.jbusres.2012.12.021

Appendix 1: Efficiency, Corruption, Red Tape, and GDP Growth Data, Countries A through M

nation	judic	redtape	corrpupt	polstab	burea	gdpgro	judic_cal	redtape_c	corrupt_c	polstab_c	bureau_c	gpd_gro_cal
algeria	7.25	2.5	5	7.71	4.92	0.01	0.63	0.08	0.23	0.51	0.24	0.09
angola	4	6.33	8.66	4.61	6	0.24	0.12	0.58	0.92	0.09	0.42	0.97
argentina	6	6.66	7.66	7.72	6.77	0.02	0.35	0.66	0.73	0.52	0.6	0.17
australia	10	9.25	10	8.5	9.75	0.05	0.99	0.96	0.99	0.86	0.98	0.53
austria	5	7.25	8	9.04	8.25	0.04	0.22	0.78	0.82	0.96	0.89	0.44
banglades	6	4	4	6.5	4.67	-0.02	0.35	0.2	0.14	0.29	0.21	0.01
belgium	9.5	8	9.76	8	9.08	0.04	0.98	0.88	0.98	0.67	0.96	0.44
brazil	5.75	4	5.75	7.54	5.17	0.05	0.32	0.2	0.32	0.47	0.28	0.53
cameroon	7	6	7	8.5	6.67	0.05	0.54	0.5	0.5	0.86	0.57	0.53
canada	9.25	9.5	10	9	9.58	0.04	0.97	0.97	0.99	0.95	0.98	0.44
chile	7.25	9.25	9.25	6.46	8.58	0.07	0.63	0.96	0.97	0.28	0.92	0.61
colomiba	7.25	4.5	4.5	6	5.42	0.08	0.63	0.26	0.18	0.22	0.32	0.65
denmark	10	9.5	9.25	8.5	9.58	0.02	0.99	0.97	0.97	0.86	0.98	0.17
dominica	6.75	6	6.5	7.58	6.42	0.06	0.48	0.5	0.43	0.48	0.5	0.57
ecuador	6.25	6	5.5	6.63	5.58	0.17	0.39	0.5	0.29	0.31	0.34	0.9
egypt	6.5	8	3.25	8.67	4.25	0.08	0.44	0.88	0.1	0.9	0.16	0.65
finland	10	8.5	9.5	8.79	9.33	0.03	0.99	0.92	0.98	0.93	0.97	0.28
france	8	6.75	10	8.92	8.25	0.02	0.83	0.68	0.99	0.94	0.89	0.17
germany	9	7.5	9.5	8.21	8.67	0.04	0.95	0.82	0.98	0.77	0.93	0.44
ghana	4.66	2.33	3.66	5	3.55	-0.04	0.18	0.07	0.12	0.11	0.1	0
greece	7	4	6.25	8.63	5.76	0.06	0.54	0.2	0.39	0.9	0.38	0.57
haiti	2	2	2	6.67	2	-0.04	0.03	0.06	0.05	0.32	0.03	0
hongkong	10	9.75	8	9.5	9.26	0.04	0.99	0.98	0.82	0.98	0.96	0.44
india	8	3.25	5.25	7	5.5	0.01	0.83	0.13	0.26	0.37	0.33	0.09
indonesia	2.5	2.75	1.5	7.46	2.25	0.03	0.05	0.09	0.04	0.46	0.04	0.28
irand	2	1.25	3.25	3.25	2.17	0.08	0.03	0.03	0.1	0.03	0.04	0.65
iraq	6	3	10	5.72	6.33	0.1	0.35	0.11	0.99	0.18	0.48	0.72
ireland	8.75	7.5	9.75	7.67	8.67	0.03	0.93	0.82	0.98	0.5	0.93	0.28
israel	10	7.5	9.25	6.25	8.92	0.06	0.99	0.82	0.97	0.25	0.95	0.57
italy	6.75	4.75	7.5	7.92	6.33	0.02	0.48	0.29	0.68	0.63	0.48	0.17
jamaica	7.33	4	5	7.5	5.44	0.16	0.66	0.2	0.23	0.46	0.32	0.88
japan	10	8.5	8.75	9.42	9.08	0.02	0.99	0.92	0.93	0.98	0.96	0.17
jordan	8.66	6.33	8.33	7.78	7.77	0.02	0.93	0.58	0.88	0.55	0.83	0.17
kenya	5.75	6	4.5	6.96	5.08	0.07	0.32	0.5	0.18	0.36	0.26	0.61
korea_s	6	6.5	5.75	7.5	6.08	0.08	0.35	0.62	0.32	0.46	0.44	0.65
kuwait	7.5	6.25	7.75	8.33	7.17	0.27	0.71	0.56	0.75	0.81	0.71	0.98
liberia	3.33	5	2.66	5	3.66	-0.09	0.08	0.33	0.07	0.11	0.11	0
malaysia	9	6	6	8.42	7	0.08	0.95	0.5	0.35	0.84	0.66	0.65
mexico	6	5.25	3.25	6.88	4.83	0.07	0.35	0.37	0.1	0.35	0.23	0.61
morocco	6.66	5.33	5.66	7.11	5.88	0.03	0.46	0.38	0.31	0.39	0.4	0.28

Appendix 1: Efficiency, Corruption, Red Tape, and GDP Growth Data, Countries M through Z

nation	judic	redtape	corrpupt	polstab	burea	gdpgro	judic_cal	redtape_cal	corrupt_cal	polstab_ca	bureau_cal	gpd_gro_cal
netherlan	10	10	10	8.83	10	0.06	0.99	0.98	0.99	0.93	0.98	0.57
newzealai	10	10	10	8.5	10	0.03	0.99	0.98	0.99	0.86	0.98	0.28
nicaragua	6	4	8.75	5.5	6.25	0.05	0.35	0.2	0.93	0.16	0.47	0.53
nigeria	7.25	2.75	3	7.29	4.33	0.21	0.63	0.09	0.08	0.42	0.17	0.94
norway	10	9	10	9.5	9.67	0.07	0.99	0.95	0.99	0.98	0.98	0.61
pakistan	5	4	4	5.33	4.33	0	0.22	0.2	0.14	0.14	0.17	0.05
panama	6.75	7.25	5	7.54	6.33	0.13	0.48	0.78	0.23	0.47	0.48	0.81
peru	6.75	5.75	7.25	6.04	6.58	0.09	0.48	0.46	0.59	0.22	0.55	0.69
philipp	4.76	5	4.5	6.08	4.75	-0.06	0.19	0.33	0.18	0.23	0.22	0
portugal	5.5	4.5	6.75	7.54	5.58	0.04	0.28	0.26	0.46	0.47	0.34	0.44
saudi_ara	6	5.25	4.75	8.33	5.33	0.15	0.35	0.37	0.21	0.81	0.3	0.86
singapor	10	10	10	10	10	0.17	0.99	0.98	0.99	0.99	0.98	0.9
s_africa	6	7	8	6.5	7	-0.02	0.35	0.73	0.82	0.29	0.66	0.01
spain	6.25	6	7	6.67	6.42	0.04	0.39	0.5	0.5	0.32	0.5	0.44
sirlanka	7	6	7	7.22	6.67	0.03	0.54	0.5	0.5	0.41	0.57	0.28
sweden	10	8.5	9.23	9	9.25	0.06	0.99	0.92	0.97	0.95	0.96	0.57
switzer	10	10	10	9.25	10	0.04	0.99	0.98	0.99	0.97	0.98	0.44
taiwan	6.75	7.25	6.75	8.58	6.92	0.06	0.48	0.78	0.46	0.88	0.64	0.57
thailand	3.25	3.25	1.5	5.63	2.67	0.01	0.08	0.13	0.04	0.17	0.05	0.09
trinidad	8	4	6.5	7.79	6.17	0.14	0.83	0.2	0.43	0.56	0.45	0.84
turkey	4	5.33	6	8.17	5.11	0.1	0.12	0.38	0.35	0.75	0.27	0.72
uk	10	7.75	9.25	8.33	9	0.03	0.99	0.85	0.97	0.81	0.95	0.28
usa	10	9.25	10	9.33	9.75	0.02	0.99	0.96	0.99	0.98	0.98	0.17
urugray	6.5	6	8	9	6.83	-0.02	0.44	0.5	0.82	0.95	0.62	0.01
venezuela	6.5	4	5.75	7.71	5.42	0.17	0.44	0.2	0.32	0.51	0.32	0.9
zimbabwe	7.5	7.75	8.75	6.5	8	-0.07	0.71	0.85	0.93	0.29	0.86	0

ILLUSTRATING MRA AND ALGORITHMS

Appendix 2 includes "gdpgro" that represents data for average annual GDP (in purchasing power parity USD, PPP) per capita for 2006-2011—hereafter GDP growth. These data are available from the annual Central Intelligence Agency (CIA) World Factbook. The CIA World Factbook publications are available online, for example, CIA World Factbook, 2012. The study here examines the issue of whether or not Mauro's data on corruption, red tape, and efficiency for 1980-1983 relates to average GDP growth for 2006-2011. Given that all variables usually change slowly, the study is likely to support the hypothesis that the corruption, red tape, and inefficiency reduce growth.

Table 3
Multiple Regression Analysis for Entire Sample of *66* Nations

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.223[a]	.050	-.029	.06819

a. Predictors: (Constant), bureu_eff, polstability, redtape, judiciary, corruption

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.015	5	.003	.630	.678[b]
	Residual	.279	60	.005		
	Total	.294	65			

a. Dependent Variable: gdp_grow_06_11_ave
b. Predictors: (Constant), bureu_eff, polstability, redtape, judiciary, corruption

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.001	.050		-.025	.980
	judiciary	.015	.012	.475	1.266	.210
	redtape	.006	.011	.197	.521	.605
	corruption	.018	.014	.659	1.261	.212
	polstability	.009	.009	.180	.989	.327
	bureu_eff	-.041	.032	-1.302	-1.277	.206

a. Dependent Variable: gdp_grow_06_11_ave

Post-print version of an article published in Journal of Business Research 66(4): 463-472. doi: 10.1016/j.jbusres.2012.12.021

Table 4
Correlations of Measures for Inefficiency and Corruption and GDP Average Annual Growth Per Capita in PPP

Correlations

		judiciary	redtape	corruption	polstability	bureu_eff	gdp_grow_06_11_ave
judiciary	Pearson Correlation	1	.747**	.751**	.675**	.891**	.080
	Sig. (2-tailed)		.000	.000	.000	.000	.522
	N	66	66	66	66	66	66
redtape	Pearson Correlation	.747**	1	.749**	.633**	.891**	.000
	Sig. (2-tailed)	.000		.000	.000	.000	.999
	N	66	66	66	66	66	66
corruption	Pearson Correlation	.751**	.749**	1	.494**	.929**	.043
	Sig. (2-tailed)	.000	.000		.000	.000	.730
	N	66	66	66	66	66	66
polstability	Pearson Correlation	.675**	.633**	.494**	1	.653**	.102
	Sig. (2-tailed)	.000	.000	.000		.000	.413
	N	66	66	66	66	66	66
bureu_eff	Pearson Correlation	.891**	.891**	.929**	.653**	1	.027
	Sig. (2-tailed)	.000	.000	.000	.000		.830
	N	66	66	66	66	66	66
gdp_grow_06_11_ave	Pearson Correlation	.080	.000	.043	.102	.027	1
	Sig. (2-tailed)	.522	.999	.730	.413	.830	
	N	66	66	66	66	66	66

**. Correlation is significant at the 0.01 level (2-tailed).

For the total available sample of 66 nations, the MRA findings in Table 3 do not support the hypothesis. Significant partial regression (b) coefficients do not occur from entering all five variables (or anyone variable—not shown) into a regression equation to predict GDP growth. A correlation matrix (Table 4) shows that all five indexes for corruption, red tape, and efficiencies relate to each other significantly and none relate to 2006-2011 GDP growth.

Table 5a shows findings from using two of the variables and an interaction term for these variables in examining the data for a randomly created subsample of nations from the total data set. These findings provide modest support of the hypothesis that judicial inefficiency and corruption affects GDP growth, if a model includes these two variables with an interaction term (adjusted $R^2 = .133$, $p < .083$). The impact of both variables meets expectations that less corruption and more efficiency serve to increase GDP growth—both variables have b coefficients with t values greater than 2.00.

Table 5b shows the findings for the remaining data in the random split of the data for testing the same model. These results are similar to the other model though the b coefficient for only judiciary efficiency is significant statistically (t = 2.267, p < .030).

Table 5
Multiple Regression Analysis for Two Random Samples of Nations

a. First Sample (n = 30)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.472[a]	.223	.133	.08149

a. Predictors: (Constant), jud_X_corrupt_data, judiciary, corruption

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.050	3	.017	2.486	.083[b]
	Residual	.173	26	.007		
	Total	.222	29			

a. Dependent Variable: gdp_grow_06_11_ave

b. Predictors: (Constant), jud_X_corrupt_data, judiciary, corruption

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.251	.123		-2.032	.053
	judiciary	.043	.022	1.017	2.009	.055
	corruption	.049	.019	1.369	2.621	.014
	jud_X_corrupt_data	-.006	.003	-2.076	-2.435	.022

a. Dependent Variable: gdp_grow_06_11_ave

b. Second Sample (n = 36)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.441[a]	.195	.119	.04115

a. Predictors: (Constant), jud_X_corrupt_data, judiciary, corruption

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.013	3	.004	2.577	.071[b]
	Residual	.054	32	.002		
	Total	.067	35			

a. Dependent Variable: gdp_grow_06_11_ave

b. Predictors: (Constant), jud_X_corrupt_data, judiciary, corruption

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.014	.050		-.283	.779
	judiciary	.021	.009	1.066	2.267	.030
	corruption	.000	.009	.011	.022	.983
	jud_X_corrupt_data	-.002	.001	-1.119	-1.397	.172

a. Dependent Variable: gdp_grow_06_11_ave

Note. Using first sample model to predict GDP growth for second sample data: r = 0.067, p < .698, n = 36;
Using second sample model to predict GDP growth for first sample data: r = 0.004, p < .984, n = 30.

However, testing for predictive validity of the first model on the second holdout sample indicates that the model does not have acceptable predictive validity. The correlation appears at the bottom of Table 5a, 5b for the comparison of predicted and actual scores, r = 0.67 (p < .698). Using the estimated model from the second sample to predict the scores of the first sample leads to the same conclusion; the model provides more noise than information.

Table 6
Analysis of the Joint Lagged Impact of Judicial Inefficiency and Corruption on GDP Growth

Report

gdp_grow_06_11_ave

Judicial by Corruption 4 Grps	Mean	N	Std. Error of Mean
Highest	-.0042	9	.01951
High	.0723	31	.01184
Low	.0750	9	.03096
Lowest	.0471	17	.00885
Total	.0558	66	.00827

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
gdp_grow_06_11_ave * Judicial by Corruption 4 Grps	Between Groups	(Combined)	.045	3	.015	3.781	.015
		Linearity	.001	1	.001	.139	.711
		Deviation from Linearity	.045	2	.022	5.603	.006
	Within Groups		.248	62	.004		
	Total		.294	65			

Measures of Association

	R	R Squared	Eta	Eta Squared
gdp_grow_06_11_ave * Judicial by Corruption 4 Grps	.043	.002	.393	.155

Table 6 follows from taking an additional look at the data to test the hypothesis the countries scoring the highest in judicial inefficiency in combination with highest scores in corruption had lower GDP growth in comparison to the countries with low scores in judicial inefficiency. Nine countries had extremely high scores for both of these two variables—that is, scores of 1.0 for each variable (recall that 10.0 is equal to excellent performance and 1.0 is extremely low performance). For these nine countries, average GDP growth is equal to -0.0042 with a standard error of the mean equal to 0.0195 while average GDP growth is substantially higher for countries with lower scores on the combination of judiciary inefficiency and corruption. Details in Table 6 include a large effect size ($\eta^2 = .155$).

Note that the findings in Table 6 indicate that high GDP growth associate with many nations with relative high, but not the highest, levels of corruption, red tape, and inefficiency.

The mean findings are suggestive that some interesting patterns among the five efficiency indices are likely to occur in regards to influencing GDP growth.

To further explore the possibility that causal recipes of two or more variables of corruption, red tape, and government inefficiencies may influence GDP growth, each of the variables in the original data were calibrated using the computer software subroutine in the fsQCA software program. The procedure is analogous to performing a z-scale transformation of original data; see Ragin (2008) for details. The researcher needs to specify three values for calibrating an original scale into a fuzzy set scale: the original value covering 5 percent of the data values, 50 percent of the values, and 95 percent of the values. Table 7 provides the original values for these three points for each of the five independent variables and GDP growth.

Table 7
Summary Data for Country Efficiency, Corruption, Political Stability, and GDP Growth Study

Statistics

		judiciary	redtape	corruption	polstability	bureu_eff	gdp_grow_06_11_ave
N	Valid	66	66	66	66	66	66
	Missing	0	0	0	0	0	0
Mean		7.0970	6.1562	6.8920	7.5305	6.6974	.0558
Std. Error of Mean		.26634	.28159	.30673	.16792	.26437	.00827
Median		6.8750	6.0000	7.0000	7.6900	6.4200	.0433
Std. Deviation		2.16376	2.28762	2.49192	1.36421	2.14772	.06721
Minimum		2.00	1.25	1.50	3.25	2.00	-.09
Maximum		10.00	10.00	10.00	10.00	10.00	.27
Calibration values at:							
95%		**9.00**	**9.00**	**9.00**	**9.00**	**9.00**	**.2200**
50%		**6.875**	**6.000**	**7.000**	**7.690**	**6.42**	**.0433**
5%		**2.50**	**1.75**	**2.00**	**3.75**	**2.50**	**.0000**

In fsQCA, configural statements proposed by theory as well as all possible configural statements are testable using the fsQCA software. The program tests "logical and" statements of the possible combinations of the independent (simple to complex) antecedent expressions. The score for a "logical and" statement is equal to the lowest value of the simple antecedents in a statement containing two or more antecedent conditions. For example, Algeria appears in the first row of Appendix 2; the score for Algeria for the conjunctive statement judic_cal AND redtape_cal AND corrupt_cal is equal to 0.08. The score 0.08 is the lowest value among the three scores for Algeria for the respective simple antecedent conditions of judic_cal (0.63), redtape_cal (0.08), and corrupt_cal (0.23).

Using the software for fsQCA to test for the occurrence of different conjunctive statements (i.e., Mauro's "mechanisms"), six conjunctive statements (causal recipes) associate with high growth. In fsQCA, a researcher usually concludes that a model is informative when consistency is above 0.74 and coverage is between .25 and .65 (see Ragin, 2008).

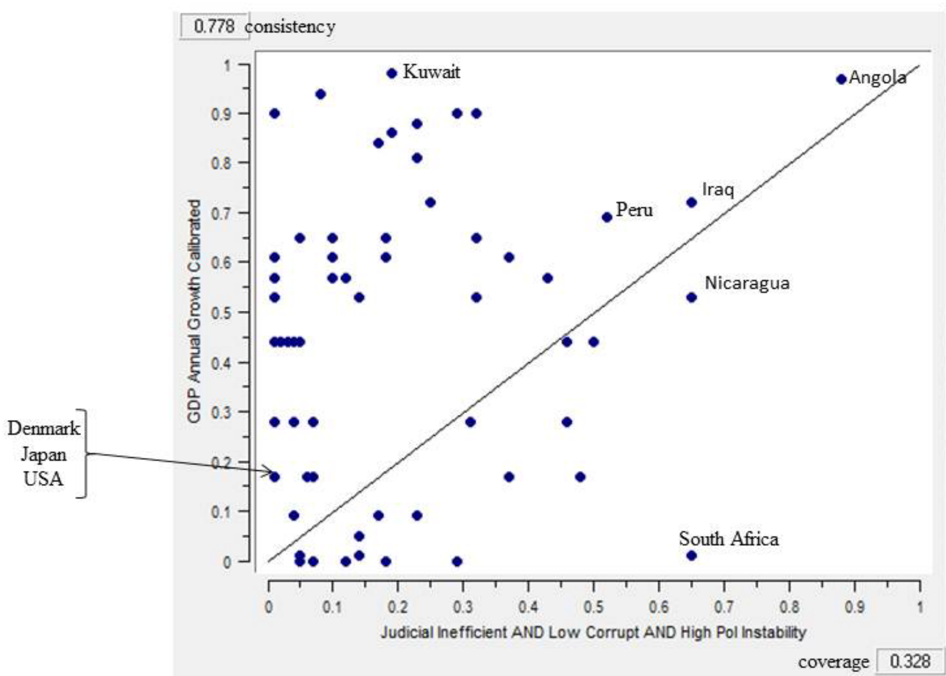
Table 8
Findings from fsQCA for Efficiency, Corruption, Red Tape, and GDP Growth

```
--- COMPLEX SOLUTION ---
frequency cutoff: 1.000000
consistency cutoff: 0.758904

                                                   raw       unique
                                                 coverage   coverage   consistency
                                                ---------- ---------- ----------
~judic_cal*corrupt_cal*~polstab_cal              0.328338   0.047343   0.778047
judic_cal*~redtape_cal*~corrupt_cal*~bureau_cal  0.370232   0.053815   0.778096
~judic_cal*redtape_cal*~corrupt_cal*~bureau_cal  0.331403   0.029292   0.819023
~judic_cal*~redtape_cal*polstab_cal*~bureau_cal  0.361035   0.030995   0.761494
redtape_cal*corrupt_cal*~polstab_cal*bureau_cal  0.349796   0.076976   0.743664
~judic_cal*redtape_cal*polstab_cal*bureau_cal    0.314033   0.019755   0.763877
solution coverage: 0.655313
solution consistency: 0.723853


Cases with greater than 0.5 membership in term ~judic_cal*corrupt_cal*~polstab_cal: angola (0.88,0.97),
  iraq (0.65,0.72), nicaragua (0.65,0.53), s_africa (0.65,0.01),
  peru (0.52,0.69)
Cases with greater than 0.5 membership in term judic_cal*~redtape_cal*~corrupt_cal*~bureau_cal: india (0.67,0.09),
  jamaica (0.66,0.88), algeria (0.63,0.09), colomiba (0.63,0.65),
  nigeria (0.63,0.94), trinidad (0.55,0.84), greece (0.54,0.57)
Cases with greater than 0.5 membership in term ~judic_cal*redtape_cal*~corrupt_cal*~bureau_cal: egypt (0.56,0.65),
  korea_s (0.56,0.65), panama (0.52,0.81)
Cases with greater than 0.5 membership in term ~judic_cal*~redtape_cal*polstab_cal*~bureau_cal: saudi_ara (0.63,0.86),
  turkey (0.62,0.72), italy (0.52,0.17), venezuela (0.51,0.9)
Cases with greater than 0.5 membership in term redtape_cal*corrupt_cal*~polstab_cal*bureau_cal: israel (0.75,0.57),
  chile (0.72,0.61), zimbabwe (0.71,0), s_africa (0.66,0.01)
Cases with greater than 0.5 membership in term ~judic_cal*redtape_cal*polstab_cal*bureau_cal: austria (0.78,0.44),
  argentina (0.52,0.17), taiwan (0.52,0.57)
```

Table 8 describes these six complex antecedent conditions. The first complex statement is the combination of high judicial inefficiency (indicated by the negation symbol, "~" for judicial efficiency, AND low corruption, AND high political instability. Negation scores in fsQCA are equal to 1 minus the original calibrated score. For example, for Algeria in Appendix 2, the negation score for the nation's judic_cal score equal to 0.63 is equal to 0.37.

Figure 4 is an XY plot for the first complex antecedent condition in Table 8 (representing not judicial efficiency AND low corruption AND not political stability) and GDP growth using calibrated scores. The bottom right quarter of the plot is nearly empty—indicating high sufficiency but not necessary condition for high GDP growth. South Africa has a relatively high score for the conjunctive statement but low GDP growth—this nation fails to support the conclusion that high GDP growth always occurs among nations with scores above 0.5 for the first complex antecedent condition. The research would want to study South Africa further to learn why this complex causal statement does not relate to high GDP growth for the country.

Figure 4
Example fsQCA Findings for Efficiency, Corruption, Red Tape, and GDP Growth



The findings in Table 8 indicate that high corruption may associate with high GDP growth and low corruption may associate with high GDP growth depending on other simple antecedents forming conjunctive statements with high and low corruption—the same conclusion applies for judicial efficiency and political stability. Understanding how corruption, red tape, and government inefficiencies affect GDP growth requires going beyond examining simple main effects and two-way interaction effects. Thinking and advancing theory using causal recipes are useful in particular in business research, as well as the study of chronic (i.e., measured) variables or a mix of chronic and manipulated (i.e., active or "experimental") variables.

The country identifications in findings for high GDP growth in Table 8 indicate that countries with consistently high calibrated scores across all five simple antecedent conditions do not have high calibrated scores for GDP growth—though they do have growth rates above the a separate group of nations with zero GDP growth rates. The conclusion is that nations low in corruption and high in all forms of government efficiencies do not appear to experience very high GDP growth rates and also avoid the bottom level of GDP growth rates. Very high growth rates may extend to a few years with a recipe that includes a high corruption while maintaining low inefficiencies or the reverse recipe—Iceland and Greece during 2002-2007 would be examples of such antecedent combinations and high GDP growth.

Consider the substantial benefit from studying the case findings in Figure 4 and Table 8. In fsQCA the researcher is able to generalize beyond the individual case but still identify individual cases in specific models relevant to her investigation.

The following observation by Ragin (2006, p. 7) relates to comparing the examination of conjunctive statements using MRA versus fsQCA: "The search for patterns of multiple conjunctural causation[s], a common concern of case-oriented researchers, poses serious practical problems for variable-oriented research." To investigate this type of causation with statistical techniques, it is necessary to examine high-level interactions (e.g., three-way interactions in the causal argument just described).

> However, these sophisticated techniques are very rarely used by variable-oriented researchers. When they are, they require at least two essential ingredients: (1) a very large number of diverse cases, and (2) an investigator willing to contend with a difficult mass of multi-collinearity. These techniques are simply not feasible in investigations with small or even moderate Ns, the usual situation in comparative social science. When Ns are small to moderate, causal complexity is more apparent, more salient, and easier to identify and interpret; yet it is also much less amenable to statistical analysis. (Ragin, 2006, pp. 7-8)

CONCLUSION

Tools shape theory as well as how a researcher goes about analyzing data. Taking time to read Gigerenzer's (1991) brilliant review on this perspective is worthwhile. Researchers need to embrace Armstrong's (2012) recommendations on testing for predictive validity and not just fit validity and not attempting to control the effects of other variables by simply adding them to produce regression equations with many terms.

Adopt McClelland (1998) approach in moving beyond the use of MRA and crafting and testing algorithms. Embrace Ragin's (2008) thinking and modeling in terms of conjunctive statements—think and test algorithms—rather than thinking only in net effects of variables on a dependent variable.

Post-print version of an article published in Journal of Business Research 66(4): 463-472. doi: 10.1016/j.jbusres.2012.12.021

Appendix 1
Computing Consistency and Coverage in Fuzzy-Set Qualitative Comparative Analysis

x_calibrated	y_calibrated	minimum (Xi, Yi)	
0.02	0.98	0.02	
0.03	0.98	0.03	
0.25	0.98	0.25	
0.97	0.98	0.97	
0.98	0.98	0.98	
0.02	0.5	0.02	
0.07	0.5	0.05	
0.1	0.5	0.05	
0.32	0.5	0.32	
0.41	0.5	0.41	
0.02	0.02	0.02	
0.03	0.02	0.02	
0.05	0.02	0.02	
0.07	0.02	0.02	
0.1	0.02	0.02	
∑ = 3.44	7.5	3.2	
Consistency (Xi ≤ Yi) =	∑ [min(Xi, Yi)]/∑(Xi)	3.2/3.44 =	0.93
Coverage (Xi≤Yi)=	∑[min(Xi,Yi)]/∑(Yi)	3.2/7.5 =	0.43

Note. Data are same as the final two columns in Table 1. The small differences in the consistency and coverage indexes in Table 2 and Figure 1b are due to rounding.

References

Armstrong, J.S. (2012). Illusions in regression analysis. International Journal of Forecasting, forthcoming.

CIA World Factbook. (2012). https://www.cia.gov/library/publications/the-world-factbook/index.html.

Cooper, R.G. (1992). The NewProd System: The industry experience. Journal of Product Innovation Management, 9, 113-127.

Cooper, R.G., & Kleinschmidt, E.J. (2007). Winning businesses in product development: The critical success factors. Research-Technology Management, 50, 52-66. Available at: http://www.proddev.com/downloads/working_papers/wp_6.pdf.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, & the ABC Research Group, Simple heuristics that make us smart (pp. 97–118). New York: Oxford University Press.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. Psychological Review, 98, 254-267.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. Topics in Cognitive Science, 1, 107–143.

Gladwell, M. (1996). The tipping point. The New Yorker, 72, 32-39.

Hotchkiss, E., Smith, D.C., & Strömberg, P. (2013). Private equity and the resolution of financial distress. Journal of Finance, forthcoming.

Lazarsfeld, P.F. (1937). Some remarks on the typological procedures in social research. Zeitschrift fur Sozialforschung, 6, 119-139.

Mauro, P. (1995). Corruption and growth. The Quarterly Journal of Economics (1995) 110 (3), 681-712. DOI: 10.2307/2946696.

McClelland, D. C. (1998). Identifying competencies with behavioral-event interviews. Psychological Science, 9, 331-3339.

Prado, A. (2011). Choosing among competing environmental and labor standards: An exploratory analysis of producer adoption. Working paper. New York: Stern School, New York University (aprado@stern.nyu.edu).

Ragin, C. (2006). Turning the tables: How case-oriented research challenges variable-oriented research. Comparative Social Research, 16, 27-42.

Ragin, C. (2008). Redesigning social inquiry: Fuzzy sets and beyond. Chicago: Chicago University Press.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. Psychological Review, 107, 358–367.

Woodside, A.G. (2011). Responding to the severe limitations of cross-sectional surveys: Commenting on Rong and Wilkinson's perspectives. Australasian Marketing Journal 19, 153–156.

Zellner, A. (2001). Keep it sophisticatedly simple. In Keuzenkamp, H. & McAleer, M. Eds. Simplicity, inference, and modeling: Keeping it sophisticatedly simple. Cambridge, UK: Cambridge University Press.