BAYESIAN ANALYSIS IN STRATEGIC MANAGEMENT RESEARCH:

TIME TO UPDATE YOUR PRIORS

BRIAN T. MCCANN
Associate Professor of Strategic Management
Owen Graduate School of Management
Vanderbilt University
401 21st Avenue South
Nashville, TN 37203
brian.mccann@owen.vanderbilt.edu

ANDREAS SCHWAB
Associate Professor and Dean's Faculty Fellow in Management
Ivy College of Business
3315 Gerdin Business Building
2167 Union Drive
Iowa State University
Ames, IA 50011
aschwab@iastate.edu

Acknowledgments
We are grateful for the advice and suggestions received from *SMR* editor Michael Leiblein and an anonymous *SMR* reviewer. We appreciate helpful feedback and comments on an earlier draft from Rahul Parsa.

BAYESIAN ANALYSIS IN STRATEGIC MANAGEMENT RESEARCH:

TIME TO UPDATE YOUR PRIORS

ABSTRACT

Bayesian statistical methods offer an important and increasingly endorsed alternative to traditional statistical significance testing. This paper presents a brief introduction to Bayesian methods, providing guidance to strategic management researchers who may wish to incorporate these methods into their research. We describe the advantages of Bayesian approaches and explain the steps involved in conducting and reporting a Bayesian analysis. For illustration, we provide a sample analysis, including all associated code using version 15 of Stata, which features significantly augmented Bayesian capabilities.

INTRODUCTION

Methodological rigor and the reliability of quantitative empirical research in strategic management continues to draw substantial attention (Bergh et al., 2017; Goldfarb and King, 2016). Recent articles have proposed and discussed a wide range of recommendations for improvements (Bettis et al., 2016). These recommendations, however, are focused on improvements within the existing paradigm of null hypothesis significance testing. In contrast, our aim is to provide guidance to strategic management researchers who are considering Bayesian statistics as an alternative paradigm for empirical investigation in their future research projects.

Limitations of traditional statistical significance tests and criticism of how they are applied have garnered increasing attention in the social sciences (e.g., Cohen, 1994; Gigerenzer, 2004; Schwab et al., 2011; Wasserstein and Lazar, 2016). Recently, several of these critiques have advocated for Bayesian analysis as an alternative to statistical significance tests (Gelman, 2015; Zyphur and Oswald, 2015). Bayesian methods have been productively applied across a wide number of research areas, leading to “a revolution in fields ranging from genetics to marketing” (Kruschke, Aguinis, and Joo, 2012: 722). Recent special issues on the topic have appeared in disciplines from psychology (Vandekerckhove, Rouder, and Kruschke, 2018) to econometrics (Kaufmann, Frühwirth-Schnatter, and van Dijk, 2019) and management (Zyphur, Oswald, and Rupp, 2015). Among other benefits, Bayesian methods produce results that integrate existing knowledge, focus greater attention on the size and the uncertainty of effects, and are easier to understand and communicate. More broadly, McKee and Miller’s (2015: 477) survey of a panel of 26 “institutional elites” (e.g., journal editors, former presidents of the Academy of Management and Strategic Management Society) reported that a “vast majority of scholars in our panel championed increased use of Bayesian methodology within the organizational sciences."

In spite of these endorsements, Bayesian studies in management research have remained rare (Zyphur, Oswald and Rupp, 2015; Kruschke, Aguinis, and Joo, 2012). Of the 849 articles listed by Web of Science as published in SMJ over the period 2010 – 2017, only four used Bayesian methods. Of the 1,467 articles published in AMJ and SMJ from 2001-2010, only three used Bayesian methods (Kruschke, Aguinis, and Joo, 2012). The lack of easy-to-use software to undertake Bayesian analysis has been a significant challenge to the application of these methods. This barrier, however, is rapidly eroding. In addition to continuing improvements in Bayes-specific software, commonly used software packages have begun to incorporate Bayesian capabilities. Stata Corporation, for example, introduced Bayesian commands into its software with the release of Version 14 in 2015. Version 15, released in 2017, included expanded Bayes functionality with increased ease-of-use (StataCorp, 2017).[1] Similarly, SPSS recently added Bayesian capabilities (IBM Corp., 2017).

A final challenge remains the limited familiarity of strategy researchers with these methods, a challenge this article targets. The discussion of Bayesian analysis and the introduction of illustrative examples in this article offers a starting point to strategic management researchers who wish to experiment with Bayesian methods. This article cannot, however, provide an exhaustive review. Instead, it provides an overview with references to more in-depth treatment of relevant topics. It begins by comparing Bayesian approaches to the existing frequentist paradigm; it then provides a review of the specific advantages of the Bayesian approach, which includes highlighting some of the limitations of statistical significance

[1] Stata released Version 16 in June of 2019 with additional improvements to its Bayesian capabilities.

applications in strategy research. It continues by discussing how Bayesian methods are particularly appropriate for research in strategic management while also noting some of the challenges associated with adoption of these methods. Finally, this article outlines the specific steps of a typical Bayesian analysis and concludes with a sample analysis that demonstrates the capabilities available in Stata to easily conduct and report Bayesian analyses. Appendices provide definitions of key terms along with all code used to conduct the sample analyses. In the end, the provided information should encourage researchers to consider Bayesian methods and to start exploiting their unique advantages.

Frequentist versus Bayesian Approaches

In general, statistical inference involves drawing conclusions based on data. More specifically, it is the process of drawing conclusions about population parameters from sample statistics. The currently dominant approach to statistical inference in management is based on a frequentist view of probability. Under this view, probability represents an estimate of the relative frequency of some outcome in a population. Say, for example, that we are interested in the relationship between prior acquisition experience and acquisition activity in a population. Frequentists assume that there is a certain specific, fixed relationship between prior experience and acquisition activity in this population. Observations of this relationship, however, are assumed to vary across samples drawn from this population (sampling variation). For randomly drawn samples, a random-sampling distribution can be used to estimate the probability of observing a specific parameter value or a more extreme parameter value in the collected data based on the assumption that the null-hypothesis of no effect is true. This probability statement is referred to as the *p*-value.

In statistical significance tests, two hypotheses are postulated, a null and an alternative. The null hypothesis typically asserts that a particular independent variable has zero or no effect on the dependent variable. The frequentists decide whether to reject this null hypothesis using the *p*-value as a test statistic. A *p*-value indicates the probability of obtaining a parameter estimate of similar or more extreme size, if the null hypothesis were true and the sample was randomly selected. If a *p*-value is very small (e.g., less than 0.05), the researcher may conclude that the difference of the observed parameter value from zero in the sample is unlikely to be just the result of random sampling. In this case, the researcher will reject the null hypothesis and concludes that the observed data provide evidence supportive of the hypothesized effect. As Bettis et al. (2016) noted, such null hypothesis significance tests are the core workhorses of quantitative empirical research in strategic management.

This frequentist approach provides probability statements for random sampling having affected observed estimates conditional on the null hypothesis being true. The objective is to reject the null hypothesis. Even if successful, frequentist approaches do not provide direct probability statements about hypothesized effects. In contrast, Bayesian approaches are not concerned with whether a null hypothesis should be rejected. Instead, Bayesian approaches use the observed statistics in the sample to estimate the probability of a hypothesized effect in the population. That is, a Bayesian would make a statement along the lines of "given prior experience and the observed data, there is a probability of 0.95 that β_1 is greater than 0.2." In this sense, Bayesian models lead to more meaningful interpretations of empirical data focused on the size and the uncertainty of hypothesized effects.

Inference in Bayesian analysis is based on posterior distributions, namely the probability of population parameters (θ) given the observed data (D), i.e., $P(\theta \mid D)$. Observed data is combined with prior information via Bayes' rule:

$$P(\theta \mid D) = \frac{P(D \mid \theta) P(\theta)}{P(D)} \tag{1}$$

In Bayesian inference, each of the four terms in Equation 1 has a specific name. $P(\theta)$ represents the expected probability for θ before data collection and is referred to as the *prior*. $P(D \mid \theta)$ is called the *likelihood* and represents the probability or likelihood of observing the particular data given the prior θ. The left-hand term of interest, $P(\theta \mid D)$, is called the *posterior*, and it is a combination of the *likelihood* and the *prior*. The term in the denominator of the right-hand side, $P(D)$, is the *marginal likelihood*, and it acts as a normalizing constant to scale the posterior density to make it a proper density. Because this term is simply a scaling constant, it is often excluded and Bayes' theorem is expressed as a proportion:

$$P(\theta \mid D) \propto P(D \mid \theta) P(\theta) \tag{2}$$

Equation 2 states that the posterior is proportional to the likelihood times the prior. That is, a posterior distribution represents a weighted average of information about the parameters in the observed data and knowledge about the parameters prior to observing the data; the data exert greater influence as the size of the sample increases (Gelman et al., 2014).

Historically, one of the major challenges to the application of Bayesian statistics has been the difficulty of computing posterior distributions, which typically involve very complex mathematical functions. The past few decades, however, have seen advances in both conceptual work and computing power that have fostered estimation via simulation. More specifically, posterior distributions of parameters in Bayesian estimation are now typically generated via Markov Chain Monte Carlo (MCMC) methods. MCMC algorithms, however, require thousands and sometime tens of thousands of iterations in order to adequately estimate posterior distributions. Hence, advancements in computing power have been critical to the growing applications of Bayesian methods. The basic idea is to home in on an approximate representation of the posterior by drawing a large number of representative random samples of parameter values from the observed data. As the number of iterations grows, the generated sets of estimates are expected to converge and to provide a reasonable approximation of the posterior distribution. Extensive reviews of MCMC simulation in Bayesian analysis are available (e.g., Gelman et al., 2014; Hahn, 2014; Kruschke, 2015; McElreath, 2016).

Benefits of Bayesian approach

Bayesian and frequentist analyses represent fundamentally different approaches to statistical inference. So why might a researcher prefer a Bayesian approach relative to the traditional approach? We contend that the Bayesian approach allows researchers to avoid several limitations of statistical significance tests while offering a variety of attractive properties, several of which are particularly relevant to strategic management research.

Limits of statistical significance testing. Numerous authors provide comprehensive discussions of the limitations of statistical significance tests and critiques of its current usage (e.g., Cohen, 1994; Gigerenzer, 2004; Schwab et al., 2011; Hubbard, 2004, 2015). For example, Schwab et al. (2011) highlight dichotomous evaluation based on a fixed *p*-value threshold and sensitivity of results to sample size. They also note that observing absolutely no effect is a highly implausible outcome of empirical tests. Hence, rejecting an implausible description of reality provides little or no valuable information. Wasserstein & Lazar (2016) emphasize the critique that researchers often apply statistical significance tests to inappropriate data and misinterpret the results. Below, we discuss a few examples of the most serious concerns in more detail. The first and perhaps most important concern with the frequentist approach is that it does not answer the question of usual interest to researchers. Typically, researchers want to know how likely an effect is given some observed data, i.e., $P(\theta \mid D)$. The frequentist approach instead assesses the probability of random sampling as an explanation for parameter values (or more extreme values) observed in the data assuming the null hypothesis were true, i.e., $P(D \mid \theta)$. These are two fundamentally different probabilities (Cohen, 1994).

Following the frequentist approach also creates challenges for communicating findings to practitioners. Researchers can either attempt to communicate the precise meaning of statistical significance (something along the lines of "if the effect we think is true is not actually true, we would expect to find an effect as large or larger than this being the result of random sampling only five percent of the time"), or run the risk that managers will misinterpret reported findings into direct probability statements about the hypothesized effect. In the end, it is also not clear whether rejecting a null hypothesis provides any profound new insights. As McShane et al. (2017: 6) summarize "given that effects are generally small and variable, the assumption of zero effect is false. Further, given that measurements are generally noisy and systematically so, even were an effect truly zero, it would not be in any study designed to test it."

An additional noteworthy concern is that statistical significance tests assume several underlying conditions apply to the analysis, many of which are infrequently met in practice. They assume, for example, that researchers have specified in advance stopping rules for collecting data (see Dienes, 2011 and Kruschke, Aguinis, and Joo, 2012: 734 for more extensive discussions). Researchers also need to specify hypotheses ex ante, and practices such as

HARKing invalidate statistical significance estimates. Underpowered studies tend to produce unreliable results (Cumming, 2011; Gelman and Carlin, 2014; Cohen 1994) and are common in psychology (Bakker, van Dijk, and Wicherts, 2012), economics (Ioannidis, Stanley, and Doucouliagos, 2017), and management research (Cashen and Geiger, 2004). It is unclear the extent to which these problems apply to strategy research, although several concerns have been raised in recent work (e.g., Bergh et al., 2017; Goldfarb and King, 2016). It may be that some of the conclusions of major studies in strategic management rely on unmet assumptions; if so, those conclusions are simply not supported by the reported empirical analyses.

In spite of these concerns and limitations, many management researchers apply and report statistical significance tests in their research reports and without much discussion or accounting for these potential limitations. Consequently, there is a general tendency to assign implicitly or explicitly more meaning to what a statistical significance test can tell us (e.g., Nuzzo, 2014). In response to a perceived widespread misapplication and misinterpretation of statistical significance tests, the American Statistical Association convened a panel of experts to draft a policy statement on *p*-values and statistical significance, which has subsequently been published in an effort to increase understanding (Wasserstein and Lazar, 2016). This policy statement tries to stop researchers from misinterpreting *p*-values and encourages researchers to start using other information to interpret consistency of data with their hypotheses.[2] The introduction to a more recent special issue of *The American Statistician* goes even further with a conclusion that "it is time to stop using the term 'statistically significant' entirely" (Wasserstein,

[2] Confidence intervals suffer from similar misinterpretation issues (Hoekstra et al., 2014; Morey et al., 2016). A 95% confidence interval produced by a frequentist analysis does not indicate that there is a 95% probability that the true parameter lies within that interval. The 95% indicates a property of the procedure not the parameter. More specifically, it indicates that if the same sampling procedure were used repeatedly, and interval estimates were computed for each sample, it would be expected that the true population parameter falls within the interval estimates in 95% of those repeated cases. This interpretation obviously runs into some of the same fundamental issues as do *p*-values, including the basic challenge of understanding and clearly communicating the meaning of these intervals.

Schirm and Lazar, 2019: 2). For those looking to move beyond the world of statistical significance testing, Bayesian methods provide a number of advantages.

Potential advantages of Bayesian analyses. In contrast to statistical significance tests, Bayesian analyses estimate the probability for hypothesized effects to occur, which is typically a primary question of interest to researchers. They quantify the level of confidence in the related hypothesis being true based on the data observed. In the process, Bayesian analyses also allow the incorporation of existing knowledge about the hypothesized effects via the prior distribution. The latter fosters a cumulative approach to research. In addition, Bayesian results lend themselves to easy communication to other researchers and practitioners (e.g., "there is a 95% probability that the effect is at least *x*"). Bayesian analyses also shift away from the dichotomous "effect or no effect" evaluation of statistical significance tests. Instead, Bayesian analyses estimate the distribution of effects. These posterior distributions provide detailed information about the size of effects and the uncertainty of effects. They can answer a wide variety of questions, such as: How strong is the central tendency of the effect? What is the probability of effects below or above any desired effect thresholds? How thick are the tails of the effect distribution? What are the range and functional form of the effect distribution?[3]

Bayesian approaches have a number of other attractive properties. They can help address multicollinearity issues (Leamer, 1973). If credible priors are available, meaningful estimates even from small samples are feasible. Kruschke, Aguinis and Joo (2012) also highlight that

[3] Questions such as these are naturally linked to issues of interest to strategy researchers. For example, a weak central tendency in an effect suggests substantial variance across firms while the probability of exceeding thresholds is potentially of substantial interest for a field concerned with firms' ability to earn returns in excess of their cost of capital. Hansen, Perry, and Reese's (2004) study of the relationship between firm performance and a variety of corporate actions (e.g., buying a new business unit, laying off staff) demonstrates how questions such as these may be answered. The central tendency of the relationship between buying a business unit and accounting returns indicated the most likely effect was an increase of approximately one-half of one percent with a probability of 0.81 of exceeding a threshold of zero effect. The distribution of the effect indicated a slightly left-skewed normal distribution that ranged from about -1.2 percent to 2.1 percent with somewhat fat tails.

Bayesian methods can deal with a variety of complex data structures and models. As Kruschke and Liddell (2018: 161) describe, "any parameterized model of data can have its parameters estimated via Bayesian methods," meaning that results can be generated from nearly all forms of statistical analysis familiar to strategy researchers.

The benefits of Bayesian approaches to strategic management research. A recent survey of senior strategy scholars reported in Leiblein and Reuer (2020: 18) noted a broad consensus among respondents that "much of the field's contribution lies in providing rigorous insights to general managers and other leaders of organizations." With their emphasis on estimating the distribution of effects, Bayesian methods are well aligned with the informational needs of practicing managers. Managers are not interested in the simple presence or absence of a particular relationship; instead they want to know about the magnitude of effects and the probability that those effects occur. Focusing on effect sizes and the distribution of effects is also consistent with recent changes in publication standards at outlets like *SMJ* as described by Bettis et al. (2016: 261): "*SMJ* will require in papers accepted for publication that authors explicitly discuss and interpret effect sizes of relevant estimated coefficients."

Bayesian methods are also flexible enough to apply across the broad topical coverage of the strategic management field. Moreover, wider adoption of these methods offers the potential to extend our thinking about the canonical problems in strategy research. Theories and methods are inherently linked. Methods affect not just how we test theories but also the questions we ask. Many of our current theories in strategic management are inherently dichotomous (i.e., effect or no effect) because that is what we test within the constraints of the existing statistical significance paradigm. Methods that devote attention to effect sizes and distribution of effects invite richer theorizing, extending and deepening our knowledge. Denrell, Feng, and Zhao

(2013), as just one recent example, demonstrate how Bayesian methods can be deployed to examine the relationship between sustained superior performance and superior capabilities, a central question in resource-based theories of strategic management. Their analyses, among other findings, indicate that sustained performance may not be a particularly reliable indicator in settings where chance events have enduring effects.

Finally, Bayesian approaches align well with one of the foundational principles of the field of strategic management. Among other questions, strategy research attempts to answer the question of why firms are different, one of the four fundamental issues of the field highlighted by Rumelt, Schendel and Teece (1991). Bayesian methods are particularly useful for examining differences in relationships for individual firms (e.g., Hansen, Perry and Reese, 2004) as opposed to frequentist methods that are more focused on average relationships. Mackey, Barney and Dotson (2017) provide a powerful example by applying Bayesian analysis to investigate differences in effects across firms within the context of the performance-diversification relationship. Their hierarchical Bayesian model allowed them to estimate firm-specific statistics and test the hypothesis that profit-maximizing firms vary in the diversification strategies they choose. Advantages such as this have led prominent researchers to note that "Bayesian methods are well matched to strategic management theory and often more appropriate than traditional frequentist approaches" (Jay Barney as quoted in McKee and Miller, 2015: 477).

Adoption Challenges

If the Bayesian approach is so useful, why has it not been more widely adopted? First, profound philosophical differences between Bayesians and frequentists have been the subject of extended historical debates (see Cohen, 1994; Gigerenzer, 2004; Gigerenzer and Marewski, 2015) and have limited the propagation of Bayesian methods. At a fundamental level, the two approaches view probability differently. Frequentist view probability in terms of relative frequency of an event in the long run. In contrast, Bayesians associate probability with degrees of belief or knowledge. These fundamentally different views imply deep and meaningful differences in how empirical data is interpreted and the corresponding statistical models to be used. McGrayne's (2011) book provides a rich narrative history of the development of Bayesian thinking and the associated conflicts between frequentists and Bayesians. Up to now, this conflict has been clearly "won" by the frequentists to the extent that statistical significance tests dominate management research. So, one challenge is the ubiquity of frequentist approaches and the comparative unfamiliarity of Bayesian methods to editors and reviewers. McKee and Miller's (2015) survey of institutional elites indicated that over 90 percent of respondents never or almost never encountered as reviewers papers that used Bayesian statistical analyses. Moreover, just over half of them indicated they would not be comfortable reviewing such articles. We do note, however, that the increasing number of Bayesian publications in the management literature suggests an increasing pool of knowledgeable reviewers.

Application of Bayesian methods by more strategy researchers will also require investments on the part of researchers to learn how to deploy these methods. Courses on Bayesian methods for organizational researchers are not nearly as widespread as those teaching frequentist methods (McKee and Miller, 2015), requiring researchers to take a more active role in searching out training opportunities. Courses in Bayesian statistics are offered by statistics departments at research universities, online course platforms and statistical software providers. In addition, collaborations with Bayesian researchers can provide rich learning opportunities. Wider adoption will also depend on the inter-relationship with theory, as strategy theories will need to evolve to align better with these methods.

Finally, we note that more practical considerations have also been a significant issue. Perhaps one of the largest barriers to increasing usage has been software-related. As Jebb and Woo noted in 2015 noted “many software packages supporting Bayesian statistics are not user-friendly relative to those commonly used in organizational research (e.g., SPSS, Stata).” Notably, however, this has substantially changed since 2015. To prove this point, the next sections describe the typical steps of Bayesian analysis and then demonstrates how those steps can be easily executed using version 15 of Stata (see Appendix C for a list of Stata resources).[4]

Steps in a Bayesian Analysis

As with many empirical research projects, a Bayesian study begins with identifying a research question of interest and collecting data appropriate to examine it. Once the data have been gathered and the researcher is ready to analyze, the first step is selection of a probability distribution and modeling approach appropriate for the particular dependent variable. The researcher should choose a distribution considering what is known about the dependent variable and the population that produced the observed data. For example, an analyst might opt for a Poisson distribution to model a count-based dependent variable or a binomial distribution to model a dichotomous dependent variable. Jebb and Woo (2015) provide a table that links various canonical forms of data (e.g., continuous, count, or categorical data) with their associated probability distributions for the likelihood, and the standard analytic approaches associated with each (e.g., linear, Poisson, or logistic regression). For researchers trained in frequentist methods, this step should be quite familiar.

[4] Zyphur and Oswald (2015) provide an excellent discussion of Bayesian approaches along with examples using Mplus. Jebb and Woo (2015) similarly demonstrate another software solution, BugsXLA an Excel add-in that works in combination with WinBUGS.

The next step in the process, however, is unique to Bayesian analyses. The researcher must specify a prior probability distribution for the effect of each independent variable in the model (typically referred to as “priors”). This requires both deciding the general distributional form (e.g., uniform, normal, lognormal) and the hyperparameters that further specify this distribution – for example, a normal distribution with the mean of zero and a variance of 1. The prior distribution represents what is known about the effect prior to observing the data (Lynch, 2007). Priors are generally classified into the two categories of informative or uninformative.

Uninformative priors assume an equal or relatively equal probability across the range of feasible effect outcomes. These priors have flat densities and have also been labeled vague, diffuse, or flat priors. Researchers tend to choose uninformative priors when they lack credible information justifying unequal outcome probabilities.

Informative priors, in contrast, specifically include existing information about parameter probability distributions. Informed priors foster parameter estimation that is a combination of both the observed data and existing knowledge represented by the informed prior. A clear benefit of informative priors is that they allow current analyses to be informed by past findings, fostering a cumulative approach to the production of knowledge. They are also particularly valuable in small sample research, given the high degree of uncertainty due to sampling error. Even vague but correct prior information about the expected distribution of effects will improve the quality of estimated posterior distributions (Gelman, 2009). Priors may be informed by theory, individual prior empirical studies, meta-analyses, or elicited expert opinion.[5] To date, the use of informative priors in Bayesian studies within management remains relatively rare.[6] The

[5] O’Hagan et al. (2006) and Gill (2015) provide advice on how to elicit expert opinions and convert that information into probability distributions.
[6] See LoPilato, Carter, and Wang (2015) for an exception that examines several different types of priors including informative ones.

development, accumulation, and use of informative priors represents an opportunity for future development as Bayesian methods spread through the field.

The dichotomous categorization of uninformative and informative priors, however, is a strong and often misleading simplification. There are no completely uninformative priors, so it makes more sense to think of how informative a prior is in continuous terms. Weakly informative priors allow the researcher to formulate priors based on rather limited information available about the phenomenon under study. In the end, the amount of influence of the prior on the posterior depends on the sample size. Hence, researchers have to be more concerned about the impact of priors when dealing with small samples. The use of large samples imply that priors tend to have relatively little influence on posterior distributions, a situation referred to as likelihood dominance (Lancaster, 2004). Ultimately, any prior selected does make assumptions and has the potential to affect posterior distributions. Hence, researchers should always explicitly describe and justify the priors they use and perform sensitivity analyses to assess and interpret their impact on posterior distributions.

As an example of how prior distribution choices can vary, consider examining the association between profitability and an indicator of whether a firm is owner-managed or managed by a hired CEO. If one assumes a normal distribution is a good representation of the distribution of the effect, specifying the prior requires selecting two hyperparameters, the mean and the variance. A relatively uninformative prior would presume no particular directionality for the effect and would specify a large variance. Such a choice might result in a prior distribution of $N(0, 10000)$, which represents the default normal prior with a mean of zero and a variance of 10,000 in the Stata software package (StataCorp, 2017). A more informative prior might recognize that profitability is generally constrained to a smaller range and specify hyperparameters reflective of this observation, say as $N(0, 9)$. This prior represents the researcher's belief, prior to data collection, that approximately 95% of the distribution of the parameter lies between -6 (mean less two standard deviations) and +6. Researchers could also draw on specific prior related research on the effect of owner management, such as Kulchina (2017), who reported a slightly positive mean and assign an even smaller variance, say $N(0.3, 0.4)$. Researchers are, of course, not constrained to priors based on the normal distribution. For example, a lognormal (0.3, 0.4) distribution represents a right-skewed distribution relative to $N(0.3, 0.4)$. Another alternative would be to select a uniform prior, e.g., $U(-1.0, 1.0)$ that assigns constant probability across a range of values. Selection of a prior might initially appear confusing to researchers only familiar with statistical significance tests. The lack of Bayesian management studies also means that existing research provides limited guidance and conventions for developing and justifying priors. This will change as Bayesian analysis becomes more established. The need to identify priors might initially feel like a challenge; in the end, however, it represents a strength of Bayesian analysis as it encourages the explicit integration of prior knowledge into current empirical investigation. This promotes knowledge accumulation and enables meaningful conclusions from smaller samples. Still, sensitivity analyses should always be performed to see how the choice of prior distribution affects estimation results.

Once the prior distributions have been specified, the observed empirical data is used to estimate model parameters. The Bayesian algorithms use the prior distribution of effects as their starting point and update this distribution using the observed data. Before interpreting results, however, it is important to conduct a number of checks of the adequacy of the estimation algorithms (see Depaoli and van de Schoot (2017) for a more extensive discussion).

Post-Estimation

As noted above, posterior distributions of parameters are produced via MCMC methods in Bayesian estimation. Kruschke (2015: 178) highlights that researchers should be concerned with both representativeness/convergence and stability/accuracy when using MCMC sampling to generate posterior distributions. For representativeness, values in the Markov chain should fully represent the range of the posterior, and they should not be influenced by arbitrary initial values of the chain. For stability and accuracy, estimates should be similar if MCMC estimation is duplicated (with a change in the seed number used for random number generation).

Representativeness and convergence. One primary concern to check with MCMC sampling is convergence (Lynch, 2007). Convergence refers to whether the Markov chain has adequately explored the targeted posterior distribution and is drawing mainly from the bulk of the distribution, but also still reasonably from the tails. A good chain will also mix rapidly; mixing refers to the rate at which a Markov chain explores the parameter space and reaches convergence. A variety of statistical tools are available to evaluate chain convergence (Sinharay, 2004). Convergence is typically first assessed by examining trace plots of estimates, which depict the value of each sampled parameter value (along the y-axis) across iterations (x-axis). Once a model has converged, the traceplot should show movement around the mode of the distribution. In other words, the chains should remain in the same general region but demonstrate vertical movement in the plot to indicate adequate sampling and representativeness.

Another recommended way to assess convergence is to use multiple chains that start from different initial values and then compare their results (Lynch, 2007). A common test using multiple chains is the Gelman-Rubin convergence diagnostic, which compares the variance between chains relative to variance within chains. The more similar the variance of each chain, the greater the degree of convergence. Fully converged chains have a value of 1. Large differences among the multiple chains are indicative of non-convergence (Gelman and Rubin, 1992; Kruschke, 2015), and non-convergence indicates that chain values are not representative of the posterior distribution. Problems with non-convergence are often addressable via increasing the length of the chain (i.e., increase the number of iterations). For more extensive discussion of methods of assessing convergence see Cowles and Carlin (1996) and Gelman et al. (2014).

Stability and accuracy. A researcher should also examine autocorrelation. Ideally, each draw from the posterior would provide a unique piece of information; however, the Markov nature of the algorithm means that it produces dependent samples by definition. Hence, samples are correlated. Autocorrelation starts at some positive value for early samples, but should decrease as the lag index increases in well-mixing chains. Whether this occurs can be assessed visually with autocorrelation plots. These plots show autocorrelation values across the sequence of generated samples. Stability of the estimates is also captured by the effective sample size (ESS). It reports the total number of independent MCMC samples out of the total number of MCMC samples. The closer the value is to the total, the less autocorrelation, and low autocorrelation indicates stability and accuracy of estimates.

Two final plots worth examining are the histogram and kernel density plots of the posterior distribution. Histograms depict the exact proportion of points in each bin while density plots average across overlapping intervals to generate a continuous depiction. In most cases, researchers expect plots to be relatively smooth and precise. Lumpy or multi-modal distributions provide important information for interpreting the stability and accuracy of hypothesized effects. Such distributions, however, can also be the result of convergence issues, which would suggest repeating estimations with longer burn-in rates and longer estimation runs (see Depaoli and van de Schoot (2017) for examples of both favorable and problematic plots). These plots can also be split across the first and second half of the samples and compared to ensure they are similar to each other and close to the overall density estimate.

Reporting of Results

The focus of reporting is the posterior distribution of the parameters of interest. In addition to representing these distributions graphically, researchers typically report both measures of central tendency and distribution for hypothesized effects. Central tendency is reported using the mean, median, and mode. The precision of the estimated posterior mean is commonly evaluated with the Monte Carlo standard error (MCSE), which reports how much error is in the estimate due to the fact that MCMC is used, and it quantifies how much one might expect the estimate to vary if the analysis were run again. As the number of MCMC iterations increases, MCSE decreases.

Bayesian researchers also typically report two types of intervals. The first is an equal-tailed credibility interval, which reports an interval where the probability of the parameter being below the interval is as likely as being above it. For example, a 95% equal-tailed interval has two tails of 2.5% and provides a range such that the probability the parameter lies within that range given the data is 95%. It is also important to note that while this credibility interval may sound similar to a 95% confidence interval produced by frequentist methods, the two are fundamentally different. Confidence intervals do not provide a probability statements for hypothesized effects. In addition, Bayesian credibility intervals consider their bounds fixed and estimated statistics as random variables within these bounds. Frequentist confidence intervals consider their bounds as random variables and statistics as fixed values. The second frequently reported interval is the highest posterior density (HPD) interval, which provides an interval of the shortest width for a particular probability level. For symmetric posterior distributions, the two intervals should be quite similar; however, the HPD interval may be preferable in cases of asymmetric or skewed posterior distributions.

Kruschke (2015: 336) suggests comparing the HPD interval to a particular "region of practical equivalence" (ROPE) for more sophisticated decisions regarding the parameter of interest. Testing a hypothesis against the baseline of absolutely no effect often represents a less meaningful evaluation (Cohen, 1990; Schwab and Starbuck, 2012). For example, a researcher might be interested in whether a regression coefficient is practically non-zero. The research might specify that in the specific context coefficient values of -0.02 to +0.02 are practically equivalent to zero (a ROPE of {-0.02, 0.02}). The non-zero hypothesis would be accepted if the 95% HPD interval falls completely outside the ROPE. In contrast, the zero effect hypothesis would be accepted if the 95% HPD interval falls completely inside the ROPE (i.e., all of the 95% most credible values are practically equal to zero). When the HPD interval and the ROPE overlap, the researcher withholds a decision. Under this approach, both non-zero effect and zero-effect hypotheses with various effect size thresholds can be accepted (or rejected) based on corresponding effect size and effect probability estimates. This is a clear departure from frequentist approaches that only center on the possible rejection, but not acceptance, of null hypotheses. In contrast, Bayesian analysis can provide probability estimates for both the zero-effect and non-zero effect hypotheses being true.

Aggregate measures of distribution can help describe and compare distribution characteristics. Researchers, however, should always also report graphs of the posterior distribution because they provide the most powerful way to communicate rich and comprehensive distributional information, information that enables the simultaneous evaluation of effect size and effect uncertainty (Schwab, 2018; Greve, 2018). Beyond the dichotomous "effect or no-effect" evaluations outlined above, graphs of posterior distributions provide detailed probability information for any hypothesized effect being larger than any effect size level considered of substantive or practical importance. The available Bayesian software solutions all offer routines for creating posterior distribution graphs.

How to Execute Bayesian Analysis with Stata: An Example

The introduction of MCMC-based Bayesian routines in commonly used software packages has substantially reduced the time and effort needed to execute Bayesian analyses. For a simple illustration, we perform a Bayesian analysis using a publicly available small data set accompanying Wooldridge (2013), which permits a straightforward study of the relationship between firm performance and CEO compensation. The data include the salaries of 209 CEOs for the year 1990, as well as firm revenue and performance metrics like return on equity and return on sales (3-year averages from 1988-1990). Additional measures include dummy variables indicating whether the firm was primarily an industrial, financial, consumer product, or utilities firm (transportation being the omitted industry). All code to conduct the analysis, which can be generated easily in Stata via menu-driven choices, is included in Appendix B. Beginning with version 15 of Stata, users may easily estimate a wide variety of models using Bayesian approaches. Model specification is as easy as typing "*bayes:*" in front of any of 46 estimation commands. A broad assortment of options can be specified using easy-to-follow, menu-driven commands.

In our example, the model is a multivariate regression, where CEO salary (logged) is predicted from sales (logged) and the consumer products industry dummy.[7]

[7] To facilitate explanation, we present a relatively simple model; however, this should not be interpreted as limitation of Stata, which can fit a wide variety of models. Logit or probit models are easily estimated in Stata as are models for count-based dependent variables (e.g., Poisson or negative binomial). Sample selection models (e.g., OLS and probit models with Heckman corrections) and multilevel models are available as well.

$$lsalary_i = \beta_0 + \beta_1 lsales_i + \beta_2 consprod_i + \varepsilon_i$$

The residuals are normally distributed with a mean of 0 and constant variance, represented by σ^2. To estimate this model using Bayesian methods, the first steps are to specify a likelihood and prior distributions for all parameters in the model. For a linear regression model, such as this one, the normal distribution is a common choice for the likelihood; an alternative approach to accommodate concerns such as possible outliers would be to fit a robust linear regression model using the t-distribution (which features heavier tails) for the likelihood. The model has four unique parameters, $\beta_0, \beta_1, \beta_2,$ and σ^2. As an initial estimation approach, the model employs the default priors specified by Stata 15.1; the default for regression coefficients are normal distributions with mean 0 and variance 10,000. The default for the variance is an inverse gamma with shape parameters of 0.01 and 0.01. These represent uninformative priors. Alternatively, researchers can (and often should) consider priors based on findings in prior studies, theories or expert judgements. Uninformative priors are used here for simplicity and illustrative reasons.

To illustrate the effects of different numbers of MCMC iterations, three different analyses were conducted with totals of 12,500 (the Stata default setting), 62,500, and 125,000 iterations. Because it can take some time for MCMC algorithms to begin sampling more heavily from the modal region of the posterior distribution (meaning toward the center and not the tails), it is routine to discard an initial set of several hundreds or thousands of iterations (referred to as the burn-in period). In each of the analyses run, the first 20% of the iterations were discarded. Table 1 shows that the mean values change very little across the analyses; however, they are estimated with increasing precision, as indicated by the decreasing MCSE values.

Examination of the trace plots for each parameter (produced with *bayesgraph trace _all*) show relatively constrained horizontal bands with adequate movement above and below the midpoint of the bands, indicating reasonable convergence and mixing. As examples, the trace plots for the *lsales* parameter for each estimation are included as Figures 1a-1c. Autocorrelation plots from Model 3 (produced with *bayesgraph ac _all*) shown in Figures 2a-2d demonstrate rapidly decreasing autocorrelations at larger lag values.

-------*Insert Table 1, Figures 1a-1c, Figures 2a-2d about here*-------

Examination of the Gelman-Rubin statistic, which assesses whether changes converge on the same space even when starting from different initial values, provided further support to the conclusion of adequate convergence. Three additional, separate chains were run with different initial starting values, and the user-written Stata command *grubin*, was used to compute the Gelman–Rubin diagnostic. The *grubin* command is unfortunately not written to be compatible with the simple approach of merely adding a *bayes* prefix to a typical regression command. It necessitates the use of the more general *bayesmh* command that requires the user to specify all aspects of the estimation.[8] As seen in Appendix B, however, this is relatively straightforward to accomplish. The Gelman-Rubin convergence statistic indicated no convergence concerns.

How should researchers interpret the results of this Bayesian analysis? Based on Model 3 of Table 1, Figure 3a shows the posterior distribution of the *lsales* coefficient, easily graphed using the Stata command of *bayesgraph histogram {lsales}.* With both the dependent and independent variable logged, the coefficient represents the estimated percent change in sales for a one percent increase in CEO salary. That is, the mean value of *lsales* of 0.259 suggests that a ten percent increase in sales is associated with a 2.6% increase in CEO salary. For a CEO operating a company at the median salary level of $1.04 million, increasing sales by 10% is associated with a salary increase of approximately $27,000. The mean value of *consprod* of

---

[8] Version 16 of Stata includes a new *bayesstats grubin* command to allow more direct post-estimation calculation of the Gelman-Rubin convergence diagnostic.

0.286 indicates that consumer-product CEOs make approximately 33% more on average relative to CEOs in other industries. Figures 3a and 3b also indicate the 95% credibility interval for each coefficient, which provides the upper and lower limits for the middle 95% of the distribution; given the relatively symmetric natures of the distributions, the highest posterior density (HPD) intervals are quite similar.

-------*Insert Figures 3a and 3b about here*-------

Bayesian analysis also permits assessing a wide variety of different hypotheses that may be of interest, and these post-estimation tests are easily implementable in Stata using the *bayestest interval* command. Say, for example, a researcher is interested in assessing the probability that the effect of being a consumer product CEO relative to other industries is 20% or higher. This corresponds to calculating the probability that $\beta_2 > 0.182$ [9], and the results indicate this probability is 0.918. Similarly, one could ask the probability that the sales effect is less than 0.2% for a 1% change in sales, i.e., the probability that $\beta_1 < 0.2$; the corresponding probability is 0.038.

Researchers can also evaluate the relative support for different models using "Bayes factors" (Kass and Raftery, 1995). A Bayes factor is a model comparison statistic that quantifies the relative evidence two models receive given the collected data. An attractive feature is that the models do not have to be nested to be compared. The interpretation of a Bayes factor (BF) is relatively straightforward. For example, if $BF(M_1, M_2)$=5, there is 5 times more evidence in the data supporting Model A compared to Model B. Estimating Bayes factors is easily accomplished in Stata using the *bayesstats ic* command after running multiple models. To demonstrate this, we

[9] With a logged dependent variable, a one-unit change in the independent variable (from 0 to 1) has an effect of exp(b) on the DV in unlogged units. A beta of 0.182 is equivalent to a 20% increase in sales (exp(0.182)=1.199).

return to the salary data, which includes two profitability measures: return on equity (ROE) and return on sales (ROS). Comparing a model that utilizes ROE to one that utilizes ROS leads to the logged $BF(M_{ROE}, M_{ROS})$ =7.14, indicating greater evidence for the ROE model. For more extensive examples of the use of Bayes factors in management models, see Andraszewicz et al. (2015), Braeken, Mulder, and Wood (2015), and Kass and Raftery (1995).

Finally, sensitivity analyses to evaluate the impact of priors on the posterior distribution often make sense. The code in Appendix B shows how alternative priors are easily accommodated by simply including an additional option in the regression command.[10] Default priors continue to be used for any parameters not listed. The results reported in Table 2 indicate that in our case reported estimates are largely unaffected by the specification of more informative priors, e.g., priors with reduced variance.

-------*Insert Table 2 about here*-------

CONCLUSION

Bayesian methods provide a powerful alternative to traditional frequentist approaches. Most importantly, they enable researchers to estimate the probability that an effect of a particular size is present. Moreover, they facilitate incorporating prior existing knowledge about the phenomenon, thereby fostering a cumulative approach to research. Bettis and Blettner (2020: 6) cite a key goal of strategy research to be making "our scholarship more meaningful to the accumulation of knowledge." Leiblein and Reuer (2020: 3) similarly note that the "construction of a robust, cumulative body of knowledge" is a key opportunity and challenge facing the field

[10] To follow the default estimation approach, the additional *block* option must also be specified. When priors are not specified, Stata's default is to sample all regression coefficients in one block. Sampling parameters simultaneously in one block generally increases the efficiency of the sampler, particularly when parameters are correlated. If a non-default prior is specified for a particular parameter, it will be sampled in a separate block unless the *block* option is included.

of strategic management. Bayesian approaches are strongly aligned with this philosophy of building upon what is already known in order to accumulate and extend knowledge.

Results from Bayesian analysis also naturally lend themselves to a stronger focus on effect sizes rather than the dichotomous question of the presence or absence of effects. This focus answers recent calls in strategy research to pay greater attention to effect sizes (Bettis et al., 2016), and it is consistent with the needs of practicing managers. The first sentence of the Division Statement for the STR Division of the Academy of Management clearly emphasizes the link to practice: "The division encourages and supports the development and dissemination of knowledge relevant to general managers." What is most relevant to practicing managers is not the simple presence or absence of effects but rather the size of effects and the uncertainty associated with those effects. Bayesian methods are much better suited to meeting the call to produce managerially relevant knowledge.

The value of Bayesian methods has contributed to their increasing prominence in several academic fields. Regardless of the state of change in other fields, the "Bayesian revolution" has yet to reach the field of strategic management. Attention to Bayesian methods, however, is growing as evidenced by the increasing number of Bayesian studies and overviews in the management literature. Still, the number of scholars with expertise in Bayesian statistics is limited (McKee and Miller, 2015). Hence, Bayesian studies have to provide clear and convincing descriptions of these less familiar methods. As such, conducting and publishing Bayesian studies requires additional time and effort on the part of researchers. Still, many positive signs indicate that these obstacles are eroding. Relatively straightforward, easy-to-follow advice on how to conduct, review, and report Bayesian studies is now available to help guide authors and reviewers (Hahn, 2014; Kruschke, 2015; Gelman et al., 2014; Gill, 2015). Support for Bayesian methods is strong among many leading management scholars in the field (McKee and Miller, 2015). Equally important, one significant obstacle that has prevented the broader adoption of these methods, namely the lack of easy-to-use, point-and-click software that implements Bayesian analysis, is quickly disappearing. In particular, strategy researchers who use Stata will find that this obstacle has been largely removed with the release of Stata version 15, which supports easy Bayesian estimation of over 45 different models using menu-driven commands. These advancements are, however, not limited to Stata as SPSS, MPlus, R and other programs are continuously introducing improved Bayesian routines. Given these positive developments, strategy scholars with skeptical prior beliefs about the potential of Bayesian studies should definitely update those beliefs in light of this new evidence.

REFERENCES

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E. J. (2015). An introduction to Bayesian hypothesis testing for management research. *Journal of Management*, 41(2), 521-543.

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, 7(6), 543-554.

Bergh, D. D., Sharp, B. M., Aguinis, H., & Li, M. (2017). Is there a credibility crisis in strategic management research? Evidence on the reproducibility of study findings. *Strategic Organization*, 15(3), 423-436.

Bettis, R., & Blettner, D. 2020. Strategic reality today: Extraordinary past success, but difficult challenges loom. *Strategic Management Review*, 1(1): forthcoming.

Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchell, W. (2016). Creating repeatable cumulative knowledge in strategic management. *Strategic Management Journal*, 37(2), 257-261.

Braeken, J., Mulder, J., & Wood, S. (2015). Relative effects at work: Bayes factors for order hypotheses. *Journal of Management*, 41(2), 544-573.

Cashen, L. H., & Geiger, S. W. (2004). Statistical power and the testing of null hypotheses: A review of contemporary management research and recommendations for future studies. *Organizational Research Methods*, 7(2), 151-167.

Cohen J. (1994) The earth is round ($p < 0.05$). *American Psychologist*, 49, 997–1003.

Cowles, M. K., & Carlin B. P. (1996). Markov chain Monte Carlo convergence diagnostic: A comparative review. *Journal of the American Statistical Association* 91(434), 883–904.

Cumming, G. (2011). *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York: Routledge.

Denrell, J., Fang, C., & Zhao, Z. (2013). Inferring superior capabilities from sustained superior performance: A Bayesian analysis. *Strategic Management Journal*, 34(2), 182-196.

Depaoli, S., & van de Schoot, R. (2017). Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological Methods*, 22(2), 240-261.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6(3), 274-290.

Gelman, A. (2009). Bayes, Jeffreys, prior distributions and the philosophy of statistics. *Statistical Science*, 24(2), 176-178.

Gelman, A. (2015). The connection between varying treatment effects and the crisis of unreplicable research: A Bayesian perspective. *Journal of Management*, 41(2), 632-643.

Gelman, A., & Carlin, J. (2014). Beyond power calculations: assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, 9(6), 641-651.

Gelman, A., Carlin, J. B. Stern, H. S. Dunson, D. B. Vehtari, A. & Rubin, D. B. (2014). *Bayesian Data Analysis*. 3rd ed. Boca Raton, FL: Chapman & Hall/CRC.

Gelman, A., & Rubin D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–511.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33(5), 587-606.

Gigerenzer, G., & Marewski, J. N. (2015). Surrogate science: The idol of a universal method for scientific inference. *Journal of Management*, 41(2), 421-440.

Gill, J. (2015). *Bayesian Methods: A Social and Behavioral Sciences Approach*. Boca Raton, FL: Taylor and Francis/CRC.

Goldfarb, B., & King, A. A. (2016). Scientific apophenia in strategic management research: Significance tests & mistaken inference. *Strategic Management Journal*, 37(1), 167-176.

Greve, H. R. (2018). Show me the data! Improving evidence presentation for publication. *Management and Organization Review*, 14(2), 423-432.

Hahn, E. D. (2014). *Bayesian Methods for Management and Business: Pragmatic Solutions for Real Problems*. Hoboken, NJ: John Wiley & Sons.

Hansen, M. H., Perry, L. T., & Reese, C. S. (2004). A Bayesian operationalization of the resource- based view. *Strategic Management Journal*, 25(13), 1279-1295.

Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E. J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, 21(5), 1157-1164.

Hubbard, R. (2004). Alphabet soup: Blurring the distinctions between p's and alpha's in psychological research. *Theory and Psychology*, 14(3), 295-327.

Hubbard, R. (2015). *Corrupt Research: The Case for Reconceptualizing Empirical Management and Social Science*. Thousand Oaks, CA: SAGE Publications.

IBM Corp. (2017). IMB SPSS Statistics for Windows version 25. Armonk, NY: IBM Corp.

Ioannidis, J., Stanley, T. D., & Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(605).

Jebb, A. T., & Woo, S. E. (2015). A Bayesian primer for the organizational sciences: The "two sources" and an introduction to BugsXLA. *Organizational Research Methods*, 18(1), 92-132.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773-795.

Kaufmann, S., Frühwirth-Schnatter, S., & van Dijk, H. (2019). Editorial introduction on complexity and big data in economics and finance: Recent developments from a Bayesian perspective. *Journal of Econometrics*, 210(1), 1-3.

Kruschke, J. K. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*, 2nd ed.. London: Academic Press.

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods*, 15(4), 722-752.

Kruschke, J. K., & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic Bulletin & Review*, 25(1), 155-177.

Kulchina, E. (2017). Do foreign entrepreneurs benefit their firms as managers?. *Strategic Management Journal*, 38(8), 1588-1607.

Lancaster, T. (2004). *An Introduction to Modern Bayesian Econometrics*. Malden, MA: Blackwell Publishing.

Leamer, E. E. (1973). Multicollinearity: A Bayesian interpretation. *The Review of Economics and Statistics*, 55(3), 3371-380.

Leiblein, M., & Reuer, J. 2020. Foundations and futures of strategic management. *Strategic Management Review,* 1(1): forthcoming.

LoPilato, A. C., Carter, N. T., & Wang, M. (2015). Updating generalizability theory in management research: Bayesian estimation of variance components. *Journal of Management*, 41(2), 692-717.

Lynch, S. M. (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer.

Mackey, T. B., Barney, J. B., & Dotson, J. P. (2017). Corporate diversification and the value of individual firms: A Bayesian approach. *Strategic Management Journal*, 38(2), 322-341.

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Boca Raton, FL: CRC Press

McGrayne, S. B. (2011). *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, & Emerged Triumphant from Two Centuries of Controversy*. New Haven, CT: Yale University Press.

McKee, R. A., & Miller, C. C. (2015). Institutionalizing Bayesianism within the organizational sciences: A practical guide featuring comments from eminent scholars. *Journal of Management*, 41(2), 471-490.

McShane, B. B., Gal, D., Gelman, A., Robert, C., & Tackett, J. L. (2017). Abandon statistical significance. arXiv preprint arXiv:1709.07588.

Morey, R. D., Hoekstra, R., Rouder, J. N., & Wagenmakers, E. J. (2016). Continued misinterpretation of confidence intervals: response to Miller and Ulrich. *Psychonomic Bulletin & Review*, 23(1), 131-140.

Nuzzo, R. (2014). Scientific method: Statistical errors. *Nature News*, 506(7487), 150.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., ... & Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*. Hoboken, NJ: John Wiley & Sons.

Schwab, A. (2018). Investigating and communicating the uncertainty of effects: The power of graphs. *Entrepreneurship, Theory and Practice*, 42(6), 823–834.

Schwab, A., Abrahamson, E., Starbuck, W. H., & Fidler, F. (2011). Researchers should make thoughtful assessments instead of null-hypothesis significance tests. *Organization Science*, 22(4), 1105-1120.

Schwab, A., & Starbuck, W. H. (2012). Using baseline models to improve theories about emerging markets. In C. L. Wang, D. J. Ketchen, & D. Bergh (eds.), *Research Methodology in Strategy and Management*, Vol. 7. Bingley, UK: Emerald, 3-33.

Sinharay, S. (2004). Experiences with Markov Chain Monte Carlo convergence assessment in two psychometric examples. *Journal of Educational and Behavioral Statistics*, 29(4), 461-488.

StataCorp (2017). *Stata Bayesian Analysis Reference Manual*, Release 15. College Station, TX: StataCorp LLC.

Vandekerckhove, J., Rouder, J. N., & Kruschke, J. K. (2018). Bayesian methods for advancing psychological science. *Psychonomic Bulletin & Review*, 25 (1), 1-4.

Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a World Beyond "p< 0.05". *The American Statistician*, 73:sup1, 1-19.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *American Statistician*, 70(2), 129-133.

Wooldridge, J.M. (2013). *Introductory Econometrics: A Modern Approach*, 5th ed.. Mason, OH: South-Western, Cengage Learning.

Zyphur, M. J., & Oswald, F. L. (2015). Bayesian estimation and inference: A user's guide. *Journal of Management*, 41(2), 390-420.

Zyphur, M. J., Oswald, F. L., & Rupp, D. E. (2015). Rendezvous overdue: Bayes analysis meets organizational research. *Journal of Management*, 41(2), 387-389.

Table 1. Bayesian Estimates

Model 1 (*MCMC iterations=12,500; MCMC sample size=10,000*)

Coefficients	Mean	SD	MCSE	Median	95% Cr. Interval		ESS
lsales	0.2579	0.0342	0.0013	0.2577	0.1934	0.3252	697
consprod	0.2848	0.0785	0.0029	0.2843	0.1245	0.4378	710
_cons	4.7304	0.2861	0.0110	4.7291	4.1569	5.2630	672
sigma2	0.2411	0.0247	0.0006	0.2400	0.1982	0.2942	1876

Model 2 (*MCMC iterations=75,000; MCMC sample size=62,500*)

Coefficients	Mean	SD	MCSE	Median	95% Cr. Interval		ESS
lsales	0.2586	0.0333	0.0004	0.2582	0.1941	0.3249	8227
consprod	0.2861	0.0741	0.0011	0.2865	0.1399	0.4323	4785
_cons	4.7239	0.2797	0.0031	4.7246	4.1680	5.2631	8092
sigma2	0.2409	0.0239	0.0002	0.2395	0.1980	0.2916	12755

Model 3 (*MCMC iterations=125,000; MCMC sample size=100,000*)

Coefficients	Mean	SD	MCSE	Median	95% Cr. Interval		ESS
lsales	0.2589	0.0333	0.0003	0.2588	0.1937	0.3245	12662
consprod	0.2862	0.0740	0.0009	0.2864	0.1394	0.4312	7379
_cons	4.7218	0.2791	0.0025	4.7207	4.1713	5.2663	12611
sigma2	0.2409	0.0240	0.0002	0.2395	0.1984	0.2924	20759

Table 2. Bayesian Estimates using Alternative Priors

Model 3 (*Default priors used for all model parameters*)

Coefficients	Mean	SD	MCSE	Median	95% Cr. Interval		ESS
lsales	0.2589	0.0333	0.0003	0.2588	0.1937	0.3245	12662
consprod	0.2862	0.0740	0.0009	0.2864	0.1394	0.4312	7379
_cons	4.7218	0.2791	0.0025	4.7207	4.1713	5.2663	12611
sigma2	0.2409	0.0240	0.0002	0.2395	0.1984	0.2924	20759

Model 4 (*Prior for lsales and consprod: normal (0,10); defaults for other parameters*)

Coefficients	Mean	SD	MCSE	Median	95% Cr. Interval		ESS
lsales	0.2592	0.0337	0.0004	0.2592	0.1931	0.3260	6902
consprod	0.2864	0.0748	0.0008	0.2861	0.1407	0.4336	8887
_cons	4.7192	0.2829	0.0034	4.7183	4.1603	5.2780	7008
sigma2	0.2408	0.0241	0.0002	0.2394	0.1981	0.2929	20077

Model 5 (*Prior for lsales and consprod: normal (0,0.1); defaults for other parameters*)

Coefficients	Mean	SD	MCSE	Median	95% Cr. Interval		ESS
lsales	0.2558	0.0335	0.0004	0.2558	0.1905	0.3215	8659
consprod	0.2714	0.0723	0.0008	0.2710	0.1286	0.4124	8965
_cons	4.7509	0.2808	0.0030	4.7513	4.1972	5.2997	8741
sigma2	0.2409	0.0240	0.0002	0.2394	0.1984	0.2921	20213

Figure 1. Trace Plots of *lsales* Parameter from Models 1 – 3

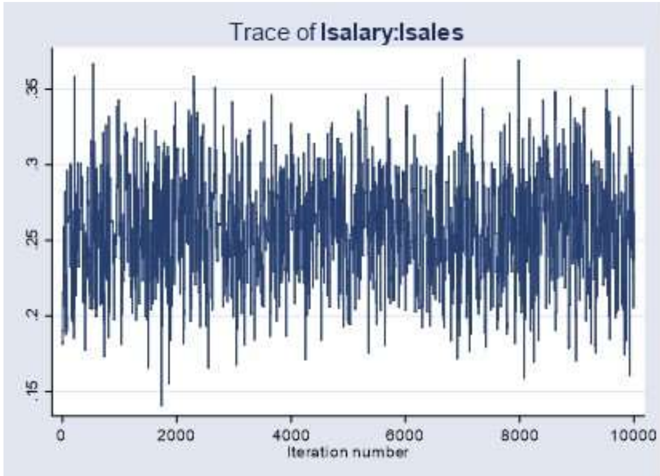

Figure 1a

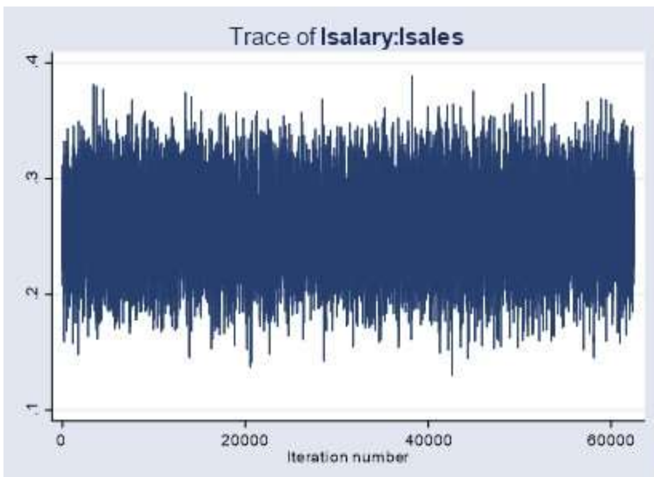

Figure 1b

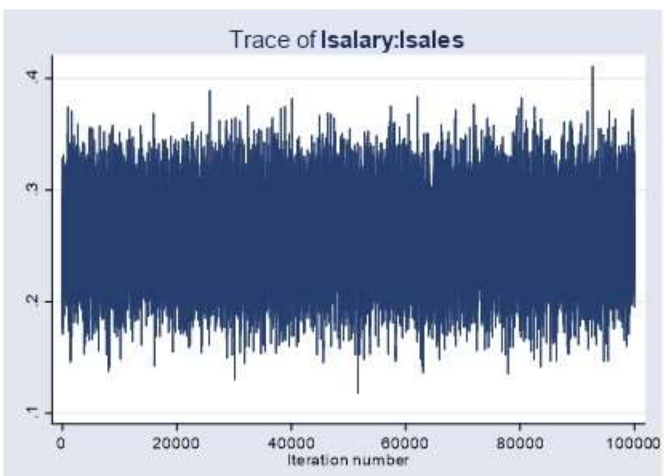

Figure 1c

Figure 2. Autocorrelation Plots of Parameters from Model 3

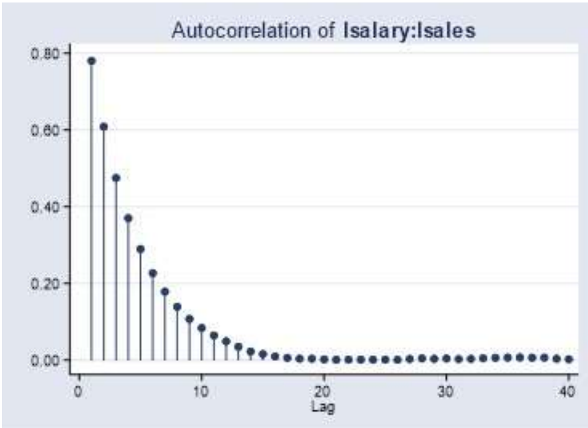

Figure 2a

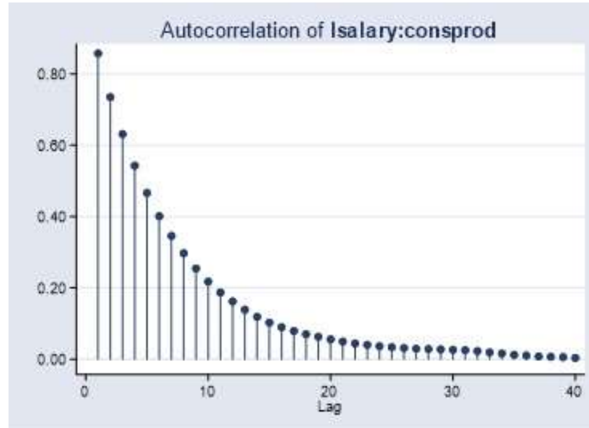

Figure 2b

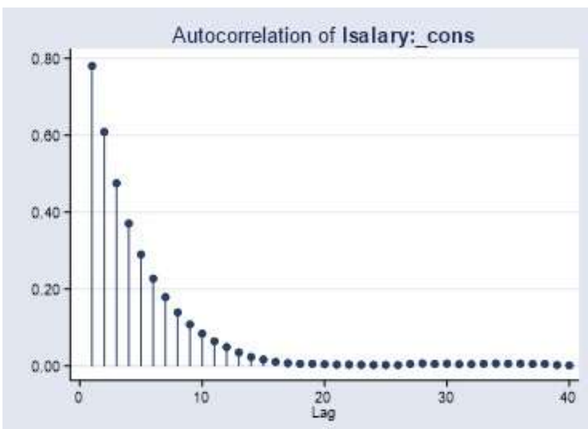

Figure 2c

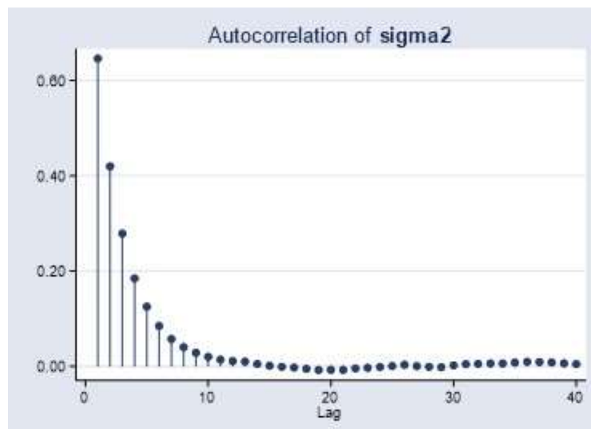

Figure 2d

Figure 3. Histograms from Model 3

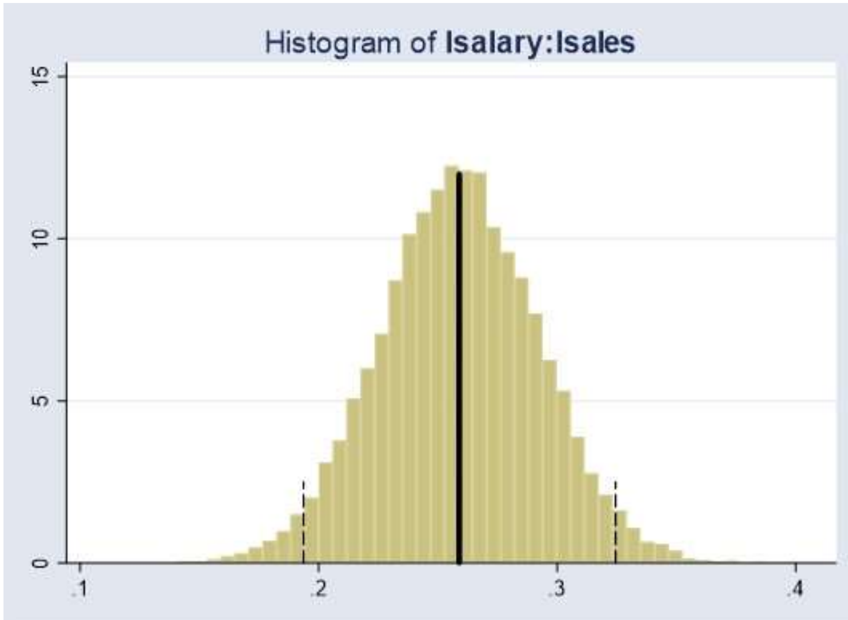

Figure 3a
(dark line indicates mean; dotted lines indicate equal-tailed 95% Cr. Interval)

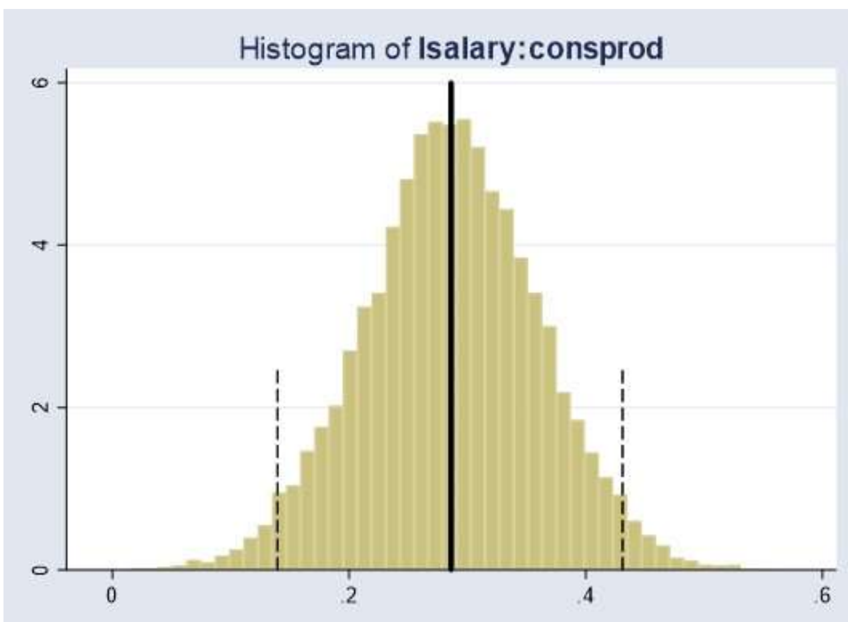

Figure 3b
(dark line indicates mean; dotted lines indicate equal-tailed 95% Cr. Interval)

Appendix A – Key Terms

Bayes Factor: An index quantifying the amount of evidence for a particular hypothesis relative to another; a ratio of the likelihood of two competing models or hypotheses

Burn-In: The preliminary period of an MCMC algorithm during which the chain moves from unrepresentative initial values to the bulk of the posterior; these iterations are discarded from the sample.

Convergence: Refers to whether an MCMC algorithm has reached a point that it is adequately sampling from the bulk of the posterior distribution of interest.

Credibility Interval: Interval estimate that indicates the parameter values that have the most probability given the data and prior distributions (a 95% credible interval is an often reported interval).

Effective Sample Size: A measure of the efficiency of MCMC that indicates the number of independent observations in an MCMC sample; small ESS relative to the MCMC sample size can be a sign of poor mixing.

Gelman-Rubin Statistic: A numerical measure of whether multiple chains converge on the same posterior distribution.

Highest Posterior Density (HPD) Interval: An interval in which the minimum density of every point within the interval is equal to or larger than the density of any point outside the interval; an HPD interval has the shortest width relative to all other credible intervals.

Hyperparameters: Parameters of the prior distribution (e.g., the mean and variance of a normal prior).

Likelihood: The probability of data given parameters in the model.

Markov Chain Monte Carlo (MCMC): A simulation approach for generating samples from probability distributions used to estimate the posterior probability distribution in Bayesian statistics. Algorithms are used to repeatedly sample from the posterior probability distribution.

Mixing: refers to the rate at which a Markov chain explores the parameter space; poor mixing refers to a slow rate to reach convergence.

Monte Carlo Standard Error (MCSE): Standard error of the posterior mean estimate; provides a measure of how much error is in the estimate due to the fact that MCMC is used.

Parameter: An unknown value in a population.

Posterior Distribution: A probability distribution determined by the likelihood of the parameters and their prior distribution.

Prior Distribution: Probability distribution of parameter values formed based on existing knowledge of the parameters prior to observing the data, characterized by hyperparameters of the distribution; priors vary in how informative they are.

Appendix B – Stata Code

```
cscript
version 15.1

use http://fmwww.bc.edu/ec-p/data/wooldridge/ceosal1, clear

/*Describe variables in data set*/
        describe

/*Included in case frequentist estimates are of interest for later comparison*/
        regress lsalary lsales consprod

/*****************************Bayesian Model Estimation****************************************
*Note: set seed # specifies the initial value of the random-number seed used by the random-number functions
*including this command ensures additional runs of the code produce same results
*
*********************************************************************************************/

        set seed 76
        bayes, saving(salreg1,replace): regress lsalary lsales consprod
        estimates store salreg1

        /*Effective sample size information*/
                bayesstats ess

        set seed 76
        bayes, mcmcsize(62500) burnin(12500) saving(salreg2,replace) : regress lsalary lsales consprod
        estimates store salreg2
        /*Effective sample size information*/
                bayesstats ess


        set seed 76
        bayes, mcmcsize(100000) burnin(25000) saving(salreg3,replace) : regress lsalary lsales consprod
        estimates store salreg3
        /*Effective sample size information*/
                bayesstats ess

        *Report results of last model with HPD interval
        bayes, hpd


/***Model Checks***/
        /*Convergence using trace plots - note curly bracket required around name of parameter*/
                estimates restore salreg1
                bayesgraph trace {lsales}

                estimates restore salreg2
                bayesgraph trace {lsales}

                estimates restore salreg3
```

```
                bayesgraph trace {lsales}

        /*Autocorrelation Plots - note that plots for all parameters produced using _all*/
                bayesgraph ac _all

        /*Other diagnostic graph - not included in paper*/
                bayesgraph kdensity _all



/*Multiple Chain Convergence - grubin requires use of bayesmh*/

        set seed 76
        bayesmh lsalary lsales consprod, likelihood(normal({sigma2})) prior({lsalary:}, normal(0,10000))
prior({sigma2}, igamma(0.01,0.01)) initial ({lsalary:} 0.9) block({lsalary:}) saving(salreg_ch1,replace)
        estimates store salreg_ch1

        set seed 76
        bayesmh lsalary lsales consprod, likelihood(normal({sigma2})) prior({lsalary:}, normal(0,10000))
prior({sigma2}, igamma(0.01,0.01)) initial ({lsalary:} 0.01) block({lsalary:}) saving(salreg_ch2,replace)
        estimates store salreg_ch2

        set seed 76
        bayesmh lsalary lsales consprod, likelihood(normal({sigma2})) prior({lsalary:}, normal(0,10000))
prior({sigma2}, igamma(0.01,0.01)) initial ({lsalary:} 2) block({lsalary:}) saving(salreg_ch3,replace)
        estimates store salreg_ch3

        *Gelman–Rubin convergence statistic
        *the following command must be run the first time the code is used: net install grubin,
from("http://www.stata.com/users/nbalov")
                grubin, estnames(salreg_ch1 salreg_ch2 salreg_ch3)


/*Extract values from stored estimates to draw graphs*/
        estimates restore salreg3
        bayes

        local ls_mean=el(e(mean),1,1)
        local ls_lower=el(e(cri),1,1)
        local ls_upper=el(e(cri),2,1)

        local cp_mean=el(e(mean),1,2)
        local cp_lower=el(e(cri),1,2)
        local cp_upper=el(e(cri),2,2)


/*Draw histograms*/
        bayesgraph histogram {lsales}, addplot(function y=`ls_mean', horizontal range(0 12) lcolor(black)
lwidth(thick) || function y=`ls_upper', horizontal range(0 2.5) lcolor(black) lpattern(dash) || function y=`ls_lower',
horizontal range(0 2.5) lcolor(black) lpattern(dash)) legend(off)
        bayesgraph histogram {consprod}, addplot(function y=`cp_mean', horizontal range(0 6) lcolor(black)
lwidth(thick) || function y=`cp_upper', horizontal range(0 2.5) lcolor(black) lpattern(dash) || function
y=`cp_lower', horizontal range(0 2.5) lcolor(black) lpattern(dash)) legend(off)
```

```
/*Post-Estimation Tests*/
*Intervals
	*Confidence that consumer products CEOs make at least 20% more
		bayestest interval {lsalary:consprod}, lower(0.182321557)
	*Confidence that sales effect is less than .2% for a 1% change in sales
		bayestest interval {lsalary:lsales}, upper(.2)


/*Model comparison using Bayes Factors*/
	set seed 76
	bayes, mcmcsize(100000) burnin(25000) saving(salreg4,replace) : regress lsalary lsales consprod ros
	estimates store salreg4

	set seed 76
	bayes, mcmcsize(100000) burnin(25000) saving(salreg5,replace) : regress lsalary lsales consprod roe
	estimates store salreg5

	bayesstats ic salreg4 salreg5


/*Alternative priors*/
	/*Base model*/
		set seed 76
		bayes, mcmcsize(100000) burnin(25000) saving(salreg6,replace) : regress lsalary lsales consprod
		estimates store salreg6
		bayesstats ess

	/*Alternative priors - note need to include block statement*/
		set seed 76
		bayes, mcmcsize(100000) burnin(25000) prior({lsalary:lsales}, normal(0,10))
prior({lsalary:consprod}, normal(0,10)) block({lsalary:}) saving(salreg7,replace) : regress lsalary lsales consprod
		estimates store salreg7
		bayesstats ess

		set seed 76
		bayes, mcmcsize(100000) burnin(25000) prior({lsalary:lsales}, normal(0,.1))
prior({lsalary:consprod}, normal(0,.1)) block({lsalary:}) saving(salreg8,replace) : regress lsalary lsales consprod
		estimates store salreg8
		bayesstats ess
```

Appendix C – Stata-Specific Bayes Resources

- Overview of features:
  - https://www.stata.com/features/bayesian-analysis/

- Stata Bayesian Analysis. Reference Manual:
  - https://www.stata.com/manuals/bayes.pdf

- Bayesian regression models using the bayes prefix:
  - https://www.stata.com/features/overview/bayes-prefix/

- Thompson, J. (2014). Bayesian analysis with Stata. College Station, TX: Stata Press:
  - https://www.stata.com/bookstore/bayesian-analysis-with-stata/
  - (note that this book does not cover the latest version of Stata)

- Stata Bayesian Analysis Videos:
  - https://www.youtube.com/playlist?list=PLN5IskQdgXWktwVOxs3vAVkI4jpMX3pIv
  - Introduction to Bayesian statistics, part 1: The basic concepts
  - Introduction to Bayesian statistics, part 2: MCMC and the Metropolis Hastings algorithm
  - Bayesian linear regression using the bayes prefix
  - Bayesian linear regression using the bayes prefix: How to specify custom priors
  - Bayesian linear regression using the bayes prefix: Checking convergence of the MCMC chain
  - Bayesian linear regression using the bayes prefix: How to customize the MCMC chain
  - (note that the first two videos cover Version 14; the last four cover Version 15)