2009-07-02

Meta-Analysis Using Bayesian Hierarchical Models in Organizational Behavior

Michael David Ulrich
Brigham Young University - Provo

BYU ScholarsArchive Citation

Ulrich, Michael David, "Meta-Analysis Using Bayesian Hierarchical Models in Organizational Behavior" (2009). *All Theses and Dissertations*. 2349.
https://scholarsarchive.byu.edu/etd/2349

META-ANALYSIS USING BAYESIAN HIERARCHICAL MODELS IN ORGANIZATIONAL BEHAVIOR

by

Michael D Ulrich

A project submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Statistics
Brigham Young University
August 2009

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a project submitted by

Michael D Ulrich

This project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	David G. Whiting, Chair
Date	Scott D. Grimshaw
Date	Del T Scott

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the project of Michael D Ulrich in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date	David G. Whiting Chair, Graduate Committee

Accepted for the Department

Scott D. Grimshaw
Graduate Coordinator

Accepted for the College

Thomas W. Sederberg
Associate Dean, College of Physical and Mathematical Sciences

ABSTRACT

META-ANALYSIS USING BAYESIAN HIERARCHICAL MODELS IN ORGANIZATIONAL BEHAVIOR

Michael D Ulrich
Department of Statistics
Master of Science

Meta-analysis is a tool used to combine the results from multiple studies into one comprehensive analysis. First developed in the 1970s, meta-analysis is a major statistical method in academic, medical, business, and industrial research. There are three traditional ways in which a meta-analysis is conducted: fixed or random effects, and using an empirical Bayesian approach. Derivations for conducting meta-analysis on correlations in the industrial psychology and organizational behavior (OB) discipline were reviewed by Hunter and Schmidt (2004). In this approach, Hunter and Schmidt propose an empirical Bayesian analysis where the results from previous studies are used as a prior. This approach is still widely used in OB despite recent advances in Bayesian methodology. This paper presents the results of a hierarchical Bayesian model for conducting meta-analysis of correlations and then compares these results to a traditional Hunter-Schmidt analysis conducted by Judge et al. (2001). In our approach we treat the correlations from previous studies as a likelihood, and present a prior distribution for correlations.

ACKNOWLEDGEMENTS

First and foremost I must thank my wife for putting up with late nights away from home, and for letting me talk to her about this project despite her not knowing anything about the subject (and acting interested nonetheless). I would also like to express thanks for the support I received from my advisor and other faculty members throughout my tenure in the program: David Whiting, Scott Grimshaw, William Christensen, Natalie Blades, Del Scott, Dennis Eggett and John Bingham.

CONTENTS

CHAPTER

APPENDIX

TABLES

Table

FIGURES

Figure

1. INTRODUCTION

As companies attempt to rein in inflated budgets and payrolls, they often look to find ways to improve the productivity of their employees to maintain company-wide performance. Numerous hypothesized factors can contribute to employee performance, such as compensation, integrity, accountability, or having a strong, charismatic leader. One of the oldest theorized facets of job performance is the level of satisfaction that workers find within their jobs. This relationship, described as the "Holy Grail" of industrial psychologists (Landy 1989; Judge et al. 2001), has been examined and described in many academic and industrial studies (Sheridan and Jr. 1975; Petty et al. 1984; Iaffaldanco and Muchinsky 1985; Judge et al. 2001).

With such a profusion of information, attempts have been made to synthesize the satisfaction/performance relationship. The first meta-analytical study on this topic was published by Petty et al. (1984) and was limited to 16 studies from 1964 to 1983. A second meta-analysis conducted by Iaffaldanco and Muchinsky (1985) offered a much broader analytic vantage point than Petty et al. (1984) by incorporating 74 studies. This provided a great boost in understanding of the relationship between performance and satisfaction, yet fell short on some requirements of a rigorous meta-analysis (Judge et al. 2001; Hunter and Schmidt 2004). One such flaw was Iaffaldanco and Muchinsky's measurement of job satisfaction. Instead of measuring overall satisfaction, they elected to use a composite score of various satisfaction components. Despite this study's shortcomings, it stood as the principal study on the relationship between performance and satisfaction for over 15 years.

In 2001, Judge et al. published an exhaustive study which examined over 1,000 prior studies on satisfaction and performance. This study attempted to definitively describe the correlation between job performance and job satisfaction. These estimates were found using procedures presented by Hunter and Schmidt (2004) and are based on a random effects meta-analysis. This technique is discussed in greater depth later in the paper.

The data that we used for this project are found in the appendix of Judge et al. (2001). These data represent 312 unique studies taken from both published and unpublished articles, and were filtered to insure that they addressed the question of correlation between job satisfaction and job performance. The total number of studies gathered by Judge et al. exceeded 1,000, but only 312 of these studies met the criteria presented by Judge et al. (2001). The total number of subjects represented in this meta-analysis is 54,417, with an average sample size per study of 174. The size of this sample emphasizes the benefit of a meta-analysis over conducting original research; gathering a sample of over 50,000 subjects would require a substantial amount of money, time, and other resources.

The Judge data set also provided variables describing the source of each study (whether from a ranked or unranked journal, or from an unpublished study), the measurement used to assess job performance (supervisory ratings, objective records, peer-subordinate ratings, or other) and job satisfaction (global measurement, facet composite measurement, not specified), the research design (cross-sectional or longitudinal), job complexity (low, medium, or high), and sample occupation (scientists or engineers, sales, teachers, managers or supervisors, accountants, clerical workers or secretaries, unskilled and semi-skilled workers, nurses, or miscellaneous and mixed).

In the original model, most of these moderator variables did not appear to substantially alter the correlation between satisfaction and performance. The largest deviates from the global average were found in high complexity jobs and with scientists and engineers, both of which measured above average, and nurses, which measured below average.

Our project is built upon previous work done by Hunter and Schmidt and presents a Bayesian hierarchical model for use in the meta-analysis of correlations.

Using the hierarchical model allows us to better account for parameter uncertainty in our model and improve the validity and ease of interpretation. Our approach results in a fully Bayesian analysis, which permits a Bayesian interpretation of the results.

1.1 Literature Review: Meta-analysis

> We find ourselves in the mildly embarrassing position of knowing less than we have proven. (Glass 1976, p. 8)

Researchers can produce findings and results faster than the information from these studies can be assimilated. In a short period of time, dozens or more studies can be conducted on a single research topic (Glass 1976), with each study demonstrating varying levels of academic rigor and results. The problem of combining studies has been debated and discussed in earnest since the 1970s. The term "meta-analysis" was first coined by Glass (1976) in his seminal work. This paper asked that greater attention be given to researching how to combine quantitative research. Glass contested that "we [as researchers] face an abundance of information. Our problem is to find the knowledge in the information. [. . .] We need more scholarly effort concentrating on the problem of finding the knowledge that lies untapped in completed research studies" (Glass 1976, p. 4). Since the birth of "meta-analysis," hundreds of articles and dozens of books have been penned to explain both its theoretical and applied concepts.

Beginning with their article in the *Journal of Applied Psychology* in 1977, Hunter and Schmidt laid the foundation for meta-analysis in sociological research (with an emphasis on organizations and industrial psychology). Hunter and Schmidt (2004) produced what has become the standard textbook in synthesizing research for the social sciences. While other books with emphasis on biology, medicine, or other hard sciences have also become highly influential, they will not be mentioned in detail here.

Hunter and Schmidt's earliest work concentrated on validity generalization, a method that attempts to generalize findings to a larger population (Rafilson 1991; Schmidt et al. 1985). This technique was used by Hunter and Schmidt to develop a quantitative measure of job performance that could be applied across different types of jobs and industries. This method utilized meta-analysis to combine existing studies in such a way that new studies can be readily integrated into previous research and the results from a meta-analysis can be applied to new situations by making use of "artifact" distributions.

Meta-analysis allows a researcher to combine the results found in other studies into one grand encompassing study. By doing this, the researcher is able to take advantage of a larger composite sample size without actually having to recruit thousands of participants. Take, for example, a study of retail workers for an international organization. In order to take a representative sample, the researcher would need to find a list of thousands of employees, devise a sampling method, write and distribute a survey, etc., all of which costs lots of money. A meta-analysis still involves a great deal of work, but would allow the researcher to focus less energy on gathering a data set and more time on analysis and interpretation. The researcher is now able to go back and draw from previous work to find new results.

Meta-analysis also allows a researcher to study "moderator" variables that are not observable in individual studies. Moderators are used to examine relationships that are only seen when combining studies. These variables are created by coding unique aspects of individual studies. For example, in Judge et al. (2001), each of the 312 studies were coded to include information on the type of journal the studies appeared in, the measurement tools used for both job performance and satisfaction, the job complexity, and the participant's occupation. With this information, Judge et al. (2001) were able to compare the performance/satisfaction correlation in differing scenarios to better understand this relationship.

Meta-analysis is not without its limitations and faults. Since the data used in a meta-analysis are taken from previous work, the results of the analysis are greatly dependent on the quality of others' research. The researcher loses control over sample design and how the original data were collected. There is also a substantial amount of effort required to find all of the relevant research findings. It is not enough to just find a few convenient studies and slap them together to produce a comprehensive body of knowledge. Hunter and Schmidt (2004) and Lipsey and Wilson (2001) both provide a thorough discussion on the steps required to populate the studies to be used in a meta-analysis. The greatest advantage that Judge et al. (2001) had over previous meta–studies came from their comprehensive gathering of previous work. By collecting over 1,000 studies from a multitude of sources, they were able to examine additional moderator variables and other facets of interest.

There are two philosophies behind meta-analytic calculations. These methods serve the same ultimate purpose but are based on very different premises. The most basic approach is called a fixed effects model, which was adopted early in meta-analysis's history due to its simple calculations and freedom from computational requirements (National Research Council 1992; Sutton and Abrams 2001; Hunter and Schmidt 2004). An assumption made by the fixed effects model is that the correlations collected from previous studies are taken from a single population.

Random effects models are based on the premise that each study carries with it a unique true correlation distribution. These distributions are combined to synthesize a Pangaea distribution from which a researcher can make inference (Sutton and Abrams 2001; Field 2005). This approach requires that additional variability terms be included in the model to account for variability between studies. This additional variation increases the realism and breadth of application at the expense of more statistical uncertainty.

A simple example is used to explain the differences between fixed and random effects models. In the fixed effects model, one can assume that each study correlation is sampled from a normal distribution centered at ρ with variance σ^2. In the random effects model, the model takes the form of a two-stage hierarchical model (National Research Council 1992):

$$\begin{aligned} r_i &\sim N(\rho_i, \sigma^2) \\ \rho_i &\sim N(\rho, \tau^2). \end{aligned} \tag{1.1}$$

In this equation, r_i represents a sample correlation taken from the population correlation for that study, ρ_i with variance σ^2. The ρ_i are then normally distributed centered around ρ with variance τ^2. This random effects model is identical to the fixed effects model when $\tau^2 = 0$. Field (2005) also conducted a simulation which showed that the random effects model is highly accurate when study population correlations vary. It is also important to note that in this model, ρ and σ are independent from one another. We will later show using the sampling distribution of r that this is an incorrect assumption for correlations.

As with most statistical techniques, advances in methodology take time to implement. In 1992 the National Research Council published a report detailing the advantages of the random effects model and stated that “the panel generally favors random effects models,” and that “fixed effects models should be used for combining information only in the presence of significant prior information that $\tau^2 = 0$” (National Research Council 1992, pp. 52, 143). Hunter and Schmidt (2004) and the National Research Council (1992) both present tests for homogeneity, checking whether $\tau^2 = 0$. Despite the recommendation from the National Research Council and the philosophical advantages of a random effects model, numerous fixed effects models still appear in the literature (see Hunter and Schmidt, pp. 395–96 for a listing).

There are three proposed formulae used to calculate point and variability estimates for random effects meta-analysis. These techniques are compared and criticized in Johnson et al. (1995) and debated by Schmidt and Hunter (1999). The empirical Bayesian approach found in Hunter and Schmidt (2004) has become a de facto standard in organizational behavior and was the basis for the calculations used by Judge et al. (2001). Deeley and Lindley (1981), Morris (1983), and Sutton and Abrams (2001) show that these empirical methods are a classical maximum likelihood method disguised as a Bayesian analysis. Hunter and Schmidt (2004) justify their use of the Bayesian paradigm by adjusting the correlations to account for statistical artifacts. After simply adjusting for statistical artifacts, Whitener (1990) claims that these correlations now represent the posterior distribution. The reliability estimates for each study used in our analysis were provided by Judge et al. (2001). These estimates are

$$r_{\text{corrected}} = \frac{r_{xy}}{\sqrt{\text{reliability}_x}\sqrt{\text{reliability}_y}} \tag{1.2}$$

and

$$\sigma^2_{\text{e corrected}} = \frac{\sigma^2_e}{\sqrt{\text{reliability}_x}\sqrt{\text{reliability}_y}}. \tag{1.3}$$

Equations (1.2) and (1.3) are used to correct for artifact distributions using reliability estimates for x and y.

In this approach, an estimate of the true observed correlation is found using

$$\bar{\rho} = \frac{\sum[N_i r_i]}{\sum[N_i]}. \tag{1.4}$$

This formula is a weighted average with the sample size of each study used as weights. Similarly, an estimate of the variance between studies is found using

$$\sigma^2_r = \frac{\sum[N_i(r_i - \bar{r})^2]}{\sum[N_i]}, \tag{1.5}$$

which again incorporates the sample sizes of the individual studies as weights.

The formula presented by Hunter and Schmidt for finding the variability attributed to sampling error is

$$\sigma^2_e = \frac{(1-\bar{r}^2)^2}{\bar{N}-1}. \tag{1.6}$$

Finally, the variance of population correlations is

$$\sigma_\rho^2 = \sigma_r^2 - \sigma_e^2. \tag{1.7}$$

2. ANALYSIS

2.1 Development

The examples used in Hunter and Schmidt (2004) did not clearly explain how to correct correlations for attenuation based on reliability coefficients. To better understand how these equations are used, we reproduced Table 1.4 from Judge et al. (2001) using the same methods described in Hunter and Schmidt (2004). The results are shown in Table 2.1. In our findings, we calculated $\bar{\rho} = 0.29$ and σ_ρ^2 to be 0.19, compared to 0.30 and 0.21 as reported in Judge et al. (2001), respectively. The `psychometric` package in `R`, which uses the Hunter and Schmidt formulae for meta-analysis, produced the same results as we found in our own calculations.

These differences do not change the conclusion reached in Judge et al. (2001); however, there is room to question whether these results should be considered significant as assumed by Judge et al. In reference to the 80% credibility interval, Judge et al. wrote the following:

> The 80% credibility interval also excluded zero, indicating that more than 90% of the individual corrected true score correlations are greater than zero. [. . .] Thus, these results indicate that the mean true score correlation between job satisfaction and job performance is moderate in magnitude (.30) and distinguishable from zero. (2001, p. 385)

One question about this conclusion revolves around the methods used to calculate a credibility interval. When Hunter and Schmidt developed their methodology for meta-analysis, Bayesian methods were still in their infancy. The Hunter-Schmidt model utilizes an empirical Bayesian approach, where the correlations from previous studies is used as the Bayes prior distribution. This empirical technique was favored due to the computational limitations of the 1970s, which prohibited researchers from utilizing a Markov chain Monte Carlo (MCMC) or similar algorithm to compute the

k	N	$\bar{r}$	SD_r	$\bar{\rho}$	SD_ρ	80% CV	95% CI	% Var	Q
312	54471	0.18	0.12	0.30	0.21	0.03–0.57	0.27–0.33	25.15	1240.51

Table 2.1: Judge et al. (2001). k=number of correlations; N=total sample size for all studies combined; mean r = average uncorrected correlation; SD_r=standard deviation of uncorrected correlation; mean ρ=average corrected correlation; SD_ρ = standard deviation of corrected (true score) correlation; 80% CV = lower and upper limits of 80% credibility interval; 95% CI = lower and upper limits of 95% confidence interval; % variance = percentage of the variance in the correlations explained by statistical artifacts; Q = statistic used to test for homogeneity in the true correlations across studies.

posterior distribution. Researchers at the time also distrusted what we would consider a true Bayesian analysis because of the subjective nature of selecting Bayesian prior distributions. A comparison between an empirical and a "fully" Bayesian meta-analysis are presented by Sutton and Abrams (2001). Sutton concludes that a "fully" Bayesian meta-analysis is preferable to the empirical analysis presented by the early meta-analysis developers.

The majority of meta-analyses that are currently being conducted in the field of industrial psychology still use an empirical Bayesian approach. Deeley and Lindley (1981) and Morris (1983) show that these empirical methods are really classical maximum likelihood methods disguised as Bayesian. Since these methods are not truly Bayesian, Raudenbush and Bryk (1985) argue that they should not be allowed a Bayesian interpretation. A method for conducting a fully Bayesian analysis in medical research was explained for by Sutton and Abrams (2001) and in general in Gelman et al. (2003). These approaches take advantage of modern MCMC methods, which allow a more robust and sophisticated statistical model as well as a more intuitive interpretation.

The full Bayesian analysis as described by Sutton and Abrams (2001) is a Bayesian hierarchical model. This approach to meta-analysis has found wide use outside of the organizational behavior literature, including in medicine, ecology, and education (Seltzer et al. 1996; Sutton and Abrams 2001; Helser et al. 2007). The Hunter-Schmidt model assumes both a normal likelihood and a normal prior distribution. These assumptions allowed Hunter and Schmidt to use a conjugate normal posterior distribution, which had little computational requirements. These normal distribution assumptions are not required for the hierarchical model. Hierarchical models permit greater flexibility by allowing researchers to use any reasonable likelihood and prior distributions to model correlation.

The Hunter-Schmidt paradigm also incorporates a knowledge of the posterior distribution of ρ using Formulae (1.7), (1.5), and (1.6). In Whitener (1990) we are told,

> In Bayesian statistics, the appropriate standard deviation used to construct credibility intervals is the standard deviation of the posterior distribution. *For meta-analysis, the posterior distribution is the distribution of effect sizes that results after corrections have been made for statistical artifacts.* (p. 316, emphasis added)

These formulae do not necessarily describe a true Bayesian posterior distribution, which is implied by the conversation about Bayesian credibility intervals found in Hunter and Schmidt (2004). Using the hierarchical model, we are able to calculate a true posterior distribution of ρ, which is then used to find a true Bayesian credibility interval.

Using the hierarchical approach, we treat the correlations found in Judge et al. (2001) as our likelihood. This differs from the Hunter-Schmidt approach, which treats these data as being taken from the prior distribution. The likelihood distributions used in our analysis were chosen because they had identical support to a correlation statistic, $(-1, 1)$. There are very few continuous distributions which are bounded on $(-1, 1)$. The most widely known, the Beta distribution, had to be generalized to range from -1 to 1. It is possible to truncate any distribution at -1 and 1, but for the purposes of this study, we chose to only examine the normal distribution. Kendall and

Stuart (1973) show that correlation statistics approximately follow a t-distribution, which is very similar to the normal distribution.

We chose to examine three candidate prior distributions for our analysis. We selected prior distributions using similar rationale to our method in selecting our likelihood distributions. For this reason our first two candidate distributions are a Beta distribution ranging from -1 to 1, and a Truncated Normal distribution with the same range. Finally, we used a Uniform$(-1, 1)$ distribution, a noninformative prior, as a comparison distribution.

The general form of our hierarchical model is

$$\begin{aligned} \pi(\rho) &\propto \prod \pi(\rho_i|r_i)\pi(\rho_i) \\ \rho_i &\sim f(\theta). \end{aligned} \tag{2.1}$$

In this equation, ρ represents the grand population correlation, ρ_i represents the population correlation from each study, r_i represents the sample correlation from each study, and f represents one of the previously mentioned prior distributions.

2.2 Methods

As previously mentioned, we elected two different likelihood distributions: the Truncated Normal (TN) distribution (2.2), and the Generalized Beta (GB) distribution (2.3) (United States Commerce Department 2006), both of which are bounded by -1 and 1. For our prior distributions on ρ_i, we used a Uniform$(-1, 1)$ distribution, in addition to the TN and GB distributions. These combinations gave us six unique candidate distributions for modeling the correlation between job satisfaction and job performance. For comparison purposes, we also modeled a Normal-Normal likelihood-prior distribution.

If $X \sim TN(\rho, \sigma)$ then

$$f(x|\rho, \sigma, a, b) = \frac{\frac{1}{\sigma} \cdot \phi\left(\frac{x-\rho}{\sigma}\right)}{\Phi\left(\frac{b-\rho}{\sigma}\right) - \Phi\left(\frac{a-\rho}{\sigma}\right)}, \qquad a < x < b. \tag{2.2}$$

If $X \sim GB(\alpha, \beta)$ then

$$f(x|\alpha,\beta,a,b) = \frac{(x-a)^{\alpha-1}(b-x)^{\beta-1}}{B(\alpha,\beta)(b-a)^{\alpha+\beta-1}}, \qquad a < x < b. \tag{2.3}$$

2.2.1 Sampling Distribution of r

Once these distributions were defined, we used the sampling distribution of r to determine values for ρ, σ, α, and β. According to Kendall and Stuart (1969), "the ordinates and distribution function of the correlation coefficient are not expressible in terms of simple mathematical functions" (p. 389). Despite this limitation, Kendall and Stuart (1969) state that "r tends to normality" as n increases. Additionally, Kendall and Stuart (1969) give equations for both the mean,

$$E(r_i) = \rho_i \left\{1 - \frac{(1-\rho_i^2)}{2n_i} + O\left(\frac{1}{n_i^2}\right)\right\}, \tag{2.4}$$

and variance,

$$Var(r_i) = \left(\frac{(1-\rho_i^2)^2}{n_i - 1}\right) \cdot \left(1 + \frac{11\rho_i^2}{2n_i}\right) + O\left(\frac{1}{n_i^3}\right), \tag{2.5}$$

of the sample correlation coefficient.

Figure 2.1 shows the effect that the size of the correlation coefficient and sample size have on the sampling distribution of r. From this we see that as the sample size increases, the variance decreases. This relationship gives greater strength to studies with larger sample sizes, which in essence gives studies with a larger sample size a greater weight in the final joint posterior distribution. We also see that as r approaches -1 or 1, the variance also asymptotically approaches 0. This association insures that the sampling distribution will not over reach the bounds of the correlation coefficient (-1, or 1), while also giving greater weight to the studies that differ the most from 0.

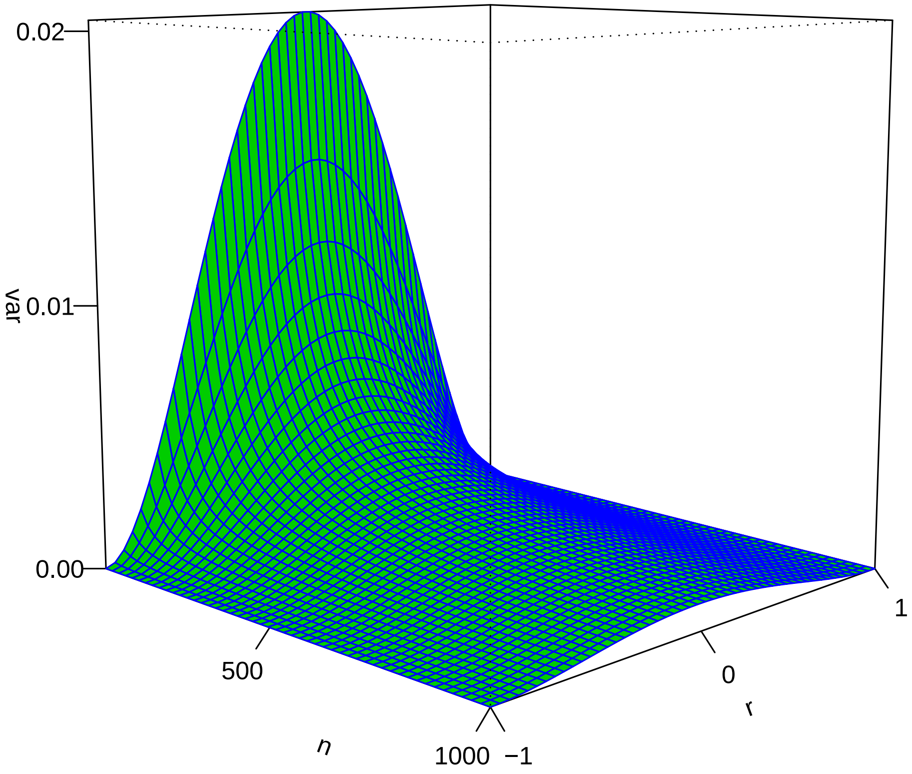

Figure 2.1: Variance of the distribution of r with different n and r

2.2.2 Prior-Likelihood Combinations

In order to implement the MCMC algorithm, we first determined the complete conditional distributions for ρ_i. Using a GB as a likelihood distribution, we find the complete conditionals of ρ_i with a Uniform$(-1, 1)$ prior to be

$$\begin{aligned}\pi(\rho_i|r) &\propto f(r_i|\rho_i) \cdot \pi(\rho_i)\\ &= \prod \left[\frac{(r_i+1)^{\alpha_i-1}(1-r_i)^{\beta_i-1}}{B(\alpha_i,\beta_i)2^{\alpha_i+\beta_i-1}}\right] \cdot \left(\frac{1}{2}\right);\end{aligned}$$

thus

$$[\rho_i] \propto \prod \left[\frac{(r_i+1)^{\alpha_i-1}(1-r_i)^{\beta_i-1}}{B(\alpha_i,\beta_i)2^{\alpha_i+\beta_i-1}}\right], \qquad -1 < r < 1. \tag{2.6}$$

Now using a TN(ρ, σ, -1, 1) prior where $\rho = 0$ and $\sigma = \sqrt{.4}$ we have

$$\begin{aligned}\pi(\rho_i|r) &\propto f(r_i|\rho_i) \cdot \pi(\rho_i)\\ &= \prod \left[\frac{(r_i+1)^{\alpha_i-1}(1-r_i)^{\beta_i-1}}{B(\alpha_i,\beta_i)2^{\alpha_i+\beta_i-1}}\right] \cdot \left[\frac{\frac{1}{\sqrt{.4}} \cdot \phi\left(\frac{\rho_i}{\sqrt{.4}}\right)}{\Phi\left(\frac{1}{\sqrt{.4}}\right) - \Phi\left(\frac{-1}{\sqrt{.4}}\right)}\right];\end{aligned}$$

thus

$$[\rho_i] \propto \prod \left[\frac{(r_i+1)^{\alpha_i-1}(1-r_i)^{\beta_i-1}}{B(\alpha_i,\beta_i)2^{\alpha_i+\beta_i-1}}\right] \cdot \phi\left(\frac{\rho_i}{\sqrt{.4}}\right), \qquad -1 < r < 1. \tag{2.7}$$

Finally, combining the GB likelihood with a GB(α, β, -1, 1) prior where $\alpha = 2$ and $\beta = 2$ we have

$$\begin{aligned}\pi(\rho_i|r) &\propto f(r_i|\rho_i) \cdot \pi(\rho_i)\\ &= \prod \left[\frac{(r_i+1)^{\alpha_i-1}(1-r_i)^{\beta_i-1}}{B(\alpha_i,\beta_i)2^{\alpha_i+\beta_i-1}}\right] \cdot \left[\frac{(\rho_i+1)^{2-1}(1-\rho_i)^{2-1}}{B(2,2) \cdot 2^{2+2-1}}\right];\end{aligned}$$

thus

$$[\rho_i] \propto \prod \left[\frac{(r_i+1)^{\alpha_i-1}(1-r_i)^{\beta_i-1}}{B(\alpha_i,\beta_i)2^{\alpha_i+\beta_i-1}}\right] \cdot \left[(\rho_i+1)(1-\rho_i)\right], \qquad -1 < r < 1. \tag{2.8}$$

The GB distribution is not parametrized by its mean and variance. Instead this distribution is characterized by two shape parameters, α and β. We desired to have the mean of our likelihoods match the mean of the sampling distribution of the correlation coefficient, which meant setting Equations (2.4) and (2.5) equal to the expected value and variance of a GB distribution. To do this, we first found the mean and variance of a GB distribution, with range -1 to 1. The mean of this distribution can be expressed as,

$$E(r) = \frac{\alpha - \beta}{\alpha + \beta} \tag{2.9}$$

and the variance,

$$Var(r) = \frac{4\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{2.10}$$

(see Appendix A for a detailed walk-through).

With these equations, we were able to solve for α_i and β_i by solving the system of equations,

$$\begin{aligned} \rho_i \left(1 - \frac{(1-\rho_i^2)}{2n_i}\right) &= \frac{\alpha_i - \beta_i}{\alpha_i + \beta_i} \\ \frac{(1-\rho_i^2)^2}{n_i - 1}\left(1 + \frac{11\rho_i^2}{2n_i}\right) &= \frac{4\alpha_i\beta_i}{(\alpha_i + \beta_i)^2(\alpha_i + \beta_i + 1)}. \end{aligned} \tag{2.11}$$

Using Maple$^{\text{TM}}$ we found for each α_i and β_i

$$\alpha = -1/8\,\frac{\left(4\,n^3 + 8\,\rho^2 n^2 - 8\,n^2 + 23\,\rho^4 n - 27\,\rho^2 n - \rho^4 + \rho^2\right)\left(\rho^2 - \rho + 2\,n\right)}{\left(-1 + \rho\right) n^2 \left(11\,\rho^2 + 2\,n\right)}, \tag{2.12}$$

and

$$\begin{aligned} \beta = 1/8\,&\left(\frac{-\rho^6 + 23\,\rho^6 n - \rho^5 + 23\,\rho^5 n - 29\,\rho^4 n + 54\,\rho^4 n^2 + \rho^4 + 8\,\rho^3 n^2 + \rho^3}{\left(11\,\rho^3 + 11\,\rho^2 + 2\,\rho n + 2\,n\right) n^2}\right) + \\ 1/8\,&\left(\frac{-27\,\rho^3 n + 20\,\rho^2 n^3 - 62\,\rho^2 n^2 + 2\,\rho^2 n + 4\,\rho n^3 - 8\,\rho n^2 - 16\,n^3 + 8\,n^4}{\left(11\,\rho^3 + 11\,\rho^2 + 2\,\rho n + 2\,n\right) n^2}\right). \end{aligned} \tag{2.13}$$

These estimates for α_i and β_i allowed us to generate a sample for each ρ_i based on the sampling distribution of r_i, and give greater weight to the studies from Judge et al. (2001) which had larger sample sizes. Using these values in an MCMC algorithm, we were able to draw 50,000 observations from each of the 312 studies. These samples could then be combined to create a joint posterior distribution of ρ with an effective sample size of 15,600,000.

Using the TN likelihood, the complete conditional distribution of ρ_i with a $\text{U}(-1,1)$ prior is

$$\begin{aligned}\pi(\rho_i|r_i) &\propto f(r_i|\rho_i)\cdot\pi(\rho_i)\\ &=\prod\left[\frac{\frac{1}{s_i}\cdot\phi\left(\frac{r_i-\rho_i}{s_i}\right)}{\Phi\left(\frac{1-\rho_i}{s_i}\right)-\Phi\left(\frac{1+\rho_i}{s_i}\right)}\right]\cdot\left(\frac{1}{2}\right);\end{aligned}$$

thus

$$[\rho_i]\propto\prod\left[\frac{\frac{1}{s_i}\cdot\phi\left(\frac{r_i-\rho_i}{s_i}\right)}{\Phi\left(\frac{1-\rho_i}{s_i}\right)-\Phi\left(\frac{1+\rho_i}{s_i}\right)}\right],\qquad -1<r<1. \tag{2.14}$$

Now, using a TN(ρ, σ, -1,1) prior where $\rho=0$ and $\sigma=\sqrt{.4}$ we have

$$\begin{aligned}\pi(\rho_i|r_i) &\propto f(r_i|\rho_i)\cdot\pi(\rho_i)\\ &=\prod\left[\frac{\frac{1}{s_i}\cdot\phi\left(\frac{r_i-\rho_i}{s_i}\right)}{\Phi\left(\frac{1-\rho_i}{s_i}\right)-\Phi\left(\frac{1+\rho_i}{s_i}\right)}\right]\cdot\left[\frac{\frac{1}{\sqrt{.4}}\cdot\phi\left(\frac{\rho_i}{\sqrt{.4}}\right)}{\Phi\left(\frac{1}{\sqrt{.4}}\right)-\Phi\left(\frac{-1}{\sqrt{.4}}\right)}\right];\end{aligned}$$

thus

$$[\rho_i]\propto\prod\left[\frac{\frac{1}{s_i}\cdot\phi\left(\frac{r_i-\rho_i}{s_i}\right)}{\Phi\left(\frac{1-\rho_i}{s_i}\right)-\Phi\left(\frac{1+\rho_i}{s_i}\right)}\right]\cdot\phi\left(\frac{\rho_i}{\sqrt{.4}}\right),\qquad -1<r<1. \tag{2.15}$$

Finally, combining the TN likelihood with a GB(α, β, -1,1) prior where $\alpha=2$ and $\beta=2$ we have

$$\begin{aligned}\pi(\rho_i|r_i) &\propto f(r_i|\rho_i)\cdot\pi(\rho_i)\\ &=\prod\left[\frac{\frac{1}{s_i}\cdot\phi\left(\frac{r_i-\rho_i}{s_i}\right)}{\Phi\left(\frac{1-\rho_i}{s_i}\right)-\Phi\left(\frac{1+\rho_i}{s_i}\right)}\right]\cdot\left[\frac{(\rho_i+1)^{2-1}(1-\rho_i)^{2-1}}{B(2,2)(2)^{2+2-1}}\right];\end{aligned}$$

thus

$$[\rho_i] \propto \prod \left[\frac{\frac{1}{s_i} \cdot \phi\left(\frac{r_i - \rho_i}{s_i}\right)}{\Phi\left(\frac{1-\rho_i}{s_i}\right) - \Phi\left(\frac{1+\rho_i}{s_i}\right)} \right] \cdot [(\rho_i + 1)(1 - \rho_i)], \qquad -1 < r < 1. \quad (2.16)$$

For the TN likelihood distributions above, we used (2.4) and (2.5) as estimates of ρ_i and s_i^2 respectively (omitting $O(\cdot)$). We then used these estimates in an MCMC algorithm to generate a sample of 50,000 from each of the 312 studies used in Judge et al. (2001). This allowed us to create a joint posterior distribution from which we could conduct our analysis.

3. RESULTS

With the samples from the joint posterior densities, our task turned to determining which candidate posterior distribution best described the relationship between job satisfaction and job performance.

Before beginning this project, we believed that the fully Bayesian analysis would have a smaller posterior standard deviation than 0.21, the standard deviation reported in Judge et al. (2001). Our thought was that by using a true hierarchical model with proper prior distributions, we could better describe each of the 312 study distributions, which would then reduce the overall posterior variance. In lieu of this, we decided one way to determine the quality of a posterior distribution would be by comparing standard deviations between each candidate distribution, and against the standard deviation reported in Judge et al. (2001).

Table 3.1 lists the standard deviations for each of our six proposed distributions. Not surprisingly, the Uniform prior distribution led to the largest posterior standard deviation. The Uniform distribution included no meaningful prior information. Since this distribution has no tails, it is easy to see that the Uniform prior will have a larger variation than either the TN or the GB distributions.

The TN prior distribution had a lower standard deviation for both the TN and the GB likelihood distributions than the GB prior distribution. The TN likelihood had a smaller standard deviation than the GB likelihood when using a TN prior, but not with the GB prior. Overall, using a TN for both prior and likelihood resulted in the smallest standard deviation.

There are two standard deviations reported in Judge et al. (2001) which we can use for comparison. The first is based solely on the correlation coefficients from each of the 312 studies. The second is based on correlation coefficients corrected for attenuation and statistical artifacts. It is difficult to say which of these measurements

		Prior		
		Uniform	TN	GB
$L(\cdot)$	TN	0.1983	0.1678	0.1930
	GB	0.1914	0.1860	0.1877

Table 3.1: Standard deviations of the joint posterior distributions

is a better comparison for our purposes. The first uses uncorrected correlations to determine the standard deviation of the distribution of r. These uncorrected correlations were also used in our analysis. The second approach attempted to describe the distribution of ρ (albeit incorrectly), which distribution we also attempted to describe. Since our analysis attempts to describe the distribution of ρ while using the uncorrected correlations, it is little surprise that the standard deviations we found fell between the standard deviations reported in Judge et al. (2001). Hence it is our belief that based on the standard deviation criteria, our approach produces a better estimate for distribution spread.

Figure 3.1 graphs each candidate posterior distribution. From this graph we notice a few things. First, these distributions are not smooth and symmetrical. Hunter and Schmidt's model assumes a Normal posterior distribution, which is both smooth and symmetrical. It is our belief that a posterior distribution which combines information from different studies is unlikely to be either smooth or symmetrical. The second thing that we notice is that our distributions are nearly identical to one another. The lone exception to this is the TN prior/likelihood distribution. This tells us that posterior distribution is primarily based on the actual data and not the prior or likelihood distributions.

For the purpose of information, Table 3.2 reports the means from each of our candidate distributions.

The next step in our analysis was calculating intervals for our posterior dis-

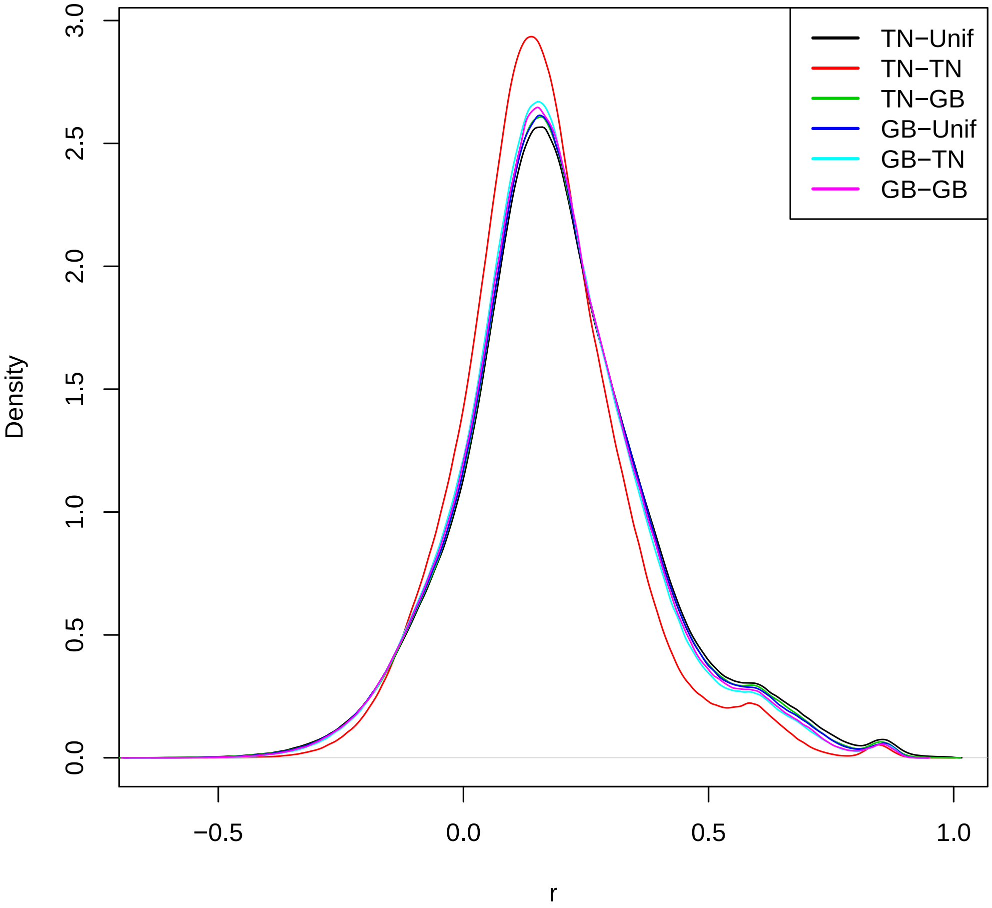

Figure 3.1: Joint posterior densities for each likelihood-prior combination

		Prior		
		Uniform	TN	GB
$L(\cdot)$	TN	0.1934	0.1606	0.1890
	GB	0.1875	0.1813	0.1836

Table 3.2: Joint posterior means for each of the six distributions

		Prior		
		Uniform	TN	GB
$L(\cdot)$	TN	(-0.062,0.417)	(-0.053,0.346)	(-0.060,0.408)
	GB	(-0.058,0.407)	(-0.055,0.401)	(-0.055,0.402)

Table 3.3: Joint posterior 80% credibility intervals for each of the six distributions

tribution. Judge et al. (2001) present what they call a Bayesian credibility interval for ρ which was calculated using Hunter and Schmidt (2004). In our analysis, we calculated credibility intervals using the highest probability density of the posterior distribution. Table 3.3 lists these intervals for each of our candidate distributions. As can be seen, these intervals differ from those found in Judge et al. (2001) (Table 2.1). Whereas Judge et al. (2001) found their interval did not overlap 0 (and therefore they conclude ρ is significantly different from 0), each of our intervals did encompass 0. This tells us that the relationship between job satisfaction and job performance is slightly positive, but not distinguishable from 0.

Our intervals were also narrower than the interval found in Judge et al. (2001). Judge's interval was 0.54 units wide, while our intervals were about 0.45 units wide.

Once we had finished comparing the six distributions we first proposed in our study, we decided to examine a Normal prior, Normal likelihood hierarchical model. This was done to better replicate the model used by Hunter and Schmidt. We used the same MCMC approach as described earlier, which differs from the methods used by Hunter and Schmidt. The Hunter and Schmidt model does not view meta-analysis as a hierarchical model; therefore, their resulting posterior distribution is conjugate to a Normal distribution which can be summarized using formulae from Hunter and Schmidt (2004). This approach does not allow each sample used in the analysis its own distribution.

The resulting posterior distribution found using an MCMC algorithm can be seen in Figure 3.2. The mean of this distribution is 0.1929, with a standard deviation of 0.1715, both of which are similar to the statistics found from our original posterior distributions. We also observe from Figure 3.2 that the general shape of the Normal-Normal posterior distribution is similar to those of our proposed distributions. The Normal-Normal posterior distribution exhibits more jaggedness than our previous posterior distributions. This jaggedness is a result of a smaller variability around ρ_i with the Normal-Normal model than in any of our alternative models.

Figure 3.2 compares the posterior distributions of the Normal-Normal hierarchical model, a GB-GB hierarchical model, and the conjugate Normal model presented by Hunter and Schmidt. As you can see, the Hunter-Schmidt model is quite different from the hierarchical models. Since the Hunter and Schmidt posterior distribution is normally distributed, it does not have the same skewness as any of the hierarchical models. The Hunter-Schmidt model is also symmetric, and centered around 0.30, whereas the hierarchical models are nonsymmetric and centered around 0.19.

From Figure 3.2 we also see the danger in Hunter and Schmidt's claim that we can arrive at the posterior distribution for correlations simply by correcting for attenuation. This correction produced a distribution with a wider range and a symmetric shape compared to the narrow yet skewed distributions produced using a fully Bayesian analysis.

We feel that skewness is an important facet of the relationship between job satisfaction and job performance. Although the correlation between these two variables is weak, there is a positive skewness to their relationship which speaks to the logical existent relationship between satisfaction and performance.

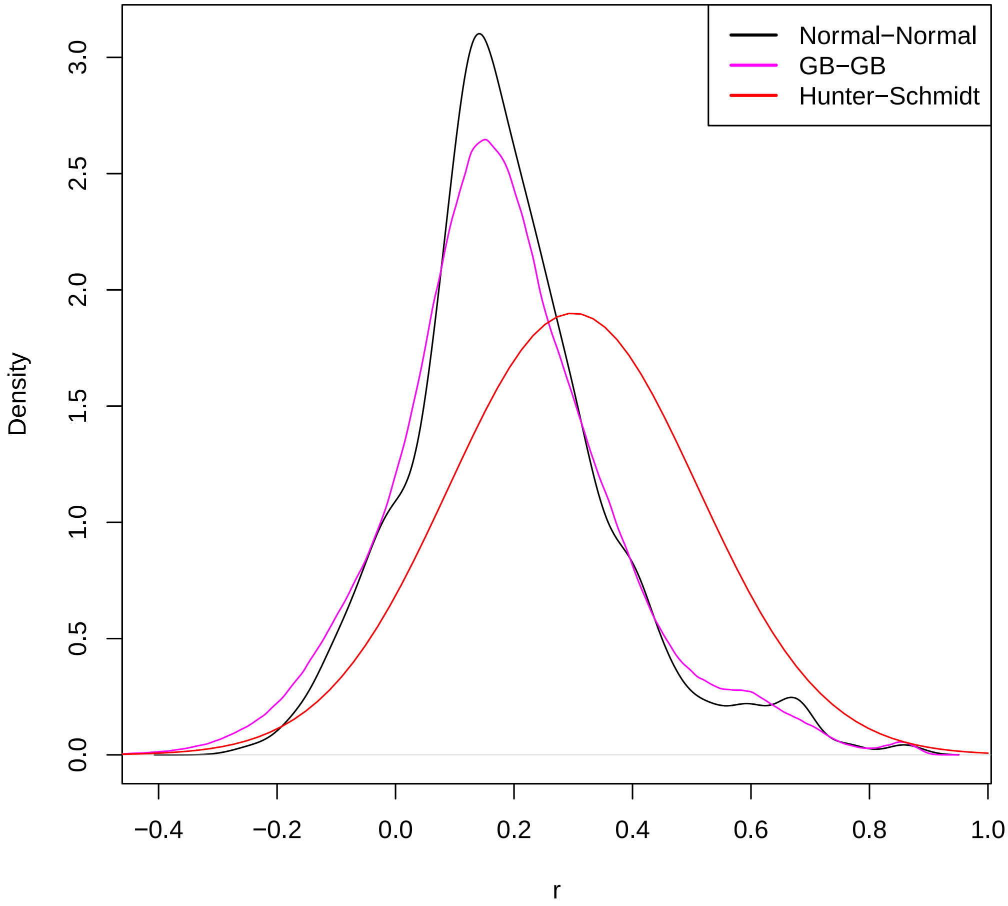

Figure 3.2: Comparing the Normal-Normal, GB-GB, and Hunter-Schmidt posterior distributions

4. CONCLUSION AND DISCUSSION

4.1 Discussion

Judge et al. (2001) produced the largest meta-analysis on the correlation between job satisfaction and job performance to date. Their methods were largely based on the random effects model presented in Hunter and Schmidt (2004). These methods are presented by Hunter and Schmidt as Bayesian despite not utilizing a fully Bayesian paradigm.

We proposed a fully Bayesian hierarchical model to verify the accuracy of Judge's results, and to improve the validity of his interpretations. Although our model did not improve over one of the standard deviations found in Judge et al. (2001), we believe that this hierarchical model is still an improvement over the current empirical Bayesian model employed in meta-analysis in the business field. This fully Bayesian approach more accurately incorporates information from previous studies by treating these studies as a likelihood. A prior distribution for these data can then be formed by eliciting expert opinion or by using noninformative distributions. This is made possible by computers utilizing power statistical algorithms.

The work done by Hunter and Schmidt is still an important tool in conducting meta-analysis, but should not be considered Bayesian. Their techniques would be more accurately described as frequentist, and therefore interpretations of any results should be done in a frequentist framework. The hierarchical Bayesian model allows greater flexibility in model specification and in interpretation.

In our models we were also able to utilize the properties of the sampling distribution of r. These properties inhibit the prior, likelihood, and posterior distributions from reaching beyond the natural limits of the correlation statistic. By correcting for attenuation, the Hunter and Schmidt model enables their statistics to go past the limits of correlations. The variance of the sampling distribution of r also allowed us to more appropriately lend greater weight to studies with larger sample sizes.

In this project we proposed six possible prior/likelihood combinations. Through the MCMC algorithm we were able to show that the resulting posterior distributions were all very similar. In Figure 3.1 we see that five of these distributions are essentially interchangeable.

Using these posterior distributions we were able to calculate credibility intervals to compare to the intervals found in Judge et al. (2001). The intervals we calculated encapsulated 0, which tells us that there is a reasonable probability that the correlation between job satisfaction and job performance is 0. This differs from the result found in Judge et al. (2001), which found this relationship to be significantly positive, albeit very slightly.

This result goes against conventional wisdom which would believe that the more satisfied employees are with their jobs, the more productive they will be. There are likely dozens of reasons for such a relationship which we will not explore at this time. We will speculate that the complexity of the satisfaction/performance relationship does not lend itself to simple studies. Each of these variables are multifaceted and complex.

We have also not been able to justify the use of an 80% credibility interval found in the Hunter-Schmidt approach to meta-analysis. A standard used in interval estimation has been to use at least a 90%, if not a 95% credibility interval. When using a 90% credibility interval as opposed to the 80% credibility interval presented in Judge et al. (2001), the given interval would no longer exclude 0, negating one of the fundamental findings of Judge et al. (2001): that there exists a nonzero correlation between job performance and job satisfaction. A justification for using an 80% interval was neither given in Judge et al. (2001) nor Hunter and Schmidt (2004) and goes against conventional statistical practices.

Since the business audience to which this research is directed may be less comfortable with advanced statistical concepts, we propose that future meta-analytic research use a uniform prior distribution and a TN likelihood distribution. These distributions provide the simplest and most easily understood interpretations. For more advanced researchers we would recommend the GB-GB model. This model will allow a researcher to more intuitively express the weight that the prior distribution will have in the analysis in a way similar to using a Beta prior.

4.2 Future Research

The purpose of this paper was mainly to introduce an approach to conduct a modern Bayesian meta-analysis. Future research should examine the influence that nondiffuse priors have on the posterior distribution. In our model we centered each of our priors at 0 with a large variance. With a well-informed prior distribution we would be able to reduce the posterior variance to produce even stronger results.

We were also limited in scope to a single data set. This prohibited us from examining the behavior of each of these models in more extreme situations where $|r| \geq .8$. It is our belief that the Hunter and Schmidt model will perform very poorly when studying data in these ranges. The models we proposed will properly adjust themselves as they approach -1 and 1 since our models were built on the sampling distribution of r. Any future simulation should also vary the numbers of previous studies used in an analysis.

Research should also be done to examine the validity of correcting correlations for reliability as proposed by Hunter and Schmidt (2004). A simulation could be conducted to examine whether this technique overinflates a correlation statistic or whether this is an appropriate adjustment.

BIBLIOGRAPHY

Copas, J. (1999), "What works? selectivity models and meta-analysis," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 162, 95–109.

Deeley, J. J. and Lindley, D. V. (1981), "Bayes Empirical Bayes," *Journal of the American Statistical Association*, 76, 833–841.

Field, A. P. (2005), "Is the Meta-Analysis of Correlation Coefficients Accurate When Population Correlations Vary?" *Psychological Methods*, 10, 444–467.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003), *Bayesian Data Analysis*, Boca Raton, FL: Chapaman and Hall/CRC, 2nd ed.

Glass, G. V. (1976), "Primary, Secondary, and Meta-Analysis of Research," *Educational Researcher*, 5, 3–8.

Helser, T. E., Stewart, I. J., and Lai, H.-L. (2007), "A Bayesian Hierarchical Meta-analysis of growth for the Genus Sebastes in the Eastern Pacific Ocean," *Canadian Journal of Fisheries and Aquatic Sciences*, 64, 470–485.

Hunter, J. and Schmidt, F. (2004), *Methods of meta-analysis: Correcting error and bias in research findings*, Thousand Oaks, CA: Sage Publications, Inc.

Iaffaldanco, M. T. and Muchinsky, P. M. (1985), "Job Satisfaction and Job Performance: A Meta-analysis," *Psychological Bulletin*, 97, 251–273.

Johnson, B. T., Mullen, B., and Salas, E. (1995), "Comparison of Three Major Meta-Analytic Approaches," *Journal of Applied Psychology*, 80, 94–106.

Judge, T., Thoreson, C., Bono, J., and Patton, G. (2001), "The Joy Satisfaction-Job Performance Relationship: A Qualitative and Quantitive Review," *Psychological Bulletin*, 127, 376–407.

Kendall, M. G. and Stuart, A. (1968), *The Advanced Theory of Statistics*, vol. 3, New York, NY: Hafner Publishing Company, 2nd ed.

— (1969), *The Advanced Theory of Statistics*, vol. 1, New York, NY: Hafner Publishing Company, 3rd ed.

— (1973), *The Advanced Theory of Statistics*, vol. 2, New York, NY: Hafner Publishing Company, 3rd ed.

Landy, F. L. (1989), *Psychology of Work Behavior*, Pacific Grove, CA: Thomson Brooks/Cole, 4th ed.

Lipsey, M. W. and Wilson, D. B. (2001), *Practical Meta-analysis*, vol. 49 of *Applied Social Research Methods Series*, Thousand Oaks, CA: Sage Publications, Inc.

Morris, C. N. (1983), "Parametric Empirical Bayes Inference: Theory and Applications," *Journal of the American Statistical Association*, 78, 47–65.

National Research Council (1992), *Combining information statistical issues and opportunities for research*, Washington D.C.: National Academy of Science Press.

Novick, M. R. and Jackson, P. H. (1974), *Statistical Methods for Educational and Psychological Research*, New York: McGraw-Hill Inc.

Petty, M. M., McGee, G. W., and Cavender, J. W. (1984), "A Meta-analysis of the Relationships Between Individual Job Satisfaction and Individual Performance," *Academy of Management Review*, 9, 712–721.

Rafilson, F. (1991), "The Case for Validity Generalization," *Practical Assessment, Research & Evaluation*, 2.

Raudenbush, S. W. and Bryk, A. S. (1985), "Empirical Bayes Meta-analysis," *Journal of Educational Statistics*, 10, 75–98.

Rosenthal, R. and DiMatteo, M. R. (2001), “Meta-Analysis: Recent Developments in Quantitative Methods for Literature Reviews,” *Annual Review of Psychology*, 52, 59–82.

Schmidt, F. L. and Hunter, J. E. (1977), “Development of a General Solution to the Problem of Validity Generalization,” *Journal of Applied Psychology*, 62, 529–540.

— (1999), “Comparison of Three Major Meta-Analytic Approaches Revisited: An Analysis of Johnson, Mullen, and Salas (1995),” *Journal of Applied Psychology*, 84, 144–148.

Schmidt, F. L., Hunter, J. E., Pearlman, K., and Hirsh, H. R. (1985), “Forty Questions About Validity Generalization and Meta-analysis,” *Personnel Psychology*, 38, 697–798.

Seltzer, M. H., Wong, W. H., and Bryk, A. S. (1996), “Bayesian Analysis in Applications of Hierarchical Models: Issues and Methods,” *Journal of Educational and Behavior Statistics*, 21, 131–167.

Sheridan, J. E. and Jr., J. W. S. (1975), “The Direction of the Causal Relationship Between Job Satisfaction and Work Performance,” *Organizational Behavior and Human Performance*, 14, 159–172.

Sutton, A. J. and Abrams, K. R. (2001), “Bayesian Methods in Meta-analysis and Evidence Synthesis,” *Statistical Methods in Medical Research*, 10, 277–303.

United States Commerce Department (2006), “Engineering Statistics Handbook,” Available at *http://www.manticmoo.com/articles/jeff/programming/latex/bibtex-types.php*.

Whitener, E. M. (1990), “Confusion of Confidence Intervals and Credibility Intervals,” *Journal of Applied Psychology*, 75, 315–321.

A. DERIVATIONS OF THE EXPECTED VALUE AND VARIANCE OF A $\beta(\alpha, \beta, -1, 1)$ DISTRIBUTION

$$\begin{aligned} X &\sim GB(\alpha, \beta, -1, 1), (-1 \leq X \leq 1) \\ f(x) &= \frac{(x+1)^{\alpha-1}(1-x)^{\beta-1}}{2^{\alpha+\beta-1}B(\alpha, \beta)} \end{aligned} \tag{A.1}$$

$$\begin{aligned} E(X) &= \int_{-1}^{1} x \cdot \frac{(x+1)^{\alpha-1}(1-x)^{\beta-1}}{2^{\alpha+\beta-1}B(\alpha, \beta)} dx \\ E(X) &= E(X+1) - 1 \end{aligned}$$

$$\begin{aligned} E(X+1) &= \int_{-1}^{1} (x+1) \cdot \frac{(x+1)^{\alpha-1}(1-x)^{\beta-1}}{2^{\alpha+\beta-1}B(\alpha, \beta)} dx \\ &= \int_{-1}^{1} \frac{(x+1)^{\alpha-1+1}(1-x)^{\beta-1}}{2^{\alpha+\beta-1}B(\alpha, \beta)} dx \\ &= \frac{2^{\alpha+1+\beta-1}B(\alpha+1, \beta)}{2^{\alpha+\beta-1}B(\alpha, \beta)} \times \int_{-1}^{1} \frac{(x+1)^{\alpha-1+1}(1-x)^{\beta-1}}{2^{\alpha+1+\beta-1}B(\alpha+1, \beta)} dx \\ &= 2 \cdot \left[\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)}\right] \cdot \left[\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right]^{-1} \\ &= 2 \cdot \left[\frac{\alpha\Gamma(\alpha)}{(\alpha+\beta)\Gamma(\alpha+\beta)}\right] \cdot \left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+\beta)}\right]^{-1} \\ &= \frac{2\alpha}{\alpha+\beta} \end{aligned}$$

$$\begin{aligned} E(X) &= E(X+1) - 1 \\ &= \frac{2\alpha}{\alpha+\beta} - \frac{\alpha+\beta}{\alpha+\beta} \\ \therefore E(X) &= \frac{\alpha-\beta}{\alpha+\beta} \end{aligned} \tag{A.2}$$

$$\mathrm{var}(X+1) = E[(X+1)^2] - [E(X+1)]^2$$

$$\begin{aligned}
E[(X+1)^2] &= \int_{-1}^{1} (x+1)^2 \cdot \frac{(x+1)^{\alpha-1}(1-x)^{\beta-1}}{2^{\alpha+\beta-1}B(\alpha,\beta)} \\
&= \int_{-1}^{1} \frac{(x+1)^{\alpha-1+2}(1-x)^{\beta-1}}{2^{\alpha+\beta-1}B(\alpha,\beta)} \\
&= \frac{2^{\alpha+2+\beta-1}B(\alpha+2,\beta)}{2^{\alpha+\beta-1}B(\alpha,\beta)} \times \int_{-1}^{1} \frac{(x+1)^{\alpha-1+2}(1-x)^{\beta-1}}{2^{\alpha+2+\beta-1}B(\alpha+2,\beta)} \\
&= 4 \cdot \left[\frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+2+\beta)}\right] \cdot \left[\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right]^{-1} \\
&= 4 \cdot \left[\frac{\alpha(\alpha+1)\Gamma(\alpha)}{(\alpha+\beta)(\alpha+\beta+1)\Gamma(\alpha+\beta)}\right] \cdot \left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+\beta)}\right]^{-1} \\
&= \frac{4\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)}
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(X+1) &= \frac{4\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} - \left(\frac{\alpha-\beta}{\alpha+\beta}\right)^2 \\
&= \frac{4\alpha(\alpha+1)(\alpha+\beta) - 4\alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{4\alpha^2(\alpha+\beta) + 4\alpha(\alpha+\beta) - 4\alpha^2(\alpha+\beta) - 4\alpha^2}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{4\alpha(\alpha+\beta) - 4\alpha^2}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{4\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(X) &= \mathrm{var}(X+1) \\
\therefore \mathrm{var}(X) &= \frac{4\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \qquad \text{(A.3)}
\end{aligned}$$

B. R CODE

See also the attached CD

B.1 Truncated Normal Likelihood

```
setwd("/Users/gradstudent/Documents/Masters Project/")


library(msm) #This library holds truncated normal distribution
functions library(MCMCpack) #Various MCMCM functions library(MASS)
library(mvtnorm) #multivariate normal library(xtable) library(sm)
library(boa) #boa.hpd() calculates credible intervals library(rgl) #3d
plots


###Read in data data<-read.csv(’data2.csv’) #names(data)


#k is the number of studies in our study k<-nrow(data)


#I just prefer little n for sample sizes data$n<-data$N N<-sum(data$n)
#Total sample size


#a and b are the bounds for correlations which will be used throughout
the program a<- -1 b<-1


#Formula for st.dev. of correlations found in The Advanced Theory of
Statistics s2<-(1-data$r^2)^2/(data$n-1)*(1+11*data$r^2/(2*data$n))
mu<-data$r*(1-(1-data$r^2)/(2*data$n))
```

```
########TN likelihood, Uniform prior(-1,1)

#Posterior distribution of TN likelihood and U(-1,1)
gTNUnif<-function(x,mu=0,s2=1/3,k=1){
log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))
}

t<-Sys.time() ###MCMC set up burn<-1000 length<-50000

rhoTNUnif<-matrix(0,ncol=k,nrow=(length+burn))

cand.sig<-.1

t<-Sys.time()

for(i in 2:(length+burn)){

#acc<-0 #Calculate the acceptance rate

#Update rhoUnif (i = 1,...,k), k=312 studies for(j in 1:k){
rhoTNUnif[i,j]<-rhoTNUnif[i-1,j] #This takes care of the else
condition for comparing the candidate value with the previous value

#Generate a candidate value from a normal distribution
```

```
cand<-rnorm(1,rhoTNUnif[i-1,j],cand.sig)

if(cand<b&cand>a){ #Make sure that we’re within the
proper bounds for correlations

accept<-gTNUnif(mu[j],cand,s2=s2[j])-
gTNUnif(mu[j],rhoTNUnif[i-1,j],s2[j])
#Find the difference between the candidate and the previous value

u<-log(runif(1))
if(u<accept){rhoTNUnif[i,j]<-cand}#;acc<-acc+1} #Test to accept or
reject the candidate value } } #if(i%%500==0){print(i)} #Keep track of
how many iterations have been completed } t<-Sys.time()-t t
#acc/(length+burn) #Calculate the acceptance rate

# ##Check mixing plots #plot(rhoTNUnif[-c(1:burn),23],type=’l’) #
##Joint posterior density plot #dev.new()
#lines(density(rhoTNUnif[-c(1:burn),]),col=12)
#abline(v=c(boa.hpd(rhoTNUnif[-c(1:burn),],.05))) #95% Credible
interval # ##Descriptive statistics #mean(rhoTNUnif[-c(1:burn),])
#median(rhoTNUnif[-c(1:burn),]) #sd(c(rhoTNUnif[-c(1:burn),])) # #


########TN likelihood, TN prior(0,sqrt(.2))

#Posterior distribution of TN likelihood and TN prior(0,sqrt(.4))
gTNTN<-function(x,mu=0,s2=1/3,k=1,sig=.2){
```

```
log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))+log(dnorm(mu/sig))
}


###MCMC set up burn<-1000 length<-50000

rhoTNTN<-matrix(0,ncol=k,nrow=(length+burn))

cand.sig<-.1

t<-Sys.time() for(i in 2:(length+burn)){

#acc<-0 #Calculate the acceptance rate

#Update rhoTN (i = 1,...,k), k=312 studies for(j in 1:k){
rhoTNTN[i,j]<-rhoTNTN[i-1,j] #This takes care of the else condition
for comparing the candidate value with the previous value

#Generate a candidate value from a normal distribution
cand<-rnorm(1,rhoTNTN[i-1,j],cand.sig)

if(cand<b&cand>a){ #Make sure that we're within the
proper bounds for correlations

accept<-gTNTN(mu[j],cand,s2=s2[j])-gTNTN(mu[j],
rhoTNTN[i-1,j],s2[j])
```

```
#Find the difference between the candidate and the previous value

u<-log(runif(1))
if(u<accept){rhoTNTN[i,j]<-cand}#;acc<-acc+1} #Test to accept or
reject the candidate value } } #if(i%%1000==0){print(i)} #Keep track
of how many iterations have been completed } t<-Sys.time()-t t
#acc/(length+burn) #Calculate the acceptance rate

# ##Check mixing plots #plot(rhoTNTN[-c(1:burn),23],type=’l’) #
##Joint posterior density plot
#plot(density(rhoTNTN[-c(1:burn),]),col=2)
#abline(v=c(boa.hpd(rhoTNTN[-c(1:burn),],.05))) #95% Credible interval
#boa.hpd(rhoTNTN[-c(1:burn),],.2) ##Descriptive statistics
#mean(rhoTNTN[-c(1:burn),]) #median(rhoTNTN[-c(1:burn),])
#sd(c(rhoTNTN[-c(1:burn),])) # #
```

```
########TN likelihood, Beta prior(2,2,-1,1)

#Posterior distribution of TN likelihood and Beta prior(2,2,-1,1)
gTNBeta<-function(x,mu=0,s2=1/3,k=1,sig=.3,alpha=2,beta=2){
#log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))+log(dnorm(mu/sig))
log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))+(alpha-1)*log(mu-a)+(beta-1)*log(b-mu)
}
```

```
###MCMC set up burn<-1000 length<-50000

rhoTNBeta<-matrix(0,ncol=k,nrow=(length+burn))

cand.sig<-.1

t<-Sys.time() for(i in 2:(length+burn)){

#acc<-0 #Calculate the acceptance rate

#Update rho (i = 1,...,k), k=312 studies for(j in 1:k){
rhoTNBeta[i,j]<-rhoTNBeta[i-1,j] #This takes care of the else
condition for comparing the candidate value with the previous value

#Generate a candidate value from a normal distribution
cand<-rnorm(1, rhoTNBeta[i-1,j],cand.sig)

if(cand<b&cand>a){ #Make sure that we’re within the
proper bounds for correlations

accept<-gTNBeta(mu[j],cand,s2=s2[j])-
gTNBeta(mu[j],
rhoTNBeta[i-1,j],s2[j]) #Find the difference between the candidate and
the previous value
```

```
u<-log(runif(1))
if(u<accept){rhoTNBeta[i,j]<-cand}#;acc<-acc+1} #Test to accept or
reject the candidate value } } #if(i%%1000==0){print(i)} #Keep track
of how many iterations have been completed } t<-Sys.time()-t t

# ##acc/(length+burn) #Calculate the acceptance rate # ##Check mixing
plots #plot(rhoTNBeta[-c(1:burn),23],type='l') # ##Joint posterior
density plot #lines(density(rhoTNBeta[-c(1:burn),]),col=13)
#abline(v=c(boa.hpd(rhoTNBeta[-c(1:burn),],.05))) #95% Credible
interval #boa.hpd(rhoTNBeta[-c(1:burn),],.2) ##Descriptive statistics
#mean(rhoTNBeta[-c(1:burn),]) #median(rhoTNBeta[-c(1:burn),])
#sd(c(rhoTNBeta[-c(1:burn),])) #
```

B.2 Generalized Beta Likelihood

```
setwd("/Users/gradstudent/Documents/Masters Project/")

library(msm) #This library holds truncated normal distribution
functions library(MCMCpack) #Various MCMCM functions library(MASS)
library(mvtnorm) #multivariate normal library(xtable) library(sm)
library(boa) #boa.hpd() calculates credible intervals library(rgl) #3d
plots

###Read in data data<-read.csv('data2.csv') #names(data)

#k is the number of studies in our study k<-nrow(data)

#I just prefer little n for sample sizes data$n<-data$N
```

```
#a and b are the bounds for correlations which will be used throughout
the program a<- -1 b<-1

#Formula for st.dev. of correlations found in The Advanced Theory of
Statistics s2<-(1-data$r^2)^2/(data$n-1)*(1+11*data$r^2/(2*data$n))
mu<-data$r*(1-(1-data$r^2)/(2*data$n))



p<-data$r n<-data$n alphaR<-
-(1/8)*(4*n^3+8*p^2*n^2-8*n^2+23*p^4*n-27*p^2*n-p^4+p^2)*(p^2-p+2*n)/
((-1+p)*n^2*(11*p^2+2*n))

betaR<-
(1/8)*(-p^6+23*p^6*n-p^5+23*p^5*n-29*p^4*n+54*p^4*n^2+p^4+8*p^3*n^2-
27*p^3*n+p^3+20*p^2*n^3-62*p^2*n^2+2*p^2*n+4*p*n^3-8*p*n^2-16*n^3+8*n^4)/
((11*p^3+11*p^2+2*p*n+2*n)*n^2)



########Beta likelihood, Uniform prior(-1,1)

#Posterior distribution of TN likelihood and U(-1,1)
gBetaUnif<-function(x,p,n=1){ #
log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))
aR<-
-(1/8)*(4*n^3+8*p^2*n^2-8*n^2+23*p^4*n-27*p^2*n-p^4+p^2)*(p^2-p+2*n)/
```

```
((-1+p)*n^2*(11*p^2+2*n))

bR<-
(1/8)*(-p^6+23*p^6*n-p^5+23*p^5*n-29*p^4*n+54*p^4*n^2+p^4+8*p^3*n^2-
27*p^3*n+p^3+20*p^2*n^3-62*p^2*n^2+2*p^2*n+4*p*n^3-8*p*n^2-16*n^3+8
*n^4)/((11*p^3+11*p^2+2*p*n+2*n)*n^2)

(aR-1)*log(x+1)+(bR-1)*log(1-x)+lgamma(aR+bR)-lgamma(aR)-
lgamma(bR)-(aR+bR-1)*log(2)
}

t<-Sys.time() ###MCMC set up burn<-1000 length<-50000

rhoBetaUnif<-matrix(0,ncol=k,nrow=(length+burn))

cand.sig<-.1

t<-Sys.time()

for(i in 2:(length+burn)){

#acc<-0 #Calculate the acceptance rate

#Update rhoUnif (i = 1,...,k), k=312 studies for(j in 1:k){
rhoBetaUnif[i,j]<-rhoBetaUnif[i-1,j] #This takes care of the else
condition for comparing the candidate value with the previous value
```

```
#Generate a candidate value from a normal distribution
cand<-rnorm(1,rhoBetaUnif[i-1,j],cand.sig)

if(cand<b&cand>a){ #Make sure that we're within the
proper bounds for correlations

accept<-gBetaUnif(data$r[j],cand,n=data$n[j])-
gBetaUnif(data$r[j],rhoBetaUnif[i-1,j],n=data$n[j])
#Find the difference between the candidate and the previous value

u<-log(runif(1))
if(u<accept){rhoBetaUnif[i,j]<-cand}#;acc<-acc+1} #Test to accept or
reject the candidate value } } #if(i%%500==0){print(i)} #Keep track of
how many iterations have been completed } t<-Sys.time()-t t
#acc/(length+burn) #Calculate the acceptance rate

# ##Check mixing plots #plot(rhoBetaUnif[-c(1:burn),23],type='l') #
##Joint posterior density plot #dev.new()
#lines(density(rhoBetaUnif[-c(1:burn),]),col=4)
#abline(v=c(boa.hpd(rhoBetaUnif[-c(1:burn),],.05))) #95% Credible
interval # ##Descriptive statistics #mean(rhoBetaUnif[-c(1:burn),])
#median(rhoBetaUnif[-c(1:burn),]) #sd(c(rhoBetaUnif[-c(1:burn),])) # #
#cbind(sort(data$r),data$n[order(data$r)],s2[order(data$r)],
(1-data$r^2)^2)
#
```

```
########Beta likelihood, TN prior(0,sqrt(.2))

#Posterior distribution of TN likelihood and TN prior(0,sqrt(.4))
gBetaTN<-function(x,p=0,n,sig=.5){
#log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))+log(dnorm(mu/sig))
#p<-data$r[197] #n<-data$n[197] aR<-
-(1/8)*(4*n^3+8*p^2*n^2-8*n^2+23*p^4*n-27*p^2*n-p^4+p^2)*(p^2-p+2*n)/
((-1+p)*n^2*(11*p^2+2*n))

bR<-
(1/8)*(-p^6+23*p^6*n-p^5+23*p^5*n-29*p^4*n+54*p^4*n^2+p^4+8*p^3*n^2-
27*p^3*n+p^3+20*p^2*n^3-62*p^2*n^2+2*p^2*n+4*p*n^3-8*p*n^2-16*n^3+
8*n^4)/((11*p^3+11*p^2+2*p*n+2*n)*n^2)

(aR-1)*log(x+1)+(bR-1)*log(1-x)+lgamma(aR+bR)-lgamma(aR)-
lgamma(bR)-(aR+bR-1)*log(2)+log(dnorm(p/sig))

}

###MCMC set up burn<-1000 length<-50000

rhoBetaTN<-matrix(0,ncol=k,nrow=(length+burn))

cand.sig<-.2
```

```
t<-Sys.time() for(i in 2:(length+burn)){ #acc<-0 #Calculate the
acceptance rate

#Update rhoTN (i = 1,...,k), k=312 studies for(j in 1:k){
rhoBetaTN[i,j]<-rhoBetaTN[i-1,j] #This takes care of the else
condition for comparing the candidate value with the previous value

#Generate a candidate value from a normal distribution
cand<-rnorm(1, rhoBetaTN[i-1,j],cand.sig)

if(cand<b&cand>a){ #Make sure that we’re within the
proper bounds for correlations

accept<-gBetaTN(data$r[j],cand,n=data$n[j])-
gBetaTN(data$r[j],rhoBetaTN[i-1,j],n=data$n[j])
#Find the difference between the candidate and the previous value

u<-log(runif(1))
if(u<accept){rhoBetaTN[i,j]<-cand}#;acc<-acc+1} #Test to accept or
reject the candidate value } } #if(i%%400==0){print(i)} #Keep track of
how many iterations have been completed } t<-Sys.time()-t t
#acc/(length+burn) #Calculate the acceptance rate

# #par(mfrow=c(1,1)) ##Check mixing plots
#plot(rhoBetaTN[c(2200:2300),213],type=’l’)
#plot(rhoBetaTN[,301],type=’l’) # # #apply(rhoBetaTN,2,mean) ##Joint
```

```
posterior density plot #lines(density(rhoBetaTN[-c(1:burn),]),col=3)
#abline(v=c(boa.hpd(rhoBetaTN[-c(1:burn),],.05))) #95% Credible
interval #boa.hpd(rhoBetaTN[-c(1:burn),],.2) ##Descriptive statistics
#mean(rhoBetaTN[-c(1:burn),]) #median(rhoBetaTN[-c(1:burn),])
#sd(c(rhoBetaTN[-c(1:burn),])) # #head(rhoBetaTN)
#plot(rhoBetaTN[,303],type='l') # #



########Beta likelihood, Beta prior(2,2,-1,1)


#Posterior distribution of TN likelihood and Beta prior(2,2,-1,1)
gBetaBeta<-function(x,p=0,n=1,alpha=2,beta=2){
#log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))+log(dnorm(mu/sig))
#log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/
sqrt(s2)))+(alpha-1)*log(mu-a)+(beta-1)*log(b-mu)
aR<-
-(1/8)*(4*n^3+8*p^2*n^2-8*n^2+23*p^4*n-27*p^2*n-p^4+p^2)*(p^2-p+2*n)/
((-1+p)*n^2*(11*p^2+2*n))
bR<-
(1/8)*(-p^6+23*p^6*n-p^5+23*p^5*n-29*p^4*n+54*p^4*n^2+p^4+8*p^3*n^2-
27*p^3*n+p^3+20*p^2*n^3-62*p^2*n^2+2*p^2*n+4*p*n^3-8*p*n^2-16*n^3+
8*n^4)/((11*p^3+11*p^2+2*p*n+2*n)*n^2)

(aR-1)*log(x+1)+(bR-1)*log(1-x)+lgamma(aR+bR)-lgamma(aR)-
lgamma(bR)-(aR+bR-1)*log(2)+(alpha-1)*log(p-a)+(beta-1)*log(b-p)
```

```
}

###MCMC set up burn<-1000 length<-50000

rhoBetaBeta<-matrix(0,ncol=k,nrow=(length+burn))

cand.sig<-.3

t<-Sys.time() for(i in 2:(length+burn)){

#acc<-0 #Calculate the acceptance rate

#Update rho (i = 1,...,k), k=312 studies for(j in 1:k){
rhoBetaBeta[i,j]<-rhoBetaBeta[i-1,j] #This takes care of the else
condition for comparing the candidate value with the previous value

#Generate a candidate value from a normal distribution
cand<-rnorm(1, rhoBetaBeta[i-1,j],cand.sig)

if(cand<b&cand>a){ #Make sure that we're within the
proper bounds for correlations

accept<-gBetaBeta(data$r[j],cand,n=data$n[j])-
gBetaBeta(data$r[j],rhoBetaBeta[i-1,j],n=data$n[j])
#Find the difference between the candidate and the previous value
```

```
u<-log(runif(1))
if(u<accept){rhoBetaBeta[i,j]<-cand}#;acc<-acc+1} #Test to accept or
reject the candidate value } } #if(i%%300==0){print(i)} #Keep track of
how many iterations have been completed } t<-Sys.time()-t t

# ##acc/(length+burn) #Calculate the acceptance rate # ##Check mixing
plots #plot(rhoBetaBeta[c(1000:1100),23],type=’l’) # ##Joint posterior
density plot #lines(density(rhoBetaBeta[-c(1:burn),]),col=1)
#abline(v=c(boa.hpd(rhoBetaBeta[-c(1:burn),],.05))) #95% Credible
interval #boa.hpd(rhoBetaBeta[-c(1:burn),],.2) ##Descriptive
statistics #mean(rhoBetaBeta[-c(1:burn),])
#median(rhoBetaBeta[-c(1:burn),]) #sd(c(rhoBetaBeta[-c(1:burn),])) # #
```

B.3 Normal Likelihood

```
setwd(’/Users/gradstudent/Documents/Masters Project/Code/’)
library(msm) #This library holds truncated normal distribution functions
library(MCMCpack)
library(MASS)
library(mvtnorm)
library(rgl)
library(xtable)
library(sm)
library(boa)
data<-(read.table("faculty.dat",col.names=c(’rating’)))$rating

########Beta likelihood, Beta prior(2,2,-1,1)
```

```
gNN<-function(x,p=0,sd,sd.p=.3){
#log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-
pnorm((a-mu)/sqrt(s2)))+log(dnorm(mu/sig))
#log(dnorm((x-mu)/sqrt(s2)))-log(pnorm((b-mu)/sqrt(s2))-
pnorm((a-mu)/sqrt(s2)))+(alpha-1)*log(mu-a)+(beta-1)*log(b-mu)

dnorm(x,mean=p,sd=sd,log=T)+dnorm(p,mean=0,sd=sd.p,log=T)
}

###MCMC set up
burn<-1000
length<-50000

rhoNN2<-matrix(0,ncol=k,nrow=(length+burn))

cand.sig<-.03

t<-Sys.time()
for(i in 2:(length+burn)){

#acc<-0 #Calculate the acceptance rate

#Update rho (i = 1,...,k), k=312 studies
for(j in 1:k){
rhoNN[i,j]<-rhoNN[i-1,j] #This takes care of the else
#condition for comparing the candidate value with the previous value
```

```
#Generate a candidate value from a normal distribution
cand<-rnorm(1, rhoNN[i-1,j],cand.sig)

if(cand<b&cand>a){  #Make sure that we're within the
proper
bounds for correlations

accept<-gNN(mu[j],cand,s2[j])-gNN(mu[j],
rhoNN[i-1,j],s2[j])
#Find the difference between the candidate and the previous value

u<-log(runif(1))
if(u<accept){rhoNN[i,j]<-cand}#;acc<-acc+1}
#Test to accept or reject the candidate value
}
}
#if(i%%300==0){print(i)} #Keep track of how many iterations have been
completed
}
t<-Sys.time()-t
t

plot(rhoNN[-c(1:burn),23],type='l')

dev.set()
lines(density(rhoNN[-c(1:burn),]))
```

```
for(i in 1:k)
rhoNN2[,i]<-rnorm((length+burn),mean=(mu[i]*.3^2)/(.3^2+s2[i])
,sd=(.3^2*s2[i])/(.3^2+s2[i]))
```

B.4 Summary Statistics

```
##Plot joint posteriors for each distribution
setwd('/Users/gradstudent/Documents/Masters Project/statthesis/')

#Plot Joint Density Distributions
pdf('jointDensities.pdf')
plot(density(rhoTNTN[-c(1:burn),]),col=2,lwd=1,main=" ",xlab="r")
lines(density(rhoTNUnif[-c(1:burn),]),col=1,lwd=1)
lines(density(rhoTNBeta[-c(1:burn),]),col=3,lwd=1)
lines(density(rhoBetaUnif[-c(1:burn),]),col=4,lwd=1)
lines(density(rhoBetaTN[-c(1:burn),]),col=5,lwd=1)
lines(density(rhoBetaBeta[-c(1:burn),]),col=6,lwd=1)
legend('topright',legend=c('TN-Unif','TN-TN','TN-GB','GB-Unif','GB-TN',
'GB-GB'),col=1:6,lty=1,lwd=2)
dev.off()


##Plot joint posterior for N-N distribution
pdf('normalnormal.pdf')
plot(density(rhoNN[-c(1:burn),],bw=.03),col=1,lwd=1,main="",xlab="r")
dev.off()
```

```
plot(n)
for(i in 1:15)
abline(v=i,col=i,lwd=2)

lines(xxx<-seq(-1,1,length=100),dnorm(xxx,mean=.3,sd=.21),type='l')

dev.set()
lines(density(rhoNN[-c(1:burn),]))

for(i in 1:k)
rhoNN2[,i]<-rnorm((length+burn),mean=(mu[i]*.3^2)/(.3^2+s2[i]),
sd=(.3^2*s2[i])/(.3^2+s2[i]))



plot(density(rhoNN2[-c(1:burn),]))
lines(density(rhoNN[-c(1:burn),]),col=2)
lines(density(rhoTNTN[-c(1:burn),]),col=3)
ls()


#Calculate the means for each posterior
meanTNUnif<-mean(rhoTNUnif[-c(1:burn),])
meanTNTN<-mean(rhoTNTN[-c(1:burn),])
meanTNBeta<-mean(rhoTNBeta[-c(1:burn),])
meanBetaUnif<-mean(rhoBetaUnif[-c(1:burn),])
```

```
meanBetaTN<-mean(rhoBetaTN[-c(1:burn),])
meanBetaBeta<-mean(rhoBetaBeta[-c(1:burn),])
meanNN<-mean(rhoNN[-c(1:burn),])
meanNN2<-mean(rhoNN2[-c(1:burn),])

cbind(meanTNUnif,meanTNTN,meanTNBeta,meanBetaUnif,meanBetaTN,meanBetaBeta,
meanNN,meanNN2)

#Calculate the SD for each posterior
sdTNUnif<-sd(c(rhoTNUnif[-c(1:burn),]))
sdTNTN<-sd(c(rhoTNTN[-c(1:burn),]))
sdTNBeta<-sd(c(rhoTNBeta[-c(1:burn),]))
sdBetaUnif<-sd(c(rhoBetaUnif[-c(1:burn),]))
sdBetaTN<-sd(c(rhoBetaTN[-c(1:burn),]))
sdBetaBeta<-sd(c(rhoBetaBeta[-c(1:burn),]))
sdNN<-sd(c(rhoNN[-c(1:burn),]))
sdNN2<-sd(c(rhoNN2[-c(1:burn),]))

cbind(sdTNUnif,sdTNTN,sdTNBeta,sdBetaUnif,sdBetaTN,sdBetaBeta,sdNN,sdNN2)

#Calculate alph level Credible intervals for each posterior
alph<-.2
hpdTNUnif<-boa.hpd(rhoTNUnif[-c(1:burn),],alpha=alph)
hpdTNTN<-boa.hpd(rhoTNTN[-c(1:burn),],alpha=alph)
hpdTNBeta<-boa.hpd(rhoTNBeta[-c(1:burn),],alpha=alph)
hpdBetaUnif<-boa.hpd(rhoBetaUnif[-c(1:burn),],alpha=alph)
hpdBetaTN<-boa.hpd(rhoBetaTN[-c(1:burn),],alpha=alph)
```

```
hpdBetaBeta<-boa.hpd(rhoBetaBeta[-c(1:burn),],alpha=alph)
hpdNN<-boa.hpd(rhoNN[-c(1:burn),],alpha=alph)
hpdNN2<-boa.hpd(rhoNN2[-c(1:burn),],alpha=alph)

credInt80<-cbind(hpdTNUnif,hpdTNTN,hpdTNBeta,hpdBetaUnif,hpdBetaTN,
hpdBetaBeta,hpdNN,hpdNN2)
credInt95<-cbind(hpdTNUnif,hpdTNTN,hpdTNBeta,hpdBetaUnif,hpdBetaTN,
hpdBetaBeta,hpdNN,hpdNN2)

#Varv and mean for the sampling distribution of r
(1-meanTNTN^2)^2/(N-1)*(1+11*meanTNTN^2/(2*N))
meanTNTN*(1-(1-meanTNTN^2)/(2*N))
```