

**BAYESIAN INFERENCE OF PREDICTORS RELATIVE
IMPORTANCE IN LINEAR REGRESSION MODEL
USING DOMINANCE HIERARCHIES**

Xiaoyin Wang^{1 §}, Philippe Duverger², Harvir S. Bansal³

¹Department of Mathematics

Towson University

8000 York Road, Towson, MD 21252, USA

²Department of Marketing

Towson University

Towson, USA

³The Conrad Business, Entrepreneurship and Technology Centre

University of Waterloo

Waterloo, CANADA

Abstract: Regression analysis is perhaps the most frequently used statistical tool for the analysis of data in practice. The purpose of regression analysis is to predict or explain response variables from several pre-selected predictive variables. In the model selection stage, the researcher identifies a subset of predictive variables from a full available set. The selected set is chosen to provide the most adequate description of the response variable. Relative importance analysis is a very useful supplement to regression analysis (Tonidandel and LeBreton 2011). The purpose of determining predictor importance is not model selection but rather uncovering the individual contributions of predictors relative to each other within a selected model. The purpose of this article is to extend the current research practice by developing a statistical modeling approach using the Bayesian framework to evaluate the relative importance of each predictor in a multiple regression model. In what follows we will first illustrate

Received: February 24, 2013

© 2013 Academic Publications, Ltd.
url: www.acadpubl.eu

[§]Correspondence author

the Dominance Analysis procedure, and will then use our critic as a starting point to introduce the Bayesian Dominance Hierarchies approach based on a statistical model of paired comparisons. Simulation studies are conducted to compare with outcomes from the Dominance Analysis procedure, and followed by an empirical example in service research. Finally, the discussion will lead to reviewing potential future research avenues.

1. Introduction

Regression analysis is perhaps the most frequently used statistical tool to characterize the nature and the degree of relationships between one response variable Y and multiple predictors (X_1, X_2, \dots, X_p) . The purpose of regression analysis is to predict or explain response values from several well-selected predictors (Courville and Thompson, 2001). When the regression analysis is used for prediction purposes, people are interested in identifying a set of "good" predictors that provide the most adequate description of the response variable. When the regression analysis is used for explanatory purposes, people are interested in the extent to which each variable contributes to the response variable. Therefore, it makes sense to distinguish between two separate stages: model selection and predictor comparison (see Azen, Budescu, and Reiser, 2001; Budescu, 1993).

While model selection is very well-known and studied in numerous statistical research, there are not many studies on the relative importance analysis, which is a very useful supplement to regression analysis (Tonidandel and LeBreton 2011). The purpose of determining predictor importance is not model selection but rather uncovering the individual contributions of the predictors relative to each other within a selected model. When one is mainly concerned with how much scores on the response variable would change based on a unit increase in a predictor while holding the other predictors constant, regression coefficients based on the selected model are well suited to address such a question. However, the research interests in predictor variables often extends beyond such simple questions to better understand the impact of a particular predictor relative to others in the model. The goal of relative importance analysis is to partition explained variance among multiple predictors so as to better understand the role played by each predictor in a regression equation. Here, the question of interest becomes understanding the extent to which each variable drives the prediction. Essentially, one wishes to understand the contribution each predictor makes towards explaining variance in the response variable. When predictors are uncorrelated, commonly used indices zero-order correlation and standardized

regression coefficient are equivalent. The relative importance of each variable is expressed as the proportion of variability in the response variable that may be attributed to the predictor. When predictors are correlated, past research has documented these indices as inadequate (Budescu, 1993, Darlington, 1968; Hoffman, 1960); and commonly used methods that deal with multicollinearity have often forced the researcher to eliminate or select variables, even if sometimes one would want to keep all variables in. In response, two alternative approaches, dominance analysis (Budescu 1993) and relative weight analysis (Fabbris 1980; Johnson 2000), have been developed that allow for more accurate variance partitioning among correlated predictors while keeping all variables in the model (Tonidandel and LeBreton 2011). However, critics have disapproved of relative importance analysis on the grounds that such analyses are atheoretic.

The purpose of this article is to extend the current research practice by developing a statistically model-based approach in the Bayesian framework to evaluate the relative importance of each predictor in a linear regression model. In what follows, we will first illustrate the state-of-the-art measures of relative importance, and will then use our critic as a starting point to introduce a Bayesian approach based on a statistical model of paired comparisons. Simulation studies are conducted to compare with the outcomes of the Dominance Analysis, followed by an empirical example in a service research study. Finally, the discussion will lead to reviewing potential future research avenues.

2. Dominance Analysis

Determining the relative importance of predictors in linear regressions has been a long debated issue, with literature being developed in a variety of fields such as, statistics, psychology, economics, organizational research, and medicine (Johnson and LeBreton, 2004). Over the years, researchers have suggested various methods for assessing relative importance of predictors; for a recent comprehensive review see Johnson and LeBreton (2004) and Retzer et al. (2009). However, the meaning of importance was ambiguous with many possible definitions (Darlington 1968, Kruskal and Majors 1989, John and LeBreton 2004, Retzer, Soofi and Soyer 2009), and arguments have been confounded because there has been a lack of agreement on the definition and meaning of relative importance of predictors in multiple regressions. Budescu (1993) proposed the Dominance Analysis method stating that relative importance should be defined "in terms of a reduction in error in predicting the criterion"; and "importance should be decomposed as a direct and indirect effect". Research comparing in-

dices from a theoretical perspective (Johnson and LeBreton 2004), Monte Carlo simulations (LeBreton, Ployhart and Ladd 2004), and real data sets (LeBreton, Ployhart and Ladd 2004), suggest that the Dominance Analysis approaches are the preferred method for assessing relative importance of predictors in linear regression analysis.

The concept of predictor "dominance" proposed by Budescu (1993) is a new way to compare predictors in a multiple regression context. Dominance Analysis (Budescu 1993) approaches the problem of relative importance by examining the change in R^2 resulting from adding a predictor to all possible subset regression models. One predictor is said to completely dominate another if its additional contribution in R^2 is greater than that of the other predictor under each of the subset models; if the average additional contribution in R^2 within each model size is greater for one predictor than the other, then that predictor is said to conditionally dominate the other. By averaging the conditional dominance weights, one obtains a predictor's general dominance weight. The three levels of dominance are related in a hierarchical fashion: complete dominance implies conditional dominance which, in turn, implies general dominance. However, for $p > 3$, the converse does not necessarily hold; that is, general dominance does not necessarily imply conditional dominance, and conditional dominance does not necessarily imply complete dominance. The general dominance weight can be interpreted as the partial effect of the variable (LeBreton, Ployhart Ladd 2004), because it addresses a variable's contribution in combination with the other predictors, while overcoming the problems associated with correlated predictors. It represents the proportion of variance of the variable explained by all other variables in the model (Azen and Budescu 2003), and has been shown to be very similar to other measures of relative importance (Johnson 2000).

However, the general dominance weight, because it is based on the averaging of conditional dominance weights, introduces certain degrees of bias. Note that the additional contribution of R^2 of a predictor is more substantial for small models; as the number of predictors increases, the additional contribution of any extra predictors will get lesser. Using an equal weighted average of conditional dominance weights across different model sizes, in fact, puts more credence on the small model. Hence, the order of the dominance tends to be influenced more by the conditional dominance weights of the models with small number of predictors. We will further illustrate this issue in the empirical example with real life data.

Also all dominance weights are generated based on observable manifest variables, and may differ from the true population values. The reproducibility process used to derive the probability of pair-wised dominance without any

particular model assumptions are based on a non-parametric bootstrapping technique. Lahiri (2003) provides a critical review of various modifications of Efrons original bootstrap to handle complex issues in survey sampling. For example, it has been shown that Efrons basic bootstrap procedure fails when the observed sample points are not independent (Lahiri 2003).

Hence, we propose to circumvent the current Dominance Analysis limitations by approaching the problem in a Bayesian framework to discover the relative importance of predictors in the true population of a linear regression model in a probabilistic fashion.

3. Bayesian Dominance Hierarchies Models

In this section, we apply the method of paired comparisons within the Bayesian framework to reveal the statistical inference of the relative importance/dominance of various predictors in a linear regression model. Once again, we call attention that this is not a model selection approach rather an extension. When a model is selected, we would like to see which predictor is more important than the others.

Statistical methods of determining relative dominance abilities based on paired comparisons have long history dated back to Bradley and Terry (1952) in the context of chess tournaments, and have been broadly applied in many areas. Algorithms used to produce maximum likelihood estimates of dominance abilities under the Bradley-Terry model fail to converge to finite values, and so cannot be used for many data sets with zero counts. Davidson and Solomon (1973) and Leonard (1977) described Bayesian versions of the method of paired comparisons. The major benefit of the Bayesian approach is that prior information can be incorporated into the analysis so that the resulting estimates of competitive abilities are always finite (Leonard 1977). Although recent research (Retzer, Soofi and Soyer 2009) has used a reduction of uncertainty in the predictability power of the variables in a model to conceptualize importance, we chose to keep and extend the original concept of dominance as introduced by Budescu (1993) which lends itself to more intuitive results, straightforward interpretations, and practical application (Azen and Budescu 2003). In doing so, we apply the statistical method of paired comparisons based on counts of "wins and losses" (David 1988).

Suppose there are p predictors in a linear regression model, X_1, \dots, X_p . The predictor X_i is said to dominate or "win" X_j if adding X_i to a model results in a greater increase in R^2 , than would be obtained by adding X_j to the model with

the same subsets of other variables. In a regression model with p predictors, X_i and X_j are compared amongst 2^{p-2} models that contain the same subsets of other predictors. Let W_{ij} be the number of times predictor X_i dominating X_j , and θ_{ij} be the probability of predictor X_i dominating X_j . Then W_{ij} has a binomial distribution with parameters $(2^{p-2}, \theta_{ij})$. The Bradley-Terry model is based on the assumption that $\theta_{ij} = \xi_i / (\xi_i + \xi_j)$, where ξ_i represents the relative dominance ability of predictor X_i . Let $d_i = \ln(\xi_i)$, then d_i can be interpreted as the relative dominance ability of predictor X_i on the logarithm scale. Here, "dominance ability" is understood as a measure of degree of intrinsic association between predictors and response variables with presence of other predictors. Based upon this reparameterization, we have $\text{logit}(\theta_{ij}) = d_i - d_j$, or $\theta_{ij} = \exp(d_i - d_j) / (1 + \exp(d_i - d_j))$, which implies that the dominance probability depends only on the difference of dominance indices. Therefore, the Bradley-Terry model assumes the probability that one predictor prevails over another is a logistic function (the link function) of the difference in dominance indices between these two predictors.

A more general approach is to assume that $\theta_{ij} = H(d_i - d_j)$, where H is a link function, which maps differences in dominance indices, ranging from negative infinity to positive infinity, to a probability that must lie between 0 and 1, inclusively. Two most commonly used link functions are the logit link, as assumed by the Bradley-Terry model, and the probit link, which assumes $\theta_{ij} = \Phi(d_i - d_j)$ with Φ being the cumulative function of the standard normal distribution. Both link functions have the following characteristics: (1) the bigger the difference, the larger the value of θ_{ij} ; that is, the more likely one predictor will dominate the other; (2) when two dominance indices are equal, θ_{ij} equals 0.5; that is each predictor has a 50% chance of prevailing when dominance indices of two predictors are the same. Note that $\theta_{ji} = 1 - \theta_{ij}$, and $w_{ji} = 2^{p-2} - w_{ij}$. The likelihood function of the results from paired comparison of the predictors can be written as

$$\prod_{i < j} \binom{2^{p-2}}{w_{ij}} \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{2^{p-2} - w_{ij}} \propto \prod_{i < j} \theta_{ij}^{w_{ij}} \theta_{ji}^{w_{ji}} \\ = \prod_{i=j} \theta_{ij}^{w_{ij}} = \prod_{i=j} H(d_i - d_j)^{w_{ij}}. \quad (1)$$

The prior distribution is a key part of Bayesian inference. It represents the information about an uncertain parameter, and is combined with the probability distribution of observed data to yield the posterior distribution, which in turn is used for future inferences and decisions. In practice, often there is no or very

limited prior information about the dominance ability of predictors. A normal prior distribution corresponds to the prior belief that modest differences in dominance abilities are nearly equally likely, and that very large differences are somewhat less probable. It is also reasonable to assume that there is the same symmetrical prior probability distribution for each predictors dominance index. Since the dominance probability θ_{ij} depends only on differences between dominance indices, d_i and d_j , we assume, without loss of generality, that $d_i \sim N(0, \sigma^2)$.

An Inverse-Gamma distribution is then applied to the variance of the normal distribution as the second stage of a hierarchical prior; that is $\sigma^2 \sim IG(a, b)$, where, a and b are hyper-parameters, and are being chosen so that the distribution of σ^2 will be very much diffused. Hierarchical (multilevel) prior models are central to modern Bayesian statistics for both conceptual and practical reasons. On the theoretical side, hierarchical models allow a more objective approach to inference by estimating the parameters of prior distributions from data, rather than requiring them to be specified using subjective information (see James and Stein, 1960, Efron and Morris, 1975, and Morris, 1983). At a practical level, hierarchical models are flexible tools for combining information and partial pooling of inferences (see, for example, Kreft and De Leeuw, 1998, Snijders and Bosker, 1999, Carlin and Louis, 2001, Raudenbush and Bryk, 2002, Gelman et al., 2003). With the advent of Markov Chain Monte Carlo (MCMC), it has become possible to do calculations on these much more complex models, thus, rendering the hierarchical Bayesian approach much more practical.

A Bayesian posterior density is a true probability density which can be used to make direct probability statements about a parameter. Writing down the posterior probability explicitly, we see that

$$f(d_1, \dots, d_p | data) \propto \prod_{i=j} H(d_i - d_j)^{w_{ij}} \prod_{i=1}^p \phi(d_i | \sigma^2) \eta(\sigma^2 | a, b),$$

where $\phi(\cdot | \sigma^2)$ is the probability density function of $N(0, \sigma^2)$, and $\eta(\cdot | a, b)$ is the probability density function of $IG(a, b)$. When the posterior density is too complicated, we can estimate posterior quantities by generating a sample from the posterior density and computing sample statistics. WinBUGS, a statistical software for Bayesian analysis using MCMC methods, is used to implement the dominance hierarchy analysis. After discarding a burn-in of 5,000 steps, the remaining 10,000 samples were considered as drawn from the posterior distribution.

Bayesian dominance method provides prosperous information about the relative importance of predictors via posterior distributions of dominance indices

d_i s, which are made available by the MCMC computation. First of all, the posterior mean of the dominance index d_i can be used as an overall measure of relative importance of the predictor X_i . Secondly, the posterior distribution dominance probability θ_{ij} can be retrieved by applying the link function to the MCMC chain, and the posterior mean of θ_{ij} can be applied as an estimate of the dominance probability to reveal the conclusion of pairwise comparison among predictors. More over, the lower and upper 2.5th percentile of the posterior distributions of d_i and θ_{ij} can be used to construct 95% confidence intervals of d_i and θ_{ij} , respectively. Last but not least, the posterior probability of any particular order of dominance can be estimated by the proportion of steps in the MCMC chain in this order.

4. Simulation Studies

To illustrate the Bayesian approach in linear regression, we conduct simulation studies in multiple linear regression settings with four predictors. In this case, a data set of 1,000 observations are generated based on the correlation matrix shown in Table 1, which was also used by Azen and Budescu (2003). The correlation table is consistent with possible strong collinearity between X_1 and X_2 , which is evidenced by the pairwise correlation coefficient of 0.80. A paired comparison of predictors based on the additional increase in R^2 was conducted among $2^{4-2} = 4$ possible subset regression models. Numbers of times that the variable X_i prevails over X_j among all possible encounters for all possible pairs of i and j are shown in Table 2. The predictor X_1 "wins" other predictors under all possible subset models. This is consistent with Azen and Budescu (2003) finds, where they classified that X_1 completely dominated all other predictors. Predictor X_2 "ties" with X_3 ; and predictor X_3 "ties" with X_4 . As is shown in Table 2, each of them "wins" half of the time, 2 out of 4. When it comes to the pair of X_2 and X_4 , X_4 "wins" 3 out of total of 4 times. It implies that the order rank of X_4 is a little high than X_2 and X_3 . However, it is not easy to determine the dominance order of X_2 and X_3 based on Table 2 because X_4 ties with X_3 , which in turn ties with X_2 .

Bayesian paired comparison methods with a vague inverse gamma prior, $a = b = 0.001$, are conducted to further study dominance hierarchies among predictors under two different link functions: logit and probit. The posterior means, and 95% confidence intervals of dominance indices d_i s as listed in Table 3. It shows that X_1 has the largest dominance ability with the largest values of dominance index under both link functions. The order rank of dominance

abilities of the remaining variables from the largest to the smallest is d_4 , d_3 , and d_2 . The predictor X_4 has larger dominance ability than X_2 and X_3 as is implied in Table 2. The predictor X_3 is more "important" than X_2 in that X_1 and X_2 are highly related. The order of Bayes dominance indexes are consistent with the order of general dominance indexes in Azen and Budescu (2003). When using the lower and upper 2.5th percentiles of the MCMC chain, 95% confidence intervals are accomplished effortlessly, as shown in Table 3. Note that 95% confidence intervals of dominance indices are overlapped with each other. It is challenging to determine which predictor is definitely dominance the other predictors.

By simply applying link functions to dominance indices, one can benefit from the Bayesian method to achieve inferences about posterior properties of dominance probability, θ_{ij} 's. The summary posterior statistics of θ_{ij} , with $i < j$, are listed in Table 4. For $i > j$, $\theta_{ij} = 1 - \theta_{ji}$. It shows that X_1 is very likely to dominate the other predictors with large probabilities of over 0.95 under logit and over 0.80 under probit link function, respectively. The 95% confidence intervals of the dominance probabilities of X_1 versus other predictors all have lower limits greater than 0.60, which also implies a large chance of dominance. The posterior estimates of the other dominance probabilities are all around 0.5, with a little less than 0.5 under logit and a little greater than 0.5 under probit link function. Also, under both link functions, all 95% confidence intervals run across 0.5 upto 0.73, which implies a moderate dominance ability.

The posterior probability of a particular order of the dominance ranking is estimated by the proportion that the order was encountered in the MCMC steps. Table 5 listed six orders of ranking of predictors with posterior probability greater than 5%. The results show that the posterior probability of dominance hierarchies are consistent within the two link functions. With ordering rank of X_1, X_4, X_3, X_2 having the largest value of posterior probability, once again it runs down to the conclusion that this dominance order is most likely true.

5. Empirical Examples

In this section, we will apply the Bayesian approach to a real life data set. Data drawn from a larger study of auto-repair service customers provide the empirical context to demonstrate the utility of Bayesian analysis of relative importance. Here, switching intention (Y) is modelled as a function of service quality (X_1), satisfaction (X_2), trust (X_3), switching costs (X_4), alternative attractiveness (X_5) and customer commitment (X_6). A sample of 356 completed surveys was

obtained.

Construct measurement used primarily existing scales, adapted for the service context examined. Service quality was measured with the three-item scale used by Taylor and Baker (1994), satisfaction with Oliver and Swans (1989) scale. Alternative attractiveness and switching cost scales were based on Ping (1993). Commitment was assessed with Meyer and Allens (1997) three-component scale of organizational commitment, adapted for the service contexts of the current study. Trust was measured by adapting Larzelere and Hustons (1980) interorganizational scale, modified by Morgan and Hunt (1994). Finally, switching intentions were measured using Oliver and Swans (1989) scale of behavioral intention. All items were made up of seven-point scales.

Using the Dominance Analysis, results show that trust (X_3) completely/conditional dominates other variables; alternative attractiveness (X_5) completely/conditional dominates the rest of variables expect customer commitment (X_6); service quality (X_1) completely/conditional dominates satisfaction (X_2); and customer commitment (X_6) completely/conditional dominates switching cost (X_4). The order of dominance based on general dominance weigh is $X_3 > X_5 > X_1 > X_2 > X_6 > X_4$. Table 6 shows the additional contribution of R^2 for each variable under different subset model sizes. When there is not other variable in the model ($k = 0$), adding any of the variables will result in a quite significant increasing in R^2 , which ranges from 0.0939 to 0.5392; while when there have been many variables already in the model ($k = 5$) adding one last variable does not contribute much change in R^2 , which ranges from 0 to 0.1085. The general dominance weight, which is the average of the conditional dominance weights, is biased towards the order of dominance when there are fewer variables in the model.

Paired comparison of predictors based on the additional increase in R^2 was conducted among $2^{(6-2)} = 16$ possible subset regression models, and the number of times that the variable X_i prevails over X_j are shown in Table 7. We can see clearly that trust (X_3) wins against all other variables, and that alternative attractiveness (X_5) also wins against all variables except (X_3). The customer commitment (X_6) "wins" most of the times against service quality (X_1), satisfaction (X_2), and switching cost (X_4); while the service quality (X_1) "wins" most of the times against satisfaction (X_2) and switching cost (X_4); and switching cost (X_4) "wins" all the times against satisfaction (X_2). The observed order of dominance is $X_3 > X_5 > X_6 > X_1 > X_4 > X_2$.

The Bayesian analysis of relative importance of these variables is conducted based on three vague prior distributions with $a = b = 0.001, 0.01$, and 0.5, respectively. Bayesian estimates of relative importance indices under different

link functions and prior distributions are shown in Table 8. The mean values of importance indices are little smaller under logistic link function than under probit link function; and increase a little as the value of hyper parameters increase. However, the relative order of mean values of importance indices remain the same under two link functions, and are the same as what we observed from Table 7. The summary posterior statistics (means and 95% confidence intervals) of dominance probabilities of θ_{ij} with $i < j$ are listed in Table 9 for logistic link function and Table 10 for probit link function. For $i > j$, $\theta_{ij} = 1 - \theta_{ji}$. In spite of discrepancies among mean values of relative importance indices, the posterior statistics of dominance probabilities remain very robust under different link functions and prior distributions. It shows that X_3 is most likely to dominate all other predictors with large probability of over 0.95 under both link functions; X_5 is most likely to dominate all other predictors except X_3 with large probability of over 0.85 under both link functions. The 95% confidence intervals of dominance probabilities of θ_{34} , θ_{35} and θ_{36} all have lower limits greater than 0.85 and upper limits close to 1.0, while the 95% confidence intervals of θ_{13} , θ_{23} have lower limits close to 0 and upper limits less than 0.01; both imply that X_3 has large chances of dominance over the other predictors. Similarly, we can find that X_5 has more chance of dominating the other predictors except X_3 . In the same way, we can reveal the dominance order of remaining variables is X_6 , X_1 , X_4 and X_2 . Because the main inferences about the posterior dominance probability mains with little changes, we can confidently say that the model is stable under the sensitivity analysis, and that the results given are robust.

The posterior probabilities of the different orders of dominance shows that among all the possible orders of dominance, the order with $X_3 > X_5 > X_6 > X_1 > X_4 > X_2$ has the largest posterior probability (0.9744 and 0.9932 for logit and probit link function, respectively), which implies that it is the most likely order of dominance.

6. Summary and Concluding Remarks

Relative importance analyses will permit a greater understanding of the particular role played by variables in a multiple regression equation. Crucially, these analyses can reveal the underlying impact of a particular predictor more accurately than standardized regression coefficients or simple correlations. This paper extends the current research practices of measuring the relative importance of predictors in linear regression models by introducing an original method built

on a Bayesian pair-wise comparison model in a multiple regression context.

The advantages of the Bayesian approach are that it allows the use of genuine prior information in addition to the information that is available in the observed data for producing better results. In general, Bayesian methods provide a better approximation to the level of uncertainty or, conversely, the amount of information provided by the model and the data than other approaches (Rossi and Allengy 2003). In addition to providing useful statistics, such as, the mean and percentiles of the posterior distribution of the unknown parameters, Bayesian methods give more reliable results for small samples (Lee and Song 2004).

More specifically the Bayesian approach presented in this paper offers several advantages over the method of Dominance Analysis in determining the relative importance of predictors in a regression model. First, instead of classifying the relative importance into three categories, the Bayesian method of dominance identifies the most likely order of relative importance amongst all ordering of predictors in a linear regression setting. Second, the Bayesian approach provides more information about the relative importance of the predictor by making straightforward statements about the posterior probability of each possible order of dominance, which is the probability of each ordering given the model and the data. Third, the Bayesian method is not limited to a simple model but offers a rich potential to incorporate more complex models.

When complete dominance or conditional dominance of Dominance Analysis can be established among the predictors, Bayesian approach reaches similar results. The general dominance weight of Dominance Analysis is biased towards the results of subset models with small size, Bayesian approach resolved this problem by modelling the number of times a predictor outranking the others. Also, a probabilistic model based approach provides more comprehensive inference about the population dominance ability and the population relative dominance ability of predictors. In conclusion, the Bayesian method overcomes several of the limitations of Dominance Analysis.

Relative importance analyses are also applicable to situations commonly confronted by organizational scholars where the criteria may not meet the meet the distributional assumptions of ordinary least squares (OLS) regression, such as predicting binary criteria like turnover, promotion decisions, or training success. Although both Dominance Analysis and Bayesian approach were developed for use with OLS regression, Azen and Traxel (2009) presented modifications of these respective analyses to handle categorical criterion variables that would typically be analyzed using logistic regression. Thus, we will extend our work to questions of the relative contribution of each of the variables in terms

Table 1: Population Correlation Matrix for Simulation

	Y	X ₁	X ₂	X ₃	X ₄
Y	1.0	0.6	0.3	0.4	0.5
X ₁	0.6	1.0	0.8	0.1	0.3
X ₂	0.3	0.8	1.0	0.1	0.1
X ₃	0.4	0.1	0.1	1.0	0.5
X ₄	0.5	0.3	0.1	0.1	1.0

Table 2: Paired Comparisons of Predictors in Simulation

	X ₁	X ₂	X ₃	X ₄
X ₁	0	4	4	4
X ₂	0	0	2	1
X ₃	0	2	0	2
X ₄	0	3	2	0

Table 3: Summary of Posterior Statistics of Dominance Index of Predictors in Simulation

	Logit Link		Probit Link	
	Mean	95% C.I.	Mean	95% C.I.
d_1	2.7350	(0.1552,7.2633)	1.8853	(0.1038,5.3091)
d_2	-1.1951	(-3.9031,0.8466)	-0.6464	(-2.5381,0.9066)
d_3	-0.8814	(-3.5831,1.1841)	-0.4691	(-2.3070,1.078)
d_4	-0.5836	(-3.1231,1.5113)	-0.2950	(-2.1111,1.2940)

of predicting the categorical criterion that can be examined in this context as well.

One of the central questions in a multivariate analysis of variance (MANOVA) concerns identifying the dependent variables that are driving the significant multivariate F-test. Unfortunately, the correlations among the various dependent variables often make it difficult to accurately identify the role being played by the various dependent variables. The Bayesian approach of relative importance analyses may also be a useful supplement to analyses other than multiple regression, such as MANOVA.

Table 4: Summary of Posterior Statistics of Dominance Probability of Predictors in Simulation

	Logit Link		Probit Link	
	Mean	95% C.I.	Mean	95% C.I.
θ_{12}	0.9807	(0.7033,0.9998)	0.8263	(0.7300,0.9999)
θ_{13}	0.9738	(0.6677,0.9998)	0.8231	(0.6932,0.9999)
θ_{14}	0.9651	(0.6277,0.9997)	0.8188	(0.6553,0.9999)
θ_{23}	0.4222	(0.1594,0.7140)	0.6689	(0.1664,0.7234)
θ_{24}	0.3517	(0.1158,0.6561)	0.6468	(0.1188,0.6603)
θ_{34}	0.4261	(0.1662,0.7268)	0.6692	(0.1651,0.7227)

Table 5: Posterior Probability of Dominance Hierarchies of Predictors in Simulation

Dominance Hierarchy	Logit Link	Probit Link
X_1, X_4, X_3, X_2	0.3489	0.3453
X_1, X_3, X_4, X_2	0.2079	0.2021
X_1, X_4, X_2, X_3	0.2043	0.2061
X_1, X_3, X_2, X_4	0.0863	0.0887
X_1, X_2, X_4, X_3	0.0799	0.0882
X_1, X_2, X_3, X_4	0.0543	0.0574

References

- [1] Adams, E.S. (2005). Bayesian analysis of linear dominance hierarchies. *Animal Behaviour*, 69: 1191-1201.
- [2] Allenby, G.M. (1990). Hypothesis Testing with Scanner Data: The Advantage of Bayesian Methods. *Journal of Marketing Research*, 27(4): 379-389.
- [3] Azen, R.; Budescu, D. V.; and Reiser, B. (2001), Criticality of predictors in multiple regression. *British Journal of Mathematical and Statistical Psychology*, 54: 201-225.
- [4] Azen, R. and Budescu, D.V. (2003). The Dominance Analysis Approach for Comparing Predictors in Multiple Regression, *Psychological Methods*, 8 (2): 129-148.
- [5] Azen, R., and Traxel, N. M. (2009). Using dominance analysis to determine

Table 6: Additional R^2 under different model size in Auto-Repair Service Example

k	X_1	X_2	X_3	X_4	X_5	X_6
0	0.3141	0.2545	0.5392	0.0939	0.3333	0.1058
1	0.1423	0.1040	0.3382	0.0463	0.1735	0.0671
2	0.0573	0.0340	0.2271	0.0274	0.0937	0.0501
3	0.0202	0.0072	0.1666	0.0178	0.0567	0.0399
4	0.0073	0.0009	0.1326	0.0112	0.0398	0.0327
5	0.0025	0.0000	0.1085	0.00614	0.0278	0.0277
Average	0.0906	0.0668	0.2520	0.0338	0.1208	0.0539

Table 7: Paired Comparisons of Predictors in Auto-Repair Service Example

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	0	16	0	9	0	4
X_2	0	0	0	5	0	2
X_3	16	16	0	16	16	16
X_4	7	11	0	0	0	0
X_5	16	16	0	16	0	13
X_6	12	14	0	16	3	0

Table 8: Summary of Posterior Statistics of Dominance Index of Predictors in Auto-Repair Service Example

		a=b=0.001		a=b=0.01		a=b=0.5	
		Mean	95% C.I.	Mean	95% C.I.	Mean	95% C.I.
Logistic Link Function	d_1	-2.1085	(-6.3971, 1.0181)	-1.7806	(-5.2730, 1.2141)	-1.4017	(-3.4852, 0.5061)
	d_2	-4.0439	(-8.4332, -0.8990)	-3.7172	(-7.2251, -0.6510)	-3.2571	(-5.4010, -1.3360)
	d_3	5.8663	(1.8969, 9.5420)	5.9161	(2.3737, 9.5841)	5.0701	(2.6250, 8.1240)
	d_4	-3.0239	(-7.3240, 0.1322)	-2.6899	(-6.1741, 0.3132)	-2.2725	(-4.3820, -0.3636)
	d_5	1.1653	(-3.0481, 4.3530)	1.4509	(-1.9401, 4.5880)	1.6159	(-0.4507, 3.6281)
	d_6	-0.6869	(-4.9471, 2.4371)	-0.3721	(-3.8502, 2.6510)	-0.0648	(-2.1230, 1.8540)
Probit Link Function	d_1	-1.0868	(-3.9380, 0.9633)	-0.9311	(-3.2110, 1.4492)	-0.8384	(-2.4562, 0.6379)
	d_2	-2.1778	(-5.0393, -0.1569)	-2.0219	(-4.3541, 0.3359)	-1.9160	(-3.5551, -0.4089)
	d_3	3.7916	(1.1839, 8.5102)	3.8284	(1.3600, 8.4391)	3.2523	(1.4510, 5.7860)
	d_4	-1.6285	(-4.5231, 0.3913)	-1.4715	(-3.7340, 0.8846)	-1.3718	(-3.0091, 0.1077)
	d_5	0.7482	(-2.0390, 2.8121)	0.9093	(-1.3161, 3.3671)	0.9404	(-0.6502, 2.4660)
	d_6	-0.3153	(-3.1820, 1.7161)	-0.1574	(-2.4185, 2.2110)	-0.0833	(-1.6900, 1.4301)

predictor importance in logistic regression. Journal of Educational and Behavioral Statistics, 34, 319-347

[6] Bansal, H.S.; Taylor, S.F.; and St. James, Y. (2005). Migrating to New Service Providers: Toward a Unifying Framework of Consumers Switching

Table 9: Summary of Posterior Statistics of Dominance Probability of Predictors in Auto-Repair Service Example under Logit Link Function

	a=b=0.001		a=b=0.01		a=b=0.5	
	Mean	95% C.I.	Mean	95% C.I.	Mean	95% C.I.
θ_{12}	0.8647	(0.7346 , 0.9515)	0.8567	(0.7243 , 0.9469)	0.8632	(0.7324 , 0.9502)
θ_{13}	0.0013	(0.0000 , 0.0075)	0.0031	(0.0001 , 0.0133)	0.0013	(0.0000 , 0.0075)
θ_{14}	0.7063	(0.5267 , 0.8560)	0.6987	(0.5255 , 0.8460)	0.7054	(0.5269 , 0.8553)
θ_{15}	0.0444	(0.0085 , 0.1161)	0.0548	(0.0129 , 0.1321)	0.0447	(0.0089 , 0.1160)
θ_{16}	0.2057	(0.0836 , 0.3687)	0.2172	(0.0942 , 0.3813)	0.2057	(0.0836 , 0.3664)
θ_{23}	0.0002	(0.0000 , 0.0012)	0.0005	(0.0000 , 0.0024)	0.0002	(0.0000 , 0.0012)
θ_{24}	0.2728	(0.1196 , 0.4601)	0.2786	(0.1299 , 0.4646)	0.2745	(0.1223 , 0.4621)
θ_{25}	0.0072	(0.0010 , 0.0218)	0.0096	(0.0017 , 0.0273)	0.0073	(0.0011 , 0.0224)
θ_{26}	0.0391	(0.0098 , 0.0948)	0.0445	(0.0123 , 0.1033)	0.0395	(0.0100 , 0.0944)
θ_{34}	0.9995	(0.9967 , 1.0000)	0.9987	(0.9938 , 1.0000)	0.9995	(0.9968 , 1.0000)
θ_{35}	0.9748	(0.8752 , 1.0000)	0.9510	(0.8323 , 0.9984)	0.9737	(0.8732 , 1.0000)
θ_{36}	0.9952	(0.9725 , 1.0000)	0.9891	(0.9556 , 0.9997)	0.9950	(0.9726 , 1.0000)
θ_{45}	0.0189	(0.0031 , 0.0518)	0.0244	(0.0051 , 0.0627)	0.0191	(0.0034 , 0.0538)
θ_{46}	0.0976	(0.0325 , 0.2032)	0.1073	(0.0379 , 0.2145)	0.0979	(0.0317 , 0.2000)
θ_{56}	0.8498	(0.6870 , 0.9601)	0.8286	(0.6607 , 0.9437)	0.8485	(0.6816 , 0.9591)

Table 10: Summary of Posterior Statistics of Dominance Probability of Predictors in Auto-Repair Service Example under Probit Link Function

	a=b=0.001		a=b=0.01		a=b=0.5	
	Mean	95% C.I.	Mean	95% C.I.	Mean	95% C.I.
θ_{12}	0.8539	(0.7211 , 0.9483)	0.8509	(0.7140 , 0.9469)	0.8536	(0.7197 , 0.9496)
θ_{13}	0.0003	(0.0000 , 0.0026)	0.0005	(0.0000 , 0.0042)	0.0003	(0.0000 , 0.0027)
θ_{14}	0.7007	(0.5299 , 0.8483)	0.6980	(0.5285 , 0.8426)	0.7003	(0.5319 , 0.8478)
θ_{15}	0.0417	(0.0050 , 0.1206)	0.0461	(0.0064 , 0.1274)	0.0412	(0.0049 , 0.1196)
θ_{16}	0.2272	(0.1012 , 0.3859)	0.2320	(0.1041 , 0.3898)	0.2266	(0.1001 , 0.3838)
θ_{23}	0.0000	(0.0000 , 0.0001)	0.0000	(0.0000 , 0.0001)	0.0000	(0.0000 , 0.0001)
θ_{24}	0.2974	(0.1447 , 0.4785)	0.2990	(0.1462 , 0.4789)	0.2970	(0.1450 , 0.4769)
θ_{25}	0.0032	(0.0001 , 0.0146)	0.0038	(0.0001 , 0.0163)	0.0032	(0.0001 , 0.0146)
θ_{26}	0.0371	(0.0067 , 0.0980)	0.0394	(0.0075 , 0.1019)	0.0369	(0.0068 , 0.0976)
θ_{34}	0.9999	(0.9995 , 1.0000)	0.9999	(0.9992 , 1.0000)	0.9999	(0.9995 , 1.0000)
θ_{35}	0.9782	(0.8737 , 1.0000)	0.9675	(0.8561 , 1.0000)	0.9773	(0.8683 , 1.0000)
θ_{36}	0.9977	(0.9808 , 1.0000)	0.9962	(0.9728 , 1.0000)	0.9976	(0.9804 , 1.0000)
θ_{45}	0.0130	(0.0008 , 0.0465)	0.0147	(0.0011 , 0.0504)	0.0128	(0.0008 , 0.0463)
θ_{46}	0.1029	(0.0310 , 0.2159)	0.1071	(0.0330 , 0.2250)	0.1027	(0.0302 , 0.2152)
θ_{56}	0.8446	(0.6817 , 0.9578)	0.8362	(0.6717 , 0.9506)	0.8454	(0.6839 , 0.9575)

Behaviors. Journal of the Academy of Marketing Research, 33(1): 96-115.

- [7] Budescu, David V. (1993), Dominance Analysis: A New Approach to the Problem of Relative Importance of Predictors in Multiple Regression, Psychological Bulletin, 114, 542-551.

- [8] Bradley, R.A.; and Terry, M.E. (1952). Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika* 39(3/4):324-34.
- [9] Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, 114, 542-551.
- [10] Carlin, B. P., and Louis, T. A. (2001). *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. London: Chapman and Hall.
- [11] Courville, T., and Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: b is not enough. *Educational and Psychological Measurement*, 61, 229-248.
- [12] Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, 69, 161-182.
- [13] David, H. A. (1988) *The Method of Paired Comparisons*, 2nd edn. London: Chapman and Hall.
- [14] Davidson, R.R.; and Solomon, D.L. (1973). A Bayesian Approach to Paired Comparison Experimentation. *Biometrika* 60(3):477-487.
- [15] Efron, B., and Morris, C. (1975). Data analysis using Steins estimator and its generalizations. *Journal of the American Statistical Association* 70, 311-319.
- [16] Fabbris, L. (1980). Measures of predictor variable importance in multiple regression: An additional suggestion. *Quality and Quantity*, 4, 787-792.
- [17] Hoffman, P. J. (1962). Assessment of the independent contributions of predictors. *Psychological Bulletin*, 59, 77-80.
- [18] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*, second edition. London: Chapman and Hall.
- [19] James, W., and Stein, C. (1960). Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium 1*, ed. J. Neyman, 361-380. Berkeley: University of California Press.
- [20] Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, 35, 1-19.

- [21] Johnson, J.W. and LeBreton, J.M. (2004). History and Use of Relative Importance Indices in Organizational Research. *Organizational Research Methods*, 7 (3): 238-257.
- [22] Kreft, I., and De Leeuw, J. (1998). *Introducing Multilevel Modeling*. London: Sage.
- [23] Kruskal, W. and Majors, R. (1989). Concepts of Relative Importance in Recent Scientific Literature. *The American Statistician*, 43(1): 2-6.
- [24] Lahiri, P. (2003). On the Impact of Bootstrap in Survey Sampling and Small-Area Estimation. *Statistical Science*, 18(2): 199-210.
- [25] Larzelere, R. E. and Huston, T.L. (1980). The Dyadic Trust Scale: Toward Understanding Interpersonal Trust in Close Relationships. *Journal of Marriage and the Family*, 42:595-604.
- [26] LeBreton, J. M.; Ployhart, R.E.; and Ladd, R.T. (2004). A Monte Carlo Comparison of Relative Importance Methodologies. *Organizational Research Methods*, 7 (3): 258-282.
- [27] Lee, S.Y. and Song X.Y. (2004). Bayesian model comparison of nonlinear structural equation models with missing continuous and ordinal categorical data. *British Journal of Mathematical and Statistical Psychology*, 57: 131-150.
- [28] Leonard, T. (1977). An Alternative Bayesian Approach to the Bradley-Terry Model for Paired Comparisons. *Biometrics* 33(1):121-132.
- [29] Meyer, J.P. and Allen, N. J. (1997), *Commitment in the Workplace: Theory, Research, and Application*. Thousand Oaks, CA: Sage.
- [30] Morgan, R. M. and Hunt, S.D. (1994). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, 58, 20-38.
- [31] Morris, C.N. (1983). Parametric Empirical Bayes Inference: Theory and Applications. *Journal of the American Statistical Association*, 78(381): 47-55.
- [32] Montgomery, A.L. and Bradlow, E.T. (1999). Why Analyst Overconfidence about the Functional Form of Demand Models Can Lead to Overpricing. *Marketing Science*, 18(4): 569-583.

- [33] Oliver, R. L. and Swan, J.E. (1989). Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach. *Journal of Marketing*, 53 (April), 21-35.
- [34] Ping Jr., R. A. (1993). The Effects of Satisfaction and Structural Constraints on Retailer Exiting, Voice, Loyalty, Opportunism, and Neglect. *Journal of Retailing*, 69 (3): 320-352.
- [35] Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical Linear Models*, second edition. Thousand Oaks, Calif.: Sage.
- [36] Retzer, J.J.; Soofi, E.S.; and Soyer. R. (2009). Information importance of predictors: Concept, measures, Bayesian inference, and Applications. *Computational Statistics & Data Analysis*, 53(6): 2363-2377.
- [37] Rossi, P.E. and Allenby, G.M. (2003). Bayesian Statistics and Marketing. *Marketing Science*, 22(3): 304-328.
- [38] Scheines, R.; Hoijsink, H.; and Boomsma, A. (1999). Bayesian estimation and testing of structural equation models. *Psychometrika*, 64(1): 37-52.
- [39] Snijders, T. A. B., and Bosker, R. J. (1999). *Multilevel Analysis*. London: Sage.
- [40] Taylor, S.A. and Baker, T.L. (1994). An Assessment of the Relationship Between Service Quality and Customer Satisfaction in the Formation of Consumers Purchase Intentions. *Journal of Retailing*, 70 (2): 163-178.
- [41] Tonidandel, S. and LeBreton J.M. (2011). Relative Importance Analysis: A Useful Supplement to Regression Analysis. *Journal of Business and Psychology*, 26(1):1-9.

