A Heuristic Bayesian Regression Approach for Causal Explanatory Study: Exemplified by an IS Impact Study

Shing-Hwang Doong
Department of Information Management, Shu-Te University

Ching-Chang Lee
Department of Information Management,
National Kaohsiung First University of Science and Technology

Abstract

Causal explanatory study is a very important research method in empirical research whereof research models are frequently validated by multiple linear regressions (MLR) with significant factors sought. An alternative to MLR is Bayesian regressions where statistical inferences are made with samples drawn from posterior distributions. Efficient simulation algorithms of the Markov chain Monte Carlo type have made Bayesian regressions practical. We propose a heuristic method based on the outputs of MLR to construct informative priors for Bayesian regressions. Data collected from two empirical studies of information systems (IS) impact on performance is used to demonstrate the proposed method. Deviance information criterion shows that this heuristic procedure significantly improves a Bayesian modeling with uninformative priors. When credible intervals are used to locate significant factors, it is found that the heuristic Bayesian approach, capable of finding delicate drivers, can help design better diagnostics for IS problems.

Key words: Research methods, casual explanatory study, IS impact, Bayesian regressions, model selection

This research is partly supported by grant from National Science Council (NSC) of Taiwan under contract number NSC 95-2221-E-366-020-MY2 & NSC 95-2416-H-327-006.

因果解釋性研究的啟發式貝氏迴歸方法—以資訊系統影響研究為例

董信煌
樹德科技大學資訊管理學系

李慶章
高雄第一科技大學資訊管理學系

摘要

因果解釋性研究是實證研究中很重要的一種研究方法，在實證研究中學者常使用複迴歸方法來驗證研究模式並找到顯著因子。貝氏迴歸是一種不同於複迴歸的分析工具，它使用事後機率抽取樣本來做統計推論，由於馬可夫鏈蒙地卡羅演算法可以有效率依機率分佈來抽取樣本，貝氏迴歸分析已變得越來越可行。本研究提出一個基於複迴歸分析結果的啟發式方法來建構貝氏迴歸分析的資訊事前機率，來自於兩個不同的實證研究資料將被用來測試此方法，這兩個實證研究皆是在探討資訊系統對績效的影響。偏離資訊法則顯示出此一啟發式方法能顯著的改善使用非資訊事前機率塑模的貝氏迴歸分析，當信任區間被用來尋找顯著因子時，我們發現此一新方法能找到更細膩的因子且可以設計出更好的方法來診斷資訊系統問題。

關鍵字：研究方法、因果解釋性研究、資訊系統影響、貝氏迴歸、模式選擇

1. INTRODUCTION

Causal explanatory study is a very important research method in empirical research. A causal explanatory study is "designed to determine whether one or more variables explain the causes or effects of one or more outcome (dependent) variables" (Cooper and Schindler, 2008). An empirical study frequently starts with a literature review to set up a research model relating various constructs of interest. Questionnaire survey or secondary data collection is then conducted. After that, various statistical tools can be used to validate the research model with the collected data. Popular statistical tools include multiple linear regression (Hogg and Tannis, 1997) and structural equation modelling (Diamantopoulos and Siguaw, 2000). This study is focused on the former technique and its Bayesian counterpart.

A multiple linear regression (MLR) expresses a dependent variable as a linear combination of multiple independent variables plus a random noise. Techniques of parameters estimation and significance detection for MLR can be found in many textbooks of statistics (Devore, 2004; Hogg and Tannis, 1997). The statistics used in MLR is called "frequentist statistics", which assumes that we can make random samples repeatedly and the observed data is just one of these realizations (Garthwaite et al., 2002).

A competing approach to frequentist statistics is called "Bayesian statistics", where the observed data is not re-sampled but rather used to update our prior belief of distributions for parameters of interest. In the Bayesian approach, regression parameters are considered random variables (in contrast to the constant value perspective of MLR) with their own probability distribution functions. Using Bayes' theorem and the observed data, prior distributions are updated to posterior distributions, of which samples are drawn to make statistical inferences (Garthwaite et al., 2002; Urbach, 1992).

A classical MLR lacks the flexibility to incorporate our prior belief into a regression analysis, and "p-values are commonly misinterpreted in this manner" (Burton et al., 1998, p. 318). On the other hand, a Bayesian approach gives us more flexibility to control regression parameters and results from a Bayesian regression are more intuitive. Indeed, the Bayesian approach "is quite similar to how the human mind works and thus it feels very natural" (Thorburn, 2005, p. 80). Bayesian statistics has gained the attention of many econometricians (Gallizo et al., 2002; Koop, 2003; Wright, 2003) and epidemiologists (Burton et al., 1998; Greenland, 2007).

Literature shows that previous information systems (IS) studies have used regression techniques other than MLR to analyze causal models. Bansal et al. (1993) compared the modeling performance of MLR and neural networks when data quality was a concern. They found that MLR outperformed neural nets in terms of forecasting accuracy, but the opposite was

true when the business value of the forecast was used to measure performance. The researchers also discovered that performance of the neural nets tended to be robust even when the data was corrupted by noise. Lately, Banerjee et al. (2005) used a dynamic Bayesian analysis to study drivers of Internet firm survival. However, they did not consider the priors selection and model comparison issues in their study.

The choice of a prior distribution reflects the information available to a user at the time of a Bayesian analysis. Unless the prior information has been overly distorted, it is convenient to choose mathematically tractable forms of priors and likelihood functions. This may be achieved through the use of conjugate priors where prior and posterior distributions are of the same type (Denison et al., 2002). In practice, Markov chain Monte Carlo (MCMC) type simulation techniques can be used to get a good insight of the posterior distributions. MCMC simulations require both an efficient algorithm and a powerful computer. With the incredible dissemination power of the Internet, WinBUGS program (Spiegelhalter et al., 2003) has become a very popular MCMC simulation package for Bayesian regression.

Cao et al. (2006) argued that different research paradigms may complement each other due to the interactions involved in different research methods and activities. In this research, we show that Bayesian regressions can become an alternative and effective tool in a causal explanatory study. The objectives of this research are (1) to study the prior distributions selection problem and to find an efficient and effective way to construct informative priors; (2) to locate and explain significant factors in Bayesian regressions with a sound foundation like the MLR approach; and (3) to compare MLR and Bayesian models with well-known model assessment criteria. Informative priors with an objective background will be sought. Two empirical studies of IS impact on performance based on the task technology fit theory will be used to show a novel priors selection method in Bayesian regressions.

This paper is organized as follows. In section 2, we introduce important concepts in MLR and Bayesian regressions. Deviance information criterion will be introduced to compare two Bayesian models. Section 3 is devoted to a discussion of the survey research that founds causal models in this study. In section 4, an empirical study is set up to collect the primary data for model validation. MLR and Bayesian regressions with uninformative priors are applied to analyze the data. We propose a heuristic method to construct informative priors in section 5. Comparison between MLR and Bayesian models, extraction of significant factors and a second empirical study are also presented in this section. Finally, we conclude in section 6 with a few remarks about implications of this study.

2. MATERIALS AND METHODS

We briefly review MLR especially in the area of parameter significance test. A careful

observation of the test leads us to a comparable parameter significance test in Bayesian regressions. Deviance information criterion will be used to compare different Bayesian models.

2.1 Multiple linear regressions

Kennedy (2003) has a good discussion on MLR from the viewpoint of an econometrician. Assumptions related to specification errors, disturbance bias, homoskedasticity, autoregression, exogeneity and multicollinearity are often checked before an MLR analysis is conducted. For the purpose of this discussion, it is assumed that the dependent variable Y is related to the independent variables $x_1, \cdots, x_k$ by a linear equation with parameters $\beta_o, \cdots, \beta_k$. It is further assumed that the output is corrupted by a random noise ε :

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \varepsilon \tag{1}$$

The regression parameters $\beta_o, \cdots, \beta_k$ are fixed but unknown, and they can be estimated by several estimators given a set of observed data$(\vec{x}_i, y_i)$, i=1,$\cdots$, n. Here $\vec{x}_i = (x_{1i}, \cdots, x_{ki})$ is the input vector for the i^{th} sample and y_i is the corresponding output value. For example, the ordinary least square (OLS) estimator is used to compute parameters $\hat{\beta}_0^{OLS}, \ldots, \hat{\beta}_k^{OLS}$ so that $SSE = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$ is minimized. Here $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1_i} + \cdots + \hat{\beta}_k x_{ki}$ is a computed output with the estimated parameters $\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_k$.

The frequentist approach assumes that the random noise ε can be realized repeatedly so that we have many realizations of data$(\vec{x}_i, y_i), i = 1, \ldots, n$. Across all realizations, the input vector $\vec{x}_i$ is unchanged and the output y_i is corrupted by a realized noise. With an estimator like OLS, each of these realizations yields an estimated $\hat{\beta}_j$ for the unknown parameter β_j. With all these estimates, interesting results about the true parameter β_j can be inferred by hypothesis testing procedures in statistics. Significant factors (drivers) are often sought in a causal explanatory study to explain how input variables affect the output variable. In MLR, the significance of a predictor variable x_j is tested against the following null hypothesis:

$$\mathrm{H}_0 : \beta_j = 0. \tag{2}$$

The variable x_j is said to be a driver if H_0 is *rejected*. When the random noise ε is normally distributed with a zero mean (unbiased) and a constant variance at all $\vec{x}_i$ (homoskedasticity), H_0 can be tested by using the OLS estimator $\hat{\beta}_j^{OLS}$ and a Student t-distribution (Hogg and Tannis, 1997). For example, given a significance level of α (typically .05 in social science studies), H_0 is rejected when the following t statistic

$$t = \frac{\hat{\beta}_j^{OLS}}{s_{\hat{\beta}_j^{OLS}}} \tag{3}$$

is greater than $t_{\alpha/2,n-(k+1)}$ or less than $-t_{\alpha/2,n-(k+1)}$. Here $t_{\alpha/2,n-(k+1)}$ denotes the α / 2 critical value of the t-distribution of n – (k + 1) degree of freedom, $s_{\hat{\beta}_j^{OLS}}$ is the estimated standard error of $\hat{\beta}_j^{OLS}$, n is the number of observed data and k is the number of predictor variables (Devore, 2004). A (1 - α) confidence interval for β_j is then given by

$$(\hat{\beta}_j^{OLS} - t_{\alpha/2,n-(k+1)} \cdot s_{\hat{\beta}_j^{OLS}},\ \hat{\beta}_j^{OLS} + t_{\alpha/2,n-(k+1)} \cdot s_{\hat{\beta}_j^{OLS}}) \tag{4}$$

This confidence interval is centered at the OLS estimate of a parameter. A simple observation shows that H_0 is rejected if and only if both endpoints in Eq. (4) are of the same sign, i.e. x_j is a driver if and only if the confidence interval does not contain zero. This can be verified as follows. Suppose H_0 is rejected because the t statistic in Eq. (3) is greater than $t_{\alpha/2,n-(k+1)}$, then

$$\frac{\hat{\beta}_j^{OLS}}{s_{\hat{\beta}_j^{OLS}}} > t_{\alpha/2,n-(k+1)}$$

$$\Leftrightarrow \hat{\beta}_j^{OLS} > t_{\alpha/2,n-(k+1)} \cdot s_{\hat{\beta}_j^{OLS}}$$

$$\Leftrightarrow \hat{\beta}_j^{OLS} - t_{\alpha/2,n-(k+1)} \cdot s_{\hat{\beta}_j^{OLS}} > 0,$$

since $s_{\hat{\beta}_j^{OLS}} > 0$. Thus, the left endpoint of Eq. (4) is positive and the confidence interval does not include zero. If H_0 is rejected because of the other reason, then both endpoints in Eq. (4) are negative. This simple observation will form our basis to locate drivers in a Bayesian approach.

Though parameters estimation in MLR is easy, it is hard to interpret the result. Frequently a statement like "a 95% confidence interval for β_j is (-.5, 1.5)" is misinterpreted as the true value of β_j lies in the interval with a probability of .95. Urbach (1992, p. 322) explains why "no such probability statement is inferable". The correct interpretation should be: if sufficiently many data sets $(\vec{x}_i, y_i), i = 1,\ldots,n$ are realized according to the frequentist procedure described above and endpoints are computed by Eq. (4), then approximately 95% of these intervals will contain the true value of β_j. Since the particular interval (-.5, 1.5) is just one of these intervals, it may or may not belong to the group of intervals containing β_j.

2.2 Bayesian regressions

Bayesian regressions consider each regression parameter a random variable with its own probability distribution function (pdf). If the random noise ε in Eq. (1) is assumed to have a zero mean and a constant but unknown variance σ^2, then the parameters of interest in a regression study include $\vec{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)$ and σ^2. Let $D = \{(\vec{x}_i, y_i), i = 1,\ldots,n\}$ denote the data collected by a researcher. According to Bayes' theorem, the posterior distribution of parameters of interest after observing data D is given by

$$p(\vec{\beta}, \sigma^2 \mid D) = \frac{p(D \mid \vec{\beta}, \sigma^2)\, p(\vec{\beta}, \sigma^2)}{p(D)}, \tag{5}$$

where $p(\vec{\beta}, \sigma^2)$ is a prior distribution of parameters given by a user, the likelihood function $p(D \mid \vec{\beta}, \sigma^2)$ is the probability of observing D with given model parameters, and $p(D) = \int p(D \mid \vec{\beta}, \sigma^2)\, p(\vec{\beta}, \sigma^2)\, d\vec{\beta} d\sigma$ is the marginal likelihood. We will handle the prior

distributions problem in a later section. But, first let us explain how to generalize the simple observation about the significance test of parameters in MLR to Bayesian regressions.

In Bayesian regressions, inferences about parameters are made by using samples drawn from the posterior distribution. After these posterior samples are tallied on a real line, interesting percentile points are marked on the real line. For example, assume that 10000 samples have been drawn for β_j according to the posterior distribution $p(\vec{\beta}, \sigma^2 \mid D)$. These 10000 numbers are lined up on a real line and the 2.5%, 50% and 97.5% sample cumulative points are determined. Let *a* denote the 2.5% sample cumulative point and *b* the 97.5% sample cumulative point. Then, out of the 10000 posterior samples, 95% of them are between *a* and *b*. Thus (*a*,*b*), called the centralized 95% *credible* interval, contains β_j with a probability of .95 according to the posterior distribution of the parameter. Based on the confidence interval perspective of significance test in MLR, we make the following definition.

Definition (significant factors): In a Bayesian regression, a predictor x_j is *significant* if the centralized 95% credible interval (*a*,*b*) does not contain zero.

Efficient drawing of posterior samples is a major task in Bayesian regressions. Markov chain Monte Carlo type algorithms have been developed to sample data efficiently. Special cases of MCMC simulations include Gibbs sampling and Metropolis-Hastings algorithm. Since MCMC simulations are iterative procedures, it is important to make sure that the iteration converges before any posterior samples are taken into consideration for inferences making. This means that samples from the *burn-in* phase must be kept away. Two criteria are commonly used for checking the convergence: (1) trace plots of the sampled data; and (2) autocorrelation function plots. For a convergent simulation, the trace plots exhibit a saw-like shape while the autocorrelation function plots drop very quickly to nearly zero after a few lag steps (Garthwaite et al., 2002).

Unlike MLR with OLS where only one set of estimated parameters is available for a given data set, Bayesian regressions have the freedom to choose different prior distributions for different results. This makes model comparison and selection an important issue in Bayesian regressions. Deviance information criterion (*DIC*) is a useful tool for comparing different Bayesian models (Spiegelhalter et al., 2002). The deviance is defined by the log likelihood function as

$$V(\vec{\beta}, \sigma^2) = -2\log(p(D \mid \vec{\beta}, \sigma^2)). \tag{6}$$

The expectation $\overline{V} = E^{\vec{\beta},\sigma^2}(V)$ taken over the parameter space measures how well the model fits the data. The larger it is, the worse the model. This $\overline{V}$ is an in-sample assessment measure like *SSE* in MLR. Additionally, the effective number of parameters is defined by

$$P_V = \overline{V} - V(\overline{\vec{\beta}}, \overline{\sigma}^2), \tag{7}$$

where $\overline{\vec{\beta}}, \overline{\sigma}^2$ are the expected values of model parameters. Here expected values are all

taken with respect to the posterior distribution of parameters. The larger this effective number is, the easier it is for the model to fit the data. This P_V is similar to a model complexity term in statistical learning theory (Vapnik, 1998). Finally, the *DIC* is defined as

$$DIC = \overline{V} + P_V , \tag{8}$$

where constituent terms on the right side work against each other. For example, as P_V goes up, $\overline{V}$ goes down and as P_V goes down, $\overline{V}$ goes up. The idea for model selection is to choose a model with small *DIC*, i.e. balancing the effects of data fit ($\overline{V}$) and model complexity (P_V). This is similar to the famous Occam's razor principle which says that, with approximately equal data fit capacity, parsimonious models are better than complex ones (Witten and Frank, 2005). Late developments in support vector machines also support this viewpoint of model selection (Schköpf and Smola, 2002). According to a recommendation from Spiegelhalter et al. (2003), if the difference in *DIC* between two models is more than 10, one might definitely rule out the model with a higher *DIC*. When the difference is between 5 and 10, it is considered significant. On the other hand, if the difference in *DIC* is less that 5, it could be misleading to report results from the model with a lower *DIC*. We take this recommendation in our later demonstration of Bayesian regressions.

3. DESIGN FOR A CASUAL RESEARCH

In order to exemplify the Bayesian regression approach, we set up a survey research based on the task technology fit theory to collect empirical data.

3.1 Task technology fit (TTF) theory

The influence of information technology (IT) on individual performance has been an ongoing issue in IS research. There are two main streams on which technology to performance chain (TPC) study can be based. The first stream focuses on utilization of technology and implicitly assumes that increased utilization results in improved performance. This stream, exemplified by technology acceptance model (Davis, 1989), employs users' attitudes and beliefs to predict utilization of IT. When IT utilization is assured because of task requirements or other reasons, the second stream of TPC study prevails. In this stream of study, the fit between adopted technology and task characteristics becomes a crucial factor impacting users' performance. For example, different types of decision making tasks (in trend or detailed analysis) often require different types of data representation technology (by graphs or tables). Using laboratory experiments, previous studies have shown the impact of TTF on performance (Dickson, et al., 1986; Vessey, 1991).

Goodhue and Thompson (1995) argued that for an IT to have positive influences on a

user's performance, the technology must have a good fit with the tasks it supports and proposed the TTF theory. According to the TTF theory, the existence of a fit among task, technology and user promotes the willingness of a user to use the technology and thus improves the user's work performance (Figure 1). Tasks are defined as the actions performed by users to turn inputs into outputs. Technologies are tools provided to users to finish their tasks. Users have different characteristics that may affect how they use technologies to carry out tasks. Task technology fit is "the correspondence between task requirements, individual abilities and the functionality of the technology" (Goodhue and Thompson, 1995, p.218).

The antecedents of TTF involve complex interactions between tasks, technologies and users characteristics. Within the domain of decision making, TTF has been successfully measured (Goodhue, 1995). Two additional IT-supported task domains are included in Goodhue and Thompson (1995), namely responding to changed business environments and conducting day to day business operations. Instruments from Goodhue (1995) were borrowed to measure TTF in the domain of decision making, and new questions were developed for measuring TTF in the two new task domains. After conducting the reliability and validity assessment, Goodhue and Thompson (1995) preserved 34 questions in their final questionnaire. Using factor analysis, these questions were categorized into eight TTF factors: data quality, data locatability, authorization, compatibility, ease of use/training, production timeliness, systems reliability, and relationship with users. The first five factors are related to user tasks in decision making. The next two are related to user tasks in conducting day to day business operations and the last one is focused on responding to changed business requirements. These factors are further explained as follows:

1. Data quality: This factor includes three dimensions which are currency of the data, proper maintenance of right data and right granularity of the data.
2. Data locatability: Two dimensions are associated with this factor, and they are ease of determining what data is available where and ease of determining the meaning of a data element.
3. Authorization: User perception of obtaining authorization to access data necessary to do the job.
4. Compatibility: User perception of consolidating data from different sources without inconsistency.
5. Ease of use/Training: This factor has two dimensions. These are ease of use of hardware and software, and obtaining proper training for using the IS.
6. Production timeliness: The IS meets task schedule.
7. System reliability: The IS is dependable and provides consistent service level of access.
8. Relationship with users: This factor has five dimensions. Altogether, they measure the ability of the IS to meet changed business requirements.
   i. IS understanding of business: Does the IS understand business mission and goal?

ii. IS interest and dedication: Does the IS have high interest and dedication to support customers?
iii. Responsiveness: Is the IS responsive to service requests?
iv. Consulting: Does the IS offer consulting service?
v. IS performance: Does the IS deliver agreed-upon solutions?

The TTF theory predicts that these eight factors have strong influences on a user's performance. Goodhue and Thompson (1995) suggested that TTF could be a diagnostic tool to evaluate whether information systems and services are meeting user requirements and that TTF might be a good surrogate measure of IS success when utilization is assured. Thus, the measure of TTF can be used to identify gaps between systems capabilities and user requirements. By understanding specific gaps, managers can make decision on stopping or redesigning systems or redesigning tasks to exploit IT potential.

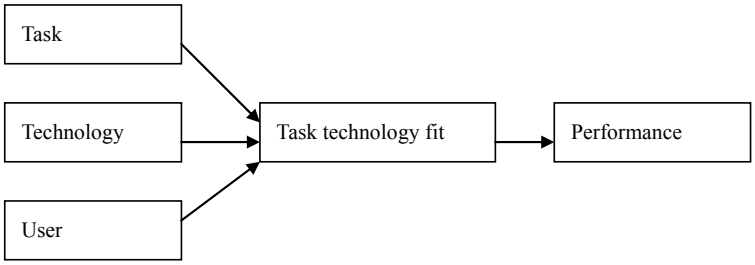

Figure 1: Task technology fit (Goodhue and Thompson, 1995)

3.2 The case company and the IS

The target system, Mobile Dr. Insurance system, is the most popular mobile insurance system developed by company G. This IS has been adopted by over twelve life insurance corporations in Taiwan for more than three years. The Mobile Dr. Insurance system has a customer base of over 30,000 agents since its inception. We conducted a survey research to collect user evaluations of TTF and perceived performance from the customer base of the system.

Insurance agents represent their companies to deal with customers for the needs in insurance. In Taiwan, an agent works in the field to visit customers for most of his (her) work. Thus, an information system capable of mobility is deemed suitable for an agent. After many discussions with senior agents, it was found that an agent had three primary tasks to fulfil:

1. Recruitment of new contracts. The most important task of an agent is to identify potential customers, provide the most appropriate policies and close the deals. Definitely, some of these goals can be helped with the technology of data mining.
2. Post-contract customer services. The market competition is so fierce that customer satisfaction is nothing short of a mandatory requirement in this industry. An agent needs to help customers change policies, beneficiaries or the payment methods from time to

time.

3. Tax and legal information services. People need tax and legal advices when they make major financial decisions. Customer relationship management tells us that supplying these services may result in cross sales of other products from an affiliated company.

Performance of the above three tasks will be used as the dependent variable in our survey research, while TTF factors are the independent variables.

3.3 The research model

The task and technology in Figure 1 are the same for every survey respondent. The task is one of the three major tasks explained in the last section, and the adopted technology is the Mobile Dr. Insurance system. Therefore, only individual characteristics can affect TTF degrees in Figure 1. To demonstrate the Bayesian approach for a causal explanatory study, we restrict our attention to the second part of the TTF model, namely the impact of TTF factors on performance. It was found that users of the subject system must pay a fee to use the services. Therefore, users of the system were properly authorized for access to data. Thus, we dropped the authorization factor as an independent variable in this study. Corresponding to the three major tasks of an insurance agent, three causal effects of seven TTF factors on task performance are modeled in Figure 2.

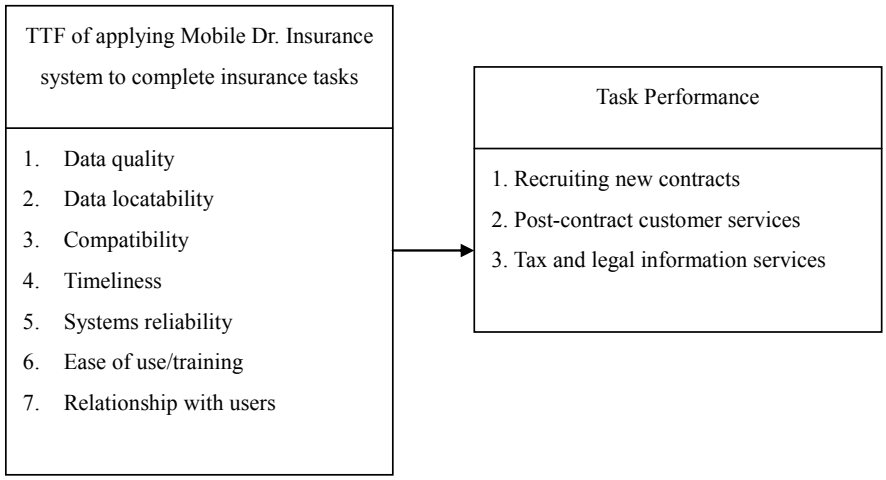

Figure 2: The research model

4. THE FIRST EMPIRICAL STUDY

An empirical study for the research model was conducted. Classical MLRs and Bayesian regressions with uninformative priors will be compared in this section.

4.1 Data description

In this research, data collected from a questionnaire survey is used to support an empirical study of IS impact on performance. The input variables include seven TTF factors (data quality, data locatability, compatibility, timeliness, system reliability, ease of use/training and relationship with users) and the output variable is user's performance of the three major tasks for an insurance agent. Three causal effects will be analyzed in total.

4.2 Samples collection

An Internet-based survey was conducted with the sponsorship and cooperation of company G, developer of the Mobile Dr. Insurance system. Users of the system were invited by e-mails to answer questionnaires published on the company's web site between October 1 and November 13 in 2005. In total, 307 agents submitted complete responses to the survey.

Out of the 307 useful samples, there were 45% male and 55% female. Regarding the age distribution, respondents in the highest percentage group (48%) were between 31 and 40 years old, while agents in the lowest percentage group (24%) were 41 years old and above. In terms of education, 41% of the respondents received associate degree, 33% received college degree and 26% received high school diploma.

The questionnaire was prepared according to established procedures in Nunnally (1978). All questions were made according to the validated TTF factors explained previously. Measurement methods are described as follows.

1. Task technology fit factors

Survey questions related to the seven TTF factors used a 5-point Likert scale to collect responders' feedback for the degree of task technology fit. A user's responses to questions of the same TTF factor were averaged to get a representative score for the factor.

2. User's task performance

Impacts of the subject system on task performance were measured by the respondents' self-assessment of how useful utilizing the system had assisted them in performing the tasks of recruiting new contracts, post-contract customer services and tax and legal information services. The user's task performance was measured by a 5-point Likert scale.

3. Reliability and validity

Cronbach's α was used to measure the reliability of research instruments. In practical applications, the value of Cronbach's α should exceed .5, preferably more than .7 (Nunnally, 1978). The Cronbach's α of our instrument ranged from .595 to .885, indicating a medium high to high reliability. To ensure content validity, our questionnaire design was based on well-established and validated instruments in the literature (Goodhue and Thompson, 1995).

4.3 Results from MLR

An MLR uses the OLS estimator to compute regression parameters. Tables 1, 2 and 3 summarize results from the MLR analysis for the three tasks respectively. The respective adjusted R^2 for these MLR models is .484, .618 and .399, and each model is significant at the level of .05. Unlike time series data, adjusted R^2 for cross sectional data is frequently low. Indeed, adjusted R^2 for the TTF model in Goodhue and Thompson (1995) was only .14. Our models show a moderate level of goodness-of-fit with the stated R^2 values. Since each variance inflation factor (VIF) is smaller than 10, the data does not have any multicollinearity issue (Devore, 2004). Lower bound of the 95% confidence interval is denoted as 95%-lb, while upper bound is denoted as 95%-ub. With a cut-off of .05 for the p-value (the Sig column), it can be seen that data quality and timeliness are both significant factors affecting user's performance for task 1 positively. This means the higher a user evaluates the fit in data quality and production timeliness, the more the user perceives a positive impact of the IS on recruiting new contracts. These two drivers belong to domains in decision making and conducting day to day business operations respectively. During the process of recruiting new contracts, an agent needs to make many decisions such as which customers to contact and what kinds of insurance products to recommend. Data quality is the main concern of agents in this domain of decision making. Recruiting new contracts is a day to day operation for most insurance agents, therefore it is important for the IS to meet task schedule (production timeliness).

Regarding the second task, data quality, timeliness, system reliability and ease of use/ training are drivers of performance. These drivers cover domains in decision making and conducting day to day operations as above. The third task has the most drivers covering all three domains considered in Goodhue and Thompson (1995). All drivers have their 95% confidence intervals with endpoints of the same sign as claimed before. Standard error (Std. err) is an estimate of the standard deviation for a regression parameter. The smaller it is, the more efficient OLS has estimated a regression parameter (Kennedy, 2003).

Table 1: Multiple linear regression for task 1

TTF factor	Coeff.	Std. err	T	Sig	95%-lb	95%-ub	VIF
(Constant)	.019	.221	.087	.930	-.415	.454	
Data quality	.452*	.078	5.814	.000	.299	.605	2.479
Data locatability	.063	.051	1.223	.222	-.038	.164	1.846
Compatibility	.102	.055	1.862	.064	-.006	.209	1.715
Timeliness	.206*	.056	3.683	.000	.096	.317	1.537
System reliability	.032	.049	.669	.504	-.063	.128	1.490
Easy of use/Training	.000	.051	.010	.992	-.099	.100	1.811
Relationship with user	.118	.071	1.660	.098	-.022	.258	2.258

Adjusted R^2 = .484*.
*: significant with p-value < .05

Table 2: Multiple linear regression for task 2

TTF factor	Coeff.	Std. err	T	Sig	95%-lb	95%-ub	VIF
(Constant)	.258	.174	1.486	.138	-.084	.600	
Data quality	.542*	.061	8.847	.000	.421	.662	2.479
Data locatability	-.031	.040	-.761	.447	-.110	.049	1.846
Compatibility	.077	.043	1.781	.076	-.008	.161	1.715
Timeliness	.095*	.044	2.144	.033	.008	.181	1.537
System reliability	.134*	.038	3.498	.001	.058	.209	1.490
Easy of use/Training	.170*	.040	4.260	.000	.091	.248	1.811
Relationship with user	-.011	.056	-.204	.839	-.122	.099	2.258

Adjusted R^2 = .618*.
*: significant with p-value < .05

Table 3: Multiple linear regression for task 3

TTF factor	Coeff.	Std. err	t	Sig	95%-lb	95%-ub	VIF
(Constant)	.057	.240	.238	.812	-.415	.529	
Data quality	.194*	.085	2.292	.023	.027	.360	2.479
Data locatability	.060	.056	1.080	.281	-.050	.170	1.846
Compatibility	.176*	.059	2.971	.003	.060	.293	1.715
Timeliness	.156*	.061	2.558	.011	.036	.276	1.537
System reliability	.024	.053	.448	.654	-.080	.128	1.490
Easy of use/Training	.116*	.055	2.110	.036	.008	.225	1.811
Relationship with user	.172*	.077	2.226	.027	.020	.325	2.258

Adjusted R^2 = .399*.
*: significant with p-value < .05

4.4 Results from Bayesian regressions with uninformative priors

WinBUGS was chosen to conduct Bayesian regressions for the casual explanatory study. This program can use various MCMC techniques including Gibbs sampling, Metropolis-Hastings algorithm and adaptive rejection sampling, and chooses the correct technique automatically (Congdon, 2003). The following code snippet specifies a linear regression of y (performance) on the seven TTF factors x1 (data quality),⋯,x7 (relationship with user).

```
for (i in 1:N) {
   y[i] ~ dnorm(mu[i], tau)
      mu[i] <- a0 + a1*x1[i] + a2*x2[i] + a3*x3[i] +
 a4*x4[i] + a5*x5[i] + a6*x6[i] + a7*x7[i]
 }
```

The above program assumes a normally distributed random noise of mean 0 and variance $\sigma^2 = \tau^{-1}$. WinBUGS uses a precision variable tau ($\tau = \sigma^{-2}$) to specify the variance. Prior

distributions are needed for the random variables a0, a1,⋯, a7 and tau. Under the assumption that no particular value of these parameters is preferred, the following *uninformative* priors were used:

a0 ~ dnorm(0, 1.0E-6)

a1 ~ dnorm(0, 1.0E-6)

(Same formulae for the other regression parameters)

tau ~ dgamma(0.001, 0.001)

Since the regression parameters a0, ⋯, a7 can take any positive or negative value with equal chances, we assume that they are normally distributed with a mean of 0 and a large variance of 10^6. The precision variable tau must be positive and therefore can only assume a distribution function with positive support. A common uninformative prior for tau is given by the gamma distribution with pdf $X \sim gamma(r, \mu), p(x) = \mu^r x^{r-1} e^{-\mu x} / \Gamma(r); x > 0$.

With the above setting for Bayesian regressions and an initial value of 0 for the parameters a0,⋯, a7 and 1 for tau, plots in Figure 3 and Figure 4 confirm that the MCMC simulation has converged after 5000 iterations. At this stage, 10000 more iterations were made to collect posterior samples for inference makings. All these steps took only a few minutes for a moderate personal computer to complete. Sample mean, sample standard deviation, Monte Carlo error (MC error), 2.5%, 50% (median) and 97.5% sample cumulative points are reported in Table 4.

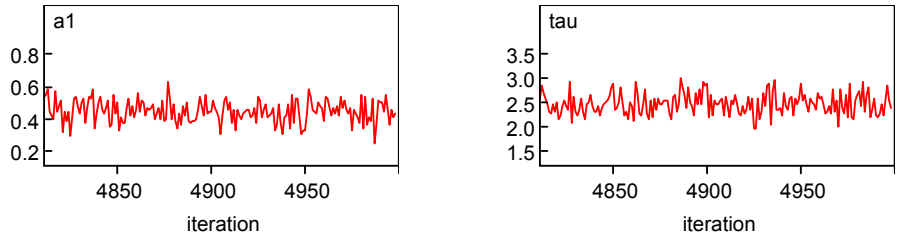

Figure 3: Trace plot of selected parameters (a1 and tau)

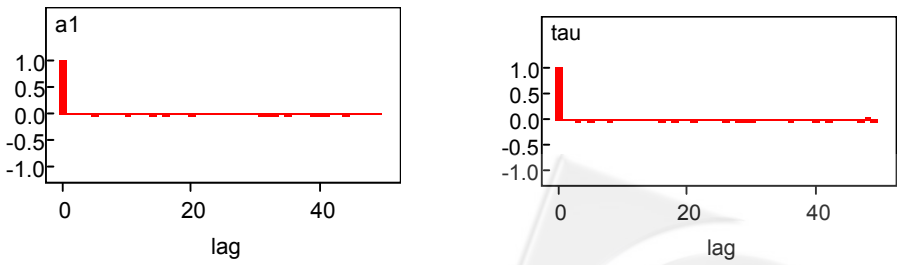

Figure 4: Autocorrelation function plot of selected parameters (a1 and tau)

Table 4: Bayesian regression with uninformative priors for task 1

TTF factor	mean	std. dev	MC error	2.5%	median	97.5%
(Constant)	.023	.223	.0023	-.413	.023	.461
Data quality	.453*	.078	.0007	.301	.453	.608
Data locatability	.063	.051	.0005	-.039	.063	.163
Compatibility	.101	.054	.0006	-.007	.101	.208
Timeliness	.206*	.056	.0004	.096	.205	.318
System reliability	.033	.049	.0005	-.063	.033	.128
Easy of use/Training	.001	.050	.0005	-.099	.001	.098
Relationship with user	.118	.072	.0006	-.020	.117	.260

*: the centralized 95% credible interval does not contain 0

How many samples should be collected after a convergence has been confirmed? In general, the more posterior samples are drawn, the more accurate the inference will be. Monte Carlo error is an estimate of the difference between a sample mean and the unknown true posterior mean. WinBUGS suggests that the simulation is run until Monte Carlo error for each parameter is less than 5% of the sample standard deviation. Based on this suggestion, it was determined that 10000 posterior samples were enough to make inferences for the parameters.

It is interesting to note that mean of each parameter in Table 4 is very close to estimated value of the same parameter in Table 1. The centralized 95% credible interval for data quality is (.301, .608). Since this interval does not contain zero, we can say that data quality is a driver (with $p < .05$ in the MLR sense). The same conclusion can be made for the timeliness factor, and the remaining five TTF factors are not significant. In this case, we observe that Bayesian regression and MLR have detected the same set of drivers for task 1.

When Bayesian regressions with uninformative priors are applied to task 2 and task 3, similar results can be observed. Table 5 summarizes results for the task 2 regression while Table 6 is for the task 3 regression. Again, sample means are close to estimated values in MLR. Drivers detected by Bayesian regressions are exactly the same as those from MLR.

Table 5: Bayesian regression with uninformative priors for task 2

TTF factor	mean	std. dev	MC error	2.5%	median	97.5%
(Constant)	.261	.175	.0018	-.082	.261	.608
Data quality	.543*	.061	.0006	.423	.542	.664
Data locatability	-.031	.040	.0004	-.111	-.031	.048
Compatibility	.076	.043	.0005	-.009	.076	.160
Timeliness	.094*	.044	.0003	.007	.094	.182
System reliability	.134*	.038	.0004	.058	.134	.029
Easy of use/Training	.170*	.040	.0004	.092	.170	.247
Relationship with user	-.012	.056	.0005	-.120	-.012	.101

*: the centralized 95% credible interval does not contain 0

Table 6: Bayesian regression with uninformative priors for task 3

TTF factor	mean	std. dev	MC error	2.5%	median	97.5%
(Constant)	.061	.242	.0025	-.413	.061	.540
Data quality	.195*	.085	.0008	.029	.194	.363
Data locatability	.060	.055	.0005	-.050	.060	.169
Compatibility	.176*	.059	.0006	.058	.175	.291
Timeliness	.155*	.061	.0005	.035	.155	.277
System reliability	.024	.053	.0006	-.081	.024	.128
Easy of use/Training	.117*	.055	.0006	.008	.117	.222
Relationship with user	.172*	.078	.0006	.022	.171	.327

*: the centralized 95% credible interval does not contain 0

5. INFORMATIVE PRIORS AND MODEL SELECTION

The Bayesian approach shines when informative priors are available. In this section, we propose a heuristic method based on the outputs of MLR to construct informative priors. This hybrid approach to Bayesian regressions improves research model significantly and makes casual explanation more interesting.

5.1 A heuristic method for constructing informative priors

Prior selection is a major issue in Bayesian regressions. When credible evidences are available, there is no reason to ignore them in the construction of prior distributions. MLR is often considered objective in regression analysis since its main goal is to minimize the sum of squared errors. We propose to utilize the outputs of MLR as credible evidences to construct prior distributions as follows.

With the OLS estimator, MLR outputs estimated values and standard errors of regression parameters. In Table 1, the intercept parameter a0 has an estimated value of .019 and a standard error of .221, which translates into a precision of 1/(.221*.221) = 20.4746. Therefore, instead of an unbiased mean of 0 and a very large variance of 106, we can set up a more precise prior distribution for this parameter as a0~dnorm(.019, 20.4746). We feel that this approach is appropriate since MLR has made a good estimate for the mean and variance of a0.

The same procedure is carried out for each of the remaining parameters a1, ⋯, a7. With these heuristic priors, MCMC simulation was conducted for task 1 as before. The simulation converged to a satisfactory stage after 5000 iterations. Results based on the 10000 posterior samples after convergence are reported in Table 7. A few observations can be made: (1) Standard deviations have been reduced substantially in the new model while MC errors are still under control. Therefore the new model is more efficient in parameter estimation than the old model; and (2) Two more factors, namely compatibility and relationship with user, become

significant in the causal model. Relationship with user belongs to the domain of responding to changed business environments. Users of the Mobile Dr. Insurance system consider recruiting new contracts a challenging task that requires them to be prepared for changed business environments. The inclusion of this driver is welcome since the highly competitive insurance industry in Taiwan has a fast changing business environment.

Table 7: Bayesian regression with heuristic informative priors for task 1

TTF factor	mean	std. dev	MC error	2.5%	median	97.5%
(Constant)	.022	.151	.0016	-.275	.021	.317
Data quality	.453*	.051	.0005	.353	.453	.554
Data locatability	.063	.035	.0003	-.006	.063	.131
Compatibility	.101*	.037	.0004	.029	.102	.175
Timeliness	.206*	.038	.0003	.132	.205	.281
System reliability	.033	.033	.0004	-.034	.033	.098
Easy of use/Training	.001	.034	.0004	-.067	.001	.067
Relationship with user	.118*	.048	.0004	.027	.118	.213

*: the centralized 95% credible interval does not contain 0

5.2 Model comparison and selection

How do we compare an MLR model and a Bayesian model with or without informative priors? The coefficient of determination $R^2 = 1 - \frac{SSE}{SST}$, where *SSE* denotes the sum of squared errors and *SST* denotes the total variation of the dependent variable, is often used to measure the goodness-of-fit of an MLR model. This is an in-sample criterion for model assessment. In order to account for model parsimony, an adjusted $R^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)}$, where n is the number of observations and k is the number of independent variables, can be used. On the other hand, the Bayesian approach uses *DIC* of Eq. (8) to measure the fit of a Bayesian model. Since MLR yields only a set of regression parameters, *DIC* is not applicable to MLR. Nevertheless, the adjusted R^2 can always be computed as long as a set of regression parameters is available. It is fair to use sample means of a Bayesian model to compute the adjusted R^2 of the Bayesian model. In the following, we will use the adjusted R^2 to compare an MLR model with a Bayesian model, and *DIC* to compare two Bayesian models.

Using adjusted R^2, the MLR model, the Bayesian model with uninformative priors (BU) and the Bayesian model with informative priors (BI) for task 1 regression virtually have the same level of goodness-of-fit. This outcome is expected since sample means of the Bayesian models are so close to their corresponding parameter values in MLR. On the other hand, *DIC* for BU is 626.239 and *DIC* for BI is 618.831. Because the difference between these two *DIC* values is plausible (> 5), heuristic informative priors improve Bayesian regressions in model assessment significantly (Spiegelhalter et al., 2003).

Similar Bayesian regressions with heuristic informative priors were conducted for task 2 (Table 8) and task 3 (Table 9) regressions. The task 2 regression has a DIC of 479.397 and 471.956 respectively for uninformative and informative priors. The task 3 regression has a *DIC* of 677.723 and 670.298 respectively for uninformative and informative priors. The difference between two *DIC* values is about 7.4 in either case, thus BI is preferred over BU. In terms of the adjusted R^2, all three models (MLR, BU and BI) have the same level of goodness-of-fit for either task 2 or task 3 regression. We decide to choose BI as the final model for a causal explanatory study because it improves BU significantly with a plausible reduction of *DIC* and is as good as MLR and BU with the same adjusted R^2.

Table 8: Bayesian regression with heuristic informative priors for task 2

TTF factor	mean	Std. dev	MC error	2.5%	median	97.5%
(Constant)	.260	.119	.0012	.026	.259	.492
Data quality	.542*	.040	.0004	.464	.542	.622
Data locatability	-.031	.027	.0003	-.085	-.030	.023
Compatibility	.076*	.029	.0003	.019	.076	.134
Timeliness	.094*	.030	.0002	.036	.094	.153
System reliability	.134*	.026	.0003	.082	.134	.185
Easy of use/Training	.170*	.027	.0003	.117	.170	.223
Relationship with user	-.012	.038	.0003	-.083	-.012	.063

*: the centralized 95% credible interval does not contain 0

Table 9: Bayesian regression with heuristic informative priors for task 3

TTF factor	Mean	std. dev	MC error	2.5%	median	97.5%
(Constant)	.060	.164	.0017	-.263	.059	.381
Data quality	.195*	.056	.0005	.087	.195	.304
Data locatability	.060	.038	.0004	-.015	.061	.135
Compatibility	.176*	.040	.0004	.097	.176	.255
Timeliness	.155*	.041	.0003	.075	.155	.237
System reliability	.024	.036	.0004	-.048	.024	.095
Easy of use/Training	.116*	.037	.0004	.043	.117	.189
Relationship with user	.172*	.052	.0004	.073	.172	.275

*: the centralized 95% credible interval does not contain 0

5.3 Drivers extraction

Using .05 as a cut-off level, we now extract drivers from the three regression models. For MLR, the criterion for a driver is its 95% *confidence* interval does not include zero, and for BU or BI this criterion becomes the centralized 95% *credible* interval does not include zero. After comparing various tables shown above, we conclude that BU and MLR have the same set of

drivers, which is subsumed in the set of drivers for BI. Table 10 summarizes this observation.

Thus, MLR and BU have the same power of detecting drivers in this empirical study. BI can usually find one or more interesting drivers than the other two models. After a careful examination of the results, we find that TTF factors with a *p*-value up to .1 in MLR can often become significant in BI. For example, relationship with user of task 1 has a *p*-value of .098 in MLR. After using heuristics from MLR, BI has claimed this factor a driver. It seems that BI is delicate in detecting border line drivers. This is probably due to the fact that BI starts with a more precise estimate of the regression parameters.

Table 10: Drivers of performance for 3 insurance tasks using MLR, BU and BI

TTF factor	Task 1			Task 2			Task 3		
	MLR	BU	BI	MLR	BU	BI	MLR	BU	BI
Data quality	X	X	X	X	X	X	X	X	X
Data locatability									
Compatibility			X			X	X	X	X
Timeliness	X	X	X	X	X	X	X	X	X
System reliability				X	X	X			
Easy of use/Training				X	X	X	X	X	X
Relationship with user			X				X	X	X

X: factor is significant

5.4 Perturbing mean and variance in the heuristic priors

In the above study, we adopted estimated parameter values and standard errors from MLR to construct informative priors. It seems that these heuristics work pretty well: the Bayesian model is improved significantly and more interesting drivers can be detected. What happens if we use other means and variances in the priors?

With a few more experiments, we find that: (1) if means and variances are set very close to the proposed heuristics, results about sample means and drivers for Bayesian regressions are very similar to those obtained in BI. However, *DIC* is increased a little bit; (2) the same statement can be said if means and variances are set very close to those in BU; and (3) if means and variances are set randomly, especially when they deviate a lot from the previous two situations, MCMC simulations may converge to models with a much higher *DIC*. With a wrong choice of means and variances in the priors, a Bayesian model can be substantially worse than BU, and thus should be rejected. Are there better choices of means and variances? This is potentially a difficult search problem that warrants advanced search algorithms.

5.5 More evidence - A second empirical study

In order to check feasibility of the proposed method, we applied the heuristic Bayesian modeling approach to a second case of IS impact study, where data was collected independent of the first empirical study. The subject company in this case, termed company N, is an international insurance company in Taiwan and belongs to the Asia Pacific group of an international conglomerate headquartered in the United States. Company N offers many insurance products including individual life insurance and accidental insurance. To cope with the fast moving pace of the insurance industry, company N has started an early planning of adopting mobile technology in October 2000. The first mobile network named PDA mobile commerce system based on Palm OS and Windows CE was implemented in February 2001. The research model of this case is very similar to that of company G except the authorization factor is now a predictor variable.

Company N had 18 branches and over 300 field offices in Taiwan. A random sampling of the company's agents was made to participate in the survey study. We sent out 450 questionnaires and received 274. Excluding incomplete and inconsistent questionnaires, there were 238 final useful samples. Data was processed as before by following established procedures in Nunnally (1978). It was found that Cronbach's α of our instrument ranged from .6347 to .9412, and each VIF in MLR was smaller than 10. For MLR, the adjusted R^2 is .306, .430 and .373 for task 1, task 2 and task 3 regressions respectively. Since sample means in BU and BI are very close to estimated parameter values in MLR, all three models have the same adjusted R^2 in each regression. With a cut-off level of .05, drivers located by MLR, BU and BI are marked in Table 11. Drivers detected by MLR and BU are exactly the same, and they are subsumed by drivers in BI. In addition, we compute *DIC* for BU and BI with results listed in Table 12. A plausible (> 5) reduction of *DIC* from BU to BI has been detected. In summary, observation from this second empirical study is pretty much the same as that from the first empirical study.

BI has again turned border line factors into drivers. A careful examination of the results shows that TTF factors with a *p*–value smaller than .15 in MLR become significant in BI. For example, authorization of the task 2 regression has an estimated value of .106 with $p = .144$ in MLR; using heuristics from MLR, BI has obtained a centralized 95% credible interval of (.0097, .2006). Thus, the authorization factor is a driver in BI, but insignificant in MLR. This delicate feature of BI can sometimes introduce surprise drivers in a causal explanatory study. In the task 3 regression, authorization has an estimated value of -.153 with p = .114. Since it is insignificant in MLR, most studies will simply discard its impact on performance. This parameter has a centralized 95% credible interval of (-.2825, -.02583) and a sample mean of -.1543 in BI. Therefore, the authorization factor is a driver for performing the task of providing tax and legal information services. This poses an interesting question since authorization *negatively* affects the task performance. Goodhue and Thompson (1995) found a similar situation in their study:

compatibility significantly and negatively impacts the task performance. After they added utilization as another independent variable in the research model, this phenomenon disappeared. The researchers concluded that when the technology is utilized and TTF evaluation is high, the performance will be improved. Since the task of providing tax and legal information services is generally considered a non-essential task by many insurance agents, we suspect that utilization was not assured for this task in company N. Previous studies have shown that voluntariness exists on a continuum (Moore and Benbasat, 1993), thus utilization may play an important role in the IS impact study of task 3. Company N can take a qualitative study to further examine this phenomenon.

Table 11: Drivers of performance for 3 insurance tasks in the second empirical study

TTF factor	Task 1			Task 2			Task 3		
	MLR	BU	BI	MLR	BU	BI	MLR	BU	BI
Data quality	X	X	X	X	X	X	X	X	X
Data locatability			X						
Authorization						X			X
Compatibility							X	X	X
Timeliness									
System reliability									
Easy of use/Training									
Relationship with user						X	X	X	X

X: factor is significant.

Table 12: DIC for Bayesian regressions in the second empirical study

	Task 1	Task 2	Task 3
BU	471.672	383.808	524.074
BI	463.271	375.407	515.689
BU - BI	8.401	8.401	8.385

6. CONCLUSIONS

In this research, a Bayesian regression is introduced for a causal explanatory study. Frequently, we are interested in finding what causes have significantly influenced a dependent variable. Drivers located in such a causal explanatory study can often play an important role in management decisions. Previous studies in empirical research often use MLR to conduct a regression analysis. MLR is fast and simple, but it lacks the flexibility to incorporate our prior belief into the building of a regression model. Results from MLR are also hard to interpret. On the other hand, Bayesian regressions have the flexibility to accommodate our prior belief,

and results from Bayesian regressions are intuitive. Even with these advantages, Bayesian regressions are still not popular in empirical research possibly due to two reasons: (1) some researchers view the selection of priors as a subjective work of arts; and (2) the computing resources associated with MCMC simulations are intensive. In this study, we answer the first issue by proposing a heuristic method to construct informative priors that are objective. The second issue is answered by today's fast computing machines and efficient algorithms.

In order to demonstrate the Bayesian approach, two survey based technology to performance impact studies were conducted in different periods. Our research model is based on the task technology fit theory, a well-known theory in IS research that relates TTF degrees to user performance. In each case, three regression models (MLR, BU and BI) have been obtained to explain the impact of TTF on performance. The findings are the same in both empirical studies: (1) sample means from BU and BI are very close to estimated parameter values in MLR, and the adjusted R2 is virtually the same for all three models; (2) BI significantly improves BU and is the recommended final model in this research; and (3) BU and MLR detect the same set of drivers, while BI can usually detect one or more drivers. BI finds additional drivers by turning border line factors into drivers.

Based on the two empirical studies, we find that the inclusion of extra drivers in BI is progressive, not radical. The proposed BI approach does not change the way that a research model is established; a research model is still founded on a careful and thorough literature review. The BI approach does not change the means of data collection either. It just changes the way that data is analyzed by shifting from frequentist statistics to Bayesian statistics.

6.1 Implications for IS impact study

Lucas (1975) suggested that utilization is an appropriate surrogate for IS success when use is voluntary, and user evaluations of IS are appropriate when use is mandatory. Goodhue and Thompson (1995) generalized this idea to include the fit between task and technology. The researchers recommended to use TTF evaluations as a surrogate when utilization is assured, otherwise TTF alone is an incomplete surrogate for IS success. The delicacy of BI in detecting drivers can potentially provide valuable supports to the TTF theory in specific or IS impact study in general. In our empirical studies, BI found more drivers than MLR and BU. Actual managerial implications of this capability have to be examined case by case. For example, in our first empirical study, BI has claimed compatibility as a driver for all three tasks of insurance agents. Therefore, company G has to pay special attention to the overall database design in the Mobile Dr. Insurance system. Also, BI has reminded company G that users of its system are concerned with the fast changing business environments, because relationship with user is a driver for the task of recruiting new contracts. Thus, company G may want to understand the insurance industry in depth by learning from insurance consumers. In the second empirical

study, the negative impact of authorization on performance should prompt company N to study whether utilization of IS for the task of providing tax and legal services is assured. In this way, BI and TTF together can help design better diagnostics for IS problems.

6.2 Implications for Bayesian study

To the best of our knowledge, the novel application of the outputs from MLR to construct informative priors has not been used in any previous studies. Based on our empirical studies, it seems that these heuristics can yield a model that is significantly better than BU. Theoretically it would be interesting to know whether the proposed heuristic approach always leads to a significantly better model. Our empirical studies seem to confirm this, but a rigorous proof needs substantial knowledge and technique in Bayesian statistics. The result may also depend on a correct type specification for the prior distributions. In this study, selection of the (normal and gamma) distribution types is subjective. Overall, the Bayesian approach allows a user to quickly change distribution types of the priors, thus future studies may focus on how to smartly choose the distribution types in order to construct a much better Bayesian model.

6.3 Limitation of the research

The two empirical studies in this research were conducted for the insurance industry in Taiwan. Unlike the broad spectrum of tasks and technologies involved in Goodhue and Thompson (1995), tasks and technologies were clearly defined in our study. This may explain why our adjusted R2 is substantially higher than theirs. Direct generalizations of the managerial implications to other industries or countries may not be appropriate. However, we believe that the proposed heuristic Bayesian regression approach can be easily applied to other empirical studies and findings should be similar to the ones presented in this study.

REFERENCES

1. Banerjee, S., Kauffman, R.J. and Wang, B. "A Dynamic Bayesian Analysis of the Drivers of International Firm Survival," *In Proceedings of International Conference on Electronic Commerce*, Xi'an, China: ACM, 2005, pp. 151-158.
2. Bansal, A., Kauffman, R.J. and Weitz, R.R. "Comparing the Modeling Performance of Regression and Neural Networks as Data Quality Varies: A Business Value Approach," *Journal of Management Information Systems* (10:1), 1993, pp. 11-32.
3. Burton, P.R., Gurrin, L.C. and Campbell, M.J. "Clinical Significance Not Statistical Significance: A Simple Bayesian Alternative to P Values," *Journal of Epidemiology and Community Health* (52), 1998, pp. 318-323.

4. Cao, J. Crews, J.M., Lin, M., Deokar, A., Burgoon, J.K. and Nunamaker, J.F. Jr. "Interactions between System Evaluation and Theory Testing: a Demonstration of the Power of a Multifaceted Approach to Information Systems Research," *Journal of Management Information Systems* (22:4), 2006, pp. 207-235.
5. Congdon, P. *Applied Bayesian Modeling*, John Wiley & Sons, New York, 2003.
6. Cooper, D.R. and Schindler, P.S. *Business Research Methods, tenth edition*, McGraw Hill, New York, 2008.
7. Davis, F.D. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), 1989, pp. 319-342.
8. Denison, D.G., Holmes, C.C., Mallick, B.K. and Smith, A.F. *Bayesian Methods for Nonlinear Classification and Regression*, John Wiley & Sons, New York, 2002.
9. Devore, J.L. *Probability and Statistics for Engineering and the Sciences, sixth edition*, Thomson Brooks/Cole, Belmont, CA, 2004.
10. Diamantopoulos, A. and Siguaw, J.A. *Introducing Lisrel: a Guide for the Uninitiated*, Sage Publications Ltd, London, 2000.
11. Dickson, G.W., DeSanctis, G. and McBride, D.J. "Understanding the Effectiveness of Computer Graphics for Decision Support: a Cumulative Experimental Approach," *Communications of the ACM* (29:1), 1986, pp. 40-47.
12. Gallizo, J.L., Jimenez, F. and Salvador, M. "Adjusting Financial Ratios: a Bayesian Analysis of the Spanish Manufacturing Sector," *OMEGA* (30), 2002, pp. 185-195.
13. Garthwaite, P.H., Jolliffe, I.T. and Jones, B. *Statistical Inference, second edition*, Oxford University Press, New York, 2002.
14. Goodhue, D.L. "Understanding User Evaluations of Information Systems," *Management Science* (41:12), 1995, pp. 1827-1844.
15. Goodhue, D.L. and Thompson, R.L. "Task-technology Fit and Individual Performance," *MIS Quarterly* (19:2), 1995, pp. 213-236.
16. Greenland, S. "Bayesian Perspectives for Epidemiological Research. II. Regression analysis," *International Journal of Epidemiology(36:1)*, advance access published February 28, 2007, pp.195-202.
17. Hogg, V.H. and Tannis, E.A. *Probability and Statistical Inference, fifth edition*, Prentice-Hall, Upper Saddle River, N.J., 1997.
18. Kennedy, P. *A Guide to Econometrics, fifth edition*, MIT Press. Cambridge, MA, 2003.
19. Koop, G. *Bayesian Econometrics*, John Wiley & Sons, New York, 2003.
20. Lucas, H. "Performance and Use of an Information System," *Management science* (21:8), 1975, pp. 908-919.
21. Moore, G.C., and Benbasat, I. "An Empirical Examination of a Model of the Factors Affecting Utilization of Information Technology by End-users," *Working paper*, University of British Columbia, Faculty of Commerce, 1993.

22. Nunnally, J. *Psychometric Theory, second edition*, McGraw Hill, New York, 1978.
23. Schköpf, B. and Smola, A. *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*, MIT Press, Cambridge, MA, 2002.
24. Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. "Bayesian Measures of Model Complexity and Fit (with discussion)," *Journal of the Royal Statistical Society: Series B* (64:4), 2002, pp. 583-640.
25. Spiegelhalter, D.J., Thomas, A., Best, N.G. and Lunn, D. "WinBUGS - an Interactive Windows Version of the BUGS Program for Bayesian Analysis of Complex Statistical Models Using Markov Chain Monte Carlo (MCMC) Techniques," 2003 (available online at http://www.mrc-bsu.cam.ac.uk/bugs).
26. Thorburn, D. "Significance Testing, Interval Estimation or Bayesian Inference: Comments to "Extracting a Maximum of Useful Information from Statistical Research Data" by S. Sohlberg and G. Andersson," *Scandinavian Journal of Psychology* (46:1), 2005, pp. 79-82.
27. Urbach, P. "Regression Analysis: Classical and Bayesian," *The British Journal for the Philosophy of Science* (43:3) 1992, pp. 311-342.
28. Vapnik, V. *Statistical Learning Theory*, John Wiley & Sons, New York, 1998.
29. Vessey, I. "Cognitive Fit: a Theory-based Analysis of the Graphics vs. Tables Literature," *Decision Sciences* (22:2), 1991, pp. 219-240.
30. Witten, I.H. and Frank, E. *Data Mining, Practical Machine Learning Tools and Techniques*, Morgan Kauffman publishers, San Francisco, 2005.
31. Wright, J.H. "Bayesian Model Averaging and Exchange Rate Forecasts," *Boards of Governors of the Federal Reserve System*, International Finance Discussion, 2003, pp. 779.