The Determinants of Technology Adoption by UK Farmers using Bayesian Model Averaging. The Cases of Organic Production and Computer Usage.

Balcombe, Kelvin and Tiffin, R

University of Reading

2010

The Determinants of Technology Adoption by UK Farmers using Bayesian Model Averaging. The Cases of Organic Production and Computer Usage.

R Tiffin and K Balcombe

June 3, 2010

Abstract

We introduce and implement a reversible jump approach to Bayesian Model Averaging for the Probit model with uncertain regressors. This approach provides a direct estimate of the probability that a variable should be included in the model. Two applications are investigated. The first is the adoption of organic systems in UK farming, and the second is the influence of farm and farmer characteristics on the use of a computer on the farm. While there is a correspondence between the conclusions we would obtain with and without model averaging results, we find important differences, particularly in smaller samples.

1. Introduction

The evaluation of the importance of explanatory variables based on measures such as statistical significance, adjusted R-square or information criteria is a common practice. The literature is replete with discussions where the importance of regressors are evaluated ostensibly by their statistical significance or lack thereof. Commonly, investigators also exclude variables from their models because their associated coefficients are not statistically significant or the exclusion of that variable leads to improvements in a given criteria, and only those that are left in the model are deemed important.

Faced with a large number of potential explanatory variables and limited sample sizes, there is often a need to choose a subset of the available regressors. This is a requirement when using time series or cross sectional data and needs to be addressed whether the aim is to estimate a purely predictive model, or to estimate a model with structural interpretations. In some cases variable selection is automated, and in others investigators choose to adopt a more 'hands on' approach.

The pitfalls of automated model selection of the 'stepwise' kind are well documented (Millar, 1984). The 'hands on' approach is open to the criticism that it is insufficiently objective. Problems exist regardless of whether the modelling strategy is 'specific to general', 'general to specific' or a mixture of both. While model selection has been the subject of extensive discussion going back many years (e.g. Pagan, 1987) adequate solutions arguably remain absent from classical perspective. Recent classical work automating model selection from a 'general to specific' point of view notwithstanding (e.g. Hendry and Krolzig; 2005), sequential reduction algorithms inevitably involve testing sequences that can pre-determine the final selection. Though a search can sometimes be conducted over the entire model space (e.g. Balcombe et al. 2005), investigators still face a difficult choice between competing ways of evaluating one model (as defined by the set of regressors included in the model) relative to many others, particularly if there are many models that perform similarly according to a given criteria.

Even if the aim is not to conduct a specification search, investigators often seek to establish whether a given variable belongs in the model on the basis of statistical significance. Thus, a model is estimated and the resulting discussion is then about which variables are significant and which are not. When conducting this type of analysis, investigators often indirectly infer that a small p-value implies a high probability of a non-zero coefficient. This makes sense to the extent that a p-value of below 0.05 implies that a 95% confidence interval for the parameter would exclude zero. Thus, in this limited sense, the investigator is at least 95% certain that the variable should be in the model. Unfortunately, as simple as this logic seems, it cannot be used to construct a formal argument that p-value below 0.05 means that there is at least 95% probability that the variable should be included the model. While investigators may know that a p-value cannot strictly be interpreted in this way, in practice, statistical significance is used as an indicator of the probability that regressor has an impact on the dependent variable. The gulf between the formal meaning of a p-value and how it is used in practice is not a mere curiosity. It may have a substantive impact on interpretation and findings.

An appealing alternative to estimating one very large model or searching for a better performing submodel, is to take an average over many models. While classical statistics struggles to give any formal basis for averaging over models[1], a Bayesian approach provides both a theoretical underpinning, along with clear methodology for implementing model averaging. Final estimates can be obtained by taking a weighted average of estimates over models, where a model that is highly supported by the data will be given a higher weight than one which is less supported by the data. Importantly, this approach can further deliver a measure of the probability that a given variable enters the model.

The construction of the weights used for Bayesian model averaging (BMA) is performed using the Bayesian 'marginal likelihood' (ML). Unlike the likelihood function, the ML is not defined on parameters since these have been integrated out of the expression.[2] Thus, it is a function of the data and the model but not the parameters. Where two models are thought equally likely, a priori, the 'posterior odds' for two models is equal to the ratio of their MLs (also known as the Bayes factor or ratio). Therefore, the ML can be used to give a weight to a given model with BMA.

[1] Sali-a-Martin et al. (2004) develop a hybrid approach to model averaging, but this could not be strictly labelled Bayesian or Classical.

[2] The likelihood can be viewed as the probability of the data conditionally on the parameters. The marginal likelihood can be viewed as the expected likelihood, given the model and model priors.

BMA can be difficult to implement because the ML is often hard to compute. Under these circumstances, BMA can practically take place only over a small number of models. For the standard linear regression model, the ML can be expressed analytically and computed quickly. Even so, where the set of models is defined by all the combinations of regressors that can enter the regression, the model space can be massive.[3] Therefore, it can still be impractical to estimate every model and assign a weight to each model. Bayesian computation can solve this problem by employing an algorithm where only a relatively small subset of the models require estimation. The ML can provide a basis for choosing models as part of the algorithm which 'jumps' between one model and another. This class of algorithms are an extension of the Monte Carlo Markov Chain (MCMC) algorithms that are employed to estimate many Bayesian models.

This paper uses and explains a Bayesian reversible jump (RJ) procedure for Probit model selection in which the probability of regressors entering the model is estimated along with the parameters that enter the Probit equation. The estimates obtained from this procedure include the probability that a variable enters the model along with model averaged estimates (and standard deviations) of the Probit parameters. The RJ method is an approach to model averaging, that can be applied to the selection of models where the number of potential models is very large. The general RJ approach to estimation of models was developed by Green (1995) and a general approach to the estimation of limited dependent variable data was outlined in Holmes and Held (2005). The RJ approach was applied to a linear time series model by Balcombe and Rapsomanikis (2010), but so far there have been no applications of the reversible jump procedure for a Probit model within the Agricultural Economics literature. Applications of the RJ Probit within the Economics literature are few. An exception is Leon-Gonzalez and Scarpa (2007) which applied this algorithm in a contingent valuation setting.

It is not our aim to compare Classical with Bayesian methods, but to compare results with and without model averaging. In introducing the RJ approach, we are mindful that agricultural economists will be less concerned with theoretical arguments for BMA, but more concerned with its practicality, and how it may change the inferences obtained from a given data set. BMA is certainly practical. The

[3] A model with k variables has 2^k submodels

models in this paper take less than half an hour to estimate on a modern computer, even though the data sets would be regarded as relatively large. Moreover, there is no doubt that BMA can sometimes have a substantive influence on inferences draw from a given set of data. For example, Balcombe and Rapsomanikis (2010) show that in the context of a time series model, the use of BMA lead to quite different conclusions. Results from a Bayesian analysis (without model averaging) are often very similar to a Classical analysis, albeit with some differences in the way that the results are presented and interpreted. Non informative priors and large sample sizes, Bayesian and Classical approaches often lead to comparable point estimates and confidence intervals (see Mittelhammer, et al. 2000, pp. 661-666). Therefore, we use Bayesian methods throughout this paper, since estimates produced for a standard Probit model using a Bayesian or Classical methods do not differ substantively.

The RJ approach to estimation model is applied to the analysis of two data sets and the results are compared to the results obtained without model averaging. First, the adoption of organic production in agriculture is analysed. This data set has already been discussed and analysed by Rigby et al. (1999). Second, we apply the Probit model to an original data set on the determinants of computer adoption in UK agriculture. Agricultural producers have lagged behind other businesses in computer ownership and use. Despite the rapid adoption of computer technology by British farmers in recent years, there has been little in the way of formal econometric analyses about why farmers purchase computers, what they use them for, and whether computers are making a positive impact on farm profitability.

The paper proceeds by discussing the estimation of the Probit using the RJ method in Section 2. Section 3 introduces the data and presents and discusses the empirical results. Section 4 concludes. Mathematical details are contained in an appendix.

2. Model and Estimation

A common Bayesian approach to estimation is to simulate the posterior distribution for the parameters of a model using Monte Carlo Markov Chain (MCMC) algorithms (e.g. Chib and Greenberg, 1995). The RJ approach is an extension of the MCMC algorithms. The difference is in that, when using the RJ approach, the model is also drawn from its posterior distribution, not just the parameters.

Using the notation $f(x|y)$ to denote the conditional distribution of x given y, MCMC algorithms operate by drawing from $f(x|y)$ then using the draw of x to draw y from $f(y|x)$. Subject to certain conditions, this leads (provided the sequence is repeated many times) to draws from $f(x,y)$. Within the standard MCMC approach, the quantities y and x can represent parameters or latent data. With the RJ approach, they can also represent models. In this section we describe the model and estimation procedure in more detail.

2.1. The Model

The model employed within this paper is of a standard binomial Probit form:

$$\begin{aligned} y &= xb + e \qquad (1) \\ y &= (y_1, ...y_T)' ;\ e = (e_1, ...e_T)' ;\ \text{and,} \\ x &= (x_1, ...x_T)' \text{ where } x_i = (1, x_{i,1}, x_{i,2}...x_{ik})' . \end{aligned}$$

It is assumed that $e_i \overset{iid}{\sim} N(0,1)$. The restriction on the variance of e_i is the usual identifying assumption for the Probit model. The data y_i is not observed for the Probit model. Instead, we observe the indicator variable $d = (d_1, ...d_n)'$where $d_i = 1$ where $y_i > 0$ and $d_i = 0$ otherwise. The Bayesian approach to estimation, requires a prior distribution for b. Where this is specified as $f(b) = N\left(\beta_0, M_0^{-1}\right)$, then:

$$f(b|y) = N\left(\beta_2, M_2^{-1}\right) \text{ and } y_i = TN^{+}(xb,1)\,\mathbf{1}(d_i = 1) + \ TN^{-}(xb,1)\,\mathbf{1}(d_i = 0) \qquad (2)$$

where $M_1 = x'x$; $M_2 = M_0 + M_1$; $\beta_1 = M_1^{-1}x'y$; $\beta_2 = M_2^{-1}(M_0\beta_0 + M_1\beta_1)$; and, $TN^{+}(TN^{-})$ denotes a positively (negatively) truncated distributed normal distribution and $\mathbf{1}(.)$ denotes and indicator function.

2.2. Estimation

Where the regressors are known, estimation can proceed through simulation by drawing b from $f(b|y)$ then y then $f(y|b,d)$ and so on, recording the draws of b so as to simulate the marginal posterior distribution. The reversible jump algorithm only involves a further step by augmenting the sequence by drawing from $f(m|y,d)$ where m denotes the model (the choice of regressors). The last step is achieved by proposing a new model m^* in a 'symmetric fashion'[4], then accepting this new model (rather than the old model m) with probability

$$p = \min \left(\frac{\frac{|M_2(m^*)|^{-\frac{1}{2}}}{|M_0(m^*)|^{-\frac{1}{2}}} \times e^{-\frac{J_{m^*}}{2}}}{\frac{|M_2(m)|^{-\frac{1}{2}}}{|M_0(m)|^{-\frac{1}{2}}} \times e^{-\frac{J_m}{2}}}, 1 \right) \tag{3}$$

where

$$J_m = \left(y - x_m \beta'_{0,m}\right)\left(y - x_m \beta_{0,m}\right) - \left(y - x_m \beta_{0,m}\right)' x_m M(m)_2^{-1} x'_m \left(y - x_m \beta_{0,m}\right) \tag{4}$$

and x_m are the regressors for model m, and $\beta_{0,m}$ $M(m)_0$ are the priors for the parameters under model m, and $M(m)_2 = (M_{0,m} + x'_m x_m)$. Derivations of equations (3) and (4) are left for an appendix (see section A2.2). The validity of the 'model step' above follows from the fact that the conditional distribution of the model $f(m|y,d)$ simplifies to $f(m|y)$ since any admissible set of latent data (y), is sufficient to deduce the observed data (d). Therefore, the model step within the RJ algorithm for the Probit model is almost identical to the model step within the normal linear model, except that the variance is set to one.

The Priors adopted in this survey are the 'G-Priors' with $\beta'_{0,m} = 0$ for all models. Using this construction, the priors are $M_{0,m} = \frac{x'_m x_m}{n}$. For the rationale behind the use of these priors, readers are referred to the discussion and further references within Fernandez et al. (2001). Within our analysis the priors over all models are uniform (each model is, a priori, equally likely as another). In principle, informative priors could be placed over the model space if some variables were thought more likely to determine adoption than others. However, we prefer to use non-informative priors over the model space.

[4]This means that the probability of proposing move from m to m^* is equal to the probability of proposing a move from m* to m.

3. Empirical Section

Our analysis within this section will examine two different data sets: organic technology adoption; and, computer adoption in agriculture. As discussed above, our analysis throughout will be Bayesian. Classical point estimates and confidence intervals for the data sets in this paper are similar to the results we present below without using BMA. Although it is not entirely consistent with the Bayesian literature, we denote significance at the 5% level if the coefficient has a 95% Bayesian confidence interval (also known a high density region) that excludes zero. Bayesian significance has a slightly different interpretation to that of classical significance[5]. However, as already noted, the Bayesian confidence intervals presented herein are similar to those obtained by a classical analysis, thus, the exclusion of zero from the Bayesian confidence interval would also indicate significance in the classical sense. Therefore, we continue to label the parameter 'significant' if its confidence interval excludes zero.

3.1. Organic Adoption

3.1.1. Data on Organics Adoption

The Organics data is composed of 237 horticultural producers from the UK, of which 151 were conventional producers and 86 had adopted organic technologies. The survey was conducted in 1996. The discussion of the sources and summary of this data is discussed in Rigby et al (1999). The data, with descriptions can also be found on the ESRC data archive.

3.1.2. Results for Organics Adoption

Rigby et al. (1999) run a Logit regression of the decision to adopt organics on a set of explanatory variables. These are listed in Table 2 of Rigby et al. (1999). Since our aim is to compare the model averaged results with those obtained by Rigby et al. (1999), we use exactly the same variables used in their analysis. We produce the Probit results for their model in Table 1 below, on the left hand side. The variables that are significant at the 5% level, are superscripted by a star. In terms of significance, these are broadly the same as those reported in Rigby et al.

[5] e.g. see Mittelhammer, 2000, chap 24.

The BMA results are reported on the right hand side of Table 1. The last column gives the probabilities that the relevant variables are included in the model. The estimates and the standard deviations are the mean of the posterior distributions for both the standard and BMA results. Apart from the intercept, we can see that 4 variables: *conin, orgff, infpss,* and *infbuy* (*conin=1* if farmer believes that current practices will sustain farm productivity, *orgff=1* if farmer believes that farming alone can satisfy societies needs for food and fibre, *infpss (infbuy)= 1* of main source of information is press (merchants)) are included with probability 0.99 or above, closely followed by *infadas*, which is included with probability of 0.972. Notably, all these are also significant at the 5% level. More generally, across most of the variables, there is a correspondence between the probability of being included in the model, and the significance of the associated coefficient. Generally, the more significant a variable (the further away the interval is from zero) the more likely that the variable is included in the model. This said, we would revise the importance of variables in the light of the probabilities in the last column. First, it would be inaccurate to conclude that a variable that is significant at the 5% level should be in the model with 95% probability. Several variables such as *hhsize*, *fem* and *inffmrs (hhsize*=household size, *fem*=female (1 for female, 0 for male) , *inffmrs*=1 if main source of information is other farmers) are significant at the 5% level, but are included in the model less than 80% of the time. Even more notable is the variable *age* which is significant at the 5% level. However, it is only included in the model around 47% of the time. Likewise, *memenv* (=1 member of environmental organisation) is significant at the 5% level, but is only included in around 55% of the time. By contrast the variable *maxcon (*=1 if maximiser of consumption of own production) was insignificant at the 5% level, but as it was included in 61% of the time, it has a higher probability of being included than either of the significant variables *age* and *memenv.* Finally, we note that the model averaged results differ substantively from the standard results. For variables that are not included with a probability close to 1, the estimates are substantially lower in absolute value, reflecting the high probability that they are zero.

Finally, with respect to the estimates, the coefficients of the model averaged results are generally smaller (in absolute value) than for the standard results, but in nearly all cases retain their original signs. Those variables that enter the models a relatively small proportion of the time, have correspondingly smaller values in absolute terms. In this sense the BMA results represent the mid ground between a model selection strategy in which only the most general model is estimated and one where insignificant variables are eliminated from the model. The strategy of excluding insignificant variables from the model is an extreme one, and arguably does not truly reflect the nature of our uncertainty about the role of the variable. Thus, the BMA results represent a more balanced approach between two polar approaches that are commonly employed in the literature.

3.2. Computer Adoption

The literature suggests various factors which may affect the diffusion of farm-based computer technology in England and Wales. The likelihood of computer adoption within a farm business depends on the characteristics of the farmer and his/her operation. The age and education of the farmer have been found to be significant determinants in the adoption process (Lazarus and Smith 1988; Putler and Zilberman, 1988; Batte, 1990; Woodburn et al. 1994; Hoag et al. 1999; Lewis 1998; Ascough et al. 2002). Older farmers have been found not to use as many sources of information as their younger colleagues and are more dependent on their experience in farming. Moreover, older and more experienced farm decision makers tend to maintain less complicated record types, which may reduce their demand for computer-based management innovation. Although (Jarvis, 1990; Baker 1992) find that the managers' age and education are insignificant in determining computer adoption among Texan rice producers and New Mexico non-farm agri-businesses, respectively. In addition, Woodburn et al. (1994); Ortmann et al. (1994) and Ascough (2002); find that farmers' self-rating of financial, computer and management skills to be significant factors in the adoption process.

Results from a number of studies (Lazarus and Smith 1988; Putler and Zilberman, 1988; Batte, 1990; Jarvis, 1990; Baker 1992; Woodburn et al. 1994; and Lewis 1998;) indicate that gross farm income or farm size is a significant factor in computer adoption. In the UK, Warren (2000) finds a clear positive relationship between increasing use of computer technology and increasing farm size, as well as a tendency for cattle and sheep farms to have lower levels of adoption than other farm types. Woodburn et al. (1994); also found that the probability of computer adoption declines with the presence of beef enterprises in Natal, South Africa. While Batte, (1990) found adoption rates among Ohio commercial farms to be highest for mixed livestock and dairy producers. The reasoning for these conflicting results may lie in the degree of livestock production intensity in the different regions and the availability of appropriate livestock production decision analysis and record-keeping software. Further significant positive factors in the decision to adopt computer technology include ownership of farm sales related businesses (Putler and Zilberman, 1988; and Baker 1992), the presence of off-farm employment and higher proportions of rented land (Woodburn et al. 1994), and reduced-levels of diversification (Putler and Zilberman, 1988) and off-farm investments (Ortmann et al., 1994).

3.2.1. Data on Computer Adoption

The Department for Environment, Food and Rural Affair's (Defra) (2001) survey of computer use in England found that 35 per cent of holdings had computer access. Moreover, 25 per cent of holdings owned a computer but do not use it for farm business. In the 2002/03 FBS survey period, 75 per cent of 1,718 farmers had access to a computer, and 76 per cent of these farmers used computers for farm business purposes. Of those farm business computer users 82 per cent made at least some use of the computer for office management functions, 69 per cent for farm management accounts, 55 per cent for livestock enterprise management, 49 per cent for statutory records, 42 per cent for tax accounts, 39 per cent for arable enterprise management, and 23 per cent for the farm's payroll. In this paper we use data from Defra's Farm Business Survey data for 1,718 farms in England and Wales over the 2002/2003 financial year. There are 335, 531 and 424 farms in the North, East and West of England, respectively, and 428 in Wales. There are 917 full owner-occupied farms, 251 full tenanted farms, and 550 have a mixed tenure status. The sample includes 622 small farms, 613 medium-sized farms and 483 large farms.

The average age of the farmers in the sample is 54 years. Of the total sample of farmers, 563 (33 per cent) have a "school only" highest education level, while 873 (51 per cent) have GCE "O" or "A" levels or the equivalent, and 211 (12 per cent) have a degree or postgraduate qualification. Of the four regions surveyed the

East of England has the lowest proportion of "school only" educated farmers and the highest proportion of farmers with GCE and university qualifications, whilst the reverse is true for Wales. The average age of farmers with a "school only" education is 60 years, while those with a GCE/College education, and university graduates, average 52 years, and 51 years, in age, respectively.

Table 2 gives a summary of the use of computers on the farms in the sample. Of the total sample of farms, 432 (25 per cent) did not have access to a computer, 314 (18 per cent) used a computer for personal/family purposes only, and 972 (57 per cent) used a computer for farm or related business use. Farmers are extending their use of computers for farm or related business use, specifically various farm, financial and record management purposes.

3.2.2. Results for Computer Adoption

The dependent variable in our analysis is whether the farmer owns and uses a computer on the farm. The explanatory variables are detailed in Table 3. The structure of Table 3 is the same as in Table 2. As with organics data, there is close correspondence between the significance of the variable, and the probability that it will enter the model. In 9 out of the 12 significant variables the probability that the variable enters the model exceeds 0.95. In a number of cases, the significant variables were deemed to be in the model with probability near 1. However, *East area*, *Cattle and Sheep Farms in the less favoured areas*, and *Net Farm income,* have probabilities of entering the model of 0.363, 0.696, and 0.763 respectively even though they are significant at the 5% level. Thus, as in the case of the organics data, the BMA results differ substantively from the standard one in some important respects. Overall, the correspondence between results with and without BMA, are closer for the computer data than for the organics data. This is unsurprising since, with the larger sample size, the power of tests increase, and important variables are more likely to be significant.

With regard to farm type, there is clear evidence that farms classified under Cereals, are more likely to use a computer than other farming types. This classification variable enters the model 100% of the times. Other farm classifications do not seem to be particularly important, except perhaps *Cattle and Sheep Farms in less favoured areas,* which has a negative association with computer ownership. A number of findings differ from the preceding literature. In contrast to previous results, *age* is not found to be an important predictor, at least when other covariates are taken into account. This variable enters the equation only 3.5% of the times, and is also insignificant at the 95% level. Next, in contrast to previous results, we do not find a positive relationship between farm size and computer use. Larger farms are found to be less likely to use computers than the median sized farm. Farm size is significant and is included in the model over 99% of the times, and its coefficient is negative. However, net farm income (which is positively related to farm size) is found to be a significant positive predictor of computer use, although it is included in the models around 75% of the times. Other variables connected to size include the *number of paid workers* which is positively related to the physical size of the farm, and is included in 100% of the models and is positively related to computer usage. Care needs to be taken in interpreting these results. Although the results are not included here, a simple probit of computer use on farm size (excluding other variables) indicates a positive relationship between farm size and computer use. Therefore, it is the inclusion of other covariates in the model that has produced this result. Our results suggests that large farms with given *net incomes*, and *number of paid workers* are less likely to use computers. However, these other covariates also reflect farm size. Nonetheless, we believe that the results correctly reflect the fact that what might be termed the 'commercial size' of the farm is positive predictor of whether a computer is used, rather than land size. The role of education is generally in line with previous findings. The 'school only' variable seems to have a significant negative impact on the use of computers in the farm, with the GCE qualification also. However, the influence of a degree, while probably being positive, is insignificant, and enters the regression only around 25% of the times.

Finally, comparing the magnitudes of the BMA coefficients and the original ones, the impact of using BMA has been similar for the Computer data set and Organic data set. In most cases the sign of the coefficients remain unchanged, but for those variables entering the model with a small probability, the coefficients are correspondingly small in absolute terms.

4. **Conclusions**

This paper outlined the BMA approach to Probit regressions with uncertain regressors and then explores its use in a comparison of two Probit regressions with and without using BMA. We found the BMA method to be fast, and added another useful layer of information when interpreting the results. While we found a high correspondence between the results across estimation with and without BMA, there were also some substantive differences. Overall, if a variable was significant at the 95% level this could not be used as reliable indicator of whether that variable should be in the model.

With regard to the results on organic adoption, while broadly in accord with Rigby et al. (1999), some differences were obtained by using BMA. Most notably, we found that using BMA produced considerably weaker evidence that age and membership of an environmental organisation were good predictors of the use of organic technology, once other covariates were taken into account.

With regard to the influence of farm and farmer characteristics on the uptake and use of computers, we found that cereal farmers were much more likely to be users of computer technology. With regard to size, contrary to previous work, the physical size of the farm was negatively associated with computer use, once covariates were taken into account. Education was found to be a useful predictor, with those farmers having only a school education being less likely to use a computer. The impact of higher levels of education (a degree) were less clear.

The use of BMA in this article has been limited to the Probit model with linear effects. However, there are other contexts in which it may have utility. One further application may be in the selection of regressors when using 'flexible functional forms', which are popular in the Agricultural Economics literature. Where the number of explanatory variables is large, flexible functional forms can suffer badly from the 'curse of dimensionality'.

Technical Appendix

A1. Preliminary Definitions

Take the model as defined in the paper. Let V be the prior variance, and M_0 be the prior precision for the parameters b that have a prior normal $b \sim N\left(\beta_0, M_0^{-1}\right)$. Also, let β and M be conformable vectors and matrices indexed by j (defined further below) and define: $Q_j = \left(b-\beta_j\right)' M_j \left(b-\beta_j\right)$ and $P_j = \beta_j' M_j \beta_j$. Using this notation, the prior distribution for β is: $\pi(b) = 2\pi^{-\frac{p}{2}} \left|M_0\right|^{\frac{1}{2}} \exp\left(\frac{-Q_0}{2}\right)$. The likelihood is $f(y|b) = (2\pi)^{\frac{-T}{2}} \exp\left(-\frac{S(b)}{2}\right)$. Further define $M_1 = x'x$ and $M_2 = (M_0 + M_1)$. Let $\beta_1 = M_1^{-1} x'y$ and $\beta_2 = M_2^{-1}(M_0\beta_0 + M_1\beta_1)$. Further define $S(.)$ $S\left(\beta_j\right) = \sum\left(y_t - x_t'\beta_j\right)^2$. Three results are of use in what follows are:

- (See Proof 1)

$$S(b) = S(\beta_1) + Q_1 \tag{5}$$

- (See Proof 2)

$$Q_0 + Q_1 - Q_2 = P_0 + P_1 - P_2 \tag{6}$$

- and (see Proof 3)

$$S(\beta_1) + P_0 + P_1 - P_2 = \underbrace{(y - x\beta_0)'(y - x\beta_0)}_{A} - \underbrace{(y - x\beta_0)' x M_2^{-1} x' (y - x\beta_0)}_{B} \tag{7}$$

A2. Deriving the Posterior and Marginal Likelihood

A2.1 The Posterior

Combining the prior with the likelihood we obtain:

$$p(\beta/y) = \left(\frac{1}{\sqrt{2\pi}}\right)^{T+p} \left|M_0\right|^{\frac{1}{2}} \exp\left(-\frac{S(b) + Q_0}{2}\right) \tag{8}$$

Using (5)

$$p(\beta/y) = (2\pi)^{-\frac{T+p}{2}} \left|M_0^{-1}\right|^{-\frac{1}{2}} \exp\left(-\frac{S(\beta_1) + Q_1 + Q_0}{2}\right) \tag{9}$$

using (6)

$$p\left(\beta/y\right)=\left(2\pi\right)^{-\frac{T+p}{2}}\left|M_0^{-1}\right|^{-\frac{1}{2}}\exp\left(-\frac{S\left(\beta_1\right)+P_0+P_1-P_2}{2}\right)\times\exp\left(-\frac{Q_2}{2}\right) \tag{10}$$

therefore $S\left(\beta_1\right)$, P_0, P_1 and P_2 are functions of the data and priors only. It is evident that this joint density is:

$$p\left(\beta/y\right)\propto\exp\left(-\frac{\left(b-\beta_2\right)'M_2\left(b-\beta_2\right)'}{2}\right) \tag{11}$$

Therefore $\beta/y\sim N\left(\beta_2, M_2^{-1}\right)$.

A2.2 The Marginal Likelihood

The marginal likelihood is $ML=\int_\beta p\left(\beta/y\right)d\beta$, therefore:

$$\begin{aligned}ML &= \left(2\pi\right)^{-\frac{T+p}{2}}\left|M_0^{-1}\right|^{-\frac{1}{2}}\exp\left(-\frac{S\left(\beta_1\right)+P_0+P_1-P_2}{2}\right)\int_\beta\times\exp\left(-\frac{Q_2}{2}\right)d\beta \\ &= \left(2\pi\right)^{-\frac{T+p}{2}}\left|M_0^{-1}\right|^{-\frac{1}{2}}\exp\left(-\frac{S\left(\beta_1\right)+P_0+P_1-P_2}{2}\right)\times\left|M_2^{-1}\right|^{\frac{1}{2}}\left(2\pi\right)^{\frac{p}{2}}\end{aligned} \tag{12}$$

Using (6):

$$J=S\left(\beta_1\right)+P_0+P_1-P_2=\underbrace{\left(y-x\beta_0\right)'\left(y-x\beta_0\right)}_{A}-\underbrace{\left(y-x\beta_0\right)'xM_2^{-1}x'\left(y-x\beta_0\right)}_{B} \tag{13}$$

Thus the marginal likelihood observes the following proportionality:

$$ML\propto\left(2\pi\right)^{-\frac{T}{2}}\frac{\left|M_2\right|^{-\frac{1}{2}}}{\left|M_0\right|^{-\frac{1}{2}}}\times e^{-\frac{J}{2}} \tag{14}$$

It is this equation that provides the basis of the Metropolis Hastings acceptance probability in the paper (equation 3).

A3. Proofs

Proof 1:

$$
\begin{aligned}
S(b) &= S(\beta_1) + Q_1 = (y - xb)'(y - xb) \qquad (15)\\
&= (y - x\beta_1 - x(b - \beta_1))'(y - x\beta_1 - x(b - \beta_1))\\
&= \underbrace{(y - x\beta_1)'(y - x\beta_1)}_{=S(\beta_1)} + \underbrace{(b - \beta_1)x'x(b - \beta_1)}_{=Q_1} + 2\underbrace{(y - x\beta_1)'}_{=0}x(b - \beta_1)
\end{aligned}
$$

Proof 2: First note that for j=1,2,3 $Q_j = b^{'}M_j b + P_j - 2b^{'}M_j\beta_j$. Using these conditions

$$
\begin{aligned}
Q_0 + Q_1 - Q_2 &= P_0 + P_1 - P_2 + K\\
&\quad \text{where}\\
K &= b^{'}M_0 b - 2b^{'}M_0\beta_0 + b^{'}M_1 b_1 - 2b^{'}M_1\beta_1 - b^{'}M_2 b + 2b^{'}M_2\beta_2
\end{aligned}
$$

We can show that K is zero since $b'M_2 b = b^{'}M_0 b + b^{'}M_1 b$ and $b'M_2\beta_2 = b'M_0\beta_0 + b'M_1\beta_1$:

$$
\begin{aligned}
K &= b^{'}M_0 b - 2b^{'}M_0\beta_0 + b^{'}M_1 b_1 - 2b^{'}M_1\beta_1 - b^{'}M_0 b - b^{'}M_1 b + 2(b'M_0\beta_0 + b'M_1\beta_1)\\
&= 0 \qquad (16)
\end{aligned}
$$

Proof 3: We need to show that:

$$
\begin{aligned}
J &= S_1 + P_0 + P_1 - P_2 = (y - x\beta_0)'\left(I - xM_2^{-1}x'\right)(y - x\beta_0)\\
&= \underbrace{(y - x\beta_0)'(y - x\beta_0)}_{A} - \underbrace{(y - x\beta_0)'xM_2^{-1}x'(y - x\beta_0)}_{B} \qquad (17)
\end{aligned}
$$

Result 3.1: $S(\beta_1) = y'y - P_1$

- Proof of 3.1

$$\begin{aligned}
S(\beta_1) &= (y - x\beta_1)'(y - x\beta_1) = y'y - P_1 \\
&= y'y + P_1 - 2\beta_1 x'y \qquad (18) \\
\text{using } \beta_1 x'y &= \beta_1 M_1 \beta_1 \Rightarrow S(\beta_1) = y'y - P_1
\end{aligned}$$

Result 3.2: $P_2 - P_0 = \beta_0 M_1 \beta_0 + (y - x\beta_0)' x M_2^{-1} x'(y - x\beta_0) + 2\beta_0' M_2 M_2^{-1} x'(y - x\beta_0)$

- Proof of 3.2

$$\begin{aligned}
\underbrace{\beta_2' M_2 \beta_2}_{P_2} &= \left(\beta_0 + M_2^{-1} x'(y - x\beta_0)\right) M_2 \left(\beta_0 + M_2^{-1} x'(y - x\beta_0)\right) \qquad (19) \\
&= \beta_0 M_2 \beta_0 + (y - x\beta_0)' x M_2^{-1} x'(y - x\beta_0) + 2\beta_0' M_2 M_2^{-1} x'(y - x\beta_0) \\
&= \underbrace{\beta_0 M_0 \beta_0}_{P_0} + \beta_0 M_1 \beta_0 + (y - x\beta_0)' x M_2^{-1} x'(y - x\beta_0) + 2\beta_0' M_2 M_2^{-1} x'(y - x\beta_0)
\end{aligned}$$

Therefore, using R3.1

$$J = S_1 + P_0 + P_1 - P_2 = y'y + P_0 - P_2 \qquad (20)$$

and R3.2

$$\begin{aligned}
J &= y'y - \underbrace{\left[\beta_0 M_1 \beta_0 + (y - x\beta_0)' x M_2^{-1} x'(y - x\beta_0) + 2\beta_0' M_2 M_2^{-1} x'(y - x\beta_0)\right]}_{P_2 - P_0} \\
&= \underbrace{y'y - \beta_0 M_1 \beta_0 - 2\beta_0' x'(y - x\beta_0)}_{A = (y - x'\beta_0)'(y - x\beta_0)} - (y - x\beta_0)' x M_2^{-1} x'(y - x\beta_0) \qquad (21)
\end{aligned}$$

References

Balcombe K. and G Rapsomanikis, An Analysis of the Impact of R&D on Productivity using Bayesian Model Averaging with a Reversible Jump Algorithm *American Journal of Agricultural Economics.* (forthcoming)

Balcombe K., Bailey A. and Fraser I. Measuring the impact of R&D on Productivity from a Econometric Time Series Perspective. *Journal of Productivity Analysis*, vol 24, (2005) pp.49—72

Baker G. Computer Adoption and Use by New Mexico Nonfarm Agribusiness. *American Journal of Agricultural Economics,* (1992) pp.737-744.

Batte M. Jones E and Schnitkey G. Computer use by Ohio Commercial Farmers, *American Journal of Agricultural Economics* vol 72 (4) (1990) 935-945.

Chib. S. and Greenberg E. Understanding the Metropolis-Hastings Algorithm. The American Statistician, Vol. 49, No. 4. (1995), pp. 327-335.

Chen M.H., Shao, Q.M and Ibrahim, J.G. . Monte Carlo Methods in Bayesian Computation. Springer Series in Statistics, Springer (2000) New York.

Fernandez, C., Ley E. Steel M. . Benchmark priors for Bayesian Model Averaging. *Journal of Econometrics* vol 100 (2001) pp. 381-427.

George E.I. and McCulloch R.E. . Variable Selection via Gibbs Sampling. *Journal of the American Statistical Association* vol 88 (1993) pp. 881-889.

Green, P.J. . Reversible Jump Markov Chain Monte Carlo Computation and Bayesian and Bayesian Model Determination. vol 82, *Biometrika* (1995) pp.711-732.

Holmes, C.C. and Held, L. . Bayesian Auxiliary Variable Models for Binary and Multinomial Regression. Bayesian Analysis, vol 1 (2005) , pp.145-168.

Hoag D. Ascough, J. and Fraser W. Farm Computer Adoption in the Great Plains, *Journal of Agricultural and Applied Economics,* vol 31 (1999). pp.57-67.

Hendry David F. and Hans-Martin Krolzig Volume (Year): 115 *Economic Journal*: vol 3, (2005) pp.502

Jarvis, A. Computer Adoption Decisions - Implications for Research and Extension: The Case of Texas Rice Producers. *American Journal of Agricultural Economics* (1990) pp.1388-1398.

Lazarus, W. and Smith Adoption of Computers and Consultant Services by New York Dairy Farmers. *Journal of Dairy Science* (1988) 70: pp.1667-1675

Leon-Gonzalez, Roberto & Scarpa, Riccardo, . Improving multi-site benefit functions via Bayesian model averaging: A new approach to benefit transfer, *Journal of Environmental Economics and Management*, vol. 56(1) (2008), pp.50-68

Lewis T Evolution of Farm Management Information Systems, *Computers and Electronics in Agriculture,* vol 19 (1998) 233-248.

Mercier W.H.S. Joint Adoption of Microprocessor Technologies: An Analysis of Farmer Decisions, *Review of Economics and Statistics,* vol 73(3) (1991) pp.541-545

Miller, A. J. Selection of Subsets of Regression Variables (with discussion) *Journal of the Royal Statistical Society* A, vol 147, (3), (1984) pp.389–425.

Mittelhammer R.C. Judge, G.G. and Miller D.J. Econometric Foundations. Cambridge University Press. (2000).

Ortman G, Patrick G. and Musser Use and Rating of Computers by Large Scale U.S. Cornbelt Farmers, *Computers and Electronics in Agriculture* (1994) pp.31-43.

Pagan A. Three Econometric Methodologies, A critical Appraisal. *Journal of Economic Surveys* (1987) 1,1, pp.3-24.

Putler D. and Zilberman D Computer Use in Agriculture: Evidence from Tulare County, California, *American Journal of Agricltural Economics.*vol 70, (1988) pp.790-802.

Rigby D and M Burton and T Young Journal of Agricultural Economics - vol 50 (1), (1999) pp.47-63

Sala-i-Martin, X, Doppelhofer, G., Miller, R.I. . Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review* vol 94(4) (2004) pp.813-35.

Warren M. E-Farming or E-Folly? Adoption of Internet Technology by Farmers in England, Technical Report University of Plymouth, Newton Abbot. Devon, UK. (2000) pp.16-21.

Woodburn M. Ortmann G. and Levin. Computer Use and Factors Influencing Computer Adoption amoung Commercial Farmers in Natal Province, South Africa, *Computers and Electronics in Agriculture*, (1994) pp.183-194.

Tables

Table 1. Organics

	Standard				Model Averaged		
	Est	Stdv	LB	UB	Est	Stdv	Prob
int	-1.316*	0.240	-1.809	-0.864	-1.161	0.245	1.000
hhsize	0.255*	0.092	0.079	0.437	0.175	0.119	0.796
fem	1.430*	0.549	0.358	2.495	0.935	0.663	0.781
age	-0.027*	0.015	-0.056	0.001	-0.012	0.015	0.469
hefe	-0.182	0.329	-0.810	0.476	-0.006	0.121	0.130
yagric	-0.025	0.377	-0.766	0.707	-0.036	0.164	0.160
toha	-0.002	0.001	-0.005	-0.000	-0.002	0.002	0.633
enviss	0.986	0.510	-0.005	1.992	0.619	0.650	0.595
conin	-2.378*	0.486	-3.336	-1.453	-2.231	0.482	1.000
orgff	1.195*	0.325	0.566	1.848	1.196	0.324	0.996
fsv	0.144	0.350	-0.542	0.828	0.052	0.184	0.183
maxcon	0.761	0.447	-0.100	1.630	0.539	0.536	0.616
memenv	0.674*	0.341	0.004	1.332	0.382	0.423	0.554
mempga	-0.937*	0.371	-1.666	-0.212	-0.740	0.453	0.836
infpss	-1.333*	0.346	-2.001	-0.650	-1.237	0.332	0.999
infbuy	-1.278*	0.340	-1.955	-0.622	-1.142	0.344	0.991
inffmrs	0.835*	0.332	0.192	1.480	0.579	0.433	0.748
infadas	-1.112*	0.359	-1.832	-0.423	-1.076	0.389	0.972

Table 2. On Farm Computer Uses

	No Use		Some Use		Total Reliance	
	Farms	%	Farms	%	Farms	%
Office Management	918	53	491	29	309	18
Farm Management	1048	61	264	15	406	24
Tax Accounts	1305	76	186	11	227	13
Payroll	1499	87	74	4	145	8
Arable Enterprise Management	1342	78	222	13	154	9
Livestock Enterprise Management	1183	69	359	21	176	10
Statutory Records	1244	72	334	19	140	8
Other	1612	94	43	3	63	4

Table 3: Computer Adoption

	Standard				Model Averaged		
Area	Est	Stdv	LB	UB	Est	Stdv	Prob
North	0.269*	0.033	0.205	0.332	0.264	0.032	1.000
East	0.207*	0.091	0.030	0.385	0.075	0.112	0.363
West	0.566*	0.098	0.377	0.757	0.514	0.092	1.000
Main Farm Activity							
Cereals	0.538*	0.090	0.356	0.708	0.491	0.089	1.000
General Crop	-0.044	0.124	-0.295	0.189	0.004	0.026	0.038
Horticultural	-0.191	0.142	-0.466	0.089	-0.007	0.040	0.056
Spec Pig and Poultry	-0.143	0.161	-0.442	0.182	-0.003	0.031	0.039
Dairy	-0.042	0.172	-0.370	0.297	0.002	0.026	0.026
Cattle and Sheep LFA	-0.271*	0.117	-0.499	-0.042	-0.157	0.122	0.696
Cattle and Sheep Lowland	-0.093	0.121	-0.335	0.137	-0.006	0.034	0.057
Farm Size							
Small	0.078	0.133	-0.176	0.345	0.022	0.067	0.125
Large	-0.299*	0.072	-0.443	-0.162	-0.298	0.077	0.994
Tenancy Status							
Fully Owned	0.271*	0.084	0.110	0.436	0.294	0.103	0.958
Fully Tenanted	0.014	0.063	-0.109	0.136	-0.000	0.010	0.027
Education							
School Only	-0.189*	0.029	-0.248	-0.134	-0.193	0.028	1.000
CGE College	-0.213	0.154	-0.517	0.085	-0.250	0.151	0.782
Degree	0.096	0.151	-0.206	0.385	0.056	0.132	0.274
Ownership							
Sole	0.717*	0.178	0.373	1.070	0.696	0.183	0.993
Partner	-0.304	0.178	-0.652	0.044	-0.095	0.104	0.531
Company	-0.133	0.178	-0.478	0.223	0.021	0.069	0.186
Other Attributes							
No of Unpaid Workers	0.056	0.076	-0.095	0.204	0.001	0.014	0.028
No of Paid Workers	1.307*	0.230	0.862	1.750	1.352	0.213	1.000
Net Farm Income	0.390*	0.177	0.042	0.732	0.299	0.222	0.762
Organic Enterprises	0.016	0.009	-0.001	0.035	0.002	0.006	0.115
Age of Farmer	0.026	0.087	-0.141	0.199	0.001	0.016	0.035
Other Activities							
Off Farm Work	0.261*	0.086	0.093	0.434	0.262	0.081	0.976
Off Farm Income	-0.024	0.071	-0.161	0.114	-0.002	0.015	0.039
Socail Payments	0.097	0.140	-0.187	0.361	0.006	0.041	0.051