Available online at www.sciencedirect.com
SCIENCE DIRECT®

JOURNAL OF Econometrics

Journal of Econometrics 124 (2005) 311–334

www.elsevier.com/locate/econbase

A Bayesian analysis of the multinomial probit model using marginal data augmentation

Kosuke Imai[a,*], David A. van Dyk[b]

[a]Department of Politics, Princeton University, Princeton, NJ 08544, USA
[b]Department of Statistics, University of California, Irvine, CA 92697, USA

Accepted 9 February 2004

Abstract

We introduce a set of new Markov chain Monte Carlo algorithms for Bayesian analysis of the multinomial probit model. Our Bayesian representation of the model places a new, and possibly improper, prior distribution directly on the identifiable parameters and thus is relatively easy to interpret and use. Our algorithms, which are based on the method of marginal data augmentation, involve only draws from standard distributions and dominate other available Bayesian methods in that they are as quick to converge as the fastest methods but with a more attractive prior specification. C-code along with an R interface for our algorithms is publicly available.[1]

JEL classification: C11; C25; C35

Keywords: Bayesian analysis; Data augmentation; Prior distributions; Probit models; Rate of convergence

1. Introduction

Discrete choice models are widely used in the social sciences and transportation studies to analyze decisions made by individuals (see, e.g., Maddala, 1983; Ben-Akiva and Lerman, 1985). Among such models, the multinomial probit model is often appealing because it lacks the unrealistic assumption of independence of irrelevant alternatives of logistic models (see, e.g. Hausman and Wise, 1978). Despite this appeal, the model is sometimes overlooked because model fitting can be computationally demanding owing

* Corresponding author. Tel.: +1-609-258-6601; fax: +1-609-258-1110.
E-mail addresses: kimai@princeton.edu (Kosuke Imai), dvd@ics.uci.edu (David A. van Dyk).

[1] R is a freely available statistical computing environment that runs on any platform. The R software that implements the algorithms introduced in this article is available from the first author's website at `http://www.princeton.edu/~kimai/`.

doi:10.1016/j.jeconom.2004.02.002

to the required high-dimensional integrations. Recent advances in Bayesian simulation, however, have shown that Gibbs sampling algorithms based on the method of data augmentation can provide reliable model fitting (Geweke et al., 1994). Hence, the development of efficient Markov chain Monte Carlo (MCMC) algorithms has been a topic of much recent work; see, e.g., McCulloch and Rossi (1994), Chib et al. (1998), Nobile (1998), McCulloch et al. (2000), Nobile (2000), and McCulloch and Rossi (2000).

The basic computational strategy of the proposed MCMC methods is to identify an underlying set of Gaussian latent variables, the relative magnitudes of which determines the choice of an individual. Because the natural parameterization of this model is unidentifiable given the observed choice data, a proper prior distribution is required to achieve posterior propriety. As proposed by McCulloch and Rossi (1994), a Monte Carlo sample of the identifiable parameters can then be recovered and be used for Monte Carlo integration in a Bayesian analysis. A complication involved in this procedure is that the prior distribution for the identifiable model parameters is determined as a byproduct. Inspection (e.g., via simulation) is therefore required to determine what prior distribution is actually being specified and how sensitive the final results are to this specification.

To improve the computational performance of McCulloch and Rossi's (1994) algorithm (but without addressing the difficulties in the prior specification), Nobile (1998) introduced a "hybrid Markov chain." This hybrid is quite similar to the original algorithm but adds an additional Metropolis step to sample the unidentifiable parameters and *appears* to dramatically improve the performance (i.e., mixing) of the resulting Markov chains. We illustrate that the improved mixing of Nobile's hybrid method seems to be primarily for the unidentifiable parameter—the gain for the identifiable model parameters is much smaller, at least in terms of the autocorrelation of their Monte Carlo draws. Nonetheless, Nobile's method has an advantage over McCulloch and Rossi (1994) in that it can be less sensitive to starting values. In addition to this clarification of the improvement offered by Nobile's method, we point out an error in Nobile's derivation which can significantly alter the stationary distribution of the resulting Markov chain and thus hamper valid inference.

A second computational innovation was introduced by McCulloch et al. (2000) and aims to address the difficulties with prior specification (but without addressing the computational speed of the algorithm). In particular, this proposal specifies a prior distribution only on the identifiable parameters and constructs a Markov chain that fixes the unidentifiable parameter. Unfortunately, as pointed out by McCulloch et al. (2000) and Nobile (2000), the resulting algorithm can be much slower to converge than either the procedure of McCulloch and Rossi (1994) or of Nobile (1998).

To clarify comparisons among existing algorithms and the algorithms we introduce, we specify three criteria: (1) the interpretability of the prior specification, (2) the computational speed of the algorithm, and (3) the simplicity of implementation. Our comparisons among the three existing algorithms appear in Table 1, which indicates that none of the algorithms dominates the others.

The primary goal of this article is to introduce new algorithms that perform better than the existing algorithms when evaluated in terms of these three criteria. That is, our