Written Exam — Economics Summer School 2017

Bayesian Econometrics

One-Week Take-Home Exam

January 8th, 2018 (10:00) — January 15th, 2018 (10:00)

This exam question consists of 6 pages in total (including this cover page).

The language of this exam is English. Your exam paper and additional material must be written in English.

Your exam must be uploaded to Digital Exam before Monday, January 15th, at 10:00.

Focus on Exam Cheating

In case of presumed exam cheating, which is observed by either the examination registration of the respective study programmes, the invigilation or the course lecturer, the Head of Studies will make a preliminary inquiry into the matter, requesting a statement from the course lecturer and possibly the invigilation, too. Furthermore, the Head of Studies will interview the student. If the Head of Studies finds that there are reasonable grounds to suspect exam cheating, the issue will be reported to the Rector. In the course of the study and during examinations, the student is expected to conform to the rules and regulations governing academic integrity. Academic dishonesty includes falsification, plagiarism, failure to disclose information, and any other kind of misrepresentation of the student's own performance and results or assisting another student herewith. For example failure to indicate sources in written assignments is regarded as failure to disclose information. Attempts to cheat at examinations are dealt with in the same manner as exam cheating which has been carried through. In case of exam cheating, the following sanctions may be imposed by the Rector:

1. A warning.
2. Expulsion from the examination.
3. Suspension from the University for a limited period of time or permanent expulsion.

The Faculty of Social Sciences
The Study and Examination Office
October 2006

Practical Information

Please observe the following formal requirements:

- This is an *individual* examination. You are not allowed to cooperate with other students or other people. Student collaboration is considered to be cheating.
- This exam consists of 12 questions. Please answer all questions.
- The exam paper should not exceed 10 standard pages (A4, font size set to 12, line spacing set to 1.5, margins (left/right/top/bottom) of at least 2 cm). All pages must be numbered consecutively. A maximum of 7 pages of supporting material (figures, estimation output, etc.) can accompany the paper. The computer program must be submitted separately and does not count towards the number of pages of the exam paper or of the additional material.
- Tables and figures displayed in the exam paper should be formatted appropriately (i.e., no raw output, tables and figures should have captions, axes should be labelled, a legend should be added when required, etc.).
- In addition to your exam paper and to the additional material, you must submit your computer program generating all tables and figures. The program must produce tables and figures in the same order as they appear in the exam paper. Comments should clearly indicate which tables or figures are produced. Make sure that the program can be executed without any errors. You are strongly encouraged to write your program in `R`, but you are allowed to use a different programming language.
- You should not write your name on the material you submit (exam paper, computer program, supporting material).
- You must submit the following files to Digital Exam, as separate files:
  - A single PDF document, named `1234.pdf`, where `1234` is your exam identification number, which contains your exam paper and additional material as appendices.
  - Your computer program as a file named `MAIN.R` (or with a different extension if you use a different programming language). If your computer program consists of several files, submit all files and make sure that the file `MAIN.R` loads the other files correctly and produces the expected results.

Bayesian Inference for a Model with Ordinal Responses

Consider the following ordinal probit model, for a sample of individuals $i = 1, \ldots, N$:

$$y_i = \begin{cases} 1 & \text{if } \tau_0 = -\infty < y_i^\star \leq \tau_1 = 0, \\ 2 & \text{if } \tau_1 = 0 < y_i^\star \leq \tau_2, \\ \vdots & \\ K & \text{if } \tau_{K-1} < y_i^\star \leq \tau_K = \infty, \end{cases} \tag{1}$$

where

$$y_i^\star = x_i'\beta + \varepsilon_i, \qquad \varepsilon_i \overset{iid}{\sim} N(0, 1). \tag{2}$$

The latent variable $y_i^\star$ is not observed, but instead, we observe the ordinal variable $y_i \in \{1, 2, \ldots, K\}$. Explanatory variables are contained in the column vector x_i and influence the latent variable $y_i^\star$ through a vector of coefficients β. An intercept term is included in this model, such that the first element of x_i is equal to 1. In compact form, the observed and latent variables are contained in the vectors $y = (y_1, \ldots, y_N)'$ and $y^\star = (y_1^\star, \ldots, y_N^\star)'$, explanatory variables are collected in the matrix X with N rows, where x_i corresponds to row i, and $\tau = (\tau_1, \tau_2, \ldots, \tau_{K-1})'$.

You will use a Bayesian approach to make inference on this model. To keep it as simple as possible, you will use flat priors for β and for the cutpoints τ, such that

$$p(\beta) \propto 1, \qquad p(\tau) \propto \mathbb{1}\{\tau_1 < \tau_2 < \ldots < \tau_{K-1}\}, \tag{3}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function that is equal to 1 if the corresponding condition is fulfilled, to 0 otherwise.

Part A: Deriving a Gibbs sampler

Question A.1 Discuss the identification of this model in the general case with K alternatives. In particular, explain all parameter restrictions made on the model. What would be an alternative identification strategy?

In the following, you will consider a three-level ordinal probit model, where τ_2 is the only unrestricted cutpoint.

Question A.2 Derive the likelihood function of the three-level ordinal probit model.

Question A.3 The prior defined in Eq. (3) is called *improper prior*. Explain *briefly* the difference between proper and improper priors, and why improper priors can be used for Bayesian analysis.

Question A.4 Derive the augmented data density function $p(y, y^\star \mid X, \beta, \tau)$, and explain how it can be used to design a sampling procedure for the ordinal probit model.

Question A.5 Propose a Gibbs sampler for the three-level ordinal probit model. Derive explicitly all the conditional distributions required, and be as precise as possible about the details of your sampler (initialization, parameter values used at each iteration, etc.).

Question A.6 Derive the marginal effect of a given continuous covariate x_{ij} on the probability of choosing alternative k for individual i. How can you use the output of the Gibbs sampler to make inference on this marginal effect?

Part B: Implementing and testing the sampler

Question B.1 Write a computer program implementing the Gibbs sampler derived in Question A.5.

[You may use the function `gibbs_linreg()` provided during the summer school (available on Absalon under `Files/code`) and adapt it to your needs.]

Question B.2 Based on the model defined in Eqs. (1) and (2), generate artifical data with $N = 500$ observations for the three-level ordinal probit model ($y_i \in \{1, 2, 3\}$), using the following data generating process:

$$y_i^\star = 0.5 + 0.3x_i + \varepsilon_i, \qquad \varepsilon_i \overset{iid}{\sim} N(0,1), \qquad x_i \overset{iid}{\sim} N(0,1), \qquad \tau_2 = 1.$$

Question B.3 Run your sampler on this data set for a total of 11,000 MCMC iterations, and discard the first 1,000 iterations as burn-in period.

[Hint: To initialize your sampler, you may use the true values of the latent variable $y^\star$, the OLS estimates for β, and the true value of τ_2.]

Show trace plots and autocorrelograms for all parameters, and report a table of summary statistics (posterior means and standard deviations, confidence intervals, inefficiency factors). Compare your results to the true values of the parameters.

Question B.4 Discuss the convergence and the mixing of the Markov chain produced in Question B.3. How do you explain this result for the cutpoint τ_2? *[Only provide intuition to answer this question, no derivations required.]*

Part C: Improving mixing

To address the issue raised in Question B.4, Cowles (1996)[1] suggested to use a Hastings-within-Gibbs algorithm to sample the latent variable $y^\star$ and the cutpoints τ jointly.

This can be implemented by sampling the cutpoints τ from their marginal distribution (i.e., independently of the latent variable), using for example a random walk based on the normal distribution to make proposals. This normal distribution can be tuned by adjusting its variance σ^2. Depending on the acceptance of this proposed cutpoints, the latent variable can then be updated from its conditional distribution.

Using data generated as in Question B.2, this Hastings-within-Gibbs algorithm was run for different values of the tuning parameter σ^2. In each case, $10,000$ MCMC iterations were saved for posterior inference, after a burn-in period of $1,000$ iterations. Table 1 and Fig. 1 present the corresponding posterior results for the cutpoint τ_2.

Question C.1 Explain briefly and intuitively the principle of the Metropolis-Hastings algorithm, and how it can be used within the Gibbs sampler to update the cutpoints and the latent variable of this model.

[Note that you are not asked to make any derivations to answer this question.]

Question C.2 Describe the results displayed in Table 1 and Fig. 1 for the cutpoint τ_2, and compare them to your own results produced in Question B.3.

Explain the role of the tuning parameter σ^2, and how to select it in practice (both in general and in this particular application).

Table 1: Acceptance rates of the Hastings-within-Gibbs algorithm for different tuning parameters σ^2.

σ^2	0.010	0.030	0.15	0.700
Acc.	0.939	0.818	0.40	0.103

[1]Reference: M. K. Cowles (1996), *"Accelerating Monte Carlo Markov Chain Convergence for Cumulative-Link Generalized Linear Models"*, Statistics and Computing, Vol. 6(2), pp. 101–111. DOI: https://doi.org/10.1007/BF00162520

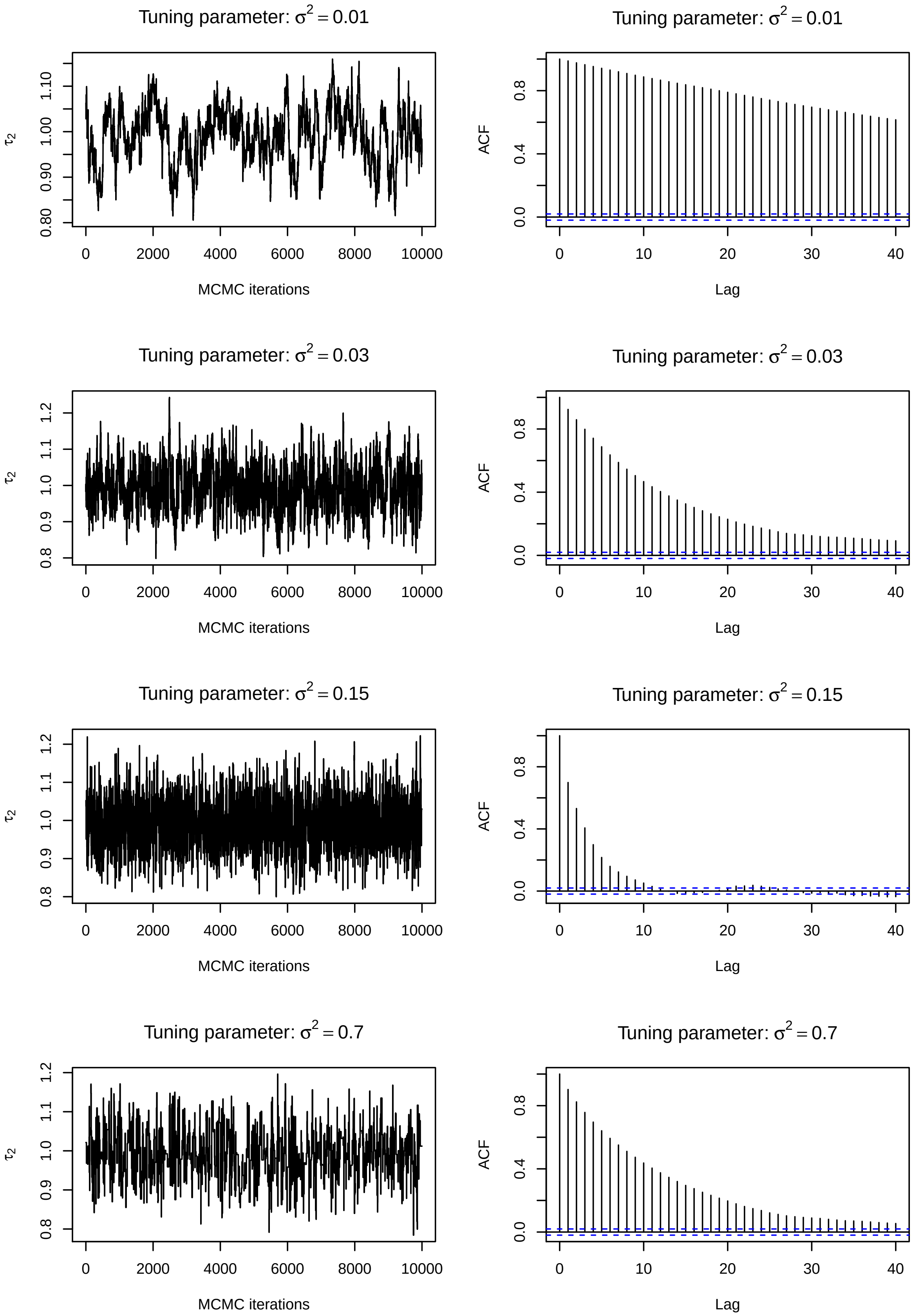

Figure 1: Trace plots (left) and autocorrelograms (right) of the parameter τ_2 from four runs of the Hastings-within-Gibbs sampler with different values of the tuning parameter σ^2. Burn-in period: 1,000 iterations (not shown).