TAKE-HOME FINAL EXAM

(1) The directory on the course web site where the exam was posted contains a data set, extracted from a larger one used by John Lott in a controversial analysis of the effect of "right to carry" laws on crime. The data set is also available on the Stock and Watson textbook's web site. This data set has observations across 51 states and the 23 years 1977-99 on measures of crime, demographic characteristics of states, and a variable called "shall" that is 1 for years and states that had right-to-carry laws. The data is available as an R data file, a `.csv` file, and an Excel file. There is also a separate file describing the variables.

Also in that directory is a set of notes on how to compute random effects estimators from so-called "between" (using data on group means only) and "within" (using deviations from group means, or equivalently fixed-effects estimates) regressions. The notes include a fairly detailed description of how one could implement MCMC inference for random effects estimators based on the between and within regressions.

You are to analyze the random-effects model

```
vio(i,t) = a * pw1064(i,t) + b * avginc(i,t) + c * density(i,t)
                        + d * shall(i,t) + e * trend(t) + eps(i,t) + eta(i)
```

where the variable names are those in the data set, `vio` being the violent crime rate. (`trend` is there only in the R version of the data set. Otherwise you have to construct it yourself.) `i` indexes states, and `t` indexes years. `eps(i,t)` is i.i.d. $N(0, \sigma^2)$ and `\eta(i)` is i.i.d. (but of course constant across t for a given i) $N(0, \tau^2)$. The two shocks `eps(i,t)` and `eta(i)` are both assumed independent of the right-hand-side variables in the regression.

(a) Estimate fixed-effects and simple pooled-data versions of the equation to contrast the results from the two approaches. [Note that variables, like `trend`, that show no variation across states in their state sample means, have to be omitted from the between regression, and variables that show no within-state variation (if there were any) would have to be omitted from the within regression.]

(b) Use a conjugate proper prior with σ^2 and τ^2 independent of each other and both inverse-gamma with one degree of freedom and scale factor 1 (i.e., $\sigma^{-4}e^{-1/\sigma^2}$ as the pdf of σ^2, for example.) Make the regression parameters given the variance parameters normal with zero mean and large variance. Find the posterior mode for the parameters and use Gibbs sampling, as described in the accompanying notes, to generate a sample from the posterior. Plot smoothed estimates of the posterior density for the `shall` coefficient. Is the point estimate large in practical terms? Does a 95% posterior probability interval include zero?

(c) Standard random effects models are often regarded with suspicion because they require the assumption that the `eta(i)` terms are independent of the right-hand-side variables. The most natural form of dependence can be accounted for, though, by entering the group averages as explanatory variables for each member of the group. Since these variables are constant within groups, they cannot be used in a fixed-effects regression, but they can be entered in a random-effects model, where they capture dependence of the group effects on other variables. Using the MLE or posterior mean from the previous regression estimates for σ^2 and τ^2, estimate the model with these additional state-mean variables as a random effects model (i.e., just do GLS). Are the group-mean variables significant by conventional criteria? Do the estimates of other parameters, particularly of `shall`, change?

(d) Extra credit: Maximize posterior density and generate an MCMC sample from the posterior for this model as you did with the previous one.