Two-Sided Matching Model

Anh Le

February 15, 2018

Since the last few days I have the following findings / improvements:

1 MCMC chains for employers' preference α

I had a hunch that the conditional logit side of the model does not work well when the choices' covariates only vary across choices and NOT across choosers. In the labor market example, it means that a job's characteristics appear the same to different workers. In the FDI market example, it means that a country's characteristics appear the same to different multinationals.

While many texts on conditional logit model say that this is fine, my simulation of one-sided conditional logit shows that the MLE is very imprecise, albeit unbiased. This could cause the MCMC to venture into very far-off region.

So I simulate my own two-sided market, with agents on two sides making, evaluating, and accepting offers, where the choices' covariates vary across choosers as well. In the job market example, it means that company offers slightly different jobs to different employees.

As shown below, the MCMC chains are remarkably better. They actually converge to the true parameter values now instead of behaving very poorly as before. This is evidence that my math and my implementation are correct.

But I'm still theoretically confused about the need for the choices' covariates to vary across choosers. If that's a requirement, the data requirement of the model is now tougher, as we need the characteristics of the jobs (or countries) to vary across workers (or multinationals) somehow.

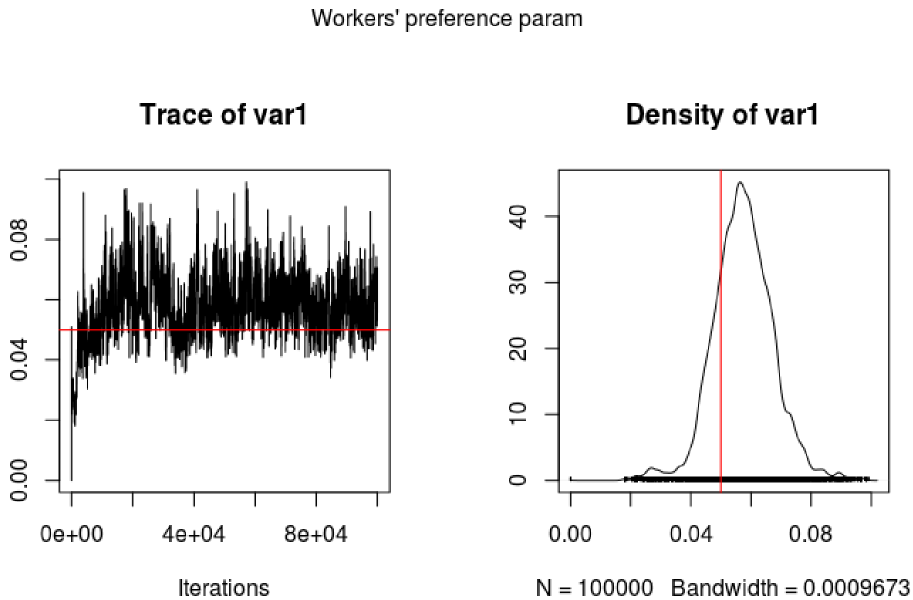

Figure 1: Using hand-crafted opportunity set

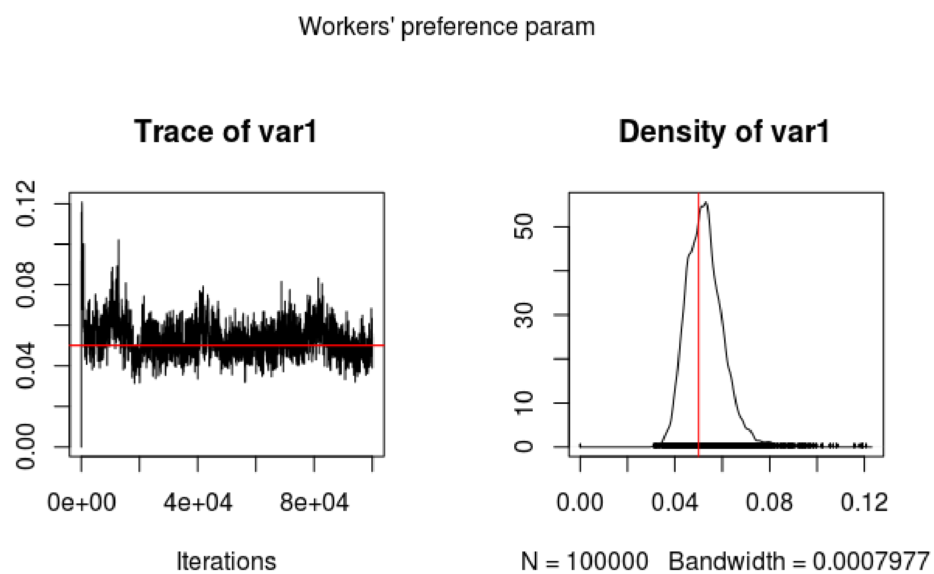

Figure 2: Using the observed opportunity set

2 MCMC chains for employers' preference β

Since I'm simulating the data, I was able to choose sensible scale for the proposal distribution, helping β slowly converge towards their true values. When the proposal values of β manage to get accepted, the proposal values of the opportunity set also do.

But I'm concerned about choosing the right scale for the β in the real dataset. Since the parameters in the random utility models like these only have meaning relative to one another, it seems difficult to reason through what their magnitude may be. This problem exacerbates when real data has 20 employers (or 30 countries), rather than 5 simulated employers here. In theory, an adaptive procedure would help, but past experience suggests that it could only help if the initial, non-adapted part of the MCMC went well.

Figure 3: Using hand-crafted opportunity set

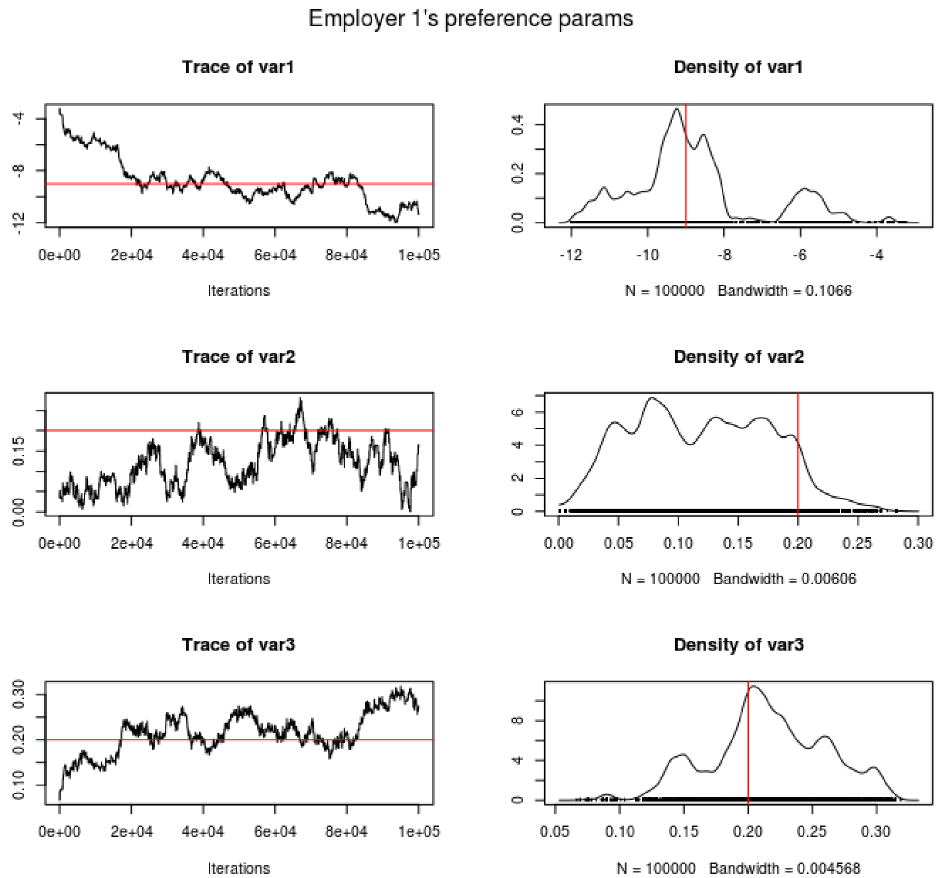

Figure 4: Using the observed opportunity set

3 Monitoring the convergence of the opportunity set

It's also difficult to check if the MCMC for the opportunity set is doing well. The opportunity set is a binary matrix, and below I check various statistics regarding how different the current set is from the true, simulated set. (It's actually quite different – which is confusing given that β seems to get closer to the true value later in the chain). However, for real data, I'm not sure how to monitor the convergence of the opportunity set.

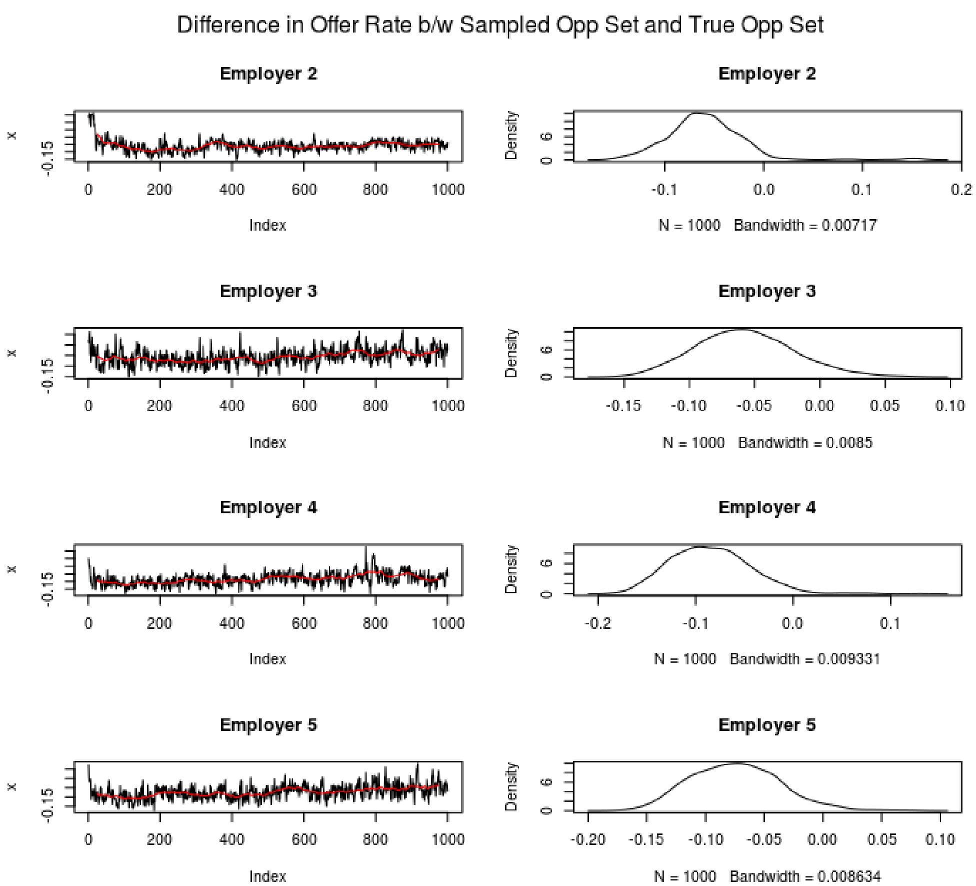

Figure 5: Using hand-crafted opportunity set

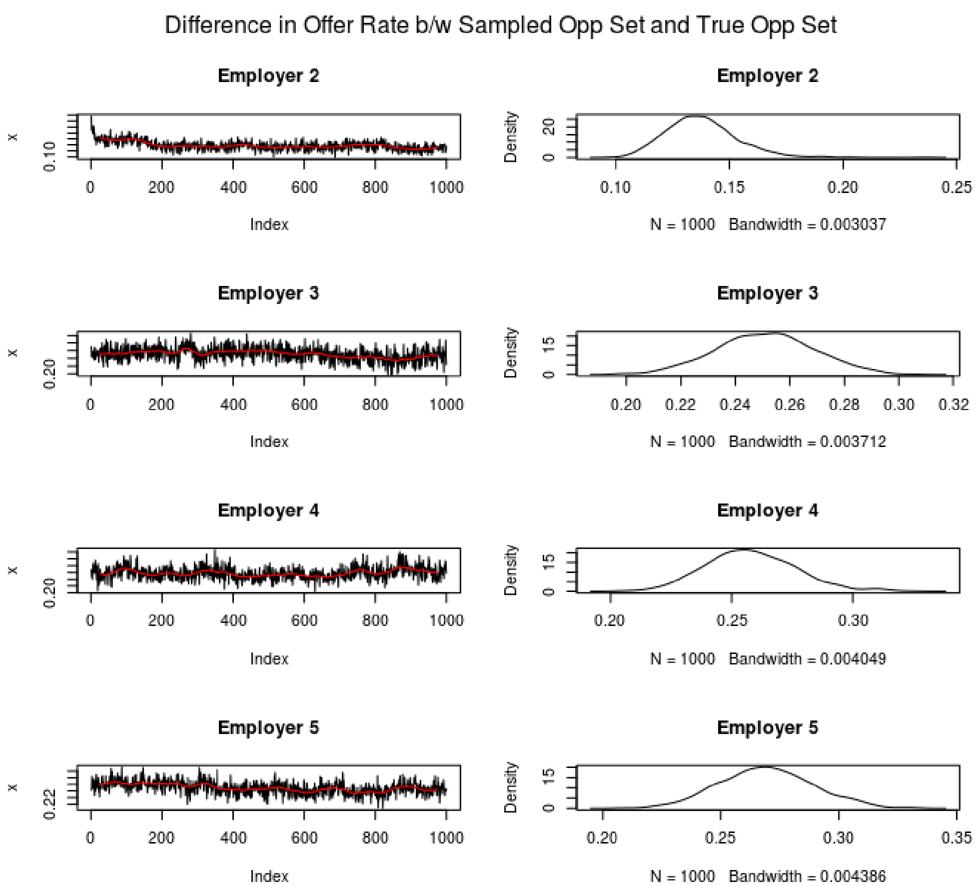

Figure 6: Using the observed opportunity set

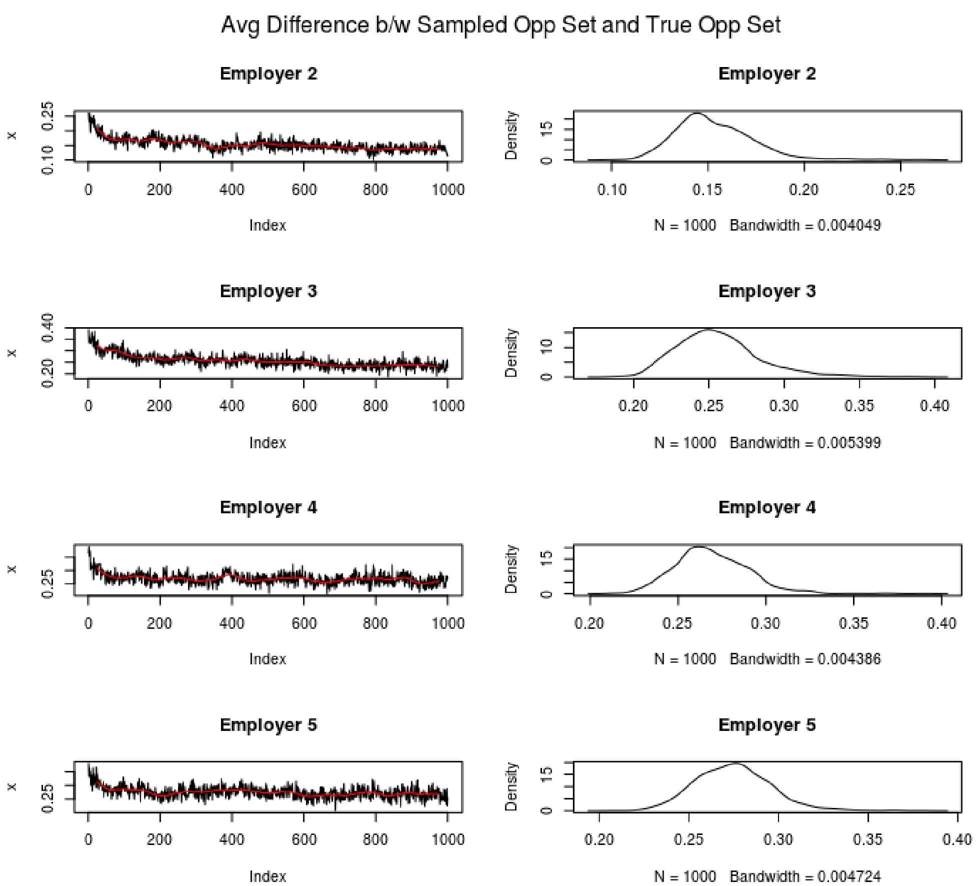

Figure 7: Using hand-crafted opportunity set

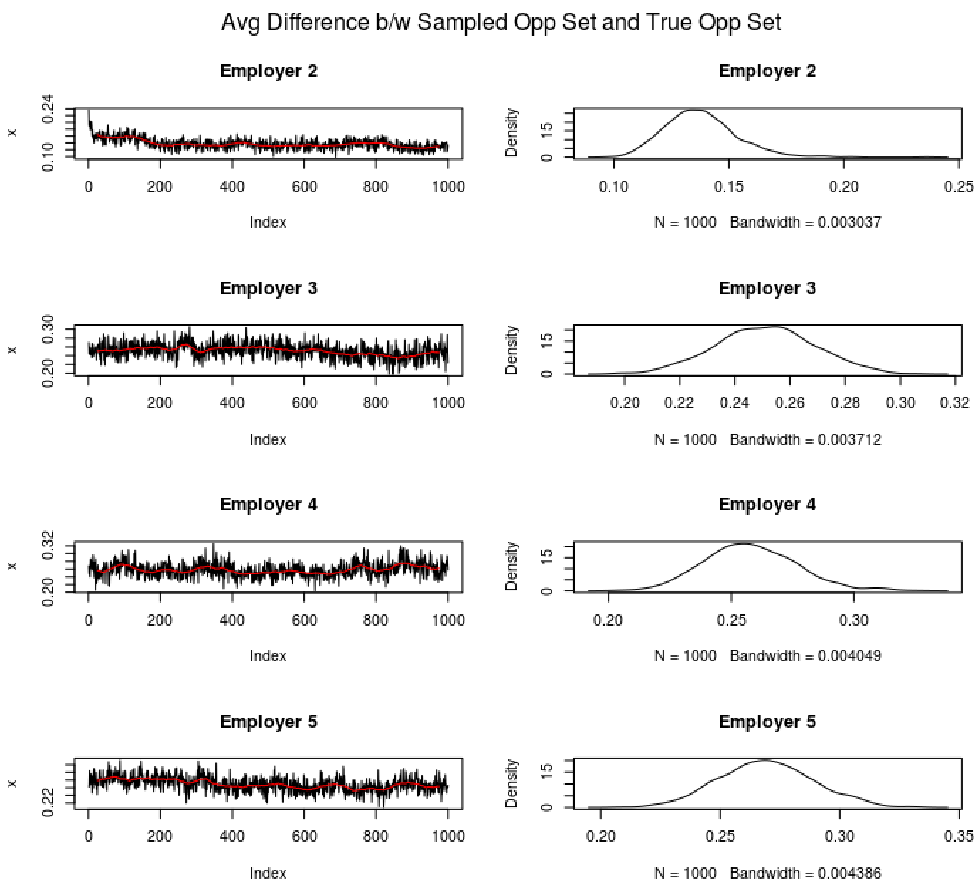

Figure 8: Using the observed opportunity set

References