The political determinants of FDI spillover

Anh Le

April 10, 2016

1 Introduction

We introduce a model that can estimate the utility of both sides, building from their utility function up.

2 Literature

1. FDI aggregate data is not clear about what it means to be foreign. IMF decides 10% is the line between foreign portfolio investment and FDI

3 The Utility Model

3.1 Officials' Utility

Following Logan (1998), we consider the utility function of the two actors, the official and the firm.[1] For official j, the utility of having firm i investing in his country is:

$$U_j(i) = \boldsymbol{\beta}_j' X_i + \epsilon_{1ij} \tag{1}$$

where

$$\begin{aligned} \beta_j &\text{ is a vector of official } j\text{'s preference for relevant characteristics of firms} \\ x_i &\text{ is a vector of firm } i\text{'s measured values on those characteristics} \\ \epsilon_{1ij} &\text{ is the unobserved component that influences official } j\text{'s utility} \end{aligned}$$

On the other hand, the utility of not having firm i investing is:

$$U_j(\neg i) = b_j + \epsilon_{0ij} \tag{2}$$

[1] For ease of exposition, in this section I will refer to country j and official j interchangeably.

where

$$b_j \text{ is the baseline utility of official } j \text{ without any firm investing}$$
$$\epsilon_{0ij} \text{ is the component that influences official } j\text{'s utility}$$

For each firm i, official j will make an offer to invest if $U_j(i) > U_j(\neg i)$. Some relevant firm characteristics (i.e. X_i) that the official may consider are: technological intensity, jobs, and capital. The corresponding β's represent the official's preference for these characteristics.

Following the discrete choice literature, we model $\epsilon_{1ij}, \epsilon_{0ij}$ as having the Gumbel distribution. Then, the probability of official j making an offer to firm i takes the familiar binomial logit form:

$$\begin{aligned} Pr(o_{ij} = 1) &= Pr(U_j(i) > U_j(\neg i)) && (3)\\ &= Pr(\epsilon_{0ij} - \epsilon_{1ij} < \boldsymbol{\beta}'_j X_i - b_j) && (4)\\ &= \frac{\exp(\boldsymbol{\beta}'_j X_i)}{1 + \exp(\boldsymbol{\beta}'_j X_i)} && (5) \end{aligned}$$

where Equation (5) is due to the fact that the difference between two Gumbel-distributed random variables has a logistic distribution. We make the constant term b_j disappear into $\boldsymbol{\beta}_j$ by adding an intercept column to the matrix of firm characteristics X_i.

The opportunity set of firm i is the set of all countries that have made firm i an offer. If we know the preferences of all countries, we can calculate the probability that firm i gets an opportunity set O_i as follows:

$$\begin{aligned} p(O_i|\boldsymbol{\beta}) &= \prod_{j \in O_i} p(o_{ij} = 1|\boldsymbol{\beta}) \prod_{j \notin O_i} p(o_{ij} = 0|\boldsymbol{\beta}) && (6)\\ &= \prod_{j \in O_i} \frac{\exp(\boldsymbol{\beta}'_{\boldsymbol{j}} X_i)}{1 + \exp(\boldsymbol{\beta}'_{\boldsymbol{j}} X_i)} \prod_{j \notin O_i} \frac{\exp(\boldsymbol{\beta}'_{\boldsymbol{j}} X_i)}{1 + \exp(\boldsymbol{\beta}'_{\boldsymbol{j}} X_i)} && (7) \end{aligned}$$

In our observed data, since we only observe the final matching of firms and countries, this opportunity set is unobserved. This is the gist of the statistical challenge. TO overcome this issue, we will use Markov chain Monte Carlo to sample from and approximate $p(O_i|\boldsymbol{\beta})$.

3.2 Firms' utility

On the other side, for firm i, the utility of investing in country j is:

$$V_i(j) = \alpha' W_j + v_{ij} \tag{8}$$

where

α is a vector of firms' preference for relevant characteristics of countries

W_j is a vector of country j measured values on those characteristics

v_{ij} is the unobserved component that influences firm i's utility

Firm i evaluates all the countries that make an offer and chooses the one that brings the highest utility. In our model, the relevant country characteristics are: labor quality, level of development, and market size. Since all firms are considered having homogeneous preferences, α does not have a subscript i. The model can be easily extended so that there is heterogeneous preference among firms.

If v_{ij} is modeled as having a Gumbel distribution, then the probability that firm i will accept the offer of official j out of all the offers in its opportunity set O_i is

$$p(A_i = a_i | O_i, \alpha_i) = \frac{\exp(\alpha' W_{a_i})}{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j)} \tag{9}$$

4 Model Estimation

While Logan (1996, 1998) successfully reformulated standard discrete choice models to a two-sided setting, the estimation of the two-sided model remains challenging. The key difficulty lies in the fact that we do not know the full sets of offers that firms receive from countries. Therefore, the likelihood function is incomplete, missing the data on firms' opportunity sets. With an incomplete likelihood function, we cannot use Maximum Likelihood Estimation to estimate firms' and countries' preferences.

To deal with this problem, Logan (1996) used the Expectation-Maximization (EM) algorithm.[2] However, an important downside of the EM algorithm is its lack of standard error. Therefore, while the algorithm is capable of producing the best estimate for the parameters of interest, it is difficult to know how good our best guess really is.

To overcome this difficulty, I use a Bayesian approach that provides the posterior distribution of the parameters. Firms' opportunity sets, firms' preferences, and countries' preferences are all considered random variables. I estimate their posterior distributions using the Metropolis-Hastings algorithm, which samples from the desired distribution by proposing new values and decide whether to those values.

For example, suppose we want to sample from the posterior distribution $p(\theta)$ and we already had a working collection of values $\{\theta^1, \ldots, \theta^{(s)}\}$. To add new values to this collection, we would propose a new value θ^*, then decide whether to keep it with the probability $\frac{p(\theta^*)}{p(\theta)}$.

[2]The EM algorithm finds the best parameter estimates by iterating between two steps. First, given the current best guess of firms' and countries' preferences, pick values for the unobserved opportunity sets so that we maximize the likelihood. Second, given the current best guess of the unobserved opportunity sets, taken from step 1, pick values for firms' and countries' preferences so that we maximize the likelihood. By iterating between these two steps, the algorithm constantly searches for parameters values that maximize the likelihood.

The intuition is that if $\frac{p(\theta^*)}{p(\theta)}$ is large, then θ^* is very likely compared to θ given $p(\theta)$. Thus, we should keep θ^* and add it to the collection. In other words, we decide to keep newly proposed values of θ at a rate proportional to how often they should appear according to $p(\theta)$. Repeating this step many times, at the end we will have a collection of θ values that approximates $p(\theta)$ as desired.

Appendix B describes the details of our Bayesian model and derives the Metropolis-Hastings acceptance ratio. The key points are that we use flat priors so that our results are driven entirely by the data. The joint distribution of the data and parameters, including the opportunity sets, the firms' and countries' preferences, factor out nicely:

$$p(A_i, O_i, \alpha, \boldsymbol{\beta}) = p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta})$$

5 Results

Host GDP is negative, host GDP per capita is positive, host democracy is positive, agree with (Eicher et al. 2012)

A EM algorithm

Our data contains a random sample of firms and the countries in which they invest. We want to find the parameters that maximize the likelihood of this observed data. This likelihood is:

$$L = \prod_{i,j:\text{i is matched with j}} Pr(A_{ij})$$

where $Pr(A_{ij})$ is the probability of a specific match between firm i and country j. $Pr(A_{ij})$ can be calculated as follows:

$$\begin{aligned}
&Pr(A_{ij}) && (10)\\
&= \sum_{k=1}^{R} Pr(A_{ij}|S_{ik})Pr(S_{ik}) && (11)\\
&= \sum_{k=1}^{R} Pr(A_{ij}|S_{ik}) \prod_{m\in O_k} Pr(o_{im}=1) \prod_{n\in \bar{O}_k} Pr(o_{in}=0) && (12)\\
&= \sum_{k:j\in O_k} \frac{\exp(\alpha w_{ij})}{\sum_{h\in O_k}\exp(\alpha w_{ih})} \prod_{m\in O_k, m>0} \frac{\exp(\beta x_i)}{1+\exp(\beta x_i)} && (13)\\
&\times \prod_{n\in \bar{O}_k, n>0} \frac{1}{1+\exp(\beta x_i)} && (14)
\end{aligned}$$

Here, the term $Pr(o_{ij}=1)$ represents the probability that country j makes an offer to firm i, through which the official's preference enters our estimation. On the other side, $Pr(A_{ij})|S_{ik}$ is the probability that firm i will accept the offer from official j, given the offering set O_k. The firm's preference is reflected in our estimation through this term, $Pr(A_{ij})|S_{ik}$.[3]

It is important to note that the offering set O_k contains *all* offers that firm i receives, only one of which is the observed match between firm i and country j. The intuition is that if we observe the full set of offers that all officials make to all firms, then by looking at the final match we can see how firms and officials reject inferior offers and thus deduce their preferences.

B Deriving the Metropolis Hasting acceptance ratio

B.1 Updating the opportunity set

Target distribution for a firm i

$$p(O_i|A_i,\alpha,\boldsymbol{\beta}) = \frac{p(O_i,A_i,\alpha,\boldsymbol{\beta})}{p(A_i,\alpha,\boldsymbol{\beta})} \tag{15}$$

[3]The appendix shows how these terms are derived.

We propose a new O_i^* by randomly sample a new offer, j^* for each firm. If the new offer is not already in the current opportunity set, we add the offer to the set. If it already is in the opportunity set, we remove it from the set.

We then calculate the Metropolis-Hasting acceptance ratio:

$$
\begin{aligned}
MH_O = \frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} &= \frac{p(O_i^*,A_i,\alpha,\boldsymbol{\beta})}{p(A_i,\alpha,\boldsymbol{\beta})} \times \frac{p(A_i,\alpha,\boldsymbol{\beta})}{p(O_i,A_i,\alpha,\boldsymbol{\beta})} && (16)\\
&= \frac{p(O_i^*,A_i,\alpha,\boldsymbol{\beta})}{p(O_i,A_i,\alpha,\boldsymbol{\beta})} && (17)\\
&= \frac{p(A_i|O_i^*,\alpha)p(O_i^*|\boldsymbol{\beta})}{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta})} && (18)\\
& && (19)
\end{aligned}
$$

where the factorization of the likelihood in Equation (18) is due to the fact that the acceptance of firm i only depends on what is offered to it and what is its preference, $p(A_i|O_i^*,\alpha)$; what is offered to i depends on the preferences of all countries, $p(O_i^*\ \boldsymbol{\beta})$.

If we plug in Equation (9) and Equation (7)

$$
\frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{\sum_{j:j\in O_i}\exp(\alpha' W_j)}{\sum_{j:j\in O_i}\exp(\alpha' W_j)+\exp(\alpha' W_{j^*})} \times \exp(\boldsymbol{\beta}_{j^*}' X_i) \tag{20}
$$

where j^* is the index of the newly sampled job. This is the case when the newly proposed job is not already offered, so it's added to the opportunity set.

When the newly proposed job is already offered, so it's removed from the opportunity set, we have

$$
\frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{\sum_{j:j\in O_i}\exp(\alpha' W_j)}{\sum_{j:j\in O_i}\exp(\alpha' W_j)-\exp(\alpha' W_{j^*})} \times -\exp(\boldsymbol{\beta}_{j^*}' X_i) \tag{21}
$$

B.2 Updating firms' parameters, α

Target distribution:

$$
p(\alpha|A,O,\boldsymbol{\beta}) = \frac{p(O,A,\alpha,\boldsymbol{\beta})}{p(A,O,\boldsymbol{\beta})} \tag{22}
$$

We propose a new α^* using a symmetric proposal distribution that sample α^* in a box whose boundary is $\alpha^* \pm \epsilon_\alpha$

Metropolis-Hasting acceptance ratio:

$$MH_\alpha = \frac{p(\alpha^*|A,O,\boldsymbol{\beta})}{p(\alpha|A,O,\boldsymbol{\beta})} = \frac{p(A_i|O_i,\alpha^*)p(O_i|\boldsymbol{\beta})}{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta})} \tag{23}$$

$$= \frac{p(A_i|O_i,\alpha^*)}{p(A_i|O_i,\alpha)} \tag{24}$$

where Equation (24) is due to the flat prior (so $\frac{p(\alpha^*)}{p(\alpha)} = 1$) and the symmetric proposal distribution (so $\frac{p(\alpha^*|\alpha)}{p(\alpha|\alpha^*)} = 1$)

If we plug in Equation (9),

$$MH_\alpha = \prod_i \left[\frac{\exp(\alpha^{*\prime} W_{a_i})}{\exp(\alpha' W_{a_i})} \times \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha^{*\prime} W_j)} \right] \tag{25}$$

$$= \prod_i \left[\exp(\epsilon'_\alpha W_{a_i}) \times \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha^{*\prime} W_j)} \right] \tag{26}$$

Finally, we log transform the MH acceptance ratio for numerical stability.

$$\log MH_\alpha = \sum_i \left[\epsilon'_\alpha W_{a_i} + \log\left(\sum_{j:j\in O_i} \exp(\alpha' W_j) \right) - \log\left(\sum_{j:j\in O_i} \exp(\alpha^{*\prime} W_j) \right) \right] \tag{27}$$

B.3 Updating countries' parameters, β

Target distribution:

$$p(\boldsymbol{\beta}|A,O,\alpha) = \frac{p(O,A,\alpha,\boldsymbol{\beta})}{p(A,O,\alpha)} \tag{28}$$

We propose a new $\boldsymbol{\beta}^*$ using a symmetric proposal distribution that sample $\boldsymbol{\beta}^*$ in a box with side length ϵ_β

Metropolis-Hasting acceptance ratio:

$$MH_\beta = \frac{p(\beta^*|A,O,\alpha)}{p(\beta|A,O,\alpha)} = \frac{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta}^*)}{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta})} \tag{29}$$

$$= \frac{p(O_i|\boldsymbol{\beta}^*)}{p(O_i|\boldsymbol{\beta})} \tag{30}$$

where Equation (29) is due to the flat prior on β and the symmetric proposal distribution.

We plug in Equation (7),

$$MH_\beta = \prod_i \left[\prod_{j \in O_j} \frac{\exp(\beta_j^{*\prime} X_i)}{\exp(\beta_j' X_i)} \times \prod_j \frac{1 + \exp(\beta_j^{*\prime} X_i)}{1 + \exp(\beta_j' X_i)} \right] \tag{31}$$

$$\log MH_\beta = \sum_i \left[\sum_{j \in O_i} \beta_j^{*\prime} X_i - \beta_j' X_i + \sum_j \log(1 + \exp(\beta_j^{*\prime} X_i)) - \log(1 + \exp(\beta_j' X_i)) \right] \tag{32}$$

In the MCMC implementation, since $\boldsymbol{\beta}$ is high dimensional, we conduct multiple block Metropolis Hastings, updating several β's at one time.

References

Eicher, T. S., L. Helfman, and A. Lenkoski (2012). Robust FDI determinants: Bayesian Model Averaging in the presence of selection bias. *Journal of Macroeconomics 34*(3), 637–651.

Logan, J. A. (1996). Opportunity and Choice in Socially Structured Labor Markets. *American Journal of Sociology 102*(1), 114.

Logan, J. A. (1998). Estimating Two-Sided Logit Models. *Sociological methodology 28*(1), 139–173.