[Two-Sided Matching Model]

by

[Anh Le]

Department of [Department of Political Science]
Duke University

Date: ____________________________

Approved:

____
[Michael Ward], Supervisor

____
[Eddy Malesky]

____
[Daniel Stegmueller]

____
[Peter Hoff]

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of [Department of Political Science]
in the Graduate School of Duke University
2018

Abstract

[Two-Sided Matching Model]

by

[Anh Le]

Department of [Department of Political Science]
Duke University

Date: ____________________________

Approved:

____
[Michael Ward], Supervisor

____
[Eddy Malesky]

____
[Daniel Stegmueller]

____
[Peter Hoff]

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of [Department of Political Science]
in the Graduate School of Duke University
2018

Contents

List of Tables

List of Figures

1
Introduction

Much of our economic and political life is governed by a two-sided matching process, in which two sets of actors evaluate each other's characteristics and voluntary form a match if deemed satisfactory. Marriage is a prominent example of such process. Others include the matching of firms and workers, federal judges and law clerks, countries and multinational firms who want to make an investment, or the *formateur* of a government and other minority parties.

Two sided matching market is substantively consequential because it involves scarce, one-of-a-kind "good", such as a life commitment to a marital partner, or a political allegiance in a government formation—certainly not the commodity like loaves of bread. It is also intellectually interesting because its different structure means that our understanding of the normal, one-sided market is inadequate to explain and analyze.

A two-sided market means that the market involves two disjoint set of actors, and one side decision affect the other side (Rysman, 2009) who evaluate each other's characteristics. For example, in a marriage market, both men and women evaluate one another's income, appearance, background, to consider a match. This stands in contrast with one-sided market in which only one side makes the judgment whether a transaction exist, e.g. a grocery shopper evaluates the plumpness of the grocer's tomato, while the grocer does not evaluate the shopper. This means that whether a transaction occurs depend not only on an individual's demand, but also on the individual's opportunity. Indeed, while we typically assume that in a grocery market, all the milk and tomato are always available to buy, and the sole decision is to buy how much, in a marriage market, the no less important question is what is available to the individual.

A matching market means that once a match is being made, the transaction is exclusive between the two individuals being matched. Once a match has been made, both sides are no longer available on the market. This results in strategic behaviors, i.e. delaying accepting an offer, falsifying preference, that causes market failures. For example, in many matching markets, such as the American physician market 1945-51, the American law graduate market, the elite Japanese university graduate markets, all suffer from the problem of offers being made aggressively earlier in order to scoop the best candidates from the market.

The study of two sided matching markets start within the market design subfield of economics, where scholars are chiefly concerned with explaining the market failures in the market, and how to design a mechanism to resolve such failure. In such studies, the preferences of the actors are given, and the goal is to find a "stable" matching.

The first work in this tradition is Gale and Shapely (1962) and Shapely and Shubik (1972). Some key proofs: - From any set of preference, a stable matching always exist - From a random matching, we will converge to a stable matching with probability one.

1.0.1 History of the game theory

(Gale and Shapley, 1962) was the first to study the one-to-one matching market, named the marriage market. In such a market, there are two finite and disjoint set of actors, men and women. Each man has preference over each woman, and vice versa. To each man and woman the option of remaining single is always available.

A matching μ is a function that matches a man with a woman. For convenience, we can say that if a man or woman decides to remain single, they are matched with themselves. We refer to $\mu(x)$ as the *mate* of x.

A matching can be improved upon in two ways. First, an individual can prefer to remain single than to be matched with his or her mate $\mu(x)$ under the current matching μ. Second, a man and a woman can improve a matching if they prefer to be with one another rather than whom they are currently matched with. Therefore, we define that a matching μ is stable if it cannot be improved upon by any individual or any pair of agent.

The first result from (Gale and Shapley, 1962) that is relevant to our purpose is that for any set of preference, there will always be a stable matching. This means that it is reasonable for us to assume that the process that generates the final match we observe in the data is one of agents maximizing their utilities. Crucially, this provides the justification for us modeling their choice using the familiar tools in the discrete choice literature.[1]

(Gale and Shapley, 1962) treats the case of many-to-one matching (which they call the "college admission" game) essentially the same. However, Roth and Sotomayor

[1] (Gale and Shapley, 1962) provides a constructive proof of stable matching. They describe the process, called *deferred acceptance*, that produces the stable matching. The process works as follows. In the first stage, every man proposes to his preferred mate. Every woman rejects all of her suitors except the one that she most prefers. However, she does not yet accept her favorite suitor (so far), but keep him along. In the second stage, every man that was rejected proposes to his second choice. Every woman then picks her favorite from the set of new proposers and the man she keeps along from the last round. The procedure continues until there is no longer any woman that is unmatched. The resulting match is stable because, throughout the process, every woman has received all the offers that would have been made to her, and she has chosen her favorite among all of those offers. If there is any other man that she would prefer to her current match, that man would not be available to her. Therefore, the final match cannot be further improved by any man or woman.

(1992) shows that a well defined many-to-one matching game is slightly different. However, hen generalize this model to many-to-one matching setting of firm and workers, all the striking results remain, but we do have to make the assumption of substitutability, meaning that firms treat workers as substitute, not complements. Formally, this means that firm never regret hiring a worker if another worker does not accept the offer. This is certainly not universally true: a football team strategizing for passing play may want to hire both a good passer and a good running back, and in the case one does not accept, will go to another strategy altogether.

So we know that a stable match always exist, meaning it is possible for agents to reach a final match that maximize their utilities. A central coordinator employing the *deferred acceptance* algorithm is guaranteed to reach this stable match. However, in a decentralized market without a central coordinator, would agents be able to reach this outcome by themselves? How likely is that this will happen. This matters for our empirical approach because we need to know whether the final outcome we observe is indeed a stable match, i.e. which would inform our estimation strategy.(Roth and Vate, 1990) shows that, starting from an arbitrary matching, the matching can converge to a stable matching with probability 1 if we allow random blocking pairs, i.e. a woman and a man that are not matched to one another but prefer each other to their current match, to match. In addition, (Adachi, 2003) shows that a random search process, in which man and woman randomly meet, evaluate, and decide to pair or not, will converge towards a stable match assuming that the search cost is negligible (specified as having a time discount = 1). This gives us confidence that the data we observed is close to a stable match, a fact that will guide us in our modeling strategy.

1.0.2 Empirical studies of two-sided matching market

Given the importance of many two sided matching markets, there have been many empirical studies that try to estimate the preference of participants in these markets. Below I discuss several examples, starting with the marriage and the labor market, where the two-sided matching literature started. Then I discuss the law graduate market and the FDI market, two markets that are relevant to politics, whose two-sided nature has not been fully appreciated and modeled accordingly.

Estimating preference is a difficult task, and throughout the discussion it will become apparent that most empirical approaches either 1) use survey to get the stated preference, not revealed preference, or 2) failing to consider the two-sided nature of the market in their estimation method, thus unable to distinguish between preference and opportunity. Our two-sided matching model aims to address both these shortcomings: it will estimate preference based on observed match, i.e. revealed preference, and will be able to distinguish between the effect of opportunity and choice.

Or they have demanding requirement on data, going further than just the simple match.

Labor market

(Abowd et al., 1999) is also estimating preference, disentangling the two-sided effect, but require salary data.

Recommender system for online two-sided markets

In recent years, many online market place proliferate such as AirBnB, dating site like eHarmony, or job board like o-lance. To save money, these sites must make recommendations to their users in a way that resembles a stable match so that their users are more engaged with the site instead of being dissatisfied with the match.[2]

Even though most of these algorithms are proprietary, there exists some academic publication addressing this problem. (Hitsch et al., 2010; Goswami et al., 2014) are two attempts using one-sided approach, ignoring any strategic component.

An interesting approach is (Tu et al., 2014) which uses the Latent Dirichlet allocation (LDA) model to uncover thethe latent types of the users and their preferences based on their observed match on an online dating platform. In the original application of LDA model on topic modeling, each document is a mixture of latent topics, and each topic is a distribution over words. In this application, each user is a mixture of latent "types," and each type is a distribution signifying relative preference over mates' feature. For example, the "outdoor type" may have higher preference for athleticism or dog ownership over other traits of their mates.

However, these kinds of models only work because they either flat out ignore the two sided nature, or have access to data of users making multiple actions (i.e. reaching out to multiple partner) instead of just the final match. However, considering that these sites typically put a limit on the number of outreach you can have in a given time, or considering the simple time cost of making an outreach, we cannot ignore the strategic component of an user only reaching out to users that they think they have a chance with. So the preference they estimate is already confounded by the type of opportunity available. While these models have decent predictive performance when the pool of applicants are static, they cannot predict well the change in matches when there is a change in the opportunity pool.

[2] To clarify, the term "recommender system" typically refers to how platform make recommendation to users based on the reviews of users like them. This is not what we are discussing here. Even though there is also a latent preference we want to uncover in that system, there is not a two sides choosing one another there.

Law graduate market

In the United States, graduates at top law schools vie for the best federal clerkship. These temporary, one-to-two-year position are the launching pad for Supreme Court clerkship, prestigious teaching jobs or positions at top law firms. On the other hand, federal judges also compete for the best law graduates, who help reduce the judges' workload, from copy-editing, to research, and even drafting the opinions (Gulati and Posner, 2016; Posner, 2001, 795). As clerkship having an outsized influence on law graduates, studying the law market has important implication for the polarization of the judicial branch.

This market has long been recognized as a classic case of a two-sided market. Clerks look for positions with not only prestige and connection but also comfortable living situation (Posner, 2001). Judges select law graduates based on not only academic credentials but also, some afraid, ideology, gender, and race (Slotnick, 1984). This market also suffers from strategic behavior emblematic of a matching market, such as offers being made as early as two years before the clerkship start date, and with a short time to accept (Posner, 2001; Posner et al., 2007).

The two sided nature of the market makes it difficult to quantitatively study the preference of the market participants. One approach has been to survey clerks and judges directly (Peppers et al., 2008). However, this only allows researchers to measure stated preference, which is unlikely to be accurate when it comes to dimensions that we care about most such as discrimination based on ideology, gender, or race.

Other quantitative approaches circumvent this problem by using observed hiring outcome to study preference. However, no study so far has properly taken into account the two-sided nature of the market, thus confusing the effects of opportunity and preference. For example, Bonica et al. (2017) use political contribution data

(DIME dataset) to measure political ideology, then find the correlation between the hiring judges' ideology and their clerks' ideology. This approach does not take into account the pool of clerk applicants, which leads to curious conclusion such as that conservative judges hire more liberal clerks than conservative clerks (Bonica et al., 2017, 31). This curious finding has a potentially simple explanation: the pool of top law graduates tend to be overwhelmingly liberal, and conservative judges may not have much choice. Despite this drawback, the authors proceed to measure judges' ideology by taking the average of their clerks' ideology. Without taking the pool of applicant into account, we may wrongly conclude that conservative judges are more liberal than they are.

Similarly, Rozema and Peng (2016) models the process as a discrete choice problem, in which clerks are differentiated products that the Supreme Court justices pick in order to maximize their utilities. However, it doesn't consider what clerks think about the offer. Rozema and Peng (2016) can make this assumption because they focus on Supreme Court clerkship, whose unparalleled prestige ensures that any offer made will be accepted. However, if we want to extend the model to the broader market of federal clerkship, such assumption is untenable.

Government formation

Besides election, government formation is probably the most consequential political process in determining the government that people are subject to. Most extant studies of government formation has either been game theoretic or thick description, "inside-the-Beltway" narrative of what happened. Potential advances be made when we consider government formation as a two-sided matching market, with the *formateur* on one side and all other minority parties on the other.[3] There are empirical

[3] The *formateur* party could be the one with the procedural power to set up the coalition, e.g. the incumbent party, or the largest party in established coalitions.

studies of policy-seeking politicians on portfolio allocation, but it's broad statement like this agragrian party will take the agriculture ministry if the party is in the coalition and that there is such a party.

An empirical study of government formation as a two-sided matching market complements the game theory literature that model politicians as policy-seeking (as opposed to office seeking). When politicians are policy-seeking, they have positions that can be modeled as their characteristics, and parties choosing one another to form coalition based on their policy positions akin to men and women choosing one another to form a marriage based on height or income Laver (1998).[4] As the game theory literature suggests, ideologically compact coalitions are more valuable because they entail fewer costs in policy compromises. With the empirical matching model, we can test if parties do indeed prefer others closer to themselves ideologically.

In addition, the two sided matching approach has the advantage of studying multidimensional policy space. It works quite naturally by considering a party's positions on various policies as their many characteristics.

The FDI market

Estimating the preference: democracy and FDI attraction, survey, comparing FDI and equity, type of FDI (greenfield vs brownfield).

Also demand for the type of FDI.

[4] In contrast, when politicians are office seeking, the only coin of the realm is the number of legislative seats that a party controls. It determines the inclusion of the party in the government, its portfolio allocation, etc. In this framework, concept like power indices and dominant parties is all about which coalition parties can join to turn it into a winning / losing coalition.

2

Two-Sided Matching Model

Here I present a behavioral model of the two-sided matching market, focusing on the case of many-to-one matching. For easier exposition, throughout the chapter I will use the example of the labor market, where many workers can be matched to one firm, to demonstrate the features of this model.

In this model, there are two separate sets of actors: firms and workers. There are two sub-components of the model, happening sequentially. In the first stage, each firm evaluates each worker in the sample, then decide whether to extend an offer or not. In the second stage,

meaning that the model rests on the behavioral assumption that people make choices by maximizing their utilities. The parameters estimated will be the preference of people, and can be interpreted as how much they value different features. Since utility is unit-less, it won't have a scale, we only care about the relative values of different parameters. Cite McFadden here

- one sided model won't be able to distinguish the effect between preference and opportunity - follows the tradition of discrete choice, in which actors choose from a set of finite and discrete alternatives (firms choosing from workers, workers choosing offers). Here we do make the assumption that the sample is representative of the population, i.e. the workers in the sample are also the workers that are presented to firm

This section describes the two-sided matching model, a behavioral model in the sense that actors which is the behavioincluding the utilities of countries' officials and of MNCs. Then, the matching process is a natural consequence of actors' choosing the best option available to them.

This model is a parametric version of the matching game by Roth

Describe the model: there are two sets of actors, one is employers, the other is employee. Employer consider every single worker, determine the utility of hiring him over not hiring him. This makes the assumption that the hiring decision does not depend on whether other candidates will accept the offer (substitutability). It also assumes that firm can theoretically hire all workers in the pool if they accept (assuming that there is no quota).

IIA the ratio of the probabilities of choosing between two alternatives is independent of the attributes of all other alternatives. (This applies to the worker side).

Stress that this is only one way to reach stable matching

For the case of FDI, if we assume that all firms have the same preference, then even if we don't include the firms who decide not to invest, that's fine.

2.0.1 Officials' Utility

Following Logan (1998), we consider the utility function of two set of actors, firms and workers. For firm j, the utility of hiring worker i is:

This section describes the behavioral model of many-to-one matching. To aid the exposition, as an example for many-to-one matching I will use the labor market where many workers can be matched with one firm.

We assume that the matching process in the labor market happens in two stage.

In the first stage, each firm evaluates each worker in the sample, then decide whether to hire that worker or not. At the end of this stage, each worker has a set of offers from firms, which we call his *opportunity set*. In the second stage, each worker evaluates the firms in his opportunity set and chooses the firm that he likes best. This constitutes a final, observed match between a worker and a firm. This is a many-to-one matching problem because a firm can make offers to multiple workers, none, some, or all of which can be accepted by workers.

If we assume that firms and workers are utility-maximizing agents, at the end of this process, no firm or worker would voluntarily change their final matches. This property is called *stability* in the game theoretic two-sided matching literature. We want the model to have this property because real life matching market tends to be stable as well. Indeed, (Roth and Sotomayor, 1992) show that for any given set of preferences, a stable match always exist. Furthermore, (Roth and Vate, 1990) and (Adachi, 2003) show that a decentralized market with agents making independent, utility-maximizing decisions can also reach a stable match by itself.

This stability property does not implies that the matches will not change. Indeed, if actors' preference shift, their characteristics change, or new actors enter the market, the matches will also change as a result of actors' recalculating their utility and adjust their decisions. Therefore, since we are estimating the actors' preference using only a snapshot of matching market, we are making the assumption that on a systemic level, the average characteristics of the actors and their preference remain sufficiently static for our estimates to be meaningful.

2.0.2 The decision making of firms

A firm j's decision on whether to hire worker i rests on two utility functions. First, firm j's utility for hiring worker i is: ¿¿¿¿¿¿¿ 5a25e367faf0028957b544b0025571f85941a03e

$$U_j(i) = \boldsymbol{\beta}_j' X_i + \epsilon_{1ij} \tag{2.1}$$

where

β_j is a vector of firm j's preference for worker characteristics

x_i is a vector of worker i's measured values on those characteristics

ϵ_{1ij} is the unobserved component that influences firm j's utility

On the other hand, the utility of not hiring worker i is:

$$U_j(\neg i) = b_j + \epsilon_{0ij} \tag{2.2}$$

where

b_j is the baseline utility of firm j

ϵ_{0ij} is the unobserved component that influences firm j's utility

For each worker i, firm j will make an offer to hire if $U_j(i) > U_j(\neg i)$. Relevant worker characteristics (i.e. X_i) that a firm may consider are age, education, or experience. The corresponding β's represent the firm's preference for these characteristics.

We make several assumptions that are standard in the discrete choice literature. First, we assume a linear utility function. Second, we assume that the error terms $\epsilon_{1ij}, \epsilon_{0ij}$ are uncorrelated. This assumption is most likely not true: unobserved factors in one firm's utility are likely to share some components with unobserved factors in another firm's utility, and thus correlated. The hope is that we have modeled the observed portion of firm's utility sufficiently well that the remaining unobserved factors are close to white noise. In any case, this issue afflicts any application of discrete choice models and is not unique to our case.[1]

Third, we assume that the as error terms $\epsilon_{1ij}, \epsilon_{0ij}$ follow the Gumbel distribution, which has a slightly fatter tail the normal distribution allowing for slightly more extreme variation in the unobserved utility. In practice, the difference between using Gumbel and independent normal error terms is small (Train, 2009). The choice of the Gumbel distribution is largely motivated by its convenience since we can derive the probability of firm j making an offer to worker i as the familiar binomial logit form:

Also discuss IIA here

$$\begin{aligned} Pr(o_{ij} = 1) &= Pr(U_j(i) > U_j(\neg i)) && (2.3) \\ &= Pr(\epsilon_{0ij} - \epsilon_{1ij} < \boldsymbol{\beta}_j' X_i - b_j) && (2.4) \\ &= \frac{\exp(\boldsymbol{\beta}_j' X_i)}{1 + \exp(\boldsymbol{\beta}_j' X_i)} && (2.5) \end{aligned}$$

$$\begin{aligned} p(O_i|\boldsymbol{\beta}) &= \prod_{j \in O_i} p(o_{ij} = 1|\boldsymbol{\beta}) \prod_{j \notin O_i} p(o_{ij} = 0|\boldsymbol{\beta}) && (2.6) \\ &= \prod_{j \in O_i} \frac{\exp(\boldsymbol{\beta_j'} X_i)}{1 + \exp(\boldsymbol{\beta_j'} X_i)} \prod_{j \notin O_i} \frac{1}{1 + \exp(\boldsymbol{\beta_j'} X_i)} && (2.7) \end{aligned}$$

In our observed data, since we only observe the final matching of firms and countries, this opportunity set is unobserved. As will discuss, we use the Metropolis-Hastings algorithm to approximate the posterior distribution of the opportunity set.

2.0.3 Firms' utility

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

[1] The discrete choice model has developed solutions for correlated error structure, such as nested logit, probit, and mixed logit, that can be applied here if we suspect that firms' unobserved utility is strong correlated.

$$V_i(j) = \alpha' W_j + v_{ij} \tag{2.8}$$

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 welcome it to invest and chooses the country that brings the highest utility. This choice of firms concludes the matching process, resulting in the observed final match between a firm and a country in our data.

In our model, relevant country characteristics can be: 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 takes the multinomial logit form (Cameron and Trivedi, 2005):

$$p(A_i = a_i | O_i, \alpha_i) = \frac{\exp(\alpha' W_{a_i})}{\sum_{j:j \in O_i} \exp(\alpha' W_j)} \tag{2.9}$$

2.1 Model Estimation

Because the opportunity set is unobserved, we have to use MCMC to estimate it.

$$U_j(i) = \boldsymbol{\beta}_j' X_i + \epsilon_{1ij} U_j(\neg i) \qquad = b_j + \epsilon_{0ij} V_i(j) = \alpha' W_j + v_{ij} \tag{2.10}$$

$$Pr(o_{ij} = 1) = Pr(U_j(i) > U_j(\neg i)) \tag{2.11}$$

$$= Pr(\epsilon_{0ij} - \epsilon_{1ij} < \boldsymbol{\beta}'_j X_i - b_j) \tag{2.12}$$

$$= \frac{\exp(\boldsymbol{\beta}'_j X_i)}{1 + \exp(\boldsymbol{\beta}'_j X_i)} \tag{2.13}$$

$$p(O_i|\boldsymbol{\beta}) = \prod_{j \in O_i} p(o_{ij} = 1|\boldsymbol{\beta}) \prod_{j \notin O_i} p(o_{ij} = 0|\boldsymbol{\beta}) \tag{2.14}$$

$$= \prod_{j \in O_i} \frac{\exp(\boldsymbol{\beta}'_{\boldsymbol{j}} X_i)}{1 + \exp(\boldsymbol{\beta}'_{\boldsymbol{j}} X_i)} \prod_{j \notin O_i} \frac{1}{1 + \exp(\boldsymbol{\beta}'_{\boldsymbol{j}} X_i)} \tag{2.15}$$

$$p(A_i = a_i|O_i, \alpha_i) = \frac{\exp(\alpha' W_{a_i})}{\sum\limits_{j: j \in O_i} \exp(\alpha' W_j)} \tag{2.16}$$

Joint likelihood:

$$p(O, A, \alpha, \beta, \mu_\beta, \tau_\beta) = p(A|O, \alpha)p(O|\beta)p(\alpha)p(\beta|\mu_\beta, \tau_\beta)p(\mu_\beta)p(\tau_\beta) \tag{2.17}$$

2.1.1 Updating the opportunity set

Target distribution for a firm i

$$p(O_i|A_i, \alpha, \boldsymbol{\beta}) = \frac{p(O_i, A_i, \alpha, \boldsymbol{\beta})}{p(A_i, \alpha, \boldsymbol{\beta})} \tag{2.18}$$

$$MH_O = \frac{p(O_i^*|A_i, \alpha, \boldsymbol{\beta})}{p(O_i|A_i, \alpha, \boldsymbol{\beta})} = \frac{p(O_i^*, A_i, \alpha, \boldsymbol{\beta})}{p(A_i, \alpha, \boldsymbol{\beta})} \times \frac{p(A_i, \alpha, \boldsymbol{\beta})}{p(O_i, A_i, \alpha, \boldsymbol{\beta})} \tag{2.19}$$

$$= \frac{p(O_i^*, A_i, \alpha, \boldsymbol{\beta})}{p(O_i, A_i, \alpha, \boldsymbol{\beta})} \tag{2.20}$$

$$= \frac{p(A_i|O_i^*, \alpha)p(O_i^*|\boldsymbol{\beta})}{p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta})} \tag{2.21}$$

$$\tag{2.22}$$

where the factorization of the likelihood in (2.21) 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 (2.16) and (2.15)

$$\frac{p(O_i^*|A_i, \alpha, \boldsymbol{\beta})}{p(O_i|A_i, \alpha, \boldsymbol{\beta})} = \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j) + \exp(\alpha' W_{j*})} \times \exp(\boldsymbol{\beta}'_{j*} X_i) \tag{2.23}$$

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

$$\frac{p(O_i^*|A_i, \alpha, \boldsymbol{\beta})}{p(O_i|A_i, \alpha, \boldsymbol{\beta})} = \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j) - \exp(\alpha' W_{j*})} \times \exp(-\boldsymbol{\beta}'_{j*} X_i) \tag{2.24}$$

2.1.2 Updating firms' parameters, α

Target distribution:

$$p(\alpha|A, O, \boldsymbol{\beta}) = \frac{p(O, A, \alpha, \boldsymbol{\beta})}{p(A, O, \boldsymbol{\beta})} \tag{2.25}$$

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:

$$MH_\alpha = \frac{p(\alpha^*|A, O, \boldsymbol{\beta})}{p(\alpha|A, O, \boldsymbol{\beta})} = \frac{p(A_i|O_i, \alpha^*)p(O_i|\boldsymbol{\beta})}{p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta})} \tag{2.26}$$

$$= \frac{p(A_i|O_i, \alpha^*)}{p(A_i|O_i, \alpha)} \tag{2.27}$$

where (2.27) 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 (2.16),

$$MH_\alpha = \prod_i \left[\frac{\exp(\alpha^{*\prime} W_{a_i})}{\exp(\alpha' W_{a_i})} \times \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha^{*\prime} W_j)} \right] \tag{2.28}$$

$$= \prod_i \left[\exp(\epsilon'_\alpha W_{a_i}) \times \frac{\sum\limits_{j:j\in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i} \exp(\alpha^{*\prime} W_j)} \right] \tag{2.29}$$

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

$$\log MH_\alpha = \sum_i \left[\epsilon'_\alpha W_{a_i} + \log\left(\sum_{j:j\in O_i} \exp(\alpha' W_j) \right) - \log\left(\sum_{j:j\in O_i} \exp(\alpha^{*\prime} W_j) \right) \right] \tag{2.30}$$

2.1.3 Updating countries' parameters, β

Target distribution:

$$p(\boldsymbol{\beta}|A, O, \alpha) = \frac{p(O, A, \alpha, \boldsymbol{\beta})}{p(A, O, \alpha)} \tag{2.31}$$

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:

$$MH_\beta = \frac{p(\beta^*|A, O, \alpha)}{p(\beta|A, O, \alpha)} = \frac{p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta}^*)p(\boldsymbol{\beta}^*|\mu_\beta, \tau_\beta)}{p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta})p(\boldsymbol{\beta}|\mu_\beta, \tau_\beta)} \tag{2.32}$$

$$= \frac{p(O_i|\boldsymbol{\beta}^*)p(\boldsymbol{\beta}^*|\mu_\beta, \tau_\beta)}{p(O_i|\boldsymbol{\beta})p(\boldsymbol{\beta}|\mu_\beta, \tau_\beta)} \tag{2.33}$$

where (2.32) is due to the flat prior on β and the symmetric proposal distribution.

We plug in (2.15),

$$MH_\beta = \prod_i \left[\prod_{j \in O_i} \frac{\exp(\beta_j^{*\prime} X_i)}{\exp(\beta_j^{\prime} X_i)} \times \prod_j \frac{1 + \exp(\beta_j^{*\prime} X_i)}{1 + \exp(\beta_j^{\prime} X_i)} \right] \times \frac{MVN(\boldsymbol{\beta}^* | \mu_\beta, \tau_\beta)}{MVN(\boldsymbol{\beta} | \mu_\beta, \tau_\beta)} \tag{2.34}$$

$$\log MH_\beta = \sum_i \left[\sum_{j \in O_i} \beta_j^{*\prime} X_i - \beta_j^{\prime} X_i + \sum_j \log(1 + \exp(\beta_j^{*\prime} X_i)) - \log(1 + \exp(\beta_j^{\prime} X_i)) \right] \tag{2.35}$$

$$+ \log MVN(\boldsymbol{\beta}^* | \mu_\beta, \tau_\beta) - \log MVN(\boldsymbol{\beta} | \mu_\beta, \tau_\beta)$$

In the MCMC implementation, since $\boldsymbol{\beta}$ is high dimensional, in each step, we randomly update several β's at one time.

2.1.4 Update μ_β, τ_β

Similar to a multivariate normal model, where β is the "data".

$$p(\mu_\beta) \sim MVN(\mu_0, \Sigma_0) \tag{2.36}$$

$$p(\mu_\beta | \beta, \tau_\beta) \sim MVN(m, V) \text{ where} \tag{2.37}$$

$$V = (\Sigma_0^{-1} + n\tau_\beta)^{-1} \tag{2.38}$$

$$m = (\Sigma_0^{-1} + n\tau_\beta)^{-1}(\Sigma_0^{-1}\mu_0 + n\tau_\beta\bar{\beta}) \tag{2.39}$$

Bibliography

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999), "High wage workers and high wage firms," *Econometrica*, 67, 251–333.

Adachi, H. (2003), "A search model of two-sided matching under nontransferable utility," *Journal of Economic Theory*, 113, 182–198.

Bonica, A., Chilton, A. S., Goldin, J., Rozema, K., and Sen, M. (2017), "The political ideologies of law clerks," *American Law and Economics Review*, 19, 96–128.

Cameron, A. and Trivedi, P. (2005), *Microeconometrics: methods and applications*, Cambridge University Press, Cambridge.

Gale, D. and Shapley, L. S. (1962), "College Admissions and the Stability of Marrige," *The American Mathematical Monthly*, 69, 9–15.

Goswami, A., Hedayati, F., and Mohapatra, P. (2014), "Recommendation systems for markets with two sided preferences," *Proceedings - 2014 13th International Conference on Machine Learning and Applications, ICMLA 2014*, pp. 282–287.

Gulati, M. and Posner, R. A. (2016), "The management of staff by federal court of appeals judges," *Vanderbilt Law Review*, 69, 479–497.

Hitsch, G. J., Hortaçsu, A., and Ariely, D. (2010), "What makes you click?-Mate preferences in online dating," *Quantitative Marketing and Economics*, pp. 1–35.

Laver, M. J. (1998), "Models of Government Formation," *Annual Review of Political Science*, 1, 1–25.

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

Peppers, T. C., Giles, M. W., and Tainer-Parkins, B. (2008), "Inside Judicial Chambers: How Federal District Court Judges Select and Use Their Law Clerks," *Albany Law Review*, 71, 8–23.

Posner, R. A. (2001), "The Market for Federal Judicial Law Clerks," *The University of Chicago Law Review*, 68, 793–902.

Posner, R. A., Avery, C., Jolls, C., and Roth, A. (2007), "The New Market for Federal Judicial Law Clerks," *The University of Chicago Law Review*, 74, 447.

Roth, A. E. and Sotomayor, M. (1992), "Two-sided matching," in *Handbook of Game Theory*, eds. S. Hart and R. Aumann, vol. 1, chap. 16, pp. 486–541, Elsevier Science Publishers, 1st edn.

Roth, A. E. and Vate, J. H. V. (1990), "Random Paths to Stability in Two-Sided Matching," *Econometrica*, 58, 1475.

Rozema, K. and Peng, S. (2016), "The Role of External Referrals in Hiring: Evidence from Judicial Law Clerks," *Available at SSRN 2822952*.

Rysman, M. (2009), "The Economics of Two-Sided Markets," *Journal of Economic Perspectives*, 23, 125–143.

Slotnick, E. E. (1984), "The paths to the federal bench: gender, race and judicial recruitment variation," *Judicature*, 67, 371–388.

Train, K. E. (2009), *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

Tu, K., Ribeiro, B., Jiang, H., Wang, X., Jensen, D., Liu, B., and Towsley, D. (2014), "Online Dating Recommendations: Matching Markets and Learning Preferences," .