Two-Sided Matching Model

by

Anh Le

Department of Department of Political Science
Duke University

Date: ____________________________
Approved:

____
Michael Ward, Co-Supervisor

____
Eddy Malesky, Co-Supervisor

____
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, Co-Supervisor

____
Eddy Malesky, Co-Supervisor

____
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 **vi**

List of Figures **vii**

List of Abbreviations and Symbols **viii**

1 Introduction **1**

1.1 Game theory models of matching markets 2

1.2 Empirical models of matching markets 4

1.2.1 US federal clerkship market 5

1.2.2 The market for forming a coalition government 7

1.2.3 The FDI market 8

1.2.4 Recommender system for online two-sided markets 8

1.2.5 Two-sided models for the labor and marriage markets 9

1.3 Conclusion 9

2 Two-Sided Matching Model **10**

2.1 Modeling firms' decision making 11

2.2 Modeling workers' decision making 14

2.3 Model estimation 16

2.3.1 Estimating the model using Metropolis-Hastings 17

2.3.2 Posterior of the opportunity set $p(O|A, \alpha, \beta)$ 18

2.3.3 Posterior of firms' preference $p(\alpha|A, O, \beta)$ 20

2.3.4 Posterior of workers' preference $p(\beta|A, O, \alpha)$ 20

2.3.5 Posterior of β's hyperparameters μ_β, τ_β 21

2.4 Results for US labor data 21

3 Simulation results **22**

4 FDI **23**

4.1 Introduction 23

4.2 Three Issues in the Literature of FDI's Political Determinants 25

4.2.1 Measuring MNCs' Activities 25

4.2.2 Estimating Countries' Demand for FDI 27

4.2.3 Estimating Countries' Preferences for FDI's Technological Intensity 29

4.3 Applying the Two-Sided Matching Model to Japanese MNCs 30

4.4 Conclusion 33

A Derivation of the Metropolis-Hastings Acceptance Ratio **34**

A.0.1 Opportunity sets O 34

A.0.2 Workers' parameters, α 35

A.0.3 Firms' parameters, β 36

List of Tables

List of Figures

4.1 China's FDI Ownership Restriction, as coded in **?**. Prior to 1986, FDI in China was limited to few experimental Special Economic Zones, and thus not mentioned in US Investment Reports. The sharp spike in 1988 also does not seem to correspond to any actual change in policy, and likely another artifact of reporting. (See **?** for a historical overview of China's FDI policy.) 29

List of Abbreviations and Symbols

Abbreviations

EM	Expectation Maximization.
FDI	Foreign Direct Investment.
MCMC	Markov Chain Monte Carlo.
MH	Metropolis-Hastings.
MLE	Maximum Likelihood Estimation.
MNC	Multinational Corporation.
MVN	Multivariate Normal.

1

Introduction

Much of our social, economic, and political life is governed by two-sided matching markets. In these matching markets, actors from two disjoint sets evaluate the characteristics of someone on the other side and voluntarily form a match if both deem each other satisfactory.[1] Marriage is a prominent example of such matching process. Others include the matching between firms and workers, federal judges and law clerks, the *formateur* of a coalition government and other minority parties, or countries and multinational corporations (MNCs) that are looking for a location to invest.

Two-sided matching market is substantively consequential because it often involves scarce, indivisible goods, such as life commitment to a marital partner or political allegiance in a coalition government. It is also intellectually interesting because the market outcome depends on the actions of both sides, demanding a different analytical approach from what's used for one-sided markets.

This chapter will proceed as follows. First, I discuss the game theory literature

[1] Throughout the dissertation, I use "two-sided matching market" and "matching market" interchangeably. On the other hand, note that a two-sided market is not necessarily a matching market (**?**).

of two-sided matching models, where much of the terminology and insight originate. I will highlight key results that are relevant to our goal of estimating actors' preference in matching markets. Second, I examine existing empirical studies of matching markets. Where existing studies have not taken into account the market's two-sided nature, I discuss how doing so can improve our understanding of the subject area. Where existing studies do model the two-sided dynamics, I discuss how they may or may not be used to study subjects that political scientists are interested in.

1.1 Game theory models of matching markets

? was the first to study the matching market, using marriage as an example. In this market, there are two finite and disjoint sets of actors: men and women. Each man has preferences over the women, and vice versa. Each man's preference can be represented as an ordered list, ranking each woman based on how much he likes her.

The outcome of this market is a set of marriages, with none of some of people prefer to remain single. We call such a set of marriages a *matching* μ, which is a one-to-one function that matches a man with a woman. We refer to $\mu(x)$ as the *mate* of x. For convenience, we say that if an individual decides to remain single, they are matched with themselves.

We define a matching μ as *stable* if it cannot be improved by any individual or any pair of agents. A matching can be improved in two ways. First, an individual may 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 may prefer to be with one another rather than whom they are currently matched with. Therefore, if a matching is stable, no one has a better option than their current situation.

The first key result from the game theory literature is that for any set of preference, there always exists a stable matching (**?**). The proof is constructive, describing the "deferred acceptance" procedure that is guaranteed to produce a stable matching.[2] This result provides some justifications for us to assume that the matching we observe in real matching market is stable, and that the agents' utility cannot be further improved. Our empirical model of two-sided matching markets thus needs to describe a process that produces a stable matching.

While a central coordinator employing the *deferred acceptance* algorithm is guaranteed to come up with a stable matching, it is unclear whether decentralized markets, such as the labor market or the FDI market, would be able to reach this outcome by themselves.[3] The second key result from the game theory literature is that stable matching in decentralized matching market is indeed possible, even likely. For example, **?** show that, starting from an arbitrary matching, the market can converge to a stable matching with probability 1 if we allow random blocking pairs, i.e. two individuals that are not matched but prefer each other to their current match, to break off and form their own match. In addition, **?** shows that a random search process, in which pairs of man and woman randomly meet and decide whether each other is better than their current mates, will converge towards a stable matching if the search cost is negligible.[4] These results further suggest that the matching we

[2] The "deferred acceptance" procedure 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 (so far) favorite suitor, but keeps him along. In the second stage, every man that was rejected in the previous round 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 previous round. The procedure continues until there is no longer any woman that is unmatched, at which point women finally accept their current favorite choices. (This procedure is called *deferred acceptance* to capture the fact that women defer accepting her favorite choice until the last round in case better options become available.) The resulting match is stable because, throughout the procedure, 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 were any other man that she would prefer to her current match, that man would not have been available to her. Therefore, the final match cannot be further improved by any man or woman.

[3] The deferred acceptance procedure was used in the market for US medical residency with enthusiastic participation from medical students and hospitals. The high participation rate indicates that the matching produced is stable enough to entice students and hospitals away from arranging their own matches outside of the centralized market.

[4] In this model, searching has a time cost. Thus, negligible search cost is modeled as agents having a time discount close to 1.

observe in decentralized markets is likely stable. Therefore, our empirical model of matching markets to describe a process that produces a stable matching.

The third key result is that all conclusions regarding the one-to-one matching market (e.g. marriage) generalize to the many-to-one matching market (e.g. college admission, labor market), albeit requiring additional assumptions **?**. One important assumption is that firms treat workers as substitutes, not complements. In other words, firms never regret hiring a worker even if another worker is no longer available. Therefore, when we conduct empirical analysis of many-to-one markets, we should focus on markets where agents have such "substitutable preference." Otherwise, a stable matching is not guaranteed, agents' utility functions are interdependent, and it becomes unclear what kind of matching process our empirical model should approximate.

1.2 Empirical models of matching markets

The game theory literature takes the agents' preference as given and proves the existence of a stable matching. In contrast, empirical models of matching markets takes the observed matching as given and attempt to estimate the agents' preference.

Unfortunately, most extant empirical models fail to adequately account for the structure of a two-sided matching market. Often, researchers simply analyze the market from one side, e.g. estimating a firm's preference by looking at the type of workers it hires. This approach does not take into account the fact that a match depends not only on the agent's preference but also his opportunity. For example, a farm may prefer to hire highly-educated workers but cannot do so because highly-educated workers do not want to work on farms. Modeling this interaction between preference and opportunity is the key contribution of this dissertation.

Alternatively, some researchers measure agents' preferences by surveying them directly (**??**). While this approach circumvents the need to disentangle preference and opportunity, it can only measure agents' *stated* preference. In addition, such surveys require a high data collection effort while data on final matching (e.g. married couples, workers' current job, country location of MNCs) are widely available. This dissertation aims to make use of such available data to estimate agents' *revealed* preference.

Below I discuss existing empirical models of matching markets. First, I discuss two markets of interest to political scientists: the US federal clerkship market and the "market" for forming a coalition government. Researchers in both subject areas have not approached the problem with an empirical model that adequately captures its two-sided dynamics.

Second, I examine models from other disciplines that do take into account the two-sided dynamics of matching markets. I start with machine learning models applied to online marketplaces and dating sites. Then, I discuss the statistical models of the labor market (**?**) and the marriage market (**?**), which are most relevant to our goal of estimating agents' preference based on observed match data. These statistical models serve as the foundation of my empirical approach.

1.2.1 US federal clerkship market

In the US, graduates at top law schools vie for the best federal clerkships every year. These temporary, one-to-two-year positions are the launching pad for Supreme Court clerkships, prestigious teaching jobs, or employment at top law firms. On the other side, federal judges also compete for the best law graduates, who help reduce the judges' workload from copy-editing to drafting opinions (**??**). Because the first clerkship tends to have an outsized ideological influence on law graduates, this matching market has important implications for the polarization of the judicial branch (**??**).

The market for US federal clerkship has been noted as a classic case of a two-sided market. Clerks look for positions that provide not only prestige and connection but also comfortable quality of life (**?**). Judges select law graduates based on not only academic credentials but also, some argue, ideology, gender, and race (**?**). This market also suffers from strategic behavior emblematic of a matching market, such as offers being made aggressively early and with a short time to accept (**??**).

One approach to estimating the preference of agents in this market is to survey clerks and judges directly (**?**). However, as discussed, this approach only measures stated preference, which is likely to suffer from social desirability bias when it comes to dimensions that we care about most such as matching based on ideology, gender, or race.

Other approaches estimate revealed preference by using observed hiring outcome. However, no existing study has properly taken into account the two-sided nature of the market, thus confusing the effects of preference and opportunity. For example, **?** use political contribution data (DIME dataset) to measure political ideology, then correlate the ideology of the hiring judge and the ideology of his clerks. This approach does not take into account the pool of applicants, leading to conclusions such as conservative judges hire more liberal clerks than conservative clerks (**?**, 31). This curious finding has a potentially simple explanation: the pool of top law graduates tend to be overwhelmingly liberal, leaving conservative judges with no choice. Despite this issue, the authors proceed to measure judges' ideology by taking the average of their clerks' ideology. Without taking the pool of applicants into account, they may wrongly conclude that conservative judges are more liberal than they actually are.

In another approach, **?** model the process as a discrete choice problem, in which clerks are differentiated products that Supreme Court justices select to maximize their utilities. Their model does not consider what clerks think about the offer because of their focus on Supreme Court clerkships, 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.

1.2.2 The market for forming a coalition government

Besides election, government formation is the most consequential political process in determining which government people are subject to. Most extant studies of government formation are either game theoretic models or thick, "inside-the-Beltway" narratives. Potential advances can be made if we consider government formation as a many-to-one matching market, with the *formateur* party on one side and other minority parties on the other.[5]

A two-sided matching model of government formation would complement the game theory literature that models politicians as policy-seeking (as opposed to office-seeking) (**?**). When politicians are policy-seeking, parties have policy positions that can be modeled as their characteristics. Then, parties choose one another to form a coalition based on their policy positions, akin to men and women choosing one another to form a marriage based on their height or income.[6] As the game theory literature suggests, ideologically compact coalitions are more valuable because they entail a smaller cost in terms of policy compromises (**?**). With the empirical matching model, we can test if parties do indeed prefer others that are ideologically close to themselves.

In addition, an advantage of the two-sided matching approach is its ability to consider multidimensional policy spaces. By considering a party's positions on various policies as their covariates, we would be able to estimate parties' relative preference

[5] 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.

[6] 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 both the inclusion of the party in the government and its portfolio allocation. In this framework, concepts like power indices and dominant parties are all about how parties can bring its controlled seats to a coalition to turn it into a winning coalition.

for ideological proximity across policy dimensions.

1.2.3 The FDI market

To be introduced here, or kept until its own empirical chapter?

1.2.4 Recommender system for online two-sided markets

In recent years, the Internet underwent a proliferation of two-sided matching markets such as online marketplaces (e.g. AirBnB), dating sites (e.g. eHarmony), or job board (e.g. Elance). To help their users discover a match quicker, these sites often build a recommender system that suggests potential matches.[7] To maximize user engagement and profitability, these sites are incentivized to make recommendations that resemble a stable matching so that their users get the best match possible. And fo find the stable matching, they have to first estimate the preferences of their users.

While most of these algorithms are proprietary, some academic publications have addressed this problem. An interesting approach is the paper by **?**, which uses the Latent Dirichlet allocation (LDA) model to uncover the latent types of users based on their activities on an online dating platform.[8] In the original application of LDA model in 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 various mates' features. For example, the "outdoor type" may have higher preference for athleticism or dog ownership over other traits.

While the LDA model works well for the online dating market, it is not applicable to most social science problems for two reasons. First, this model requires data of

[7] To clarify, the term "recommender system" typically refers to systems that recommend items to users based on the reviews of users like them. That is not what we are discussing here. Instead, we focus on matching markets where the recommender system recommends users to one another.

[8] Besides **?**, **??** are two other attempts to estimate users' preference in online matching markets. However, these papers take a simple one-sided approach, ignoring the interplay between preference and opportunity. Therefore, I don't discuss them further here.

users reaching out to multiple partners rather than just the final match. Second, while the LDA model uncovers users' latent types, most social scientists want to estimate the preference of specific, known types (e.g. how different regime types may prefer different characteristics of an MNC).

1.2.5 Two-sided models for the labor and marriage markets

To be introduced here, or in their own chapters?

1.3 Conclusion

Roadmap for the rest of the dissertation

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, proposed by **?**. For easier exposition, throughout the chapter I will use the example of the labor market, where many workers can be matched to one firm.

We assume that the matching process in the labor market happens in two stages. In the first stage, each firm evaluates each worker in the sample, deciding whether to hire that worker or not. At the end of this stage, each worker will have received 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 the 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.

Our model only needs data on 1) the covariates of firms and workers, and 2) the job that workers accept. Such data is widely available in many social science surveys of the job market. Importantly, we do not need to observe the opportunity set. Therefore, our model obviates the need to follow the matching process and record who makes offer to whom, which is rarely possible for researchers.

If we assume that firms and workers are utility-maximizing agents, at the end of the matching process, no firm or worker would voluntarily change their final matches. As discussed in Section 1.1, this property is called *stability* in the game theoretic two-sided matching literature. We want our model to have this property because matching markets tend to produce stable matching. Indeed, **?** show that for any given set of preferences, a stable match always exist. Furthermore, **?** and **?** 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 imply that the matches will never change. Indeed, if actors' preference shifts, their characteristics change, or new actors enter the market, the matches will also change as a result of actors' recalculating their utility and adjusting their decisions. Therefore, since we are estimating 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.

This chapter will proceed as follows. First, I discuss the utility model for how firms make offers to workers. Second, I discuss the utility model for how workers choose the best offer among those extended by firms. Third, I show how we can use a Bayesian MCMC approach to estimate the model. Fourth, I analyze US labor data and demonstrate how to interpret the model's result.

2.1 Modeling firms' decision making

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:

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

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

Firm j will make an offer to hire worker i 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.

This model makes two important assumptions about firms' hiring process. First, whether a firm decides to hire worker A depends on the characteristics of worker A alone, and it will continue to hire worker A even if worker B is no longer available. In other words, firms regard workers as substitutes rather than complements.[1] This assumption is not universally true. A Hollywood producer may want to hire two specific lead actors for their chemistry, and if one is unavailable, the other also has to be replaced. However, for large firms where workers are closer to swappable cogs than unique superstars, this assumption is reasonable.

Second, the model assumes that the utility of hiring a worker does not depend on how many other workers accept the offer. In other words, the firm is large enough

[1] In the terminology of **?**, firms are assumed to have "substitutable preference," or firms' preference is assumed to have the property of substitutability. As discussed in Section 1.1, this assumption is necessary to prove the existence of stable matching in the case of many-to-one matching.

to employ all the workers to whom it extends offer without feeling the effect of diminishing marginal productivity of labor. This assumption is less restrictive than it may seem. Indeed, we can model the fact that the workers under consideration are less productive than the previous batch of workers by allowing firm j to have a high baseline utility b_j. Therefore, we are not assuming that there is never any diminishing marginal productivity of labor, only that there is negligible diminishing effect between the first and the last of the workers under consideration. This assumption is a reasonable approximation if the firm's labor force is large compared to the number of workers being considered.[2]

In addition to the two above assumptions about the process of firm's decision making, we make three parametric 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 with one another and across firms. Third, we assume that the as error terms $\epsilon_{1ij}, \epsilon_{0ij}$ follow the Gumbel distribution.[3] The choice of the Gumbel distribution is largely motivated by convenience since it allows us to derive the probability of firm j making an offer to worker i as the familiar binomial logit form:

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

[2] While not concerned with diminishing marginal productivity, **?** also assume that firms' quota, i.e. the number of workers they can accept, is sufficiently large to hire everyone in the set of workers under consideration. This assumption simplifies the proof that a stable match always exists in the case of many-to-one matching.

[3] The Gumbel distribution is very similar to the normal, only with a slightly fatter tail that allows for slightly more extreme variation in the unobserved utility. Its density function is $\exp^{-(x+\exp^{-x})}$, with mode 0, mean 0.5772, and fixed variance $\frac{\pi^2}{6}$. In practice, the difference between using Gumbel and independent normal error terms is small (**?**).

The term b_j is absorbed into $\boldsymbol{\beta}$ when we add an intercept term to the covariate matrix X.

Once firms have made their offers, each worker i will have a set of offers from which to pick her favorite. We call this set of offers the *opportunity set* of worker i, denoted O_i. Since unemployment is always an available option, every opportunity set includes unemployment as an "offer."[4]

The probability of worker i obtaining the opportunity set O_i is:

$$p(O_i|\boldsymbol{\beta}) = \prod_{j\in O_i} p(o_{ij} = 1|\boldsymbol{\beta}) \prod_{j\notin O_i} p(o_{ij} = 0|\boldsymbol{\beta}) \tag{2.6}$$

$$= \prod_{j\in O_i} \frac{\exp(\boldsymbol{\beta_j'} X_i)}{1 + \exp(\boldsymbol{\beta_j'} X_i)} \prod_{j\notin O_i} \frac{1}{1 + \exp(\boldsymbol{\beta_j'} X_i)} \tag{2.7}$$

2.2 Modeling workers' decision making

Worker i's utility for the accepting an offer from firm j is:

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

where α is a vector of workers' preference for relevant characteristics of firms, W_j is a vector of firm j's measured values on those characteristics, and v_{ij} is the unobserved component that influences worker i's utility.

Worker i evaluates all the firms in her opportunity set and selects the offer that brings the highest utility. This decision of worker i concludes the matching process, resulting in the observed final match between a worker and her chosen firm in our data.

[4] In our model setup, firms and workers decide sequentially, with firms making offers first in order for workers to have opportunity sets to choose from. While firms and workers in real life certainly do not act in this sequential manner, the idea of the opportunity set is still applicable. Workers in the real labor market may not know their exact set of offers, but they can certainly guess which firms are within their reach based on their characteristics and on guesses about firms' preference.

We make two assumptions in modeling the worker's decision making. First, for simplicity, we assume that all workers share the same set of preferences–hence α does not have a subscript i. The model can be extended so that there is heterogeneous preference among workers, either by estimating a separate model for each worker type (i.e. no pooling) or by building a hierarchical model for worker preference (i.e. partial pooling).

Second, we assume that the error term v_{ij} are uncorrelated across j. In other words, the unobserved factors in the utility of one job offer is uncorrelated to the unobserved factors in the utility of another job offer.[5] This assumption is most likely not true: if worker i values some unobserved factors of an offer, she is likely to consider those same factors in another offer as well. The hope is that we have modeled the observed portion 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 setup.[6]

Similar to our model of firm's utility, our model of worker's utility has three additional parametric assumptions that are standard in the literature. First, we assume that utility is linear. Second, the error term v_{ij} are uncorrelated across i. Third, we model v_{ij} having a Gumbel distribution so that the probability that worker i will accept the offer of firm j out of all the offers in its opportunity set O_i takes the conditional logit form (**?**):

[5] This assumption also gives rise to the Independence of the Irrelevant Alternatives (IIA) property. IIA implies that the relative odds of choosing between two alternatives depend only on the two alternatives under consideration. It does not depend on whether other alternatives are available or what their characteristics may be. Hence, other alternatives are considered "irrelevant."

[6] The discrete choice literature has developed solutions for such correlated error structure, such as nested logit, probit, and mixed logit, that can be applied here if we suspect that the unobserved portion is strongly correlated.

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

where a_i is the index of the firm that i accepts to to work for. Unemployment is indexed as 0.

2.3 Model estimation

Our goal is to estimate the preference of firms and workers, i.e. β_j and α. The key insight is that, conditional on the opportunity set being observed, the model of firms' and workers' decision making is a straightforward application of the binary logit and conditional logit model. Both models can be estimated with familiar tools like Maximum Likelihood Estimation (MLE).

However, in most social science research problems, the researcher only observes the final match A and not the opportunity set O. For example, labor market data typically does not include the set of offers a worker received (or would have received if she had applied), while data on her current job is widely available. Similarly in the marriage market or the FDI market, researchers often do not have the data on the offers being made, and only observe the final matching between men and women (i.e. who is married to whom) and between MNCs and countries (i.e. which factory is located where).

?'s solution to this problem is to use the Expectation-Maximization (EM) algorithm, an iterative method capable of finding the maximum likelihood estimates when the model depends on unobserved latent variables (i.e. the unobserved opportunity set in this case) (**?**). Our innovation is to estimate the model using a Bayesian MCMC approach, which offers several advantages. First, our MCMC approach produces the full posterior distribution, making inference easy. In contrast, EM only produces point estimates out of the box.[7] Second, our MCMC approach can be faster than EM when the latent variable, i.e. the opportunity set, is high dimensional (**?**).[8] Third, within the Bayesian framework, we can naturally put a hierarchical structure on firms' preference. This allows us to borrow information across firms, producing more precise estimates even when there is not a lot of data for a specific firm.

The rest of this section describes how we conduct model estimation.

2.3.1 Estimating the model using Metropolis-Hastings

We are interested in the posterior distribution of workers' and firms' preference given the observed final match, i.e. $p(\alpha, \beta|A)$. Unconditioned on the opportunity set, this posterior is difficult to derive or sample from. Therefore, we instead sample from the augmented posterior $p(\alpha, \beta, O|A)$, whose density is much simpler to derive.[9] Specifically,

$$p(\alpha, \beta, O|A) = \frac{p(A|\alpha, \beta, O)p(\alpha, \beta, O)}{p(A)} \tag{2.10}$$

$$\propto p(A|O, \alpha)p(O|\beta)p(\alpha)p(\beta) \tag{2.11}$$

where $p(A|O, \alpha)$ is derived in (2.9), $p(O|\beta)$ is derived in (2.7), $p(\alpha)$ and $p(\beta)$ are prior distributions for α and β. A key insight of this equation is that the acceptance of offers, i.e. $p(A|O, \alpha)$, depends only on the opportunity set and on the workers' preference. Similarly, the opportunity sets, i.e. $p(O|\beta)$, depend only on firms' preference.

---

[7] **?** propose a method for estimating the standard error of EM estimates. However, for hypothesis testing, we need further assumptions about the distribution of the EM estimates.

[8] Indeed, our opportunity set O is a $(I \times J)$ matrix of 0s and 1s, where I is the number of workers and J is the number of firms. Thus, there are 2^{IJ} potential values for the opportunity set, which quickly becomes untenable even for a small number of I and J. The high dimension of O forces **?** to reduce the data dimension by aggregating 17 employers in the data into 5 employer types, e.g. professional or blue collar jobs.

[9] See **?** for a discussion of such data augmentation techniques.

Because the opportunity set O is a discrete matrix of 0's and 1's, there is not any convenient conjugate model for (2.11), making Gibbs sampling impossible. Therefore, we use Metropolis-Hastings instead, a technique to sample from an arbitrary distribution $p(\theta)$ using the following steps:

1. Start from an arbitrary value of θ
2. Generate a proposal value θ^* from the proposal distribution $q(\theta^*|\theta)$
3. Calculate the acceptance ratio $MH_\theta = \frac{p(\theta^*)q(\theta|\theta*)}{p(\theta)q(\theta^*|\theta)}$
4. Accept the proposed value θ^* with probability $\max(1, MH_\theta)$
5. Repeat step 2-4 until convergence

In our case, we will use symmetric proposal distributions, i.e. $p(\theta^*|\theta) = p(\theta|\theta^*) \forall \theta, \theta^*$, so that the MH acceptance ratio simplifies to $MH_\theta = \frac{p(\theta^*)}{p(\theta)}$. In addition, because preference parameters tend to be correlated, we use an adaptive proposal distribution so that our MCMC samples have a faster convergence rate (**??**).[10]

Below we describe how to sample from the posterior of each parameter in the model using the Metropolis-Hastings (MH) algorithm. More detailed derivation of the Metropolis acceptance ratio is included in Appendix A. We ensure that our derivation and implementation of the acceptance ratio is correct using the unit-testing approach suggested by **?**.[11]

2.3.2 Posterior of the opportunity set $p(O|A, \alpha, \beta)$

For each worker i, we propose a new value O_i^* by flipping random cells in the current value O_i from 0 to 1 and 1 to 0. Substantively, this is equivalent to perturbing the

[10] Description of the Adaptive Metropolis procedure?

[11] Describe the unit-testing framework to ensure the correctness of MCMC code?

opportunity set by randomly making new offers or withdrawing existing offers. Note that this proposal distribution is indeed symmetric because proposing O_i^* from O_i and proposing O_i from O_i^* both involve flipping the same cells in the opportunity set. Hence, $p(O_i^*|O_i) = p(O_i|O_i^*) =$ the probability of selecting these particular cells out of the opportunity set.

The Metropolis acceptance ratio for the proposed opportunity set O_i^* is

$$MH_O = \frac{p(O_i^*|A_i, \alpha, \boldsymbol{\beta})}{p(O_i|A_i, \alpha, \boldsymbol{\beta})} \tag{2.12}$$

$$= \frac{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j) \pm \exp(\alpha' W_{j*})} \times \exp(\pm \boldsymbol{\beta}'_{j*} X_i) \tag{2.13}$$

where $\pm$ evaluates to $+$ if j^* is a new offer being added to the current opportunity set, and evaluates to $-$ if j^* is an existing offer being withdrawn from the current opportunity set.

To understand the intuition behind this formula for MH_O, consider the scenario in which we propose a new opportunity set for worker i by adding an offer from firm j. Since worker i now has one more choice to choose from, it becomes less likely that worker i's accepted job is the best choice. This makes the proposed opportunity set less consistent with the observed data than the current opportunity set, and MH_O should decrease accordingly. This is reflected in the formula for MH_O by the $\exp(\alpha' W_{j*})$ term in the denominator.

On the other hand, whether we should add the offer to the opportunity set also depends on firm j's preference for worker i. If hiring worker i brings firm j net positive utility (i.e. $\boldsymbol{\beta}'_{j*} X_i > 0$), we should add the offer. This is reflected in the formula for MH_O by the multiplier $\exp(\boldsymbol{\beta}'_{j*} X_i)$, which is larger than 1 when $\boldsymbol{\beta}'_{j*} X_i > 0$.

2.3.3 Posterior of firms' preference $p(\alpha|A, O, \beta)$

At the beginning of the MCMC chain, we propose a new α^* using a Normal proposal distribution centered on the current value α with a hand-tuned diagonal covariance matrix. Later in the MCMC chain, the covariance matrix of the proposal distribution is adapted based on past samples to take into account the correlations across preference parameters (**?**).

The Metropolis acceptance ratio for the proposed α^* is[12]

$$MH_\alpha = \frac{\alpha^*|A, O, \beta)}{p(\alpha|A, O, \beta)} \tag{2.14}$$

$$\begin{aligned}\log MH_\alpha &= \sum_i \Big[(\alpha^* - \alpha)' W_{a_i} + \\ &\quad \log\left(\sum_{j:j\in O_i} \exp(\alpha' W_j)\right) - \log\left(\sum_{j:j\in O_i} \exp(\alpha^{*\prime} W_j)\right)\Big] + \\ &\quad \log p(\alpha^*) - \log p(\alpha)\end{aligned} \tag{2.15}$$

2.3.4 Posterior of workers' preference $p(\beta|A, O, \alpha)$

We propose a new $\boldsymbol{\beta}^*$ using a Normal, adaptive proposal distribution similar to α. Because $\boldsymbol{\beta}$ is high dimensional, with one set of β for each employer, in each MCMC iteration we randomly choose and update only a part of $\boldsymbol{\beta}$.

The Metropolis acceptance ratio for the proposed β is

$$MH_\beta = \frac{p(\beta * |A, O, \alpha)}{p(\beta|A, O, \alpha)} \tag{2.16}$$

$$\begin{aligned}\log MH_\beta = \sum_i \left[\sum_{j\in O_i} \left(\beta_j^{*\prime} X_i - \beta_j' X_i\right) + \sum_j \left(\log(1 + \exp(\beta_j' X_i)) - \log(1 + \exp(\beta_j^{*\prime} X_i))\right)\right] \\ + \log p(\boldsymbol{\beta}^*) - \log p(\boldsymbol{\beta})\end{aligned} \tag{2.17}$$

[12] We log-transform the Metropolis acceptance ratio for better numerics.

2.3.5 *Posterior of β's hyperparameters μ_β, τ_β*

As discussed above, the Bayesian approach to estimating our two-sided model allows us to put a hierarchical structure on the preference parameter. Here, we model firms' preference $\boldsymbol{\beta}$ as being drawn from the multivariate normal distribution $MVN(\mu_\beta, \tau_\beta)$, where μ_β is the mean and τ_β is the precision.

When the prior $p(\beta)$ is also normal, we have a conjugate multivariate normal model, where μ_β and τ_β are the parameters while β is considered the "data".

Since the model is conjugate, we can sample from the posterior of μ_β and τ_β with Gibbs sampling. Their full conditional distribution of μ_β is:

$$p(\mu_\beta) \sim MVN(\mu_0, \Sigma_0) \tag{2.18}$$

$$p(\mu_\beta|\beta, \tau_\beta) \sim MVN(m, V) \text{ where} \tag{2.19}$$

$$V = (\Sigma_0^{-1} + n\tau_\beta)^{-1} \tag{2.20}$$

$$m = (\Sigma_0^{-1} + n\tau_\beta)^{-1}(\Sigma_0^{-1}\mu_0 + n\tau_\beta\bar{\beta}) \tag{2.21}$$

The full conditional distribution of τ_β is:

$$p(\tau_\beta) \sim \text{Wishart}(\nu_0, S_0^{-1}) \tag{2.22}$$

$$p(\tau_\beta|\beta, \mu_\beta) \sim \text{Wishart}(\nu, S^{-1}) \text{ where} \tag{2.23}$$

$$\nu = \nu_0 + n \tag{2.24}$$

$$S^{-1} = \left(S_0 + \sum(\beta - \mu_\beta)(\beta - \mu_\beta)'\right)^{-1} \tag{2.25}$$

2.4 Results for US labor data

To be written ...

3

Simulation results

4

FDI

4.1 Introduction

The political science literature on Foreign Direct Investment (FDI) has focused largely on how politics shapes the flow of FDI across countries. The central insight of this literature is that multinational corporations (MNCs) face an "obsolescing bargain" against the host government. Once the MNC has sunk its investment, it is vulnerable to the host government's changing regulations, backtracking on deals, or even expropriating its properties (**??**). Certain institutional and political characteristics, such as numerous veto players, executive constraint, or strong property rights, allow the host government to make a credible commitment and thus ameliorate the severity of the "obsolescing bargain" problem (**???**). According to the literature, MNCs should invest more in countries with these characteristics.

This dominant approach in the literature has three long-standing issues that my paper will address. First, the majority of the literature relies on FDI stock and flow data as the outcome of interest even though they are often not an appropriate measure for the scale of MNCs' activities (**?**). While it would be ideal to use firm-level data instead, both the lack of cross-national firm-level data and a suitable statistical model have posed a challenge.

Second, while there has been much focus on MNCs choosing host countries, the literature has largely neglected the other side of the investment decision: what are countries' preferences regarding MNCs? Consider the established finding that democracies receive more FDI. Without controlling for countries' preferences, it is difficult to interpret this fact as democracies actively pursuing MNCs or as MNCs finding democracies attractive. Not only are countries' preferences central to the modeling of investment decision, arguably it is also more steeped with politics and deserves more attention. **?** and **?** are two pioneering works in this area of research, proposing partisan politics and regime types as factors shaping countries' preferences for FDI. However, while their theories are ground-breaking, the empirical estimation of countries' preferences remains difficult.

Third, in addition to empirical issues raised above, I propose that we need to theorize about countries' preferences for FDI quality. While the political science literature has largely focused on the quantity of FDI, national policies and discourses pay much attention to the quality of FDI, using various incentives and restrictions to target certain types of FDI. Indeed, MNCs come with varying capital, demand for labor, and technology, all of which have different effects on the host country's economy. For example, policy makers and scholars have highlighted high-tech MNCs as a source of technological transfer for developing host countries, allowing them to upgrade their technical capacity and improve their productivity (**??**). While such high-quality FDI has been enthusiastically endorsed by the development community, I argue that only governments with a long time horizon want to attract high-tech FDI because technological transfer takes time to pay off.

In sum, the current literature would benefit from an analysis that is capable of using firm-level data to estimate both firms' and countries' preferences for each other's characteristics. To accomplish this goal, I adapt the two-sided matching model originally designed for the labor market and the marriage market. In this model, both firms and countries evaluate their available options and choose the best according to their utility functions. As in many social science contexts, we only observe the final firm-country matches and not the full set of available options (also known as the opportunity set). I solve this problem by using the Metropolis Hastings algorithm, a Markov chain Monte Carlo (MCMC) approach that repeatedly samples new opportunity sets and rejects them at an appropriate rate to approximate their true distribution. Since the two-sided matching model is derived explicitly from actors' utility functions, their parameters also enjoy a straightforward interpretation in the utility space instead of some aggregate outcomes.

The paper proceeds as follows. Section 4.2 discusses the three long-standing issues with the literature and how they can be improved. Section **??** lays out the utility structure in the two-sided matching model and describes the matching process. Section 4.3 shows an application of the model on a census of Japanese firms overseas. Section presents the result. Section 4.4 concludes.

4.2 Three Issues in the Literature of FDI's Political Determinants

4.2.1 Measuring MNCs' Activities

For the majority of political science theory regarding FDI, the quantity of interest is the scale of MNCs' activities in a country, and not necessarily how much FDI crosses its border. Indeed, we theorize about how MNCs may reduce their activities for fear of expropriation, and how the host country's political factors can induce MNCs to invest more with a credible commitment not to expropriate. It is also the scale of MNCs' activities that determines how many jobs are created or how much of the domestic market is competed away, engendering labor's support and local business' lament.

However, to measure the scale of MNCs' activities, the vast majority of works uses how much FDI crosses the border, specifically FDI stock and flow (**????**). As **?** points out, these measures, whose original purpose is to monitor balance of payments, are often misleading about MNCs' activities. FDI flow does not count locally raised capital and reinvested earnings since they do not cross any border. FDI stock calculated at market value fluctuates based on market price, unrelated to firms' behavior. FDI stock calculated at historical value, which records asset value at the time it was acquired, is more stable and appropriate to measure the scale of MNCs' activities. Unfortunately, due to onerous data requirements, most countries measures FDI stock by simply adding up FDI flow across years.

Given the interest of political science theory in MNCs' activities, **?** suggests less use of FDI stock and flow and more use of firm-level statistics. For example, consider the hypothesis that countries with more veto players have more stable policies and are thus more attractive to FDI (**?**). Instead of using FDI stock and flow into a country to measure its attractiveness, we can study whether more MNCs are located there.

While firm-level data has become more abundant in recent years,[1] it is not clear how to analyze this type of data appropriately. Given the data structure of a set of firms interacting with a set of countries, one may consider a dyadic-based analysis, frequently used in the International Relations literature. In such analysis, the unit of observation is a firm-country dyad, and the model used is typically OLS regression. Each dyad is assumed to be independent of each other, and any bias caused by interdependency is fixed via post-estimation procedures, such as clustered standard errors (**?**).

Unfortunately, this dyadic approach is inappropriate to analyze MNCs' investment location. Once a firm chooses to invest in a country, it is by definition not

[1] Examples of firm-level data include the US Bureau of Economic Analysis (BEA)'s survey of all US firms abroad, Tokyo Keizai's Overseas Japanese companies database (*Kaigai Sinshutsu Kigyou Souran*), World Bank's Enterprise Survey, and Orbis database of companies worldwide.

investing in another. Therefore, the values of firm-country dyads deterministically constrain one another and cannot be modeled as independent draws from a common distribution.

The two sided matching model solves this problem by considering one firm-country match as the unit of observation. The intuition is as follows. If we observe that a firm is welcome to invest in countries $j_1, j_2, \ldots, j_n$ but ends up investing in country j^*, it must mean country j^* offers the highest utility to firms. Continuing the previous example, if country j^* has more veto players than average, we can infer that MNCs indeed prefer countries with more veto players.

4.2.2 Estimating Countries' Demand for FDI

Recognizing that our model of investment location has not taken into account countries' demand for FDI, **?** and **?** recently broke ground in this area. Similar to the rich IPE literature in trade and exchange rate, these studies argue that countries' demand for FDI varies according to FDI's distributive effect on their domestic constituencies (**??**). In this theoretical framework, labor supports FDI because foreign firms bring capital that increases the demand for labor and raises productivity, both of which lead to higher wage. On the other hand, domestic firms oppose FDI because foreign firms compete for local labor, inputs, and markets. Both **?** and **?** formulate their theories as a variant of this labor-vs-business tension, which surfaces in the former work as left-vs-right governments, and in the latter as democratic-vs-authoritarian regimes.

While these pioneering works have enriched our understanding of the relationship between politics and FDI, their empirical approaches do not satisfactorily measure countries' demand for FDI, leaving their theoretical arguments untested.

Consider **?**'s approach, which controls for economic and institutional factors that affect FDI flow into a country. The author then claims that the country's openness towards FDI is what's left in the residual.[2] For this approach to be valid, every economic, institutional, and endowment factors that affect FDI flow have to be controlled for, leaving only the country's demand in the error term. This claim is much stronger than the regular assumption of exogenous and normally distributed error, which is valid as long as the omitted factors are uncorrelated with the independent variable of interest. Framed substantively, since the residual is likely to contain more than just the country's demand for FDI, if we observe an abnormally high level of FDI, we do not know whether it is because the country welcomes FDI or because MNCs find something attractive in the country.[3]

In contrast to **?**'s statistical approach, **??** substantively measures countries' demand for FDI, using the annual US Investment Climate Reports to code the number of industries that have foreign ownership restrictions or face investment screening. The advantages of this measurement are its ease of interpretation and its availability for many countries. However, two problems remain. First, adding up the raw count of restricted industries is not appropriate because industries are not all the same. For example, given the reach of the banking sector into all corners of the economy, a country's opening up its financial industry indicates much more FDI-friendliness than, say, allowing foreign furniture makers to set up shops. Since the theoretical argument is driven by FDI's distributive effect, it is suspect to ignore the varying impact of FDI across sectoral constituencies.

Second, according to the coding rules, an industry is coded as free if there is no

[2] Specifically, the estimation of FDI openness involves two steps. First, the author runs a gravity model explaining bilateral FDI flows, estimating the intercept as the host country-year fixed effect. Second, this fixed effect is then regressed on several economic and endowment factors of that country-year (i.e. GDP, GDP per capita, average school years, arable land). The residual in the second stage is considered the country's "FDI openness" in that year.

[3] In addition, the data requirement of bilateral FDI flows, ideally disaggregated by sectors, is very demanding. Therefore, this approach is limited to OECD countries only (**?**). During the period the authors study, 1980-2000, OECD countries accounted for 95% of global FDI outflow and 90% of inflow. However, since then the role of the developed world in global FDI has declined sharply, reduced to 60.8% of outflow and 40.6% of inflow in 2014 (**?**).

mention of restriction. If an industry receives little FDI, it may not be worth mentioning as being restrictive and yet still coded as open. Therefore, "zero restriction" in the dataset can either mean that a country is very closed or very open to FDI. This concern is not hypothetical. Figure 4.1 shows that, following the coding of the US Investment Climate Reports, China seemed 100% open to FDI up until 1986 when it started imposing restrictions. The reality is the opposite. Prior to 1986, only limited FDI was allowed as joint-venture in Special Economic Zones (SEZ). The year of 1986 was, in fact, the first time China allowed any wholly owned FDI outside of SEZs.

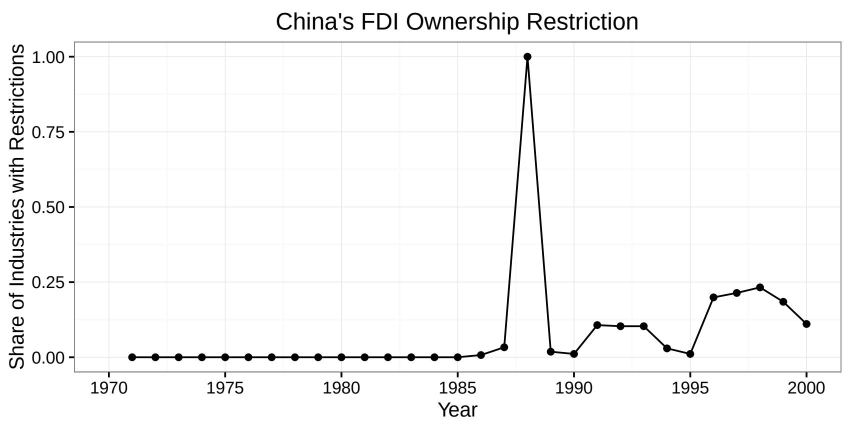

FIGURE 4.1: China's FDI Ownership Restriction, as coded in **?**. Prior to 1986, FDI in China was limited to few experimental Special Economic Zones, and thus not mentioned in US Investment Reports. The sharp spike in 1988 also does not seem to correspond to any actual change in policy, and likely another artifact of reporting. (See **?** for a historical overview of China's FDI policy.)

The two-sided matching model circumvents these thorny measurement issues by incorporating countries' utility function directly into the model. If we observe that country j welcomes firms $i_1, i_2, \dots, i_n$ to invest but not others, we can compare the characteristics of firms $i_1, i_2, \dots i_n$ with the others to infer country j's preference.

4.2.3 Estimating Countries' Preferences for FDI's Technological Intensity

While the political science literature has focused almost exclusively on the quantity of FDI, treating all FDI as one homogeneous flow of capital, policy makers seem to pay much more attention to distinguishing types of FDI. Commenting on the role of International Investment Agreements (IIAs), **?** says, "Today, increasing the quantity of investment is not enough. What matters is its quality, i.e. the extent to which investment delivers concrete sustainable development benefits." Governments in developing countries, from Ghana to China, all offer various forms of tax incentives and fee waivers to attract FDI that invests in a remote region, brings new technology, or focuses on exporting (**?**). Since 2006, China's official FDI policy has been "quality over quantity," promoting FDI with intense R&D in high-productivity sectors (**?**). Indeed, for developing countries, the hope is that MNCs will transfer their technologies to the domestic economy by training workers or partnering with local suppliers.

Despite the importance of disaggregating FDI by its quality, data unavailability remains the bottleneck. The few existing attempts use detailed data from only one country or limit the sample to OECD countries (**???**). With cross-country firm level data now available, we often have information on the firms' industry or even research and development (R&D) expenditure. With the two-sided matching model, I will be able to estimate countries' preferences for firms' technological intensity. I hypothesize that, since MNCs' technologies takes time to diffuse to local businesses, a country' preference of high-tech FDI is shaped by its time horizon.

4.3 Applying the Two-Sided Matching Model to Japanese MNCs

In this section, I apply the two-sided matching model to study the investment location of Japanese firms overseas. The data comes from the *Kaigai Shinshutsu Kigyou Souran*, an annual publication that contains information about the foreign affiliates of Japanese firms, including their location, industry, capital, and labor size.[4] This database is reputed to include all Japanese firms overseas (**?**). The final sample

[4] I thank Professor Andrew Delios for generously sharing the data.

includes 6474 Japanese foreign affiliates in 2003, spreading across 37 countries, with China and the US leading as the two top destinations for Japanese MNCs (Table **??**).

For countries' characteristics that firms consider, I include:

- Market size: MNCs are expected to prefer countries with a large market size, which present MNCs with many potential customers. Indeed, this has been often cited as the allure of China to MNCs (**?**). I follow the standards in the literature and include log GDP (constant 2005 US$), taken from the World Bank's World Development Indicators.

- Level of development: MNCs are expected to prefer countries with a high level of development. A developed economy has consumers with high purchasing power and better infrastructure. To measure development, I use log GDP per capita (constant 2005 US$) from World Development Indicators.

- Labor quality: As one primary factor of production, labor matters greatly to firms' productivity and profit. To measure labor quality, I use the average years of schooling of adult, taken from the UNDP's Human Development Report.[5]

- Democracy: Democracy has been a mainstay in the political science literature on FDI. Scholars have argued that MNCs want to invest in democratic regimes for various reasons, including stable policy, credible commitment, and strong property rights (**???**). On the other hand, recent works have also argued that democratic regimes want FDI more than autocratic regimes (**?**). Thus, it is unclear whether the observed high level of FDI in democracies is due to the push or the pull factors. By controlling for countries' preference in the two-sided

[5] Since Taiwan is not included in UNDP's and World Bank's data, I collected its statistics from the Taiwanese Statistical Website.

matching model, I can better estimate the effect of democracies on firms' utility. I measure democracy using the binary Demoracy & Dictatorship, developed by **?**.

For firms' characteristics that countries consider, I include:

- Capital size (in US$): A main argument for the benefit of FDI is that it brings capital to the country, improving labor productivity. MNCs' capital is especially important for developing countries, which cannot muster much domestic capital from their poor population. The capital size of a firm is included in the Japanese Overseas Business dataset.

- Labor size: Similarly, a reputed benefit of FDI is that it creates jobs, generating not just economic growth but also increasing the government's popularity among the populace. The total number of employees of a firm is included in the Japanese Overseas Business dataset.

- Technology intensity: I proxy for a firm's technology intensity by the industry to which it belongs. **?** categorizes ISIC industries into four levels of technology intensity—low, medium low, medium high, and high—according to the level of R&D expenditure divided by sales. I convert the industry classification of firms in my data from SIC 3 to ISIC and categorize their technology intensity from 1 to 4, with 1 being low and 4 being high. On several occasions, one industry in SIC 3 matches to multiple ISIC (rev 3) industries or none at all. In the former case, I take the average across matched ISIC industries. In the latter case, the data is missing and later removed from the analysis.

4.4 Conclusion

In this paper, I propose the two-sided matching model to estimate firms' and countries' preferences, solving three persistent issues in the literature of FDI's political determinants. The results indicate that, for Japanese MNCs, only a country's level of development matters and not its market size, labor quality, or regime type. This finding suggests that we should take a closer look at the relationship between democracies and MNCs. Since previous works in the literature have not controlled for countries' preferences, they may have mistaken democracies' love for FDI as FDI's fondness for democracies.

On the other hand, the model's estimation of countries' preference remains lacking. Since each country has its own set of parameters, the parameter space seems too large for the current implementation of the Metropolis-Hastings algorithm to fully explore. Several solutions are possible. First, we can collapse countries into categories of interest, e.g. regime types, (categorical) time horizon length. Second, we can build a hierarchical model, modeling countries' preferences as draws from a common distribution. Such model will allow us to pool information across countries and reduce the parameter space.

Appendix A

Derivation of the Metropolis-Hastings Acceptance Ratio

A.0.1 Opportunity sets O

Target distribution for a firm i

$$p(O_i|A_i,\alpha,\boldsymbol{\beta}) = \frac{p(O_i,A_i,\alpha,\boldsymbol{\beta})}{p(A_i,\alpha,\boldsymbol{\beta})} \tag{A.1}$$

$$MH_O = \frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{p(O_i^*,A_i,\alpha,\boldsymbol{\beta})}{p(A_i,\alpha,\boldsymbol{\beta})} \times \frac{p(A_i,\alpha,\boldsymbol{\beta})}{p(O_i,A_i,\alpha,\boldsymbol{\beta})} \tag{A.2}$$

$$= \frac{p(O_i^*,A_i,\alpha,\boldsymbol{\beta})}{p(O_i,A_i,\alpha,\boldsymbol{\beta})} \tag{A.3}$$

$$= \frac{p(A_i|O_i^*,\alpha)p(O_i^*|\boldsymbol{\beta})}{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta})} \tag{A.4}$$

$$\tag{A.5}$$

where the factorization of the likelihood in (A.4) 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 (2.9) and (2.7)

$$\frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{\sum\limits_{j:j\in O_i}\exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i}\exp(\alpha' W_j) + \exp(\alpha' W_{j*})} \times \exp(\boldsymbol{\beta}'_{j*}X_i) \tag{A.6}$$

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

$$\frac{p(O_i^*|A_i,\alpha,\boldsymbol{\beta})}{p(O_i|A_i,\alpha,\boldsymbol{\beta})} = \frac{\sum\limits_{j:j\in O_i}\exp(\alpha' W_j)}{\sum\limits_{j:j\in O_i}\exp(\alpha' W_j) - \exp(\alpha' W_{j*})} \times \exp(-\boldsymbol{\beta}'_{j*}X_i) \tag{A.7}$$

A.0.2 Workers' parameters, α

Target distribution:

$$p(\alpha|A,O,\boldsymbol{\beta}) = \frac{p(O,A,\alpha,\boldsymbol{\beta})}{p(A,O,\boldsymbol{\beta})} \tag{A.8}$$

Metropolis-Hasting acceptance ratio:

$$MH_\alpha = \frac{p(\alpha^*|A,O,\boldsymbol{\beta})}{p(\alpha|A,O,\boldsymbol{\beta})} = \frac{p(A_i|O_i,\alpha^*)p(O_i|\boldsymbol{\beta})p(\alpha^*)}{p(A_i|O_i,\alpha)p(O_i|\boldsymbol{\beta})p(\alpha)} \tag{A.9}$$

$$= \frac{p(A_i|O_i,\alpha^*)p(\alpha^*)}{p(A_i|O_i,\alpha)p(\alpha)} \tag{A.10}$$

where (A.10) is due to the symmetric proposal distribution (so $\frac{p(\alpha^*|\alpha)}{p(\alpha|\alpha^*)} = 1$)

If we plug in (2.9),

$$MH_\alpha = \prod_i \left[\frac{\exp(\alpha^{*\prime} W_{a_i})}{\exp(\alpha' W_{a_i})} \times \frac{\sum_{j:j \in O_i} \exp(\alpha' W_j)}{\sum_{j:j \in O_i} \exp(\alpha^{*\prime} W_j)} \right] \times \frac{p(\alpha^*)}{p(\alpha)} \tag{A.11}$$

$$= \prod_i \left[\exp(\epsilon'_\alpha W_{a_i}) \times \frac{\sum_{j:j \in O_i} \exp(\alpha' W_j)}{\sum_{j:j \in O_i} \exp(\alpha^{*\prime} W_j)} \right] \times \frac{p(\alpha^*)}{p(\alpha)} \tag{A.12}$$

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

$$\log MH_\alpha = \sum_i \left[\epsilon'_\alpha W_{a_i} + \log \left(\sum_{j:j \in O_i} \exp(\alpha' W_j) \right) - \log \left(\sum_{j:j \in O_i} \exp(\alpha^{*\prime} W_j) \right) \right] + \log p(\alpha^*) - \log p(\alpha) \tag{A.13}$$

A.0.3 Firms' parameters, β

Target distribution:

$$p(\boldsymbol{\beta}|A, O, \alpha) = \frac{p(O, A, \alpha, \boldsymbol{\beta})}{p(A, O, \alpha)} \tag{A.14}$$

Metropolis-Hasting acceptance ratio:

$$MH_\beta = \frac{p(\beta^*|A, O, \alpha)}{p(\beta|A, O, \alpha)} = \frac{p(A_i|O_i, \alpha) p(O_i|\boldsymbol{\beta}^*) p(\boldsymbol{\beta}^*|\mu_\beta, \tau_\beta)}{p(A_i|O_i, \alpha) p(O_i|\boldsymbol{\beta}) p(\boldsymbol{\beta}|\mu_\beta, \tau_\beta)} \tag{A.15}$$

$$= \frac{p(O_i|\boldsymbol{\beta}^*) p(\boldsymbol{\beta}^*|\mu_\beta, \tau_\beta)}{p(O_i|\boldsymbol{\beta}) p(\boldsymbol{\beta}|\mu_\beta, \tau_\beta)} \tag{A.16}$$

where (A.15) is due to the symmetric proposal distribution.

We plug in (2.7),

$$MH_\beta = \prod_i \left[\prod_{j \in O_i} \frac{\exp(\beta_j^{*\prime} X_i)}{\exp(\beta_j' X_i)} \times \prod_j \frac{1 + \exp(\beta_j^{*\prime} X_i)}{1 + \exp(\beta_j' X_i)} \right] \times \frac{MVN(\boldsymbol{\beta}^* | \mu_\beta, \tau_\beta)}{MVN(\boldsymbol{\beta} | \mu_\beta, \tau_\beta)} \quad \text{(A.17)}$$

$$\log MH_\beta = \sum_i \left[\sum_{j \in O_i} \beta_j^{*\prime} X_i - \beta_j' X_i + \sum_j \log(1 + \exp(\beta_j^{*\prime} X_i)) - \log(1 + \exp(\beta_j' X_i)) \right] \quad \text{(A.18)}$$

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