# 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

# List of Figures

# List of Abbreviations and Symbols

## Abbreviations

| | |
|---|---|
| EM | Expectation Maximization. |
| FDI | Foreign Direct Investment. |
| IPA | Investment Promotion Agency. |
| IPE | International Political Economy. |
| MCMC | Markov Chain Monte Carlo. |
| MH | Metropolis-Hastings. |
| MLE | Maximum Likelihood Estimation. |
| MNC | Multinational Corporation. |
| MVN | Multivariate Normal. |

# 1

# Introduction

In recent decades, the global flow of Foreign Direct Investment (FDI) has steadily increased, rising from almost nothing in the 1970s to over $2.3 trillion dollars in 2016, becoming an important source of global capital. For developing countries especially, capital from Multinational Corporations (MNCs) is robust to global economic downturns, prompting major international organizations to endorse FDI as a key factor to economic development and poverty reduction (Figure 1.1) (Mallampally and Sauvant, 1999; World Economic Forum, 2013). Within the International Political Economy (IPE), much of the literature also starts with the view that FDI bring various benefits to the host countries, and that these countries will always seek FDI (Jensen, 2008a). These works focus on *how* countries can attract FDI, and do not question *whether* they want to do so (Jensen, 2003; Li and Resnick, 2003; Li, 2006; Ahlquist, 2006).[1]

At first glance, the benefits of FDI do not seem controversial. In addition to bringing capital to and creating jobs in the host economy, FDI holds an important

[1] Two recent exceptions are Pinto (2013); Pandya (2016), which are the first to examine countries' demand for FDI.

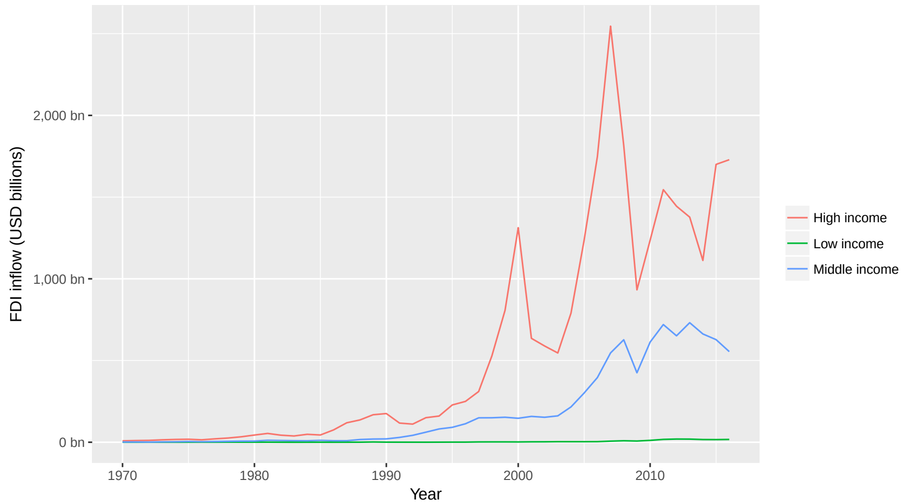


FIGURE 1.1: FDI global inflow, 1970-2006. The last four decades witness the growth of FDI into the most important source of global capital. Source: World Bank's World Development Indicators.

promise that is the spillover of productivity from foreign to domestic firms. As well-known from neoclassical growth theory, diminishing returns to capital will at one point stop capital from accumulating further, preventing long-run economic growth from being driven by capital accumulation alone (Solow, 1956). Findlay (1978)'s groundbreaking model of FDI and growth shows how FDI can counteract this dynamics. In this model, technology spillover from foreign firms shift the domestic factor-price frontier to the right, allowing more output from the same input, ultimately resulting in a continually increasing capital stock for the domestic sector. In this view, FDI is welfare-enhancing, providing spillover benefits to local firms in ways that foreign firms do not take into account in their private calculations. This claim about the positive effect of FDI provides a justification for countries' using investment promotion to rectify the "undersupply" of FDI (Moran, 1998).

Despite this prevailing theoretical argument, recent empirical evidence shows that not all FDI are the same and that its effects are highly conditional. There is no conclusive evidence of FDI having a positive effect on growth (Nair-Reichert and

Weinhold, 2001; Carkovic and Levine, 2002) or poverty reduction (Guerra et al., 2009). This puzzle opens a substantial literature on how the growth-enhancing and spillover effect of FDI is conditional on the absorptive capacity of the host economies, i.e. its level of human capital, technological sophistication, and financial market development (Durham, 2004; Nunnenkamp and Spatz, 2004; Fu, 2008; Willem, 2004). In addition, while the capital brought and jobs created by FDI are unconditionally good for the overall economy, its distributional effects across different constituencies in the host economy, creating political cleavage across both sectoral and geographical divides (Chintrakarn et al., 2012; Goldberg and Pavcnik, 2007; Nunnenkamp et al., 2007).

Given the recent evidence on the conditional effect of FDI, it is no longer tenable to hold the assumption that countries' preference for FDI is largely homogeneous. By holding this assumption, we neglect the role of the state in shaping global capital flow, falling prey to the discredited "race to the bottom" thesis of globalization (Mosley, 2005). Arguably, examining countries' preference for FDI should be of more interest to political scientists than the current focus on MNCs' location preference, which often amounts to adding a political variable to an existing economic model of FDI flow. Plus, even if we only care about MNCs' location preference, we must still take into account countries' preference in order to get an accurate estimate. For example, consider the received wisdom that democracies receive more FDI (Jensen, 2008b). Without controlling for countries' preferences, it is difficult to interpret this fact as democracies actively pursuing MNCs or as MNCs finding democracies attractive.

## 1.1 Goal of the Dissertation

In essence, this dissertation aims to estimate the preference of both countries and MNCs for each other. It develops an empirical strategy that takes into account the two-sided nature of the FDI market, i.e. a FDI project can only exist if both the MNC and the host country agree. Recognizing that this two-sided matching dynamics can also be found in the labor or the marriage markets, I adapt the statistical models first developed in Sociology for these markets and apply them to the study of FDI (Logan, 1996; Logan et al., 2008).

In doing so, I simultaneously addresses three long-standing issues in the FDI literature.

First, I bring the state back in, filling the current gap in the literature on the variation of countries' preference for FDI. Two notable exceptions are Pinto (2013) and Pandya (2016), whose pioneering works propose partisan politics and regime types as factors shaping preferences for FDI. However, while their theories are groundbreaking, the empirical estimation of countries' preferences remains inadequate. In addition, these researchers have not used their findings to re-estimate the preference of MNCs, separating out the "push" and "pull" factors of FDI flow.[2] Using a two-sided matching model, I will naturally be able to estimate both sides' preference.

Second, I propose that we need to theorize about countries' preferences for different types of FDI. While the IPE literature has largely focused on the quantity of FDI flows, 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 amount of capital, labor demand, and technological sophistication, all of which have different effects on the host country's economy. Just as the two-sided matching model can estimate MNCs' utility functions for countries' characteristics (e.g. market size, level of development), it can also estimate countries' utility functions for MNCs' characteristics (e.g. technological sophistication, export strategy).

Third, while the majority of the literature uses FDI stock and flow data, these

[2] "Push factors" refer to characteristics of the home country and of the MNC, pushing capital out from its origin. "Pull factors" refer to the characteristics of the host country, pulling capital towards its destination.

statistics are accounting constructs created to keep track of countries' balance of payment and thus map poorly to concepts in Political Science theories. Very often, the variable of interest in our theories is the scale of MNCs' activities in the host country, which can be very different from the amount of border-crossing capital thanks to MNCs' complex financial and tax strategies (Kerner, 2014). Therefore, we would do much better testing our theories with firm-level operational data. Unfortunately, even when firm-level FDI data is available, the lack of a suitable statistical model poses a big barrier to this approach.

These three issues are related and represent the status quo in the FDI literature. Data limitation forces scholars to look at country-level aggregate FDI flow, making it difficult to study countries' preference for different types of FDI. Without taking into account countries' preference, models of MNCs' location choice are suspect.

In sum, my dissertation benefits the field by developing a empirical strategy 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 MNCs and countries evaluate their available options according to their utility functions, choose the best alternative, culminating in a subsidiary being built by a MNC in a host country. Our goal is to estimate MNCs' and countries' utility functions, and the challenge is to do so using only data on which subsidiary is located where. Indeed, it would be straightforward to estimate the utility functions if we observed not only the MNC's location decision but also the set of options presented to the MNC.[3] (Following the matching literature, we call this set of options the "opportunity set."). Unfortunately, while data on subsidiaries' location are available,

[3] Discrete choice models can be used to estimate the utility functions when both the choice and the set of options are observed. Indeed, this remains the dominant empirical approach in the industrial location literature, effectively ignoring the fact that not all firms have the same set of location options (Arauzo-Carod et al., 2010).

the opportunity set is generally unobservable as researchers cannot peek into the negotiation process between countries and MNCs. The two-sided matching model solves 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, the estimated parameters also have a convenient interpretation as the effect of different variables on MNCs' and countries' utility. This allows us to make statements such as: "In evaluating MNCs, China values a 2% increase in the firm's capital as much as a 1% increase in labor demand."

## 1.2 Roadmap

In the rest of this introductory chapter, I review in-depth the three issues in the literature of FDI's political determinants, outlining the current attempts to address them and how my approach can contribute to the solution.

In Chapter 2, I describe the two-sided matching model, including both its game-theoretic origin and its statistical estimation. Chapter 3 uses simulations to demonstrate the correctness of the model and explores its characteristics. Chapter 4 applies the model on US labor market data, the original domain of the two-sided matching approach, in order to compare with and expand upon previous results. Chapter 5 brings us back to the study of FDI, applying the model on firm-level data of Japanese MNCs in East and Southeast Asia. Chapter 6 concludes and explores potential applications of the two-sided matching model in other areas of Political Science.

## 1.3 Three Issues in the Literature of FDI's Political Determinants

### *1.3.1 Estimating countries' demand for FDI*

Despite earlier pessimism about countries engaging in a race to the bottom to attract footloose global capital, empirical evidence shows that the nation-state still maintains significant autonomy. The effect of globalization seems to be mediated by both domestic politics and ideational force, allowing for variation in labor standards, environmental regulation, and tax policies (Drezner, 2001).

While the broader IPE literature has recognized the variation of countries' policies in the face of globalization, we have surprisingly much less work to explain variation in countries' FDI policies. Recognizing this gap in the literature, Pinto (2013) and Pandya (2016) recently broke ground in this area. Similar to the rich IPE literature in international trade, these studies argue that countries' demand for FDI varies according to FDI's distributive effect on their domestic constituencies (Broz and Frieden, 2001; Milner and Kubota, 2005). 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 Pinto (2013) and Pandya (2016) 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 Pinto (2013)'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.[4] 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.[5]

In contrast to Pinto (2013)'s statistical approach, Pandya (2014, 2016) 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

[4] 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.

[5] 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 (Pinto and Pinto, 2008). 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 (UNCTAD, 2015).

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 1.2 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 1.2: China's FDI ownership restriction, as coded in Pandya (2010). 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 Zebregs and Tseng (2002) 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. Intuitively, if we observe that country $j$ welcomes firms $i_1, i_2, \ldots, i_n$ to invest but not others, we can compare the characteristics of firms $i_1, i_2, \ldots i_n$ with the others to infer country $j$'s preference.

### *1.3.2 Estimating countries' preferences for FDI quality*

Laura Alfaro: is all FDI equal? - Use human capital from German firms as a proxy for all sectors (this works for the OECD sample of that paper, but not more generally) - Use IPA policy, but this could just be image building by the country (everyone says that they want advanced manufacturing (cite the picture))

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, UNCTAD (2015) 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 (Ricupero, 2000). Since 2006, China's official FDI policy has been "quality over quantity," promoting FDI with intense R&D in high-productivity sectors (Guangzhou, 2011). 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, two data limitations prevents researchers from doing so. First, FDI flow data typically does not disaggregate into types of FDI. Limiting the sample to the OECD allows Alfaro (2003) to use FDI data broken into sectors—yet it remains problematic to say that FDI from an entire sector is high quality or not. Alfaro and Charlton (2007) attempt to get around this problem by using German's skill intensity by sector as a proxy for the FDI quality of that sector. To do so is to assume that 1) Germany's sectoral variation is the same as everyone else's in the OECD, and 2) there is little variation within a sector. Both assumptions are untenable.

Second, even if we can differentiate different types of FDI, it remains an open question how to estimate countries' preference for them. Alfaro and Charlton (2007) use information from IPAs' website and survey response as a proxy for their countries' preference—if an IPA lists an industry as a "target industry," it is considered that they want to attract that type of FDI. While this approach seems reasonable at first blush, Figure 1.3 shows that there is little variation in what IPAs' claim to be their target industries: because investment promotion is mainly a marketing exercise, almost everyone claims that they target high-tech manufacturing. In addition, if we use IPAs as a proxy for countries' preferences, we should also model the selection process in which the countries that decide to establish an IPA may not be the same as those who do not. Both of these issues are not addressed by (Alfaro and Charlton, 2007), and we are still in need of a good way to estimate countries' preference for different types of FDI.

We can address this challenge using firm-level data, giving us information on not only a firm's sector but also its Research and Development (R&D) expenditure or export intensity. These measures are firm-specific and get closer to what countries are looking for in FDI projects. Using R&D expenditure or export intensity as firms' characteristics in the two-sided matching model, I will be able to estimate countries' preferences for these traits.

### *1.3.3 Measuring MNCs' Activities*

As Kerner (2014) argues, the IPE literature on FDI is a bit of a misnomer. Political scientists are rarely interested in FDI *per se*—rather, they are interested in the activities of MNCs, which in turn, affect other important issues such as nation-state autonomy (Mosley, 2005), economic development (Moran, 1998), labor standards (Mosley and Uno, 2007), and environmental policies (Prakash, 2007). However,

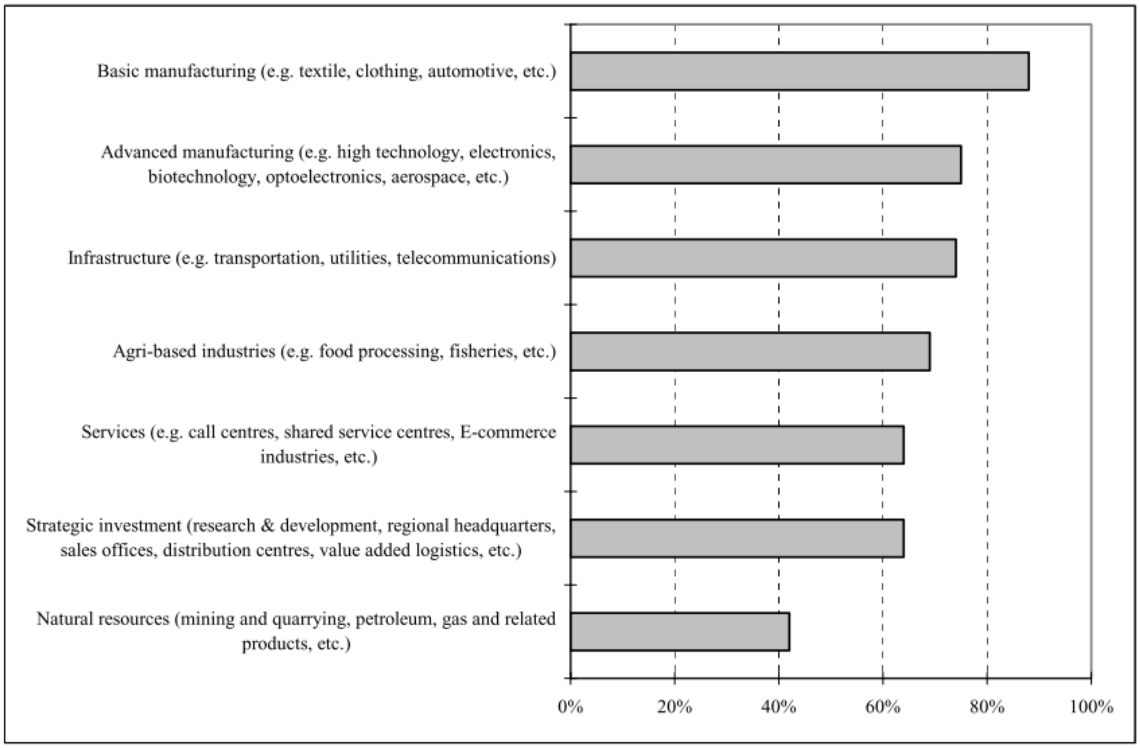


FIGURE 1.3: Target industries by IPAs around the world. Because of the image building aspect of investment promotion, almost all IPAs say that they want to attract "manufacturing," "advanced manufacturing," and "infrastructure." Therefore, using what is listed as investment priorities may not be a reliable way to measure countries' preference for FDI. Source: UNCTAD (2001)

while the theory involves MNCs as the central actor in the causal mechanism, the empirics often uses FDI flow as the variable of interest. These two concepts—the level of MNCs' affiliate activities in a country and FDI inflow into a country—are not the same.

Consider the definition of FDI from UNCTAD, the main producer of FDI data widely used by researchers:

> FDI has three components: equity capital, reinvested earnings and intra-company loans.
>
> - Equity capital, i.e. the foreign investors purchase of shares of an enterprise [in the host country].

- Reinvested earnings, i.e. the foreign investors share . . . of earnings not distributed as dividends by affiliates, or earnings not remitted to the foreign investor.
- Intra-company loans between direct investors and affiliate enterprises.

(UNCTAD, 2007, 245)

In essence, FDI data captures the amount of capital that crosses border. It is a poor proxy for the scale of MNCs' activities in the host countries because it overlooks important components of MNCs' activities while including components that are only relevant for balance of payment statistics (Beugelsdijk et al., 2010).

Consider the argument that FDI is the driver for the diffusion of labor standards across countries. Mosley and Uno (2007) theorizes that FDI can have this effect through three channels. First, MNCs may pressure the host governments for better rule of law and social programs. For MNCs to be able to effectively pressure the host governments, they must prove themselves valuable to the government by providing jobs or tax revenues. Both of these factors are only tenuously related to the amount of foreign capital inside the host country. Indeed, a MNC can employ thousands of employees, pay millions in tax, but show up as a net 0 on FDI flow data because the profit is repatriated to the foreign investor or through intra-company loans.[6] The size of the MNCs' operation is further understated because FDI statistics does not take into account capital raised locally. It also does not take into account the productivity of MNCs, which acts as an important multiplier when translating the amount of capital to the amount of output.

[6] The issue of intra-company loans is particularly fraught with issues because companies very frequently use intra-company loans to get out of paying tax in a country. These loans will be recorded on the book as a massive outflow, even though the MNC still has a large presence on the ground.

Second, scholars argue that MNCs may bring along best practices for workers' rights and spread it to local firms. If this channel operates via competition as MNCs provide better working condition forcing local firms to compete, then MNCs must employ a lot of labor for this effect to be noticeable. If this channel operates via demonstration, then it must form a lot of linkages with local firms, as suppliers and buyers, for the diffusion of norms to happen. Both the size of the labor force and the type of linkages with the local economy are not captured by FDI flow statistics.

Third, scholars argue MNCs may care more about labor quality than its cost, and thus may invest in higher wages, better benefits, or more training. Once again, for this effect to be noticeable, the size of the MNCs' labor force matters, its industry, and its investment in productivity, matters a lot more than how much capital it brings in and out of the country. In addition, non-equity transactions between the foreign parent company and the local subsidiary are not counted in FDI flow statistics, such as transfer of knowledge, technology, and management practices, thus excluding a component that is arguably much more important to labor quality than the amount of capital.[7]

This mismatch may also be a reason behind the still unsettled debate on the effect of FDI on poverty reduction. Scholars have theorized that FDI can lead to economic development and through three channels. First, MNCs can simply provide cheaper and better goods by being more productive. Second, MNCs may improve the productivity of local economy through technology transfer. Finally, MNCs can bring tax revenue to the host government, which can then spend on the poor via investment into social programs. Once again, these causal mechanisms only work depending the scale and the type of MNCs activities in the host country, not on the amount of equity capital that crosses the border. For example, productivity spillover

[7] These issues are not isolated to studies of FDI and labor standards, but are common to the whole IPE literature of the effect of FDI on policy convergence, such as environmental policies (Prakash, 2007).

is highly conditional on how thick the linkages between the MNCs and the local suppliers are as well as how technologically advanced the MNCs' activities are on the ground. The effect of FDI via tax revenue is particularly fraught with issues, as MNCs frequently engage in transfer pricing to get out of paying tax, especially via intra-company transactions of goods and services, such as charging for internal IP, whose price can be set arbitrarily by the firm (Malesky, 2015), which are not recorded in FDI flow statistics, [8]

What about studies that use FDI as the dependent variable, and are thus perhaps interested in flow of capital in and of itself?[9] The vast majority of theories on the political determinants of FDI flow relies on the "obsolescing bargain" model. Originally developed by Vernon (1971), the model is so named because the bargaining dynamics between the MNC and the host government changes over time, initially favoring the MNC and gradually tips towards the host government as the MNC commits more fixed capital on the ground. Indeed, knowing that it is costly for the MNC to uproot its increasingly large and immobile operation, the host government can unilaterally alter the original bargain, most egregiously by expropriating the MNC's asset and profit, but more often via "creeping expropriation," e.g. increased tax or tougher regulation (Li, 2009). Political economists argue that MNCs are acutely aware of the "obsolescing bargain," and thus prefer to invest in countries whose governments can make a credible commitment that they will not alter the original bargain. This means MNCs prefer countries with democratic accountability (Jensen, 2003), a federal system (Jensen and McGillivray, 2005), membership in international trade agreements (Büthe and Milner, 2008), less political risk (Beazer

[8] A similar argument is about the relationship of FDI on economic development, especially on the technology spillover and tax revenue.

[9] Arguably, political scientists are not interested in the flow of capital in and of itself, but also because of its implications for development, state autonomy, and other effects on policy. The discussion above has shown how problematic it is to study these effect of FDI using FDI flow data.

and Blake, 2011; Graham, 2010), or more veto points (Choi and Samy, 2008).[10].

The linchpin of this argument is the assumption that FDI capital is illiquid and cannot be quickly removed from the host country at will. This assumption is not fully warranted. According to the US Bureau of Economic Analysis (BEA)'s 2004 survey, 43% of US MNCs' balance sheet comprises of liquid assets that can be liquidated within one year under normal operating situations. Among the 57% of the balance sheet that are illiquid, 24% are "other non-current assets," which include non-tangible assets like brand names, trademarks, and patents—some of which are not expected to be liquidated but can be removed from the host countries. Only another 24% of the balance sheet is made up of physical capital, i.e. Plant, Property, and Equipment (PPE), which cannot be easily moved and match most closely to what we have in mind as the "illiquid capital" in the obsolescing bargain model (Kerner and Lawrence, 2014, 113).

Besides the conceptual mismatch between FDI flow and MNCs' activities, from a statistical standpoint, this measurement error may also be a contributing factor to the still unsettled debate on the effect of FDI. Even when the measurement error is random, it will inflates the standard error of our estimate when FDI is the dependent variable, and bias our estimate towards 0 when FDI is the independent variable. These effects may explain Jensen (2012)'s surprising finding that lower corporate tax rate does not lead to more FDI flow, or the mixed empirical evidence for the relationship between FDI and development (Mold, 2004, 108).

Even more worryingly, the measurement error is unlikely to be random. For example, the amount of locally raised capital, something we care about but FDI statistics does not capture, is likely to correlate with how developed the local capital market or the fluctuation in the exchange rate. On the other hand, repatriated

[10] The fact that FDI is understood as illiquid capital subject to the obsolescing bargain is the central theoretical difference between FDI and footloose equity capital (Ahlquist, 2006; Mosley and Singer, 2008)

earnings, something that does not necessarily indicate reduced MNCs' activities but is recorded as an outflow in FDI statistics, is likely to correlate with the tax rate of not only the host country but also the tax rate of tax havens that the MNC may have an affiliate in.[11] [12]

To deal with the measurement error problem, scholars have tried to use measurements that are closer to the theory than FDI flow. Given that political scientists are interested in MNCs' activities, recent work emphasizes using MNCs' operational data directly. These firm-level data allow researchers to measure directly the quantities of interest. For example, re-visiting Li (2009)'s hypothesis that democracies are more attractive to MNCs, Kerner (2014) uses data on US MNCs' fixed capital expenditures to more precisely test the relationship between democratic institutions and *illiquid* capital, not just FDI in general. The author finds that there is no relationship between democratic institutions and FDI flow and stock, but there is a positive relationship between democracy and MNCs' fixed capital expenditures, confirming the theoretical argument [13]

In another example, Arel-Bundock (2017) uses ORBIS data to study the location decision of firms. However it only does one sided (i.e. only looking at the characteristics of the host countries to predict incidence of investment). This is a bit of a missed opportunities because using his Random Forest / non-parametric approach, it would

[11] See Gallop and Weschle (2017) for a recent and more comprehensive discussion of measurement error in political science research.

[12] 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. See Kerner (2014, 809) for a more in-depth discussion of FDI stock and flow data.

[13] Another alternative is to use other variable, for example when Jensen (2008b) re-examines whether MNCs favor democratic regimes because they pose less political risk, the author avoids using FDI flow and use price data of political risk insurance agencies instead. In other areas of IPE, scholars are also paying more attention to using the data that maps more closely to the theoretical argument, e.g. (Karcher and Steinberg, 2013).

have been possible to incorporate characteristics from firms. Then the random forest would be able to take into account the interactions between the firms' characteristics and country characteristics (in the form of sequential tree split). Even then, since random forests do not produce interpretable coefficients, this black-box approach does not allow us to understand the preference of actors, how these preference are correlated with other characteristics, and how they may evolve over time. The only claim he can make is whether some factors add predictive power over other factors.

In sum, while the need to use better data is clear, and while firm-level data has become more abundant in recent years,[14] political scientists have not developed a model to estimate this 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 (Dorff and Ward, 2013).

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 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

[14] 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. However, there are concerns about its data quality, given that the data is collected via public governmental or municipal sources. Due to the differences of reporting across jurisdiction, the data quality is much less consistent than the Japanese Overseas Survey. For the US, there is a census of US firms overseas, which should be similarly high quality. However, this dataset requires citizenship.

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.

# 2

# Two-sided matching model

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 of

[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 (Rysman, 2009).

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 describe the two-sided logit model, first developed by Logan (1996) to study the labor market, and how I use a Bayesian MCMC approach to estimate it.

## 2.1 Game theory models of matching markets

Gale and Shapley (1962) 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* $\mu$, 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 $\mu$ 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 $\mu$. 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 (Gale and Shapley, 1962). 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, Roth and Vate (1990) 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, Adachi (2003) 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

[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.

further suggest that the matching we 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 Roth and Sotomayor (1992). 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.

## 2.2 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 (Posner, 2001; Sprecher et al., 1994). 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.

In addition to academic research in two-sided markets, recently there has also been substantial commercial interest in studying them as the Internet witnesses 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.[5] 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 Tu et al. (2014), which uses the Latent Dirichlet allocation (LDA) model to uncover the latent types of users based on their activities on an online dating platform.[6] 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

[5] 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.

[6] Besides Tu et al. (2014), Hitsch et al. (2010); Goswami et al. (2014) 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.

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 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).

## 2.3 Two-sided logit model

Here I present a behavioral model of the two-sided matching market, focusing on the case of many-to-one matching, proposed by Logan (1996). 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 2.1, 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, 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 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 section 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.

## 2.4 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 $\beta_j$ 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 $\epsilon_{1ij}$ 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 $\epsilon_{0ij}$ 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 $\beta$'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.[7] 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

[7] In the terminology of Roth and Sotomayor (1992), firms are assumed to have "substitutable preference," or firms' preference is assumed to have the property of substitutability. As discussed in Section 2.1, 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.[8]

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.[9] 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:

[8] While not concerned with diminishing marginal productivity, Roth and Sotomayor (1992) 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.

[9] 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 (Train, 2009).

$$Pr(o_{ij} = 1) = Pr(U_j(i) > U_j(\neg i)) \tag{2.3}$$

$$= Pr(\epsilon_{0ij} - \epsilon_{1ij} < \boldsymbol{\beta}_j' X_i - b_j) \tag{2.4}$$

$$= \frac{\exp(\boldsymbol{\beta}_j' X_i)}{1 + \exp(\boldsymbol{\beta}_j' X_i)} \tag{2.5}$$

$$Pr(o_{ij} = 0) = 1 - Pr(o_{ij} = 1) = \frac{1}{1 + \exp(\boldsymbol{\beta}_j' X_i)} \tag{2.6}$$

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."[10]

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

$$p(O_i|\boldsymbol{\beta}) = \prod_{j \in O_i} p(o_{ij} = 1|\boldsymbol{\beta}) \prod_{j \notin O_i} p(o_{ij} = 0|\boldsymbol{\beta}) \tag{2.7}$$

$$= \prod_{j \in O_i} \frac{\exp(\boldsymbol{\beta_j'} X_i)}{1 + \exp(\boldsymbol{\beta_j'} X_i)} \prod_{j \notin O_i} \frac{1}{1 + \exp(\boldsymbol{\beta_j'} X_i)} \tag{2.8}$$

## 2.5 Modeling workers' decision making

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

$$V_i(j) = \alpha' W_j + v_{ij} \tag{2.9}$$

[10] 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.

where $\alpha$ 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.

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 $\alpha$ 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.[11] 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.[12]

Similar to our model of firm's utility, our model of worker's utility has three

[11] 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."

[12] 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.

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 (Cameron and Trivedi, 2005):

$$p(A_i = a_i | O_i, \alpha_i) = \frac{\exp(\alpha' W_{a_i})}{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j)} \tag{2.10}$$

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

## 2.6 Model estimation

Our goal is to estimate the preference of firms and workers, i.e. $\beta_j$ and $\alpha$. 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).

Logan (1998)'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) (Dempster et al., 1977). 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.[13] Second, our MCMC approach can be faster than EM when the latent variable, i.e. the opportunity set, is high dimensional (Rydén, 2008).[14] 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.

### 2.6.1 *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.[15] Specifically,

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

$$\propto p(A|O, \alpha)p(O|\beta)p(\alpha)p(\beta) \tag{2.12}$$

[13] Jamshidian and Jennrich (2000) 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.

[14] 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 Logan (1998) to reduce the data dimension by aggregating 17 employers in the data into 5 employer types, e.g. professional or blue collar jobs.

[15] See Tanner and Wong (1987) for a discussion of such data augmentation techniques.

where $p(A|O, \alpha)$ is derived in (2.10), $p(O|\beta)$ is derived in (2.8), $p(\alpha)$ and $p(\beta)$ are prior distributions for $\alpha$ and $\beta$. 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.

Because the opportunity set $O$ is a discrete matrix of 0's and 1's, there is not any convenient conjugate model for (2.12), 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 $\theta$
2. Generate a proposal value $\theta^*$ 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 $\theta^*$ 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 (Haario et al., 1999, 2001).[16]

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

[16] Description of the Adaptive Metropolis procedure?

derivation and implementation of the acceptance ratio is correct using the unit-testing approach suggested by Grosse and Duvenaud (2014).[17]

### 2.6.2 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 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

$$MH_O = \frac{p(O_i^*|A_i, \alpha, \boldsymbol{\beta})}{p(O_i|A_i, \alpha, \boldsymbol{\beta})} \tag{2.13}$$

$$= \frac{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j) \pm \exp(\alpha' W_{j*})} \times \exp(\pm \boldsymbol{\beta}'_{j*} X_i) \tag{2.14}$$

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

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

$MH_O$ 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 $\boldsymbol{\beta}'_{j*} X_i > 0$.

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

At the beginning of the MCMC chain, we propose a new $\alpha^*$ using a Normal proposal distribution centered on the current value $\alpha$ 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 (Haario et al., 2001).

The Metropolis acceptance ratio for the proposed $\alpha^*$ is[18]

$$MH_\alpha = \frac{\alpha^*|A, O, \beta)}{p(\alpha|A, O, \beta)} \tag{2.15}$$

$$\begin{aligned}\log MH_\alpha = \sum_i \Big[ &(\alpha^* - \alpha)' W_{a_i} + \\ &\log\left(\sum_{j:j \in O_i} \exp(\alpha' W_j)\right) - \log\left(\sum_{j:j \in O_i} \exp(\alpha^{*\prime} W_j)\right)\Big] + \\ &\log p(\alpha^*) - \log p(\alpha)\end{aligned} \tag{2.16}$$

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

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

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

The Metropolis acceptance ratio for the proposed $\beta$ is

$$MH_\beta = \frac{p(\beta * |A, O, \alpha)}{p(\beta|A, O, \alpha)} \tag{2.17}$$

$$\log MH_\beta = \sum_i \left[ \sum_{j \in O_i} \left( \beta_j^{*\prime} X_i - \beta_j^\prime X_i \right) + \sum_j \left( \log(1 + \exp(\beta_j^\prime X_i)) - \log(1 + \exp(\beta_j^{*\prime} X_i)) \right) \right]$$
$$+ \log p(\boldsymbol{\beta}^*) - \log p(\boldsymbol{\beta}) \tag{2.18}$$

### *2.6.5 Posterior of $\beta$'s hyperparameters $\mu_\beta, \tau_\beta$*

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 $\mu_\beta$ is the mean and $\tau_\beta$ is the precision.

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

Since the model is conjugate, we can sample from the posterior of $\mu_\beta$ and $\tau_\beta$ with Gibbs sampling. Their full conditional distribution of $\mu_\beta$ is:

$$p(\mu_\beta) \sim MVN(\mu_0, \Sigma_0) \tag{2.19}$$

$$p(\mu_\beta|\beta, \tau_\beta) \sim MVN(m, V) \text{ where} \tag{2.20}$$

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

$$m = (\Sigma_0^{-1} + n\tau_\beta)^{-1}(\Sigma_0^{-1}\mu_0 + n\tau_\beta\bar{\beta}) \tag{2.22}$$

The full conditional distribution of $\tau_\beta$ is:

$$p(\tau_\beta) \sim \text{Wishart}(\nu_0, S_0^{-1}) \tag{2.23}$$

$$p(\tau_\beta|\beta, \mu_\beta) \sim \text{Wishart}(\nu, S^{-1}) \text{ where} \tag{2.24}$$

$$\nu = \nu_0 + n \tag{2.25}$$

$$S^{-1} = \left(S_0 + \sum(\beta - \mu_\beta)(\beta - \mu_\beta)'\right)^{-1} \tag{2.26}$$

## 2.7 Parameter interpretation

The coefficients can be interpreted as the relative influence of a factor on the utility of the decision maker.

# 3

# Simulation results

In this chapter, I simulate the matching process of a labor market, using real data on workers' and firms' characteristics and assigned model parameters. I then apply the two-sided logit model to show that the model is able to recover the underlying parameters and to diagnose the properties of the MCMC sampling. I compare the results of the two-sided logit model with the one-sided conditional logit model, showing that the one-sided approach produces biased estimates of workers' and firms' preference. This result demonstrates how the one-sided approach, despite being the default method for analyzing two-sided markets, can be misleading and unable to disentangle the effect of one side's preference from the other side's.

## 3.1 Labor market data

To ensure that my simulation result generalizes to real situations, I use real data on workers' and firms' characteristics from the US General Social Survey (GSS), 1982-1990.[1] On one side of the matching market is 2149 workers, a representative sample of US male workers between 25 and 44 years old. Table 3.1 shows the summary

[1] I thank Professor Michael Newton and Professor John Allen Logan for sharing the dataset.

statistics for workers. On average, a worker is 33 ($\pm$ 5.7) years old and has 13 years ($\pm$ 3.1) of education. 11% of workers in our sample are non-white.

Table 3.1: Summary statistics of workers' education, age, and race. The data come from the GSS, 1982-1990, for male workers in the US between 25 and 44 years old.

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| Years of education | 2,149 | 13.103 | 3.111 | 2 | 20 |
| Age | 2,149 | 33.524 | 5.716 | 25 | 44 |
| Non-white | 2,149 | 0.113 | 0.316 | 0 | 1 |

On the other side of the matching market are five firms, representing five job categories: professional, managerial, sales / clerical / services, manufacturing blue collar, and other blue collar. Table 3.2 shows their characteristics. *Prestige* is the Hodge-Seigel-Rossi score, used in the GSS to measure the prestige of a job (Hodge et al., 1964; NORC, 2014). *Autonomy* is calculated as the odds of having a supervisor, multiplied by -1 so that a higher score is associated with a higher level of autonomy.[2] The prestige and the autonomy scores of a firm in our dataset is the average scores reported by workers who currently work in that job categories. Unemployment by itself does not have a prestige and autonomy score. Following Logan (1996)'s study on the labor matching market, I calculate them as 50% of the prestige of the last job held and as the average autonomy scores of all workers (Logan, 1996).[3]

[2] In other words, autonomy $= -\frac{P(\text{having a supervisor})}{P(\text{not having a supervisor})}$

[3] Our measurement of firms' prestige may be biased if workers that have a high regard for a certain type of job tend to work in that kind of job. Here we are making the assumption that workers are largely homogeneous so that workers in different jobs still have the same opinion regarding a job's prestige. While it is a strong assumption, it is not an extra burden since we have already assumed that workers have the same set of preferences in the modeling step. In addition, in other applications, such as the FDI market where MNCs are matched with countries, we can obtain objective data on countries' GDP, growth, and human capital, etc.

Table 3.2: Characteristics of 5 firm types in the US, 1982-1990.

| Firm category | Prestige | Pr(Supervisor) | Autonomy |
|---|---|---|---|
| Unemployment | 18 | 0.204 | -0.256 |
| Professional | 59.670 | 0.163 | -0.483 |
| Managerial | 48.141 | 0.442 | -3.237 |
| Sales, Clerical, Services | 34.545 | 0.100 | -0.112 |
| Manufacturing Blue Collar | 34.330 | 0.071 | -0.077 |
| Other Blue Collar | 34.035 | 0.175 | -0.214 |

## 3.2 Simulated matching process

I assign values to workers' and firms' preference parameters, choosing values to achieve some level of realism and to have a sample of workers in each job. Table 3.3 describes the utility functions. I normalize firms' utility of not hiring to 0 so that firm $j$ will extend and offer to worker $i$ if the utility of hiring is positive (i.e. $U_j(i) > 0$. The magnitude of the intercept can thus be interpreted as how selective a firm is in making an offer. For example, professional and managerial firms are highly selective with intercepts of $-24$ and $-22$, while the other firms are less so with intercepts of $-9$, $-8$ and $-6$. The coefficients represent how much a firm values a worker's trait. For example, managerial and professional firms have a similar preference for a worker's education, with coefficients of 1 and 1.3. On the other hand, managerial firm values a worker's age twice as much as professional firm does, with coefficients of 0.2 and 0.1.

While the two-sided logit model can be extended to accommodate different worker types as well, for this simulation I assume that workers have homogeneous preference, sharing one utility function.

Each utility function has a random component, represented by the Gumbel-distributed error term $\epsilon$. While normally distributed error is more common in simulations, justified by the claim that the error term is the sum of many independent unobserved variables, here I use the Gumbel distribution so that the coefficient estimates from the two-sided logit model can be directly compared with the true preference parameter values.[4] Practically the Gumbel distribution is very similar to the normal distribution, and I discussed the implication of using the Gumbel distribution in further depth in Section 2.4.

Table 3.3: Utility functions of firms and workers used in labor market simulation. $x_{i1}, x_{i2}, x_{i3}$ are worker $i$'s education, age, and race (nonwhite is coded as 1). $w_{j1}, w_{j2}$ are firm $j$'s prestige and autonomy, with $j \in \{1, \ldots, 5\}$. The $\epsilon$'s are Gumbel-distributed error terms.

| | |
|---|---|
| **Firms' utility functions** | |
| Professional | $U_j(i) = -24 + 1.3x_{i1} + 0.1x_{i2} + 1x_{i3} + \epsilon$ |
| Managerial | $U_j(i) = -22 + 1x_{i1} + 0.2x_{i2} + 1x_{i3} + \epsilon$ |
| Sales, Clerical, Services | $U_j(i) = -9 + 0.75x_{i1} + -0.05x_{i2} + 0x_{i3} + \epsilon$ |
| Manufacturing blue collar | $U_j(i) = -8 + 0.5x_{i1} + 0.02x_{i2} + 0x_{i3} + \epsilon$ |
| Other blue collar | $U_j(i) = -6 + 0.5x_{i1} - 0.01x_{i2} + 1x_{i3} + \epsilon$ |
| **Workers' utility function** | $V_i(j) = 0.01w_{j1} + 0.1w_{j2} + \epsilon$ |

With these utility functions, I simulate the matching process as follows:

- First stage: Each firm $j$ evaluates each worker $i$, calculating the utility of hiring. If the utility of hiring is positive, it will extend an offer. Unemployment is always an option for workers, and is thought of as a "firm" that extends an offer to every worker. After this stage, we generate an opportunity set that is a $2149 \times 6$ matrix, in which cell $(i, j)$ is 1 if worker $i$ receives an offer from firm $j$ and 0 if not. We typically do not observe this opportunity set matrix in real datasets.

- Second stage: Worker $i$ evaluates each firm that extends him an offer in the first stage, calculating the utility of working for that firm. Worker $i$ then chooses a firm (or unemployment) if it offers the highest utility. After this stage, we

[4] If I use normally distributed error terms, the coefficient estimates have to be divided by $\frac{\pi^2}{3}$ to be comparable with the true values.

generate a choice vector that is a $2149 \times 1$ vector, in which cell $i$ equals $j$ if worker $i$ decides to work at firm $j$. This choice vector is what we observe in real datasets such as the GSS.

## 3.3 Simulation results

I estimate the two-sided logit model using the MCMC approach described in Chapter 2. I use $N(0, \mathrm{sd} = 10)$ as a weak and proper prior for the preference parameters. The results below are from a MCMC chain of 50,000 iterations with a thinning interval of 5, resulting in 10,000 saved iterations. To propose new samples in the Metropolis-Hastings algorithm, I use normal distributions with a scale hand-picked by examining the trace plots of discarded runs.

Figure 3.1 shows the trace plots of the MCMC samples for workers' preference. We see that the MCMC chain mixes well and quickly converge to the true value indicated by the red line. This result gives us confidence that the MCMC algorithm is implemented correctly and can achieve convergence within a reasonable time frame.

Figure 3.2 shows the trace plots of the MCMC samples for professional firm's preference. We see that the MCMC chain is also able to converge to the true parameter values, albeit slower and with more autocorrelation between iterations.[5] There are several reasons for this poorer mixing.

First, while we can use the entire sample to estimate the preference of workers, only a subset of the sample works at a particular firm, resulting in a smaller sample that we can use to estimate each firm's preference. This problem is clearest in the trace plots for the managerial employer, which only has a sample of 40 workers, or 1.9% of the total sample. To partially combat this issue, I use a hierarchical model

[5] To improve the mixing of the MCMC chain, I standardize workers' characteristics so that they have mean 0. Therefore, the intercept term has to be changed accordingly. The true intercept values displayed in the plots are the standardized intercepts, which is different from those reported in Table 3.3.

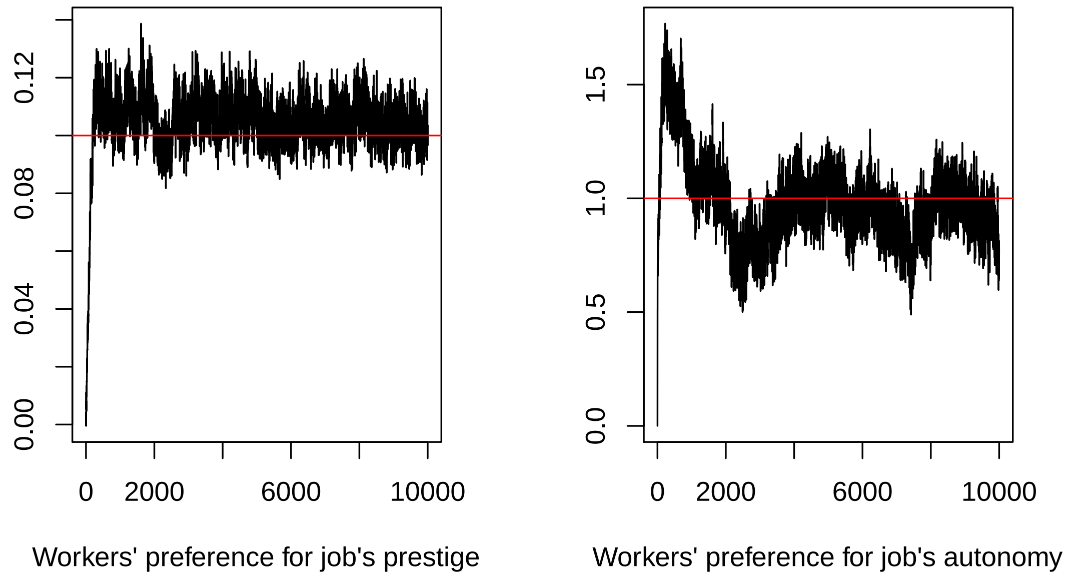


FIGURE 3.1: Two-sided logit estimates for workers' preference. The black line plots are the trace plots of the MCMC samplers, and the red line indicates the true parameter values. The trace plots show that the MCMC chain is able to converge to the true value after 10,000 iterations (2,000 saved iterations $\times 5$ thinning interval).

in which firms' preference parameters are drawn from a common distribution. By doing so, I "partially pool" the sample across firms, pulling the estimate for firms with small sample sizes towards the common mean, and thus producing estimates that have more predictive power (Gelman and Hill, 2006). On a related note, the MCMC chain of the preference parameter for *non-white* has a particularly poor mixing, likely because *non-white* is a binary variable, having less variation and thus information that our model can use.

Second, while workers only have two preference parameters (for firm's prestige and autonomy), each firm has four preference parameters (for worker's education, age, race, and an intercept term), resulting in a total of 24 parameters. Updating the MCMC chain in such high dimension is inherently difficult—to update one parameter

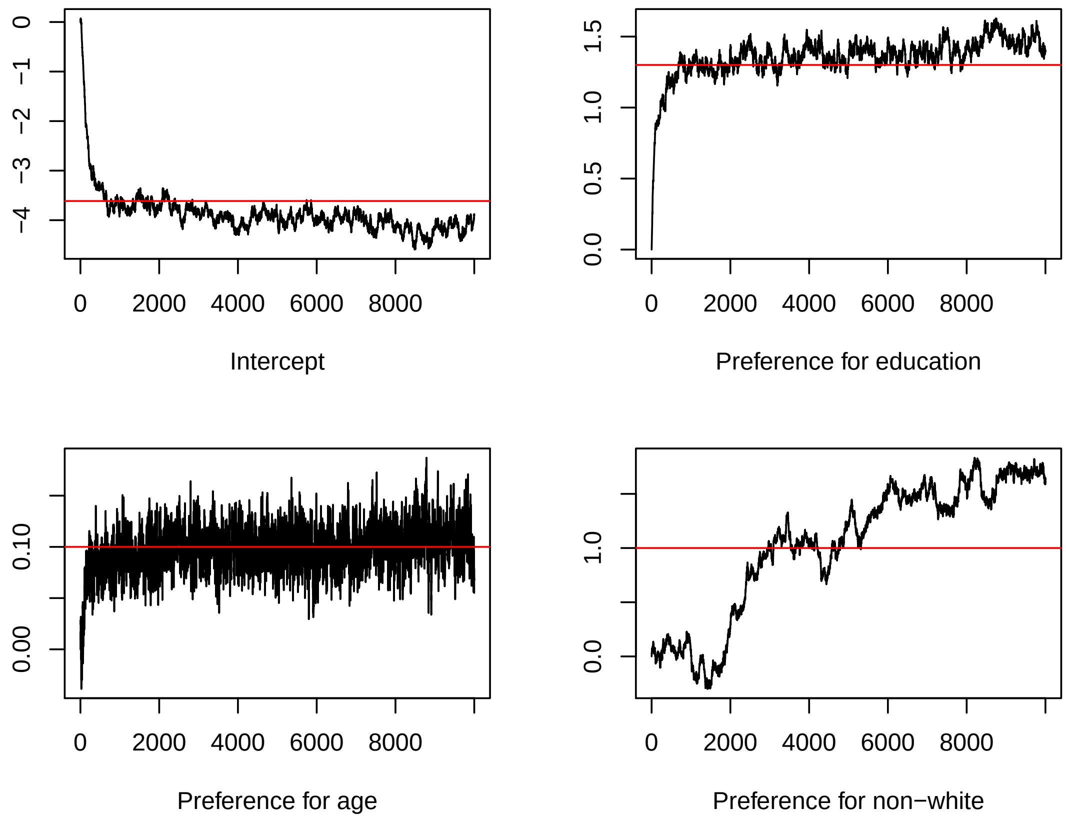


Figure 3.2: Two-sided logit estimates for professional firm's preference. The MCMC chain is able to converge to the true parameter value, indicated by the red line, albeit with more autocorrelation than the MCMC chain for worker's preference in Figure 3.1.

we only need to come up with one good proposal, but to update 24 parameters we need to come up with good proposals for each of them.

Third, while firms' preference and the opportunity set are highly correlated, our proposals for these parameters are independent, not take into their correlation, and thus causing the MCMC to get stuck at local mode. Section 3.5 discusses this issue and potential remedies in more details.

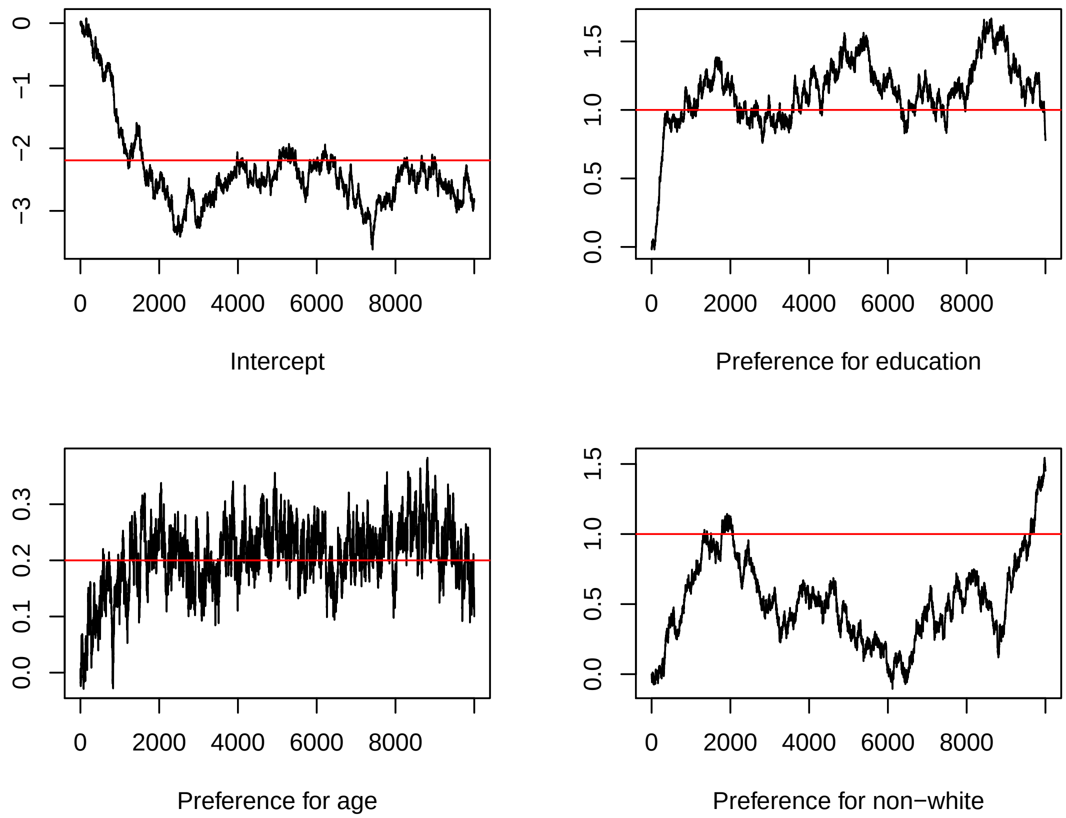


FIGURE 3.3: Two-sided logit estimates for managerial firm's preference. Because the managerial firm only has a small sample size of 40 workers, or 1.9% of the total sample, its preference is estimated more poorly than other's.

## 3.4 Comparing two-sided logit model and one-sided models

In this section, I demonstrate that, without taking into account the two-sided nature of the matching market, one-sided models produce biased estimates of the actors' preference. While it may be unsurprising that one-sided models fail when the data generating process is so different from their assumptions, this is a worthwhile exercise given that many empirical researches rely on these models. For example, using discrete choice models (multinomial logit, conditional logit), Cheng and Stough (2006) models Japanese MNCs' location choice across Chinese provinces and Aw and Lee (2008) models Taiwanese firms' decision to stay home or to open a factory in China and the US.[6] Using count models (Poisson, negative binomial), Wu (1999) models MNCs' location choice in Guangzhou, China. Arauzo-Carod et al. (2010) provides a literature review of how these methods are used in studying the location choice of firms.

I estimate a conditional logit model in which workers choose the best firm to work for as if all firms were available in their opportunity set. This assumption is not satisfied by our two-sided data generating process. Figure 3.4 shows that the one-sided conditional logit model is not robust when this assumption is violated, producing biased estimates of workers' preference. Worse yet, its estimate has little uncertainty and can cause researchers to be overly confident in the wrong result.[7]

Examining the big difference between the two-sided and one-sided estimates for *prestige* demonstrates a situation in which the one-sided approach confounds one side's preference with the other's. Figure 3.5 (left) shows the binary heat map for the true opportunity set—a dark blue cell indicates that an offer is made by firm in column $j$ to worker in row $i$. The columns for professional and managerial firms are quite similar, reflecting the fact that they make offers to the same kind of workers. In contrast, in the observed choice (Figure 3.5, right), the columns for professional and managerial firms are very different, reflecting the fact that the professional firm is slightly more desirable, causing workers that receive offers from both firms to overwhelmingly choose to work for the professional firm over the managerial firm. Therefore, there are very few workers at the managerial firm. To the one-sided conditional

[6] In the empirical literature, researchers often use the term "multinomial logit" and "conditional logit" interchangeably to refer to a discrete choice model of unordered choices. In this discussion, I follow the terminology in McFadden (1974)'s seminal paper on discrete choice models, distinguishing "multinomial logit" as the model whose independent variables are the choosers' characteristics, and "conditional logit" as the model whose independent variables are the choices' characteristics.

[7] This conditional logit model is equivalent to a Poisson model in which the dependent variable is the count of workers at each firm, as shown in Guimaraes et al. (2003). Both models, estimated with MLE, would produce exactly the same estimates for the coefficients and their covariance matrix. Therefore, the argument against one-sided logit applies fully to Poisson.

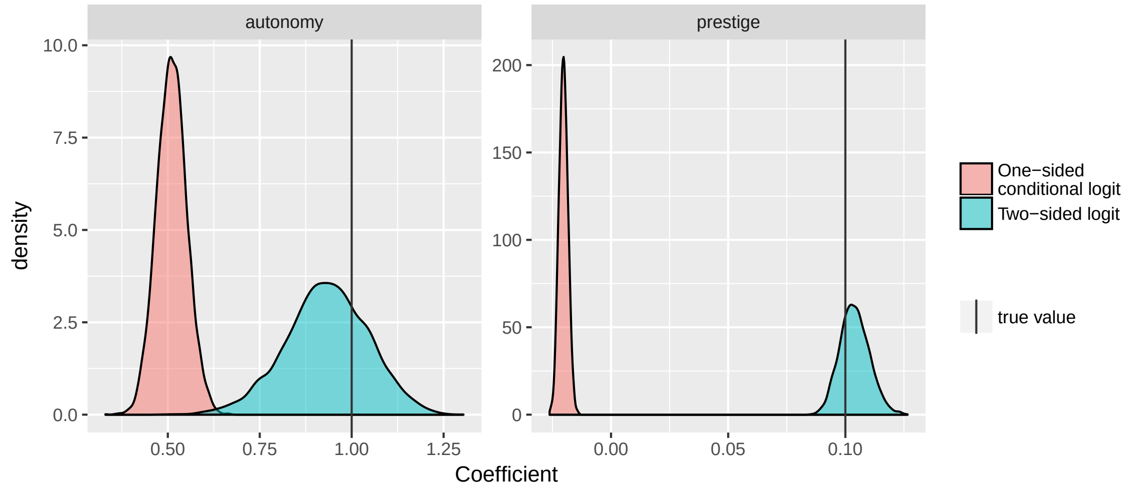


FIGURE 3.4: Estimates of workers' preference, produced by two-sided logit and conditional logit. The density plots show that the two-sided logit's 95% credible interval includes the true value, indicated by the black line, while the conditional logit's 95% confidence interval is far from it.

logit model, it looks as if the managerial firm—a highly prestigious job—were less desirable than even the services and blue collar firms. Therefore, it severely underestimate workers' preference for *prestige* to such an extent that *prestige* is considered a negative trait. This example shows how misleading it can be to estimate workers' preference by assuming that all the choices are available. Indeed, the managerial firm is rarely chosen not because it is undesirable, but because it has to compete with the professional firm for the same pool of highly educated and experienced workers.

## 3.5 Issues with MCMC convergence

The MCMC chain for firms' preference parameters $\beta$ is poorly mixed because it is highly correlated with the opportunity set. Intuitively, at any point during the MCMC, we cannot propose a new opportunity set that is very different from the current one because it would be too unlikely given the current value of $\beta$. Likewise, we cannot propose too different a value for $\beta$ because it would be rejected given the current opportunity set.

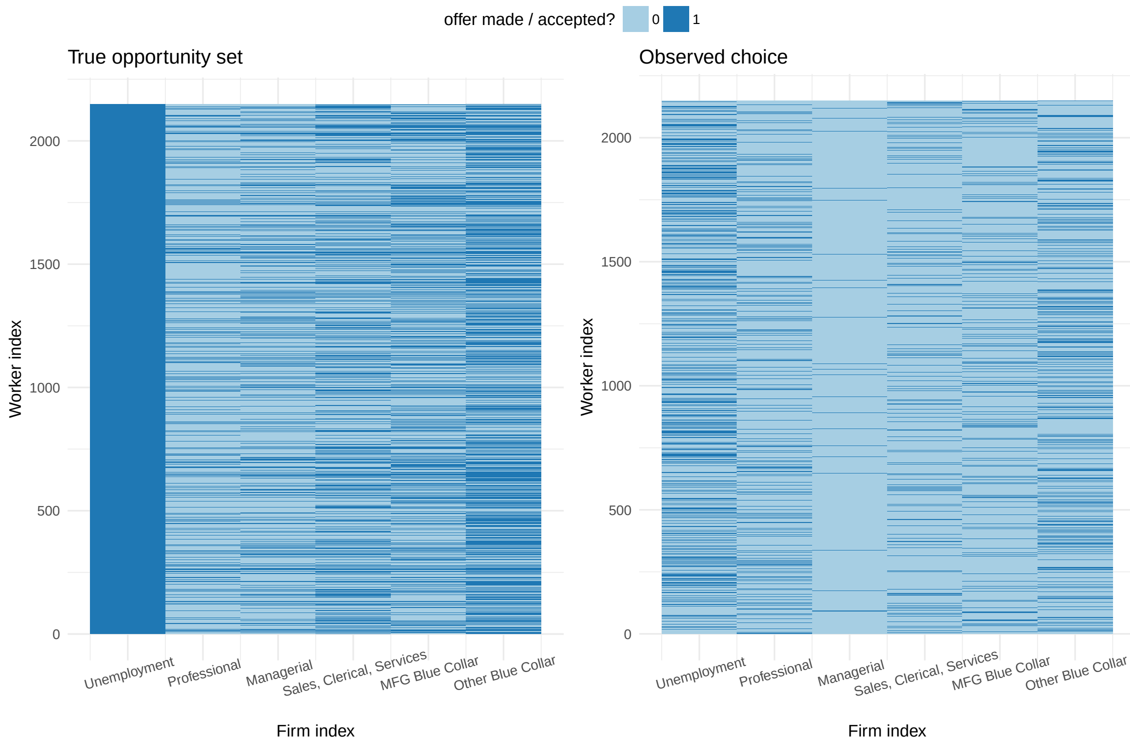


FIGURE 3.5: Binary heat map for the true opportunity set (left) and observed choice (right). A dark blue cell indicates that an offer was made (or accepted) between the firm in the corresponding column and the worker in the corresponding row.

This problem is especially severe for firms with only a few workers. If we propose a new opportunity set in which those firms now extends the offer to a new worker, then this new worker will heavily affect the estimate for $\beta$ especially if it is different from the current workers. In contrast, for firms with a large sample size, there is already a lot of information to precisely estimate their preference. Making one new offer in these cases will not substantially change the estimate.

Currently, I make random-walk proposals for $\beta$ and the opportunity set, which insufficiently takes into account this correlation, causing poor mixing. A potential solution to this problem is to make a correlated proposal for $\beta$ and for the opportunity set: if we propose a new $\beta$ that puts a high emphasis on workers' education, then we should also perturb the opportunity set to make more offers to highly-educated

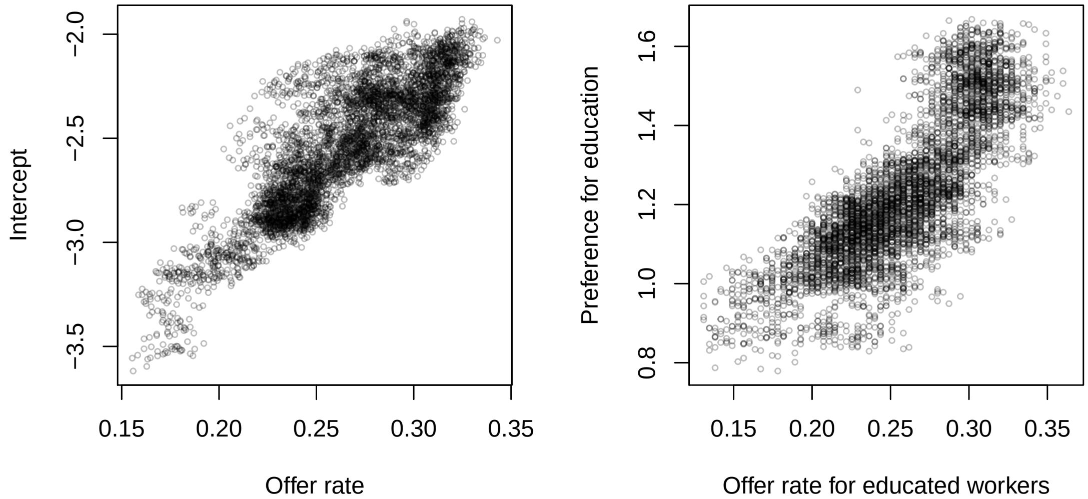


FIGURE 3.6: Correlation between the opportunity set and $\beta$

worker. While conceptually simple, this approach is not straightforward to implement, and is left for future research.[8]

## 3.6 Estimation issues when the reservation choice is unobservable

An important difference when we apply the two-sided matching model from the labor market to the FDI market is that in the latter we often will not observe the "reservation choice," i.e. the choice that will always be available to the worker or the MNC, regardless of what the other side offers.[9] In the labor market, this "reservation choice" is unemployment. In the FDI market, it is staying in the home country and not opening up a subsidiary abroad. Since most FDI data is collected by surveying firms that have made an investment abroad, we are not observing the firms who

[8] Alternatively, we may reparameterize the model entirely and eliminate the opportunity set, whose binary nature makes it impossible to use more modern MCMC approach such as Hamiltonian Monte Carlo. A potential alternative parameterization is Logan et al. (2008)'s, which samples directly from the utility space.

[9] I call unemployment the "reservation choice" in reference to the "reservation wage" in game theory and economic models.

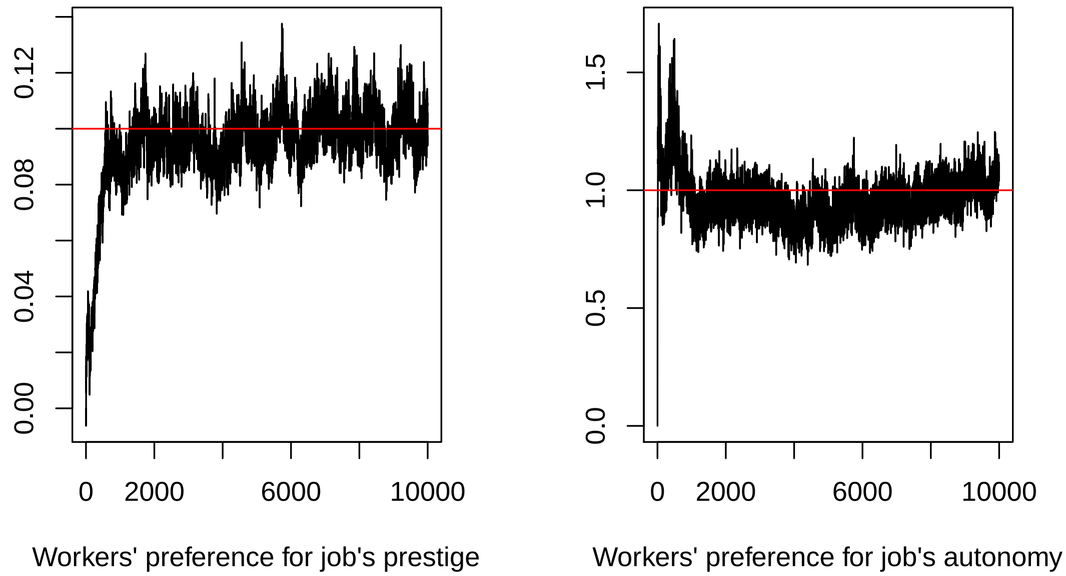


FIGURE 3.7: Estimates of workers' preference are unchanged even when we do not observe workers that choose unemployment (i.e., the reservation choice).

consider investing abroad but decide to stay put. This section investigates how missing this information affects the estimates of our model.

Intuitively, we have a sample selection problem where we only observe firms who have made it abroad. This problem has difference consequences depending on whether we are estimating firms' preference or countries' preference.

Figure 3.7 shows that our estimates of workers' preference are unaffected. This makes sense because just looking at the decision of employed workers choosing one offer over another, we still get information about their preference. Not observing the workers who decide to stay unemployed simply reduces our sample size, but otherwise does not pose any problem.[10]

[10] In a sense, we avoid this problem by assuming that all firms have the same preference. Thus, it makes sense that even though we do not observe some firms, there is only one set of preference parameters and it can be estimated from the firms that we do observe.

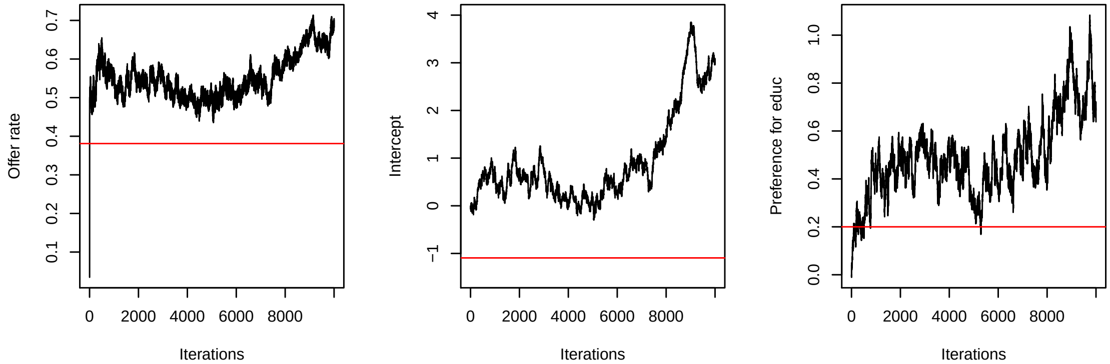


FIGURE 3.8: Estimates of firms' preference are biased when we do not observe workers that choose unemployment (i.e. the reservation choice).

On the other hand, our estimates for firms' preference can have serious bias. Figure 3.8 shows the trace plots of the preference of the managerial firm, which no longer overlaps with the true values, indicated by the red line. The estimate of the rate at which the managerial firm makes an offer is too high, as demonstrated both by the high offer rate in the first panel and the high intercept term in the second panel.[11]

This happens because the sample selection problem means that we are only observing the kind of high quality workers who receive good enough offers that they decide to stay and work instead of remaining unemployed. Since our sample only include only these good workers, our estimate of the firms' preference will make the firms look overly generous: it looks as if they extend an offer to every worker. Consequently, in our estimates, every firm will make more offers than they actually do, resulting in an estimated opportunity set that has more offers than the true opportunity set.

Another way to get the intuition around this problem is to examine the Metropolis-

[11] The firm makes an offer if the utility function is positive. Hence, if the intercept term is too high, we have too high an estimate for how often this firm makes an offer.

Hastings formula for how the opportunity set is sampled. Whenever a new offer is proposed to be added in the opportunity set, if the worker chooses to work at a bad job (or remains unemployed, which offers really low utility), then it is unlikely that this good job was really offered (otherwise the worker would have taken it). This is how the sampling of the opportunity set avoids adding spurious offers.

For this process to work well, the unemployment needs to be an option so that we can anchor other jobs against it. Essentially, by observing a lot of people who are unemployed, we know that other firms have not extended an offer to these people, and thus allowing us to better estimate their preference.

In the current situation where we do not observe the reservation choice, we no longer have this information, and thus our estimate tends to overshoot and make firms look more generous than they actually are.

$$
\begin{aligned}
MH_O &= \frac{p(O_i^*|A_i, \alpha, \boldsymbol{\beta})}{p(O_i|A_i, \alpha, \boldsymbol{\beta})} && (3.1)\\
&= \frac{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j)}{\sum\limits_{j:j \in O_i} \exp(\alpha' W_j) \pm \exp(\alpha' W_{j*})} \times \exp(\pm \boldsymbol{\beta}'_{j*} X_i) && (3.2)
\end{aligned}
$$

These findings suggest that we can make the following conclusion when the reservation choice is not available. First, the estimate of the workers' preference is unaffected. In the FDI application, it means that we can still rely on the estimates of MNCs' preference without any change. Second, given that we need a "bad" choice to anchor the estimate of firms' preference, we can still rely on the estimates for the good jobs. For these jobs, even without unemployment, there are still other worse jobs to compare to, thus allowing us a good estimate. For example, given the estimated workers' preference, professional firm is the most highly coveted job, and indeed Figure 3.9 shows the our estimates for its parameters are still correct, unlike the estimate for other blue collar.

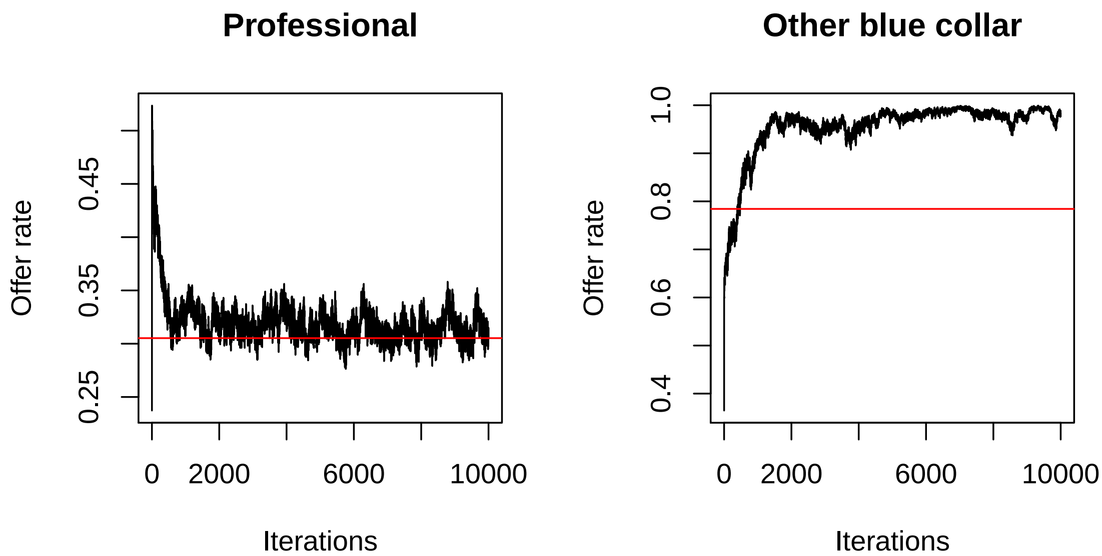


FIGURE 3.9: The estimate for the best firm is still accurate

# 4

# US labor market

In this chapter I apply the two-sided logit model on the US labor market data, where the two sided logit model is originally developed for.

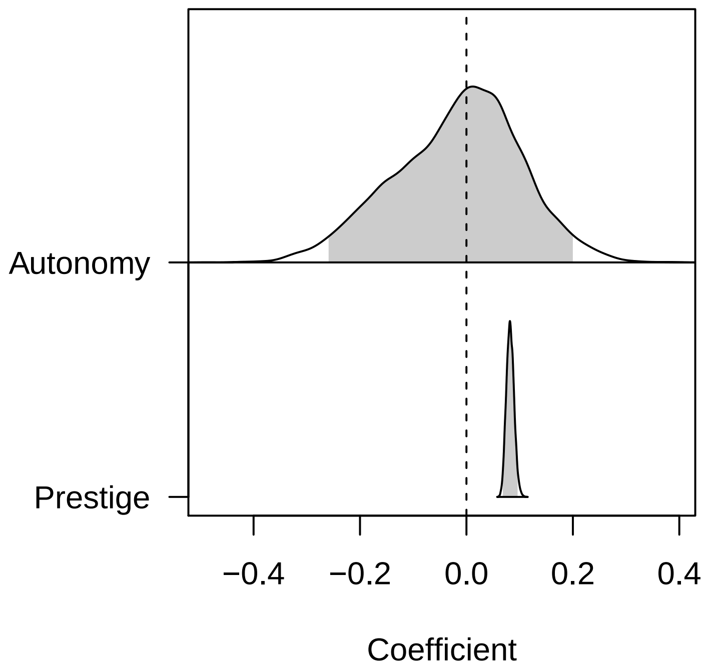


FIGURE 4.1: Preference of workers for firms' characteristics. The density plot and the shaded region show the posterior distribution and the 95% credible interval after burn-in. While prestige is highly valued by workers, autonomy seems to be less importance.

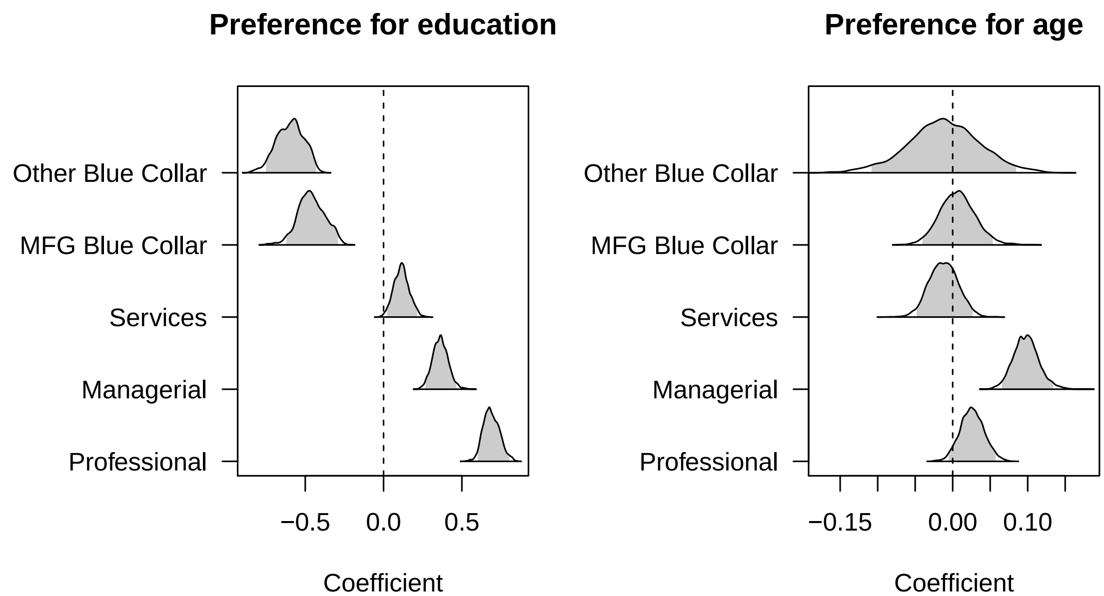


FIGURE 4.2: Preference of firms for workers' education and age. Professional and managerial firms have a strong and positive preference for highly educated workers. While most firms do not highly value older workers, except managerial firm stands out in their preference for age (likely as a proxy for experience).

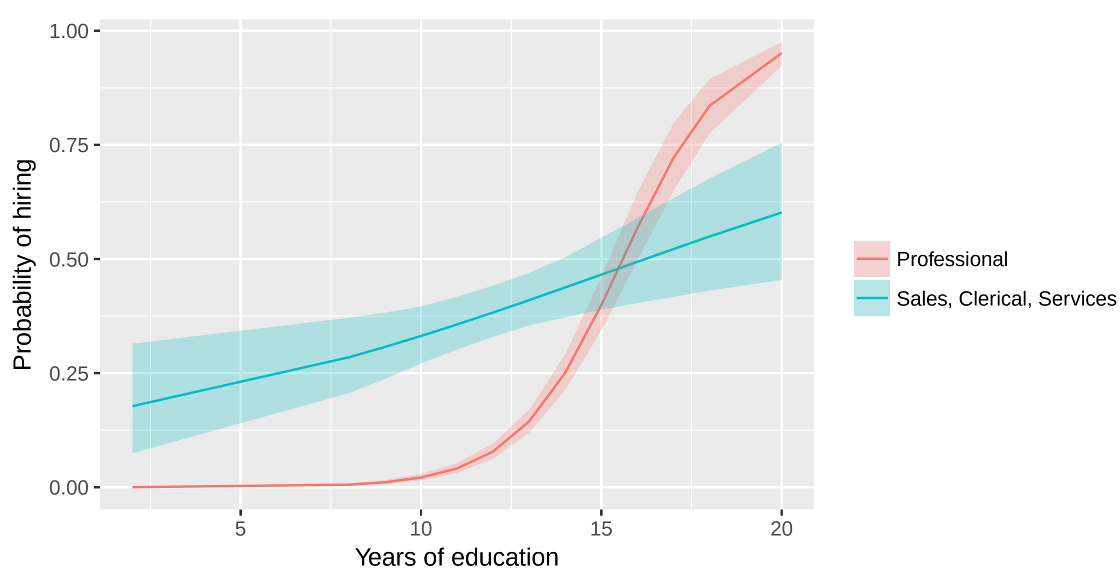


FIGURE 4.3: The effect of education on the probability of a worker being hired into a professional and a services job.

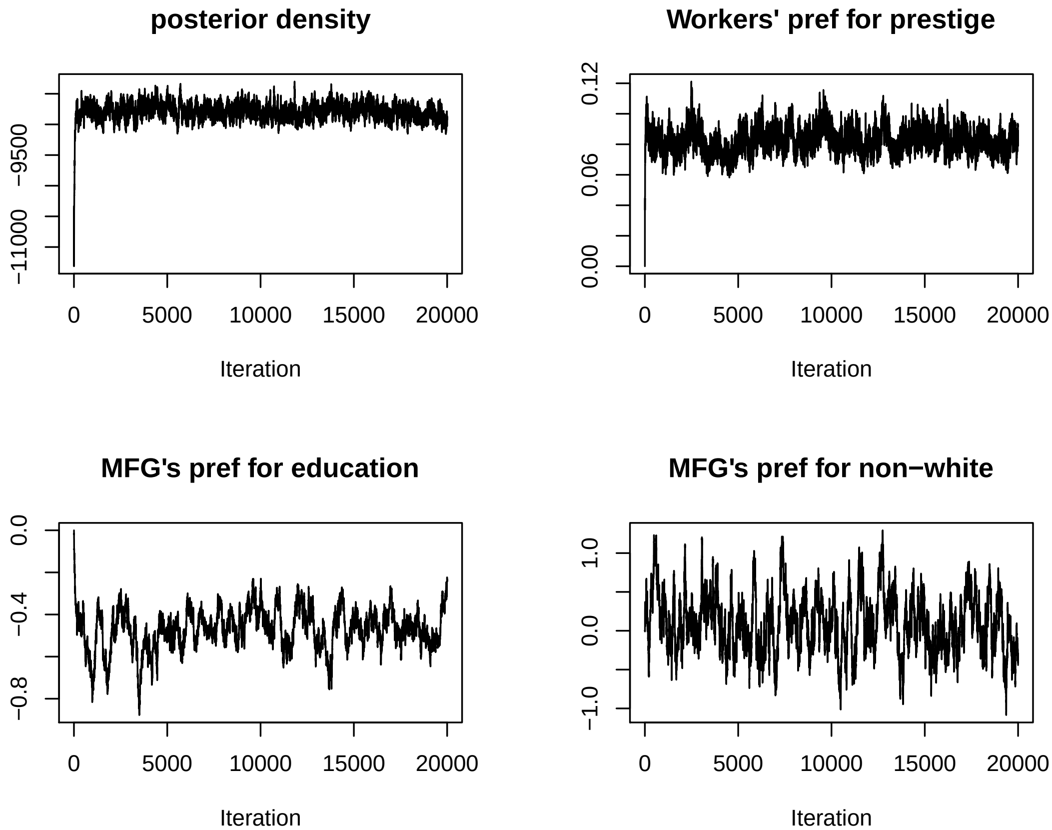


FIGURE 4.4: Trace plots of the posterior density and parameter samples, showing quick convergence.

# 5

# Two-sided matching model on Japanese FDI

In the 1980s and 1990s, Japan FDI began to surge, becoming the second largest source of FDI flow behind only the United States. An important factor behind this development is the to float the hitherto undervalued yen. As the yen began to rise in the 1980s and peak in 1995 against the dollar, Japanese companies invested heavily abroad as they could now buy foreign assets for cheap Delios and Keeley (2001).

During this period, Japanese FDI also shifted its focus away from the US and Europe, making Asia its top destination. Some scholars have argued that this wave of Japanese FDI was instrumental to Asia's economic growth by bringing not just capital but also technological know-how and the opportunity to become integrated in the global production network. In this so-called "flying geese" model of economic development, industrial development spread from Japan, the leading goose, to the rest of Asia, e.g. the Four Asian Tigers (Hong Kong, Singapore, South Korea, Taiwan), ASEAN, China, etc. (Bernard and Ravenhill, 1995; Kojima, 2000).

In this section, I apply the two-sided matching model to analyze the location pattern of manufacturing Japanese MNCs in Asia, estimating the MNCs' and countries' preference for each other.

## 5.1 Data and sample choice

The dataset was compiled by Andrew Delios from the *Kaigai Shinshutsu Kigyou Souran* (Japanese Overseas Investments-by Country), 1986-1999 editions, a biennial publication that contains operational data on all foreign affiliates of Japanese firms.[1] Tokyo Keizai, Inc. collects this data via a survey of these affiliates, which is reputed to include all Japanese firms overseas (Yamawaki, 1991). Comparing the Japanese Overseas Investment data with other data sources on publicly listed firms, (Delios and Keeley, 2001) finds that the dataset covers 98.5% of public firms, which in turn control 99.5% of the foreign subsidiaries. This high level of coverage ensures that our data captures the entire set of options available to countries and firms, obviating any worry about whether the choice set in the data represents the choice set in reality.[2]

From this dataset, I make several choices restricting the sample to better fit with the assumption of the two-sided matching model.

First, I limit the sample to manufacturing subsidiaries in Asia so that it is reasonable for our model to assume that all subsidiaries have the same set of preference parameters. Indeed, (Pak and Park, 2005) finds that Japanese FDI in the West and in the East are fundamentally different—While subsidiaries in the West seek to augment their asset via R&D and marketing, the subsidiaries in the East seek to exploit their asset by setting up local production with Japanese management. In addition, manufacturing FDI mainly consists of capital in the forms of property, plant, and

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

[2] The mismatch between the choice set in the sample and in the population is an unexplored theoretical aspect of two-sided matching models. Consider an example where we analyze a sample of 1000 men and women in the US to estimate their mate preferences. How are our estimates affected by our assumption that the potential choice set of each man includes all the women (and vice versa)? Not only does an individual not have that many acquaintances, his social circle is also not a representative sample of the entire dataset (Logan et al., 2008, 568). Fortunately, this is not a problem for our application. Given that there are only 9 Asian investment locations and approximately 200 Japanese subsidiaries being built each year, we can reasonably assume that they are all available to one another as potential options.

equipment. Therefore, our data on the size of their capital maps more precisely to the concept of illiquid capital subjecting to the "obsolescing bargain" in Political Science theories.

Second, I limit the sample to subsidiaries that are founded in the year 1996. The reason to limit the sample to subsidiaries founded in a particular year instead of including subsidiaries that have already invested is because the MNCs' utility function in our model does not capture the fixed cost of relocating. Indeed, as a linear combination of only country covariates, the utility function does not take into account the fact that, if the moving cost is too high, a subsidiary may not relocate to a new country even if the new country is available and is a better option. Therefore, the utility function in our model is not appropriate for subsidiaries that have already invested.[3]

While the decision to limit the sample to subsidiaries founded in a particular year is theoretically motivated, the decision to choose the particular year of 1996 is simply to get the largest sample size. There may be concerns about 1996 being unique as a boom time leading up to the 1997 Asian Financial Crisis. However, our sample only includes manufacturing FDI, which, unlike equity investors and land developers, were largely unaffected by the financial crisis. In addition, FDI trend remains stable before and after the crisis in terms of inflow, exit rate, and profitability (Delios and Keeley, 2001; UNCTAD, 1998). In hindsight, this is not surprising as FDI firms focus on countries' fundamentals and were thus unaffected by the fluctuations in the financial markets (Ahlquist, 2006).

The final sample includes 269 Japanese manufacturing subsidiaries in 1996, spreading across 9 Asian economies. China is the top destination, attracting 136 or 51% of

[3] Past applications of the two-sided matching model ignore this point and do not limit their sample to only agents that are participating in the matching market around the same time (Logan, 1996; Logan et al., 2008). However, given that a defining characteristic of FDI is its relative immobility compared to other form of global capital, the model's assumption of zero switching cost would be too unreasonable if we do not limit our sample.

Table 5.1: Number of Japanese manufacturing subsidiaries founded in 1996, by countries.

| nation | n | percent |
|---|---|---|
| China | 136 | 50.56 |
| Indonesia | 37 | 13.75 |
| Malaysia | 11 | 4.09 |
| Philippines | 13 | 4.83 |
| Singapore | 12 | 4.46 |
| South Korea | 8 | 2.97 |
| Taiwan | 9 | 3.35 |
| Thailand | 32 | 11.90 |
| Vietnam | 11 | 4.09 |

Japanese subsidiaries (Table 5.1).

## 5.2 Variables

For subsidiaries' characteristics that countries consider, I include:

- Capital size (in real 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 often cannot muster much domestic capital from their poor population and underdeveloped financial market.

- Labor size (number of employees): 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.

- R&D intensity (amount spent in R&D as a percentage of revenue): To measure a subsidiary's technological capability, I use the R&D intensity of the parent firm, calculated as the amount spent on R&D as a percentage of revenue. Using the parent firm's R&D intensity is a reasonable proxy because Japanese FDI in

Asia is mainly asset exploitation, i.e. implementing the know-how developed at the parent firm to the production at the subsidiary (Pak and Park, 2005).[4]

- Export intensity (amount earned via export as a percentage of revenue): Scholars have argued that economic growth in Asia is fueled by export and FDI as two mutually reinforcing forces (Liu et al., 2002). The subsidiary of an export-focused parent firm may help local suppliers become integrated into the global production network, improving the quality of their goods to match global standards and eventually being able to export themselves. Therefore, it would be interesting to test if countries actively look for investment from Japanese firms with an export focus.

For countries' characteristics that MNCs consider, I include the following variables from the Penn World Table:

- Market size (log GDP, constant 2005 USD): MNCs are expected to prefer countries with a large market size, which present MNCs with many potential customers and suppliers. In addition, market size is a key variable in the gravity model, a standard model for analyzing FDI flows Bergstrand and Egger (2007).

- Level of development (log GDP per capita, constant 2005 USD): As a measure of country income, development should attract more MNCs as MNCs prefer countries with more disposable income to consume. On the other hand, as a measure of countries' capital abundance, development should reduce FDI inflow as MNCs' capital is no longer a big advantage.

- Human capital (Penn World Table index): As one primary factor of production, labor matters greatly to firms' productivity and profit. To measure labor

[4] Another measurement of a firm's intangible asset frequently used in the FDI literature is marketing intensity (Girma, 2005). Here I focus on R&D intensity because it is the more important factor for manufacturing firms.

quality, the human capital index developed for the Penn World Table, which includes not only years of education but also labor productivity (Feenstra et al., 2015). Because the human capital index does not have a substantively interpretable unit, I standardize the variable so that it has a standard deviation of 1.

In sum, the model for the utility functions are:

- MNC $i$ utility for country $j$: $U_i(j) = \boldsymbol{\alpha}' W_j$, where $W_j$ includes log GDP, log GDP per capita, and human capital index.
- Country $j$ utility for MNC $i$: $V_j(i) = \beta_{0j} + \boldsymbol{\beta_j'} X_i$, where $X_i$ includes log number of employees, log capital size, R&D intensity, and export intensity.

## 5.3 Result

The results below are produced by an MCMC chain with $4e6$ iterations with a thinning interval of 10, resulting in $4e5$ saved iterations. The starting values for all preference parameters are set at 0. I put a diffuse prior on $\alpha$, specifically a Normal distribution with mean 0 and variance 100.

As discussed in Section 3.6, since our sample does not include MNCs that choose the reservation choice, i.e. staying home instead of investing abroad, the estimate for the $\beta$ intercept will be too high. To combat this problem, I use an informative prior so that $\beta_{0j}$ approximately follows a Normal distribution with mean -1 and variance 10.[5]

Figure 5.1 shows the posterior distribution and the 95% credible interval for MNCs' preference parameters. We can interpret the parameters as the relative weight

[5] Specifically, the prior for $\beta_{0j}$'s mean is Normal with mean -1 and variance 1. The prior for $\beta_{0j}$'s variance is Inverse-Wishart with $\nu = 7$ and $S^{-1} = 10$ so that the variance is loosely centered around 10.

MNCs attach to countries' characteristics when they decide where to invest. For example, the posterior mean for log GDP and for log GDP per capita is 0.72 and 0.66—this means that to MNCs, a 1% increase in GDP is equivalent to 0.72 / 0.66 = 1.09% increase in log GDP per capita.

The coefficients for log GDP and log GDP per capita are both positive and significant, suggesting that Japanese MNCs are looking for large markets with a lot of disposable income. On the other hand, the coefficient for human capital is negative, corroborating earlier findings in the literature that Japanese MNCs in Asia do not aim to be innovative and thus have no need for strong human capital. On the contrary, since the human capital index includes not only years of education but also labor productivity, a high human capital index may signify high labor cost, explaining why Japanese manufacturing MNEs weigh it negatively.

In addition to interpreting the coefficients as weights in the MNCs' utility function, we can also simulate and visualize their impact on MNCs' location choice. For example, we may ask if country A's GDP increases by 20%, what will be the new share of MNCs that invest in country A? Like in the one-sided conditional logit model, the share of MNCs investing in a country depends not only on country A's characteristics but also on others'. In addition, in this two-sided model, the share of MNCs investing in a country also depends on the preference of countries. For example, even if country A becomes highly desirable, we may not see much change in the share of MNCs located there if country A's preference is also highly demanding. The interaction can be much more complicated. For example, consider a scenario in which country A and country B have similar preferences and compete for the same set of MNCs—even if country A becomes more desirable than the rest of the countries, as long as country A is less attractive than country B, the share of MNCs investing in country A will still remain unchanged.

Fortunately, we can easily simulate these effects in the Bayesian framework. To

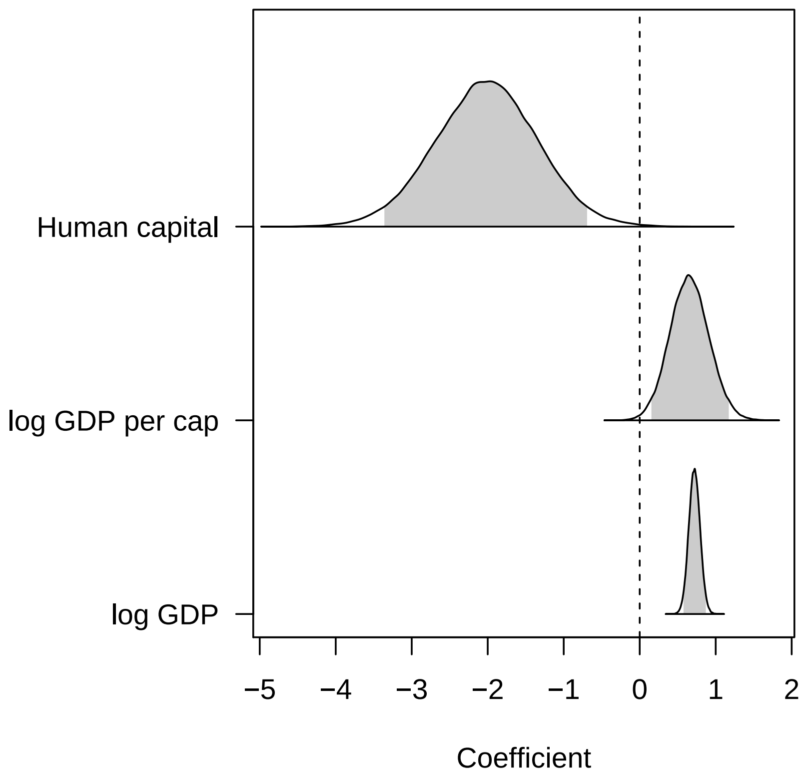


FIGURE 5.1: MNCs' preference for countries' characteristics. The density plot and the shaded region show the posterior distribution and the 95% credible interval.

see how the share of MNCs investing in Thailand changes along with hypothetical values of Thailand's GDP, we can take the following steps:

1. Construct a scenario in which Thailand has a new GDP, while all other characteristics remain for Thailand and other countries

2. Make one draw for each parameter in the model from its posterior distribution

3. Simulate the matching process in which countries first make offers to MNCs, and MNCs then choose the best option

4. Record the share of MNCs investing in Thailand after the matching process

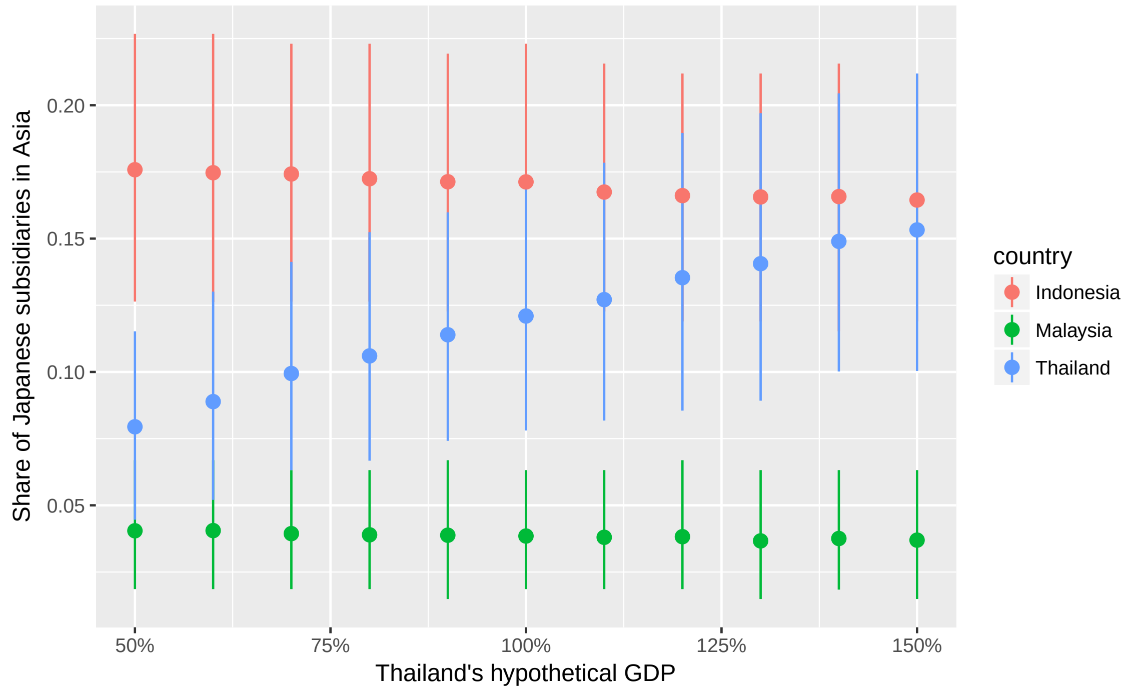


FIGURE 5.2: Effect on Thailand's GDP on its share of MNCs. The point and line range show the mean and the 95% credible interval.

5. Repeat step (2)-(4) to get a distribution for the share of MNCs in Thailand[6]

Following this process, I calculate the share of Japanese MNCs in Thailand at different hypothetical values of Thailand's GDP. Figure 5.2 shows that as Thailand's GDP increases from 50% to 150% of its true value, the share of MNCs in Thailand increases from 8% to 15%. In addition, the share of MNCs investing in Indonesia and Malaysia, two competitors of Thailand in ASEAN, declines slightly as Thailand becomes more attractive.

Similar to the interpretation of MNCs' preference parameters, we can also interpret countries' preference parameters as the relative weight that countries attach to MNCs' characteristics. Figure 5.3 shows that Taiwan wants to attract MNCs that are export-focused and discourage MNCs that employ a lot of employees. On the

[6] Essentially we are constructing the posterior predictive distribution for the share of MNCs in Thailand, integrating out the preference parameters via simulation.

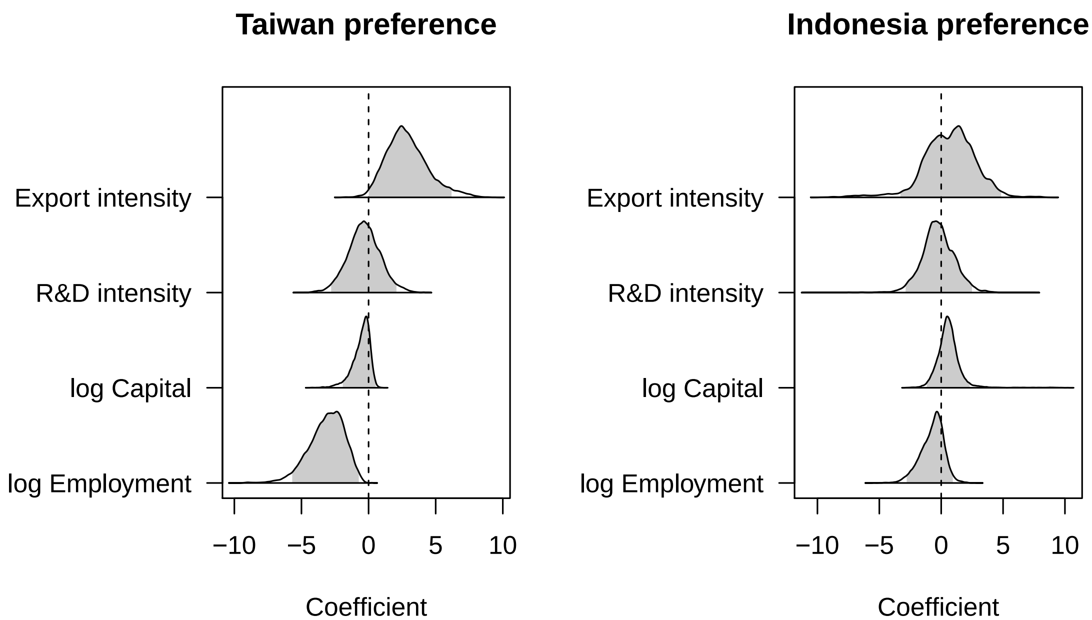


FIGURE 5.3: Preference of Taiwan and Indonesia for MNCs' characteristics.

other hand, the estimates for Indonesia's preference parameters are not statistically significant, perhaps due to our small sample size.[7]

While the relative size of preference parameters within a country is meaningful, we cannot readily compare preference parameters across countries. This is because coefficients in the logit model is normalized so that the variance of the unobserved factors in the utility function becomes 1—therefore, the relative size of coefficients across two countries may be affected by the relative variance of the unobserved factors in the utility functions of those two countries (Train, 2009, chap. 2). Since there is no guarantee that the unobserved factors will be the same or have the same variance across countries, we cannot say that, for example, Taiwan values export intensity more than Vietnam because Taiwan's coefficient for export intensity is larger.

We can, however, look at the sign of the estimates to see whether countries

[7] I present the results for Taiwan and Indonesia here because they are two countries highly desired by MNCs. As discussed in Section 3.6, the estimate for highly desirable countries are more accurate.

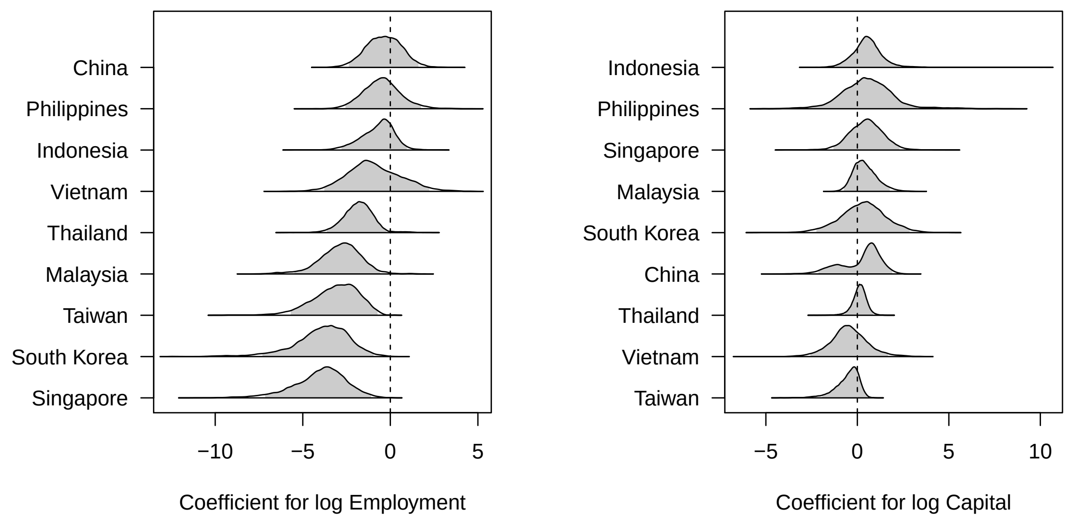


Figure 5.4: Countries' preference for subsidiaries' size, measured by their labor force (left) and capital (right). The density plot and the shaded region show the posterior distribution and the 95% credible interval.

evaluate an MNC's trait positively or negatively. Figure 5.4 and Figure 5.5 show countries' preference for labor size, capital size, R&D intensity, and export intensity. Most countries seem to dislike subsidiaries with a large labor force—this is likely because labor size is correlated with other factors that we are not capturing in our model, e.g. certain sectors in manufacturing or the level of productivity. On the other hand, several countries including Taiwan, Malaysia, and Thailand, seem to prefer subsidiaries that focus on export. This finding affirms our understanding of these economies as being export-driven.

In addition, we can simulate and visualize the impact of countries' preference on MNCs' location choice. Using the steps described above to simulate the posterior predictive distribution, Figure 5.6 shows how the distribution of Japanese MNCs across Asian countries will change if China becomes more stringent in its evaluation of MNCs. As China becomes as picky as South Korea (i.e. China's intercept is

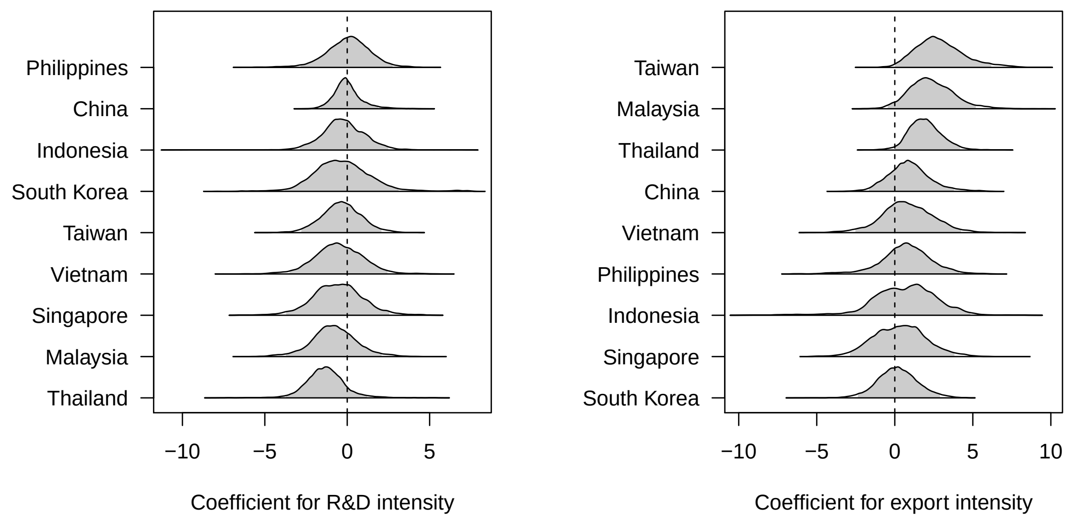


FIGURE 5.5: Countries' preference for firms' intangible assets, i.e R&D intensity (left) and export intensity (right). The density plot and the shaded region show the posterior distribution and the 95% credible interval.

decreased to match South Korea's), the share of MNCs located in China decreases from 47% to 20%, a decrease of 27 percentage points. Of these 27 percentage points, 20 go to Indonesia and Thailand, the two biggest beneficiaries of China's becoming more demanding. On the other hand, the share of MNCs in Singapore, South Korea, and Taiwan remains virtually unchanged, suggesting that China is not competing for the same MNCs as these countries do.

## 5.4 Model fit

To check that our model fits well to the data, we can conduct posterior predictive checks, generating simulated matchings and comparing them with observed matching.

Since the location of MNCs is something we model directly, we should at least expect our model to have a good fit with the observed location. Figure 5.7 shows

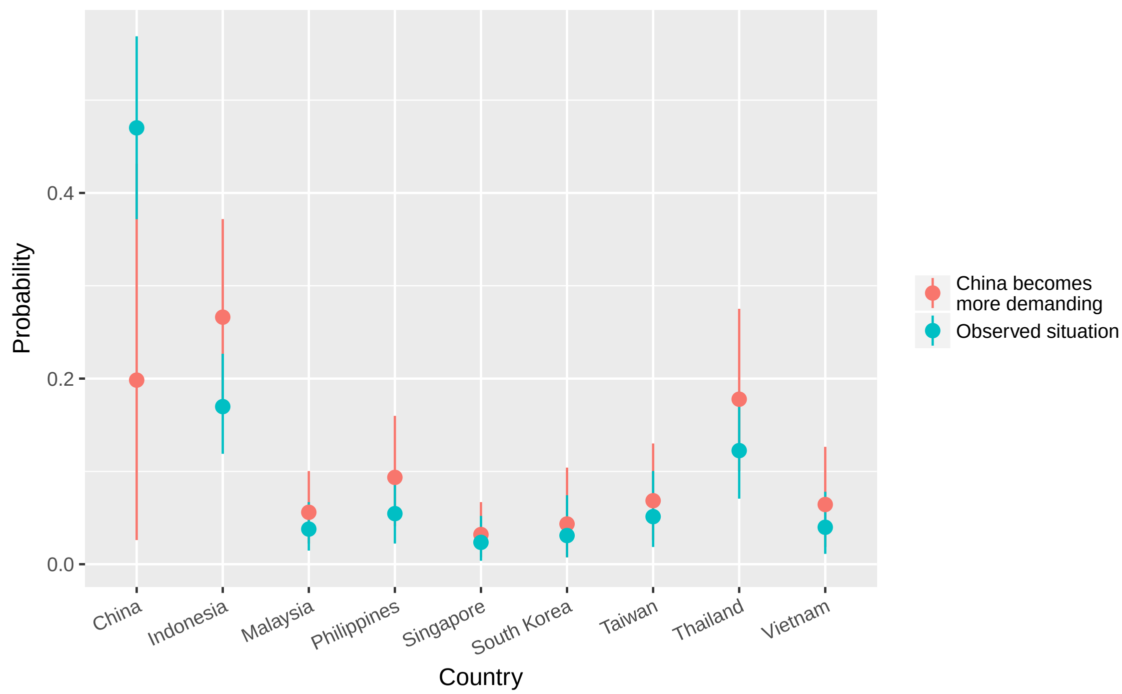


FIGURE 5.6: Effect of China's demand for MNCs on MNCs' location across Asia. If China becomes as picky as South Korea, its share of MNCs will decrease from 47% to 20%. The majority of these MNCs will relocate to Indonesia and Thailand.

that our model performs well in this regards as the predicted share of MNCs across countries match the observed share exactly in most cases and well within the 95% credible interval for all cases.

In addition, by conducting the posterior predictive checks for aspects of the data that we do not model directly, we gain a deeper understanding into what part of reality our model does not capture. Given that we may be interested in not only the share of MNCs across countries, but also which types of firms are located in which countries, I conduct the posterior predictive checks for MNCs' characteristics. Figure 5.8 and Figure 5.9 shows that our model captures the mean and the variance of of MNCs' size across countries relatively well, with the observed mean of log Employee lying within the 95% interval for all cases. Admittedly the 95% interval is really wide for many countries, reflecting the lack of precision in the estimates for

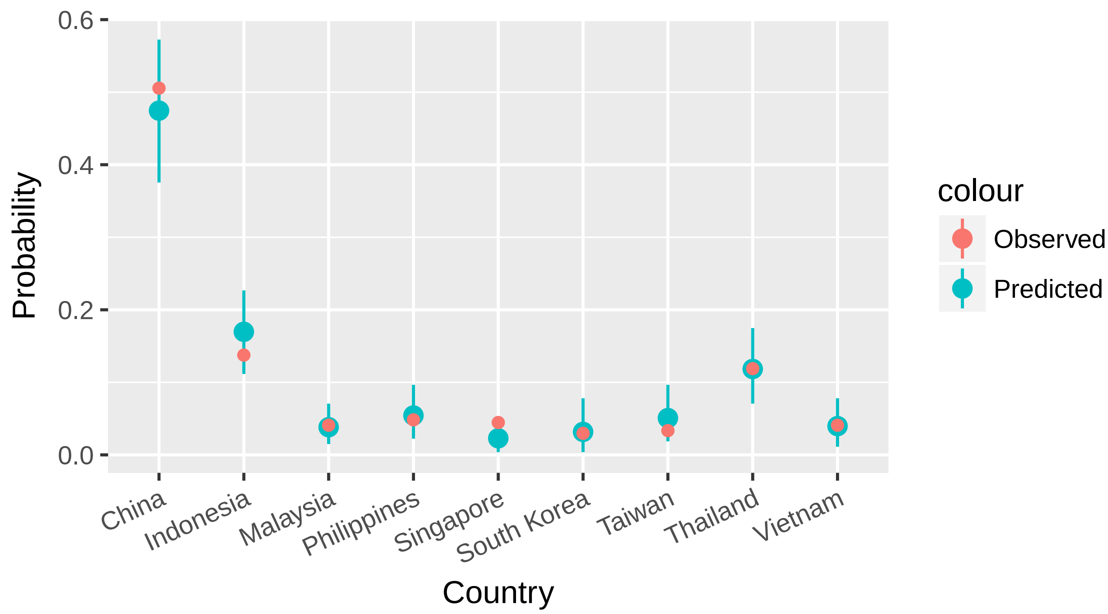


FIGURE 5.7: Predicted and observed share of MNCs across countries. The point and the error bar show the posterior mean and the 95% credible interval.

countries' preference, likely due to the small sample size.

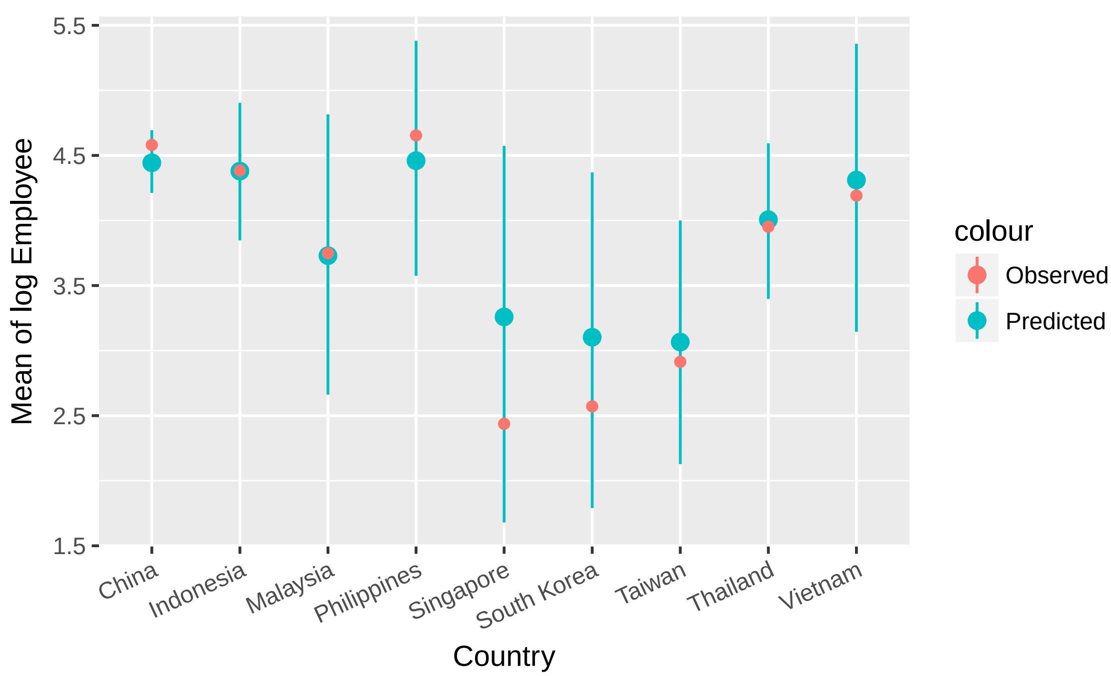


FIGURE 5.8: Average of MNCs' labor size across countries.

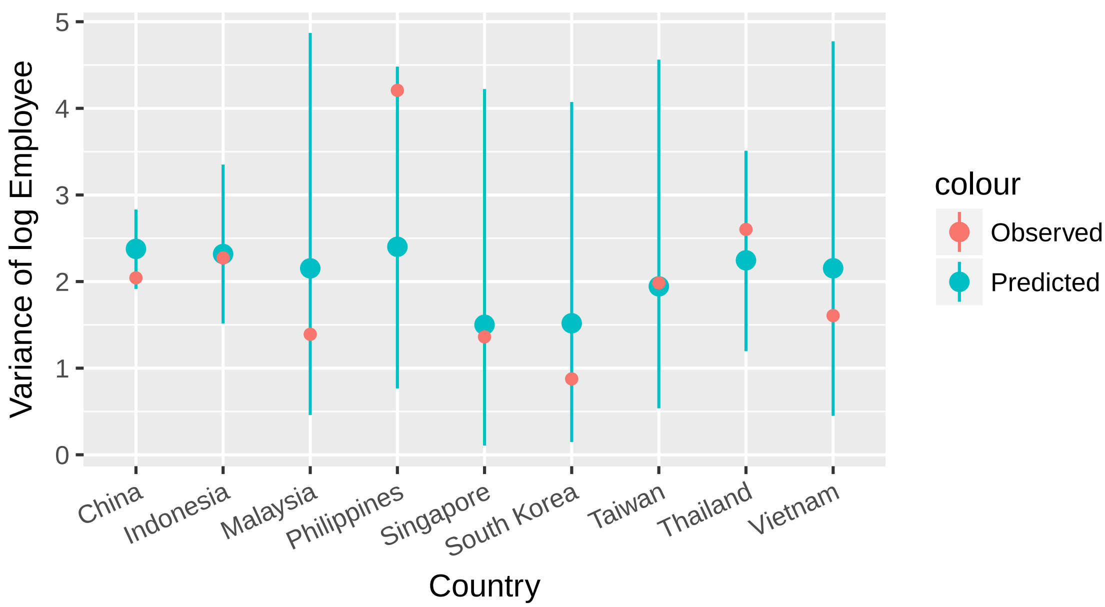


FIGURE 5.9: Variance of MNCs' labor size across countries.

# 6

# Conclusion

## 6.1 Other applications for two-sided matching model

### *6.1.1 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 (Gulati and Posner, 2016; Posner, 2001). 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 (Ditslear and Baum, 2001; Liptak, 2007).

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 (Posner, 2001). Judges select law graduates based on not only academic credentials but also, some argue, ideology, gender, and race (Slotnick, 1984). 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 (Posner, 2001; Posner et al., 2007).

One approach to estimating the preference of agents in this market is to survey clerks and judges directly (Peppers et al., 2008). 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, Bonica et al. (2017) 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 (Bonica et al., 2017, 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, Rozema and Peng (2016) 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.

### 6.1.2 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.[1]

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) (Laver, 1998). 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.[2] As the game theory literature suggests, ideologically compact coalitions are more valuable because they entail a smaller cost in terms of policy compromises (De Swaan, 1973). 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 for ideological proximity across policy dimensions.

[1] 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.

[2] 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.

# Appendix A

# Derivation of the Metropolis-Hastings Acceptance Ratio

#### *A.0.1 Opportunity sets O*

Target distribution for a firm $i$

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

$$\begin{aligned}
MH_O = \frac{p(O_i^*|A_i, \alpha, \boldsymbol{\beta})}{p(O_i|A_i, \alpha, \boldsymbol{\beta})} &= \frac{p(O_i^*, A_i, \alpha, \boldsymbol{\beta})}{p(A_i, \alpha, \boldsymbol{\beta})} \times \frac{p(A_i, \alpha, \boldsymbol{\beta})}{p(O_i, A_i, \alpha, \boldsymbol{\beta})} && \text{(A.2)} \\
&= \frac{p(O_i^*, A_i, \alpha, \boldsymbol{\beta})}{p(O_i, A_i, \alpha, \boldsymbol{\beta})} && \text{(A.3)} \\
&= \frac{p(A_i|O_i^*, \alpha)p(O_i^*|\boldsymbol{\beta})}{p(A_i|O_i, \alpha)p(O_i|\boldsymbol{\beta})} && \text{(A.4)} \\
& && \text{(A.5)}
\end{aligned}$$

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 (2.10) and (2.8)

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

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, $\alpha$*

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 (2.10),

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

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

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, $\beta$*

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 (2.8),

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

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

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

# Appendix B

## FDI terminology

Samsung, headquartered in South Korea, opened a factory in Vietnam called Samsung Electronics Vietnam.

In this scenario, Samsung Korea is the *parent company*, and Samsung Electronics Vietnam is the *subsidiary* (also the *(foreign) affiliate*). Empirically, the parent company and the subsidiary are two distinct entities (i.e. having a different number of employees, revenue size, profitability, etc.) Theoretically, they negotiate with and are evaluated by Vietnam as one entity. Therefore, in theoretical discussion, I refer to Samsung Korea and Samsung Electronics Vietnam jointly as the *MNC*.

South Korea is the *home country*, and Vietnam is the *host country* or the *local country*. I thus refer to Vietnam's economy and firms as the local economy and local firms.

# Bibliography

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

Ahlquist, J. (2006), "Economic policy, institutions, and capital flows: portfolio and direct investment flows in developing countries," *International Studies Quarterly*, 50, 681–704.

Alfaro, L. (2003), "Foreign Direct Investment and Growth : Does the Sector Matter?" *Harvard Business School*, pp. 1–32.

Alfaro, L. and Charlton, A. (2007), "Growth and the quality of foreign direct investment: is all FDI equal," .

Arauzo-Carod, J. M., Liviano-Solis, D., and Manjón-Antolín, M. (2010), "Empirical studies in industrial location: An assessment of their methods and results," *Journal of Regional Science*, 50, 685–711.

Arel-Bundock, V. (2017), "The Political Determinants of Foreign Direct Investment: A Firm-Level Analysis," *International Interactions*, 43, 424–452.

Aw, B. Y. and Lee, Y. (2008), "Firm heterogeneity and location choice of Taiwanese multinationals," *Journal of International Economics*, 76, 403–415.

Beazer, Q. and Blake, D. (2011), "It's All Relative: Home Country Risk and FDI Flows," .

Bergstrand, J. H. and Egger, P. (2007), "A knowledge-and-physical-capital model of international trade flows, foreign direct investment, and multinational enterprises," *Journal of International Economics*, 73, 278–308.

Bernard, M. and Ravenhill, J. (1995), "Beyond product cycles and flying geese: regionalization, hierarchy, and the industrialization of East Asia," *World Politics*, 47, 171–209.

Beugelsdijk, S., Hennart, J. F., Slangen, A., and Smeets, R. (2010), "Why and how FDI stocks are a biased measure of MNE affiliate activity," *Journal of International Business Studies*, 41, 1444–1459.

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.

Broz, J. L. and Frieden, J. (2001), "The political economy of international monetary relations," *Annual Review of Political Science*, pp. 317–343.

Büthe, T. and Milner, H. (2008), "The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI through International Trade Agreements?" *American Journal of Political Science*, 52, 741 – 762.

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

Carkovic, M. V. and Levine, R. (2002), "Does foreign direct investment accelerate economic growth?" *U of Minnesota Department of Finance Working Paper*.

Cheng, S. and Stough, R. R. (2006), "Location decisions of Japanese new manufacturing plants in China: A discrete-choice analysis," *Annals of Regional Science*, 40, 369–387.

Chintrakarn, P., Herzer, D., and Nunnenkamp, P. (2012), "Fdi and income inequality: Evidence from a panel of U.S. states," *Economic Inquiry*, 50, 788–801.

Choi, S.-W. and Samy, Y. (2008), "Reexamining the Effect of Democratic Institutions on Inflows of Foreign Direct Investment in Developing Countries," *Foreign Policy Analysis*, 4, 83–103.

De Swaan, A. (1973), *Coalition Theories and Cabinet Formations: A Study of Formal Theories of Coalition Formation Applied to Nine European Parliaments after 1948*, Jossey-Bass.

Delios, A. and Keeley, T. D. (2001), "Japanese Foreign Direct Investment In Thailand: An Empirical And Qualitative Post-Crisis Analysis," *Journal of International Business and Economy*, 1, 91–118.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977), "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 39, 1–38.

Ditslear, C. and Baum, L. (2001), "Selection of Law Clerks and Polarization in the U.S. Supreme Court," *The Journal of Politics*, 63, 869–885.

Dorff, C. and Ward, M. (2013), "Networks, Dyads, and the Social Relations Model," *Political Science Research and Methods*.

Drezner, D. W. (2001), "Globalization and Policy Convergence," *International Studies Review*, 3, 53–78.

Durham, J. B. (2004), "Absorptive capacity and the effects of foreign direct investment and equity foreign portfolio investment on economic growth," *European Economic Review*, 48, 285–306.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015), "The next generation of the penn world table," *American Economic Review*, 105, 3150–3182.

Findlay, R. (1978), "Relative Backwardness, Direct Foreign Investment, and the Transfer of Technology: A Simple Dynamic Model," *Quarterly Journal of Economics*, 92, 1–16.

Fu, X. (2008), "Foreign Direct Investment, Absorptive Capacity and Regional Innovation Capabilities: Evidence from China," *Oxford Development Studies*, 36, 89–110.

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

Gallop, M. and Weschle, S. (2017), "Assessing the Impact of Non-Random Measurement Error on Inference: A Sensitivity Analysis Approach," *Political Science Research and Methods*, pp. 1–18.

Gelman, A. and Hill, J. (2006), *Data analysis using regression and multilevel/hierarchical models*, Cambridge University Press.

Girma, S. (2005), "Absorptive capacity and productivity spilloversfrom FDI: A threshold regression analysis," *Oxford Bulletin of Economics and Statistics*, 67, 281–306.

Goldberg, P. K. and Pavcnik, N. (2007), "Distributional Effects of Globalization in Developing Countries," *Journal of Economic Literature*, 45, 39–82.

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.

Graham, B. (2010), "Political Risk and Diaspora Direct Investment," *APSA 2010 Annual Meeting Paper*, 42, 622–696.

Grosse, R. B. and Duvenaud, D. K. (2014), "Testing MCMC code," *CoRR*, abs/1412.5, 1–8.

Guangzhou (2011), "Implementation Opinions on Further Promoting the Work of Utilizing Foreign Investment," .

Guerra, E., de Lara, J., Malizia, A., and Díaz, P. (2009), "Supporting user-oriented analysis for multi-view domain-specific visual languages," .

Guimaraes, P., Figueirdo, O., and Woodward, D. (2003), "A Tractable Approach to the Firm Location Decision Problem," *The Review of Economics and Statistics*, 85, 201–204.

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

Haario, H., Saksman, E., and Tamminen, J. (1999), "Adaptive proposal distribution for random walk Metropolis algorithm," *Computational Statistics*, 14, 375.

Haario, H., Saksman, E., and Tamminen, J. (2001), "An Adaptive Metropolis Algorithm," *Bernoulli*, 7, 223.

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.

Hodge, R. W., Siegel, P. M., and Rossi, P. H. (1964), "Occupational Prestige in the United States, 1925-63," *American Journal of Sociology*, 70, 286–302.

Jamshidian, M. and Jennrich, R. I. (2000), "Standard Errors for EM Estimation," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 62, 257–270.

Jensen, N. (2008a), *Nation-states and the multinational corporation: A political economy of foreign direct investment*, Princeton University Press, Princeton, NJ.

Jensen, N. (2008b), "Political risk, democratic institutions, and foreign direct investment," *Journal of Politics*, 70, 1040–1052.

Jensen, N. and McGillivray, F. (2005), "Federal institutions and multinational investors: Federalism, government credibility, and foreign direct investment," *International Interactions*, 31, 303–325.

Jensen, N. M. (2003), "Democratic Governance and Multinational Corporations: Political Regimes and Inflows of Foreign Direct Investment," *International Organization*, 57.

Jensen, N. M. (2012), "Fiscal Policy and the Firm: Do Low Corporate Tax Rates Attract Multinational Corporations?" *Comparative Political Studies*, 45, 1004–1026.

Karcher, S. and Steinberg, D. A. (2013), "Assessing the Causes of Capital Account Liberalization: How Measurement Matters," *International Studies Quarterly*, 57, 128–137.

Kerner, A. (2014), "What We Talk About When We Talk About Foreign Direct Investment," *International Studies Quarterly*, pp. 804–815.

Kerner, A. and Lawrence, J. (2014), “What’s the risk? Bilateral investment treaties, political risk and fixed capital accumulation,” *British Journal of Political Science*, 44, 107–121.

Kojima, K. (2000), “The ”flying geese” model of Asian economic development: origin, theoretical extensions, and regional policy implications,” *Journal of Asian Economics*, 11, 375 – 401.

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

Li, Q. (2006), “Democracy, autocracy, and tax incentives to foreign direct investors: A cross-national analysis,” *Journal of Politics*, 68, 62–74.

Li, Q. (2009), “Democracy, Autocracy, and Expropriation of Foreign Direct Investment,” *Comparative Political Studies*, 42, 1098–1127.

Li, Q. and Resnick, A. (2003), “Reversal of fortunes: Democratic institutions and foreign direct investment inflows to developing countries,” *International organization*.

Liptak, A. (2007), “A Sign of Court’s Polarization - Choice of Clerks,” .

Liu, X., Burridge, P., and Sinclair, P. J. N. (2002), “Relationships between economic growth, foreign direct investment and trade: evidence from China,” *Applied Economics*, 34, 1433–1440.

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

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

Logan, J. A., Hoff, P. D., and Newton, M. a. (2008), “Two-Sided Estimation of Mate Preferences for Similarities in Age, Education, and Religion,” *Journal of the American Statistical Association*, 103, 559–569.

Malesky, E. J. (2015), “Transfer pricing and global poverty,” *International Studies Review*, 17, 669–677.

Mallampally, P. and Sauvant, K. P. (1999), “Foreign Direct Investment in Developing Countries,” *Finance and Development*, 36.

McFadden, D. (1974), “Conditional logit analysis of qualitative choice behavior,” in *Frontiers in Econometrics*, ed. P. Zarembka, pp. 105–142, Academic Press, New York.

Milner, H. V. and Kubota, K. (2005), *Why the Move to Free Trade? Democracy and Trade Policy in the Developing Countries*, vol. 59.

Mold, A. (2004), "Fdi and Poverty Reduction: a Critical Reappraisal of the Arguments," *Weltbank, New York.*

Moran, T. H. (1998), *Foreign Direct Investment and Development: the New Policy Agenda for Developing Countries and Economies in Transition.*

Mosley, L. (2005), "Globalisation and the state: Still room to move?" *New Political Economy*, 10, 355–362.

Mosley, L. and Singer, D. (2008), "Taking Stock Seriously: EquityMarket Performance, Government Policy, and Financial Globalization," *International Studies Quarterly*, 3, 405–425.

Mosley, L. and Uno, S. (2007), "Racing to the Bottom or Climbing to the Top ?" *Comparative Political Studies*, 40, 923–948.

Nair-Reichert, U. and Weinhold, D. (2001), "Causality tests for cross-country panels: a new look at FDI and economic growth in developing countries," 2, 153–171.

NORC (2014), "Appendix G: Prestige Scores and Socioeconomic Index Distributions," .

Nunnenkamp, P. and Spatz, J. (2004), "FDI and economic growth in developing economies: how relevant are host-economy and industry characteristics," *Transnational Corporations*, 13.

Nunnenkamp, P., Schweickert, R., and Wiebelt, M. (2007), "Distributional effects of FDI: How the interaction of FDI and economic policy affects poor households in Bolivia," *Development Policy Review*, 25, 429–450.

Pak, Y. S. and Park, Y. R. (2005), "Characteristics of Japanese FDI in the East and the West: Understanding the strategic motives of Japanese investment," *Journal of World Business*, 40, 254–266.

Pandya, S. (2010), "Labor markets and the demand for foreign direct investment," *International Organization.*

Pandya, S. S. (2014), "Democratization and Foreign Direct Investment Liberalization, 1970-2000," *International Studies Quarterly*, 58, 475–488.

Pandya, S. S. (2016), *Trading Spaces Foreign Direct Investment Regulation, 19702000*, Cambridge University Press, New York.

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.

Pinto, P. (2013), "Partisan Investment in the Global Economy," .

Pinto, P. M. and Pinto, S. M. (2008), "The politics of investment partisanship: And the sectoral allocation of foreign direct investment," *Economics and Politics*, 20, 216–254.

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.

Prakash, A. (2007), "Investing Up: FDI and the Cross-Country Difusion of ISO 14001 Management Systems," *International Studies*, 51, 723–744.

Ricupero, R. (2000), "Tax incentives for foreign direct investment: A Global Survey," Tech. Rep. 16, UNCTAD, Geneva.

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*.

Rydén, T. (2008), "EM versus Markov chain Monte Carlo for Estimation of Hidden Markov Models: A Computational Perspective," *Bayesian Analysis*, 3, 659–688.

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.

Solow, R. M. (1956), "A Contribution to the Theory of Economic Growth," *The Quaterly Journal of Economics*, 70, 65–94.

Sprecher, S., Sullivan, Q., and Hatfield, E. (1994), "Mate selection preferences: gender differences examined in a national sample." *Journal of personality and social psychology*, 66, 1074–1080.

Tanner, M. A. and Wong, W. H. (1987), "The calculation of posterior distributions by data augmentation," *Journal of the American Statistical Association*, 82, 528–540.

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," *CoRR*, abs/1401.8.

UNCTAD (1998), "The Financial Crisis in Asia and Foreign Direct Investment : an Assessment," Tech. rep.

UNCTAD (2001), "World Investment Report 2011 - Promoting Linkages," Tech. rep., UNCTAD, New York and Geneva.

UNCTAD (2007), "World Investment Report 2007: Transnational Corporations, Extractive Industries and Development," Tech. rep., UNCTAD, New York and Geneva.

UNCTAD (2015), "World Investment Report 2015: Reforming International Investment Governance," Tech. rep., UNCTAD, Geneva.

Vernon, R. (1971), *Sovereignty at Bay: The Multinational Spread of U.S. Enterprises*, Basic Books, New York.

Willem, D. (2004), "Foreign Direct Investment and Income Inequality in Latin America by and Income Inequality in Latin America," *ODI Research Papers*.

World Economic Forum (2013), "Foreign Direct Investment as key driver for Trade, Growth and Prosperity: The case for a multilateral agreement on investment," *World Economic Forum*, p. 36.

Wu, F. (1999), "Intrametropolitan FDI firm location in Guangzhou, China: A Poisson and negative binomial analysis," *The Annals of Regional Science*, 33, 535–555.

Yamawaki, H. (1991), "Exports, and Foreign Distributional Activities: Evidence on Japanese Firms in the United States," *The Review of Economics and Statistics*, 73, 294–300.

Zebregs, M. and Tseng, M. (2002), *Foreign direct investment in China: Some lessons for other countries*.