The Pennsylvania State University
The Graduate School

TWO ESSAYS ON OFFSHORING AND OUTSOURCING

A Dissertation in
Economics
by
Wenzhuo Lu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2019

The dissertation of Wenzhuo Lu was reviewed and approved* by the following:

James Tybout
Professor of Economics
Dissertation Advisor, Chair of Committee

Jonathan Eaton
Distinguished Professor of Economics

Mark Roberts
Liberal Arts Professor of Economics

Zhe Wang
Assistant Professor of Finance

Barry Ickes
Professor of Economics
Department Head, Department of Economics

*Signatures are on file in the Graduate School.

Abstract

Chapter 1: Tax, Offshoring, and Income Distributions

I study the distributional consequences of income tax policy in a world with two countries and two heterogeneous factors of production. Productivity in each production unit reflects the ability of the manager and the abilities of the workers, with complementarity between the two. I show how changes of relative wages, generated by one country's choice of income tax policy, affect the reallocation of managers across the border, the matching of workers and managers, the distributions of wages and firms' profits, and national welfare in two countries. I find that when the identity of two production factors is endogenous, the optimal income tax rate in the South is slightly lower than the globally efficient individual tax rate.

Chapter 2: Globalization, Innovation, and Firm Dynamics

Outsourcing has led to significant cost savings, but, less obviously, it has changed importing firms' incentives to invest in innovation. In particular, with better access to foreign inputs, importing firms use cheaper imported inputs as a substitute for self-made inputs, and thus have less incentive to develop their own in-house varieties. In this paper, I first show that Chinese firms associated with lower *input* tariffs undertake less innovation. I then develop, estimate, and simulate a forward-looking model that captures this margin of adjustment and show how it responds to the changing costs of outsourcing. Counterfactual exercises show that a permanent reduction in trade-variable costs decreases domestic innovation participation rate, and this negative impact will accumulate and grow over time even as far into the future as twenty years later.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am deeply indebted to my advisor, James Tybout, whose continuous guidance and support has helped me immensely in the preparation and execution of this work. I would like to thank Mark Roberts, Jonathan Eaton, Russell Cooper, and Stephen Yeaple for insightful conversation and comments. I am grateful to seminar participants at the Trade/Development Reading Group, Industrial Organization Brownbag, and Women's Workshop at Penn State for useful feedback. All errors are of my own.

Dedication

To my parents.

Chapter 1

Tax, Offshoring, and Income Distributions

1.1 Introduction

The purpose of the current paper is to understand how changes in the rates of taxes on production factor in one country might affect the organization of work, the size distribution of firms, the structure of individual earnings, and national welfare in a two-country model.

I characterize the effect of taxation in a framework developed by Antràs et al. (2006) in which production requires physical inputs and knowledge; lower-skilled agents specialize in production work while more highly-skilled agents specialize in knowledge. Besides, I allow the productivity of workers to increase in their managers' skills, and firm size of managers to increase in their workers' ability. Complementarity between manager and worker ability determines the agents' occupational choices, and there exists positive sorting among agents in the equilibrium.

Moreover, I consider a simple one-sector, two-country model in which countries differ in their skill distributions. In the model, the "skill overlapping" parameter captures the skill differences between the two countries; call the country with better-skilled agents "North" and the other country "South." The other key parameter is the cost of communicating technology, as this determines the extent to which managers can leverage their knowledge within their teams.

Regarding taxes, specifically, I differentiate between and study the effects of three types of tax rates: the tax on managers' profits, the tax on the Southern workers' wages, and the tax on the Northern workers' wages. Initially, I discuss a case in which the tax regimes are common across countries, a situation I refer to as representing a "symmetric tax equilibrium". Then, using the equilibrium outcome in a symmetric tax equilibrium as a benchmark, I discuss how changes in the rates of taxation on Southern workers affect the organization of work, the size distribution of firms, and the structure of individuals' earnings in the two countries. In particular, I demonstrate numerically that when the tax on workers in two countries is no longer symmetrical, the lowest- and highest-ability managers match with better workers, compared to the match outcomes in a symmetric tax equilibrium, and that these managers' firms expand because of the improved matching potential. At the same time, some middle-ability managers are left having to work with less-skilled workers in the new equilibrium, and thus their firms shrink. Moreover, I show that taxation changes in one country affect individuals' earnings in each country. I find that a rise in the tax on southern workers increases the returns-to-skill for lower-skilled workers, depresses the returns-to-skill for higher-skilled workers down, and vice-versa.

Furthermore, I show the effect of tax on national consumption. In a symmetric tax equilibrium, the tax on workers' wages goes up as the gross consumption of the South increases. As the tax on workers increases, workers' pre-tax wages rise, and managers' profits go down. Since the South is a net exporter of workers' output, its welfare increases in workers' wages. However, in an asymmetric tax equilibrium, the gross consumption of the South is strictly concave in the tax on southern workers. Note that the identity of two production factors, i.e., workers and managers, is endogenous. High taxes on southern workers induce substitution by northern managers toward workers. Hence, as the tax on southern workers increases, southern workers' earnings increase as well, but the tax revenue collected from offshoring foreign firms decreases. As a result, we find that the national welfare of the South reaches its maximum when the tax on its workers is slightly lower than the globally efficient individual tax rate.

The current paper's hypotheses and design are grounded in three critical strands of literature. First, the paper concerns itself with the optimal tariff policy. Venables (1987) and Ossa (2001) discuss optimal import tariffs when a homogeneous-technology, differentiated-products sector is complemented by a numeraire sector with costless transportation of goods, perfect competition, and linear technology. In such a framework, wage rates are fixed by technology, and import tariffs allow the country to attract new firms into the sector that is afflicted by trade costs. Consumers benefit from import tariffs when the tariff-induced dislocation effect dominates over the direct impact of a tariff on the price index. In addition to this framework, Demidova et al. (2009) and Felbermayr et al. (2013) explore the implications of firm heterogeneity for optimal trade policy in an environment with a CES utility function; constant fixed costs for exporting; Pareto-distributed, firm-level productivity; and uniform trade tariffs across firms. Costinot et al. (2014, 2015, 2016) relax these assumptions, deriving optimal trade policy according to a more generalized version of the conclusions of Melitz (2003).

The second strand of previous research from which this paper draws concerns itself with tax competition, namely, how one country's choice of tax policy imposes fiscal externalities in another country. Most literature on transnational income tax competition has placed its policy focus on how double-tax rules affect capital movement and national welfare [see Bond et al. (1989) and Gresik (2001)]. More recent papers, meanwhile, consider the linkages between tax competition and other standard policies, including transfer pricing regulations, interest allocation rules, debt/equity financing, and so on [see Davies (2000)].

The third of these strands of previous research that the current paper examines pertains to the implications of sorting and matching for income distribution. Grossman et al. (2015) develop a model where income distribution is affected by the sorting of heterogeneous managers and workers into industries, as well as by the matching of managers and workers in production units within each one. In particular, it is observed that trade introduces changes in relative prices, through which managers and workers are re-matched within each industry and thus within- occupation-inequality is affected.

The remainder of the paper is structured as follows: Section 2 describes the primary environment. Section 3 characterizes the equilibria in different circumstances. Section 4 discusses the effects of the income tax. Section 5 concludes.

1.2 The Model

1.2.1 Basic Environment

Endowment. Consider a world economy with two countries, indexed by $i = S, N$; one factor of production, labour; and homogeneous outputs. In particular, a country i is endowed with a continuum of agents with ability z derived from a cumulative distribution function $G_i(z)$ with density $g_i(z)$ and support in $[0, \bar{z}_i]$. Agents make an occupation choice between managers and workers: workers are immobile, but managers are mobile across countries.

Preference and Market Structure. Assume that utility is linear with respect to the consumption of homogeneous outputs and that trade cost is zero, so that agents are income-maximizers. Both the goods and the labor markets are perfectly competitive in these equilibria.

Technology. The model follows Garicano (2000) that each production team consists of one manager and a few production workers. Each agent has one unit of time and a skill level z. Workers have to solve a problem in order to carry out production. If a worker knows the solution to the problem, she solves it and produces one unit of output. If she does not know the solution, she can ask her manager. If the manager knows the solution, she then produces one unit of output; otherwise, no production happens at all. It takes a manager $0 < h < 1$ units of time to communicate about this with the worker, whether the manager knows the solution or not. Thus, the skill level of an agent determines the set of problems she can solve. An agent with skill z can solve all problems that require knowledge between 0 and z. Here, we normalize the set of problems so that the skill level z also represents the proportion of problems that an agent can solve. Hence, a manager in a team with n workers of skill z_p faces the following time constraint:

$$h(1 - z_p)n(z_p) = 1,$$

and so can deal with $n(z_p) = 1/[h(1 - z_p)]$ workers.

Lemma 1. *Each manager hires workers of homogeneous ability.*

$r_i^j(z)$ denotes the rent a manager from country j with ability z receives using labor in country i. That is, we use superscripts for managers' nationality while using subscripts for workers' origins. Then given wages $w_i(z_m)$ and production location i, a manager from country j with ability z_p chooses the ability level of her/his workers to maximize local rents:

$$r_i^j(z_m) = \max_{z_p} \left(z_m - w_i(z_p)\right)n(z_p),$$

such that

$$h(1 - z_p)n(z_p) = 1.$$

1.2.2 Existence of the Decentralized Equilibrium with Taxes

In the current system of international taxation, all countries with tax income earned by multinational corporations within their borders − indeed most countries, including all countries in the G7 (Canada, France, Germany, Italy, Japan, and the United Kingdom) − use a territorial system that exempts from taxation most of the foreign-source income earned by their multinationals. [1] Therefore, the current paper considers a source-based taxation system, wherein managers pay corporate income taxes exclusively to the countries where they conduct production, regardless of the formers' nationality.

I allow ad-valorem income taxes to vary across markets and across occupations. Formally, let $t_i^o \in \mathbb{R}$ denote the tax rate charged by country $i = \{S, N\}$ on agents with occupation $o = \{M, W\}$. Here, $t_i^o > 0$ corresponds to an income tax, while $t_i^0 < 0$ refers to an income subsidy. Tax revenues are rebated to domestic consumers through a lump-sum transfer, T_i.

I follows Antràs et al. (2006) to let the equilibrium occupational-choice decision be such that agents in country i with skill levels in $[0, z_i^*]$ become workers, and agents in $[z_i^*, \bar{z}_i]$ become managers. Agents with knowledge z_i^* are indifferent. This restriction on equilibrium occupations turns out to be without loss of generality, as Lemma 3 below shows.

Let $m_i^j(z_p)$ be the skill level of the manager coming from country j who is matched with a worker with ability z_p in country i, and let $w_i^j(z_p)$ represent the pre-tax wages for production workers with ability z_p working for a manager from j in country i.

Lemma 2. *Suppose m_i^j is real-valued on $[a, b]$, then for any $z_p \in [a, b]$,*

$$w_i^S(z_p) = w_i^N(z_p) \text{ for } i \in \{S, N\},$$
$$m_i^S(z_p) = m_i^N(z_p) \text{ for } i \in \{S, N\}.$$

Lemma 2 implies that match quality and wage are independent of a worker's nationality. Hence, for simplicity, the superscripts of wage and matching functions are omitted from now on.

let $I_i^j(z_m)$ be the percentage of managers with ability z_m from country i who choose to conduct production in location j in the equilibrium; thus, the managers' market clearing condition implies that

$$I_i^j(z_m) + I_j^j(z_m) = 1,$$

for all $z_m \in [\min\{z_i^*, z_j^*\}, \max\{\bar{z}_i, \bar{z}_j\}]$. Therefore, I characterize a decentralized equilibrium with taxes as schedules of wages $\boldsymbol{w_i} \equiv w_i(z)$, schedules of rents $\boldsymbol{r_i^j} \equiv r_i^j(z)$, sched-

[1] However, the United States imposes a minimum tax on the income US-based multinationals earn in low-tax foreign countries, with a credit for 80 percent of foreign income taxes they've paid, so as to limit their resident corporations' ability to shift profits to low-income countries by taxing foreign "passing" income on an accrual basis.

ules of matches $\boldsymbol{m_i} \equiv m_i(z)$, schedules of location assignments $\boldsymbol{I_i^j}(z)$ and occupation cutoffs z_i^* such that

(i) Given wages and rents, an agent from country i with ability z maximizes her after-tax income $E_i(z)$:

$$E_i(z) = \begin{cases} (1 - t_i^W)w_i(z) & \text{if } z \leq z_i^*, \\ \max\{(1 - t_i^M)r_i^i(z), (1 - t_j^M)r_j^i(z)\} & \text{if } z > z_i^*. \end{cases}$$

(ii) Given wages, a manager from country i with ability z_m chooses the ability of workers to maximize the enterprise income in country j subject to the time constraint:

$$m_j^{-1}(z_m) \in \arg\max r_j^i(z_m) = \arg\max \max_{z_p}(z_m - w_j(z_p))n(z_p),$$

subject to

$$h(1 - z_p)n(z_p) = 1.$$

(iii) Workers' market clears in country i

$$\int_0^{z_p} g_i(z)dz$$
$$= \int_{m_i(0)}^{m_i(z_p)} I_i^i(z)n(m_i^{-1}(z))g_i(z)dz + \int_{m_i(0)}^{m_i(z_p)} I_i^j(z)n(m_i^{-1}(z))g_j(z)dz,$$

subject to

$$I_i^j(z_m) + I_j^j(z_m) = 1,$$

for all $z_p \leq \max\{z_i^*, z_j^*\}$.

(iv) government's budget is balanced in both countries:

$$T_i = t_i^W \int_0^{z_i^*} w_i(z)g_i(z)dz + t_i^M \left[\int_{z_i^*}^{\bar{z}_i} I_i^i(z)r_i^i(z)g_i(z)dz + \int_{z_j^*}^{\bar{z}_j} I_i^j(z)r_i^j(z)g_j(z)dz \right],$$

where T_i is a lump-sum transfer.

Lemma 3. *There exists a unique competitive equilibrium of this economy for some (α, h). In the equilibrium, the set of managers and the set of workers are connected, the equilibrium exhibits positive sorting, and the earnings function is strictly convex.*

1.3 Decentralized Equilibrium

Consider a world formed by two independent countries, wherein agents can form teams with agents in their own countries or with agents in the other country. Assume that the first one, which we call "the North" (or "N"), has a uniform distribution of skills in the population, $G_N(z) = z$ for $z \in [0,1]$ with density $g_N(z) = 1$; while the second economy, which we call "the South" (or "S"), has a uniform distribution of skills $G_S(z) = z/\alpha$ for $z \in [0, \alpha]$ with density $g_S(z) = 1/\alpha$ and we require $\alpha < 1$. The North, therefore, is relatively better endowed with skilled agents, but both countries are identical in all other respects, including population size. We use the following two lemmas to further simplify the notations.

Lemma 4. *The images of assignment functions m_S and m_N are connected intervals.*

Denote M_S and M_N as the set of managers who establish firms in South and North. So, $z_m \in M_S$ means an agent with ability z_m is willing to establish a firm in South to maximize her earnings in the equilibrium. Also, $z_m \in M_S \cap M_N$ indicates an agent with ability z_m is indifferent between South and North to form her production team, although she might match with workers possessing different skill levels in two countries. Technically, we could write $M_S = [m_S(0), m_S(z_S^*)]$ and $M_N = [m_N(0), m_N(z_N^*)]$.

Lemma 5. *If $M_S \cap M_N = [a, b]$ where $a < b$, then*

1. $a = \min\{z_S^*, z_N^*\}$ *and* $b = \min\{m_S(z_S^*), m_N(z_N^*)\}$.
2. *For any* $c \in [a, b]$,
$$m_S^{-1}(c) = m_N^{-1}(c),$$
$$w_S(m_S^{-1}(c)) = w_N(m_N^{-1}(c)).$$
3. $t_N^M = t_S^M$.

Part (2) of Lemma 5 suggests that when the skill sets of managers in South and North are overlapping, then workers of the similar ability are matched with the same types of managers and receive equal pre-tax wages, regardless of the workers' nationalities. Hence, I further omit the subscripts of pre-tax wage and matching functions and instead use $w(.)$ and $m(.)$ from now on. Part (3) of Lemma 5 implies that $t_N^M = t_S^M$ (that is, the event that corporate income taxes are equal across borders) is a necessary condition for the existence of overlapping skill sets of managers from two countries.

1.3.1 Symmetric (Tax) Equilibrium

We start with a symmetric equilibrium as a benchmark, wherein two governments impose tax rates $t_S^W = t_N^W$ on workers' wages and $t_S^M = t_N^M$ on managers' rents. Depending on the value of the communication cost h, skill measure α, and tax rate t_i^o, we can show that there exist two types of equilibria in the economy. The first one is an equilibrium in which all agents in the South are workers. Since there are no managers in the South in this scenario, all of the South's agents work for northern managers. Positive sorting implies that, because they are the lowest-quality workers in the world, they work for the worst managers in the North. Hence, firms in the South consist of the worst workers and managers in the integrated economy. We call this the *Low-Quality Offshoring Equilibrium (LQE)*.

On the contrary, the second type of equilibrium is one in which some of the agents in the South are managers. This equilibrium features the South's less-skilled working for southern managers, and its more able ones working for better managers in the North. We call this the *High-Quality Offshoring Equilibrium (HQE)*.

Analogous to the approach in Antràs et al. (2006), I derive all results under the assumption that international teams are formed only if managers strictly prefer to hire foreign workers than domestic ones, to guarantee the least amount of offshoring. In general, as long as tax regimes are common across countries, the set of parameter values that determines the boundary between these two equilibria is a nonlinear function of h and α, which is plotted in Figure 1.1 and discussed in turn.

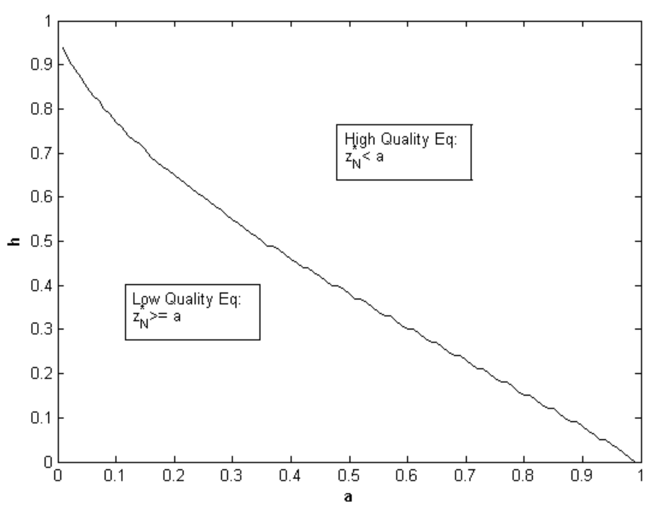

Figure 1.1: Types of Equilibria in a Symmetric Equilibrium

As shown in Figure 1, we find that an economy with a higher α or a higher h is an economy with a *High-Quality Offshoring Equilibrium (HQE)*, where managers are coming from both countries. Intuitively, as the communication cost h increases, managers can have smaller teams, so in equilibrium, there are more managers and fewer workers. Thus, some skilled agents in the South are more likely to become managers to meet the demand. Meanwhile, as α increases, it is profitable for some high-ability agents to become managers in the South. In all, the possibility of the existence of a high-quality equilibrium increases as α or h goes up.

Next, denote by z_i^* the threshold that separates the ability of the agents who choose to be workers or managers in country i in the equilibrium.

Lemma 6. *In a symmetric equilibrium where tax rates on managers and workers are symmetric in two countries, occupation thresholds are identical such as*

$$z_S^* = z_N^* = \begin{cases} \frac{1+h-\sqrt{1+h^2(3-\alpha)}}{h} & \text{if } h \leq \frac{2(1-\alpha)}{2+\alpha-\alpha^2} \\ \frac{1+h-\sqrt{1+h^2+((1-\alpha)/(1+\alpha))2h}}{h} & \text{if } h > \frac{2(1-\alpha)}{2+\alpha-\alpha^2} \end{cases}$$

Lemma 6 suggests that if $h \leq \frac{2(1-\alpha)}{2+\alpha-\alpha^2}$, then there exists a low-quality equilibrium in the model; otherwise, the economy exhibits a high-quality equilibrium. Moreover, the boundary between the two equilibria, which is determined by occupation thresholds z_i^*, is independent of tax rates. The assignment function $m(.)$ (and thus thresholds z_i^*) is derived from labor market clearing conditions and from two boundary conditions: $m(0) = z_i^*$ and $m(z_i^*) = 1$. Hence, the span of managers' control is a technological restriction of the model, and I can compute the assignment function without knowing wages, rents, or tax rates.

1.3.2 Asymmetric (Tax) Equilibrium

Now I consider an asymmetric tax equilibrium in which, although taxes on managers' rents remain the same, taxes on workers' wages are different such as $t_S^W \neq t_N^W$.

Lemma 7. *If $M_S \cap M_N = [a, b]$ where $a < b$ and $t_S^W \neq t_N^W$, then*

1. $z_S^* < z_N^* \leq \alpha < 1$ *if* $t_S^W > t_N^W$.
2. $z_N^* < z_S^* \leq \alpha < 1$ *if* $t_S^W < t_N^W$.

Lemma 7 indicates that occupation cutoffs are different in an asymmetric tax equilibrium. It follows part (2) of Lemma 5 that pre-tax wage functions are identical across the border and with respect to the fact that managers are mobile. Thus, once the tax on southern wages increases, there are more managers and fewer workers in the South. As a result, $z_S^* < z_N^*$ whenever $t_S^W > t_N^W$. The same logic applies to the case when $t_S^W < t_N^W$.

Similar to the symmetric equilibrium, depending on the value of (h, α, t), there also exist two types of equilibria in the economy: the Low-Quality Equilibrium in which all agents in the South are workers; and the High-Quality Equilibrium in which some of the agents in the South are managers.

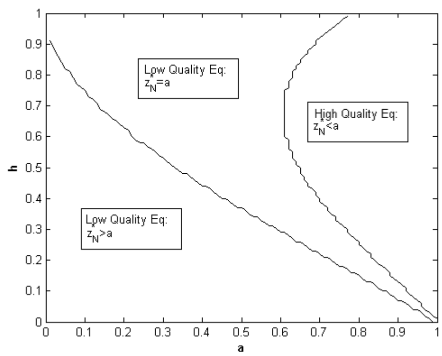

Figure 1.2: Types of Equilibria with Higher Tax on Southern Workers

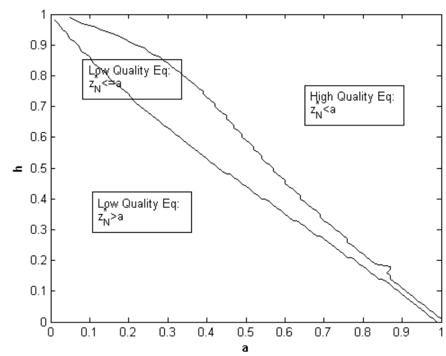

Figure 1.3: Types of Equilibria with Lower Tax on Southern Workers

Figure 1.2 presents types of equilibria with a higher tax on southern workers than on northern workers. In symmetric equilibria, the case where $z_S^* = \alpha$ is plotted as a line between two types of equilibria; see Figure 1.1. However, as the tax on southern workers increases, the region where $z_N^* = \alpha$ expands as well. When the tax on southern workers increases, there are fewer workers and more managers in the South. To clear the labour market, the occupation cutoff z_N^* goes up to increase the supply of workers in the global market. Meanwhile, z_N^* is smaller or equal to α; otherwise, the image of the assignment function, that is $M = [z_S^*, \alpha] \cup [z_N^*, 1]$, is not connected, which violates Lemma 4. As a result, the region where $z_N^* = \alpha$ holds increases in the difference between t_N^W and t_S^W.

In opposition to Figure 1.2, here Figure 1.3 characterizes types of equilibria with a lower tax on southern workers than on northern workers. In a symmetric equilibrium, $z_N^* < \alpha$ is a sufficient condition for the existence of a high-quality equilibrium. However, when the tax on wage income is lower than the tax on corporate income, southern agents with ability $z \in [z_N^*, \alpha]$ prefer to become workers rather than managers for some given parameters. As a result, even though $z_N^* <= \alpha$, the global economy exhibits a low-quality equilibrium when $(\alpha, h, t_S^W, t_N^W, t^M)$ satisfy

$$(1 - t_S^W)w(z) \geq (1 - t^M)r(z) \geq (1 - t_N^W)w(z) \text{ for any } z \in [z_N^*, \alpha].$$

Moreover, this region where the unique equilibrium exists also increases in the difference between t_N^W and t_S^W.

1.4 Effects of Labor Tax

According to the 2015 Organization for Economic Cooperation and Development (OECD) Tax Database, the global average corporate tax rate was 39%, and single workers without children paid an average tax rate of 35.9%. The US worker, in particular, paid an average tax rate of 31.7%, while the average tax rate was as high as 49.4% in Germany. The current analysis uses these four numbers to study the impact of individual taxation policies on the composition of teams, the occupational choices, the size distribution of firms, and the rewards structure in a two-country economy.

1.4.1 Matching and Occupation Choices

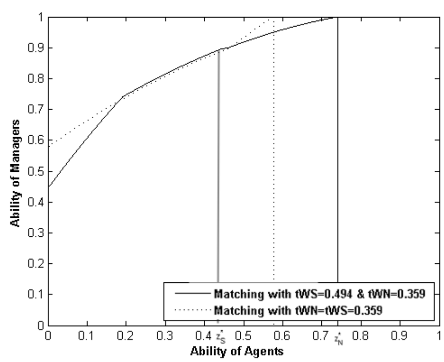

Figure 1.4: Matching with Higher Worker Tax in the South vs. Symmetric Worker Tax rates in two Countries

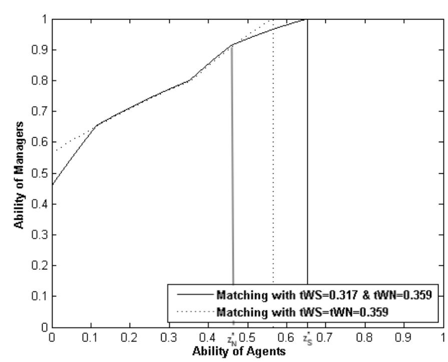

Figure 1.5: Matching with Lower Worker Tax in the South vs. Symmetric Worker Tax rates in two Countries

First, let the North's government impose a global average corporate tax rate of 39% on managers' rents and an individual tax rate of 35.9% on workers' wages. Figure 1.4 presents the matching function when the South's government collects higher worker tax rates (indicated by the solid line), in comparison with the scenario in which the South's government also follows the global average individual tax rate (indicated by the dotted line).

Analogously, Figure 1.5 presents the matching function when the South's government collects a lower worker tax than the North. Note that the tax rates do not affect the assignment function in a symmetric equilibrium, as indicated by Lemma 6.

Figures 1.4 and 1.5 suggest that changes in relative worker tax rates in two countries shift two occupation thresholds z_N^*, z_S^* toward different directions. When the tax on southern workers increases, there are fewer workers and more managers in the South. Hence, z_S^* shifts to the left. Meanwhile, the North supplies more workers to clear the labour market, and thus z_N^* shifts to the right. It follows the same argument that when the tax on southern workers decreases, there are more workers in the South and more managers in the North. In each scenario, the difference between two occupation thresholds brings less-productive managers (who used to be workers in one country) and higher-ability workers (who used to be managers in the other country) into the new integrated economy. As a result, some of the marginal workers become managers of low-skilled agents, and matches of agents with sufficiently low skill necessarily become worse. On the other hand, some previous managers turn out to be high-ability workers and are matched with the best managers, so highly skilled workers now have more competition and match with worse managers. However, some middle-skilled agents now are better off since their competitors become managers and have less competition.

1.4.2 Size Distribution of Firms

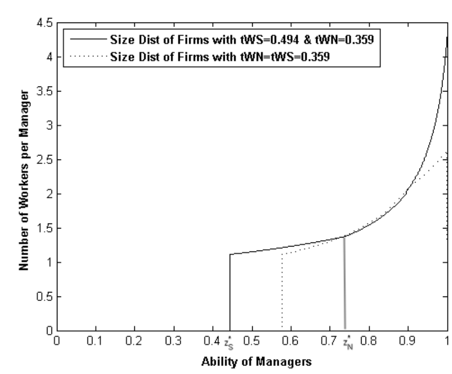

Figure 1.6: Size Distribution of Firms with Higher Worker Tax in the South vs. Symmetric Worker Tax rates in two Countries

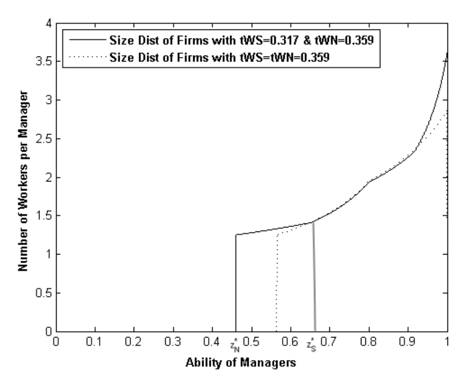

Figure 1.7: Size Distribution of Firms with Lower Worker Tax in the South vs. Symmetric Worker Tax rates in two Countries

Recall that the time constraint of the model

$$n(z_p) = 1/[h(1 - z_p)]$$

suggests that the number of workers per manager $n(z_p)$ decreases in the communication cost h and increases in workers' ability z_p. Hence, given h, firm size is increasing in the matching quality. Figure 1.4 and 1.5 suggest that as the lowest- and highest- ability managers are matched with better workers in an asymmetric tax equilibrium than in a symmetric equilibrium, their firms expand in the new equilibrium. Meanwhile, the medium-ability managers match with worse workers in the new asymmetric equilibrium, so their firms shrink. The changes in firm size distributions are presented in Figures 1.6 and 1.7.

1.4.3 Wage Structure of the Economy

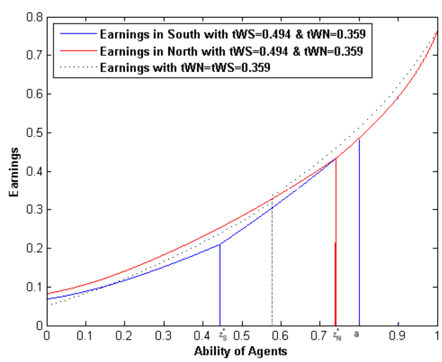

Figure 1.8: Earnings with Higher Worker Tax in the South vs. Symmetric Worker Tax rates in two Countries

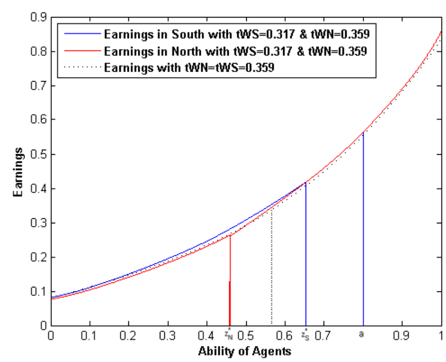

Figure 1.9: Earnings with Lower Worker Tax in the South vs. Symmetric Worker Tax rates in two Countries

There are two determinants of the marginal returns to skill. One is the quality of the match, which I refer to as the *complementarity effect*. The other determining force is the supply and demand of workers, which I call the *competition effect*.

Figure 1.8 presents a situation when the southern government imposes a higher tax rate on domestic workers' wages. The equilibrium effect on the marginal returns to skill can be decomposed in two parts. First, when a higher tax on workers decreases the supply of southern workers, z_S^* goes down, thus decreasing the competition among workers in the South and causing the baseline returns per unit of skill to go up. Second, a decrease in z_S^* brings the lowest-ability managers (who used to be workers without the shock) to the market, and these lowest-ability managers are matched with the worst workers in the world. In this way, the new policy decreases the match quality of those less-skilled workers, and the returns on their skill go down.

Similarly, a rise in z_N^* increases the supply of workers in the North, with the result that the enhanced competition hurts all workers' wages. Meanwhile, the most productive firms are now matched with the best workers (who used to be managers in the symmetric equilibrium), and as a result, the returns to skill for higher-skilled agents go up.

As shown in Figure 1.8, the competition effect dominates the complementarity effect in the case of $t_S^W > t_N^W$. Hence, we observe the returns to skill for lower-skilled workers go up, and the returns to skill for higher-skilled workers go down. However, in the case of $t_S^W < t_N^W$, Figure 1.9 exhibits a different story, such that the complementarity effect dominates. We find that the returns to skill for lower-skilled workers go down, and the returns to skill for higher-skilled ones go up.

What ultimately determines the joint impact of the competition effect and the complementarity effect are the communication cost parameter h and the skill-overlapping parameter α between two countries. If communication is more efficient (i.e., h is low), all agents can be matched with the top managers available in the world and the quality of matches is more stable. Thus, the complementarity effect decreases in the communication cost parameter h. Furthermore, the assumption $1/\alpha > 1$ implies that the measure of southern workers per skill is higher than that of the northern agents. In this way, we see that occupation choices in the South have a more substantial impact on the total supply of workers in the broader world. In other words, the supply of workers responds more significantly to the changes in tax on southern workers in particular. Hence, the smaller the skill-overlapping parameter α, the stronger the competition effect.

1.4.4 Production in a Symmetric Equilibrium

We define a country's total production as the outputs produced within its borders. Hence, in a symmetric tax equilibrium, the total production in the South and North are

$$Y_S = \begin{cases} \int_0^{\alpha} m(z) g_S(z) dz & \text{if } h \leq \frac{2(1-\alpha)}{2+\alpha-\alpha^2}, \\ \int_0^{z^*} m(z) g_S(z) dz & \text{if } h > \frac{2(1-\alpha)}{2+\alpha-\alpha^2}, \end{cases}$$

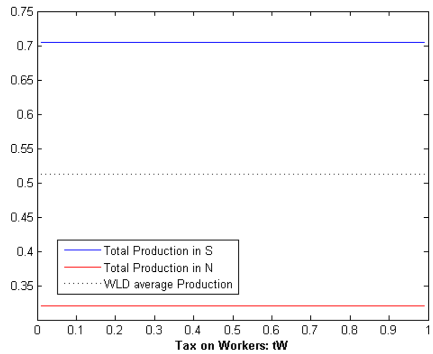

Figure 1.10: The Effect of Taxation on Production in an LQE

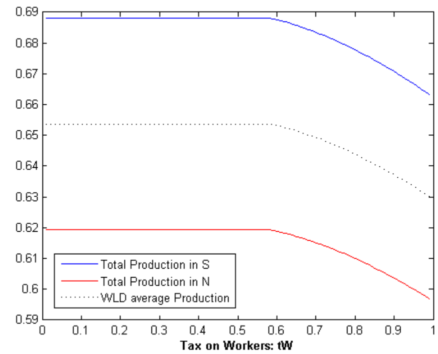

Figure 1.11: The Effect of Taxation on Production in an HQE

and

$$Y_N = \begin{cases} \int_0^{\alpha} m(z) g_N(z) dz & \text{if } h \leq \frac{2(1-\alpha)}{2+\alpha-\alpha^2}, \\ \int_0^{z^*} m(z) g_N(z) dz & \text{if } h > \frac{2(1-\alpha)}{2+\alpha-\alpha^2}, \end{cases}$$

where z^* comes from Lemma 6. Since the assignment function $m(.)$ and the occupation cutoff z^* are independent of tax rates in an LQE, it is straightforward to show that the ratio of Y_S and Y_N is also independent of tax rates such as

$$\frac{Y_S}{Y_N} = \frac{1}{\alpha}$$

in an LQE (see Figure 1.10) and in an HQE (see Figure 1.11). Furthermore, the total production is independent of tax rates in an LQE and in an HQE up to a threshold tax rate on workers. As the tax on labour rises, low-ability agents become self-employed managers in order to enjoy low corporate tax rates. However, these new managers produce less than they would as employees due to the limitations of their own knowledge. This threshold decreases in h and α. If h is low, then all agents can be matched with top managers and gains from matches are high. Also, if α is low, unskilled agents benefit more from better matches, and the improvement in match quality undermines the tax incentives.

1.4.5 Consumption in a Symmetric Equilibrium

Recall that we denote $I_i^j(z_m)$ as the percentage of managers with ability z_m from country i who offshores to country j where $I_i^j(z_m) + I_j^j(z_m) = 1$ for all $z_m \in [\min\{z_S^*, z_N^*\}, \max\{\bar{z}_S, \bar{z}_N\}]$.

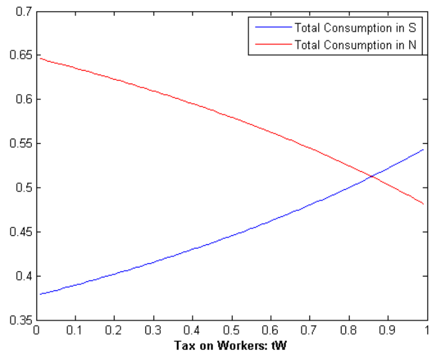

Figure 1.12: The Effect of Taxation on Consumption in an LQE

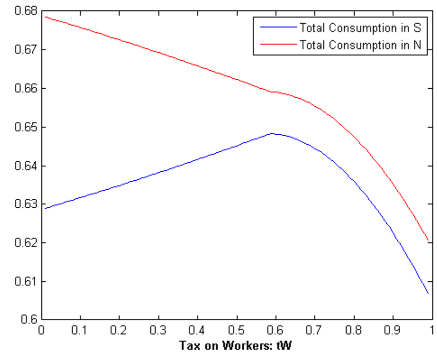

Figure 1.13: The Effect of Taxation on Consumption in an HQE

Given the ad-valorem tax rates $\{t^M, t^W\}$, the total consumption in country i is given by

$$C_i = \underbrace{\int_0^{z_i^*} w_i(z) g_i(z) dz}_{\text{domestic workers' wages}} + \underbrace{\int_{z_i^*}^{\bar{z}_i} I_i^i(z) r_i(z) g_i(z) dz}_{\text{domestic firms' rents}}$$
$$+ \underbrace{(1 - t^M) \int_{z_i^*}^{\bar{z}_i} I_j^i r_i(z) g_i(z) dz}_{\text{after-tax royalties from abroad}} + \underbrace{t^M \int_{z_j^*}^{\bar{z}_j} I_i^j(z) r_j(z) g_j(z) dz,}_{\text{tax collected from foreign firms' profits}}$$

and

$$\underbrace{C_S + C_N = Y_S + Y_N.}_{\text{goods market clearing condition}}$$

Hence, the gross consumption of a nation consists of its own domestic workers' wages, domestic firms' profits, after-tax royalties from abroad and corporate tax collected from foreign firms' profits. Figures 1.12 and 1.13 show that for a given corporate income tax, the gross consumption of the South is increasing in the tax on workers in an LQE and in an HQE up to a certain threshold.

Though the assignment function is independent of tax rates, the wage and rent functions are affected by tax rates. For an agent with ability z^* who is indifferent between two occupations, they receive the same after-tax earnings such that

$$(1 - t^W) w(z^*) = (1 - t^M) r(z^*),$$

or

$$\frac{w(z^*)}{r(z^*)} = \frac{(1 - t^M)}{(1 - t^W)}.$$

Since z^* is independent of tax rates and t^M is exogenously given, $\frac{w(z^*)}{r(z^*)}$ increases in t^W and so do the workers' pre-tax wages. The South is a net exporter of manufacturing goods, while the North is a net exporter of knowledge services. Hence, the gross consumption in the South rises when workers' job becomes more valuable, and vice versa. This result is consistent with a variant of the argument in the existing literature, which posits large countries should tax the goods in which they have a comparative advantage in order to secure more favorable terms of trade.

However, global production decreases once the tax on workers exceeds a certain threshold; as plotted in Figure 1.11. Hence, the gross consumption of the South increases along with tax on workers only up to a threshold, due to the goods-market clearing condition.

1.4.6 Production in an Asymmetric Equilibrium

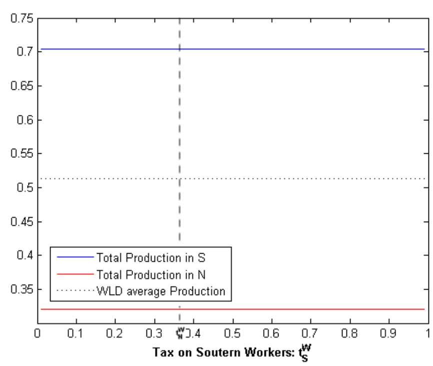

Figure 1.14: The Effect of Taxation on Production in an LQE

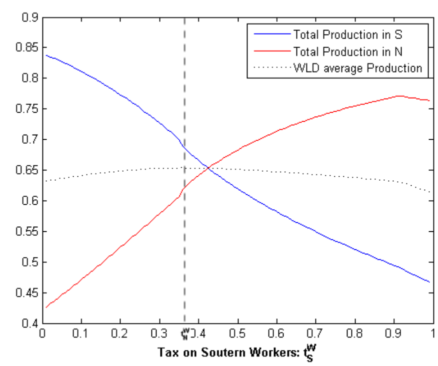

Figure 1.15: The Effect of Taxation on Production in an HQE

Now let us turn to an asymmetric tax equilibrium with $t_S^W \neq t_N^W$. In particular, I am interested in how the tax on southern workers affects production and consumption in two countries while all other variables remain fixed.

In an LQE where all the southern agents are workers, there exists only one occupation cutoff, z^*, which is pinned down by the labour-market clearing conditions. So z^* and the assignment function are independent of tax rates, and thereby production is constant, as shown in Figure 1.14.

However, Figure 1.15 tells a different story. Here I measure the average global production as $(Y_S + Y_N)/2$. It is not surprising that global output reaches its maximum when the tax on southern workers is identical to the tax on northern workers: $t_S^W = t_N^W = 0.36$. When tax rates on workers are symmetric, the global assignment function is independent of discretionary taxation, and thus the world economy is efficient. Note that at $t_S^W = t_N^W = 0.36$, the output in the South is larger than the output in the North, which is consistent with the argument of comparative advantage in the trade literature.

Besides, when two countries apply asymmetric tax rates on workers, two occupation cutoffs, z_S^* and z_N^*, shift toward opposite directions. As the tax on southern workers increases, z_S^* decreases and fewer southern agents become workers $-$ so, the total output in the South shrinks. On the contrary, z_N^* increases in the tax on southern workers. Hence, as the tax on southern workers becomes more substantial, there are more workers in the North and the gross output in the North also rises.

1.4.7 Consumption in an Asymmetric Equilibrium

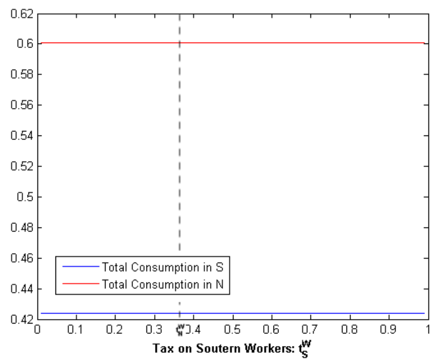

Figure 1.16: The Effect of Taxation on Consumption in an LQE

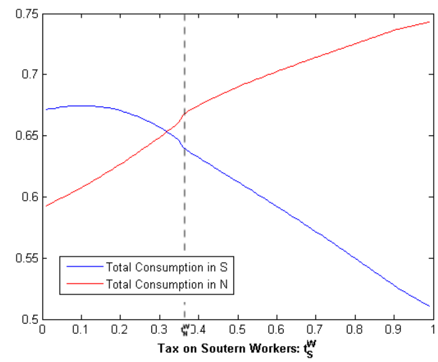

Figure 1.17: The Effect of Taxation on Consumption in an HQE

In an LQE, there are no managers in the South and the tax rates affect the earnings function through

$$(1 - t_N^W)w(z^*) = (1 - t^M)r(z^*)$$

In this case, the earnings function is independent of the tax on southern workers and national consumption curves are flat as shown in Figure 1.16.

Figure 1.17 shows that the total consumption in the South is concave in the tax on southern workers while keeping all the other variables constant. In particular, given given $\alpha = 0.9, h = 0.5, t_N^W = 36\%$, and $t^M = 39\%$, the gross consumption of the South is maximized at $t_S^W = 10\%$. Though the world economy is efficient if and only if $t_S^W = 36\%$, the national welfare of the South is maximized if the tax on southern workers is slightly lower than the globally efficient tax rate. This concavity comes from two forces. First, the South is a net exporter of workers. It gains more favourable terms of trade once workers' jobs become more valuable. As such, the consumption of the South increases in tandem with the tax on workers. Second, national consumption also depends on the tax collected from foreign firms' profits. A higher tax on southern workers induces substitution by northern managers towards workers. As fewer high-skilled agents from the North offshore their knowledge services to the South, the national consumption of the South declines. In all, the gross consumption of the South is strictly concave in t_S^W and reaches its maximum at $t_S^W = 10\%$.

1.5 Conclusion

Based upon the framework of Antràs, et al. (2006), this paper analyzes the effects of individual taxation on the organization of work, the size distribution of firms, and the structure of individuals' earnings within a source-based taxation system, as well as how these outcomes affect national production and consumption in turn.

An exciting feature of the model presented by the current paper is that the identity of two production factors $-$ that is, workers and managers $-$ is endogenous. Because of the flexibility of endowment in the environment, changes in tax rates on one factor relative to another factor ultimately affect the demand and supply of both factors. First, in a symmetric tax equilibrium, the gross consumption of the South increases in the tax on workers. As the tax on workers increases, workers' pre-tax wages increase too, at the cost of managers' rents. Since the South is a net exporter of manufacturing goods, its national welfare is enhanced when its workers' jobs become more valuable. Next, in an asymmetric tax equilibrium, the gross consumption of the South is strictly concave in the tax on southern workers because, while southern workers' earnings increase, taxes collected from foreign firms decrease in the tax on southern workers. As a result, we find that the national welfare of the South is maximized if the tax on its workers is slightly lower than the globally efficient individual tax rate.

I conclude by pointing out one limitation of the present analysis that could be relaxed in future research. Namely, there is a lack of fixed offshoring/production/entry costs in this model. I assume the perfect mobility of managers across borders, which implies the identical tax rates on firms' profits in two countries. However, in practice, governments collect varying corporate income taxes depending on firms' ownership, industry, capital stock, etc. One possible framework for relaxing this restriction would be to construct a trade model with monopolistic competition and firm-level heterogeneity.

Chapter 2

Globalization, Innovation, and Firm Dynamics

2.1 Introduction

As markets have become globally integrated, firms have increasingly relied on foreign sources for their inputs. This has led to significant cost savings, but less obviously, it has changed outsourcing firms' incentives to invest in innovation. In particular, with better access to foreign inputs, these firms have less incentive to develop their in-house varieties. In this paper I first show that this linkage is potentially critical, suggested by the finding that those Chinese firms experiencing the largest *input* tariff cuts between 2001 and 2006 were also the most likely to reduce their R&D. I then develop, estimate, and simulate a forward-looking model that captures this margin of adjustment and shows how it responds to the changing costs of outsourcing.

A key feature of my model is that it accounts for the fact that firms optimize their choice of countries from which they source inputs, recognizing that they can produce any input themselves in the event that in-house production proves cheaper. To solve for their optimal innovative effort, thus, firms must first determine which fraction of their input bundle they will import, and where each input will come from. Given that the decision of importing from one market will *decreases* the return to import from another market, this is a complicated problem, previously deemed incapable of fitting a model. However, inspired by Arkolakis and Eckert (2017), I introduce a new algorithm and show how it can indeed be used to tackle this problem.

Using the simulated method of moments, I fit my model to Chinese plant-level census data, customs data, and patent filing data, all spanning the period 2001-2006. Then I conduct counterfactual exercises to quantify the effect of a permanent reduction in the trade costs associated with Japanese imports. I am interested in Japan because it is the largest imported input supplier from the perspective of Chinese importers. I find that eliminating all tariffs on imports from Japan decreases the R&D participation rate by 1.65 percent and increases firms' average value by 0.17 percent. In an extreme situation - in which fixed outsourcing costs in dealings with Japan are zero yet the tariff rate on Japanese imports remains at 7.44 percent - the R&D participation rate falls by 19 percent and firms' average value increases by 4.97 percent.

My estimates of outsourcing and innovation costs are also of interest. Concerning the former, I find that fixed sourcing costs differ substantially across potential trading partners. As a result, relatively unproductive firms are unaffordable to import from high fixed cost countries, even if these countries are particularly attractive sources of inputs. Concerning the costs of innovation, I find that the cost of initializing innovation is much more substantial than the per-period costs of maintain investment, indicating that firm's R&D decision today not only affects it future productivity but also determines whether the firm will undertake R&D or not tomorrow.

The paper complements to the literature on how trade liberalization affects firms' innovation incentives. Most such works focus on demand-related concerns, with topics such as increases in market size, global competition, and realizing scale economies. For example, Bloom et al. (2016) find that import competition from China leads to more innovations among European firms, while imports in Europe from other developed countries have no significant effect. Meanwhile, in a South American context, Bustos (2011) argues that the increase in foreign market revenues generated by the MERCOSUR can induce Argentine exporters to upgrade their technology. My paper, by contrast, brings further attention to adjustments made by firms on the supply side: namely, with the intention of minimizing costs, the removal of trade barriers makes firms outsource more intermediates as substitutes for self-made inputs, while investing less in in-house production technology. This strategy translates into an improvement in value function but also a decrease in mean productivity over time.

The paper also complements studies on the effects of intermediate input imports on firms' performance. Several studies find that imported intermediate inputs are conducive to productivity growth through three channels: learning or transferring embedded technology; improvements in input quality; and increases in input varieties. Almeida and Fernandes (2008) report that on average 53% of technological innovations are embodied in new machinery or equipment and are transferred from developed to developing countries through exports and multinational firms. Goldberg et al. (2010) indicate that Indian firms, in particular, increase their product scope because of the fact that they can access previously unavailable new input varieties. Bas and Strauss-Kahn (2014) also argue that using more varieties of imported inputs results in higher TFP and greater export scope. However, some works also find that the use of imported inputs plays only a minor role in productivity improvement; see van Biesebroeck (2003) and Muendler (2004). My paper, similarly, points out an additional channel through which imported intermediates retard productivity growth. In doing so, it partly explains why not all estimated results about the effects of imported inputs on productivity are positive in the literature.

2.2 Data Sources

The primary firm-level data used in this paper are extracted from three sources. First, China's General Administration of Customs provides all Chinese trade transactions by both importing and exporting firms at the HS 8-digit level for the years 2000-2006. I aggregate the raw customs data to the HS 6-digit level for the concord of the product codes consistently over time. I drop firms from the sample whose ratios of processing exports over total exports are higher than 20 percent because processing trade has been exempted from import tariffs on imported inputs and materials in both the pre-WTO and post-WTO accession periods. The second data source used in this study is the Annual Survey of Manufacturing Enterprises (ASME), which is maintained by the National Bureau of Statistics of China. These data cover all manufacturing plants with annual sales greater than US$800,000 during the years 1998 to 2007, and it contains establishment-level information on sales, input usages, total R&D expenditure, etc. I treat a plant as a "firm" throughout the paper. Finally, the third data source is the China's State Intellectual Property Office (SIPO). The SIPO dataset contains detailed information on each patent filing from1985 until now.

It is important here to clarify some concerns regarding the use of patent data. The OECD report (2009) explains well the advantages and drawbacks of using the patents as a proxy for something as abstract as innovation. An alternative would be to use R&D expenditure (from the ASME) as a proxy. However, various subsidy schemes can severely distort R&D expenditure. For example, companies classified as High and New Technology Companies can receive a corporate income tax rate (CIT), which cuts taxes from the standard rate of 25% down to 15%. Companies meeting the government's criteria can get 150% of eligible R&D expenses deducted before the CIT. PricewaterhouseCoopers report (2015) also points out that a slight changes in R&D behavior might result in substantial R&D tax benefits. On the contrary, according to Li (2012), subsidy programs are subsidizing patent filing, which varies across regions, averaging around just a few hundred USD per patent filing. Hence, we expect that the distortion of R&D data is much more severe than that of patent data; moreover, the quality R&D data is also worrisome. Among R&D performers (with positive individual R&D expenditure), around 78.59 percent of R&D expenditure reported in the ASME is smaller than US$100, and about 7.83 percent of them are even less than US$1. Hence, I use the patent instead as a measure of innovation efforts in the main estimation. [1]

Next, I construct a two-digit industry-level dataset consistent with the ISIC Rev 3. The import tariff data I use comes from the WTO website, available as most-favored-nation (MFN) applied tariff at the HS 6-digit level for 2000-2006. I calculate the average industry-level import tariff rates weighted by the import share constructed from Customs data. The average capital intensity and skill intensity in a two-digit-ISIC industry in the US between 2001 and 2006 are constructed using the data from the NBER productivity database.

Finally, I use various sources to compile a dataset with the key country characteristics from 2006. Country-level R&D spending from 2006 comes from the World Bank Development Indicators. Wage data are extracted from ILO data described by Oostendorp (2005). Distance, land area and common language come from CEPII. Country-level measured TFP, physical capital, population, and total workforce come from the Penn World Table described by Heston, Summers, and Aten (2011). Control of corruption comes from the World Bank's Worldwide Governance Indicators. Finally, years of schooling, as a proxy for human capital, is obtained from Barro and Lee (2010).

2.3 Descriptive Evidence

In this section, I assess the model's predictions and assumptions. I use the differential changes in Chinese tariffs across industries to show that firms undertake less innovation in industries where import tariffs drop more sharply.

[1] There is an ongoing debate on the correlation between R&D expenditure and patents in China. Hu and Jefferson (2009) point out the weak linkage between patent and R&D cost, but Dang and Motohashi (2015) suggest that patent count is correlated with R&D input and financial output and so patents statistics are meaningful indicators of innovations. Liu and Qiu (2016) find that firms' R&D investment increases the patent filings.

My analysis is based on a panel data from the period of 2001-2006. I overcome the identification problem by exploiting the presence of large variations in input tariff reduction across industries before and after China's WTO accession. The magnitude of tariff reduction was predetermined in the pre-WTO period and thus treated as an exogenous shock; see Liu and Qiu (2016). Since I expect firms that experience larger input tariff cuts to make bigger adjustments, I examine the correlation between (weighted) average import tariffs and firms' probability of investing in patent applications. Specifically, I use the probit model for the estimation:

$$y_{ijt} = \beta_\tau \tau_{it} + X_{jt} + \lambda_i + \lambda_t + \epsilon_{ijt},$$

where j indexes firms; i indexes two-digit ISIC industries; t indexes the year between 2001 and 2006; y_{ijt} is a dummy variable that takes the value of 1 if firm j filed a patent in year t ; τ_{it} represents China's imported tariffs, weighted by import share and varying across both industry i and year t; [2] X_{jt} includes firm controls varying over time; λ_i represents a composite of industry-level controls for those industry-level characteristics that are likely to affect innovation activities. These latter characteristics are measured using the U.S. data to avoid endogeneity problems, following the example of Bustos (2011). Finally, λ_t represents the year fixed effects, controlling for all yearly shocks common to all industries; and ϵ_{ijt} is the error term.

Table 2.1 presents the regression results based on this main specification, with control variables introduced step by step. All estimations show that import tariff reduction does correlate with firms' decline in innovation participation of firms. In Column 1, with only industry fixed effects and year fixed effects being controlled for, I find a statistically significant and positive estimate for τ_{it}. Next, in Column 2, I include several time-varying firm characteristics that may influence innovation activities, such as total sales, capital-labor ratio, and export status. Column 3 then introduces industry control variables: capital-intensity and skill-intensity, which are measured as the average capital per employee and average labor compensation, respectively, per employee in two-digit ISIC industries in the U.S. from 2001- 2006. The coefficients of τ_{it} in the regression that includes firm and industry controls are still positive and significant, indicating that a reduction of input tariffs decreases innovation probability.

[2] Each firm is assigned to a 4-digit Chinese industry Code (CIC) in ASME every year, and there is a concordance between 4-digit CIC and 4-digit ISIC. I use a balanced panel data for the estimation, but only around 16 percent of firms (7352 out of 45,966) kept the same four-digit ISIC between 2001 and 2006. Instead, all the firms used the same two-digit ISIC throughout six years. So, I apply two-digit instead of four-digit ISIC industry tariffs here.

Table 2.1: Innovation Participation and Import Tariffs

	Patent Status			ln(Number of Patents+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Import Tariffs	0.612***	0.537***	0.551***	0.139***	0.108***	0.118***
	(0.146)	(0.186)	(0.191)	(0.0314)	(0.0258)	(0.0234)
Export Status		0.264***	0.264***		0.0653***	0.0663***
		(0.0331)	(0.0315)		(0.00706)	(0.00751)
ln(Sales)		0.299***	0.300***		0.0865***	0.0881***
		(0.0124)	(0.0124)		(0.00522)	(0.00532)
ln(Capital/Labor)		0.0579***	0.0592***		0.00778***	0.00852***
		(0.0102)	(0.0119)		(0.00192)	(0.00194)
ln(Capital Intensity)			0.266			0.0521***
			(0.164)			(0.00874)
ln(Skill Intensity)			0.252			0.0444***
			(0.307)			(0.0103)
Constant	-1.866***	-7.370***	-8.602***	0.0815***	-1.466***	-1.727***
	(0.0482)	(0.201)	(1.223)	(0.0107)	(0.105)	(0.0993)
Year FE	yes	yes	yes	yes	yes	yes
Firm FE	no	no	no	yes	yes	yes
Observations	275,796	275,794	266,139	275,796	275,794	266,139

Notes: Standard errors are clustered at the industry level. *** p<0.01, ** p<0.05, * p<0.1.

Since firm fixed effects cannot be consistently estimated by probit due to an incidental parameters problem, I consider a fixed-effect linear model with an alternative dependent variable $y_{ijt} = \ln(Y_{ijt} + 1)$ where Y_{ijt} represents the total number of patent filings by firm j in industry i in year t. This log-like transformation allows me to avoid the problem of too many zeros. The results in Columns 4-6 of Table 2.1 indicate that the negative effect of input tariff reduction on innovation remains.

Liu and Qiu (2016) were the first to document that input tariff reduction negatively impacts firms' innovation decisions. Though both Liu and Qiu (2016) and my own paper use Chinese firm-level data from 2001-2006 and adopt patent data as a proxy for innovation efforts, their research focuses on the effects of imported inputs on intensive margin adjustment of R&D, while mine brings attention to the extensive margin response. In particular, Liu and Qiu (2016) show that the average 4.64% input tariff cut between 2001 and 2006 reduces the number of patent filings of a firm by 0.0132. My specification results emphasize that share of R&D performers in the economy is positively related to input tariffs. Hence, better access to imported inputs appears to lead fewer firms to participate in R&D. A more fundamental difference between my work and Liu and Qiu (2016) is that I develop and estimate a structural model to quantify the impact of trade liberalization on firms' innovation participation rate, which is something the earlier work did not do.

2.4 Theoretical Framework

The theoretical model developed in this section essentially combines the static, multi-country sourcing model from Antràs, Fort, and Tintelnot (2017) and the dynamic investment model by Aw, Roberts, and Xu (2011). I allow firms to be heterogeneous in productivity, which, together with the exogenous trade shocks, determines each firm's global sourcing strategy and degree of incentive to invest in R&D. In turn, each R&D activity has feedback effects that can alter the path of future productivity and long-run profits for any given firm. In a nutshell, I divide each firm's decision-making into a static component - whereby the multi-country sourcing decision (along with the firm's productivity) affects the current profit - and into a dynamic component in which R&D decision determines productivity evolution in the next period.

2.4.1 Static Decision

There is a set of N countries, designated individually as c, n, or m, with c always referring to the home country of China. Each country is populated by an exogenous measure of workers, wherein each worker supplies her unit of labor inelastically, and labor is the only factor of production. I further simplify matters by assuming that there exists a perfectly competitive non-manufacturing sector that competes for labor with manufacturing firms and that its production technology is linear in labor. I assume that the non-manufacturing sector is large enough to guarantee that the wage rate w_n in each country n is pinned down by labor productivity in that sector. I also assume that the non-manufacturing sector's output is homogeneous, freely tradable across borders, and serves as a numeraire in the model. I can thus treat wages as exogenous in solving for firms' problem in each country's manufacturing sector.

As in the Krugman (1980) model, an individual in a country n values the consumption of differentiated varieties of manufactured goods, according to a standard symmetric CES aggregator:

$$U_n = \Big(\int_{\omega \in \Omega_n} y_n(\omega)^{(\sigma-1)/\sigma} d\omega \Big)^{\sigma/(\sigma-1)}, \sigma > 1,$$

where $y_n(\omega)$ is the quantity consumed in country n of variety ω; and Ω_n is the set of manufacturing varieties available to consumers in country $n \in N$. These preferences are assumed to be common worldwide and give rise to the following demand for variety ω

in country n:

$$y_n(\omega) = E_n P_n^{\sigma-1} p_n(\omega)^{-\sigma},$$

where $p_n(\omega)$ is the price of variety ω; P_n is the Dixit-Stiglitz price index; and E_n is aggregate spending on manufacturing goods in country n. For simplicity, we will assume that final-good varieties are prohibitively costly to trade across borders.

The market structure of final-good production is characterized by monopolistic competition, and there is free entry into the industry. There exists a fixed measure of final-good producers in each country $n \in N$, and each of these producers owns a blueprint to produce a single differentiated variety ω. Following Melitz (2003) and Chaney (2008), a firm learns its productivity φ after incurring a fixed entry cost equal to f_n^e units of labor in country n. This productivity is drawn from a country-specific Pareto distribution $F_n(\varphi)$, with support in $[\underline{\varphi_n}, \infty]$:

$$F_n(\varphi) = 1 - (\varphi / \underline{\varphi_n})^{-\theta}.$$

Thus, the heterogeneity of producers is decreasing in θ. Since all firms with the same productivity within a particular country will act the same way, I will then identify each firm by its productivity φ instead of by its distinct final variety ω.

In addition to final-good producers, there is a competitive fringe of intermediate suppliers who sell their products (to final-good producers) at marginal cost. Consumers do not purchase intermediates and these intermediates are traded internationally. Production of final-good varieties requires the assembly of a continuum of measure one of firm-specific intermediates, denoted by $v \in [0, 1]$. These intermediates are assumed to be imperfectly substitutable with each other, with a constant and symmetric elasticity of substitution equal to ρ. A firm's productivity φ has no impact on the mapping between the bundle of inputs and final-good production, but φ does affect firms' productivity with respect to producing each intermediate. For each intermediate v, a firm must decide whether to produce the intermediate by itself or to outsource it to another intermediate supplier. I model the boundaries of firms to be determined by the existence of inter-firm transaction costs, as in the transaction-cost economics literature on the vertical integration problem (see Williamson, 1985). I discuss each of the two options in turn.

In-house Production. A firm can employ labor to produce intermediates. Each unit of labor generates $z_n(\varphi, v)$ units of v in country n by firm φ, and $z_n(\varphi, v)$ is a stochastic productivity realization that follows a Fréchet distribution,

$$\Pr(z_n(\varphi, v) < z) = \exp(-\varphi z^{-\theta}).$$

The input-specific productivity draws $z_n(\varphi, v)$ are independent across φ and v. The term φ captures a firm's productivity relative to the technology of the state in country n. Higher φ will, on average, lead to higher realizations of the productivity parameters $z(\varphi, v)$. The parameter θ governs the distribution of productivities across inputs within countries: as the value of θ increases, the heterogeneity of productivity across intermediates declines. Given the positive wage vector w_n, we can denote the unit cost of v under in-house production by $p_n^I(\varphi, v)$. Then,

$$p_n^I(\varphi, v) = \frac{w_n}{z_n(\varphi, v)}.$$

Since the productivity draws $z_n(\varphi, v)$ are independent across v, the probability that firm φ is able to produce an intermediate v for a price less than p will be the same for all inputs $v \in [0, 1]$. That is:

$$G_n^I(\varphi, p) = 1 - \exp(-\varphi w_n^{-\theta} p^{\theta}).$$

Outsourcing. As an alternative to in-house production, a firm φ can outsource intermediate production to other firms. Although intermediates are produced worldwide, a firm only acquires the capability to outsource inputs from country n after incurring a firm-specific fixed outsourcing cost equal to $f_m(\varphi)$ units of labor in home country n. Although it is not important how we sort countries, I place a list of N countries in a fixed order throughout the estimation. Denote the *global sourcing strategy* of firm φ based in country n by $\mathbf{D}_n(\varphi)$ such as $\mathbf{D}_n(\varphi) = \{D_n^1(\varphi), ..., D_n^N(\varphi)\} = \{0, 1\}^N$. When $D_n^m(\varphi) = 1$, firm φ has paid the associated fixed outsourcing cost $w_n f_m(\varphi)$ and outsource inputs from supplier(s) in country m. By contrast, when $D_n^m(\varphi) = 0$, the firm φ does not pay the fixed outsourcing cost; hence, it cannot source any inputs from country m.

Building on Eaton and Kortum (2002), we shall assume that *homogeneous* input suppliers in country m use a_m units of labor to produce an intermediate v and the value of a_m comes from the Fréchet distribution, as follows:

$$Pr(a_m(v) \geq a) = \exp(-T_m a^{\theta}), \text{ with } T_m > 0.$$

where T_m governs the state of technology in country m. Since the productivity draws $a_m(v)$ are independent across v, the probability that firm φ is able to purchase an intermediate v from country m for a price lower than p will be the same for all inputs $v \in [0,1]$. That is:

$$G_m(p) = 1 - \exp(-T_m w_m^{-\theta} p^{\theta}). \tag{2.1}$$

Shipping intermediates from country m to country n entails iceberg trade cost d_{nm} and import tariffs τ_{nm}. Let $d_{nn} = \tau_{nn} = 1$. Also, we will assume further that outsourcing is subject to contracting frictions that take the form of a multiplicative wedge $\delta_{mn} = \delta_{nn} = \delta \geq 1$ on the cost of intermediates. One should think of δ as being determined by the efficiency loss associated with a hold-up problem: when the traded intermediates are firm-specific, the buyer has an incentive to hold up the seller by withholding payment. However, when enforcement of the contract ensures that the supplier receives her full return on the investment, then equilibrium is efficient with $\delta = 1$, and firms outsource all the intermediate inputs. In other words, although the intermediates market is perfectly competitive, the existence of contracting frictions in current institutions leads to too much in-house production, and the outcome is below the first-best in a frictionless market. Taking these frictions into account, the price that firm φ with sourcing strategy $\mathbf{D}_n(\varphi)$ based in country n pays for input v can be denoted by:

$$p_n^x(\varphi, v) = \min_{m \in \mathbf{D}_n(\varphi)} \tau_{nm} d_{nm} \delta p_m(v),$$

where the price $p_m(v)$ of intermediate v from country m is drawn from equation (2.1).

Therefore, given the perfect substitutability between two options, the realized price that firm φ based in country n pays for intermediate $v \in [0,1]$ is

$$p_n(\varphi, v) = \min(p_n^l(\varphi, v), p_n^x(\varphi, v)).$$

2.4.2 Dynamic Decision

Following Aw, Roberts, and Xu (2011), there are two intertemporal linkages. First, there is a sunk startup cost for a firm when it begins investing in R&D, which makes the firm's past R&D status a state variable in the investment function. Second, productivity is endogenous, meaning that the firm's R&D choices can affect its future productivity.

I begin with a description of firm's state variables: productivity φ_t, and past R&D choice d_{t-1}. I assume that productivity evolves as a Markov process that depends on the firm's R&D participation along with a random shock:

$$\varphi_t = \alpha_0 + \alpha_1 \varphi_{t-1} + \alpha_2 \varphi_{t-1}^2 + \alpha_3 d_{t-1} + \epsilon_t, \tag{2.2}$$

where d_{t-1} is a firm's R&D participation in the previous period, and ϵ_t is a i.i.d. shock with zero mean and variance, σ_ϵ^2. The inclusion of d_{t-1} recognizes that innovations are allowed to shift the mean of the distribution of future firm productivity systematically. The stochastic shocks ϵ_t reflect the inherent randomness in the evolution of a firm's productivity; that is, such evolution is not anticipated by the firm and, by construction, is not correlated with d_{t-1} or φ_{t-1}. Because of the persistence in productivity, the shocks in any period (and their consequences) affect future productivity, rather than having isolated, transitory effects.

I model the R&D choice as a discrete 0/1 variable, which implies that any firm that undertakes R&D experiences the same expected incremental increase in its productivity, as captured by the parameter α_3. This assumption is consistent with the evidence reported by Aw, Roberts, and Winston (2007), who find that productivity evolution for Taiwanese electronics producers is affected by the discrete R&D variables. They also find that firm productivity is a significant determinant of the discrete decision to undertake R&D activity at all, although they find little evidence that productivity is correlated with the level of R&D spending. This finding makes sense, given that, as Mairesse, Mohnen, and Kremp (2005) point out, measurement error in the level of R&D expenditure is more substantial than the error in the discrete participation variable - a fact which can result in R&D expenditure's having small estimated impacts on the productivity improvement. Therefore, the discrete R&D participation variable is a robust indicator of a firm's investment strategy and clearly distinguishes firms that choose to invest in uncertain R&D projects from those who do not - a core distinction, given that the focus of this empirical model is to measure how a firm's value might correlate with its investment strategy over the long term.

I observe that 47.76 percent of firms continue investing in R&D in the current period if they have invested in R&D in the preceding period; while only 3.06% of firms without R&D experiences in the preceding period start investing in R&D today. It is likely that a firm that conducts R&D continuously over time can generate a given innovation with lower expenditures than a firm that is just beginning to invest in R&D; this is because the former can rely on past expertise or synergy effects from its previous projects. To capture heterogeneity in innovation costs, I model a firm's innovation cost C as a draw from an exponential distribution, where the mean of the distribution varies with the firm's R&D

experience yesterday:

$$C(d_{t-1}) \sim \exp(d_{t-1}\gamma_F + (1 - d_{t-1})\gamma_S). \tag{2.3}$$

A firm engaging in R&D experience in the preceding period has to pay a fixed cost drawn from a known distribution with a mean of γ_F, while a firm without R&D experience in the preceding period has to pay a startup cost drawn from a distribution with a mean of γ_S. The parameters γ_F and γ_S capture differences in the fixed and startup cost distribution. Before it decides to invest, a firm observes its innovation cost, and that cost acts as an additional source of unobserved heterogeneity driving the firm's investment behavior.

Assume that, at the start of period t, a firm based in country n observes its current relative productivity level φ_t and knows both its short-run profit function $\pi_n(\varphi_t)$ and its productivity evolution function. The firm's state variables $s_t = (\varphi_t, d_{t-1})$ evolve endogenously as the firm makes its decision to conduct R&D, d_t. Given the state variables and the discount factor β, the firm's value function $V_n(s_t)$ - before it realizes the innovation cost $C(d_{t-1})$ - can be written as:

$$V_n(s_t) = \pi_n(\varphi_t) + \int_C \max_{d_t} \left(\beta E_t V_n(s_{t+1}|\varphi_t, d_t = 1) - C(d_{t-1}); \beta E_t V_n(s_{t+1}|\varphi_t, d_t = 0)\right) dC,$$

where the short-run profit $\pi_n(\varphi_t)$ depends only on current productivity φ_t and some exogenous aggregate parameters.

2.5 Equilibrium

I solve for the equilibrium of the model in four steps. First, I describe optimal firm behavior conditional on a given sourcing strategy $\mathbf{D}_n(\varphi)$. Second, I characterize the choice of this sourcing strategy. Third, I solve for firm-level investment decisions. Finally, I aggregate the firm-level decisions and derive the general equilibrium of the model.

2.5.1 Firm Behavior Conditional on a Sourcing Strategy

A firm based in country n with relative productivity φ can purchase inputs from country m if and only if it pays a firm-country-specific fixed outsourcing cost $f_n^m(\varphi)$. Suppose firm φ has incurred all fixed outsourcing costs associated with a given sourcing strategy:

$$\mathbf{D}_n(\varphi) = \{D_n^1(\varphi), ..., D_n^N(\varphi)\} = \{0, 1\}^N.$$

Using the properties of the Fréchet distribution, one can show that this firm buys a positive measure of intermediates from each country in its sourcing strategy $\mathbf{D}_n(\varphi)$, and the share of intermediate input purchases sourced from any country m (including the home country n) is simply given by:

$$\chi^x_{nm}(\varphi) = \frac{T_m(\tau_{nm}d_{nm}\delta w_m)^{-\theta}}{\Theta_n(\varphi)}, \text{ if } D^m_n(\varphi) = 1, \tag{2.4}$$

and $\chi^x_{nm}(\varphi) = 0$ otherwise, where

$$\Theta_n(\varphi) \equiv \varphi w_n^{-\theta} + \sum_{m=1}^{N} D^m_n(\varphi) T_m(\tau_{nm}d_{nm}\delta w_m)^{-\theta}. \tag{2.5}$$

Similarly, the share of in-house produced inputs satisfies

$$\chi^l_n(\varphi) = \frac{\varphi w_n^{-\theta}}{\Theta_n(\varphi)}. \tag{2.6}$$

The term $\Theta_n(\varphi)$ summarizes the *sourcing capability* of firm φ from country n. Note that each country m's market share in firm φ's purchases of intermediates corresponds to the country's contribution to the firm's sourcing capability. Countries in the set $\mathbf{D}_n(\varphi)$ with lower wages w_m; more advanced technology T_m; lower import duty τ_{nm}; or lower trade costs d_{nm} will have higher market shares in the intermediate input purchases of firms based in country n. Higher contraction friction δ leads to more in-house production and fewer outsourcing activities. I will refer to the term

$$\xi_{nm} \equiv T_m(\tau_{nm}d_{nm}\delta w_m)^{-\theta}, \tag{2.7}$$

as the *sourcing potential* of country m from the perspective of firms in n.

After choosing the least-cost source of supply (including in-house production) for each intermediate v, the overall unit cost faced by firm φ from country n can be expressed as:

$$c(\varphi) = (\gamma\Theta_n(\varphi))^{-1/\theta}$$

where $\gamma = \left[\Gamma(\frac{\theta+1-\rho}{\theta})\right]^{\theta/(1-\rho)}$ and Γ is the gamma function. [3] Adding an additional new location to the set $\mathbf{D}_n(\varphi)$ increases the sourcing capability $\Theta_n(\varphi)$ of the firm and

[3] To sure a well-defined unit cost index, we need to assume $\theta > \rho - 1$ as in Eaton and Kortum (2002). Otherwise, if the elasticity of substitution among inputs ρ is very high and the distribution of input prices has a fat lower tail, final good producers will shove their expenditure into inputs with prices arbitrarily close to zero, and then the unit cost of final varieties will also go to zero.

necessarily lowers its effective unit cost. Intuitively, a new location grants the firm an additional cost draw for all varieties v, and it is thus natural that this greater competition among suppliers will reduce the expected minimum sourcing cost per intermediate. The addition of a country to the set $\mathbf{D}_n(\varphi)$ lowers the expected price paid for all varieties v, not just for those that are ultimately sourced from the country being added to $\mathbf{D}_n(\varphi)$. Second, Equation (2.6) implies that so long as the sourcing strategy remains fixed, a firm with higher productivity draw φ tends to have a larger share of in-house produced inputs. However, a more productive firm also sources inputs from a bigger set of countries, as $\Theta_n(\varphi)$ is nondecreasing on φ. Hence, a productive firm is not necessarily associated with a higher cost share of self-made inputs.

Given a sourcing strategy $\mathbf{D}_n(\varphi)$, the profit function of firm φ from n is

$$\pi_n(\varphi; \mathbf{D}_n(\varphi)) = \left[\varphi w_n^{-\theta} + \sum_{m=1}^{N} D_n^m(\varphi)\xi_{nm}\right]^{\frac{\sigma-1}{\theta}} B_n - w_n \sum_{m=1}^{N} D_n^m(\varphi) f_{nm}(\varphi), \tag{2.8}$$

with

$$B_n = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} E_n P_n^{\sigma-1} \gamma^{\frac{\sigma-1}{\theta}}, \tag{2.9}$$

where the sourcing potential ξ_{nm} is defined in Equation (2.7) . Note that while the fixed outsourcing costs are independent of firms' productivity level, these are country-firm pairwise and thus indexed by φ. Besides, it is clear that when deciding whether to add a new country m to the set $\mathbf{D}_n(\varphi)$, a firm trades off the reduction in unit costs associated with including that country in the set $\mathbf{D}_n(\varphi)$ against the payment of additional fixed cost $w_n f_{nm}(\varphi)$.

2.5.2 Optimal Sourcing Strategy

Each firm's optimal sourcing strategy is a combinatorial optimization problem in which a set $\mathbf{D}_n = \{D_n^1, ..., D_n^N\} = \{0,1\}^N$ is chosen to maximize the firm's profits $\pi_n(\varphi, \mathbf{D}_n)$ in Equation (2.8). that is,

$$\max_{\mathbf{D}_n} \pi_n(\varphi; \mathbf{D}_n) = \left[\varphi w_n^{-\theta} + \sum_{m=1}^{N} D_n^m \xi_{nm}\right]^{\frac{\sigma-1}{\theta}} B_n - w_n \sum_{m=1}^{N} D_n^m f_{nm}(\varphi). \tag{2.10}$$

where the sourcing potential ξ_{nm} is defined in equation (2.7) and the demand residual B_n is given in equation (2.9). The problem in equation (2.10) is not straightforward to solve. Countries are heterogeneous in two dimensions: sourcing potential ξ_{nm} and fixed sourcing cost f_{nm}. As long as $\frac{\sigma-1}{\theta} \neq 1$, the decision to include an additional country

m in the set $\mathbf{D}_n(\varphi)$ depends on the number and characteristics of the other countries in the current set. One might use brute force to calculate firms' profits for all possible combinations of locations and then pick the best strategy yielding the highest level of profits. However, this would amount to computing profits for 2^N strategies, which is infeasible unless one were to choose a small enough number N of candidate countries.

Antràs, Fort and Tintelnot (2017) solve this combinatorial optimization problem in the case of $\sigma - 1 \geq \theta$, and they propose making additional assumptions to characterize optimal sourcing problems when $\sigma - 1 < \theta$. In particular, the additional assumption they suggest is that the fixed outsourcing costs are common across all sourcing countries - an assumption that appears particularly unlikely. When $\sigma - 1 \geq \theta$, the extensive margin of sourcing at the firm level (i.e., the number of sourcing countries) necessarily increases weakly in firm productivity. That is because if $\sigma - 1 \geq \theta$, the marginal gain from adding a new location to the set $\mathbf{D}$ cannot possibly be reduced by the addition of other countries to the set. However, once $\sigma - 1 < \theta$, the number of sourcing countries is not necessarily increasing in firm productivity. For example, a high-productivity firm might pay a large fixed cost to outsource to a country m with an extremely high sourcing potential, after which the firm might not have any incentives to add further locations to its sourcing strategy; meanwhile, instead, a low-productivity firm might not find it profitable to outsource to m, yet, at the same time, may well find it optimal to source from multiple foreign countries associated with lower fixed costs. Next, consider an alternative case wherein the country-specific fixed outsourcing costs are identical across firms: in such a situation, it is easy to show that when $\sigma - 1 \geq \theta$, there still exists a strict hierarchy of import sources, thus making the computational problem easier.

It is worthwhile to mention that the case of $\sigma - 1 < \theta$ does occur in practice. First, $\rho - 1 < \theta$ is the only technical condition required to guarantee that the final goods' prices are bounded away from zero, and so consumers' utility does not explode to infinity. The model does not impose any restrictions on the relationship between $\sigma - 1$ and θ. Second, the case of $\sigma - 1 < \theta$ is more likely to apply whenever intermediate efficiency level is NOT very heterogeneous across markets (high θ), or whenever profits are NOT very responsive to variable cost reductions (lower σ); hence, marginal gain from adding a new location to the set $\mathbf{D}$ is reduced by the addition of other countries to the set. For example, suppose there are two countries: a country m with considerable sourcing potential and high fixed outsourcing costs, and a country n with low sourcing potential and small fixed outsourcing costs. Now, a highly productive firm pays a substantial fixed cost to outsource to country m, after which that firm might have no further incentive to outsource from n. On the contrary, a low-productivity firm might find it unprofitable to outsource to country m, as its sales are small, so it decides to outsource to n instead. In this circumstance, the low-productivity firm's sourcing country set is not a subset of the highly productive firm's sourcing strategy. This entire scenario is a particular phenomenon under the condition of $\sigma - 1 < \theta$.

Next, I will introduce how I solve this combinatorial discrete choice problem under $\sigma - 1 < \theta$ without extra assumptions. As in Antràs, Fort and Tintelnot (2017), let the fixed outsourcing costs be country-firm pairwise. For notation simplicity, I suppress the origin country subscript n and use $\sum_{m=1}^{N} D_m \xi_{nm}$ instead of $\varphi w_n^{-\theta} + \sum_{m=1}^{N} D_m \xi_{nm}$ throughout this subsection. The profit-maximization problem in equation (2.10) can be rewritten as

$$\max_{D^1,\ldots,D^N \in \{0,1\}} \pi(\varphi, \mathbf{D}) = \Big[\sum_{m=1}^{N} D^m \xi_m \Big]^{\frac{\sigma-1}{\theta}} B - w \sum_{m=1}^{N} D^m f_m(\varphi). \tag{2.11}$$

where the sourcing strategy $\mathbf{D} = \{D^1, \ldots, D^N\}$ and the *optimal* sourcing strategy for firm φ is denoted by $\mathbf{D}(\varphi) = \{D^1(\varphi), \ldots, D^N(\varphi)\}$.

Define an indicator function $V^m(\mathbf{D})$ as:

$$V^m(\mathbf{D}) = 1\Big[B\big[\big(\sum_{l \neq m} D^l(\varphi)\xi_l + \xi_m\big)^{\frac{\sigma-1}{\theta}} - \big(\sum_{l \neq m} D^l(\varphi)\xi_l\big)^{\frac{\sigma-1}{\theta}}\big] - w f_m(\varphi) \geq 0\Big]. \tag{2.12}$$

where the function $V^m(\mathbf{D})$ takes a value of either zero or one. The right-hand side of equation (2.12) is the difference of equation (2.11) at $D^m = 1$ relative to its value at $D^m = 0$. The intuition of equation (2.12) is that whenever including country m in the sourcing strategy $\mathbf{D}$ raises firm-level profits $\pi(\varphi, \mathbf{D})$, a firm will source from country m and the function $V^m(\mathbf{D})$ takes a value of one.

Let the comparison between vectors be coordinatewise: a vector $\mathbf{D}$ is bigger than a vector $\mathbf{D}'$ if and only if every element of $\mathbf{D}$ is weakly larger: $\mathbf{D} \geq \mathbf{D}'$ iff $D_n \geq D'_n$ $\forall i$. The vectors are identical if both $\mathbf{D} \geq \mathbf{D}'$ and $\mathbf{D} \leq \mathbf{D}'$.

Proposition 8. *An indicator function $V^m(.)$ is a weakly decreasing function such that $V^m(\mathbf{D}') \leq V^m(\mathbf{D})$ whenever $\mathbf{D}' > \mathbf{D}$ and $\sigma - 1 < \theta$.*

Proposition 8 says that the value of adding source m is always smaller when starting from an expanded set of suppliers. The usefulness of this proposition is best demonstrated with two examples. If $V^2(\varphi, \mathbf{D}) = 0$ and $\mathbf{D} = \{0, 0, 0\}$, then Proposition 8 implies that firm φ does not source from the second country in its optimal sourcing strategy. Otherwise, suppose the optimal strategy is $\mathbf{D}(\varphi) = \{0, 1, 1\}$. There exists a

$\mathbf{D}' = \{0, 1, 1\} > \{0, 0, 0\} = \mathbf{D}$ and $V^2(\mathbf{D}') > V^2(\mathbf{D})$, which violates the Proposition 8. In the second example, let us start with $V^2(\varphi, \mathbf{D}) = 0$ and $\mathbf{D} = \{1, 0, 0\}$. Then Proposition 8 rules out all possible strategies whose first and second elements take the value of one, that is, firm φ shall not outsource to the first and the second countries simultaneously. Otherwise, there exists a $\mathbf{D}'$ such that $\mathbf{D}' > \mathbf{D}$ and $V^2(\mathbf{D}') > V^2(\mathbf{D})$. In the structural analysis section, I will leverage the property of $V^m(.)$ to devise an iterative algorithm to solve the problem defined in Equation (2.10) efficiently.

2.5.3 Firm's R&D Choices

Recall that given the state variables $s_t = (\varphi_t, d_{t-1})$ and discount factor β, the firm's value function $V(s_t)$, before it realizes the innovation cost $C(d_{t-1})$ is:

$$V_n(s_t) = \pi_n(\varphi_t) + \int_C \max_{d_t} \big(\beta E_t V_n(s_{t+1}|\varphi_t, d_t = 1) - C(d_{t-1}); \beta E_t V_n(s_{t+1}|\varphi_t, d_t = 0)\big) dC,$$

where the current-period profit $\pi_n(\varphi_t)$ is given by Equation (2.8), and the innovation cost $C(d_{t-1})$ is drawn from an exponential distribution, where the mean of distribution varies with its prior R&D experience, as in Equation (2.3). Define the expected future values of the firm conditional on R&D choice d_{jt} as:

$$E_t V_n(s_{t+1}|\varphi_t, d_t) = \int_\varphi V_n(s_{t+1}) dH(\varphi_{t+1}|\varphi_t, d_t).$$

In this equation, the evolution of productivity $dH(\varphi_{t+1}|\varphi_t, d_t)$ is conditional on d_t because of the assumption in Equation (2.2).

Next, define the marginal benefit of conducting R&D as the differences in the expected future returns between choosing and not choosing R&D:

$$\Delta E_t V_n(\varphi_t) \equiv \beta E_t V_n(s_{t+1}|\varphi_t, d_t = 1) - \beta E_t V_n(s_{t+1}|\varphi_t, d_t = 0). \tag{2.13}$$

Therefore, in the equilibrium, firm φ chooses to invest in R&D if $\Delta E_t V_n(\varphi_t) \geq C(d_{t-1})$. This is a condition used in the empirical model to explain a firm's observed R&D choice.

2.5.4 Comments on the Equilibrium

I want to begin with some properties of the profit function when holding constant the market demand-level B_n. First, not surprisingly, a reduction in an import tariff τ_{nm} or a fixed outsourcing cost f_{nm} weakly increases the firm's sourcing capability $\Theta_n(\varphi)$ and thus firm-level profits. Second, whenever $\sigma - 1 < \theta$, a reduction of any τ_{nm} or f_{nm} weakly decreases firm-level input purchases from all sources, including the self-made inputs. To see this, CES demand implies that the total value of self-made inputs is a fraction $(\sigma-1)\xi_n^l(\varphi)$ of firm profits, and using Equations (2.6) and (2.8), they can thus be expressed as:

$$M_n^l(\varphi) = (\sigma-1)B_n[\Theta_n(\varphi)]^{\frac{\sigma-1-\theta}{\theta}} T_n \varphi w_n^{-\theta}.$$

When $\sigma - 1 < \theta$, $M_n^l(\varphi)$ is decreasing in all the terms in $\Theta_n(\varphi)$. Intuitively, when the demand is sufficiently elastic (i.e., σ is high enough), outsourcing firms can grab more market share in response to a reduction of τ_{nm} or f_{nm}; then, after expansion of the market, they want to decrease marginal production costs through outsourcing to a greater number of countries and through investing more in R&D to boost in-house production efficiency. Hence, in response to lower trade cost, the scale effect through the demand encourages more R&D spending and more input usages from all sources. But when the strength of the comparative advantage in the intermediate-good market across countries is sufficiently high (i.e., θ is low enough), firms shift input spending toward the locations whose sourcing costs have been reduced during the trade liberalization. There is a direct substitution effect between outsourcing inputs and self-made inputs. As firms use more outsourcing and fewer self-made inputs, the gains from investing in in-house production efficiency are lower. As a result, the substitution effect related to input cost shares discourages R&D participation and has a negative impact on self-made input cost share $\xi_n^l(\varphi)$. Whenever $\sigma - 1 < \theta$, the substitution effect dominates the scale effect. Thus, we observe a lower R&D participation rate in response to trade liberalization in China between 2001 and 2006.

Proposition 9. *Holding constant the market demand-level B_n, whenever $\sigma - 1 < \theta$, an increase in the sourcing potential ξ_{nm} or a reduction in fixed cost f_{nm} of any country m, (weakly) decreases the self-made input usage as well as the R&D participation rate.*

Note that Proposition 9 applies when holding market demand-level fixed. In the general equilibrium, firms' decisions affect the market demand-level as well as the aggregate price index. As we shall see in the counterfactual exercise, the endogenous response of the market demand is quantitatively important in the estimation. Despite this limitation, Proposition 9 proves to be very useful in interpreting the stylized facts as well as the counterfactual results.

Now consider what would happen if the mass of firms is endogenously determined by a free entry condition. Since firms have to incur an fixed entry cost $f_n^e > 0$ prior to learning their relative productivity φ, then the free entry condition requires that the expected value function is equal to the entry cost.

$$\int_{\underline{\varphi}}^{\infty} \left[V_n(\varphi, d_{-1} = 0; B_n)\right] dF(\varphi) = w_n f_n^e. \tag{2.14}$$

In the lower bound of the integral, $\underline{\varphi}$ denotes the productivity of the least productive active firm. Firms with productivity $\varphi < \underline{\varphi}$ exit upon observing their productivity level. The demand level B_n affects the value function through its impact on the current profit function. In the Appendix, one can appeal to monotone comparative statics arguments to prove Proposition 3 (see next):

Proposition 10. *Equation (2.14) delivers a unique market demand-level B_n in each country $n \in N$.*

Proposition (10) ensures the existence of a unique equilibrium in the manufacturing sector. Given a market demand B_n, exogenous parameters ξ_{nm}, f_{nm}, T_n and w_n, the firm-level combinatorial problem in (2.10) delivers a unique solution.

2.6 Structural Analysis

In this section, I use the firm-level data in conjunction with country-level data to estimate the main parameters of the model. The structural analysis is performed in five distinct steps. First, I use a simple linear regression to estimate each country m's sourcing potential ξ_{cm} from a Chinese importer's perspective. In the second step, I estimate the productivity dispersion parameter θ, by projecting the estimated sourcing potential values on observed cost shifters and other controls. I also measure the demand elasticity σ, from observed firm-level markups. In the third step, I solve for the sourcing strategy and estimate the fixed costs of sourcing and other distributional parameters via the method of simulated moments, when the intermediate goods are substitutable. In the fourth step, I apply equilibrium conditions to construct the measure of firm productivity and recover the parameters of the productivity evolution process. In the fifth and final step, I adopt nested fixed point technique developed in Rust (1987) to characterize firms' R&D decisions and back out cost parameters of innovation.

Because the customs data is from a single country, in what follows, I drop the subscript n from the notation, as China is the sole importing country throughout the estimation. Also, to facilitate the estimation, I include only those countries that exported to least 200 Chinese importing firms (out of 250,208 active Chinese non-processing firms [4]) in 2006. This criterion leaves me with a total of 32 sample countries, including China.

2.6.1 Estimation of a Country's Sourcing Potential

The first step is to estimate each country's sourcing potential, and I allow these sourcing potential terms vary over time. Assume what we observe in the data is a firm's optimal sourcing strategy. Taking the firm's $\mathbf{D}_t(\varphi)$ as a given, country m's sourcing potential ξ_{mt} in year t is measured as the average share of inputs sourced from m, relative to the share of domestically sourced inputs among all Chinese importers at time t. Besides, the *adjusted* sourcing potential $\hat{\xi}_{mt}$ is defined as a country m's sourcing potential, relative to China's sourcing potential in year t. Taking logs and normalizing the share of inputs purchased from country m in equation (2.4) by each firm's share of domestic inputs leads to:

$$\log \chi^x_{mt}(\varphi) - \log \chi^x_{ct}(\varphi) = \log \hat{\xi}_{mt} + \log \epsilon_{mt}(\varphi), \tag{2.15}$$

where

$$\log \hat{\xi}_{mt} \equiv \log \xi_{mt} - \log \xi_{ct}.$$

In Equation (2.15), firms are indexed by productivity φ and the origin country's notation c has been dropped. The domestic sourcing potential in 2006 is set to be one, i.e., $\xi_{c2006} = 1$, and the dependent variable of (2.15) is the difference between a firm's share of inputs sourced from country m and its share of inputs sourced domestically in year t. I measure these shares using data on a firm's total input expenditure, employee wages, and imports from each country. Since I include countries with at least 200 Chinese importers in 2006, firms' total input usage is adjusted by subtracting their imports from any of the excluded countries.

Intuitively, this specification allows us to identify a country's average sourcing potential by observing how much a firm imports from that country relative to the same firm's domestic input purchases, thus restricting attention to countries included in the firm's sourcing strategy. Hence, in order to guarantee that there is no selection bias in the regression, assume that firms only learn their country-specific efficiency shocks $\epsilon_{mt}(\varphi)$ after selecting the sourcing strategy; otherwise, assume the term $\epsilon_{mt}(\varphi)$ simply

[4] As processing-trade firms enjoy zero input tariff rate during the entire period, their innovation decisions should not be affected much by the input tariff reductions resulting from China's WTO accession. So, I first drop all processing firms to process the data before running the estimation. Also, I classify a firm as a processing-trade firm if its average ratio of processing exports over total exports is larger than 20 percent. The average processing export ratios are calculated based on the Customs Data during the entire periods.

represents the measurement errors. I estimate Equation (2.15) via ordinary least squares (OLS) and use country fixed effects to capture the $\hat{\xi}_{mt}$ terms. The estimated fixed effects represent each country's adjusted sourcing potential, normalized by China's sourcing potential in the same year, which is merely the average share difference by country.

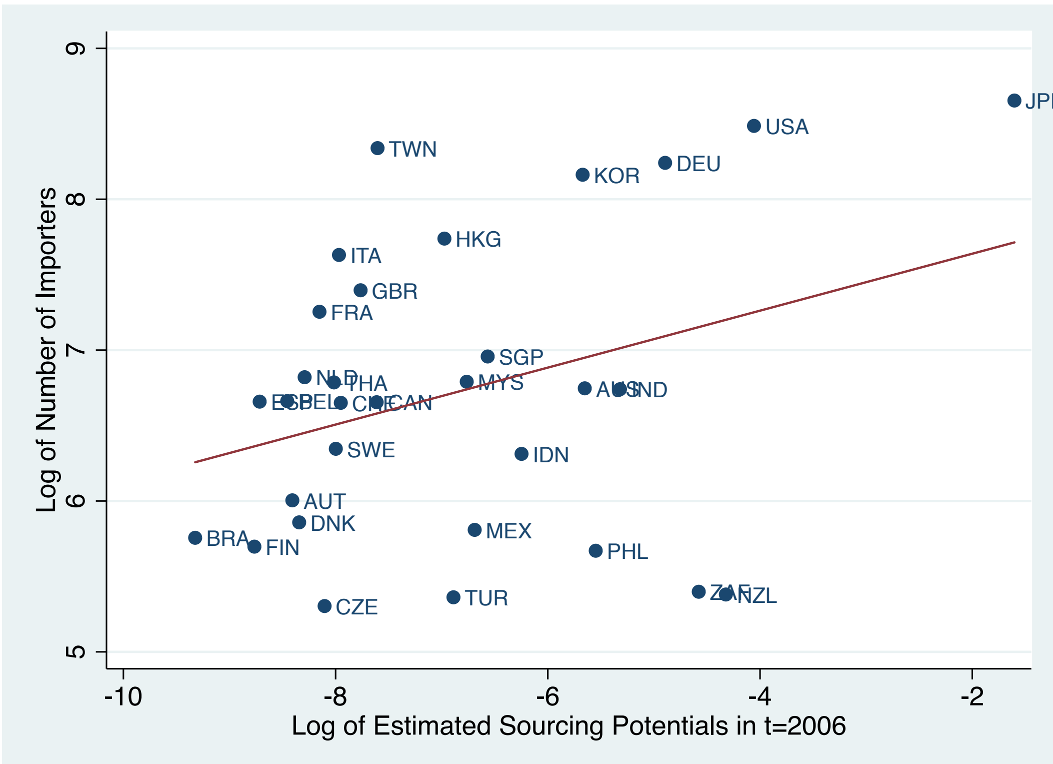

Figure 2.1: Country Sourcing Potential Parameters and the Extensive Margin

Figure 2.1 plots the estimated sourcing potential fixed effects against the number of firms importing from that country in 2006. The parameter estimates suggest that Japan has the highest sourcing potential for Chinese firms, followed by the USA and New Zealand. However, regarding the latter two, many more firms import from the USA than from New Zealand, suggesting that fixed costs of sourcing likely differ across source country characteristics like technology, transport costs etc.

Suppose a firm sources from all 32 countries and then that estimates tell us that the summation of all the foreign sourcing potential terms is 0.278 in 2006. Since the domestic sourcing potential in 2006 is normalized to be one, this result implies that the sourcing capability of a firm that purchases inputs from all 32 countries is 27.8 percent larger than that of a firm buying input only domestically. Interestingly, the summation of all the foreign sourcing potential terms, relative to domestic sourcing potential, is as small as

0.048 in 2001, which means for a firm sourcing worldwide, its sourcing capability is only 0.48 percent higher than that of a non-importing firm. Though the impact of a firm's sourcing capability on its marginal costs also depends on the firms' productivity φ, the dispersion parameter θ of the intermediates activities, and the elasticity of substitution σ, this comparison result suggests that sourcing becomes an increasingly crucial cost-saving channel after trade liberalization.

2.6.2 Estimation of the Elasticity of Demand and Input Productivity Dispersion

A key parameter of the model is demand elasticity σ of final goods. CES preferences, along with monopolistic competition imply that the ratio of sales to variable input costs is $\sigma/(\sigma-1)$. Thus, I calculate the average rate of sales to a sum of total expenditures on inputs and wages, using Chinese establishment-level data in 2006. I find that $\sigma/(\sigma-1) = 1.352$, which implies that $\sigma = 3.84$. Though it is infeasible to distinguish neatly between fixed and variable costs in the plant-level census data, 3.84 is a plausible estimate that is similar to previous findings. [5]

The other key parameter of the model is the dispersion of the productivity shocks of the intermediate inputs. The impact of a firm's sourcing capability Θ_j on its marginal costs decreases in the dispersion parameter, so a larger θ discourages sourcing and investment behavior in the model. Recall that sourcing potential ξ_{mt} is a function of a country's technology parameter, import tariff, iceberg trade costs, and wages, and that $\hat{\xi}_{m2006} = \xi_{m2006}$, as the domestic sourcing potential in 2006 is treated as a numeraire. Thus, I assume that θ is time-invariant and obtain a parameter value for θ by projecting the estimated sourcing potential on country characteristics such as R&D stock, estimated TFP, capital per worker, distance, a measure of control of corruption, wages, and common language. In particular, I estimate the following equation:

$$\begin{aligned}\log \xi_m =& \alpha_0 + \alpha_r \log \text{R\&D}_m + \alpha_p \text{TFP}_m + \alpha_k \log \text{capita/worker}_m - \theta \log \text{wage}_m \\ & - \theta(\log \alpha_c + \alpha_d \log \text{distance}_{cm} + \text{common language}_{cm} \log \alpha_l \\ & + \alpha_C \text{control corruption}_m) + \varsigma_m. \end{aligned} \tag{2.16}$$

In this equation, the national technology levels are proxied by country-level R&D spending, estimated TFP from the Penn World Table, and capital per employee; and the iceberg

[5] Broda and Weinstein (2006) estimate a mean elasticity of 4 at the SITC-3 level for 1990-2001. Antràs et al. (2017) estimate a similar median elasticity of 3.85.

trade costs are proxied by log distance, common language, and control of corruption.

Equation (2.16) shows that the parameter θ can be recovered from the estimated coefficient on wages. A potential concern with the use of country-level wage data is the fact that variation in wages partly reflects differences in worker productivity across countries. Hence, I follow Eaton and Kortum (2002) and adjust wages by worker quality, setting $\text{wage}_m = (\text{compensation}_m)e^{-gH_m}$, where H_m represents the average years of schooling for the population aged 15 and older in country m, and g represents the return to education estimates from Bils and Klenow (2000). Also, I use the total workforce and population density as instruments: given technology T_m, a country with more workers has a lower wage. Population density acts as a proxy (inversely) for productivity outside of manufacturing. Table 2.2 reports the results.

Table 2.2: Estimation of Firm and Aggregate Trade Elasticities

	Intermediate Inputs		All Inputs	
	log ξ	log imports	log ξ	log impots
log wage	-3.406***	-5.357***	-2.864**	-3.880***
	(1.318)	(1.286)	(1.369)	(1.055)
log R&D	0.391**	1.114***	0.385**	0.882***
	(0.154)	(0.160)	(0.152)	(0.124)
TFP	2.495*	3.273***	2.203	2.029*
	(1.297)	(1.267)	(1.395)	(1.079)
log capital/worker	2.110*	3.758***	1.637	2.670***
	(1.163)	(1.142)	(1.228)	(0.959)
log distance	-0.670	-0.185	-1.011*	-0.262
	(0.460)	(0.452)	(0.524)	(0.408)
control corruption	0.823	1.246**	0.813	1.023**
	(0.537)	(0.524)	(0.602)	(0.464)
common language	0.231	1.448	0.228	1.215
	(1.012)	(1.072)	(1.156)	(0.966)
constant	41.00***	61.43***	35.76**	47.80***
	(15.86)	(15.51)	(17.23)	(13.27)
observations	53	53	55	55
R-squared	0.118	0.325	0.178	0.433

Note: Standard errors in parentheses. The adjusted wage is instrumented by the total workforce and population density.

All imports are classified into three categories, according to the classification of Broad Economic Categories (BEC): capital goods, intermediate goods, and consumption goods. The dependent variables in the first two columns of Table 2.2 are backed out from firms' intermediate imports only, and the last two columns use information on both capital imports and intermediate imports. The results of Columns 3 and 4 exhibit a much larger dispersion in productivities across countries than the estimates in Columns 1 and 2, thus suggesting that capital goods producers are more responsive to changes in marginal costs and are more likely to participate in outsourcing and R&D activities. Besides, the data on firm-level trade flows display a larger dispersion in productivity than what is obtained with the aggregate trade flow, which is consistent with the estimate results in Antràs, Fort, and Tintelnot (2017). I treat my Column 1 coefficient, $\theta = 3.406$, as the benchmark estimate of θ for two reasons: first, the objective of this regression is sourcing patterns of intermediate input producers worldwide; second, after accounting for the endogeneity of wages, Eaton and Kortum (2002) suggest an estimate of θ from 2.86 to 3.60, and my estimate falls within this range.

With estimates for country sourcing potential, firm-level trade elasticity, and elasticity of demand, I can calculate how global sourcing affects costs and market size. $\theta = 3.406$ means that a firm enjoyed around 6.9 percent $(1.278^{-1/3.406})$ lower input costs and 18.5 percent $(1.278^{-2.84/3.406})$ more sales if it sourced from all countries rather than sourcing domestically in 2006. On the contrary, using the estimates for country sourcing potential in 2000, a firm benefited from around 1.4 percent lower input costs and 3.8 percent more sales if it sourced from all countries instead of sourcing domestically in 2000. Hence, the impact of outsourcing on sales is increasingly significant after trade liberalization.

2.6.3 Estimation of Fixed Costs of Outsourcing

In this section, I estimate the fixed costs of outsourcing via the method of simulated moments (MSM). First, I explain how I simulate a set of artificial Chinese importers. Next, I describe how I empirically recover firms' equilibrium behavior given a particular set of parameter values, with each firm assigned a productivity draw $\varphi(s)$ and a fixed sourcing cost draw $f_n(s)$ in each market. Third, I discuss how I calculate a set of moments from these artificial data to compare with moments from the actual data. Finally, I explain the estimation procedure and report the results.

2.6.3.1 Simulation

Denote an artificial Chinese importer by s and the number of simulated importers by S. The number S does not bear any relationship to the number of firms in real data. A massive S means less sampling variation in simulations. As I search over different parameters Θ, I want to hold fixed the realizations of the stochastic components of the model. Hence, prior to running simulations, I (i) draw S realizations of $u(s)$ independently from the uniform distribution $U[0,1]$, and (ii) draw $S \times N$ realizations of $\eta_m(s)$ independently from the standard normal distribution $N(0,1)$, where N represents the number of importing countries. Each firm is assigned with an exogenous productivity shock $u(s)$ and a set of exogenous fixed sourcing cost shocks $\eta_m(s)$ in each market m.

Recall that I assume that a firm's productivity $\varphi(s)$ is drawn from a Pareto distribution characterized by θ and with support on $[\underline{\varphi}, \infty]$. Hence, I construct: [6]

$$\varphi(s) = \underline{\varphi}(1 - u(s))^{-\frac{1}{\theta}},$$

using the $u(s)$'s that were drawn prior to the simulation, and $\underline{\varphi} = \frac{1}{\sum_s (1-u(s))^{-\frac{1}{\theta}}}$.

Next, I construct firm-country pairwise fixed sourcing costs by assuming that the fixed sourcing cost of country m for a firm s is:

$$\log f_m(s) = \log \beta_c + \beta_d \log \text{dist}_{cm} + \text{lang}_m \log \beta_l + \text{control corrup}_m \log \beta_C + \beta_{disp} \eta_m(s),$$

where β_{disp} is a dispersion parameter of interest, and $\eta_m(s)$'s were drawn earlier from a standard normal distribution. Moreover, since all firms use domestic inputs, I set fixed costs of domestic sourcing equal to zero; i.e., $f_c(s) = 0, \ \forall s$.

2.6.3.2 Estimation Algorithm

The purpose of the simulation is to characterize a firm's equilibrium expenditure on self-made inputs, outsourcing inputs from all countries, and total sales given a set of common parameters $\Omega = \{B, \beta_c, \beta_d, \beta_l, \beta_C, \beta_{disp}, \delta\}$, a firm-specific productivity draw $\varphi(s)$, and a firm-country pairwise fixed sourcing cost draw $f_m(s)$ in market $m \in N$. Given that a firm faces idiosyncratic fixed costs in each source market it purchases from, and given that inputs are substitutes in production, this is a complicated computational problem. I handle it by adapting the algorithm first introduced in Jia (2008) and exploiting the decreasing property of a function V^m defined in Equation (2.12).

[6] It is straitforward to show that $\varphi(s)$ is drawn from a Pareto distribution with $u(s)$ following a uniform distribution $U[0,1]$, that is,

$$\Pr(\varphi(s) \leq \varphi | \underline{\varphi}) = \Pr\left[\underline{\varphi}(1 - u(s))^{-\frac{1}{\theta}} \leq \varphi\right] = \Pr\left[u(s) \leq 1 - (\varphi/\underline{\varphi})^{-\theta}\right] = 1 - (\varphi/\underline{\varphi})^{-\theta}.$$

Moreover, for any $u(s) \in (0,1]$, we have

$$\varphi(s) = \underline{\varphi}(1 - u(s))^{-\frac{1}{\theta}} \geq \underline{\varphi}.$$

Step 1: Iteratively apply the V^m to find the optimal sourcing strategy $\mathbf{D}_*(s)$. Please see the computation Appendix for details.

Step 2: Calculate total sales as:

$$Y(s) = B\Big[\varphi(s)\delta^{\theta} + \sum_{m=1}^{N} D_*^m(s)\xi_m\Big]^{\frac{\sigma-1}{\theta}}.$$

Step 3: Recover the share of intermediates purchased from country m as:

$$\chi_m^x(s) = \frac{\xi_m}{\varphi(s)\delta^{\theta} + \sum_{m=1}^{N} D_*^m(s)\xi_m}, \text{ if } D_*^m(s) = 1,$$

and $\chi_m^x = 0$ otherwise. Similarly, the share of in-house produced inputs is

$$\chi^l(s) = \frac{\varphi(s)\delta^{\theta}}{\varphi(s)\delta^{\theta} + \sum_{m=1}^{N} D_*^m(s)\xi_m}.$$

Analogously, the expenditure on inputs sourced from country m is

$$M_m^x(s) = \frac{\sigma-1}{\sigma}\chi_m^x(s)Y(s),$$

and firm s spends $M^l(s)$ on self-made inputs with

$$M^l(s) = \frac{\sigma-1}{\sigma}\chi^l(s)Y(s).$$

Intuitively, searching for an optimal sourcing strategy starts from $\mathbf{D}_0(s) = \{0, ..., 0\}$ containing no foreign country. In the first loop, we apply the W-operator to the $\mathbf{D}_0(s)$, asking if firm s is allowed to add one country to the $\mathbf{D}_0(s)$, add whichever countries can improve the profits. Under the empirically relevant condition $\sigma - 1 < \theta$, Proposition 8 implies that, if marginal gain to source from a country m (compared to using self-made inputs) is even lower than the fixed cost of outsourcing to the country m, then firm s does not outsource to m in the optimal sourcing strategy. Hence, $\mathbb{D}_1(s)$ is a set of sourcing strategy from which firm s sources from only one country. In the second loop, I add one country at a time to each sourcing strategy in $\mathbb{D}_1(s)$ and answer the question of whether outsourcing to one more country increases firm s's profits, derived from the current sourcing strategy in $\mathbb{D}_1(s)$. I then stack strategies such that outsourcing to two countries is better than outsourcing to either of them alone, and construct a new set $\mathbb{D}_2(s)$. Similarly, the set $\mathbb{D}_3(s)$ includes a list of strategies, in each of which outsourcing to three selected countries is better than outsourcing to any two of them. I then apply the W-operator iteratively until firm s has no incentives to add further locations to (any of) the sourcing strategies in $\mathbb{D}_T(s)$. Note that an iterative application of the W-operator leads to a reduced set instead of a fixed point, so I must evaluate all the strategies in $\{D_0(s), \mathbb{D}_1(s), ..., \mathbb{D}_T(s)\}$ so as to locate the optimal $\mathbf{D}_*(s)$. For example, firm s finds it unprofitable to outsource to any other countries after it pays fixed outsourcing costs to the US. In the meantime, although firm s has already paid the fixed costs to Japan, it still finds it optimal to add Korea to its sourcing strategy. As the firm's profit does not increase in a straightforward fashion in the number of sourcing countries, I need to evaluate both strategies to figure out firm s's optimal one.

This procedure characterizes the behavior of S artificial Chinese importers. We know four things about each firm: where it imports $\mathbf{D}^*(s)$; input cost share $\chi(s)$; input expenditure $M(s)$; and sales $Y(s)$. From these terms, I can compute any moments that could have been constructed from the actual Chinese data.

2.6.3.3 Moments

For a candidate Ω, I use the algorithm above to simulate the behavior of S artificial Chinese firms sourcing from N potential markets. I compute a vector of $\hat{m}(\Omega)$ from these artificial data, analogous to particular moments m in the actual data.

The first set of moments includes: (i) the share of importers for all manufacturing firms, and (ii) the share of importers with firm sales below the median. This is a 2×1 vector, which I label as $m1$ in the actual data and as $\hat{m}_1(\Omega)$ for the simulated data. The second set of moments is the share of firms that imports from each country, excluding the home country. I label this $(N-1) \times 1$ vector of moments in the data as m_2 and the simulated moment vector as $\hat{m}_2(\Omega)$. The third set of moments includes: (i) the median cost share of self-made inputs; and (ii) the median cost share of domestic inputs. This is a 2×1 vector, for which I label the simulated moments as $\hat{m}_3(\Omega)$ and the actual moments as m_3. The last moment is the median sales among all manufacturing firms. This is a scalar variable, for which I label the simulated moment as $\hat{m}_4(\Omega)$ and the actual median sales as m_4.

The first set of moments is informative about the dispersion parameter as an increase in the dispersion of fixed cost draws will lead to a greater number of importers whose firm sales are below the median. The second set of moments are informative about how the overall magnitude of the fixed outsourcing costs varies with distance, common language, and control of corruption. The third sets of moments is indicative of the contracting friction parameter δ. If there is no contracting friction, the share of self-made inputs relative to the share of domestic inputs will go to zero. Similarly, if δ goes to infinity, then firms use in-house produced inputs only. Finally, the fourth moment helps pin down the scale parameter B, as B governs the level of sales and input purchases.

Stacking the four sets of moments gives me a 36-element vector of deviations between the moments of actual and artificial data:

$$f(\Omega) = m - \hat{m}(\Omega) = \begin{bmatrix} m_1 - \hat{m}_1(\Omega) \\ m_2 - \hat{m}_2(\Omega) \\ m_3 - \hat{m}_3(\Omega) \\ m_4 - \hat{m}_4(\Omega) \end{bmatrix}.$$

Assume the following moment condition is hold at the true parameter value Ω_0

$$\mathbf{E}[f(\Omega_0)] = 0,$$

and thus we seek a $\hat{\Omega}$ that achieves

$$\hat{\Omega} = \arg\min_{\Omega}\{f(\Omega)^T \mathbf{W} f(\Omega)\},$$

where $\mathbf{W}$ is a 36×36 identity matrix. I use the identity matrix, instead of the inverse of the estimated variance-covariance matrix of the 36 moments, as the weighting matrix. The identity matrix yields a better fit of the import shares of the most popular countries - especially Japan - which are most relevant for the counterfactual analysis later at the expense of a worse fit of the shares of less popular importing countries.

The parameters estimated are displayed in Table 2.3. The fixed costs of sourcing are 1.52 percent ($1.1^{0.158}$) higher when the distance increases by 10 percent, and sourcing from countries with a common language reduces fixed costs by 2.7 percent. A country with a better institution, measured by control of corruption, has lower fixed sourcing costs. Furthermore, the median fixed cost estimate ranges from US$166,000 - US$343,000. Analogously, the estimated fixed cost range is from US$10,000 - US$56,000 from a US importer's perspective in Antràs, Fort, and Tintelnot (2017). Hence, higher fixed costs of outsourcing partially explained the fact that only about 6.7 percent of Chinese firms imported foreign inputs in 2006; while as much as 25.8 percent of US firms outsourced their inputs overseas in 2007.

Table 2.3: Fixed Cost Parameter Estimates

Variable	Description	Coefficient
B	demand level	0.194
		(0.000)
β_c	constant	0.158
		(0.010)
β_d	log distance	0.109
		(0.021)
β_l	language	0.973
		(0.005)
β_C	control of corruption	-0.172
		(0.004)
β_{disp}	dispersion	0.858
		(0.023)
δ	contract friction	18.621
		(0.020)

Notes: Table reports coefficients and standard errors from estimating the model via simulated method of moments. Standard errors based on 25 bootstrap samples drawn with replacement.

2.6.3.4 Estimation of Productivity Evolution Parameters

I begin the dynamic estimation by measuring firms' productivity. Assume the demand elasticity σ, trade elasticity θ, contracting friction δ, and fixed sourcing cost $\{\beta_c, \beta_d, \beta_l, \beta_C, \beta_{disp}\}$ to be time-invariant. Let the demand-level parameter B, firm's relative productivity φ, country-level sourcing potential ξ, and firm's equilibrium choices vary over time and be indexed by year t from now on. Recall that in equilibrium, firm φ_t's cost share on domestic inputs in period t is:

$$\chi^x_{ct}(\varphi_t) = \frac{\xi_{ct}}{\varphi_t \delta^\theta + \sum_{m=1}^N D^m_t(\varphi_t)\xi_{mt}},$$

where the sourcing indicator $D^m_t(\varphi_t)$ is equal to one if firm φ_t purchases inputs from country m in year t, and it is equal to zero otherwise. The variables ξ_{mt} captures the country m's sourcing potential in year t, and it is mean ratio of firms' cost share on inputs sourcing from country m to cost share on domestic inputs in year t among all manufacturing firms. The parameters θ, δ, and $\{\xi_{mt}\}^N_{m=1}$ are estimated from previous steps. Extracting firm φ_t's domestic cost share $\chi^x_{ct}(\varphi_t)$ and its sourcing portfolio $\{D^m_t(\varphi_t)\}^N_{m=1}$

Table 2.4: Productivity Evolution Parameters

φ_t	Balanced Panel	Full Sample
φ_{t-1}	0.882***	0.850***
	(0.00571)	(0.00361)
φ^2_{t-1}	-7,660***	-7,095***
	(233.1)	(144.3)
d_{t-1}	1.47e-08**	5.99e-08**
	(6.78e-09)	(2.76e-08)
Constant	6.68e-07***	8.19e-07***
	(1.72e-08)	(1.03e-08)
SE(ϵ)	3.877e-6	4.142e-6
Observations	254,000	731,874
R-squared	0.537	0.485

directly from the data, I can recover a firm's productivity in year t as:

$$\varphi_t = \delta^{-\theta}\left(\frac{1}{\chi^x_{ct}(\varphi_t)} - \sum_{m=1}^{N} D^m_t(\varphi_t)\xi_{mt}\right). \tag{2.17}$$

Next, substitute Equation (2.17) into the productivity evolution Equation (2.2) to get a set of parameters $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$.

Table 2.4 reports the estimates of the productivity evolution process from Equation (2.17). The coefficients in Column 1 are derived from a balanced data panel between 2001 and 2006. The coefficients α_1 and α_2 measure the effect of lagged productivity on current productivity. They imply a clear significant nonlinear relationship between φ_{t-1} and φ_t. The coefficient α_3, then, measure the effect of the lagged discrete R&D investment on current productivity, and it is positive and significant. Plants that engaged in R&D investment have $1.47e^{-8}$ higher productivity than those who didn't in the preceding period. $1.47e^{-8}$ seems reasonable in magnitude, given that, in the panel data, the range of estimated productivity is from $9.64e^{-10}$ to $6.61e^{-5}$, and the median productivity is $3.62e^{-6}$.

Estimates in Column 2 are based upon an imbalanced full sample in the same period of 2001 - 2006, and I find that the estimate of α_3 in Column 2 is higher than the one in Column 1. Firms in the panel data are, on average, more productive than those in the full sample, as the former had already stayed in the market for at least six years. The higher estimate of α_3 in Column 2 somehow indicates that it takes more efforts for a productive firm to improve its productivity than it takes a less productive firm. My estimation on innovation cost parameters uses the panel data, so I treat Column 1 coefficients as the benchmark estimates.

2.6.3.5 Estimation of Innovation Cost Parameters

Estimating the innovation cost parameters is based on the likelihood function for the observed discrete patterns of firm patent applications d_t. Regardless of patent publication date, denote $d_t = 1$ if a firm submits a new patent application in year t. Assume that firm φ_t draws its sunk and fixed innovation costs from two separate, independent, exponential distributions with the mean (γ^F, γ^S). Then the firm compares the current-period cost of R&D with the increase in future value that it expects R&D might bring. Its conditional probability of investing in R&D is equal to:

$$P(d_t = 1|s_t) = P\big(d_{t-1}\gamma^F + (1 - d_{t-1})\gamma_t^S \leq \Delta E_t V(\varphi_t)\big) \quad (2.18)$$

where the term $\Delta E_t V(\varphi_t)$ is defined in Equation (2.13). I apply the nested fixed-point algorithm developed by Rust (1987) to solve for the value function $\Delta E_t V(\varphi_t)$. First, the state space $s_t = (\varphi_t, d_{t-1})$ is discretized into 1000 grid points for productivity and two values for lagged R&D choice, and I use value function iteration to solve for the value function at each element of this discretized state space. Then I apply a cubic spline to interpolate the firm value and payoff to R&D across actual data space: there are 200,028 firms in total. The objective of the estimation is to maximize the likelihood function using the observed patterns of firm patent application dummy d_t:

$$\mathcal{L}(\gamma) = \int_s \big[d_t P(d_t = 1|s_t) + (1 - d_t)P(d_t = 0|s_t)\big] ds.$$

The vectors d and s contain every firm's observed R&D choice and state variables for each period. The conditional choice probabilities is given in equation (2.18).

Table 2.5: Innovation Cost Parameter Estimates

	Mean	Std Errors
γ^F (R&D FC)	1,364	15.050
γ^S (R&D SC)	27,384	32.608

Notes: The parameters are in USD, and standard errors are derived from the Hessian matrix.

The results are reported in Table 2.5. The estimated mean of fixed cost is around US$1,364, and the average sunk innovation cost is about US$27,385. That is, the estimated fixed cost parameter is much lower than the sunk cost parameter, indicating that the startup cost is much more substantial than the per-period costs of maintaining investment. These estimates are consistent with the fact that, among firms that do not conduct R&D in the year $t-1$, only 3.06 percent of them suddenly start investing in R&D in the year t. In the meantime, conditional on investing in R&D in $t-1$, around 47.76 percent of firms continue R&D in the year t.

2.6.3.6 Fit of the Model

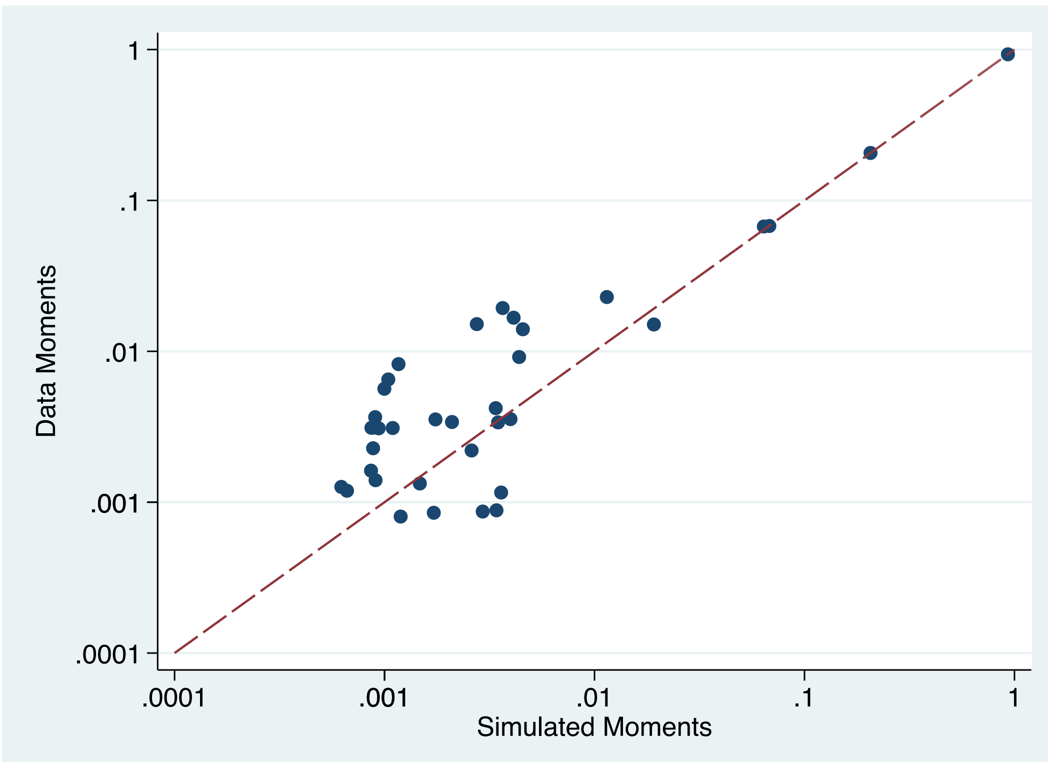

Figure 2.2: Model Fit - Estimation on Fixed Outsourcing Cost Parameters

Overall, the model fits the data reasonably well. I begin with the estimation of fixed outsourcing cost parameters via the simulated method of moments in the third step. The correlation between the predicted moments and the actual moments is 0.99. Figure 2.2 depicts this relationship. The model does a good job of matching the first, third, and fourth sets of moments: around 6.7 percent in the data and 6.4 percent of Chinese firms in the model import foreign inputs. For firms with sales below the median, around 1.5 percent in the data and 1.9 percent in the prediction do so. Median sales are approximately equal (US$2,586,776 in the data; US$2,587,756 in the model). Moreover, the self-made input share is around 6.76 percent in the data and 6.80 percent in the prediction, and the ratio of domestic inputs to the total input expenditure is about 9.29 percent in the data and 9.31 percent in the parameterized model. However, the model underpredicts the share of firms importing from countries that are far away from China. Take the US as an example: the distance between the US and China is 9.57 times as large as the distance between Korea and China. Since the fixed costs are increasing with distance with an elasticity of 0.158, fixed costs of outsourcing to the US are on average 142.88 percent ($9.57^{0.158}$) higher than fixed costs of outsourcing to Korea, holding other country characteristics constant. Even though the control of corruption and sourcing potential in Korea are slightly lower than their US equivalent terms, firms still prefer to outsource to Korea in the simulation. The variation in distance among all sample countries is more spread than the difference in other country controls, so the distance between a potential outsourcing country and China plays a dominant role in the predicted model.

Table 2.6: Predicted Transition Rates for Continuing Plants

		$d_t = 0$	$d_t = 1$
$d_{t-1} = 0$	Predicted	96.92%	3.08%
	Actual	96.94%	3.06%
$d_{t-1} = 1$	Predicted	52.76%	47.24%
	Actual	52.24%	47.76%

In the dynamic estimation of innovation parameters, there are around 5.2 percent of firms investing in knowledge (5.23 in the data and 5.22 in the baseline model). Moreover, the model also does an excellent job at matching the transition patterns I did not target directly in the estimation. Table 2.6 shows that if a plant does not conduct innovation activities in year $t-1$, it has a probability of about 3 percent (3.08 percent in the data and 3.06 percent in the model) of investing in knowledge in year t, which is lower than the 47 percent probability (47.24 percent in the data and 47.46 percent in the simulation) that a plant investing during year $t-1$ will continuously conduct R&D in year t.

2.7 Counterfactual

In this section, I conduct counterfactual exercises quantifying the impact of a reduction of trade barriers between China and Japan on domestic plants' innovation participation and value function, and the effect this reduction has on the plants' productivity trajectory. I focus on Japan, specifically, because it is the largest supplier of imported inputs for Chinese firms.

2.7.1 Baseline Predictions

Table 2.7: Firm Response to Exogenous Trade Shocks in $t = 20$

	$\tau_{jpn} = 0$	$\mathrm{fc}_{jpn} = 0$	90% dec in fc_{jpn}
Change in prop of R&D performers	0.9835	0.8117	0.9701
Change in mean value	1.0017	1.0497	1.0035
Change in prop of importers	1.0938	10.4167	1.8125
Change in prop of importers from JPN	1.1897	17.2414	2.4138

I use the parameter estimates from Section 6 to assess how plant-level innovation decisions, import decisions, and value functions respond to a reduction of trade barriers between Japan and China. These exercises hold the following three circumstances respectively: (i) import tariffs from Japan to China is zero; (ii) the median fixed cost of outsourcing to Japan from China is zero; and (iii) the median fixed cost of outsourcing to Japan decreases by 90 percent. Throughout the counterfactual exercises, assume aggregate variables (including sourcing potential, and median of fixed outsourcing costs) are constant, both at home and abroad. A firm knows that its innovation decision affects its own productivity evolution endogenously, but this decision simultaneously has no impact on (country-level) aggregate variables. I allow the demand-level B to be endogenous, such that the free-entry condition is satisfied, and I solve for the optimal innovation decisions for a set of 250,508 simulated firms for 20 years. I focus on results 20 years later to reflect the fact that the negative impact(s) of outsourcing may accrue over time in a more obscure form of slow productivity growth. A firm's R&D decision today not only affects it future productivity; it also determines whether the firm will undertake R&D or not tomorrow. For example, the analysis shows that permanently eliminating all tariffs on imports from Japan decreases R&D participation rates by 0.57 percent today, but that this decision to not invest in R&D also correlates with a reduction in the share of R&D performers by 1.65 percent even as far into the future as twenty years later (as the culmination of changes that have been growing over the years).

Table 2.7 reports the change in the proportion of R&D performers, mean value function, share of importers overall, and the share of importers outsourcing to Japan, relative to the predictions without trade shocks in year 20. After 20 years, a permanent zero import tariff between Japan and China decreases the R&D participation rate by 1.5 percent and increases the average value function by 0.13 percent. Meanwhile, a 90 percent decrease in the median fixed cost of outsourcing to Japan decreases the R&D participation rate by 3 percent and increases the average value function by 0.35 percent. In an extreme situation in which Chinese firms no longer pay any fixed outsourcing costs to Japan imports and the import tariff level remains the same as 7.44%, the R&D participation rate is decreased by 19 percent, and the average value function is increased by 4.97 percent twenty years later.

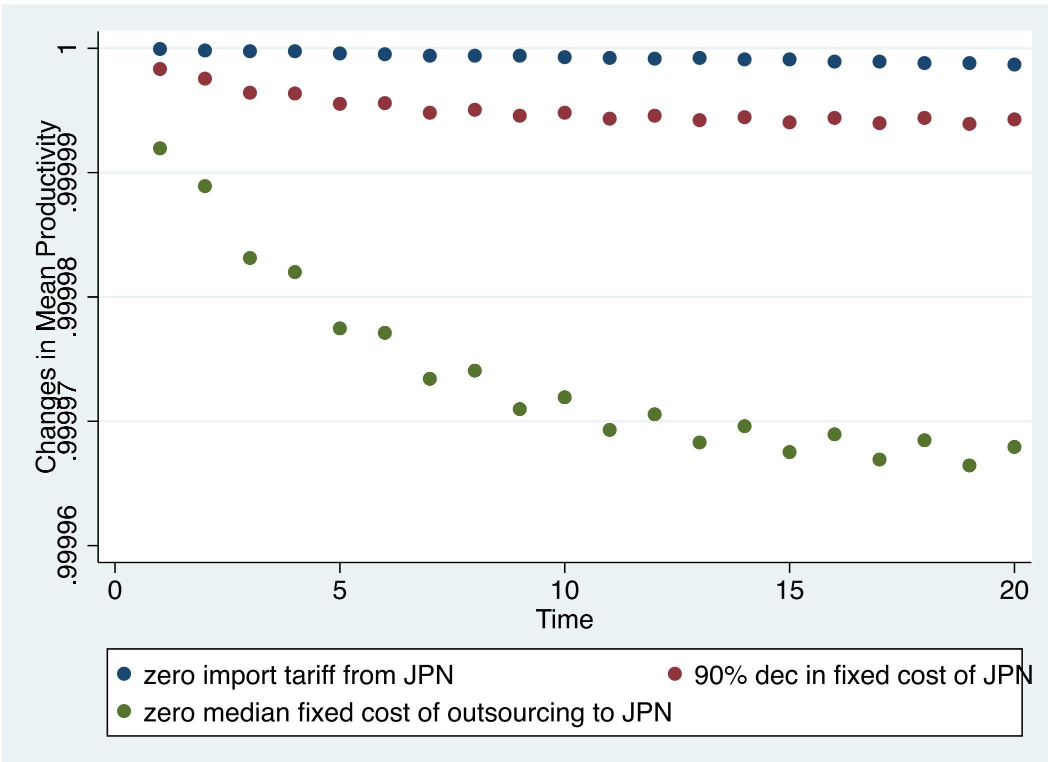

Figure 2.3: Productivity Trajectory for 20 years

Figure 2.3 plots the transition dynamics of average productivities under three circumstances: the trends of average productivity resulting in trade liberalization are all decreasing, the average productivity is the lowest in case of the zero fixed costs of outsourcing to Japan. The negative impact of trade liberalization on productivity increases in the share of importers as well as the proportion of importers sourcing from Japan. Though all firms' profits are affected by trade shocks through changes of the aggregate price index (which is an increasing function of the demand-level B), importers are more responsive to changes in exogenous trade variable costs because trade shocks affect their production costs directly. When the fixed outsourcing cost is zero, all firms are importers, as they import at least from one country - Japan. In fact, in a world without any trade frictions and fixed outsourcing costs, a firm outsources all intermediates to the lowest-input suppliers and use virtually no self-made inputs; hence, the probability of investing in in-house production technology goes to zero. On the contrary, when the fixed outsourcing cost is decreased by 90 percent, only 8.4 percent of firms are importers and about 5.1 percent of firms outsource inputs to Japan. Correspondingly, the impact of a 90 percent reduction in the median fixed outsourcing cost to Japan on firms' innovation participation rate and average value function is much smaller than an outsourcing cost fixed permanently at zero.

2.8 Conclusion

I develop a dynamic structural model of a producer's decision to invest in knowledge and to outsource intermediate input to multiple countries, allowing both choices to affect the production cost endogenously. Using Chinese plant-level data from 2001 - 2006, which feature a drastic input tariff cut in 2002 due to China's WTO accession, I find that trade liberalization results in higher revenue but less innovation among Chinese firms. The driving force underlying the negative effect of trade shocks on firm innovation is that intermediate imports substitute for self-made inputs, a strategy that decreases marginal production costs but also lowers gains that would otherwise arise from investing in in-house production technology. This argument is compatible with results obtained in other studies on the effects that input-tariff reductions have on firms' innovation participation; for example, see Liu et al. (2016).

The framework can be extended to include the intensive margin of innovation decisions. Define the R&D concentration measure as a sum of squaring the patent share of each firm in a market with a range from 0 - 10,000. I find that the R&D concentration measure in China increases from 56.205 in 2001 to 133.604 in 2006 in China. A reduction of import tariffs induces some less-productive firms to use foreign inputs as a substitute for self-made inputs and undertake less innovation, so I observe the R&D participation rate increases through import tariffs. However, in the mean time, trade liberalization pushes highly productive firms to invest even more in R&D, as they now serve more foreign markets. Therefore, taking the intensive margin adjustment of R&D decisions into account helps us to understand how trade liberalization affects productivity growth in the long run.

Appendix A

Supplemental Material for Chapter 1

A.1 Proofs

Proof of Lemma 1. See Appendix B in AGR working paper (2005). ■

Proof of Lemma 2. First, perfect mobility of workers in their own countries implies that

$$(1 - t_i^W) w_i^S(z) = (1 - t_i^W) w_i^N(z) \text{ for } i \in \{S, N\}, z \in [a, b], a \leq b$$

or

$$w_i^S(z) = w_i^N(z) \text{ for } i \in \{S, N\}, z \in [a, b], a \leq b$$

that is, workers' wages are independent of their matched managers' nationality. If not, WLOG, suppose $w_S^S(z_p) > w_N^S(z_p)$ for a particular $z_p \in [a, b]$, then $m_N^S(z_p)$ is not real-valued on $[a, b]$ as no workers are willing to work for managers from North with ability $m_N^S(z_p)$.

Next to show that

$$m_i^S(z) = m_i^N(z) \text{ for } i \in \{S, N\}, z \in [a, b], a \leq b$$

WLOG, suppose by contradiction that there exists a $z_p \in (a, b)$ such that $m_N^N(z_p) \neq m_N^S(z_p)$ where both $m_N^N(z_p)$ and $m_N^S(z_p)$ are real-valued on $[a, b]$, that is, managers from both South and North establish firms in North in the equilibrium.

Manager's rent-maximization problem yields:

$$w_N^{N'}(z_p) = \frac{m_N^N(z_p) - w_N^N(z_p)}{1 - z_p}$$

$$w_N^{S'}(z_p) = \frac{m_N^S(z_p) - w_N^S(z_p)}{1 - z_p}$$

By definition of differentiability,

$$\begin{aligned} w_N^{N'}(z_p) &= \lim_{z \to z_p} \frac{w_N^N(z_p) - w_N^N(z)}{z_p - z} \\ &= \lim_{z \to z_p} \frac{w_N^S(z_p) - w_N^S(z)}{z_p - z} \\ &= w_N^{S'}(z_p) \end{aligned}$$

where the second equality comes from the perfect labor mobility assumption.

Then

$$\begin{aligned} & w_N^{N'}(z_p) = w_N^{S'}(z_p) \\ \Leftrightarrow\ & \frac{m_N^N(z_p) - w_N^N(z_p)}{h(1 - z_p)} = \frac{m_N^S(z_p) - w_N^S(z_p)}{h(1 - z_p)} \\ \Leftrightarrow\ & m_N^N(z_p) = m_N^S(z_p) \end{aligned}$$

Symmetrically, we have

$$m_S^N(z_p) = m_S^S(z_p)$$

■

Proof of Lemma 3. See the proof of Theorem 1 in AGR. ■

Proof of Lemma 4. Denote M_i as the set of managers forming team in country i. WLOG, suppose by contradiction that $M_S = [a, b] \cup [c, d]$ with $b < c$. Managers' market clearing condition implies that $[b, c] \in M_N$.

Workers' market clearing condition implies that $m_S^{-1}(b) = m_S^{-1}(c) \equiv z_1$. Managers' before-tax rents are given by

$$r_S(b) = \frac{b - w_S(z_1)}{h(1 - z_1)}$$

$$r_S(c) = \frac{c - w_S(z_1)}{h(1 - z_1)}$$

where $r_S(z)$ denotes managers' possible rents if build up teams in S. Denote $e = \frac{1}{2}b + \frac{1}{2}c \in M_N$. Positive sorting means

$$r_S(e) = \frac{e - w_S(z_1)}{h(1-z_1)} = \frac{1}{2}r_S(b) + \frac{1}{2}r_S(c)$$

because a manager with skill levels e can hire workers with ability no higher than $m_S^{-1}(c)$ and no lower than $m_S^{-1}(b)$ in South.

On the other hand, $b, c \in M_S \cap M_N$ imply that $r_S(b) = r_N(b)$ and $r_S(c) = r_S(b)$ in an equilibrium. Lemma 3 says that r_N is strictly convex, that is,

$$r_N(\frac{1}{2}b + \frac{1}{2}c) \leq \frac{1}{2}r_N(b) + \frac{1}{2}r_N(c) = \frac{1}{2}r_N(b) + \frac{1}{2}r_N(c) = R_S(\frac{1}{2}b + \frac{1}{2}c)$$

In other words, $e \in M_S$ instead. Contradiction. ■

Proof of Lemma 5 (1). $M_S \cap M_N = [a, b]$ means that for any manager with ability $c \in [a, b]$, they receive the same after-tax payoff building up teams in two countries, that is,

$$(1 - t_S^M) r_S(c) = (1 - t_N^M) r_N(c)$$

Lemma 3 implies that $r_S(.)$ and $r_N(.)$ are continuous functions, so both functions are differentiable for any interior points. Then rewrite

$$r_S'(c) = \frac{r_S(c) - r_S(c-\epsilon)}{\epsilon} = \frac{1 - t_N^M}{1 - t_S^M} \frac{r_N(c) - r_N(c-\epsilon)}{\epsilon} = \frac{1 - t_N^M}{1 - t_S^M} r_N'(c)$$

Using the envelope condition, the marginal return to skill for managers is given by

$$r_i'(z) = \frac{1}{h(1 - m_i^{-1}(z))}$$

Thus,

$$\frac{1 - t_N^M}{1 - t_S^M} = \frac{r_S'(c)}{r_N'(c)} = \frac{1 - m_N^{-1}(c)}{1 - m_S^{-1}(c)}$$

Since $t_S^M, t_N^M, m_N^{-1}, m_S^{-1} \in (0, 1)$, we have

$$\frac{t_N^M}{t_S^M} = \frac{m_N^{-1}(c)}{m_S^{-1}(c)}$$

where denote $M \equiv \frac{t_N^M}{t_S^M}$ which is a constant.

In an equilibrium, $m_i^{-1}(z)$ can be expressed in a form such that $m_i^{-1}(z) = \sqrt{h^2 + 2hk(c-z)}$ where k and c are some constants inherited from the matching function $m(.)$ and independent of z. Thus, we can rewrite the above function as

$$M \equiv \frac{t_N^M}{t_S^M} = \frac{m_N^{-1}(c)}{m_S^{-1}(c)} = \frac{\sqrt{h^2 + 2hk_N(c_N - z)}}{\sqrt{h^2 + 2hk_S(c_S - z)}}, \text{ for } z \in (a,b)$$

Solve the equation to get $k_N = k_S$, $c_N = c_S$ and $M \equiv \frac{t_N^M}{t_S^M} = 1$, that is, $t_N^M = t_S^M$. ■

Proof of Lemma 5 (2). Since $t_N^M = t_S^M$ and $(1 - t_S^M)r_S(c) = (1 - t_N^M)r_N(c)$, we get

$$r_S(c) = r_N(c), \text{ for any } c \in (a,b)$$

Since $r_i(c) = \frac{c - w_i(m_i^{-1}(c))}{h(1 - m_i^{-1}(c))}$ and $m_N^{-1}(c) = m_S^{-1}(c)$, we have

$$w_S(m_S^{-1}(c)) = w_N(m_N^{-1}(c))$$

■

Proof of Lemma 5 (3). Denote $\underline{z} \equiv \min\{z_S^*, z_N^*\}$. Positive sorting and labor market clearing condition imply that $m_S(0) = \underline{z}$ or/and $m_N(0) = \underline{z}$. WLOG, suppose that $m_S(0) = \underline{z}$. Since $M_S \cap M_N = [a,b]$ and both M_S and M_N are connected intervals resulting from Proposition 3, then we have $m_S(0) = a$ and/or $m_N(0) = a$. If $m_S(0) = a$, then $m_S^{-1}(a) = m_N^{-1}(a) = 0$ where the first equality comes from Proposition 4; vice versa. Thus $m_S(0) = m_N(0) = a$ where $a = \underline{z} = \min\{z_S^*, z_N^*\}$.

Similarly, we have $m_S(z_S^*) = b$ or/and $m_N(z_N^*) = b$. If $m_S(z_S^*) = b$, then $m_S^{-1}(b) = m_N^{-1}(b) = z_S^*$, which is feasible only if $b \in M_N$, or $m_N(z_S^*) \leq m_N(z_N^*)$. Using the same logic, we have $m_S^{-1}(b) = m_N^{-1}(b) = z_N^*$ if $m_S(z_N^*) \leq m_S(z_S^*)$. In all, $b = \min\{m_S(z_S^*), m_N(z_N^*)\}$.
■

Proof of Lemma 6. First to show that $z_S^* = z_N^*$. Assume both governments impose tax rates $t^W \in (-1,1)$ on workers' wages and $t^M \in (-1,1)$ on managers' rents in symmetric equilibrium. WLOG, suppose by contradiction that $z_S^* < z_N^*$. Pick an arbitrary $z \in (z_S^*, z_N^*)$. Since $z > z_S^*$, then an agent with ability z in South finds it more profitable to be a manager in South, that is,

$$(1 - t^M)r(z) > (1 - t^W)w(z)$$

On the contrary, an agent with ability z is more wiling to be a worker rather than a manager, that is,

$$(1 - t^M) r(z) < (1 - t^W) w(z)$$

Since $t^W < 1$, $t^M < 1$, and both $w(z)$ and $r(z)$ are strictly positive, we show a contradiction here so that $z_S^* \geq z_N^*$. Similar, we could show $z_S^* \leq z_N^*$. In all, $z_S^* = z_N^*$ in the equilibrium.

Since z_i^* is independent of t, then value of z^* is the same as the result of the integrated equilibrium in AGRH. ∎

Proof of Lemma 7. First show $z_S^* < z_N^*$. Agents with z_S^* are indifferent between workers and managers in South, so

$$(1 - t^M) r(z_S^*) = (1 - t_S^W) w(z_S^*) < (1 - t_N^W) w(z_S^*)$$

where the inequality comes from the condition of $t_S^W > t_N^W$. Hence, agents with z_S^* strictly prefer become workers in North. Equivalently, $z_S^* < z_N^*$.

Next to show that $z_N^* \leq \alpha$. Suppose by contradiction that $\alpha < z_N^*$, then there is a gap in $M_S \cup M_N$ such that $[\alpha, z_N^*] \not\subset M_S \cup M_N$. Due to positive sorting and labor market clearing condition, there exists a z_p such that $m^{-1}(\alpha) = z_p$ and $m^{-1}(z_N^*) = z_p$. Thus, there exists an equilibrium where workers with ability z_p in a particular country i can be assigned to managers with skill levels either α or z_N^* and receive the same wages, or

$$w_i'(z_p) = \frac{\alpha - w_i(z_p)}{1 - z_p} = \frac{z_N^* - w_i(z_p)}{1 - z_p}$$

which yields $\alpha = z_N^*$. Contradiction. Hence, we have $z_S^* < z_N^* \leq \alpha$.

Using the same logic, we could show $z_N^* < z_S^* \leq \alpha$ if $t_S^W < t_N^W$. ∎

A.2 Characterize the Equilibria

A.2.1 Decentralized Symmetric Equilibrium

A.2.1.1 (SA) Low Quality Symmetric Equilibrium

Hence, a low quality symmetric equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w_S(z) = w_N(z)$, schedules of matches $\boldsymbol{m} \equiv m_S(z) = m_N(z)$, schedules of rents $\boldsymbol{r} \equiv r_S(z) = r_N(z)$, and an occupation cutoff, z_N^*, such that

1. matching function

$$m(z) = \begin{cases} \frac{1+\alpha}{\alpha} hz(1-\frac{z}{2}) + z_N^* & \text{if } 0 < z < \alpha \\ hz(1-\frac{z}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } \alpha < z < z_N^* \end{cases}$$

where

(i)

$$h(1-\frac{\alpha}{2}) + z_N^* = 1 - hz_N^*(1-\frac{z_N^*}{2})$$

2. wage function

$$w(z) = \begin{cases} \frac{1}{2}\frac{1+\alpha}{\alpha} hz^2 - A_1(1-z) + z_N^* & \text{if } 0 < z < \alpha \\ \frac{1}{2} hz^2 - A_2(1-z) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } \alpha < z < z_N^* \end{cases}$$

where

(ii)

$$A_2 = A_1 + h$$

(iii)

$$\frac{1}{2}h(z_N^*)^2 - A_2(1-z_N^*) + 1 - hz_N^*(1-\frac{z_N^*}{2}) = t\frac{A_1}{h}$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_N^*, m(\alpha)] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (m(\alpha), 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2\alpha}{1+\alpha}h(C_1 - z)}$$

$$\phi_2(z) = \sqrt{h^2 + 2h(C_2 - z)}$$

and

$$C_1 = z_N^*$$

$$C_2 = 1 - hz_N^*(1 - \frac{z_N^*}{2})$$

Note conditions (i)-(iii) pin down $\{z_N^*, A_1, A_2\}$.

A.2.1.2 (SB) High Quality Symmetric Equilibrium

Denote z_α as $m(z_\alpha) = \alpha$. Hence, a HQSE can be characterized as schedules of wages $\boldsymbol{w} \equiv w_S(z) = w_N(z)$, schedules of matches $\boldsymbol{m} \equiv m_S(z) = m_N(z)$, schedules of rents $\boldsymbol{r} \equiv r_S(z) = r_N(z)$, and an occupation cutoff, z_N^*, such that

1. matching function

$$m(z) = \begin{cases} hz(1-\frac{z}{2}) + z_N^* & \text{if } 0 < z < z_\alpha \\ \frac{1+\alpha}{\alpha}hz(1-\frac{z}{2}) + 1 - \frac{1+\alpha}{\alpha}hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z_\alpha < z < z_N^* \end{cases}$$

   where

   (i)
$$z_N^* = \frac{1}{\alpha}hz_\alpha(1 - \frac{z_\alpha}{2}) + 1 - \frac{1+\alpha}{\alpha}hz_N^*(1 - \frac{z_N^*}{2})$$

   (ii)
$$hz_\alpha(1 - \frac{z_\alpha}{2}) + z_N^* = \alpha$$

2. wage function

$$w(z) = \begin{cases} \frac{1}{2}hz^2 - A_1(1-z) + z_N^* & \text{if } 0 < z < z_\alpha \\ \frac{1}{2}\frac{1+\alpha}{\alpha}hz^2 - A_2(1-z) + 1 - \frac{1+\alpha}{\alpha}hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z_\alpha < z < z_N^* \end{cases}$$

   where

   (iii)
$$A_1 = A_2 + \frac{1}{\alpha}hz_\alpha$$

   (iv)
$$\frac{1}{2}\frac{1+\alpha}{\alpha}h(z_N^*)^2 - A_2(1 - z_N^*) + 1 - \frac{1+\alpha}{\alpha}hz_N^*(1 - \frac{z_N^*}{2}) = t\frac{A_1}{h}$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1}{2}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_N^*, \alpha] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (\alpha, 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + 2h(C_1 - z)}$$
$$\phi_2(z) = \sqrt{h^2 + \frac{2\alpha}{1+\alpha}h(C_2 - z)}$$

and

$$C_1 = z_N^*$$
$$C_2 = 1 - \frac{1+\alpha}{\alpha}hz_N^*(1 - \frac{z_N^*}{2})$$

Note conditions (i)-(iv) pin down $\{z_\alpha, z_N^*, A_1, A_2\}$.

A.2.2 Asymmetric Equilibrium with Higher Tax on Southern Workers

A.2.2.1 (IA) In case of $z_S^* < m(z_S^*) < z_N^* < \alpha < 1$

Denote z_α s.t. $m(z_\alpha) = \alpha$ and denote z_N s.t. $m(z_N) = z_N^*$. The type-IA equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and occupation cutoffs, $\{z_N^*, z_S^*\}$, such that

1. matching function

$$m(z) = \begin{cases} (1+\alpha)hz(1-\frac{z}{2}) + z_S^* & \text{if } z \in [0, z_S^*] \\ \alpha hz(1-\frac{z}{2}) + hz_S^*(1-\frac{z_S^*}{2}) + z_S^* & \text{if } z \in (z_S^*, z_N] \\ \frac{\alpha}{1+\alpha}hz(1-\frac{z}{2}) + \frac{\alpha^2}{1+\alpha}hz_N(1-\frac{z_N}{2}) + hz_S^*(1-\frac{z_S^*}{2}) + z_S^* & \text{if } z \in (z_N, z_\alpha] \\ hz(1-\frac{z}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z \in (z_\alpha, z_N^*] \end{cases}$$

where

(i)

$$\frac{1}{1+\alpha}hz_\alpha(1-\frac{z_\alpha}{2})+1-hz_N^*(1-\frac{z_N^*}{2})=\frac{\alpha^2}{1+\alpha}hz_N(1-\frac{z_N}{2})+hz_S^*(1-\frac{z_S^*}{2})+z_S^*$$

(ii)

$$\alpha hz_N(1-\frac{z_N}{2})+hz_S^*(1-\frac{z_S^*}{2})+z_S^*=z_N^*$$

(iii)

$$hz_\alpha(1-\frac{z_\alpha}{2})+1-hz_N^*(1-\frac{z_N^*}{2})=\alpha$$

2. wage function

$$w(z)=\begin{cases}\frac{1}{2}(1+\alpha)hz^2-A_1(1-z)+z_S^* & \text{if } z\in[0,z_S^*]\\ \frac{1}{2}\alpha hz^2-A_2(1-z)+hz_S^*(1-\frac{z_S^*}{2})+z_S^* & \text{if } z\in(z_S^*,z_N]\\ \frac{1}{2}\frac{\alpha}{1+\alpha}hz^2-A_3(1-z)+\frac{\alpha^2}{1+\alpha}hz_N(1-\frac{z_N}{2})+hz_S^*(1-\frac{z_S^*}{2})+z_S^* & \text{if } z\in(z_N,z_\alpha]\\ \frac{1}{2}hz^2-A_4(1-z)+1-hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z\in(z_\alpha,z_N^*]\end{cases}$$

where

(iv)

$$A_2=A_1+hz_S^*$$

(v)

$$A_3=A_2+\frac{\alpha^2}{1+\alpha}hz_N$$

(vi)

$$A_4=A_3-\frac{1}{1+\alpha}hz_\alpha$$

(vii)

$$\frac{1}{h(1-z_N)}[z_N^*-(\frac{1}{2}\alpha hz_N^2-A_2(1-z_N)+hz_S^*(1-\frac{z_S^*}{2})+z_S^*)]$$
$$=\frac{1}{t_N}[\frac{1}{2}h(z_N^*)^2-A_4(1-z_N^*)+1-hz_N^*(1-\frac{z_N^*}{2})]$$

(viii)

$$\frac{A_1}{h}=\frac{1}{t_S}[\frac{1}{2}(1+\alpha)h(z_S^*)^2-A_1(1-z_S^*)+z_S^*]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1+\alpha}{2}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_S^*, m(z_S^*)] \\ \frac{1}{\phi_2(z)}[z - \frac{\alpha}{2}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (m(z_S^*), z_N^*] \\ \frac{1}{\phi_3(z)}[z - \frac{1}{2}\frac{\alpha}{1+\alpha}\frac{(h-\phi_3(z))^2}{h} + A_3\frac{\phi_3(z)}{h} - C_3] & \text{if } z \in (z_N^*, \alpha] \\ \frac{1}{\phi_4(z)}[z - \frac{1}{2}\frac{(h-\phi_4(z))^2}{h} + A_4\frac{\phi_4(z)}{h} - C_4] & \text{if } z \in (\alpha, 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2h}{1+\alpha}(C_1 - z)}$$
$$\phi_2(z) = \sqrt{h^2 + \frac{2h}{\alpha}(C_2 - z)}$$
$$\phi_3(z) = \sqrt{h^2 + 2h\frac{1+\alpha}{\alpha}(C_3 - z)}$$
$$\phi_4(z) = \sqrt{h^2 + 2h(C_4 - z)}$$

and

$$C_1 = z_S^*$$
$$C_2 = hz_S^*(1 - \frac{z_S^*}{2}) + z_S^*$$
$$C_3 = \frac{\alpha^2}{1+\alpha}hz_N(1 - \frac{z_N}{2}) + hz_S^*(1 - \frac{z_S^*}{2}) + z_S^*$$
$$C_4 = 1 - hz_N^*(1 - \frac{z_N^*}{2})$$

Note conditions (i)-(viii) pin down $\{z_N^*, z_S^*, z_N, z_\alpha, A_1, A_2, A_3, A_4\}$.

A.2.2.2 (IB) In case of $z_S^* < z_N^* < m(z_S^*) < \alpha < 1$

Denote z_α s.t. $m(z_\alpha) = \alpha$ and denote z_N s.t. $m(z_N) = z_N^*$. The type-IB equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and occupation cutoffs, $\{z_N^*, z_S^*\}$, such that

1. matching function

$$m(z) = \begin{cases} (1+\alpha)hz(1-\frac{z}{2}) + z_S^* & \text{if } z \in [0, z_N] \\ hz(1-\frac{z}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_N, z_S^*] \\ \frac{\alpha}{1+\alpha}hz(1-\frac{z}{2}) + \frac{1}{1+\alpha}hz_S^*(1-\frac{z_S^*}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_S^*, z_\alpha] \\ hz(1-\frac{z}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z \in (z_\alpha, z_N^*] \end{cases}$$

where

(i)

$$\frac{1}{1+\alpha}hz_\alpha(1-\frac{z_\alpha}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2}) = \frac{1}{1+\alpha}hz_S^*(1-\frac{z_S^*}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^*$$

(ii)

$$hz_\alpha(1-\frac{z_\alpha}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2}) = \alpha$$

(iii)

$$(1+\alpha)hz_N(1-\frac{z_N}{2}) + z_S^* = z_N^*$$

2. wage function

$$w(z) = \begin{cases} \frac{1}{2}(1+\alpha)hz^2 - A_1(1-z) + z_S^* & \text{if } z \in [0, z_N] \\ \frac{1}{2}hz^2 - A_2(1-z) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_N, z_S^*] \\ \frac{1}{2}\frac{\alpha}{1+\alpha}hz^2 - A_3(1-z) + \frac{1}{1+\alpha}hz_S^*(1-\frac{z_S^*}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_S^*, z_\alpha] \\ \frac{1}{2}hz^2 - A_4(1-z) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z \in (z_\alpha, z_N^*] \end{cases}$$

where

(iv)

$$A_2 = A_1 + \alpha h z_N$$

(v)

$$A_3 = A_2 + \frac{1}{1+\alpha}hz_S^*$$

(vi)

$$A_4 = A_3 - \frac{1}{1+\alpha}hz_\alpha$$

(vii)

$$\frac{1}{h(1-z_N)}][z_N^* - [\frac{1}{2}(1+\alpha)hz_N^2 - A_1(1-z_N) + z_S^*]] = \frac{1}{t_N}[\frac{1}{2}h(z_N^*)^2 - A_4(1-z_N^*) + 1 - hz_N^*(1-\frac{z_N^*}{2})]$$

(viii)

$$\frac{A_1}{h} = \frac{1}{t_S}[\frac{1}{2}h(z_S^*)^2 - A_2(1-z_S^*) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^*]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1+\alpha}{2}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_S^*, z_N^*] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (z_N^*, m(z_S^*)] \\ \frac{1}{\phi_3(z)}[z - \frac{1}{2}\frac{\alpha}{1+\alpha}\frac{(h-\phi_3(z))^2}{h} + A_3\frac{\phi_3(z)}{h} - C_3] & \text{if } z \in (m(z_S^*), \alpha] \\ \frac{1}{\phi_4(z)}[z - \frac{1}{2}\frac{(h-\phi_4(z))^2}{h} + A_4\frac{\phi_4(z)}{h} - C_4] & \text{if } z \in (\alpha, 1] \end{cases}$$

where

$$\begin{aligned} \phi_1(z) &= \sqrt{h^2 + \frac{2h}{1+\alpha}(C_1 - z)} \\ \phi_2(z) &= \sqrt{h^2 + 2h(C_2 - z)} \\ \phi_3(z) &= \sqrt{h^2 + 2h\frac{1+\alpha}{\alpha}(C_3 - z)} \\ \phi_4(z) &= \sqrt{h^2 + 2h(C_4 - z)} \end{aligned}$$

and

$$\begin{aligned} C_1 &= z_S^* \\ C_2 &= \alpha h z_N(1-\frac{z_N}{2}) + z_S^* \\ C_3 &= \frac{1}{1+\alpha}hz_S^*(1-\frac{z_S^*}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* \\ C_4 &= 1 - hz_N^*(1-\frac{z_N^*}{2}) \end{aligned}$$

Note conditions (i)-(viii) pin down $\{z_N^*, z_S^*, z_N, z_\alpha, A_1, A_2, A_3, A_4\}$.

A.2.2.3 (IC) In case of $z_S^* < z_N^* < \alpha < m(z_S^*) < 1$

Denote z_α s.t. $m(z_\alpha) = \alpha$ and denote z_N s.t. $m(z_N) = z_N^*$. The type-IC equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and occupation cutoffs, $\{z_N^*, z_S^*\}$, such that

1. matching function

$$m(z) = \begin{cases} (1+\alpha)hz(1-\frac{z}{2}) + z_S^* & \text{if } z \in [0, z_N] \\ hz(1-\frac{z}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_N, z_\alpha] \\ \frac{1+\alpha}{\alpha}hz(1-\frac{z}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* - \frac{1}{\alpha}hz_\alpha(1-\frac{z_\alpha}{2}) & \text{if } z \in (z_\alpha, z_S^*] \\ hz(1-\frac{z}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z \in (z_S^*, z_N^*] \end{cases}$$

where

(i)

$$\alpha h z_N(1-\frac{z_N}{2}) + z_S^* - \frac{1}{\alpha}hz_\alpha(1-\frac{z_\alpha}{2}) = 1 - hz_N^*(1-\frac{z_N^*}{2}) - \frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2})$$

(ii)

$$hz_\alpha(1-\frac{z_\alpha}{2}) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* = \alpha$$

(iii)

$$(1+\alpha)hz_N^*(1-\frac{z_N^*}{2}) + z_S^* = z_N^*$$

2. wage function

$$w(z) = \begin{cases} \frac{1}{2}(1+\alpha)hz^2 - A_1(1-z) + z_S^* & \text{if } z \in [0, z_N] \\ \frac{1}{2}hz^2 - A_2(1-z) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_N, z_\alpha] \\ \frac{1}{2}\frac{1+\alpha}{\alpha}hz^2 - A_3(1-z) + \alpha h z_N(1-\frac{z_N}{2}) + z_S^* - \frac{1}{\alpha}hz_\alpha(1-\frac{z_\alpha}{2}) & \text{if } z \in (z_\alpha, z_S^*] \\ \frac{1}{2}hz^2 - A_4(1-z) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z \in (z_S^*, z_N^*] \end{cases}$$

where

(iv)

$$A_2 = A_1 + \alpha h z_N$$

(v)

$$A_3 = A_2 - \frac{1}{\alpha}hz_\alpha$$

(vi)

$$A_4 = A_3 + \frac{1}{\alpha} h z_S^*$$

(vii)

$$\frac{A_1}{h} = \frac{1}{t_S}[\frac{1}{2}h(z_S^*)^2 - A_4(1 - z_S^*) + 1 - hz_N^*(1 - \frac{z_N^*}{2})]$$

(viii)

$$\frac{1}{h(1-z_N)}[z_N^* - (\frac{1}{2}(1+\alpha)hz_N^2 - A_1(1-z_N) + z_S^*)] = \frac{1}{t_N}[\frac{1}{2}h(z_N^*)^2 - A_4(1-z_N^*) + 1 - hz_N^*(1 - \frac{z_N^*}{2})]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1+\alpha}{2}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_S^*, z_N^*] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (z_N^*, \alpha] \\ \frac{1}{\phi_3(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_3(z))^2}{h} + A_3\frac{\phi_3(z)}{h} - C_3] & \text{if } z \in (\alpha, m(z_S^*)] \\ \frac{1}{\phi_4(z)}[z - \frac{1}{2}\frac{(h-\phi_4(z))^2}{h} + A_4\frac{\phi_4(z)}{h} - C_4] & \text{if } z \in (m(z_S^*), 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2h}{1+\alpha}(C_1 - z)}$$
$$\phi_2(z) = \sqrt{h^2 + 2h(C_2 - z)}$$
$$\phi_3(z) = \sqrt{h^2 + 2h\frac{\alpha}{1+\alpha}(C_3 - z)}$$
$$\phi_4(z) = \sqrt{h^2 + 2h(C_4 - z)}$$

and

$$C_1 = z_S^*$$
$$C_2 = \alpha h z_N(1 - \frac{z_N}{2}) + z_S^*$$
$$C_3 = \alpha h z_N(1 - \frac{z_N}{2}) + z_S^* - \frac{1}{\alpha}hz_\alpha(1 - \frac{z_\alpha}{2})$$
$$C_4 = 1 - hz_N^*(1 - \frac{z_N^*}{2})$$

Conditions (i)-(viii) pin down $\{z_N^*, z_S^*, z_N, z_\alpha, A_1, A_2, A_3, A_4\}$.

A.2.2.4 (ID) In case of $z_N^* > \alpha$

The type-ID equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and an occupation cutoff, z_N^*, such that

1. matching function

$$m(z) = \begin{cases} z_N^* + \frac{1+\alpha}{\alpha} hz(1-\frac{z}{2}) & \text{if } 0 < z < \alpha \\ z_N^* + h(1-\frac{\alpha}{2}) + hz(1-\frac{z}{2}) & \text{if } \alpha < z < z_N^* \end{cases}$$

where

(i)

$$z_N^* + h(1-\frac{\alpha}{2}) + hz_N^*(1-\frac{z_N^*}{2}) = 1$$

2. wage function

$$w(z) = \begin{cases} z_N^* + \frac{1+\alpha}{2\alpha} hz^2 - A_1(1-z) & \text{if } 0 < z < \alpha \\ z_N^* + h(1-\frac{\alpha}{2}) + \frac{1}{2}hz^2 - A_2(1-z) & \text{if } \alpha < z < z_N^* \end{cases}$$

where

(ii)

$$A_2 = A_1 + h$$

(iii)

$$z_N^* + h(1-\frac{\alpha}{2}) + \frac{1}{2}h(z_N^*)^2 - A_2(1-z_N^*) = t\frac{A_1}{h}$$

Besides,

$$w(\alpha) > \frac{t_S^W}{t^M} R(\alpha) = t_S R(\alpha)$$

Equivalently,

$$\frac{1}{t_S}[z_N^* + \frac{1+\alpha}{2\alpha} h\alpha^2 - A_1(1-\alpha)] \geq \frac{1}{h}[\alpha - (z_N^* - A_1)]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_N^*, m(\alpha)] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (m(\alpha), 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2\alpha}{1+\alpha}h(C_1 - z)}$$
$$\phi_2(z) = \sqrt{h^2 + 2h(C_2 - z)}$$

and

$$C_1 = z_N^*$$
$$C_2 = z_N^* + h(1 - \frac{\alpha}{2})$$

Note conditions (i)-(iv) pin down $\{z_N^*, A_1, A_2\}$.

A.2.2.5 (IE) In case of $z_S^* < z_N^* = \alpha < m(z_S^*) < 1$

Denote z_α s.t. $m(z_\alpha) = \alpha$ and denote z_N s.t. $m(z_N) = z_N^*$. The type-IA equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and occupation cutoffs, $\{z_N^*, z_S^*\}$, such that

1. matching function

$$m(z) = \begin{cases} (1+\alpha)hz(1-\frac{z}{2}) + z_S^* & \text{if } z \in [0, z_N] \\ \frac{1+\alpha}{\alpha}hz(1-\frac{z}{2}) + \frac{\alpha^2-1}{\alpha}hz_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_N, z_S^*] \\ hz(1-\frac{z}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z \in (z_S^*, z_N^*] \end{cases}$$

where

(i)
$$\frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2}) + \frac{\alpha^2-1}{\alpha}hz_N(1-\frac{z_N}{2}) + z_S^* = 1 - hz_N^*(1-\frac{z_N^*}{2})$$

(ii)
$$(1+\alpha)hz_N(1-\frac{z_N}{2}) + z_S^* = z_N^*$$

(iii)
$$z_N^* = \alpha$$

2. wage function

$$w(z) = \begin{cases} \frac{1}{2}(1+\alpha)hz^2 - A_1(1-z) + z_S^* & \text{if } z \in [0, z_N] \\ \frac{1}{2}\frac{1+\alpha}{\alpha}hz^2 - A_2(1-z) + \frac{\alpha^2-1}{\alpha}hz_N(1-\frac{z_N}{2}) + z_S^* & \text{if } z \in (z_N, z_S^*] \\ \frac{1}{2}hz^2 - A_3(1-z) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } z \in (z_S^*, z_N^*] \end{cases}$$

where

(iv)

$$A_2 = A_1 + \frac{\alpha^2 - 1}{\alpha}hz_N$$

(v)

$$A_3 = A_2 + \frac{1}{\alpha}hz_S^*$$

(vi)

$$\frac{1}{h(1-z_N)}[z_N^* - (\frac{1}{2}(1+\alpha)hz_N^2 - A_1(1-z_N) + z_S^*)]$$
$$= \frac{1}{t_N}[\frac{1}{2}h(z_N^*)^2 - A_3(1-z_N^*) + 1 - hz_N^*(1-\frac{z_N^*}{2})]$$

(vii)

$$\frac{A_1}{h} = \frac{1}{t_S}[\frac{1}{2}h(z_S^*)^2 - A_3(1-z_S^*) + 1 - hz_N^*(1-\frac{z_N^*}{2})]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1+\alpha}{2}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_S^*, z_N^*] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (z_N^*, m(z_S^*)] \\ \frac{1}{\phi_3(z)}[z - \frac{1}{2}\frac{(h-\phi_3(z))^2}{h} + A_3\frac{\phi_3(z)}{h} - C_3] & \text{if } z \in (m(z_S^*), 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2h}{1+\alpha}(C_1 - z)}$$
$$\phi_2(z) = \sqrt{h^2 + 2h\frac{\alpha}{1+\alpha}(C_2 - z)}$$
$$\phi_3(z) = \sqrt{h^2 + 2h(C_3 - z)}$$

and

$$C_1 = z_S^*$$
$$C_2 = \frac{\alpha^2 - 1}{\alpha} h z_N (1 - \frac{z_N}{2}) + z_S^*$$
$$C_3 = 1 - h z_N^* (1 - \frac{z_N^*}{2})$$

Note conditions (i)-(vii) pin down $\{z_S^*, z_N, A_1, A_2, A_3\}$.

A.2.3 Asymmetric Equilibrium with Lower Tax on Southern Workers

A.2.3.1 (DA) In case of $z_N^* < m(z_N^*) < z_S^* < \alpha < 1$

Denote z_α s.t. $m(z_\alpha) = \alpha$ and denote z_N s.t. $m(z_S) = z_S^*$. The type-DA equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and occupation cutoffs, $\{z_N^*, z_S^*\}$, such that

1. matching function

$$m(z) = \begin{cases} \frac{1+\alpha}{\alpha} h z (1 - \frac{z}{2}) + z_N^* & \text{if } z \in [0, z_N^*] \\ \frac{1}{\alpha} h z (1 - \frac{z}{2}) + h z_N^* (1 - \frac{z_N^*}{2}) + z_N^* & \text{if } z \in (z_N^*, z_S] \\ \frac{1}{1+\alpha} h z (1 - \frac{z}{2}) + \frac{1}{\alpha(1+\alpha)} h z_S (1 - \frac{z_S}{2}) + h z_N^* (1 - \frac{z_N^*}{2}) + z_N^* & \text{if } z \in (z_S, z_\alpha) \\ \frac{1}{\alpha} h z (1 - \frac{z}{2}) + 1 - \frac{1}{\alpha} h z_S^* (1 - \frac{z_S^*}{2}) & \text{if } z \in (z_\alpha, z_S^*] \end{cases}$$

where

(i)

$$\frac{1}{\alpha(1+\alpha)} h z_S (1 - \frac{z_S}{2}) + h z_N^* (1 - \frac{z_N^*}{2}) + z_N^* = \frac{1}{\alpha(1+\alpha)} h z_\alpha (1 - \frac{z_\alpha}{2}) + 1 - \frac{1}{\alpha} h z_S^* (1 - \frac{z_S^*}{2})$$

(ii)

$$\frac{1}{\alpha} h z_S (1 - \frac{z_S}{2}) + h z_N^* (1 - \frac{z_N^*}{2}) + z_N^* = z_S^*$$

(iii)

$$\frac{1}{\alpha} h z_\alpha (1 - \frac{z_\alpha}{2}) + 1 - \frac{1}{\alpha} h z_S^* (1 - \frac{z_S^*}{2}) = \alpha$$

2. wage function

$$w(z) = \begin{cases} \frac{1}{2}\frac{1+\alpha}{\alpha}hz^2 - A_1(1-z) + z_N^* & \text{if } z \in [0, z_N^*] \\ \frac{1}{2}\frac{1}{\alpha}hz^2 - A_2(1-z) + hz_N^*(1-\frac{z_N^*}{2}) + z_N^* & \text{if } z \in (z_N^*, z_S] \\ \frac{1}{2}\frac{1}{1+\alpha}hz^2 - A_3(1-z) + \frac{1}{\alpha(1+\alpha)}hz_S(1-\frac{z_S}{2}) + hz_N^*(1-\frac{z_N^*}{2}) + z_N^* & \text{if } z \in (z_S, z_\alpha) \\ \frac{1}{2}\frac{1}{\alpha}hz^2 - A_4(1-z) + 1 - \frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2}) & \text{if } z \in (z_\alpha, z_S^*] \end{cases}$$

where

(iv)

$$A_2 = A_1 + hz_N^*$$

(v)

$$A_3 = A_2 + \frac{1}{\alpha(1+\alpha)}hz_S$$

(vi)

$$A_4 = A_3 - \frac{1}{\alpha(1+\alpha)}hz_\alpha$$

(vii)

$$\begin{aligned} &\frac{1}{h(1-z_S)}[z_S^* - (\frac{1}{2}\frac{1}{\alpha}hz_S^2 - A_2(1-z_S) + hz_N^*(1-\frac{z_N^*}{2}) + z_N^*)] \\ &= \frac{1}{t'}[\frac{1}{2}\frac{1}{\alpha}h(z_S^*)^2 - A_4(1-z_S^*) + 1 - \frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2})] \end{aligned}$$

(viii)

$$\frac{A_1}{h} = \frac{1}{t}[\frac{1}{2}\frac{1+\alpha}{\alpha}h(z_N^*)^2 - A_1(1-z_N^*) + z_N^*]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1+\alpha}{2\alpha}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_N^*, m(z_N^*)] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2\alpha}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (m(z_N^*), z_S^*] \\ \frac{1}{\phi_3(z)}[z - \frac{1}{2}\frac{1}{1+\alpha}\frac{(h-\phi_3(z))^2}{h} + A_3\frac{\phi_3(z)}{h} - C_3] & \text{if } z \in (z_S^*, \alpha] \\ \frac{1}{\phi_4(z)}[z - \frac{1}{2\alpha}\frac{(h-\phi_4(z))^2}{h} + A_4\frac{\phi_4(z)}{h} - C_4] & \text{if } z \in (\alpha, 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2h\alpha}{1+\alpha}(C_1 - z)}$$

$$\phi_2(z) = \sqrt{h^2 + 2h\alpha(C_2 - z)}$$
$$\phi_3(z) = \sqrt{h^2 + 2h(1+\alpha)(C_3 - z)}$$
$$\phi_4(z) = \sqrt{h^2 + 2h\alpha(C_4 - z)}$$

and

$$C_1 = z_N^*$$
$$C_2 = hz_N^*(1 - \frac{z_N^*}{2}) + z_N^*$$
$$C_3 = \frac{1}{\alpha(1+\alpha)}hz_S(1 - \frac{z_S}{2}) + hz_N^*(1 - \frac{z_N^*}{2}) + z_N^*$$
$$C_4 = 1 - \frac{1}{\alpha}hz_S^*(1 - \frac{z_S^*}{2})$$

Note conditions (i)-(viii) pin down $\{z_N^*, z_S^*, z_S, z_\alpha, A_1, A_2, A_3, A_4\}$.

A.2.3.2 (DB) In case of $z_N^* < z_S^* < m(z_N^*) < \alpha < 1$

Denote z_α s.t. $m(z_\alpha) = \alpha$ and denote z_N s.t. $m(z_S) = z_S^*$. The type-DB equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and occupation cutoffs, $\{z_N^*, z_S^*\}$, such that

1. matching function

$$m(z) = \begin{cases} \frac{1+\alpha}{\alpha}hz(1-\frac{z}{2}) + z_N^* & \text{if } z \in [0, z_S] \\ hz(1-\frac{z}{2}) + \frac{1}{\alpha}hz_S(1-\frac{z_S}{2}) + z_N^* & \text{if } z \in (z_S, z_N^*] \\ \frac{1}{1+\alpha}hz(1-\frac{z}{2}) + \frac{\alpha}{1+\alpha}hz_N^*(1-\frac{z_N^*}{2}) + \frac{1}{\alpha}hz_S(1-\frac{z_S}{2}) + z_N^* & \text{if } z \in (z_N^*, z_\alpha) \\ \frac{1}{\alpha}hz(1-\frac{z}{2}) + 1 - \frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2}) & \text{if } z \in (z_\alpha, z_S^*] \end{cases}$$

where

(i)

$$\frac{\alpha}{1+\alpha}hz_N^*(1-\frac{z_N^*}{2}) + \frac{1}{\alpha}hz_S(1-\frac{z_S}{2}) + z_N^* = \frac{1}{\alpha(1+\alpha)}hz_\alpha(1-\frac{z_\alpha}{2}) + 1 - \frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2})$$

(ii)

$$\frac{1+\alpha}{\alpha}hz_S(1-\frac{z_S}{2}) + z_N^* = z_S^*$$

(iii)

$$\frac{1}{\alpha}hz_\alpha(1-\frac{z_\alpha}{2})+1-\frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2})=\alpha$$

2. wage function

$$w(z)=\begin{cases}\frac{1}{2}\frac{1+\alpha}{\alpha}hz^2-A_1(1-z)+z_N^* & \text{if } z\in[0,z_S] \\ \frac{1}{2}hz^2-A_2(1-z)+\frac{1}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^* & \text{if } z\in(z_S,z_N^*] \\ \frac{1}{2}\frac{1}{1+\alpha}hz^2-A_3(1-z)+\frac{\alpha}{1+\alpha}hz_N^*(1-\frac{z_N^*}{2})+\frac{1}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^* & \text{if } z\in(z_N^*,z_\alpha) \\ \frac{1}{2}\frac{1}{\alpha}hz^2-A_4(1-z)+1-\frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2}) & \text{if } z\in(z_\alpha,z_S^*]\end{cases}$$

where

(iv)

$$A_2=A_1+\frac{1}{\alpha}hz_S$$

(v)

$$A_3=A_2+\frac{\alpha}{1+\alpha}hz_N^*$$

(vi)

$$A_4=A_3-\frac{1}{\alpha(1+\alpha)}hz_\alpha$$

(vii)

$$\frac{1}{h(1-z_S)}[z_S^*-(\frac{1}{2}\frac{1+\alpha}{\alpha}hz_S^2-A_1(1-z_S)+z_N^*)]$$
$$=\frac{1}{t'}[\frac{1}{2}\frac{1}{\alpha}h(z_S^*)^2-A_4(1-z_S^*)+1-\frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2})]$$

(viii)

$$\frac{A_1}{h}=\frac{1}{t}[\frac{1}{2}h(z_N^*)^2-A_2(1-z_N^*)+\frac{1}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^*]$$

3. rent function

$$r(z)=\begin{cases}\frac{1}{\phi_1(z)}[z-\frac{1+\alpha}{2\alpha}\frac{(h-\phi_1(z))^2}{h}+A_1\frac{\phi_1(z)}{h}-C_1] & \text{if } z\in[z_N^*,z_S^*)] \\ \frac{1}{\phi_2(z)}[z-\frac{1}{2}\frac{(h-\phi_2(z))^2}{h}+A_2\frac{\phi_2(z)}{h}-C_2] & \text{if } z\in(z_S^*,m(z_N^*)] \\ \frac{1}{\phi_3(z)}[z-\frac{1}{2}\frac{1}{1+\alpha}\frac{(h-\phi_3(z))^2}{h}+A_3\frac{\phi_3(z)}{h}-C_3] & \text{if } z\in(m(z_N^*),\alpha] \\ \frac{1}{\phi_4(z)}[z-\frac{1}{2\alpha}\frac{(h-\phi_4(z))^2}{h}+A_4\frac{\phi_4(z)}{h}-C_4] & \text{if } z\in(\alpha,1]\end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2h\alpha}{1+\alpha}(C_1 - z)}$$
$$\phi_2(z) = \sqrt{h^2 + 2h(C_2 - z)}$$
$$\phi_3(z) = \sqrt{h^2 + 2h(1+\alpha)(C_3 - z)}$$
$$\phi_4(z) = \sqrt{h^2 + 2h\alpha(C_4 - z)}$$

and

$$C_1 = z_N^*$$
$$C_2 = \frac{1}{\alpha}hz_S(1 - \frac{z_S}{2}) + z_N^*$$
$$C_3 = \frac{\alpha}{1+\alpha}hz_N^*(1 - \frac{z_N^*}{2}) + \frac{1}{\alpha}hz_S(1 - \frac{z_S}{2}) + z_N^*$$
$$C_4 = 1 - \frac{1}{\alpha}hz_S^*(1 - \frac{z_S^*}{2})$$

Note conditions (i)-(viii) pin down $\{z_N^*, z_S^*, z_S, z_\alpha, A_1, A_2, A_3, A_4\}$.

A.2.3.3 (DC) In case of $z_N^* < z_S^* < \alpha < m(z_N^*) < 1$

Denote z_α s.t. $m(z_\alpha) = \alpha$ and denote z_N s.t. $m(z_S) = z_S^*$. The type-DC equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and occupation cutoffs, $\{z_N^*, z_S^*\}$, such that

1. matching function

$$m(z) = \begin{cases} \frac{1+\alpha}{\alpha}hz(1-\frac{z}{2}) + z_N^* & \text{if } z \in [0, z_S] \\ hz(1-\frac{z}{2}) + \frac{1}{\alpha}hz_S(1-\frac{z_S}{2}) + z_N^* & \text{if } z \in (z_S, z_\alpha] \\ \frac{1+\alpha}{\alpha}hz(1-\frac{z}{2}) - \frac{1}{\alpha}hz_\alpha(1-\frac{z_\alpha}{2}) + \frac{1}{\alpha}hz_S(1-\frac{z_S}{2}) + z_N^* & \text{if } z \in (z_\alpha, z_N^*] \\ \frac{1}{\alpha}hz(1-\frac{z}{2}) + 1 - \frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2}) & \text{if } z \in (z_N^*, z_S^*] \end{cases}$$

where

(i)

$$hz_N^*(1 - \frac{z_N^*}{2}) - \frac{1}{\alpha}hz_\alpha(1 - \frac{z_\alpha}{2}) + \frac{1}{\alpha}hz_S(1 - \frac{z_S}{2}) + z_N^* = 1 - \frac{1}{\alpha}hz_S^*(1 - \frac{z_S^*}{2})$$

(ii)

$$\frac{1+\alpha}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^*=z_S^*$$

(iii)

$$hz_\alpha(1-\frac{z_\alpha}{2})+\frac{1}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^*=\alpha$$

2. wage function

$$w(z)=\begin{cases}\frac{1}{2}\frac{1+\alpha}{\alpha}hz^2-A_1(1-z)+z_N^* & \text{if } z\in[0,z_S]\\ \frac{1}{2}hz^2-A_2(1-z)+\frac{1}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^* & \text{if } z\in(z_S,z_N^*]\\ \frac{1}{2}\frac{1+\alpha}{\alpha}hz^2-A_3(1-z)-\frac{1}{\alpha}hz_\alpha(1-\frac{z_\alpha}{2})+\frac{1}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^* & \text{if } z\in(z_N^*,z_\alpha)\\ \frac{1}{2}\frac{1}{\alpha}hz^2-A_4(1-z)+1-\frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2}) & \text{if } z\in(z_\alpha,z_S^*]\end{cases}$$

where

(iv)

$$A_2=A_1+\frac{1}{\alpha}hz_S$$

(v)

$$A_3=A_2-\frac{1}{\alpha}hz_\alpha$$

(vi)

$$A_4=A_3+hz_N^*$$

(vii)

$$\frac{1}{h(1-z_S)}[z_S^*-(\frac{1}{2}\frac{1+\alpha}{\alpha}hz_S^2-A_1(1-z_S)+z_N^*)]$$
$$=\frac{1}{t'}[\frac{1}{2}\frac{1}{\alpha}h(z_S^*)^2-A_4(1-z_S^*)+1-\frac{1}{\alpha}hz_S^*(1-\frac{z_S^*}{2})]$$

(viii)

$$\frac{A_1}{h}=\frac{1}{t}[\frac{1}{2}h(z_N^*)^2-A_2(1-z_N^*)+\frac{1}{\alpha}hz_S(1-\frac{z_S}{2})+z_N^*]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1+\alpha}{2\alpha}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_N^*, z_S^*)] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (z_S^*, \alpha] \\ \frac{1}{\phi_3(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_3(z))^2}{h} + A_3\frac{\phi_3(z)}{h} - C_3] & \text{if } z \in (\alpha, m(z_N^*)] \\ \frac{1}{\phi_4(z)}[z - \frac{1}{2\alpha}\frac{(h-\phi_4(z))^2}{h} + A_4\frac{\phi_4(z)}{h} - C_4] & \text{if } z \in (m(z_N^*), 1] \end{cases}$$

where

$$\begin{aligned} \phi_1(z) &= \sqrt{h^2 + \frac{2h\alpha}{1+\alpha}(C_1 - z)} \\ \phi_2(z) &= \sqrt{h^2 + 2h(C_2 - z)} \\ \phi_3(z) &= \sqrt{h^2 + 2h\frac{\alpha}{1+\alpha}(C_3 - z)} \\ \phi_4(z) &= \sqrt{h^2 + 2h\alpha(C_4 - z)} \end{aligned}$$

and

$$\begin{aligned} C_1 &= z_N^* \\ C_2 &= \frac{1}{\alpha}hz_S(1 - \frac{z_S}{2}) + z_N^* \\ C_3 &= -\frac{1}{\alpha}hz_\alpha(1 - \frac{z_\alpha}{2}) + \frac{1}{\alpha}hz_S(1 - \frac{z_S}{2}) + z_N^* \\ C_4 &= 1 - \frac{1}{\alpha}hz_S^*(1 - \frac{z_S^*}{2}) \end{aligned}$$

Note conditions (i)-(viii) pin down $\{z_N^*, z_S^*, z_S, z_\alpha, A_1, A_2, A_3, A_4\}$.

A.2.3.4 (DD) In case of $z_N^* > \alpha$

The type-DD equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and an occupation cutoff, z_N^*, such that

1. matching function

$$m(z) = \begin{cases} \frac{1+\alpha}{\alpha}hz(1 - \frac{z}{2}) + z_N^* & \text{if } 0 < z < \alpha \\ hz(1 - \frac{z}{2}) + 1 - hz_N^*(1 - \frac{z_N^*}{2}) & \text{if } \alpha < z < z_N^* \end{cases}$$

where

(i)

$$(1+\alpha)h(1-\frac{\alpha}{2}) + z_N^* = h\alpha(1-\frac{\alpha}{2}) + 1 - hz_N^*(1-\frac{z_N^*}{2})$$

2. wage function

$$w(z) = \begin{cases} \frac{1+\alpha}{2\alpha}hz^2 - A_1(1-z) + z_N^* & \text{if } 0 < z < \alpha \\ \frac{1}{2}hz^2 - A_2(1-z) + 1 - hz_N^*(1-\frac{z_N^*}{2}) & \text{if } \alpha < z < z_N^* \end{cases}$$

where

(ii)

$$A_2 = A_1 + h$$

(iii)

$$\frac{A_1}{h} = \frac{1}{t}[\frac{1}{2}h(z_N^*)^2 - A_2(1-z_N^*) + 1 - hz_N^*(1-\frac{z_N^*}{2})]$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_N^*, m(\alpha)] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (m(\alpha), 1] \end{cases}$$

where

$$\phi_1(z) = \sqrt{h^2 + \frac{2\alpha}{1+\alpha}h(C_1 - z)}$$
$$\phi_2(z) = \sqrt{h^2 + 2h(C_2 - z)}$$

and

$$C_1 = z_N^*$$
$$C_2 = 1 - hz_N^*(1-\frac{z_N^*}{2})$$

Note conditions (i)-(iv) pin down $\{z_N^*, A_1, A_2\}$.

A.2.3.5 (DE) In case of $z_N^* < \alpha$

The type-DE equilibrium can be characterized as schedules of wages $\boldsymbol{w} \equiv w(z)$, schedules of matches $\boldsymbol{m} \equiv m(z)$, schedules of rents $\boldsymbol{r} \equiv r(z)$, and an occupation cutoff, z_N^*, such that

1. matching function

$$m(z) = \begin{cases} \frac{1+\alpha}{\alpha} hz(1-\frac{z}{2}) + z_N^* & \text{if } 0 < z < z_N^* \\ \frac{1}{\alpha} hz(1-\frac{z}{2}) + 1 - h(1-\frac{\alpha}{2}) & \text{if } z_N^* < z < \alpha \end{cases}$$

   where

   (i)

$$hz_N^*(1-\frac{z_N^*}{2}) + z_N^* = 1 - h(1-\frac{\alpha}{2})$$

2. wage function

$$w(z) = \begin{cases} \frac{1}{2}\frac{1+\alpha}{\alpha} hz^2 - A_1(1-z) + z_N^* & \text{if } 0 < z < z_N^* \\ \frac{1}{2}\frac{1}{\alpha} hz^2 - A_2(1-z) + 1 - h(1-\frac{\alpha}{2}) & \text{if } z_N^* < z < \alpha \end{cases}$$

   where

   (ii)

$$\frac{A_1}{h} = \frac{1}{t}[\frac{1}{2}\frac{1+\alpha}{\alpha} h(z_N^*)^2 - A_1(1-z_N^*) + z_N^*]$$

   and

   (iii)

$$A_2 = A_1 + hz_N^*$$

3. rent function

$$r(z) = \begin{cases} \frac{1}{\phi_1(z)}[z - \frac{1}{2}\frac{1+\alpha}{\alpha}\frac{(h-\phi_1(z))^2}{h} + A_1\frac{\phi_1(z)}{h} - C_1] & \text{if } z \in [z_N^*, m(z_N^*)] \\ \frac{1}{\phi_2(z)}[z - \frac{1}{2\alpha}\frac{(h-\phi_2(z))^2}{h} + A_2\frac{\phi_2(z)}{h} - C_2] & \text{if } z \in (m(z_N^*), 1] \end{cases}$$

   where

$$\phi_1(z) = \sqrt{h^2 + \frac{2\alpha}{1+\alpha} h(C_1 - z)}$$

$$\phi_2(z) = \sqrt{h^2 + 2h\alpha(C_2 - z)}$$

and

$$C_1 = z_N^*$$
$$C_2 = 1 - h(1 - \frac{\alpha}{2})$$

Besides, in order to guarantee the existence of competitive equilibrium, it requires (iv)

$$tR(z) \geq w(z) \geq t'R(z) \text{ if } z_N^* < z < \alpha$$

where

$$R(z) = [z - \frac{1}{2}\frac{1}{\alpha}\frac{(h - \phi(z))^2}{h} + A_2\frac{h - \phi(z)}{h} - 1 + h(1 - \frac{\alpha}{2})]\frac{1}{\phi(z)}$$

with

$$\phi(z) = \sqrt{h^2 + 2h[1 - h(1 - \frac{\alpha}{2}) - \frac{1}{\alpha}[\frac{1}{\alpha}hz(1 - \frac{z}{2}) + 1 - h(1 - \frac{\alpha}{2})]]}$$

In all, conditions (i)-(iv) pin down $\{z_N^*, A_1, A_2\}$.

Appendix B

Supplemental Material for Chapter 2

B.1 Derivation of $\mathbf{D}_n^*$

Let the profit maximizer be denoted by $\mathbf{D}^* = \arg\max_{\mathbf{D}\in\mathbb{D}} \Pi(\mathbf{D})$. The optimality of $\mathbf{D}^*$ implies that profit at $\mathbf{D}^*$ must be (weakly) higher than the profit at any one-market deviation, that is

$$\Pi(D_1^*, ..., D_n^*, ...D_N^*) \geq \Pi(D_1^*, ..., D_n, ...D_N^*) \; \forall n, D_n^* \neq D_i.$$

Let $\hat{\mathbf{D}} = \{D_1^*, ..., D_{n-1}^*, D_n, D_{n+1}^*, ...D_N^*\}$, then

$$\Pi(\mathbf{D}^*) - \Pi(\hat{\mathbf{D}}) = \tilde{B}[(\sum_{l\neq n} D_l^* \xi_l + D_n^* \xi_n)^{(\sigma-1)/\theta} - (\sum_{l\neq n} D_l^* \xi_l + D_n \xi_n)^{(\sigma-1)/\theta}] - D_n^* f_n + D_n f_n \geq 0.$$

Hence, it must be that $D_n^* = 1$ and $D_n = 0$ if and only if $\tilde{B}[(\sum_{l\neq n} D_l^* \xi_l + \xi_n)^{(\sigma-1)/\theta} - (\sum_{l\neq n} D_l^* \xi_l)^{(\sigma-1)/\theta}] - f_n \geq 0$, and $D_n^* = 0$ and $\hat{D}_n = 1$ if and only if $\tilde{B}[(\sum_{l\neq n} D_l^* \xi_l)^{(\sigma-1)/\theta} - (\sum_{l\neq n} D_l^* \xi_l + \xi_n)^{(\sigma-1)/\theta}] + f_n \geq 0$. Together we have $D_n^* = 1[\tilde{B}[(\sum_{l\neq n} D_l^* \xi_l + \xi_n)^{(\sigma-1)/\theta} - (\sum_{l\neq n} D_l^* \xi_l)^{(\sigma-1)/\theta}] - f_n \geq 0]$.

B.2 Proof of Proposition 1

Proof. The proposition is proved by contradiction. Let $\mathbf{D} > \mathbf{D}'$, and assume that there exists a n such that $D_n \geq D'_n$, $V_n(\mathbf{D}) = 0$ and $V_n(\mathbf{D}') = 1$. By definition, $V_n(\mathbf{D}) = 0$ means

$$\tilde{B}[(\sum_{l \neq n} D_l \xi_l + \xi_n)^{(\sigma-1)/\theta} - (\sum_{l \neq n} D_l \xi_l)^{(\sigma-1)/\theta}] - f_n < 0,$$

and $V_n(\mathbf{D}') = 1$ is equivalent to

$$\tilde{B}[(\sum_{l \neq n} D'_l \xi_l + \xi_n)^{(\sigma-1)/\theta} - (\sum_{l \neq n} D'_l \xi_l)^{(\sigma-1)/\theta}] - f_n \geq 0.$$

Together we have

$$(\sum_{l \neq n} D'_l \xi_l + \xi_n)^{(\sigma-1)/\theta} - (\sum_{l \neq n} D'_l \xi_l)^{(\sigma-1)/\theta} > (\sum_{l \neq n} D_l \xi_l + \xi_n)^{(\sigma-1)/\theta} - (\sum_{l \neq n} D_l \xi_l)^{(\sigma-1)/\theta}. \tag{B.1}$$

Let $u(x) = x^{(\sigma-1)/\theta}$, so the function $u : I \to \mathbb{R}$ is continuous, differentiable and strictly convex whenever $(\sigma - 1)/\theta > 1$. Rewrite the equation (B.1) as

$$u(\sum_{l \neq n} D'_l \xi_l + \xi_n) - u(\sum_{l \neq n} D'_l \xi_l) > u(\sum_{l \neq n} D_l \xi_l + \xi_n) - u(\sum_{l \neq n} D_l \xi_l). \tag{B.2}$$

Lemma 11. *A convex function $u(y)$ is continuous, differentiable from the left and from the right. The derivative is increasing, i.e. for $x < y$ we have $u'(x-) \leq u'(x+) \leq u'(y-) \leq u'(y+)$. If $u(y)$ is strictly convex then $u'(x+) < u'(y-)$.*

Therefore, we have

$$\begin{aligned} & u(\sum_{l \neq n} D'_l \xi_l + \xi_n) - u(\sum_{l \neq n} D'_l \xi_l) \\ &= \int_{\sum_{l \neq n} D'_l \xi_l}^{\sum_{l \neq n} D'_l \xi_l + \xi_n} u'(x) dx \leq \int_{\sum_{l \neq n} D_l \xi_l}^{\sum_{l \neq n} D_l \xi_l + \xi_n} u'(x) dx = u(\sum_{l \neq n} D_l \xi_l + \xi_n) - u(\sum_{l \neq n} D_l \xi_l) \end{aligned}$$

where the inequality comes from the fact that $D'_l \leq D_l$ for any l and the derivative of a strictly convex function is strictly increasing. However, this contradicts with the equation (B.1). So we prove that $V(.)$ is an increasing function under the assumption of $(\sigma - 1)/\theta > 1$. ∎

B.3 Computational appendix

For simplicity, let us think of a sourcing strategy $\mathbf{D}$ as a $1 \times N$ row vector instead of a set though each element of $\mathbf{D}$ is still either zero or one. Then we can treat a list of M sourcing strategies as $M \times N$ matrix. Furthermore, to distinguish a single sourcing strategy from a set of sourcing strategies, I use the math font $\mathbb{D}$ to denote a $M \times N$ matrix and the font $\mathbf{D}$ to denote a $1 \times N$ row vector.

Step 1: Start with $\mathbf{D}_0(s) = \inf(\mathbf{D}) = \{0, ..., 0\}$. Define a sequence $\{\mathbb{D}_t(s)\} : \mathbb{D}_1(s) = W(\mathbf{D}_0(s))$ and $\mathbb{D}_{t+1}(s) = W(\mathbb{D}_t(s))$ where $W(\mathbb{D}_t(s))$ means apply the mapping $W(.)$ to all the row vectors $\mathbf{D}_t(s) \in \mathbb{D}_t(s)$ such as

$$W(\mathbf{D}_t(s)) = W(\begin{bmatrix} D_t^1(s) & D_t^2(s) & D_t^3(s) & ... & D_t^N(s) \end{bmatrix})$$
$$= \begin{bmatrix} V^1(\mathbf{D}_t(s)) & D_t^2(s) & D_t^3(s) & ... & D_t^N(s) \\ D_t^1(s) & V^2(\mathbf{D}^t(s)) & D_t^3(s) & ... & D_t^N(s) \\ D_t^1(s) & D_t^2(s) & V^3(\mathbf{D}_t(s)) & ... & D_t^N(s) \\ & & ... & & \\ D_t^1(s) & D_t^2(s) & D_t^3(s) & ... & V^N(\mathbf{D}_t(s)) \end{bmatrix},$$

where

$$V^m(\mathbf{D}_t(s)) = 1\big[B[(\sum_{l \neq m} D_t^l(s)\xi_l + \xi_m)^{\frac{\sigma-1}{\theta}} - (\sum_{l \neq m} D_t^l(s)\xi_l)^{\frac{\sigma-1}{\theta}}] - f_m(s) \geq 0\big].$$

In order to avoid unnecessary estimation, extract row vectors in $W(\mathbb{D}_t(s))$ that are unique and do not belong to $\mathbb{D}_t(s)$ to construct the new matrix $\mathbb{D}_{t+1}(s)$. The iteration process stops whenever $W(\mathbb{D}_T(s)) = \emptyset$, which means changing any element in $\mathbb{D}_T(s)$ from zero to one will not improve the profits.

Step 2: Evaluate all row vectors that belong to $\{D_0(s), \mathbb{D}_1(s), ..., \mathbb{D}_T(s)\}$ to find the profit-maximizing vector $\mathbf{D}_*(s)$. On average, there are around 522 possible choices for a simulated Chinese firm.

Intuitively, starting from the set $\inf(\mathbf{D}) = \{0, ..., 0\}$ which contains no sourcing country, applying the operator V^m to each country and assess whether it makes more profits to add this country individually. This process leaves us a few countries such that outsourcing to each of these countries separately yields higher profits than relying on self-made inputs only. Denote strategies such that firms offshore to one of those countries as as $\mathbb{D}_1(s)$. Next, apply the operator V^m to each sourcing strategy in $\mathbb{D}_1(s)$ and test whether outsourcing to two countries makes more profits than outsourcing to either of them. If so, let this strategy containing two countries belongs to $\mathbb{D}_2(s)$. Let this process continue until applying V^m to any sourcing strategy in $\mathbb{D}_N(s)$ makes no more differences.

Bibliography

[1] Almeida and Fernandes (2008), "Openness and Technological Innovations in Developing Countries: Evidence from Firm Level Surveys", *Journal of Development Studies*, 44(5), 701-727.

[2] Antràs, Fort and Tintelnot (2017), "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms", *American Economic Review*, 107(9), 2514-2564.

[3] Antràs, Garicano, and Rossi-Hansberg (2006), "Offshoring in a Knowledge Economy", *Quarterly Journal of Economics*, 121(1), 31-77.

[4] Arkolakis and Eckert (2017), "Combinatorial Discrete Choice", Working Paper.

[5] Aw, Roberts, and Xu (2011), "R&D Investment, Exporting, and Productivity Dynamics", *American Economic Review*, 101(4), 1312-1344.

[6] Balassa (1963), "An Empirical Demonstration of Classical Comparative Cost Theory", https://www.nber.org/papers/w15902.pdf.

[7] Barro and Lee (2010), "A New Data Set of Educational Attainment in the World, 1950 - 2010", *Journal of International Economics*, 95(2),250-262.

[8] Bas and Strauss-Kahn (2015), "Input-trade Liberalization, Export Prices and Quality Upgrading", *Journal of International Economics*, 95(2),250-262.

[9] Bloom, Draca, and Reenen (2016), "Trade Induced Technical Change: the Impact of Chinese Imports on Innovation, Diffusion and Productivity", *Review of Economic Studies*, 83(1), 87-117.

[10] Boler, Moxnex and Ulltveit-Moe (2016), "R&D, International Sourcing, and the Joint Impact on Firm Performance", *American Economic Review*, 105(12), 3704-3739.

[11] Bond, and Samuelson (1999), "Strategic Behavior and the Rules for International Taxation of Capital", *Economic Journal*, 109-111.

[12] Bustos (2011), "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms", *American Economic Review*, 101(1), 304-340.

[13] Chan (2017), "How Substitutable Are Labor and Intermediates?", Working Paper.

[14] Chaney (2008) "Distorted Gravity: The Intensive and Extensive Margins of International Trade", *American Economic Review*, 98(4), 1707-1721.

[15] Chen, Zhang and Zheng (2017), "Import and Innovation: Evidence from Chinese Firms", *European Economic Review*, 94, 205-220.

[16] Costinot, Donaldson, Vogel, and Werning (2015), "Comparative Advantage and Optimal Trade Policy", *Quarterly Journal of Economics*, 130, 659-702.

[17] Costinot, Lorenzoni, and Werning (2014), "A Theory of Capital Controls as Dynamic Terms-of-Trade Manipulation", *Journal of Political Economy,*, 122, 77-128.

[18] Costinot, Rodriguez-Clare, and Werning (2016), "Micro to Macro: Optimal Trade Policy with Firm Heterogeneity", *Working Paper*.

[19] Dang and Motohashi (2015), "Patent Statistics: A Good Indicator for Innovation in China? Patent Subsidy Program Impacts on Patent Quality", *China Economic Review*, 35, 137-155.

[20] Davies, and Gresik (1998), "Tax Competition and Foreign Debt", *Mideo*.

[21] Demidova, and Rodriguez-Clare (2009), "Trade Policy Under Firm-Level Heterogeneity in a Small Economy", *Journal of International Economics*, 78,100-112.

[22] Eaton and Kortum (2002), "Technology, Geography, and Trade", *Econometrica*, 70(5), 1741-1779.

[23] Fan, Li and Yeaple (2015), "Trade Liberalization, Quality, and Export Prices", *Review of Economics and Statistics*, 97(5), 1033-1051.

[24] Felbermayr, Jung, and Larch (2013), "Optimal Tariffs, Retaliation and the Welfare Loss from Tariff Wars in the Melitz Model", *Journal of International Economics*, 89, 13-25.

[25] Goldberg, Khandelwal, Pavcnik and Topalova (2010), "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India", *Quarterly Journal of Economics*, 125(4), 1727-1767.

[26] Gresik (2001), "The Taxing Task of Taxing Transnational", *Journal of Economic Letter*, 39, 800-838.

[27] Grieco, Li and Zhang (2018), "Input Prices, Productivity and Trade Dynamics: Long-run Effects of Liberalization on Chinese Paint Manufacturers", Working Paper.

[28] Grossman, Helpman, and Kircher (2015), "Matching, Sorting, and the Distributional Effects of International Trade", *Journal of Political Economics*, 125,224-264.

[29] Heston, Summers and Aten (2011), Penn World Table Version 7.0. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, Philadelphia, PA.

[30] Hu and Jefferson (2009), "A Great Wall of Patents: What is Behind China's Recent Patent Explosion", *Journal of Development Economics*, 90(1), 57-68.

[31] Hummels, Jorgensen, Munch, and Xiang (2014), "The Wage and Employment Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", *American Economic Review*, 104(6), 1597-1629.

[32] Jia (2008), "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry", *Econometrica*, 76(6), 1263-1316.

[33] Li (2012), "Behind the Recent Surge of Chinese Patenting: an Institutional View", *Research Policy*, 41, 236-249.

[34] Li and Whalley (2013), "China and the Trans-Pacific Partnership: A numerical Simulation Assessment of the Effects Involved", *World Economy*, December.

[35] Lileeva and Trefler (2010), "Improved Access to Foreign Markets Raises Plant-Level Productivity ... for Some Plants", *Quarterly Journal of Economics*, 125(3), 1051-1099.

[36] Liu and Qiu (2016), "Intermediate Input Imports and Innovations: Evidence from Chinese Firms' Patent Filings", *Journal of International Economics*, 103, 166-183.

[37] Melitz (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71(6), 1695-1725.

[38] Oostendorp (2005), "The Standardized ILO October Inquiry 1983-2003", http://www.nber.org/oww/Technical_document_1983-2003_standardizationv3.pdf.

[39] Ossa (2011), "A 'New Trade' Theory of GATT/WTO Negotiations", *Journal of Political Economy*, 119, 122-152.

[40] Peters, Roberts, Vuong, and Fryges (2017), "Estimating Dynamics R&D Choice: An Analysis of Costs and Long-run Benefits", *RAND Journal of Economics*, 48(2), 409-437.

[41] Roberts and Tybout (1997), "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs", *American Economic Review*, 87(4), 545-564.

[42] Rust (1997), "Using Randomization to Break the Curse of Dimensionality", *Econometrica*, 65(3), 487-516.

[43] Venables (1987), "Trade and Trade Policy with Differentiated Products: A Chamberlainian-Ricardian Model", *The Economic Journal*, 97, 700-717.

Vita

Wenzhuo Lu

Education

- Ph.D. in Economics, The Pennsylvania State University, 2013-2019

- B.A. in Mathematics and Economics (with distinction), University of Virginia, 2010-2013

Research Fields

- Primary field: International Trade
- Secondary fields: Industrial Organization, Economic Growth

Research Papers

- Tax, Offshoring, and Income Distributions
- Globalization, Innovation, and Firm Dynamics

Research in Progress

- "The Linkage between Export and Import"

Teaching Experience

- Summer 2017: Instructor for Introduction to Econometrics (online), The Pennsylvania State University
- 2013-2018: TA for Industrial Organization, Public Economics, Intermediate Microeconomics, and Intermediate Macroeconomics, The Pennsylvania State University

Skills

- Language: Chinese (native), English (fluent)
- Computer: Matlab, Python, Stata