Thesis on Industrial Organization and Development of China

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Jing Fang

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Prof. Thomas J. Holmes

August, 2011

Acknowledgments

I would like to express my sincere thanks to my advisor Thomas Holmes for his constant support and encouragement. I would like to thank my committee members Timothy Kehoe, Erzo Luttmer, Amil Petrin, Rodney Smith and Joel Waldfogel for their suggestions that greatly improved the paper. I would also like to thank Patrick Bajari and Robert Town for their suggestions, in addition to Jose Asturias, Konstantin Golyaev, Yenan Hou, Christina Marsh, Nan Xu, and all participants of the University of Minnesota Applied Microeconomics Workshop.

Dedication

This dissertation is dedicated to my parents, Lizhong Fang and Shuangqin Cai. I am deeply indebted to them for their continued support and unwavering faith in me.

Contents

List of Tables

List of Figures

1 Introduction

- Chapter 2 develops a theory for why a big country, like China, might be special when it comes to pursuing strategic trade policy.
- Chapter 3 studies the rise of patenting in China, specifically, the rise and dispersion of Chinese inventors.
- Chapter 4 concludes and discuss the future research.

2 Country Size and Strategic Trade Policy

2.1 Introduction

Perhaps the most controversial issue today regarding trade policy concerns China and its exports. There is wide agreement that China is pursuing a policy to promote exports by manipulating its exchange rate to subsidize exports and tax imports. If China is behaving strategically it begs the question: Why is it special? Why don't we hear similar accusations being leveled at small countries such as Vietnam or Thailand?

This paper develops a theory for why a particular country might be special when it comes to pursuing strategic trade policy. It considers an environment where countries are identical, except in terms of their scale. There is one exporting country (think of this as "China") that is an enormous proportional expansion of all the other exporting countries. On a per capita basis the exporting countries are the same in endowments.

Given its enormous size, the large country behaves strategically, and internalizes the impacts of its actions, in ways that the smaller exporting countries do not.

In this paper there is a set of exporting countries including the large country that sells to another set of importing countries. In principle, with the large country acting as a dominant firm, we might see that it holds back its exports, relative to the small exporting countries, to maintain high export prices. However, suppose there is some kind of learning by doing (following the general idea of Arrow (1962)) that operates at the country level and that depends upon massive scale to be operative. Then, to exploit these scale economies the large country has an incentive to promote exports. As the small countries are not in a position to attain massive scale, the incentive for them to promote exports to exploit learning by doing is reduced. If these scale economies are external to the firm, then the large country will adopt export subsidies to promote exports. These subsidies will raise the real exchange rate for the large country compared to the small exporting countries.

This paper develops a formal model that captures this story. It makes explicit the assumptions that deliver the result that exports are higher in the large exporting country, compared to the small exporting countries, and the result that the real exchange rate is higher. The small exporting countries modeled as "fringe firms"[1] as in the dominant firm model used in industrial organization, while the large country is a "dominant firm." (Like in Riordan (1998) and Gowrisankaran and Holmes (2004)). The paper takes this industrial organization structure and embeds it into a trade model with three goods. Two goods are tradable and what we call the exporting countries sell good Y in return for good X. The third good is nontradable. With the nontradable good in the model, it is possible to analyze differences in real exchange rates.

This paper has some difference with the previous literature. Beginning with Brander and Spencer (1985), there is a large literature on strategic trade theory, like

---

[1] In the trade literature, they usually call it a "small open economy".

Brander (1995) and Bagwell and Staiger (2001). There is also huge literature on the question that why the export country can manipulate the exchange rate and push export, starting from Krugman (1986), and followed by Knetter (1989), Knetter (1993), Goldberg and Knetter (1997), Goldberg and Hellerstein (2008), Atkeson and Burstein (2007) and Atkeson and Burstein (2008). But all of them do not explain the question I asked above: why is a big exporter special? This paper tries to answer this question and contributes to this literature. It is different from what has been previously done in several respects.

The first difference is that the main focus of this paper is competition between countries that are asymmetric in size, such as China compared to Vietnam. In the prototypical model of the literature, such as in Brander and Spencer, competing countries are symmetric. A classic example is the Baldwin and Krugman (1988) analysis of competition in wide-bodied aircraft. This is a model of Europe through Airbus competing with the United States through Boeing. A symmetric model is a reasonable starting place when comparing the United States and Europe, and comparing Boeing and Airbus. But now let's turn to a comparison of China and Vietnam. China is on the order of fifteen times as big as Vietnam is in terms of population. It is a dramatic asymmetry such as this that paper aims to capture through the use of its dominant firm/competitive fringe structure.

The second difference from typical strategic trade analysis is that the scale economies that are modeled are external to the firm but internal to the countries. For Airbus versus Boeing analysis, it may be sufficient to rule out external economies. However for China, there are good reasons to think there are external economies. Usually, the technology improvement is spillover from one firm to other firms, by trading the blueprint or patent of new technology. If a firm in a small country invents some new skill, it will easily spread out to other firms in other countries, but this is not true for a big country like China. According to Hessler (2010), in some industries, the first company imported a European-made machine to his shop in Fujian, a province in southeast of China. An unskilled worker, became an expert in the maintenance of the machine and secretly created a detailed blueprint of the machine. After that, he moved to Guangdong, a province nearly seven hundred miles away from Fujian, and custom built the machine for a series of other companies. Clearly, the fruit of learning by doing travels seven hundred miles, but remained in China. This distance is similar to the one from Vietnam to Thailand, and it even crosses Cambodia or Laos. Of course, this story is not the usual way that the new technology spillover occurs, but it is a good example to show that the scale factor of economics is very important in my analysis, and this factor is external to firms but internalized by the big country. There is also several literature (Backus et al. (1992), Keller (2002) and Keller (2004)) provides the evidence that the benefits from technology spillover are related with distance, which again proves that the scale is the key factor at the country level.

The third difference is that this paper explicitly incorporates nontradable goods, (Like in Chipman (2007)) so it is conceptually possible to include a discussion of the real exchange rate in the analysis. Other papers may also incorporate the nontradable goods but not use this structure to analyze the change of real exchange rate.

In this paper, I found that, with the learning-by-doing effect, the big country exports more in both time periods, comparing the no learning-by-doing case. The real exchange rate is higher and increases as the learning-by-doing effect goes up. This is mainly because, with learning-by-doing, inputs move to the export sector and significantly increase the price of non-tradable goods. The utility of big exporter and importers increases dramatically but remains nearly the same for small exporters, since they take the price of export goods as given.

The rest of the paper is organized as follows. Section 2 provides some background information about China's trade policies and in particular compares it to Thailand. Section 3 describes the model and characterizes the equilibrium. Section 4 provides some numerical findings. Section 5 concludes.

2.2 Export growth about Trade Policies

Beginning in 1978, China started its economic reform and opening up to the world. It tried to build an export-oriented economy to stimulate economic growth. To illustrate the explosion in trade, I take two industries, textile and toys and three export destinations, U.S., Japan, and European Union as examples, as shown in Figure 2.1 and Figure 2.2.

Exports from China in these two industries doubled over this period. In some countries, like United States, the import of textiles is nearly four times as big as ten years ago. Comparing with exports from Thailand, we can see clearly that its export to the U.S. did not increase very much. How could this huge difference happen?

Usually, there are two ways to promote exports. The first and well-known method is subsidies. The New York Times once reported of a solar panel firm in Hunan province, China (Bradsher (2010)). This firm received subsidies from the local government instead of the central government, because the subsidies from the central government is not allowed by WTO rules. This is a very typical example of this method.

Figure 2.1: Textile Exports

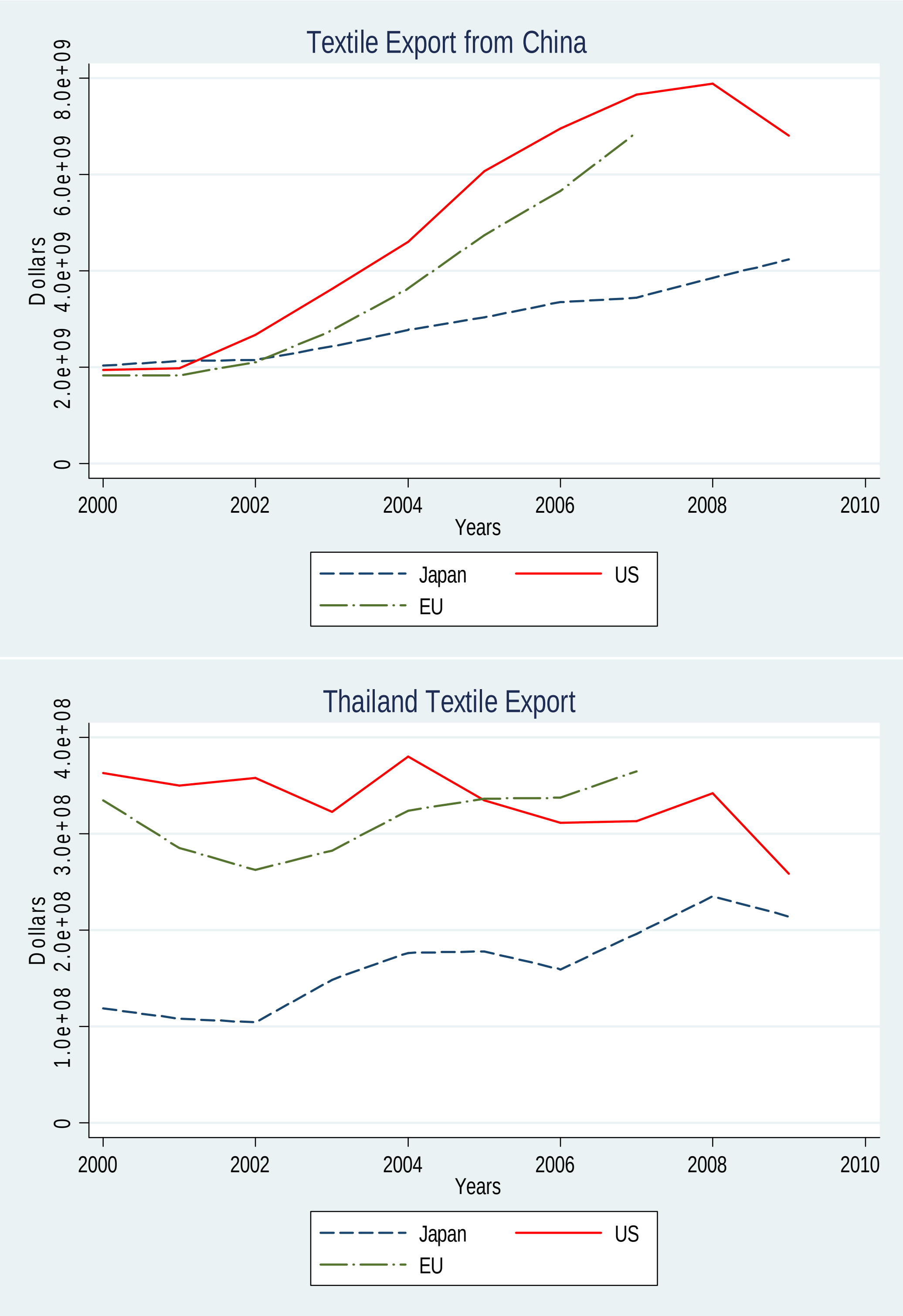
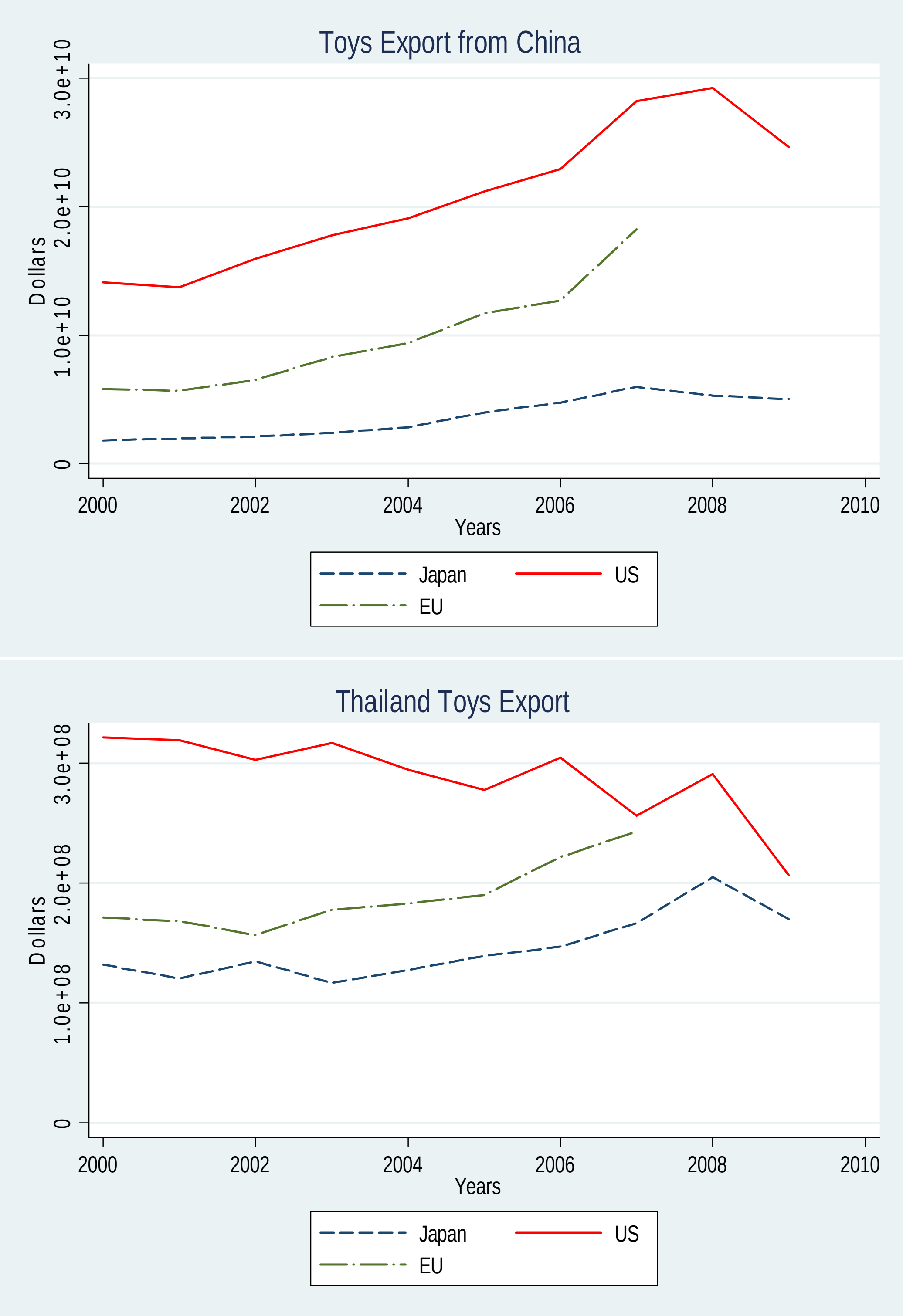
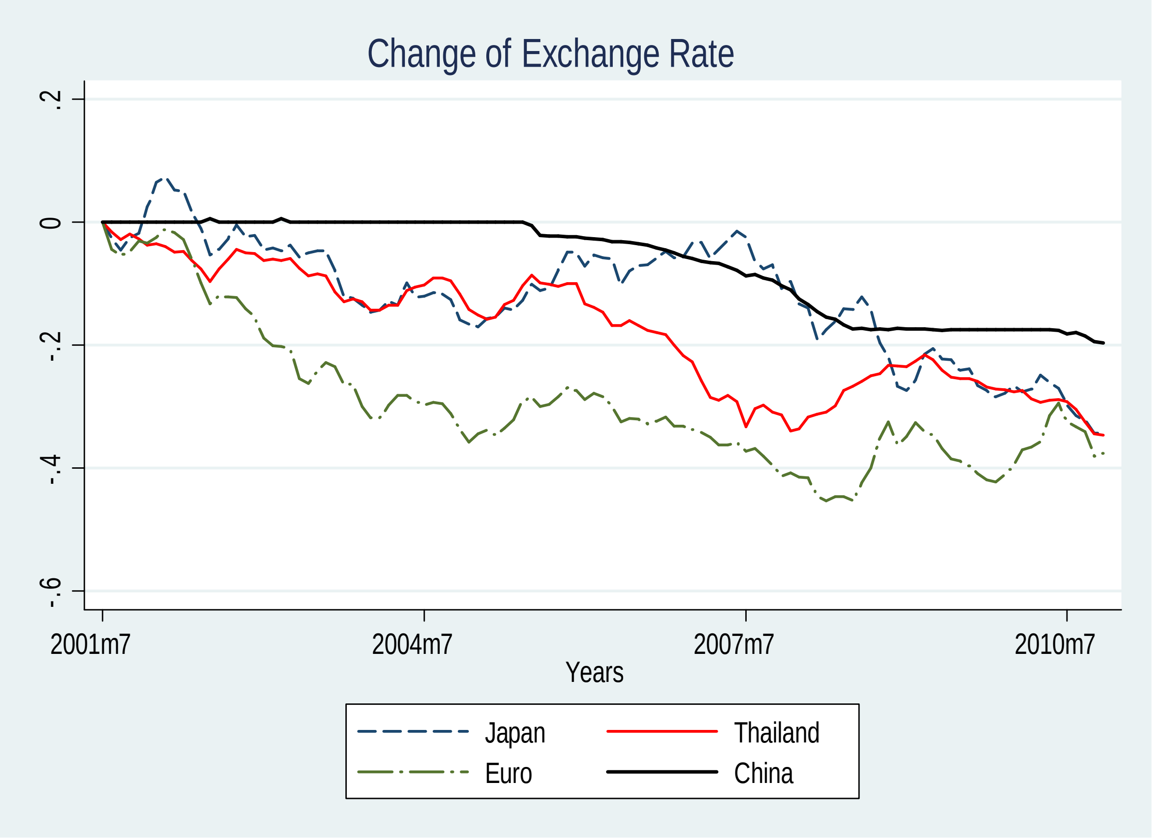


Figure 2.2: Toy Exports



The other way to do so is to manipulate the exchange rate. Here I present the change of exchange rate from 2001 to the present in Figure 2.3.

Figure 2.3: Exchange Rate



This graph shows the change of how much we can exchange 1 U.S. dollar. So when the curve goes up, it means U.S. dollar appreciated to this currency, and vice versa. In these ten years, it is clear that U.S. dollar depreciated a lot relative to some other currency, like euro or Japanese yen, but not for Chinese Yuan. China kept its rate pegged for the most of time. When U.S. dollar depreciated, so did Chinese Yuan. This is obviously helpful for the export firms in China. On the other hand, I use the Thai baht as a comparison. When U.S dollars depreciated to other currencies, it also depreciated to Thai baht. That means, the Thai baht appreciated and of course, this did not help export firms in Thailand.

2.3 Model

In this part, I want to present my model. It is a two-period model, with a learning-by-doing effect in the second period. First, I present a simple version, in which the consumer only cares about the first period, so the learning-by-doing effect doesn't influence consumers behavior. I want to compare my model's result to the first simple version, to see the difference between these two versions. I believe the two-period version is more helpful in explaining a big exporter's behavior.

In my model, I assume there is a population of N in the world. Every consumer, with 1 unit of labor, owns a firm. The firm could produce tradable goods X, Y and non-tradable goods Z.[2] Every firm has the same technology and for n units of labor, they can produce n unit of X, or $f(n)$ unit of Y or n unit of Z. (Suppose $f' > 0$ and $f'' < 0$, $\lim_{y\to 0} f(y) = +\infty$ and $\lim_{y\to +\infty} f(y) = 0$). All of them are exporters and sell some Y to get X back. I normalize the price of X to 1. Price of Y is $P_y > 1$ and price of Z is decided by firms in their countries.

The consumer has the utility function:

$$U = x_{c1}^{\alpha_x} y_{c1}^{\alpha_y} z_{c1}^{\alpha_z} + \rho x_{c2}^{\alpha_x} y_{c2}^{\alpha_y} z_{c2}^{\alpha_z}$$

where ρ is the discount between periods, x_{ci}, y_{ci}, z_{ci} is the consumption of that goods in period i.

At the country level, the total measure of export countries is 1. There is a big

---

[2] I will use the uppercase letters to denote the name of goods and lowercase letters to quantity of that goods.

country, with measure μ of land. The rest of the exporting countries together have $1-\mu$ land. Each small exporting country is measure zero.

2.3.1 One-period version

Here I want to present a simple version of my model, as a baseline to compare. In this version, consumers only care about the first period, that is, $\rho = 0$.

For a small country, since it is really small, there is some competitive firms in it, and the firms could decide the p_{zs1}, and also they are competitive in the whole world export market. This means, the firms will take p_{y1} as given. The representative consumer supplies the labor that the firm need and owns this firm, so the consumer's income is equal to the wage plus the profit of firm. Since wage is the cost of the firm, the consumer will only care about the total revenue of his/her firm. In fact, at the location level, this is a social planner's problem. The planner needs to decide how much labor should be supplied to produce Y (n_{ys1}) and to Z (n_{zs1}), and how many Y (y_{se1}) should be sold to trade X $(p_{y1}y_{se1})$ back.

So formally, the planner's problem is:

$$\max_{n_{sy1},y_{se1}} U = x_{sc1}^{\alpha_x} y_{sc1}^{\alpha_y} z_{sc1}^{\alpha_z} \tag{2.1}$$

where

$$\begin{aligned} x_{sc1} &= p_{y1}y_{se1} \\ y_{sc1} &= f(n_{sy1}) - y_{se1} \\ z_{sc1} &= 1 - n_{sy1} \end{aligned}$$

such that

$$\begin{aligned} 0 \leqslant n_{sy1} \leqslant 1 \\ y_{se1} \leqslant f(n_{sy1}) \end{aligned}$$

If I know the form of production function $f(n_y)$, for example, assume $f(n_y) = \beta n_y^{\lambda}$, where $0 < \lambda < 1$, then I could get first order condition and solve it.

F.O.C

$$\begin{aligned} \alpha_y(1-n_{sy1})\beta\lambda n_{sy1}^{\lambda-1} &= \alpha_z(\beta n_{sy1}^{\lambda} - y_{se1}) \\ \alpha_x(\beta n_{sy1}^{\lambda} - y_{se1})p_{y1} &= \alpha_y p_{y1} y_{se1} \end{aligned}$$

Solve it, and we get

$$\begin{aligned} n_{sy1} &= \frac{\lambda(\alpha_x+\alpha_y)}{\lambda(\alpha_x+\alpha_y)+\alpha_z} \\ y_{se1} &= \frac{\alpha_x}{\alpha_x+\alpha_y}\beta n_{sy1}^{\lambda} \end{aligned}$$

Since the ratio of marginal utility should be equal to the ratio of price of that good, i.e.

$$\frac{U_x}{U_y} = \frac{1}{p_y}, \frac{U_x}{U_z} = \frac{1}{p_z}, \text{ and } \frac{U_z}{U_y} = \frac{p_z}{p_y}$$

we could solve p_{zs1}:

$$p_{zs1} = \frac{U_z}{U_x} = \frac{\alpha_z x_{sc1}}{\alpha_x z_{sc1}}$$

For the big country, it has a large area of land with measure of μ. It has more labor endowment and more firms in this country. Of course it can produce much more to export. So obviously, this big country can decide the price level (p_{y1}) as it wants. All the rest small countries could only take the price (p_{y1}) as given and solve the question (Equation 2.1).

On the other hand, the big country knows its followers response to different price and will solve a similar problem as small ones do and get n_{by1}, n_{bz1} and y_{be1}. Here I want to write down the big country's problem.

I also need to know the outside demand for the export tradable good Y to do this, so I assume the outside demand for Y is a constant elasticity demand: $p_y = cQ_d^{\varepsilon}$ $(c > 0, -1 < \varepsilon < 0)$. And in the equilibrium, the quantity demand should be equal to the quantity of export: $Q_d = \mu N y_{be1} + (1-\mu) N y_{se1}$. On the firm side, again assume that $f(n_y) = \beta n_y^{\lambda}$. Remember that y_{se1} is a function of p_{y1} and now we can see that p_{y1} is a function of y_{be1} and y_{se1}, so we know that y_{se1} is a function of y_{be1}. This function $y_{se1}(y_{be1})$ is the response function of the small countries.

The planner's problem of the big country is similar to (Equation 2.1):

$$\max_{n_{by1}, y_{be1}} U = x_{bc1}^{\alpha_x} y_{bc1}^{\alpha_y} z_{bc1}^{\alpha_z}$$

where

$$\begin{aligned}
x_{bc1} &= p_{y1}y_{be1} \\
&= c[\mu N y_{be1} + (1-\mu)N y_{se1}(y_{be1})]^{\varepsilon} y_{be1} \\
y_{bc1} &= \beta n_{by1}^{\lambda} - y_{be1} \\
z_{bc1} &= 1 - n_{by1}
\end{aligned}$$

such that

$$0 \leqslant n_{by1} \leqslant 1$$

$$y_{be1} \leqslant \beta n_{by1}^{\lambda}$$

F.O.C

$$\begin{aligned}
\alpha_y(1-n_{by1})\beta\lambda n_{by1}^{\lambda-1} &= \alpha_z(\beta n_{by1}^{\lambda} - y_{be1}) \\
\alpha_x(\beta n_{by1}^{\lambda} - y_{be1})[p_{y1} + y_{be1}p_{y1}'(y_{be1})] &= \alpha_y p_{y1} y_{be1}
\end{aligned}$$

I cannot solve for an analytical solution. But I will try to get some numerical results in next part. Mainly, I want to use this one-period model as a baseline, so that I can compare the two-period model result with this and to see which one is better for fitting the real world situation and why.

2.3.2 Two-period version

Now I present my model, where $\rho > 0$, so all the countries face a two-period problem. And all of them have some learning-by-doing effect. Here I always assume that the spillover of technology could not be prevented from one firm to another. The workers of small countries can easily spread the technology improvement to firms in other countries. However, workers in the big country can only spread knowledge to firms inside the big country.

The second period production will be multiplied by an extra learning-by-doing effect parameter A, where

$$A = g(\int_{countryarea} (TotalOutput)dl).$$

For small countries, firms still produces Y and Z, and trade some Y for X. Firms in the country could decide the price of non-tradable good at each period, p_{zsi}. But since it is still a price taker in the world export market, the firms will take p_{yi} as given for any period. Again, the consumer's income is still equal to the wage plus the profit of firm and therefore they will only care about the total revenue of his/her firm.

So in each period, the planner still need to decide how much labor (n_{syi}) should be used to produce Y and rest (n_{szi}) to Z, and how many Y (y_{sei}) should be sold to trade X ($p_{yi}y_{sei}$) back.

More formally, the question is:

$$\max_{n_{syi},y_{sei}} U = x_{sc1}^{\alpha_x} y_{sc1}^{\alpha_y} z_{sc1}^{\alpha_z} + \rho x_{sc2}^{\alpha_x} y_{sc2}^{\alpha_y} z_{sc2}^{\alpha_z}$$

where

$$
\begin{aligned}
x_{si} &= p_{yi}y_{sei} \\
y_{s1} &= f(n_{sy1}) - y_{se1} \\
y_{s2} &= Af(n_{sy2}) - y_{se2} \\
&= g(\int_{countryarea} f(n_{sy1})dl)f(n_{sy2}) - y_{se2} \\
z_{si} &= 1 - n_{syi}
\end{aligned}
$$

such that:

$$
\begin{aligned}
0 \leqslant n_{syi} \leqslant 1 \\
y_{se1} \leqslant f(n_{sy1}) \\
y_{se2} \leqslant Af(n_{sy2})
\end{aligned}
$$

Again, I need to assume that $f(n_y) = \beta n_y^{\lambda}$. And for $g(x)$, I assume:

$$
g(x) = \begin{cases} 1+\delta \text{ if } x < \beta \\ (1+\delta)(1+\ln(1-\beta+x)) \text{ if } x \geqslant \beta \end{cases}
$$

This functional form means that there is a threshold level to export. If a country's exports are less than that level, the learning-by-doing effect will be small. On the other hand, if exports are more than that level, the learning-by-doing effect will be large, and increasing as the exporting quantity increases.

The learning-by-doing effect has this functional form for several reasons. First, every country should have some learning-by-doing effect, regardless its size and how much it produces. Thus the learning-by-doing effect parameter $A = g(x)$

should be bigger than 1 across countries. Second, small producers have a small and constant learning-by-doing effect. It is small because they cannot learn much from production, and it is constant because of spillovers across small countries.

But for the big country, the learning-by-doing effect is large and increasing as the production and exports go up. This is because, the big country is big enough to internalize the learning-by-doing effect.

Now go back and look at the small countries learning-by-doing effect. The small country only has an area with measure 0, so $g(\int_{countryarea} f(n_{sy1})dl) = g(0) = 1+\delta$, since small country's production does not exceed the threshold level.

Now we can see that, the small countries consumers don't have any dynamic decision to make. They solve the maximization problem in each period, so I can restate the small country's problem as following:

In the first period, consumers solve:

$$\max_{n_{sy1}, y_{se1}} U = x_{sc1}^{\alpha_x} y_{sc1}^{\alpha_y} z_{sc1}^{\alpha_z}$$

where

$$\begin{aligned} x_{sc1} &= p_{y1} y_{se1} \\ y_{sc1} &= \beta n_{sy1}^{\lambda} - y_{se1} \\ z_{sc1} &= 1 - n_{sy1} \end{aligned}$$

such that:

$$0 \leqslant n_{sy1} \leqslant 1$$

$$y_{se1} \leqslant \beta n_{sy1}^{\lambda}$$

And in the second period, the consumers solve:

$$\max_{n_{sy2},y_{se2}} U = \rho x_{sc2}^{\alpha_x} y_{sc2}^{\alpha_y} z_{sc2}^{\alpha_z}$$

where

$$\begin{aligned}
x_{sc2} &= p_{y2} y_{se2} \\
y_{sc2} &= A\beta n_{sy2}^{\lambda} - y_{se2} \\
&= (1+\delta)\beta n_{sy2}^{\lambda} - y_{se2} \\
z_{sc2} &= 1 - n_{sy2}
\end{aligned}$$

such that:

$$\begin{aligned}
0 \leqslant n_{sy2} \leqslant 1 \\
y_{se2} \leqslant (1+\delta)\beta n_{sy2}^{\lambda}
\end{aligned}$$

Solve this question and I could get:

$$\begin{aligned}
n_{sy1} &= \frac{\lambda(\alpha_x+\alpha_y)}{\lambda(\alpha_x+\alpha_y)+\alpha_z} \\
y_{se1} &= \frac{\alpha_x}{\alpha_x+\alpha_y}\beta n_{sy1}^{\lambda} \\
n_{sy2} &= \frac{\lambda(\alpha_x+\alpha_y)}{\lambda(\alpha_x+\alpha_y)+\alpha_z} \\
y_{se2} &= \frac{\alpha_x}{\alpha_x+\alpha_y}(1+\delta)\beta n_{sy1}^{\lambda}
\end{aligned}$$

Similar to the one-period model, we could

$$p_{zsi} = \frac{U_z}{U_x} = \frac{\alpha_z x_{sci}}{\alpha_x z_{sci}}$$

For the big country, it has a large area of land with measure of μ and sets the price level of tradable goods (p_{yi}). The big country knows small ones response of different price, $y_{sei}(p_{y1}, p_{y2})$ and will solve the similar problem as small ones and get n_{byi}, n_{bzi} and y_{bei}.

Again I assume the outside demand for Y is still $p_y = cQ_d^{\varepsilon}$ $(c > 0, -1 < \varepsilon < 0)$ and the quantity demand is $Q_d = \mu N y_{bei} + (1-\mu)N y_{sei}$. On the firm side, assume that $f(n_y) = \beta n_y^{\lambda}$ again, but the learning-by-doing effect for the big country is

$$\begin{aligned} A &= g(x) = g(\int_{countryarea} TotalOutput\ dl) \\ &= g(\int_{countryarea} \mu N f(n_{by1})) \\ &= g(\mu^2 N f(n_{by1})) \\ &= (1+\delta)(1+\ln(1-\beta+\mu^2 N f(n_{by1}))) \end{aligned}$$

The big country has an area (measure μ) and population (N) that is big enough to make big country's production cross the threshold level. Again, notice that y_{sei} is a function of p_{yi} and we can see now, that p_{yi} is a function of y_{bei} and y_{sei}, so we know that y_{sei} is a function of y_{bei}.

Then planner's problem for the big country is:

$$\max_{n_{byi}, y_{bei}} U = x_{bc1}^{\alpha_x} y_{bc1}^{\alpha_y} z_{bc1}^{\alpha_z} + \rho x_{bc2}^{\alpha_x} y_{bc2}^{\alpha_y} z_{bc2}^{\alpha_z}$$

where

$$\begin{aligned}
x_{bci} &= p_{yi}y_{bei} \\
&= c[\mu N y_{bei} + (1-\mu)N y_{sei}(y_{bei})]^{\varepsilon} y_{bei} \\
y_{bc1} &= f(n_{by1}) - y_{be1} = \beta n_{by1}^{\lambda} - y_{be1} \\
y_{bc2} &= Af(n_{by2}) - y_{be2} \\
&= g(\int_{countryarea} f(n_{by1})dl)f(n_{by2}) - y_{be2} \\
&= [(1+\delta)(1+\ln(1-\beta+\mu^2 N f(n_{by1})))]f(n_{by2}) - y_{be2} \\
&= [(1+\delta)(1+\ln(1-\beta+\mu^2 N \beta n_{by1}^{\lambda}))]\beta n_{by2}^{\lambda} - y_{be2} \\
z_{bci} &= 1 - n_{byi}
\end{aligned}$$

such that

$$0 \leqslant n_{byi} \leqslant 1$$

$$y_{be1} \leqslant \beta n_{by1}^{\lambda}$$

$$y_{be2} \leqslant (1+\delta)(1+\ln(1-\beta+\mu^2 N \beta n_{by1}^{\lambda}))]\beta n_{by2}^{\lambda}$$

I can only solve for the optimal n_{byi}, n_{bzi} and y_{bei} numerically. I will show this in the next part.

2.4 Numerical Result

In this part, I want to show some numerical results to see the difference between these two versions. I expect that in the one-period setting, the big exporter should hold back the quantity of production to increase the price. In the two-period version, since there is the learning-by-doing effect, even a big exporter will increase its quantity and lower the price. I will also do a counterfactual analysis at the end of this section, to analyze the effect of an export quota on the big country.

In the utility function, I assume that $\alpha_x = 0.3, \alpha_y = 0.4$ and $\alpha_z = 0.3$. And in the two-period version, consumers have the same weight on the first and second period, that is, $\rho = 1$. So the utility function will be $U = X_1^{\alpha_x} Y_1^{\alpha_y} Z_1^{\alpha_z} + X_2^{\alpha_x} Y_2^{\alpha_y} Z_2^{\alpha_z} = X_1^{0.3} Y_1^{0.4} Z_1^{0.3} + X_2^{0.3} Y_2^{0.4} Z_2^{0.3}$. About the production function, I want to use a Cobb-Douglas production function, $f(k,l) = \beta k^{1-\lambda} l^{\lambda}$. But in my model, there is no capital, so I assume $k = 1$ here and so that $f(n) = \beta n^{\lambda}$. The average share of labor in national income in U.S. is about 0.64. Considering that exporter countries products are generally more labor-intensive, I assume $\beta = 1$ and $\lambda = 0.7$. So the production function is $f(n_y) = \beta n_y^{\lambda} = n_y^{0.7}$. The outside demand from import countries has constant elasticity. And mainly they import necessities, so the absolute elasticity will be lower than 1. So here I assume $c = 10$ and $\varepsilon = -0.5$, and the demand function is $p_y = 10 Q_d^{-0.5}$. Finally, I assume there is a population of 100 in the export world, and half of them belong to the big country. That is $N = 100$ and $\mu = 0.5$.

2.4.1 One-period version

For small countries, the problem is:

$$\max_{n_{sy1}, y_{se1}} U = x_{sc1}^{0.3} y_{sc1}^{0.4} z_{sc1}^{0.3}$$

where

$$\begin{aligned}
x_{sc1} &= p_{y1}y_{se1} \\
y_{sc1} &= \beta n_{sy1}^{\lambda} - y_{se1} \\
&= n_{sy1}^{0.7} - y_{se1} \\
z_{sc1} &= 1 - n_{sy1}
\end{aligned}$$

such that

$$0 \leqslant n_{sy1} \leqslant 1$$

$$y_{se1} \leqslant n_{sy1}^{0.7}$$

For the big country, the question is:

$$\max_{n_{by1}, y_{be1}} U = x_{bc1}^{0.3} y_{bc1}^{0.4} z_{bc1}^{0.3}$$

where

$$\begin{aligned}
x_{bc1} &= p_{y1}y_{be1} \\
&= c[\mu N y_{be1} + (1-\mu) N y_{se1}(y_{be1})]^{\varepsilon} y_{be1} \\
&= 10[50 y_{be1} + 50 y_{se1}(y_{be1})]^{\varepsilon} y_{be1} \\
y_{bc1} &= n_{by1}^{0.7} - y_{be1} \\
z_{bc1} &= 1 - n_{by1}
\end{aligned}$$

such that

$$0 \leqslant n_{by1} \leqslant 1$$
$$y_{be1} \leqslant n_{by1}^{0.7}$$

Solve it and get the result in Table 2.1:

Table 2.1: One Period

	The big country	Small countries
n_{y1}	0.5959	0.6203
n_{z1}	0.4041	0.3797
y_{e1}	0.2554	0.3068
p_{y1}	1.8862	1.8862
p_{z1}	1.1919	1.5237
x_{c1}	0.4817	0.5786
y_{c1}	0.4406	0.4090
z_{c1}	0.4041	0.3797
U	0.4410	0.4439

And real exchange rate (big/small) is:

$$\text{real exchange rate}(RER) = \frac{p_{bz1}\frac{z_{bc1}}{x_{bc1}+y_{bc1}+z_{bc1}} + p_{y1}\frac{y_{bc1}}{x_{bc1}+y_{bc1}+z_{bc1}} + \frac{x_{bc1}}{x_{bc1}+y_{bc1}+z_{bc1}}}{p_{sz1}\frac{z_{sc1}}{x_{sc1}+y_{sc1}+z_{sc1}} + p_{y1}\frac{y_{sc1}}{x_{sc1}+y_{sc1}+z_{sc1}} + \frac{x_{sc1}}{x_{sc1}+y_{sc1}+z_{sc1}}} \approx 0.9591$$

Since I know the demand curve, I can calculate the consumers surplus:

$$\begin{aligned} CS &= \int_0^{Q_d} c * Q^{\varepsilon} dQ \\ &= 53.0171 \end{aligned}$$

2.4.2 Two-period version

I still assume the same parameter values as the one-period version. Now I have the learning-by-doing effect in the second period:

$$A = g(x) = \begin{cases} 1+\delta \text{ if } x < 1 \\ (1+\delta)(1+\ln(0.3+x)) \text{ if } x \geqslant 1 \end{cases}$$

Here I will give every country a 10% increase for the second period if it is below the threshold level, that is $\delta = 0.1$. This is not a very high forecast, considering the rapid growth of technological innovation.

For small countries, in the first period consumers solve:

$$\max_{n_{sy1}; y_{se1}} U = x_{sc1}^{0.3} y_{sc1}^{0.4} z_{sc1}^{0.3}$$

where

$$\begin{aligned} x_{sc1} &= p_{y1} y_{se1} \\ y_{sc1} &= \beta n_{sy1}^{\lambda} - y_{se1} \\ &= n_{sy1}^{0.7} - y_{se1} \\ z_{sc1} &= 1 - n_{sy1} \end{aligned}$$

such that:

$$0 \leqslant n_{sy1} \leqslant 1$$
$$y_{se1} \leqslant n_{sy1}^{0.7}$$

In the second period, consumers solve:

$$\max_{n_{sy2}, y_{se2}} U = \rho x_{sc2}^{\alpha_x} y_{sc2}^{\alpha_y} z_{sc2}^{\alpha_z}$$

where

$$\begin{aligned} x_{sc2} &= p_{y2} y_{se2} \\ y_{sc2} &= A\beta n_{sy2}^{\lambda} - y_{se2} = (1+\delta)\beta n_{sy2}^{\lambda} - y_{se2} \\ &= 1.1 n_{sy2}^{0.7} - y_{se2} \\ z_{sc2} &= 1 - n_{sy2} \end{aligned}$$

such that:

$$0 \leqslant n_{sy2} \leqslant 1$$

$$y_{se2} \leqslant 1.1 n_{sy2}^{\lambda}$$

The big country's problem is:

$$\max_{n_{byi}, y_{bei}} U = x_{bc1}^{0.3} y_{bc1}^{0.4} z_{bc1}^{0.3} + x_{bc2}^{0.3} y_{bc2}^{0.4} z_{bc2}^{0.3}$$

where

$$
\begin{aligned}
x_{bci} &= p_{yi}y_{bei} = c[\mu N y_{bei} + (1-\mu)N y_{sei}(y_{bei})]^{\varepsilon} y_{bei} \\
&= 10[50y_{bei} + 50y_{sei}(y_{bei})]^{\varepsilon} y_{bei} \\
y_{bc1} &= f(n_{by1}) - y_{be1} = \beta n_{by1}^{\lambda} - y_{be1} = n_{by1}^{0.7} - y_{be1} \\
y_{bc2} &= Af(n_{by2}) - y_{be2} = g(\int_{countryarea} f(n_{by1})dl) f(n_{by2}) - y_{be2} \\
&= [(1+\delta)(1+\ln(1-\beta+\mu^2 N\beta n_{by1}^{\lambda}))]\beta n_{by2}^{\lambda} - y_{be2} \\
&= [1.1(1+\ln(25n_{by1}^{0.7}))]n_{by2}^{0.7} - y_{be2} \\
z_{bci} &= 1 - n_{byi}
\end{aligned}
$$

such that

$$
\begin{aligned}
0 \leqslant n_{byi} \leqslant 1 \\
y_{be1} \leqslant n_{by1}^{0.7} \\
y_{be2} \leqslant 1.1(1+\ln(25n_{by1}^{0.7}))]n_{by2}^{0.7}
\end{aligned}
$$

The solution to the above problem is reported in Table 2.2:

The real exchange rate (big/small) is:

$$
\begin{aligned}
RER_1 &\approx 1.0667 \\
RER_2 &\approx 1.2030
\end{aligned}
$$

Comparing to the one-period case, I find that consumers in the big country get a higher price of non-tradable goods Z and a lower price of tradable goods Y, and have a higher real exchange rate. The utility goes up considerably in the

Table 2.2: Two Periods

	$i=1$		$i=2$	
	The big country	Small countries	The big country	Small countries
n_y	0.6994	0.6203	0.5789	0.6203
n_z	0.3006	0.3797	0.4211	0.3797
y_e	0.2827	0.3068	0.9559	0.3375
p_y	1.8419	1.8419	1.2435	1.2435
p_z	1.7324	1.4880	2.8227	1.1050
x_c	0.5208	0.5651	1.1887	0.4196
y_c	0.4959	0.4090	2.0216	0.4499
z_c	0.3006	0.3797	0.4211	0.3797
U	0.4330	0.4407	1.0768	0.4188

second period, because of increased production destined for export market. For small countries, their utility is nearly the same as one-period model.

Again, we can calculate the consumers surplus in both periods:

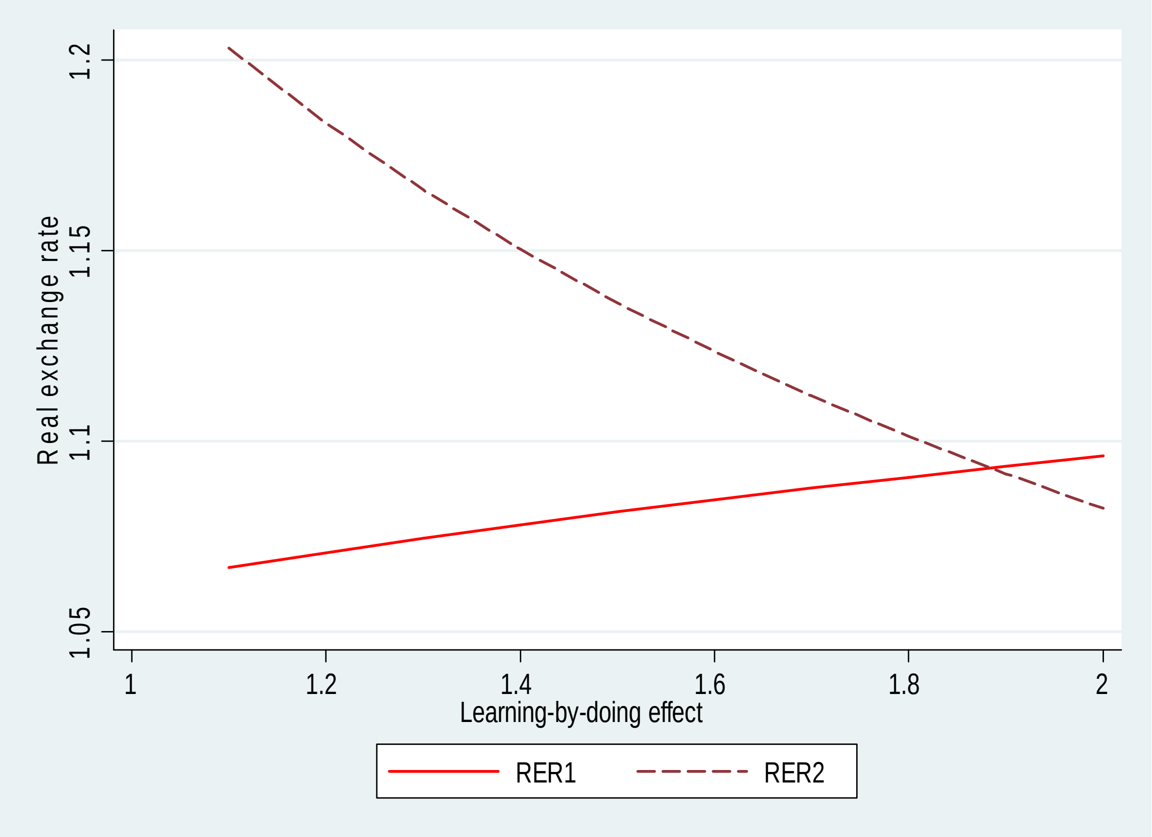
$$\begin{aligned} CS_1 &= 54.2908 \\ CS_2 &= 80.4159 \end{aligned}$$

It is very clear that, comparing with the one-period model, consumers of the importer country are better off in both periods, especially for the second one, because they can buy more tradable goods Y at a lower price.

Here, I also want to show a picture about the relationship between learning-by-doing effect and the real exchange rate in Figure 2.4.

In the graph, we see that the real exchange rate in period 1 goes up, but in period 2 goes down, as the learning-by-doing effect A increases.

Figure 2.4: Learning-by-doing and Real Exchange Rate



2.4.3 Counterfactual Analysis

Although the consumers of the import country are better off, we should notice that the utility of small exporters decreases a little bit. Now suppose a small country asks WTO to apply a new policy of export quotas. No firm can export more than $\theta = 0.5$ unit of Y in any period. This is nearly 1.5 times quantity of export from a small country firm. I want to see what will happen when restrict the big country's export.

Now all the countries face nearly the same problem as before. The only difference is that all the export quantity y_{bei} and y_{sei} should be less than $\theta = 0.5$. Solve this policy questions, and I get the following results in Table 2.3:

Table 2.3: Counterfactual Analysis

	$i=1$		$i=2$	
	The big country	Small countries	The big country	Small countries
n_y	0.6821	0.6203	0.5320	0.6203
n_z	0.3179	0.3797	0.4680	0.3797
y_e	0.2783	0.3068	0.5000	0.3375
p_y	1.8489	1.8489	1.5454	1.5454
p_z	1.6183	1.4937	1.6510	1.3733
x_c	0.5145	0.5672	0.7727	0.5215
y_c	0.4868	0.4090	2.2941	0.4499
z_c	0.3179	0.3797	0.4680	0.3797
U	0.4356	0.4412	1.0274	0.4470

And real exchange rate (big/small):

$$\begin{aligned} RER_1 &\approx 1.0487 \\ RER_2 &\approx 1.1194 \end{aligned}$$

I find that, with the export cap, comparing to the two-period model, consumers of the big country suffer from a higher price of tradable goods Y, and also, their welfare decreases a large amount in the second period. The big country has a lower real exchange rate, comparing to the one-period case. For small countries, they still have the similar utility level.

Once again, I calculate the consumers surplus in both periods:

$$\begin{aligned} CS_1 &= 54.0851 \\ CS_2 &= 64.7091 \end{aligned}$$

Comparing with the one-period model, consumers of the importer country are worse off in both periods, especially in the second one, because of the increased price of goods Y.

2.5 Conclusion and Future research

This paper develops a theory for why a big country might be special when it comes to pursuing strategic trade policy. I present an industrial organization model to capture the story of international trade with one big exporter and many small exporters. The first and simple version of my model embeds the basic static dominate-fringe firm structure to describe the one period case. Then I expand it to a two-period version with a learning-by-doing effect. All the countries have this kind of effect, but only the big country can internalize it and has a much larger effect. This is the main reason why a big country should subsidies exports by controlling the nominal exchange rate.

From the numerical result part, we can see that, with the learning-by-doing effect, the big country exports much more than before. Consumers in both the big country and the importer country are better off, and the welfare of small exporters remains roughly the same. The real exchange rate rises for the large country compared to the small exporting countries.

The counterfactual analysis shows that, an export quota policy does not benefit the small countries much but make both big exporters and importers significantly worse off. This means that export cap for big countries reduces international welfare.

For future research, there are three possible directions. The first is to expand the model from a two-period to a multi-period dynamic version. The firms have similar learning-by-doing effects in each period, and the spillover could not be prevented from firm to firm at any time. Of course, this new version is much more complicated but fits the real world much better and maybe reveal some new feature of the big exporter. The second is to use the same method to the world import market. Currently, the big country, like China, is not only a big exporter but also a large importer, especially natural resources. I want to see what is optimal for a big importer, and what is the effect of its behavior on other countries. The last way is to think about the micro base of the behavior I describe in this paper. Why do people prefer to purchase from big exporter? Is this because it supplies cheap products or a variety of products? The question involves a totally new model.

3 Invention and Super Inventors

3.1 Introduction

In less than three decades, China has grown from playing a negligible role in world trade to being one of the world's largest exporters, a substantial importer of raw materials, intermediate outputs, and both a recipient and source of foreign investment. Not surprisingly, China's rise in manufacturing and exports has received much attention and concern in the United States and elsewhere. However, less attention has been placed on China's growing role as a location of invention. As recently as 1984 there did not exist a patent system; now China has the largest patent system across all countries in the world. State Intellectual Property Office (SIPO) currently receives roughly 1.2 million applications a year and grants 0.8 million and the rate of growth is more than 30 percent. This number is large compared to industrialized countries - the U.S. patent office receives nearly 520,000 patent applications and grants 240,000, and Europe receives 230,000 patent applications and grants 130,000. One might try to make a case that typical Chinese patents might involve less of an inventive step than U.S. patents or European patents. But it is worth noting that patent filed in the United States or Europe by inventors from China is also exploding. While only in the most recent year Chinese inventors accounted for

6879 of the applications in the U.S. and 3303 of granted patents, and the rate of growth is nearly 50% in U.S. and more than 90% in Europe.

This paper studies the rise of patenting in China. It makes a use of a data set of all patents granted in China since the inception of the system in 1985. This data was downloaded from the SIPO web site and as far as I know it is the first academic paper to make use of the entire set of Chinese patents.

The particular question I address is the role of individual inventors in the growth of Chinese patenting. One fact I document is that the distribution of patents across inventors is highly skewed and becoming more so over time. For example, I find that the top 1% of all inventors in 2010 accounted for 22.8% of all patents. In 1990-1991, by contrast, the top 1% share was only 5.5%. This increased skewness is something special that is happening in China.

A variety of phenomena in economics and nature occurs in a skewed fashion and relationships are often illustrated in a plot of log rank on log size. Gabaix (1999) is a famous paper on cities, where the slope is -1, called Zipf's law in the literature on the size distribution of cities. Luttmer (2007) and Axtell (2001) report analogous relationships for firm size and Axtell (2001) also comments about phenomena in nature. I plot the same kind of relationship between for invention counts of individuals. To a remarkable degree, I obtain a straight-line relationship for a large portion of the data set. This reflects the fact that the right tail is relatively fat, that is there is a surprising number of superstar inventors. I find that the slope is flattening over time, meaning the tail is getting fatter, the distribution is getting more skewed.

I develop a theory that can jointly account for why patenting levels are increasing overall and why the distribution is becoming more skewed. The idea is based on the theory of the optimal allocation of inputs, just like Hsieh and Klenow (2009). The idea is that in the early period, the Chinese economy was heavily regulated. Consider that in 1985 when the first patent systems started, the government controlled nearly all resources, making it very difficult for inventors to operate. For example, even those inventors that had funding often found it impossible to buy necessary inputs since all the inputs are limited supplied and needed both money and tickets to get them. Over the last 30 years, there has been substantial freeing of the economy, which means that productive resources are being reallocated to the highest ability people. With this better matching of resources and talent, the distribution of output has become more skewed. Moreover, with better allocation of resources, the incentive for entry into patents has gotten higher.

The rest of the paper proceeds as follows. Section 2 reviews the relevant literature on patent, skewness (Zipf's law), and China. Section 3 discusses the data. Section 4 present empirical results. Section 5 presents the theoretical results.

3.2 Review of the Literature

3.2.1 Patents

The literature about patents mainly focuses on the relation between patents and R&D spending, like Hausman et al. (1981), Hall et al. (1983) and Hall et al. (1984), or on the reason of patent explosion like Hall (2004). Others focus on the innovation spillover, like Jaffe et al. (1993), Hu and Jaffe (2001) and Bottazzi and Peri (2003), or explain what factors will decide the patent transfer, like Serrano (2010). About this topic, Azoulay et al. (2011) also argues that article-to-article citation from the scientific community at the superstar inventor's origin location are barely affected by their departure. In contrast, article-to-patent citations, and especially patent-to-patent citations, decline at the origin location following a star's departure. This suggest that the communication between inventors is not that easy as we think, at least less easy than the authors of academic papers. In all the literature, there are few paper about inventors, and none using Chinese data.

3.2.2 Zipf's law literature

There also are a lot of papers about the Zipf's law (skewness). Many of them show that a lot of phenomena in economics and nature follows the Zipf's law, for example, Gabaix (1999), Axtell (2001), Helpman et al. (2004), Luttmer (2007) and Holmes and Lee (2008). Gabaix (1999) is a very influential paper about this topic. It shows the relationship between large cities size and their rank in Figure 3.1 . Notice that the slope of the regression line is -1 and called Zipf's law in his paper. Other papers like Helpman et al. (2004) instead address the firm size and use the slope of regression to measure the dispersion. I will also follow their idea in my paper. But all of them mainly talked about the firms size or cities size instead of the patents. In fact, there is no literature about the dispersion of inventors using the entire Chinese patent data set.

3.2.3 Innovation in China

There are very few papers about innovation in China. Cheung and Lin (2003) find that foreign direct investment has positive effects on the patents application, using the provincial data in China. There are several other papers about this topic but published on journals written in Chinese instead of English, so they are not well known by researchers.

Figure 3.1: Zipf's law of cities

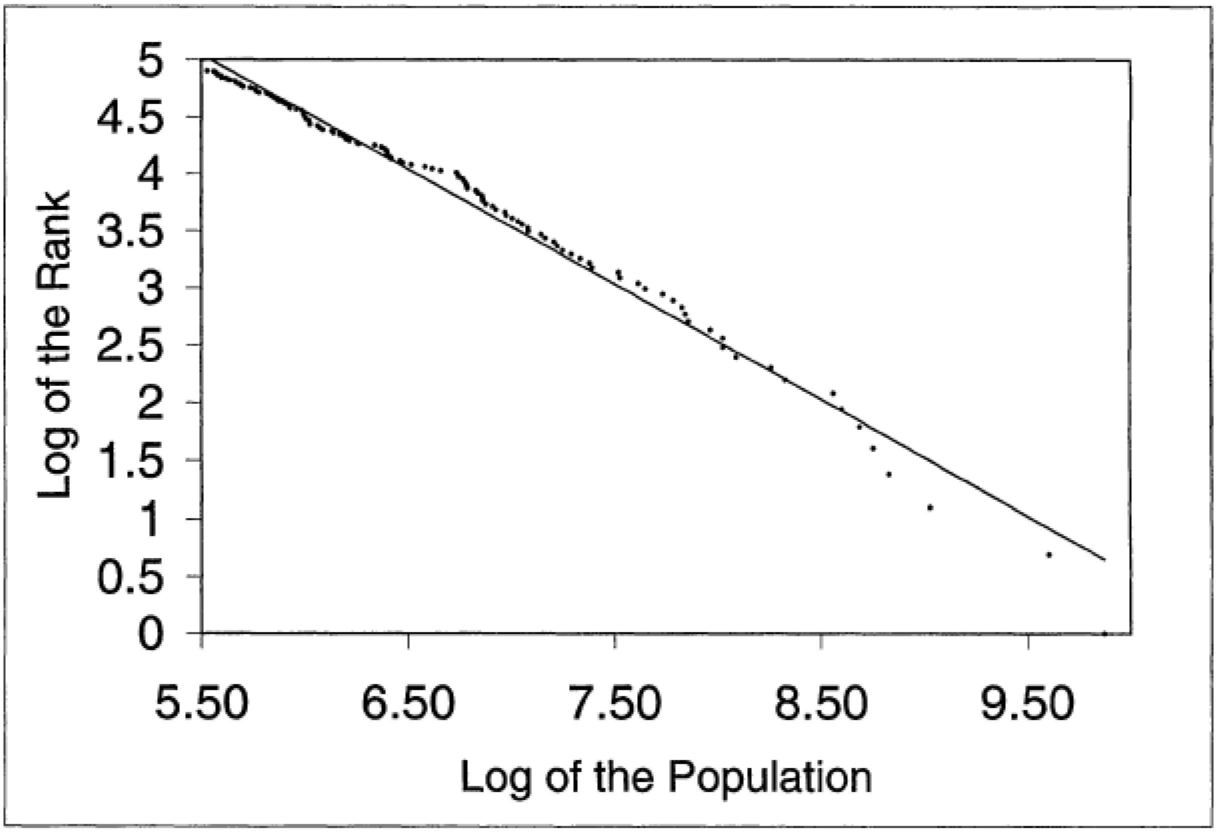


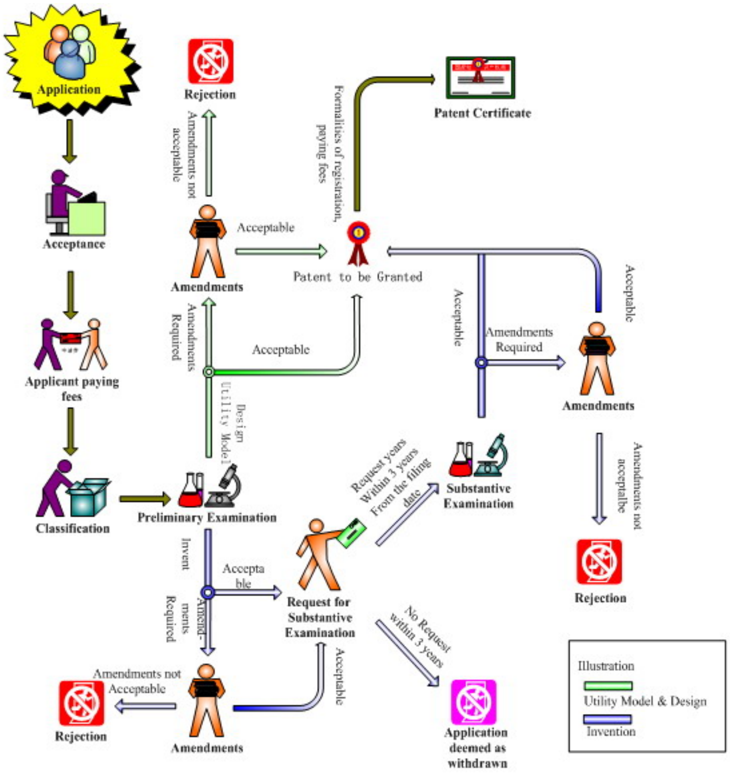
FIGURE I

Log Size versus Log Rank of the 135 largest U. S. Metropolitan Areas in 1991
Source: Statistical Abstract of the United States [1993].

3.3 Chinese patent system

The Chinese patents can be divided into three categories: invention, utility and design, following the definition from State Intellectual Property Office (SIPO) of the People's Republic of China (PRC). SIPO's website in English does not provide data on design patents, so I only discuss invention and utility patents. The difference between these two, roughly speaking, is that an invention patent is a big innovation and the utility patent is small innovation. The term of patent for invention patent is 20 years and for utility patent is only 10 years. In addition, these two kinds of patent have different application processes. The Figure 3.2[1] provides more detail.

Figure 3.2: Patent Application



---

[1] Source: SIPO website `http://english.sipo.gov.cn`

3.4 Data

I download the Chinese patent data from SIPO database. It contains all the patents granted in China from 1985 to 2010. Totally, there are nearly 3.5 million patents. For each patent, I have the following information: the title, application date and number, granted date and number, international class, applicant and inventor's name, agent's name and code, and also the abstract of the patent. The Table 3.1 summarize the basic info about the patent by year, from which we can see the trend of this Chinese patent data set. [2]

Table 3.1: Patent Granted Summary

year	N	Mean	Min	Max	Growth Rate
1985-2000	654,740	40,921.25	201	94,188	14.41%
2001-2010	2,933,159	293,315.9	106,276	631,660	21.33%

Source: Authors calculations from patent data downloaded from SIPO website.

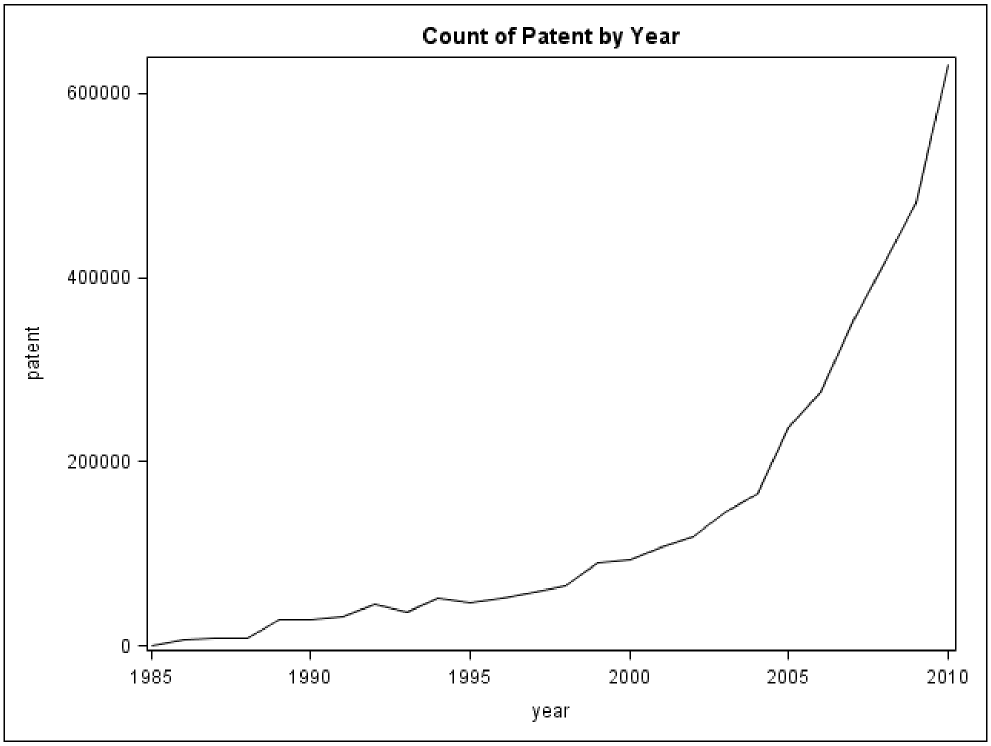
The Figure 3.3 shows the huge increase of the Chinese patents during these 26 years. We can see that, especially after 2001, the number of patent exploded. Compare this with the U.S patents in Figure 3.4.

We find that, the number of patents granted in the U.S. increased by a large and stable rate, but not as much as in China. Sanyal and Jaffe (2004) shows that the explosion of U.S. patents is not because the change of the patent law (making it easier to acquire a patent), but driven to a significant extent by an increase in the underlying invention rate. I believe this is also the situation in China. China entered

---

[2] In the first several year, there is huge change since the quantity is pretty small, so I just calculate growth of 1991-2000.

Figure 3.3: Chinese Patent by year

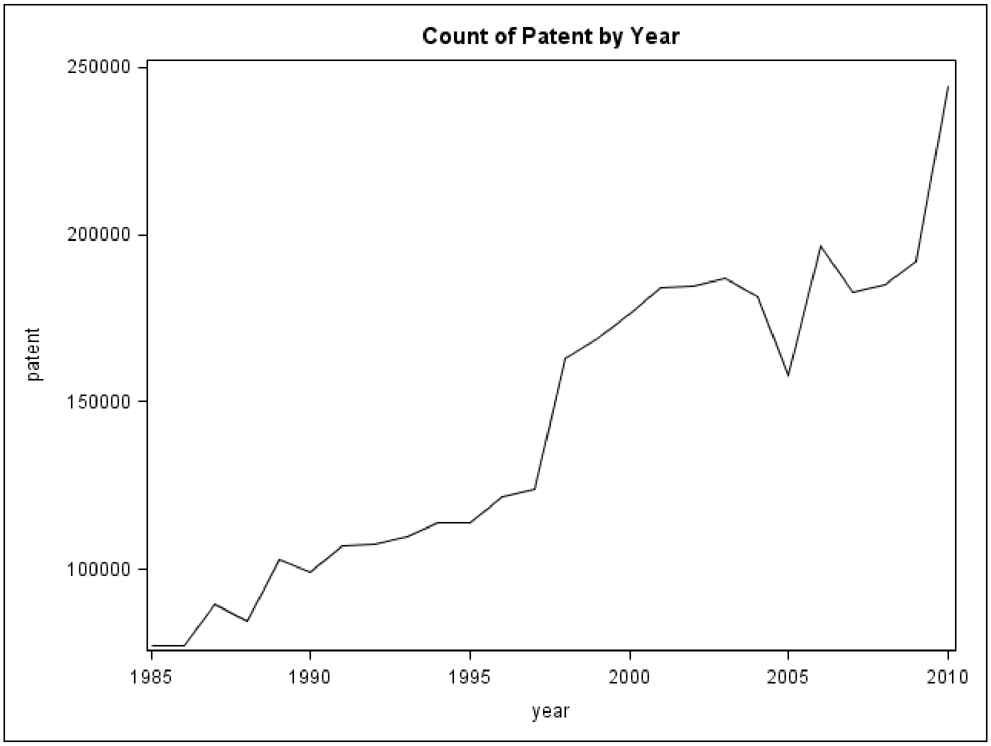


World Trade Organization (WTO) in 2001 and after that every inventor has more incentive to protect his/her intellectual property. Of course, these kind of increases in patents is accompanied by a large increase of inventors. I show the increase of inventors in Figure 3.5. Thus, next step, I want to focus on the inventors side.

First of all, we need to define what an inventor is. I am going to start by assuming each name corresponds to an inventor. Note that this is subject to two kinds of errors. First, there may be two or more people with the same name. Second, the same person might fill in his or her name a little differently. One time, the person might abbreviate the first name, another time the full name may be used. I will be considering several strategies to deal with these errors.

The data has 1,772,717 unique names for all patents granted, which means nearly there are 1.7 million inventors from 1985 to now. Table 3.2 shows the sum-

Figure 3.4: US patent by year



mary statistics of inventors. I want to see how serious the name error is. I created the list of the top10 superstars inventors in China in Table 3.3.

This is the 10 superstars with more than 3000 patents for each of them. For example the third superstar, Qiu Zeyou, is very famous for the number of patent that he has in China. Both his first and last names are relatively rare in China. I check the Chinese version of his patent and nearly no inventor use the same name. Another example is the 6th superstar Xie Yi. This is a not very common last name but a very common first name. I check the Chinese version and do find some inventors share the same name, but not much, because one of them is the co-author with the 7th superstar Mao Yumin (a real superstar) and a lot of patents are under their company.[3] The quantity of other inventors under the name Xie Yi is pretty small

---

[3] We see that the first large company under their two names only has half of their patent. But in fact, the 2nd and 3rd large companies are also belongs to them. They report their company very

Figure 3.5: Counts of Inventors by Year

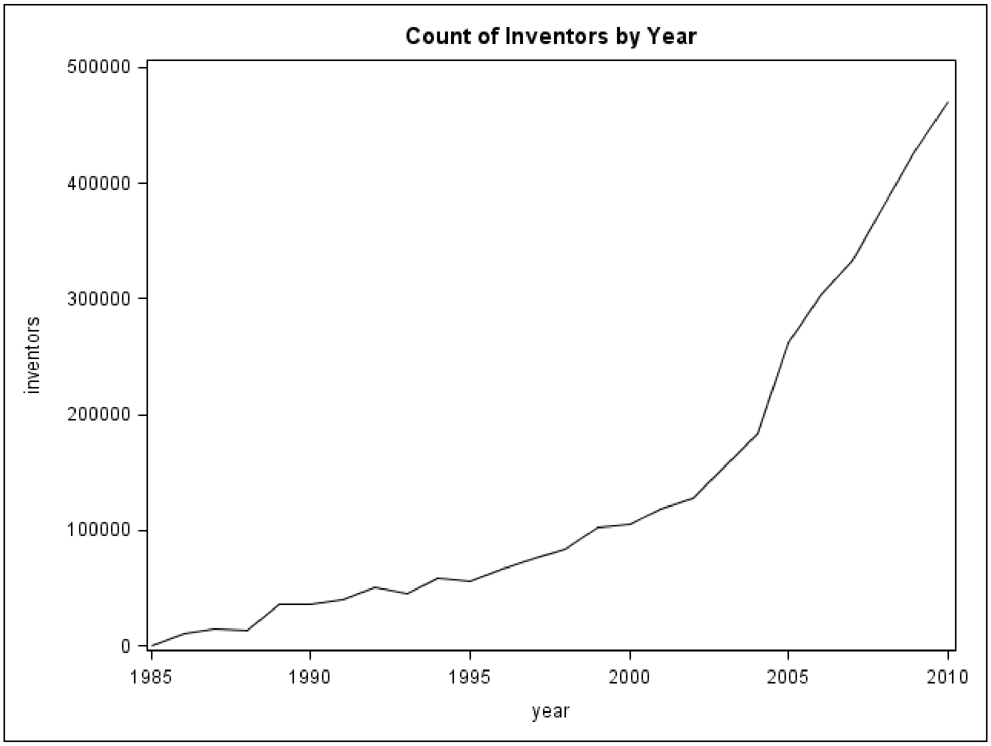


Table 3.2: Inventors Summary

	Total	1985-2000	2001-2010
Inventors w/ 1 patent	972890	327119	777558
Inventors w/ 2 patent	272637	82205	228894
Inventors w/ 3-5 patent	271741	65632	235052
Inventors w/ 6-10 patent	128053	20433	112820
Inventors w/ 11-20 patent	71380	7251	63251
Inventors w/ 21-50 patent	40066	2371	35172
Inventors w/ 51-100 patent	10361	333	8894
Inventors w/ more than 100 patent	5665	132	4882
Average patents per inventor	2.02	1.30	2.00

Table 3.3: Superstar inventors

Inventor	Total	1st Org	2nd Org	Total Org
Wang, Wei	5581	729	102	2262
Zhang, Wei	5245	1156	90	1965
Qiu, Zeyou	4879	4862	5	11
Wang, Jun	4319	794	83	1611
Li, Wei	4267	472	62	1811
Xie, Yi	3999	2018	443	139
Mao, Yumin	3889	2008	443	74
Li, Jun	3607	514	71	1531
Wang, Yong	3325	534	109	1308
Liu, Wei	3318	427	35	1500

and with very few patent. This does not pose a problem in identifying the number of patents for each inventor. However, there are still some real problems with the first kind error on names, like the first superstar, Li Wei, using a very common last and first name in China. For example, the last name Li. This is the first largest last name in China. About 8% Chinese people use this last name. (Remember China has more than 1.3 billion people!) Thus, this name is not a guy, but a large group of inventors. That's why there are so many patents under this name.

This problem is also reflected in the fact that the patents-inventors ratio is abnormally high for China relative to other countries. The result is showed in Table 3.4 . Chinese inventors also have the least fraction of inventors but the highest fraction of patents. This tells us that there is a large first kind of error about inventors name.

People may worry about this name problem, especially for the large group of inventors under the same name. However, this problem is mitigated by the fact that

---

informal so the name of the company is a little bit different in data set and be viewed as different one by the program. This is the same as the second kind error of name.

Table 3.4: Top10 inventors of countries in Chinese market

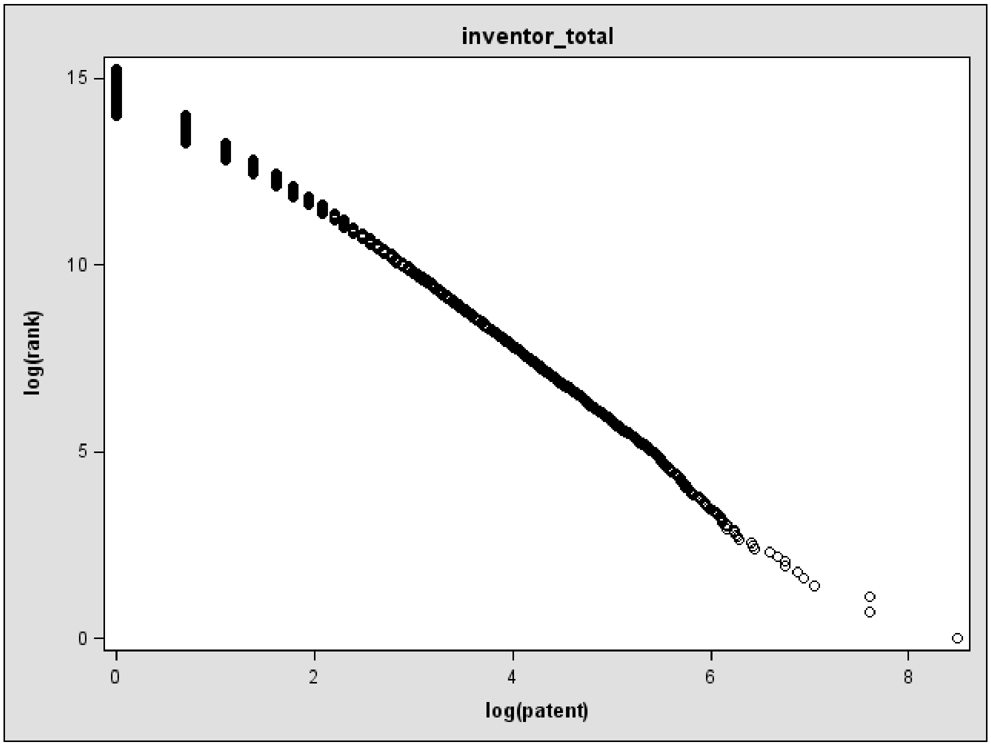
	Top10 inventors			All inventors
Inventors	patents	%patent	%inventors	patents/inventor
China	41487	0.7145%	0.0012%	6.91
US	2402	0.3711%	0.0028%	1.79
Japan	2391	0.3361%	0.0038%	2.69

my interest is in comparing inventors over time. I expect this issue of common names has not change over time, since China naming convention have not changed very much. The effect I will be examining over time will then be due to differences in skewness by inventors. Secondly, I will employ a strategy to reduce the severity of the problem. I check the applicant's name, which is mainly filled by inventor's organization (university, institute or company). Thus, I will define an inventor as a person with the same name and organization name. Of course, there is the chance that several inventors have the same name and are in the same organization, but it is really small and should not have much effect on my result.

I rank the Chinese inventors by their patents quantity, so the inventor who has most patents is rank 1 and who has second most patents is rank 2 and so on. Then, I plot them in a log scale graph as in Figure 3.6. I found that this graph has a linear portion except the left fat tail. So I estimate the slope of this linear portion. I will leave in the right tail since this does not involve many inventors so will not be weighted much anyway. The result is showed in the Figure 3.7.

I find it fits the data very well. The inventor I dropped is the one who has

Figure 3.6: Inventors in Chinese Market

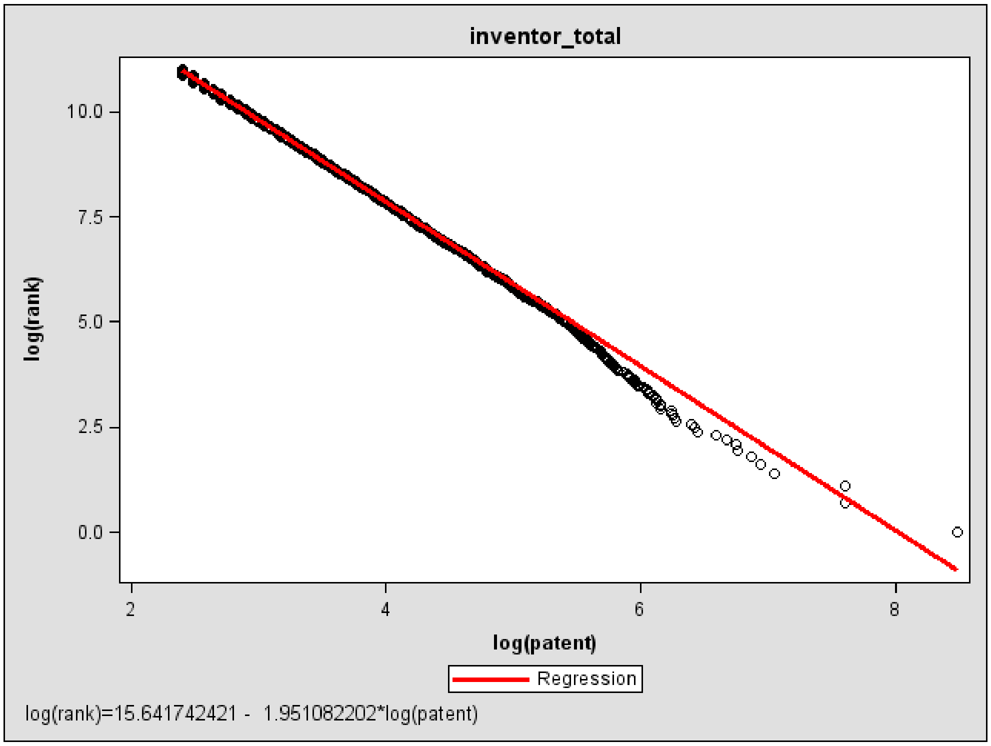


few patents (less than 10 patents). In total, I dropped 4,126,814 inventors (98.54% of total inventors) and 6,575,041 patents (83.22% of total patents)[4]. Here we can see the regression equation and the slope of the regression line is nearly -1.95. Generally speaking, the slope of this straight line measures the dispersion of the inventors. Imagine a very special case: if the regression line is very steep and nearly vertical (a very large slope value). It means that, no matter what rank the inventor is, he or she has nearly the same amount of patents. That is, inventors are pretty similar in the quantity of the patents. But, of course, people are different in the ability to invent, so one possible explanation is that the easier communication skill nowadays helps inventor to understand (maybe sometimes steal) each others idea and get the new ideas. On the other hand, if the slope value is pretty small

---

[4] Here, I define a patent as a patent with a single inventor. Thus if a patent has four inventors, then it will be counted as four patents.

Figure 3.7: Regression of inventors in Chinese Market



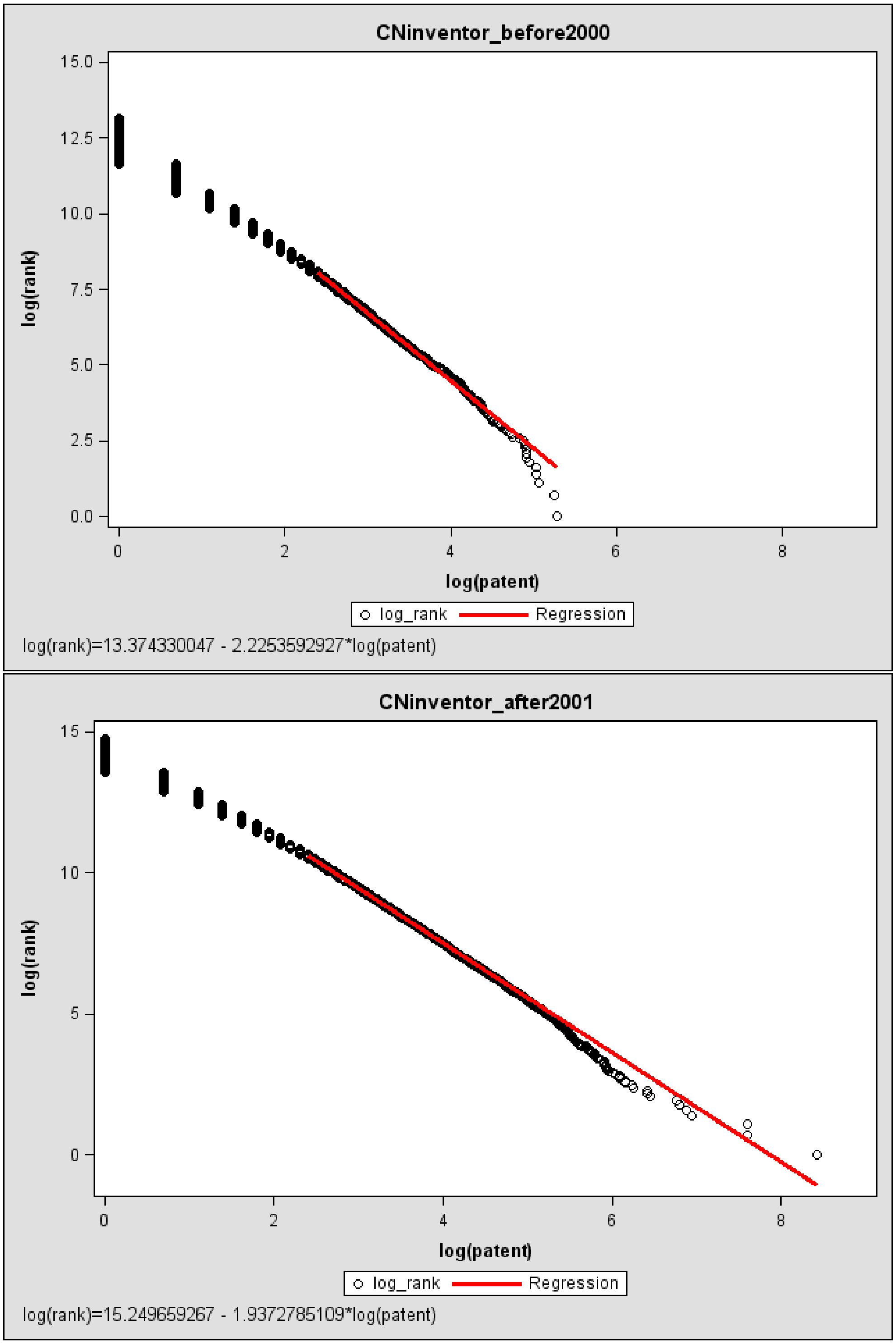
and the regression line is flat, it just means that the inventors have much difference. One possible reason for this, is that some patent is harder to invent than others, for example, invent patent vs. utility patent, so only the smarter inventors can do that.

We will see how the distribution of inventors has evolved in China. The first period is from 1985 to 2000 and the next period is 2001-2010. I plot the two graphs in the same scale so we can easily see the slope, as shown in Figure 3.8.

The regression result is summarized in Table 3.5. The slope of regression line in the second period is nearly -1.99. The absolute value is much less than the slope of first period-2.26, nearly 12% less. So what we get is a flatter regression line, which means, the difference between inventors are much more than before. In next section, I will develop a model to explain this.

I also do the same for other countries' inventors in China, as shown in Figure 3.9

Figure 3.8: Chinese inventors



log(rank)=13.374330047 - 2.2253592927*log(patent)

log(rank)=15.249659267 - 1.9372785109*log(patent)

and Figure 3.10. All countries show the same pattern, i.e., there is more dispersion than before. But from Table 3.6 we can see that, different countries have different slope change. I could not generalize them in one model, so I only want to supply a theory to answer the question about Chinese inventors. The other countries puzzles will be tried to figure out in future research.

3.5 Theory

3.5.1 Model

In my model, I assume that the production function of patent is: $q = \gamma a^{\alpha} s^{\beta}$, where q denotes the quantity of patent, a is the ability of inventors, s is the resource each inventor gets, α, β, γ are the parameters, and $\alpha, \beta > 0$. I suppose the fixed cost of entry ϕ varies across inventors. Inventors do not know their ability until after they enter this market. Let $E(q)$ be the expected number of patents, then the free entry condition will be $\phi = E(q)$. That is, anyone who has entry fixed cost less than the expected quantity of patents, will enter the market.

The most important assumption is about the ability to invent. It is not good to assume the ability follows the normal distribution, since normal distribution takes on negative values. Instead, following the idea from Helpman et al. (2010), I assume that the ability of inventors follows a Pareto distribution, which has the cumulative distribution function (cdf)

$$F(x) = \begin{cases} 1 - (\frac{x_m}{x})^k & for \quad x > x_m \\ 0 & for \quad x < x_m \end{cases}$$

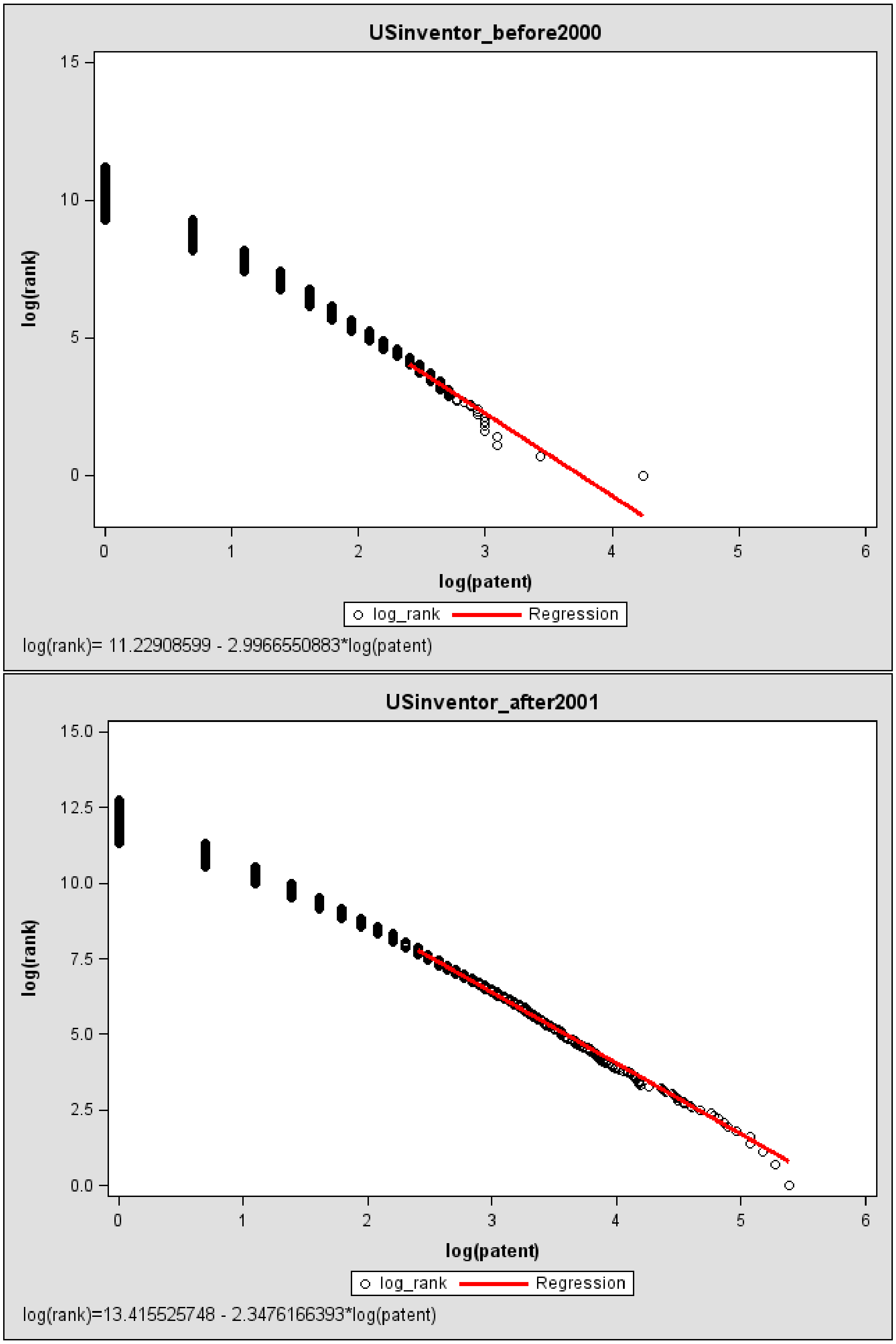
Table 3.5: Chinese inventors

variable	obs	estimate	std err	Adj R^2	p value
$log(patent_{total})$	61070	-1.9511	0.0003	0.9982	< 0.0001
$log(patent_{2000})$	3310	-2.2254	0.0026	0.9955	< 0.0001
$log(patent_{2001})$	42514	-1.9373	0.0004	0.9979	< 0.0001

Table 3.6: Different Countries Inventors

Countries	$log(patent_{2000})$	$log(patent_{2001})$	$\%\triangle$
China	-2.2254 (0.0026)	-1.9373 (0.0004)	12.95%
US	-2.9967 (0.1102)	-2.3476 (0.0028)	21.66%
Japan	-2.5530 (0.0439)	-2.4520 (0.0023)	3.96%

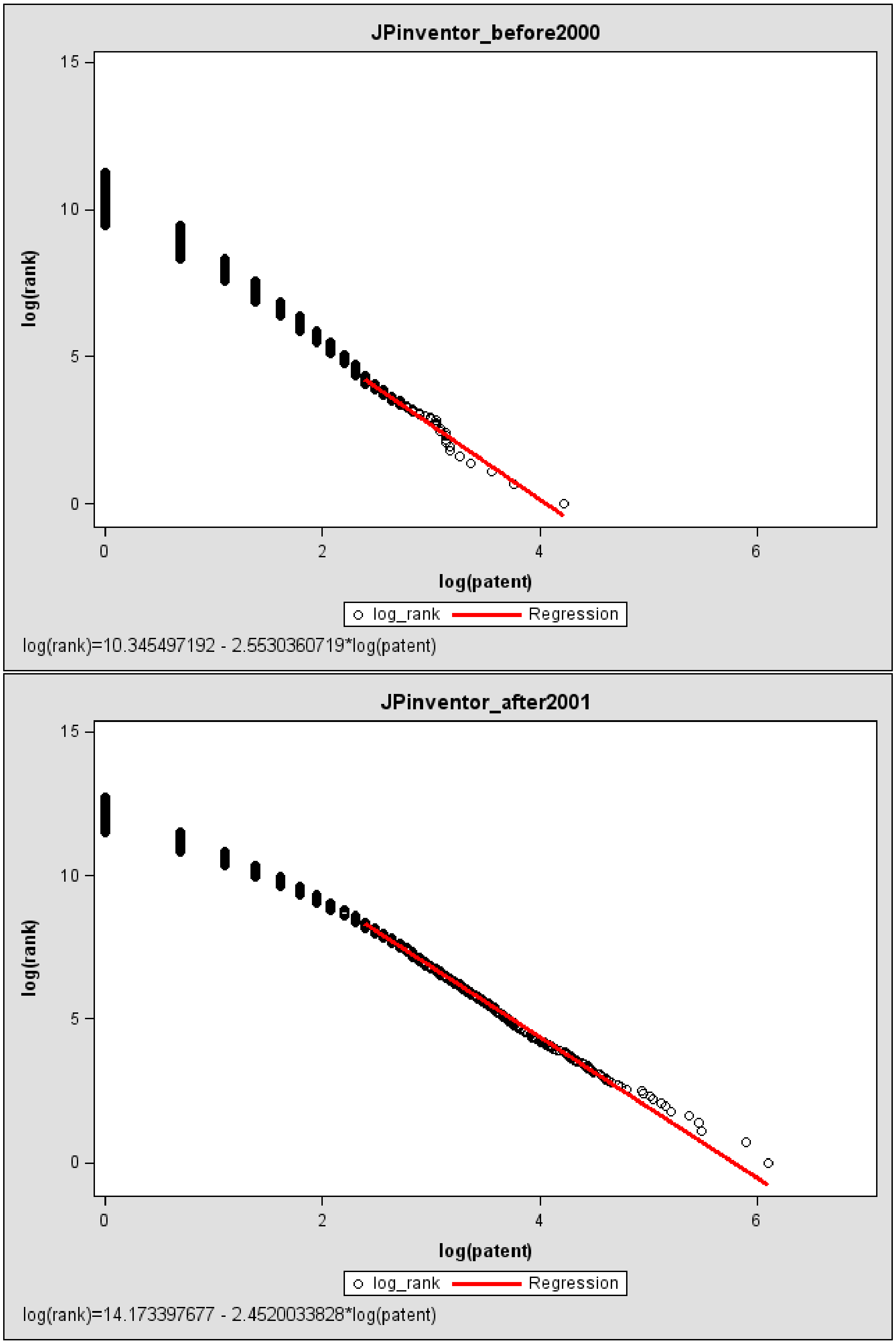
Figure 3.9: United States Inventors



log(rank)= 11.22908599 - 2.9966550883*log(patent)

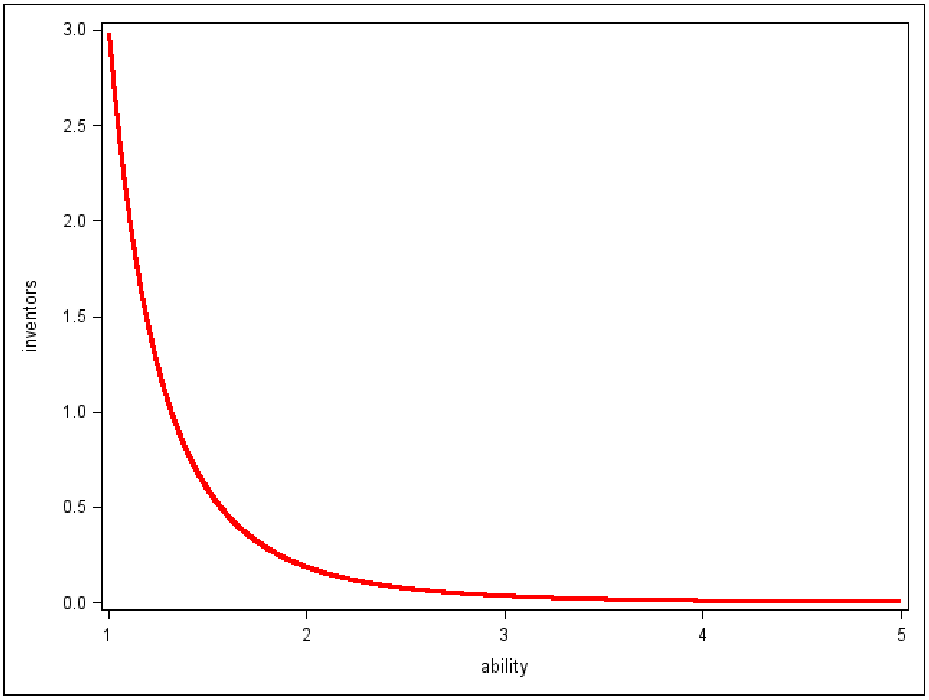
log(rank)=13.415525748 - 2.3476166393*log(patent)

Figure 3.10: Japan Inventors



where, x_m is the minimum possible value of x and k is the parameter of this distribution. The graph of probability density function (pdf) of Pareto distribution is in Figure 3.11. With this distribution, the requirement of an ability threshold is naturally satisfied.

Figure 3.11: pdf of Pareto distribution



Now, think about two different case of resource distribution. The object of the government is, of course, to maximize the welfare from technology innovation, i.e., maximize the quantity of patent. At first, China has few resources and the planner does not know the ability of each inventor. Thus, the resource will be divided equally to see the ability first. I assume every inventor just gets 1 unit of them ($s_i = 1$). The total resource is n units if we assume there are n inventors. After several year's development, China has more resources for each inventor (here I suppose the resource for each inventor times z, so totally yzn units) so the expected number of patents, $E(q)$ goes up. Then more and more people with fixed cost

$\phi_i \leqslant E(q)$ enter this market, so there will be more inventors. Now planner decides to distribute the resource more efficiently, so every inventor gets the resource depends on their ability.

In the first case,

$$\begin{aligned} q_i &= \gamma a_i^{\alpha} s_i^{\beta} = \gamma a_i^{\alpha} 1^{\beta} \\ &= \gamma a_i^{\alpha} \end{aligned}$$

Since ability follows the Pareto distribution, then we can find the distribution of patent. We only look at the situation with $x > x_m > 0$, so the cdf of ability is:

$$F_a(x) = 1 - (\frac{x_m}{x})^k$$

The cdf of the patent is:

$$\begin{aligned} F_q(x) &= Pr(q_i \leqslant x) = Pr(\gamma a_i^{\alpha} \leqslant x) \\ &= Pr(a_i \leqslant (\frac{x}{\gamma})^{\frac{1}{\alpha}}) = F_a((\frac{x}{\gamma})^{\frac{1}{\alpha}}) \\ &= 1 - (\frac{x_m}{(\frac{x}{\gamma})^{\frac{1}{\alpha}}})^k = 1 - (\frac{x_m \gamma^{\frac{1}{\alpha}}}{x^{\frac{1}{\alpha}}})^k \end{aligned}$$

Thus, the pdf of the patent is: $f_q(x) = \frac{k}{\alpha}(x_m \gamma^{\frac{1}{\alpha}})^k x^{-\frac{k}{\alpha}-1}$. Plot in the log scale axis and we get

$$\log f(x) = \log(\frac{k}{\alpha}(x_m \gamma^{\frac{1}{\alpha}})^k) - (\frac{k}{\alpha} + 1)\log(x)$$

So, the slope of the log scale graph should be $\frac{k}{\alpha} + 1$.

And in the second case, we need to find efficient quantity of resource first. The object of the planner is:

$$\begin{aligned} \max \quad & \textstyle\sum_{i=1}^{yn} q_i = \sum_{i=1}^{yn} \gamma a_i^{\alpha} s_i^{\beta} \\ s.t. \quad & \textstyle\sum_{i=1}^{yn} s_i = yzn \end{aligned}$$

F.O.C

$$\begin{aligned} a_i^{\alpha} s_i^{\beta-1} &= a_j^{\alpha} s_j^{\beta-1} \\ \frac{s_i}{s_j} &= \left(\frac{a_j}{a_i}\right)^{\frac{\alpha}{\beta-1}} \end{aligned} \qquad for \quad \forall i,j$$

Thus

$$s_i = \left(\frac{a_1}{a_i}\right)^{\frac{\alpha}{\beta-1}} s_1$$

And we know that $\sum_{i=1}^{yn} s_i = yzn$, i.e.

$$\begin{aligned} \textstyle\sum_{i=1}^{yn} \left(\frac{a_1}{a_i}\right)^{\frac{\alpha}{\beta-1}} s_1 &= yzn \\ s_1 &= \frac{yzn}{a_1^{\frac{\alpha}{\beta-1}} \sum_{i=1}^{yn} \left(\frac{1}{a_i}\right)^{\frac{\alpha}{\beta-1}}} \end{aligned}$$

Finally we get that, for each inventor:

$$\begin{aligned} s_i &= (\frac{a_1}{a_i})^{\frac{\alpha}{\beta-1}} \frac{yzn}{a_1^{\frac{\alpha}{\beta-1}} \sum_{i=1}^{yn} (\frac{1}{a_i})^{\frac{\alpha}{\beta-1}}} \\ &= \frac{yzn(\frac{1}{a_i})^{\frac{\alpha}{\beta-1}}}{\sum_{i=1}^{yn} (\frac{1}{a_i})^{\frac{\alpha}{\beta-1}}} \end{aligned}$$

Then we could know the quantity of patent:

$$\begin{aligned} q_i = \gamma a_i^{\alpha} s_i^{\beta} &= \gamma a_i^{\alpha} (\frac{yzn(\frac{1}{a_i})^{\frac{\alpha}{\beta-1}}}{\sum_{i=1}^{yn} (\frac{1}{a_i})^{\frac{\alpha}{\beta-1}}})^{\beta} \\ &= \frac{\gamma (yzn)^{\beta} a_i^{\frac{\alpha}{1-\beta}}}{(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\beta}} \end{aligned}$$

Here I want to show that with better allocation of resource, the total output is higher, just like Hsieh and Klenow (2009). This is trivial to prove in these two cases, since in the second case we have more resource and more inventors. But even with the same resource and same inventors, I can also show that the total output increases. Remember that, this q_i is the solution of maximization problem $(y = 1, z = 1)$:

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} q_i = \sum_{i=1}^{n} \gamma a_i^{\alpha} s_i^{\beta} \\ s.t. \quad & \sum_{i=1}^{n} s_i = n \end{aligned}$$

And this is also the problem of the planner in the first period. The resource for each is 1 and the total resource is equal to the number of inventors, n. Thus, this q_i in the second period, even with $y = 1$ and $z = 1$, makes the total output higher than the first period.

Let's go back to see the cdf of the patent:

$$\begin{aligned}
F_q(x) = Pr(q_i \leqslant x) &= Pr(\frac{\gamma(yzn)^{\beta} a_i^{\frac{\alpha}{1-\beta}}}{(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\beta}} \leqslant x) \\
&= Pr(a_i \leqslant (\frac{x(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\beta}}{\gamma(yzn)^{\beta}})^{\frac{1-\beta}{\alpha}}) = F_a((\frac{x(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\beta}}{\gamma(yzn)^{\beta}})^{\frac{1-\beta}{\alpha}}) \\
&= 1 - (\frac{x_m}{(\frac{x(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\beta}}{\gamma(yzn)^{\beta}})^{\frac{1-\beta}{\alpha}}})^k = 1 - (\frac{x_m(\gamma(yzn)^{\beta})^{\frac{1-\beta}{\alpha}}}{(x(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\beta})^{\frac{1-\beta}{\alpha}}})^k
\end{aligned}$$

Thus, the pdf of the patent is: $f_q(x) = \frac{k(1-\beta)}{\alpha}(\frac{x_m(\gamma(yzn)^{\beta})^{\frac{1-\beta}{\alpha}}}{(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\frac{(1-\beta)\beta}{\alpha}}})^k x^{-\frac{k(1-\beta)}{\alpha}-1}$. Plot in the log scale axis and we get

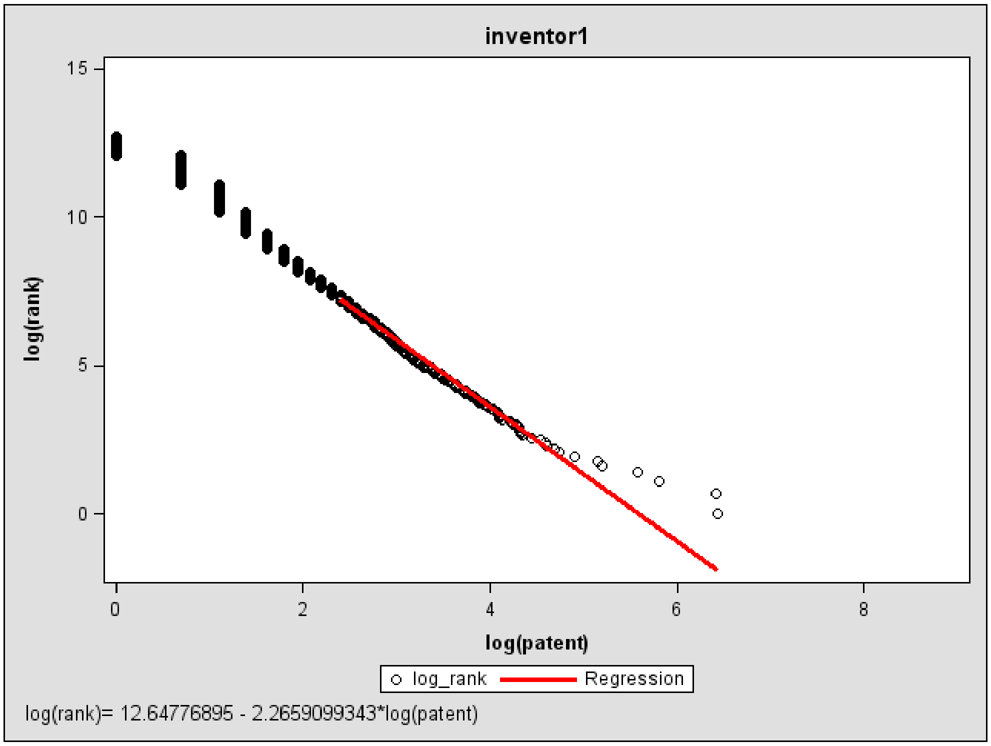
$$\log f(x) = \log(\frac{k(1-\beta)}{\alpha}(\frac{x_m(\gamma(yzn)^{\beta})^{\frac{1-\beta}{\alpha}}}{(\sum_{i=1}^{yn} a_i^{\frac{\alpha}{1-\beta}})^{\frac{(1-\beta)\beta}{\alpha}}})^k) - (\frac{k(1-\beta)}{\alpha} + 1)\log(x)$$

So, the slope of the log scale graph should be $\frac{k(1-\beta)}{\alpha} + 1$. Because $\beta > 0$, $\frac{k(1-\beta)}{\alpha} + 1 < \frac{k}{\alpha} + 1$. This explains why the absolute value of the slope of regression line is less, comparing the period after 2001 to before 2000 in China.

3.5.2 Numerical Example

We can also see this by a numerical examples of the model. Here I assume that the ability follows the Pareto distribution with minimum value $x_m = 1$ and parameter $k = 3$. I first draw $n = 0.5 million$ sample from it, and using the patent function with parameters $\alpha = 1\frac{1}{4}$, $\beta = \frac{1}{6}$ and $\gamma = 1$. The resource at first is not enough, so will be shared equally by all the inventors ($s = 1$). Here I also suppose there is an error item ε in the patent, ε is an independent variable to contains all the other factors, which follows the standard normal distribution. I can calculate their patent number and rank them. All these process are the same as in the data section. I get the Figure 3.12.

Figure 3.12: Inventor in the first period



Then in the second period, China's economic growth increases available resources ($z = 2$) and the way it is allocated depends on ability. Smarter inventors should get more resources. For ability data, because now more inventors ($y = 3$) come into this market for some other reason, I draw a 1.5 million sample from the same Pareto distribution, and I get the Figure 3.13.

Figure 3.13: Inventors in the second period

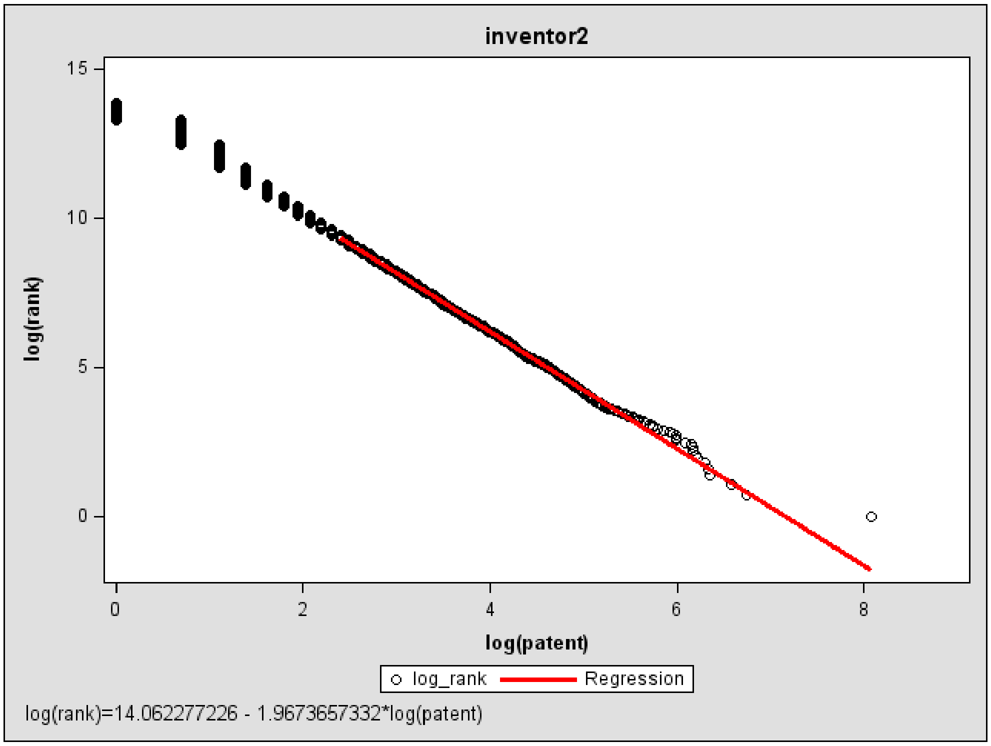


Table 3.7: Numerical Example of the model

variable	obs	estimate	std err	Adj R^2	p value
$log(patent_1)$	1581	-2.2659	0.0079	0.9813	< 0.0001
$log(patent_2)$	12476	-1.9674	0.0008	0.9980	< 0.0001

In both graphs, I dropped the fat tail like the above process in the data section. The regression result is in Table 3.7. We can easily see that the absolute value of slope of period 2 is 1.97, nearly 13% less than the first period (2.27). That means the inventors are more dispersed since resources are allocated more efficiently and more inventors enter this market as education and ability increase. This pattern is exactly the same as what I got in the real Chinese patent data.

I also do some comparative statics analysis with β, which impacts the effect of efficient distribution of resource. I find that, if I make β larger, the slope difference between these two periods will be larger than the original numerical example. Here I change β to $\beta = \frac{1}{4}$ with all other parameters the same as before. I get the Figure 3.14. The regression result is in Table 3.8. We can see that the change of slope now is -2.27 to -1.77, 22% difference and greater than the original numerical example.

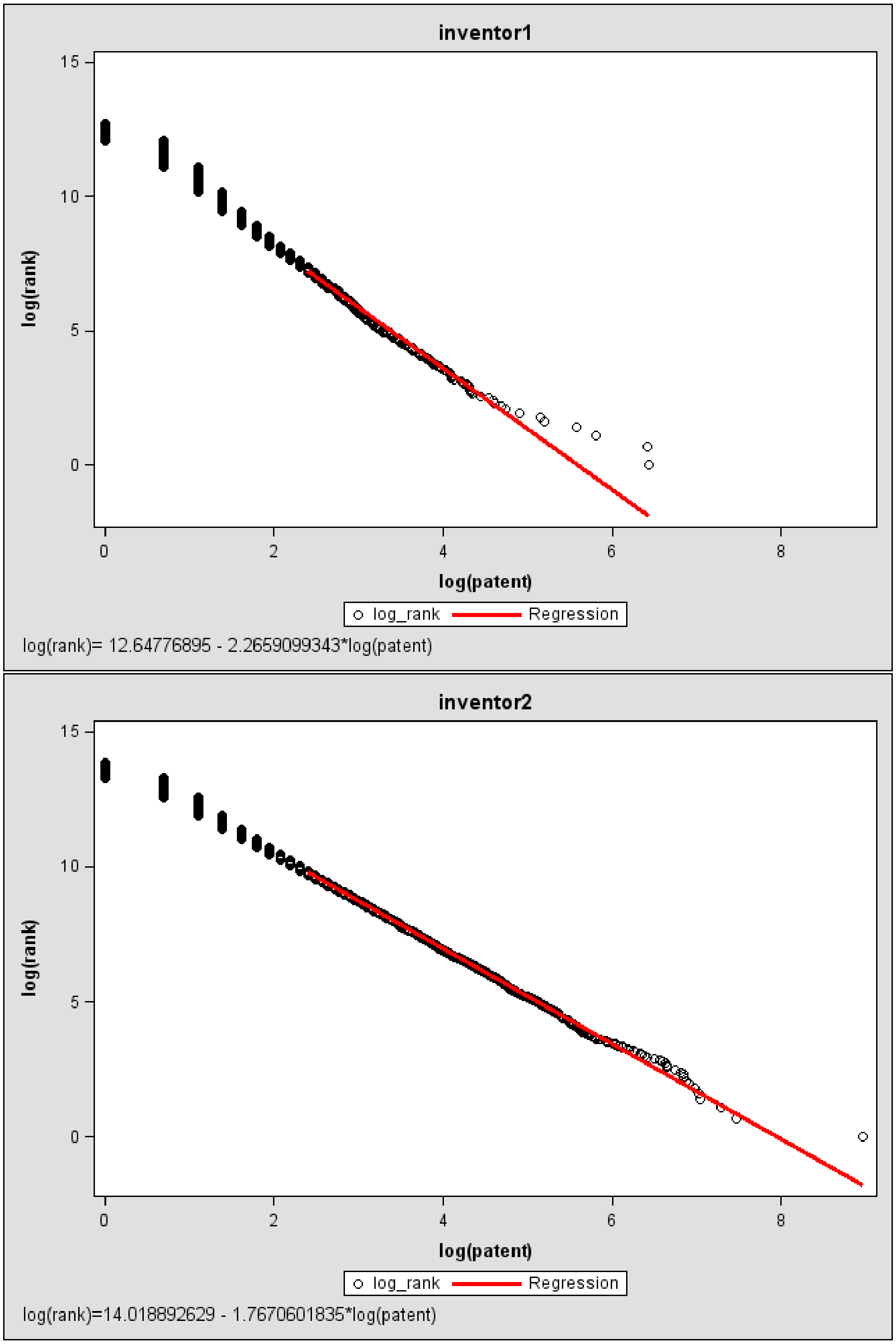
Table 3.8: Comparative Statics

variable	obs	estimate	std err	Adj R^2	p value
$log(patent_1)$	1581	-2.2659	0.0079	0.9813	< 0.0001
$log(patent_2)$	19200	-1.7671	0.0005	0.9985	< 0.0001

3.6 Conclusion

From 1985, especially after 2001, the Chinese patents quantity increase dramatically. Also, an increasing number of inventors entered this market. This paper studies the rise of patenting in China. The particular question I address is the role

Figure 3.14: Comparative Statics



inventor1

log(rank)= 12.64776895 - 2.2659099343*log(patent)

inventor2

log(rank)=14.018892629 - 1.7670601835*log(patent)

of individual inventors in the growth of Chinese patenting. A variety of phenomena in economics and nature occurs in a skewed fashion and relationships are often illustrated in a plot of log rank on log size. In this paper, I do the similar work and document the pattern of Chinese inventors. I use this regression to measure the dispersion of the inventors and find that, there is more and more dispersion among all inventors in China. Thus, I develop a theory that can jointly account for why patenting levels are increasing overall and why the distribution is becoming more skewed. The idea is based on the theory of the optimal allocation of inputs. The main reason for this phenomenon is that the resource distribution is more efficient than before. Another factor is that there are much more research resource now in the Chinese patent market as economic growth. I also calibrate the model with some parameter and get the same pattern for dispersion of inventors in different time periods.

This finding raises several other interesting questions. The first one I already mentioned in data section. It seems all countries inventors have the same pattern and more dispersion than before. But the question is why some has more than 20% change and the other only have 3%. My model answers the question of Chinese inventors. In the future, I want to answer similar questions for American and Japanese inventors. The other question is to see if this pattern of the dispersion happens in different type of patents. Here I only finish the analysis for different time periods. But because utility patents are usually easier to invent than invention patent, what happens to the dispersion of inventors in different type? Usually we will think that utility inventors are less dispersion since they are easy, but is that match the real data? Another question is to see the specific industry. I have the data with classification numbers and can know patents belongs to which industry. If we only look at the patents, for example, in textile industry, what kind of patterns can we get? Should I see the same dispersion of inventors? All these questions are pretty interesting and a lot of work can be done.

4 Concludsion

- Chapter 2 shows that all the countries have the learning-by-doing effect, but only the big country can internalize it and has a much larger effect. This is the main reason why a big country should subsidies exports by controlling the nominal exchange rate. The numerical results also prove the concludsion. For future research, there are three possible directions. The first is to expand the model from a two-period to a multi-period dynamic version. The second is to use the same method to the world import market. And the last way is to think about the micro base of the behavior I describe in this chapter.
- Chapter 3 documents the pattern of Chinese inventors and using the log scale regression to measure the dispersion of them. I found that the distribution is becoming more skewed. The main reason for this phenomenon is that the resource distribution is more efficient than before. For future research, there are three possible directions. The first one is to answer the similar question for American and Japanese inventors. The second one is to see if this pattern of the dispersion happens in different type of patents. And the last question is to see the specific industry.

Bibliography

Arrow, K. J., 1962. The economic implication of learning by doing. The Review of Economic Studies 29, 155–173.

Atkeson, A., Burstein, A., 2007. Pricing-to-market in a ricardian model of international trade. The American Economic Review 97 (2), 362–367.
URL `http://www.jstor.org/stable/30034476`

Atkeson, A., Burstein, A., 2008. Pricing-to-market, trade costs, and international relative prices. The American Economic Review 98 (5), 1998–2031.

Axtell, R., 2001. Zipf distribution of u.s. firm sizes. Science 293, 1818–1820.

Azoulay, P., Zivin, J. S. G., Sampat, B. N., 01 2011. The diffusion of scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine.

Backus, D. K., Kehoe, P. J., Kehoe, T. J., 1992. In search of scale effects in trade and growth. Journal of Economic Theory 58, 377 409.

Bagwell, K., Staiger, R. W., 2001. Strategic trade, competitive industries and agricultural trade disputes. Economics & Politics 13, 113 – 128.

Baldwin, R., Krugman, P. R., 1988. Trade Policy Issues and Empirical Analysis. University of Chicago Press, Ch. Industrial Policy and International Competition in Wide-Bodied Jet Aircraft, pp. 45 – 78.

Bottazzi, L., Peri, G., 2003. Innovation and spillovers in regions: Evidence from european patent data. European Economic Review 47, 687–710.

Bradsher, K., 9 2010. On clean energy, china skirts rules. New York Times.
URL `http://www.nytimes.com/2010/09/09/business/global/09trade.html?pagewanted=1`

Brander, J. A., 1995. Handbook of International Economics. Elsevier, Ch. Strategic trade policy, pp. 1395 – 1455.

Brander, J. A., Spencer, B. J., 1985. Export subsidies and international market share rivalry. Journal of International Economics 18, 83 – 100.

Cheung, K.-y., Lin, P., 2003. Spillover effects of fdi on innovation in china: Evidence from the provincial data. China Economic Review 15, 25–44.

Chipman, J. S., 2007. Production and exchange rates in a small open economy. Review of Development Economics 11, 205 – 216.

Gabaix, X., 1999. Zipf's law for cities: An explanation. Quarterly Journal of Economics 114, 739–767.

Goldberg, P. K., Hellerstein, R., 2008. A structural approach to explaining incomplete exchange-rate pass-through and pricing-to-market. The American Economic Review 98 (2), 423–429.

Goldberg, P. K., Knetter, M. M., 1997. Goods prices and exchange rates: What have we learned? Journal of Economic Literature 35 (3), 1243–1272.
URL `http://www.jstor.org/stable/2729977`

Gowrisankaran, G., Holmes, T. J., 2004. Mergers and the evolution of industry concentration: Results from the dominant-firm model. The RAND Journal of Economics 35 (3), 561–582.
URL `http://www.jstor.org/stable/1593708`

Hall, B. H., 06 2004. Exploring the patent explosion.

Hall, B. H., Griliches, Z., Hausman, J. A., 11 1983. Patents and r&d: Searching for a lag structure.

Hall, B. H., Griliches, Z., Hausman, J. A., 09 1984. Is there a lag?

Hausman, J. A., Hall, B. H., Griliches, Z., 08 1981. Econometric models for count data with an application to the patent-r&d relationship.

Helpman, E., Itskhoki, O., Redding, S., July 2010. Inequality and unemployment in a global economy. Econometrica 78 (4), 1239–1283.

Helpman, E., Melitz, M. J., Yeaple, S. R., 2004. Export versus fdi with heterogeneous firms. The American Economic Review 94 (1), pp. 300–316.
URL `http://www.jstor.org/stable/3592780`

Hessler, P., 2010. Country Driving. HarperCollins.

Holmes, T. J., Lee, S., 12 2008. Cities as six-by-six-miles squares: Zipf's law?

Hsieh, C.-T., Klenow, P. J., 2009. Misallocation and manufacturing tfp in china and india. Quarterly Journal of Economics 124, 1403–1448.

Hu, A. G. Z., Jaffe, A. B., 10 2001. Patent citations and international knowledge flow: the cases of korea and taiwan.

Jaffe, A. B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 108 (3), pp. 577–598.
URL `http://www.jstor.org/stable/2118401`

Keller, W., 2002. Geographic localization of international technology diffusion. The American Economic Review 92, 120–142.

Keller, W., 2004. International technology diffusion. Journal of Economic Literature 42, 752–782.

Knetter, M. M., 1989. Price discrimination by u.s. and german exporters. The American Economic Review 79 (1), 198–210.
URL `http://www.jstor.org/stable/1804781`

Knetter, M. M., 1993. International comparisons of pricing-to-market behavior. The American Economic Review 83 (3), 473–486.
URL `http://www.jstor.org/stable/2117529`

Krugman, P., 5 1986. Pricing to Market when the Exchange Rate Changes.

Luttmer, E., 2007. Selection, growth, and the size distribution of firms. Quarterly Journal of Economics 122, 1103–1144.

Riordan, M. H., 1998. Anticompetitive vertical integration by a dominant firm. The American Economic Review 88, 1232–1248.

Sanyal, P., Jaffe, A. B., 04 2004. Peanut butter patents versus the new economy: Does the increased rate of patenting signal more invention or just lower standards?

Serrano, C. J., 2010. The dynamics of the transfer and renewal of patents. The Rand Journal of Economics 41 (4), 686–708.