International Fragmentation and Work Effort: Networks, Loyalty and Wages*

Soonhee Park**

University of Michigan

Abstract

This study examines how physical networks, loyalty and wages affect international fragmentation by a multinational firm and the work effect of its employees in a world composed of a developed country and a developing country. Fragmentation improves the firm's ability to monitor. This increases employees' effort. However, the headquarters cost for coordinating and monitoring the whole production process increases with the number of production stages. This cost limits the total stages. The international differences in networks, border barriers, loyalty and wages split the stages into the North and the South. *(i)* An increase in the Southern network increases outsourcing from North to South. The effort of the South relative to the North does not change. *(ii)* With globalization, Southern employees' loyalty falls. Their effort relative to Northern employees' effort falls. This decreases outsourcing. *(iii)* A rise in the Southern wage increases effort of the South relative to the North, and does not affect outsourcing. *(iv)* A rise in the Northern wage increases Northern effort relative to Southern effort. Its effect on outsourcing is ambiguous. This paper provides a new perspective on the motive of production fragmentation and firms' organization decision by introducing the possibility of shirking and imperfect monitoring of employee effort.

Keywords: international fragmentation, work effort, shirking, networks, loyalty of employees, wages

JEL Classification: F12, F16

---

* I am particularly grateful to Alan Deardorff for his guidance and insightful comments. I would like to thank John Laitner, Robert Stern and Katherine Terrell for valuable comments and suggestions. I also thank Eunyoung Hwang and Rahul Mukherjee for helpful discussions and comments.

** Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109, USA, Tel. (734) 665-6563, E-mail: soonheep@umich.edu

[1] Introduction

The world has seen a marked increase in global economic activities among countries. One of these trends is international fragmentation – firms shift some part of production to a foreign country out of the home country to produce the same product more efficiently. Factor endowment differences, factor intensity differences and increasing returns to scale have been identified as the main reasons for the international fragmentation of production.

Harris (2001) emphasizes increasing returns to scale of global trade networks – a set of links connecting a large number of related demanders for and suppliers of intermediate inputs – that the firm runs, in order to explain the phenomenon of international fragmentation among developed countries. The fixed cost of running these networks is an important determinant of the extent of international fragmentation.[1]

The innovative development of physical network technology has given the world's economies new infrastructure for communications, such as the Internet and telecommunication networks. These have made massive flows of information possible, and coordination of business partners scattered worldwide easier. Thus the development of physical networks increases international trade. Freund and Weinhold (2002) find a positive correlation between the Internet in a foreign country, and U.S. service imports and service exports.[2] Also, the development of physical networks makes it possible to control from one location many business divisions scattered geographically. This prompts the multinational firm's activities. I study how physical networks in a developed country (the North) and a developing country (the South) affect international fragmentation of the vertical multinational firm.

Another strand of literature examines how work effort of workers influences factor prices, employment, trade and welfare.[3] However, I explore how fragmentation of production affects work effort.[4] In an environment where ability of the headquarters to monitor employees is imperfect, employees would tend to shirk, since they feel disutility in making an effort, and since there is the possibility that shirking is not discovered by the headquarters. When the headquarters splits a single-stage production process into multiple-stages, the split plays the role of dividing employees into smaller groups and assigns one group to each stage.

[1] Deardorff (2001 a) explains that business and social networks reduce the costs of trade, and their welfare effect is likely to be beneficial for the world as a whole.
[2] A 10% increase in internet use in a foreign country leads to about a 1.7% increase in the growth of U.S. service imports and a 1.1% increase in the growth of U.S. service exports from 1995 to 1997.
[3] Copeland (1989) examines the effect of cross-country differences in the disutility of effort on the pattern of trade and welfare. Brecher (1992) analyzes the effect of a tariff on imports on employment and social welfare in an efficiency-wage model with explicit monitoring. Matusz (1996) shows that international trade leads to increased employment in both countries in a model in which intermediate inputs are produced under monopolistic competition and wages are determined according to efficiency wages of Shapiro-Stiglitz type (1984). Leamer (1999) examines the effects of factor intensity on effort level and wage at the sectoral level.
[4] As another issue, Deardorff (2001 b) analyzes how fragmentation affects specialization and trade.

This makes it easier to detect the group of employees in which shirking arises. Employees raise their work effort to avoid the penalty of a wage cut. Thus, fragmentation can lead to an increase in work effort in the corresponding country. However, present studies associated with fragmentation seem to have overlooked this characteristic of fragmentation.

As international economies become globalized and free trade zones resulting from an increase in free trade agreements expand, the new competitive business environment makes employment relationships between employees and firms unstable. Firms compete intensely with domestic and foreign firms for markets. The volatility of their business increases and retaining employment becomes more difficult for the worker. Thus the employment contracts of firms are easily broken, or the duration of employment shortens. Hence, loyalty of employees to their firms falls.[5] Since disutility of work effort of the employees with low loyalty is high, their effort level decreases and thus their productivity becomes low. However, the new competitive environment makes the employment relationship more unstable in the South compared to that in the North, since the South, under a more protection that limits free trade, becomes relatively more exposed to international competition than the North. This means loyalty in the South, and thus Southern work effort is relatively reduced. Therefore, loyalty emerges as a factor affecting the determination of outsourcing to the South. I explore this relationship between loyalty, work effort and international fragmentation.[6]

Wages in developing countries have fast increased. Can this increasing wage cause a Northern firm to reduce outsourcing to the South? Since the increasing wage in the South increases production costs, the Northern firm will tend to reduce outsourcing. However, since a higher wage gives Southern employees the incentive to increase effort, the level of effort also rises. This reduces production costs in the South, which will lead to the Northern firm increasing outsourcing. These opposite effects offset each other, so that the Southern wage does not influence the degree of outsourcing. If the wage in developed countries rises, the effect on the degree of outsourcing would depend on whether the wage effect on both the headquarters cost and the production cost dominates the effort effect on the production cost.

Finally, when the Northern firm changes its mode of production from national fragmentation to international fragmentation, I investigate how the number of total stages would change. I show that, compared with national fragmentation, international fragmentation can increase, decrease or not change the number of total stages according to differences in the amount of employment between the modes of production.

[5] See Reichheld and Teal (1996), and Minkler (2004).
[6] I will compare average employee tenures between the developed countries and the developing countries for inferring loyalty differences between the two.

As for the contribution of this paper to methodology, contrary to the existing method that focuses on finding the level of output for an exogenously given number of stages, this paper focuses on how the firm optimally splits a single-stage production process into multiple-stages, for exogenously given output. Also, under international fragmentation, I explain how the firm optimally determines the number of total stages and how the firm splits the stages between the North and the South.

[2] Model

[2-1] Autarkic Economy

This section examines how national fragmentation is determined in the North. National fragmentation is defined as a production pattern locating all production stages within the national border, but all stages are involved in producing a single final good.

Output increases more when the number of production stages increases though the technologies for different stages do not have different factor intensities, so that a firm has an incentive to fragment its production into many stages. This brings on specialization in production.

Increasing returns with respect to specialization can be represented by output increasing by more with an increase in the variety of inputs (components). I develop a Leontief production function that has the characteristic of increasing returns due to specialization.[7] In products assembled from many components – for example, electronics, automobiles and airplanes – each component is produced by technology specific to the respective component, and all components are complementary to each other. If a component is removed, the assembled product does not work. This description matches well with the property of the Leontief production function.

There is a homogeneous final consumption good X which is supplied by a competitive industry. The good X is produced by a production technology that consists of multiple production stages. The primary production factor is labor. First, I consider the production function for good X.

$$x = \min\{z, \frac{el}{b}\}.$$

x is output of the good X. z is units of the intermediate good Z used. l is units of labor employed. b is units of labor that are required to produce one unit of the good X. e is the level of work effort.

[7] The increasing returns production functions of CES form (Ethier 1979) and Cobb-Douglas form (Edwards and Starr 1987) have been used in the literature. These functions have the feature that the inputs are substitutable for each other.

When the production process is split into multiple stages, each stage produces a distinct intermediate good, using labor and an intermediate input produced at the previous stage. The final stage produces the finished final good X . Let the production function for an intermediate good of the nth stage be of the Leontief form.

$$z_n = \min\{z_{n-1}, \frac{e(\bar{n})l_n}{b(\bar{n})}\}, \quad n = 2, \ldots, \bar{n}.$$

z_n is output of an intermediate good Z_n produced by the nth stage in the production with $\bar{n}$ stages for X . z_{n-1} is units of an intermediate input provided by the $(n-1)$th stage. l_n is units of labor employed at the nth stage.

The productivity of each stage depends on the degree of labor division; greater division of labor makes labor more productive. If the degree of labor division is identical to the number of total stages, labor productivity $(= \frac{1}{b(\bar{n})})$ rises with the number of total stages $\bar{n}$. The reason is that if $\bar{n}$ increases, the units of labor $b(\bar{n})$ that each stage uses to produce one unit of the intermediate good become smaller. I define a functional form for $b(\bar{n})$ as follows:

$$b(\bar{n}) = (\frac{1}{\bar{n}})^{\delta}, \quad \text{where } \delta \geq 1, \ \frac{\partial b(\bar{n})}{\partial \bar{n}} < 0. \tag{1}$$

The condition $\delta \geq 1$ stands for the increasing returns due to labor division – specialization. It is well known that the division of labor raises the productivity of the laborers working at each stage and leads to increasing returns to specialization. However, for a given number of total stages $\bar{n}$, the labor productivity in each stage is the same, $\frac{1}{b(\bar{n})}$.

The work effort is explained. I assume that the firm’s ability to monitor the employees who work in production is imperfect. For better checking of the employees’ performance, the firm splits its production into multiple-stage production. The split plays the role of dividing the employees into groups with a narrower range of activities. This makes it easier to identify the group of employees in which shirking arises. If the number of total stages increases, the firm can monitor the performance of the employees more effectively. This causes the employees to work harder since they know that they will be penalized if idleness is detected. Thus their effort is associated positively with the number of total stages $\bar{n}$: $\frac{\partial e(\bar{n})}{\partial \bar{n}} > 0$.

I now address how labor is employed by each stage when the firm has $\bar{n}$ stages. The first stage produces its intermediate good Z_1 with only labor. The production function of the first stage is

$$z_1 = \frac{e(\bar{n})l_1}{b(\bar{n})}.$$

z_1 is output of the intermediate good produced at the first stage. l_1 is labor employed by this stage. The amount of labor of the first stage is $l_1 = \frac{b(\bar{n})z_1}{e(\bar{n})}$. The labor employed per unit output of the intermediate good Z_1 is $\frac{b(\bar{n})}{e(\bar{n})}$.

The second stage produces an intermediate good Z_2. This stage uses the intermediate good Z_1 produced in the first stage as an intermediate input and labor. Its production technology is

$$z_2 = \min\{z_1, \frac{e(\bar{n})l_2}{b(\bar{n})}\}.$$

z_2 is output of Z_2. z_1 is units of intermediate good Z_1 used. Suppose that the second stage produces $\bar{z}_2$ units of Z_2. Since the input-output coefficient on Z_1 is 1, the second stage uses Z_1 in the amount $\bar{z}_2$. Thus z_1 becomes $\bar{z}_2$ units: $z_1 = \bar{z}_2$. This means that the first stage will provide $\bar{z}_2$ units of Z_1. Then the intermediate input used at the second stage embodies $\frac{b(\bar{n})\bar{z}_2}{e(\bar{n})}$ units of labor since one unit of Z_1 is produced with $\frac{b(\bar{n})}{e(\bar{n})}$ units of labor. The embodied labor $\frac{b(\bar{n})\bar{z}_2}{e(\bar{n})}$ becomes the labor demand of the first stage. The second stage also employs labor, $l_2 = \frac{b(\bar{n})\bar{z}_2}{e(\bar{n})}$. Thus total labor embodied for the production of Z_2 is the sum of these two labor demands, $\frac{b(\bar{n})\bar{z}_2}{e(\bar{n})} + l_2 = \frac{2b(\bar{n})\bar{z}_2}{e(\bar{n})}$. The labor employed per unit of Z_2 is $\frac{2b(\bar{n})}{e(\bar{n})}$.

When this logic is generalized to the nth stage, the labor employed is explained as follows. The production function of the nth stage is

$$z_n = \min\{z_{n-1}, \frac{e(\bar{n})l_n}{b(\bar{n})}\}, \quad n \geq 2. \tag{2}$$

If output of the nth stage is $\bar{z}_n$, this stage demands $\bar{z}_n$ units of Z_{n-1} as an intermediate input: $z_{n-1} = \bar{z}_n$. Since one unit of Z_{n-1} at the $(n-1)$th stage is produced with $\frac{(n-1)b(\bar{n})}{e(\bar{n})}$ units of labor,[8] $\bar{z}_n$ units of Z_{n-1} embody $\frac{(n-1)b(\bar{n})\bar{z}_n}{e(\bar{n})}$ units of labor that are accumulated up to the $(n-1)$th stage. Also the nth stage employs labor, $l_n = \frac{b(\bar{n})\bar{z}_n}{e(\bar{n})}$. Thus the total labor employed, which is accumulated until the nth stage, is $\frac{(n-1)b(\bar{n})\bar{z}_n}{e(\bar{n})} + l_n = \frac{nb(\bar{n})\bar{z}_n}{e(\bar{n})}$.

The production cost at the nth stage PC_n is the product of the total labor employed and the wage.

$$PC_n = \frac{nb(\bar{n})\bar{z}_n w}{e(\bar{n})}, \quad \text{where } n \le \bar{n}. \tag{3}$$

This cost is a cumulative cost from the first stage to the nth stage. As n becomes larger, PC_n increases. As the number of total stages $\bar{n}$ increases, PC_n decreases since labor productivity rises with $\bar{n}$ (in other words, $b(\bar{n})$ is decreasing in $\bar{n}$). As $\bar{z}_n$ becomes larger and w rises, production cost increases. If the work effort $e(\bar{n})$ rises, the cost falls since the labor input to produce one unit of good is reduced. Its production cost per unit output of the good is

$$pc_n = \frac{nb(\bar{n})w}{e(\bar{n})}. \tag{3'}$$

I call this the unit production cost.

The input-output coefficients on intermediate goods Z_n, $n \in \{1, \dots, \bar{n}\}$, are one, so that the demand equals the supply for the respective intermediate goods. This makes the relationships among outputs of intermediates Z_n, $n = 1, \dots, \bar{n}-1$, and final good X at $\bar{n}$ be

$$\bar{z}_1 = \bar{z}_2 = \dots = \bar{z}_n = \dots = \bar{z}_{\bar{n}-1} = \bar{z}_{\bar{n}}. \tag{4}$$

The output at the final stage $\bar{n}$ equals the output of the final good X: $\bar{z}_{\bar{n}} = \bar{x}$.

---

[8] Since one unit of Z_1 is produced with $\frac{b(\bar{n})}{e(\bar{n})}$ units of labor, one unit of Z_2 is produced with $\frac{2b(\bar{n})}{e(\bar{n})}$ units and Z_3 is produced with $\frac{3b(\bar{n})}{e(\bar{n})}$ units, one unit of Z_{n-1} should be produced with $\frac{(n-1)b(\bar{n})}{e(\bar{n})}$ units of labor.

However, the firm cannot expand endlessly the number of stages so as to increase the productivity of labor employed. Difficulty in operating all the production stages would rise as the number of production stages increases. Coordinating and monitoring the whole production process are necessary in order to lead to the optimal performance of production. The role of the headquarters is to provide the services of coordination and monitoring. The headquarters cost for providing these services rises as the number of production stages increases. This would limit the number of stages.[9] Also, I assume that the headquarters cost of providing the services increases if the output of each stage $\bar{z}_n$ increases. As a variable indexing the headquarters service, I use total output which is increasing both in the number of stages and in the output level of each stage:

$$\sum_{n=1}^{\bar{n}} \bar{z}_n \ . \tag{5}$$

The development of physical networks like the Internet and telecommunications network makes it easier for a firm to access to networks, and to exchange information faster and in greater quantities. Traveling time and cost of managers for coordination and monitoring decrease. Thus the network improves the efficiency of the headquarters services. The home country and foreign country are both assumed to have networks, which are considered public goods. The development of network technology has greatly reduced the cost of accessing networks, so that the cost is very small relative to other costs. Thus every firm is assumed to use the networks free.

To reflect the effect of the network on the headquarters cost, I need to define accessibility to the network. N is the network size of the economy. This should be a large value such that $N > 1$. At the nth stage, the degree of accessibility to the network for each stage of the firm is assumed to be $\alpha_n(N)$. This has a value between zero and one. The assumption of partial accessibility, $0 < \alpha_n(N) < 1$, implies that the individual firm uses only a part of the network in the economy. If the economy has a larger network, the accessibility of each stage to the network increases: $\frac{\partial \alpha_n(N)}{\partial N} > 0$. An increase in accessibility causes improvement in efficiency. I assume that the stage n obtains improvement in efficiency by $\alpha_n(N)$ times its output $\bar{z}_n$ which is the index of the headquarters services of the corresponding stage: $\alpha_n(N)\bar{z}_n$. Then the total improvement in efficiency

[9] Adam Smith (1937) explains that the division of labor (specialization) is limited by market size. Edwards and Starr (1987) explain that it is limited by indivisibilities of labor or setup costs in the transition of labor between production tasks. In Becker and Murphy (1992), specialization depends on the cost of coordinating specialized workers who perform complementary tasks, and on the extent of knowledge available.

for the firm with $\bar{n}$ stages is the sum of the improvement in efficiency of individual stages: $\sum_{n=1}^{\bar{n}} \alpha_n(N)\bar{z}_n$. For simplicity, I assume that every stage has the same degree of accessibility $\alpha_n(N) = \alpha(N)$ for all $n \in \{1, \ldots, \bar{n}\}$. Then the headquarters gets efficiency of

$$\sum_{n=1}^{\bar{n}} \alpha_n(N)\bar{z}_n = \alpha(N)\sum_{n=1}^{\bar{n}} \bar{z}_n \ . \qquad (6)$$

I represent the degree of accessibility $\alpha(N)$ as an index and specify the function constructing the index.

$$\alpha(N) = (1 - \frac{1}{N}), \quad \text{where } N > 1,\ 0 < \alpha(N) < 1,\ \frac{\partial \alpha(N)}{\partial N} = \frac{1}{N^2} > 0\ . \qquad (7)$$

When the effect of the network is considered, the index for actual headquarters services is the difference between the headquarters services (5) and the improvement in efficiency (6): $\{1 - \alpha(N)\}\sum_{n=1}^{\bar{n}} \bar{z}_n$. This is re-expressed, using $\{1 - \alpha(N)\} = \frac{1}{N}$ from (7),

$$\frac{1}{N}\sum_{n=1}^{\bar{n}} \bar{z}_n \ . \qquad (8)$$

The employees at the headquarters play the roles not only of coordinator but also of monitor. These jobs are given many responsibilities for performance similar to those of the owner or principal. This makes employees at headquarters become highly loyal to the firm and work honestly. Thus, I assume that these employees do not shirk.[10] This means that the headquarters employees provide effort $e^{\max}$, which is the upper bound of the set of possible effort levels that they can provide physically: $e^{\max} > e$, for all e. Let θ be the input requirement of effective labor – effective labor is defined as units of labor times an effort level per unit of labor – per unit of headquarters service. Then, the number of units of labor included in θ is $\frac{\theta}{e^{\max}}$ since labor in the headquarters provides effort level $e^{\max}$. Note that θ has a small value relative to the network size of the economy N: $\theta < N$. The amount of labor employed for the headquarters is, from (8),

$$\frac{\theta}{e^{\max} N}\sum_{n=1}^{\bar{n}} \bar{z}_n \ , \quad \text{where } \theta < N\ . \qquad (9)$$

[10] In reality, the employees who work in the headquarters may shirk. However, I do not consider this case. The reason is that if they are shirkers, their role as monitor of the employees who work in production could be compromised.

The headquarters cost HC is the wage cost for coordination and monitoring. Using (9) and (4),

$$HC = \frac{\theta w \bar{z}_n \bar{n}}{e^{\max} N} . \tag{10}$$

This cost rises if the unit effective labor input requirement for coordination and monitoring θ, the wage w, the output of each stage $\bar{z}_n$ and the number of total stages $\bar{n}$ increase. However, the cost falls if the network size N is larger. The headquarters cost per unit of output is

$$hc = \frac{\theta w \bar{n}}{e^{\max} N} , \tag{10'}$$

which is the unit headquarters cost.

The total cost of the firm for production up to the nth stage is the sum of the headquarters cost (10) and the production cost (3).

$$TC_n = HC + PC_n = w\bar{z}_n \{ \frac{\theta \bar{n}}{e^{\max} N} + \frac{nb(\bar{n})}{e(\bar{n})} \} . \tag{11}$$

The total cost per unit of output of the good up to the nth stage becomes

$$tc_n = w\{ \frac{\theta \bar{n}}{e^{\max} N} + \frac{nb(\bar{n})}{e(\bar{n})} \} . \tag{11'}$$

I call this the unit total cost. The first term in the square brackets is the unit headquarters cost. This is affected by the parameters characterizing a country, such as θ and N. This is a country-specific cost, which is neutral across the interim stages. The second term in the square brackets is the unit production cost. Since it is cumulative, the unit total cost tc_n (11') is also a cumulative cost.

For clear understanding of the unit total cost, I will graphically explain tc_n with Figure 1. The unit total cost at the first stage, tc_1, is the sum of $hc = \frac{\theta w \bar{n}}{e^{\max} N}$ and pc_1 $(= \frac{b(\bar{n})w}{e(\bar{n})})$. The unit total cost accumulated until the second stage, tc_2, is hc plus pc_2 $(= \frac{2b(\bar{n})w}{e(\bar{n})})$. Then the unit total cost accumulated until the nth stage, tc_n, is the sum of hc and pc_n $(= \frac{nb(\bar{n})w}{e(\bar{n})})$.

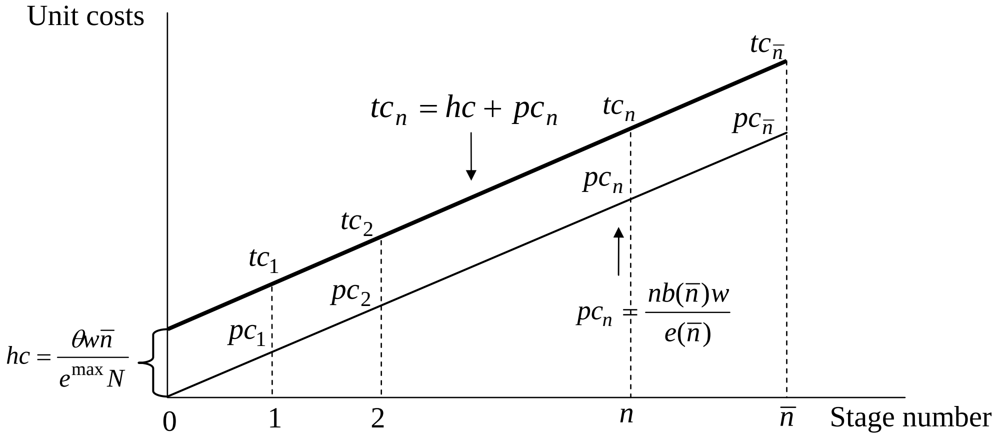

Figure 1

I explain the costs of producing the final good. Recall that the output at the $\bar{n}$ th stage is $\bar{x}$. Since the final good X is produced at the final stage $\bar{n}$ of the production process, the total cost and the unit total cost of the final stage are obtained by substituting $\bar{n}$ for n in TC_n (11) and tc_n (11'). Substituting $b(\bar{n}) = (\frac{1}{\bar{n}})^{\delta}$ in (1), $TC_{\bar{n}}$ and $tc_{\bar{n}}$ are

$$TC_{\bar{n}} = w\bar{x}\{\frac{\theta\bar{n}}{e^{\max}N} + \frac{\bar{n}^{1-\delta}}{e(\bar{n})}\}, \tag{12}$$

$$tc_{\bar{n}} = w\{\frac{\theta\bar{n}}{e^{\max}N} + \frac{\bar{n}^{1-\delta}}{e(\bar{n})}\}. \tag{12'}$$

[2-1-1] Decision of Effort by Employee

An employee provides effort to the firm. The employee is compensated by a wage, which finances consumption. Thus, she obtains higher utility, but feels disutility from effort. She determines a level of effort where the difference between the utility and disutility is maximized. The employee has consumption c and her utility is $u(c)$. She makes an effort e, $0 \le e \le e^{\max}$, and feels disutility $v(e)$. A low level of effort means that the employee is very idle. $e = 0$ stands for complete idleness. If the employee provides full effort without shirking, the level of effort is $e^{\max}$ and she becomes a fully effective worker.

The disutility of effort is also affected by the employee's work ethic. Sennett (1998) and Minkler (2004) contend that work ethic is influenced by the economic environment. With the development of the economy and globalization, the firm faces a more competitive environment in domestic and international markets. The volatility of its business increases, and its survival becomes more difficult. This situation makes for unstable employment relations between the firm and its employees. Employment contracts with the firm are easily broken, for example, the possibility of being unemployed rises, or the duration of employment shortens. This causes the employees' work ethic to be diminished: their loyalty to the firm weakens so that their motivation and willingness to work hard are reduced (Reichheld and Teal 1996, and Minkler 2004).

In this paper, I do not deal with how the firm makes employment contracts to cope with the new competitive business environment. Instead, I will focus on how the lower loyalty affects productivity and production cost. Assume that employees with low loyalty have high disutility of effort. This assumption implies that the disutility depends not only on effort but also on loyalty: $v(e;\eta)$. η is the index for loyalty, $0<\eta\leq 1$; a high value of η represents a high level of loyalty.

The firm expects its employees to work with maximal effort, $e^{\max}$. It monitors their performance and punishes the employees who are caught shirking. However, the headquarters' ability to monitor is imperfect. The employees would shirk on the job since effort involves disutility, and since there is a possibility that shirking is not detected by the headquarters. Their performance would be lower compared with that in the absence of shirking. Let P be the probability of an employee's performance being checked by the headquarters. If the performance of an employee is checked and if her level of effort is revealed to be below the level of full effort such that $0\leq e<e^{\max}$, she is penalized.[11] Let $(1-P)$ be the probability of an employee's performance not being checked. If the performance of an employee is not checked, the firm treats the employee as a non-shirker, and thinks that she works with full effort $e^{\max}$.[12] Under this scheme, if the detection probability is $P=0$, the employee provides no effort. The equilibrium level of effort is zero. If the probability is $P=1$, her performance is perfectly revealed and thus she will provide full effort. The equilibrium level of effort is $e^{\max}$.

The employee can face two states. One is that she is not checked, with probability $(1-P)$, and is considered a non-shirker working with full effort $e^{\max}$. She receives a wage w from the firm, which is

[11] Copeland (1989), Brecher (1992), and Matusz (1996) based on the efficiency wage model assume that if shirking of the employee is caught, she is fired. However, I follow the assumption of Calvo and Wellisz (1978), that is, when the shirker is caught, she receives the penalty that her wage is cut.

[12] Though the firm treats the employee as a non-shirker, it acknowledges that she can still shirk. However, without any evidence on her performance, if the firm regards her as a shirker, this incurs complaints from her and lowers her productivity. However, if the firm regards her as a non-shirker (i.e., a fully effective worker), this can (or cannot) increase her motivation to work hard. Anyway, the latter case of regarding a worker as the non-shirker can become beneficial to the firm than the former case of regarding the worker as the shirker.

measured in units of the final good X . Assuming that the wage income is consumed, consumption is $c = w$. The net utility is $u(w) - v(e;\eta)$.

The other state, which occurs with probability P , is that she is checked and is penalized if her effort is less than the full effort, $e < e^{\max}$. She is then paid a wage that is prorated to her effort. Since the wage per unit of effort that the firm considers is $\frac{w}{e^{\max}}$, the paid wage corresponding to her effort e is $\frac{ew}{e^{\max}}$. Her consumption is $c = \frac{ew}{e^{\max}}$. Her net utility is $u(\frac{ew}{e^{\max}}) - v(e;\eta)$.

Specific functional forms for $u(c)$ and $v(e;\eta)$ are defined.

$$u(c) = 2\sqrt{c}\ , \quad \text{where } \frac{\partial u}{\partial c} = \frac{1}{\sqrt{c}} > 0\ , \ \frac{\partial^2 u}{\partial c^2} = \frac{-1}{2\sqrt{c^3}} < 0\ .$$

I define the disutility of the employee in the North with a level of loyalty η as

$$v(e;\eta) = e(2-\eta)\ , \quad 0 < \eta \le 1\ ,$$

$$\text{where } \frac{\partial v(e;\eta)}{\partial e} = 2 - \eta > 0\ , \ \frac{\partial v(e;\eta)}{\partial \eta} = -e < 0\ . \tag{13}$$

The constant 2 in the disutility function makes the level of disutility positive for a given positive effort level. The first sign in (13) means that if effort increases, the disutility increases. The second sign in (13) says that a fall in loyalty η leads to an increase in the disutility of effort.

The employee chooses a level of effort maximizing her expected net utility. Substitute $c = \frac{ew}{e^{\max}}$ for the shirker or $c = w$ for the non-shirker into $u(c) = 2\sqrt{c}$. Using $v(e;\eta) = e(2-\eta)$, the maximization of the expected utility is

$$\underset{e}{Max}\ EU = P\{2\sqrt{\frac{ew}{e^{\max}}} - e(2-\eta)\} + (1-P)\{2\sqrt{w} - e(2-\eta)\}\ , \quad \text{s.t. } 0 \le e \le e^{\max}\ , \tag{14}$$

$$\text{where } \frac{\partial EU}{\partial e} = Pe^{\frac{-1}{2}}(\frac{w}{e^{\max}})^{\frac{1}{2}} - (2-\eta)\ , \ \frac{\partial^2 EU}{\partial e^2} = -\frac{1}{2}Pe^{\frac{-3}{2}}(\frac{w}{e^{\max}})^{\frac{1}{2}} < 0\ . \tag{15}$$

The maximization problem restricts effort to be $0 \le e \le e^{\max}$. In view of the restriction, three possible situations may arise: (i) EU is maximized at $e = 0$, (ii) EU is maximized at $0 < e < e^{\max}$ and (iii) EU is maximized at $e^{\max}$.

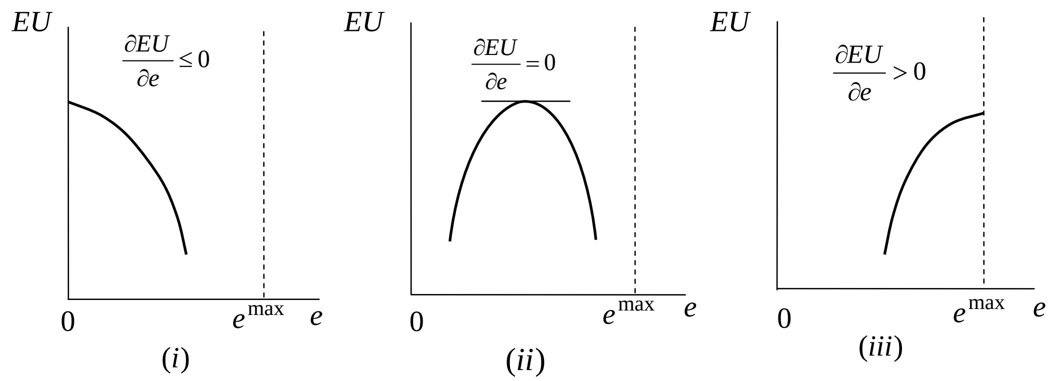

Figure 2

As e approaches zero, $\frac{\partial EU}{\partial e}$ in (15) approaches positive infinity. This sign is not what case (i) requires as shown in Figure 2. Thus the optimal effort is not zero. The optimal effort would be determined in either of the cases (ii) or (iii). From the constrained maximization problem of expected utility, case (ii) yields a higher expected utility than case (iii).[13] Thus case (ii) is chosen for the optimal effort.

The optimal effort is an interior solution and is smaller than $e^{\max}$:

$$e = \frac{wP^2}{e^{\max}(2-\eta)^2} < e^{\max}, \quad \text{where } 0 < P \le 1,\ 0 < \eta \le 1. \tag{16}$$

A higher wage gives the employee the incentive to be willing to raise her level of effort. Also, a higher probability of being checked raises the possibility that the employee is penalized, thus she raises her work effort. If the employee has high loyalty, she provides high effort.

I now address the detection probability. When the headquarters splits the single-stage production process into multiple-stages, the split gives it more detailed information about the stage in which shirking occurs. This raises the detection probability. A probability function showing a positive relation between the detection probability and the number of total stages can be explained as follows. First, imagine a single stage production process that produces the final good. The firm is able to check the employees' performance only after the production of the good is completed, since the output is observed only at that time. However, all employees produce the same good in one stage, so the firm cannot observe exactly each individual's performance. Instead, it observes the performance of the pool of employees. Thus the possibility that a single employee's performance is checked depends on her portion of the employee pool, that is, the number of employees. Since

[13] The details are in Appendix 1.

the firm competes in a perfectly competitive environment, its output is assumed to be exogenous.[14] Thus the number of employees, l, can be considered as given. This number is assumed to be large. The employees are assumed to be distributed uniformly over the line in the range $(0, l]$ as (i) in Figure 3. Then the size of the employee pool is represented by this line.

Consider a single employee in the pool who is represented as one point, such as an arbitrary point, t, on the line. The length of the range in which she is included is $\frac{l}{1}$, where one in the denominator is the number of total stages. This length of the range affects the probability that she is checked; I denote the probability simply as a function of the length of the range: $P(\frac{l}{1})$.

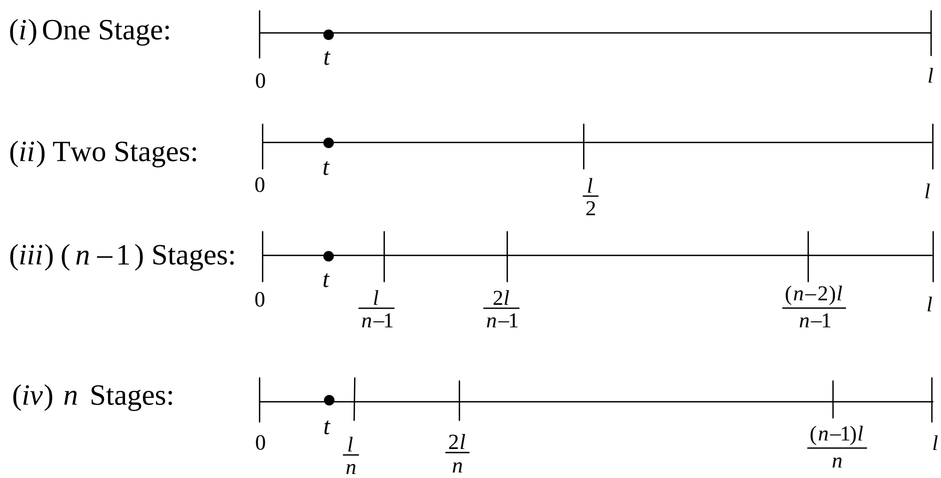

Figure 3

If the firm changes the production mode from single-stage production to two-stage production, the probability that she is checked would change. The first stage of the two-stage production produces a distinct intermediate good that is assembled into the final good. The second stage produces the final good. The employees engaged in production would be rearranged in two groups on the assumption that each employee is assigned to only one stage. One group is for the first stage that is represented by the range $(0, \frac{l}{2}]$ as (ii) in

[14] In perfect competition, price equals average cost and marginal cost. Since the firm in my model has constant marginal cost, this equilibrium condition does not pin down a level of output. Also, this model is a partial equilibrium model that does not consider the factor market. Thus, the output level should be treated as given exogenously.

Figure 3. The other is for the second stage which is represented by the range $(\frac{l}{2},\ l]$. Since the headquarters observes the output of each stage, if shirking occurs, it knows the group in which shirking occurs. Her location at t belongs to the range of $(0,\ \frac{l}{2}]$. The range that the firm has to scrutinize, conditional on the shirker belonging to the first stage, would decrease to half of the size of the pool, $\frac{l}{2}$. Then the probability that she is detected becomes higher than the probability in the case of the single-stage production mode: $P(\frac{l}{1}) < P(\frac{l}{2})$.

The same logic works for the general case of n stages. If production is split into n stages, the range is split into n parts as (iv) in Figure 3. Her location at t is in $(0,\ \frac{l}{n}]$. The length of the range in which she is included is $\frac{l}{n}$. This is shorter than $\frac{l}{n-1}$ that is the length under the $(n-1)$-stages as (iii). Then the probability that she is detected under the n-stage production becomes higher than that under the $(n-1)$-stage production: $P(\frac{l}{n-1}) < P(\frac{l}{n})$. Thus the probability is increasing in the number of stages. $P(\frac{l}{1}) < P(\frac{l}{2}) < ... < P(\frac{l}{n-1}) < P(\frac{l}{n})$. This also means that the detection probability has an inverse relation to the length of the range in which the employee is included.

As the length of the range in which the employee is included becomes shorter, the probability that she is detected becomes higher. However, the degree of the increase in the detection probability decreases as the length of the range shortens. The reason is as follows. The increase in the number of stages makes the range in which she is included smaller, so that the risk of detection becomes higher; the higher risk makes her become more cautious and shirk less; thus the rate of increase in the detection probability decreases if the number of stages increases. To capture all these characteristics, the probability is expressed as

$$P(\frac{l}{n}) = \sqrt{\frac{1}{l/n}}, \quad \text{where } \frac{\partial P}{\partial n} = \frac{n^{\frac{-1}{2}}}{2\sqrt{l}} > 0, \ \frac{\partial^2 P}{\partial n^2} = \frac{-n^{\frac{-3}{2}}}{4\sqrt{l}} < 0. \tag{17}$$

If $n = l$, each stage has only a single employee. This says that monitoring is perfect and the detection probability is one, $P(\frac{l}{n}) = 1$. If $n < l$, the detection probability lies in the range $0 < P(\frac{l}{n}) < 1$. If $n > l$, this means that a single employee works in multiple stages. However, the firm is assumed to allocate more than one of its employees to each of the stages, so that I exclude this case of $n > l$.

Since a larger number of total stages increases the detection probability, the work effort of employees increases with the number of total stages. Replacing n in (17) by $\overline{n}$ and substituting (17) for P in (16),

$$e(\overline{n}) = \frac{w}{e^{\max}(2-\eta)^2}(\frac{\overline{n}}{l}) , \tag{18}$$

where $$\frac{\partial e(\overline{n})}{\partial \overline{n}} = \frac{w}{e^{\max}(2-\eta)^2 l} > 0 , \quad \frac{\partial^2 e(\overline{n})}{\partial \overline{n}^2} = 0 . \tag{19}$$

The first equation in (19) shows that the effort level increases if the number of total stages increases. And the marginal increase in effort with respect to the number of total stages is constant. Thus the effort function (18) is linear with respect to $\overline{n}$. If $\overline{n} = 1$, $e = \frac{w}{e^{\max}(2-\eta)^2 l} > 0$, where $0 < \eta \leq 1$. Also, since an employee cannot provide more than $e^{\max}$, (18) can determine the upper bound of the range of the number of total stages in which effort is provided. This happens where the obtained effort becomes equalized to the maximum effort: $\overline{n} = \frac{\{e^{\max}(2-\eta)\}^2 l}{w}$. Figure 4 illustrates the above relation between the number of total stages and the effort level in the range, $1 \leq \overline{n} \leq \frac{\{e^{\max}(2-\eta)\}^2 l}{w}$.

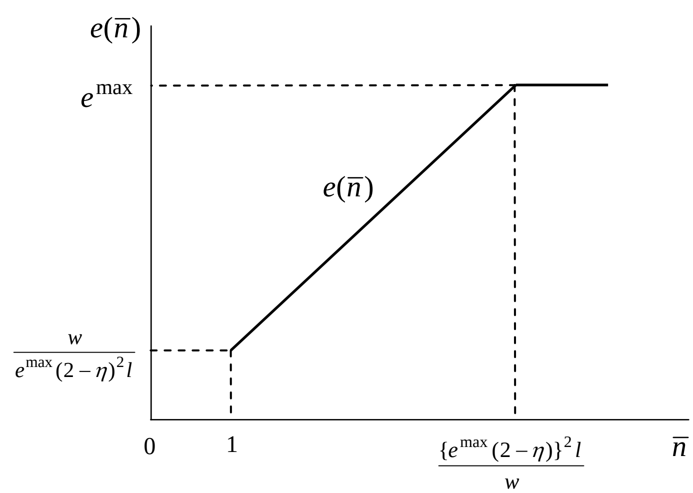

Figure 4

[2-1-2] Equilibrium in National Fragmentation

Where the unit total cost of producing the final good is minimized, the firm optimally determines the number of the final stage, that is, the number of total stages. From (12’), the unit total cost for the final good is $tc_{\bar{n}} = \frac{w\theta\bar{n}}{e^{\max}N} + \frac{w\bar{n}^{1-\delta}}{e(\bar{n})}$. The first term on the RHS is the unit headquarters cost. This rises with the number of the final stage, and thus hc is linearly increasing in $\bar{n}$ as shown in Figure 5. The second term is the unit production cost. Since effort and labor productivity rise as the number of the final stage increases, the unit production cost $pc_{\bar{n}}$ is decreasing in $\bar{n}$ as in the following equation (20). Here, I define elasticity of effort with respect to the number of the final stage as $\varepsilon = \frac{\partial e(\bar{n})}{\partial \bar{n}} \frac{\bar{n}}{e(\bar{n})}$. Since $\frac{\partial e(\bar{n})}{\partial \bar{n}} > 0$ from (19), ε is positive.

Then the first derivative of the unit production cost at $\bar{n}$ is expressed as

$$\frac{\partial pc_{\bar{n}}}{\partial \bar{n}} = \frac{(1-\delta-\varepsilon)w\bar{n}^{-\delta}}{e(\bar{n})} < 0, \quad \text{where } \delta \geq 1,\ \varepsilon > 0,\ \bar{n} \leq \frac{\{e^{\max}(2-\eta)\}^2 l}{w}. \tag{20}$$

The second derivative is

$$\frac{\partial^2 pc_{\bar{n}}}{\partial \bar{n}^2} = \frac{-(\delta+\varepsilon)(1-\delta-\varepsilon)w\bar{n}^{-(\delta+1)}}{e(\bar{n})} > 0. \tag{21}$$

The conditions, (20) and (21), say that $pc_{\bar{n}}$ is downward sloping and convex as seen in Figure 5.

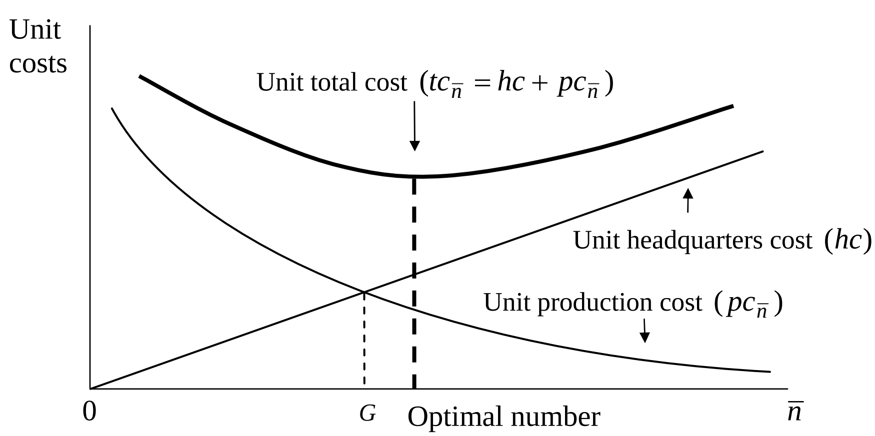

Figure 5

When hc and $pc_{\bar{n}}$ are summed, the graph of $tc_{\bar{n}}$ is convex. The optimal number of stages is determined where $tc_{\bar{n}}$ is minimized with respect to $\bar{n}$ as shown in Figure 5.[15] The first and second order conditions for the cost minimization are

$$\frac{\partial tc_{\bar{n}}}{\partial \bar{n}} = w\{\frac{\theta}{e^{\max} N} + \frac{(1-\delta-\varepsilon)\bar{n}^{-\delta}}{e(\bar{n})}\} = 0\ , \ \frac{\partial^2 tc_{\bar{n}}}{\partial \bar{n}^2} = \frac{\partial^2 pc_{\bar{n}}}{\partial \bar{n}^2} > 0\ .$$

The cost-minimizing number of the final stage $\bar{n}$ is

$$\bar{n}(e(\bar{n})) = \{\frac{(\delta+\varepsilon-1)e^{\max} N}{e(\bar{n})\theta}\}^{\frac{1}{\delta}}\ . \tag{22}$$

The number of stages is decreasing in $e(\bar{n})$ and is convex: $\frac{\partial \bar{n}(e(\bar{n}))}{\partial e(\bar{n})} < 0$ and $\frac{\partial^2 \bar{n}(e(\bar{n}))}{\partial e(\bar{n})^2} > 0$. Note that the number of stages is greater than one.[16] The curve for $\bar{n}(e(\bar{n}))$ is depicted in Figure 6. This means that if the effort of employees increases, the necessity of dividing the production process would be reduced, and thus the firm would decrease the number of total stages.

The equilibrium number of total stages $\tilde{\bar{n}}$ and the equilibrium effort $\tilde{e}$ are determined by equations (18) and (22). Substituting (18) for $e(\bar{n})$ in (22),

$$\tilde{\bar{n}} = [\frac{(\delta+\varepsilon-1)\{e^{\max}(2-\eta)\}^2 lN}{\theta w}]^{\frac{1}{\delta+1}}\ , \tag{23}$$

$$\tilde{e} = (e^{\max})^{\frac{1-\delta}{\delta+1}}\{\frac{(\delta+\varepsilon-1)N}{\theta}\}^{\frac{1}{\delta+1}}\{\frac{w}{(2-\eta)^2 l}\}^{\frac{\delta}{\delta+1}}\ .$$

Also, $\tilde{\bar{n}}$ and $\tilde{e}$ can be shown, by putting the graph for $e(\bar{n})$ in Figure 4 and the curve for $\bar{n}(e(\bar{n}))$ together in the same space, as in Figure 6. $\tilde{\bar{n}}$ is smaller than the number of total stages where effort is maximal, $\frac{\{e^{\max}(2-\eta)\}^2 l}{w}$. The employee provides $\tilde{e}$ that is less than $e^{\max}$.

[15] The optimal number can also lie on the LHS of G, at the point G or on the RHS of G. The location depends on functional forms of the unit headquarters cost and the unit production cost. For convenience, the optimal number is located on the RHS of G in Figure 5.

[16] The reasons are: $\delta \geq 1$; $\varepsilon = 1$ from (18) since $e(\bar{n})$ is proportional to $\bar{n}$; $e^{\max} > e(\bar{n})$; and $N > \theta$ in (9).

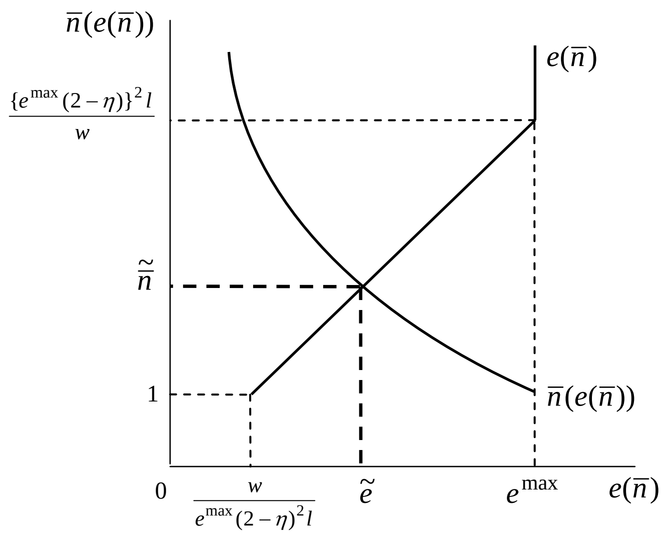

Figure 6

[2-2] Open Fragmented Economies

Revolutionary progress in communication technologies has weakened the link between specialization and geographic concentration.[17] Such progress makes communication across borders possible without physical traveling and loss of time. Thus, regardless of production location, Northern firms can coordinate and monitor all production stages remotely. This causes Northern firms to get the benefit of the low Southern wage, and gives them incentives to relocate their production to the South.

I assume that the firm's headquarters is located in the home country, the North. However, the headquarters faces more difficulty running the Southern stages than the Northern stages. The first reason is Southern border barriers, such as different culture, language and legal system. These require more headquarters services. Second, since the network size of the South is assumed to be smaller than that of the North, the small Southern network makes the Northern firm less efficient. Thus the firm faces a higher headquarters cost. However, the low wage of the South reduces the firm's production cost.

If the reduction in the production cost is larger than the increase in the headquarters cost, and thus if the total cost with international fragmentation becomes lower than that with national fragmentation, the firm would change the mode of production from national fragmentation to international fragmentation. The firm will split the production stages across the two regions as a way of incurring the lowest total cost. I define

[17] See Grossman and Rossi-Hansberg (2007 and 2009). They also say that a revolution in transportation makes it easy to transport partially processed goods quickly, and at lower cost. Thus specialization can be achieved without geographic concentration.

international fragmentation as a production pattern locating some production stages in the North and some production stages in the South, but all stages are involved in producing a single final good. The firm keeps the headquarters in the North and might shift all production stages to the South. I exclude this case, but this is explained in subsection [2-2-4].

To show that the total cost with international fragmentation is lower than that with national fragmentation, we need to know the cost of the Northern firm as a multinational firm. For this, I first suppose ex-ante that international fragmentation arises as follows: $\hat{n}$ is the cutoff stage number at which the stages are split between the North and the South; m is the number of total stages across both regions.

[2-2-1] Effort of Employee under International Fragmentation

The effort level of employees affects the multinational firm's cost, therefore, I address how effort is provided by employees under international fragmentation. The firm's total employment in the North and the South is l^f. f denotes international fragmentation. The employees' performance is monitored in each of the m stages by the same headquarters regardless of which region they reside in. The size of any group in which an individual employee is included is $\frac{l^f}{m}$. The detection probabilities for the Northern and Southern employees are associated inversely with the size of group, and are the same.

The effort function of the Northern employees under international fragmentation has the same form as that in (18) under national fragmentation, except that the detection probability is different. From (17), the detection probability in the North changes from $\sqrt{\frac{\bar{n}}{l}}$ to $\sqrt{\frac{m}{l^f}}$.

$$e(m) = \frac{w}{e^{\max}(2-\eta)^2}(\frac{m}{l^f}) . \tag{24}$$

Additionally, we need to know the disutility function of effort in order to pin down the effort level of the employees in the South. Letting an asterisk $*$ denote the South, I define it as $v^* = e^*(2-\eta^*)$, $0 < \eta^* \le 1$ and $\frac{\partial v^*}{\partial e^*} = 2 - \eta^*$. Using (24), the effort in the South is

$$e^*(m) = \frac{w^*}{e^{*\max}(2-\eta^*)^2}(\frac{m}{l^f}) . \tag{25}$$

When the number of total stages m changes, Southern effort changes by the same proportion, $\frac{1}{m}$, as Northern effort.[18]

I address the degrees of loyalty in the North and the South to compare the effort levels in both regions. More global and more advanced economies (i.e., the North) have experienced more competition than less global and less advanced economies (i.e., the South). The competitive environment in the North makes the employment contracts between the firm and its employees more fragile, and thus duration of employment is shortened. Under more protection that limits free trade, the South has been less exposed to international competition, and has kept relatively stable employment relationships. To show these employment relations in the North and the South, I compare average employment tenure of nine countries: Australia, Germany, the Netherlands, the United Kingdom and the United States as the North, and Malaysia, the Philippines, the Republic of Korea and Taiwan as the South.

Table 1 shows the distribution of workers' average tenure from one country to another in 1991. Unlike the developed countries and Taiwan, Malaysia, the Philippines and Korea did not collect official data on tenure, so that I use data from Bai and Cho who conducted their survey in 1991. They selected metropolitan cities such as Kuala Lumpur in Malaysia, Manila in the Philippines and Seoul in Korea, and surveyed average tenure in the manufacturing industry. Since the economic powers of Malaysia, the Philippines and Korea are concentrated in each capital city and its vicinity,[19] their survey data are likely to be representative of the nationwide tenure of these countries.[20]

---

[18] From (25), $\frac{\partial e^*(m)}{\partial m} = \frac{w^*}{e^{*\max}(2-\eta^*)^2 l^f}$. Using this and (25), $\frac{\frac{\partial e^*(m)}{\partial m}}{e^*(m)} = \frac{1}{m}$.

[19] Kuala Lumpur and Selanger made up 33 percent of GDP in 1988 (Lee and Sivananthiran, 1991); Metro Manila provided 30.8 percent of GDP in 1989 (Institute of Labor Studies, 1991). In the case of Korea, the value added of both Seoul and its vicinity Kyongki Province was 45.1 percent of GDP in 1990 (Report on Mining and Manufacturing Census for Korea, 1990). I cite these data from Bai and Cho (1995).

[20] Rural areas may be very different from cities. However, since the year 1965, the importance of agriculture has sharply decreased in Asian-Pacific countries. The shares of agriculture in GDP in 1990 were 9 percent in Korea, 19 percent in Malaysia, and 22 percent in the Philippines (Asian Development Bank, Key Indicator of Developing Member Countries of ADB, 1992). Also the proportion of employment in agriculture has decreased, during the period 1980-1989, in Asia-Pacific countries. At the same time, the share of the manufacturing industry in GDP has increased. The share of the manufacturing industry in 1991 in Korea was 31 percent, 27 percent in Malaysia, and 25 percent in 1990 in the Philippines (Asian Development Bank, Key Indicator of Developing Member Countries of ADB, 1992). I cite these data from Bai and Cho (1995). Thus, since the agriculture share of GDP is lower than the manufacturing share of GDP, the tenures in table 1 can be representative of the nationwide tenure.

Table 1 **Distribution of Employment by Enterprise Tenure, 1991**

	Australia	Germany *a*	Netherlands *a*	United Kingdom	United States	Malaysia *b*	Philippines *c*	Republic of Korea *d*	Taiwan *e*
Average tenure (years)									
All persons	6.8	10.4	7	7.9	6.7				7.7
Men	7.8	12.1	8.6	9.2	7.5	8.6	10.0	5.2	
Women	5.4	8.0	4.3	6.3	5.9	7.6	9.0	3.5	

(*a*) Data were collected in 1990.
(*b*) In Kuala Lumpur, Malaysia, 739 men and 706 women were surveyed from May to August, 1991.
(*c*) In Manila, the Philippines, 784 men and 786 women were surveyed from May to August 1991.
(*d*) In Seoul, Korea, 701 men and 689 women were surveyed in the first survey took place in February and March, 1991, and in the second survey took place in August 1991.
(*e*) Data for men are not available. Data for women are not available.

Source: OECD Employment Outlook (1993) for developed countries, Bai and Cho's survey report (1995) for Malaysia, the Philippines and Korea, and Report on the Manpower Utilization Survey (2005) for Taiwan

As shown in table 1, for all persons, Taiwan has an average tenure that is longer than the average tenure in Australia, the Netherlands and the United States. However, Taiwan and the United Kingdom have similar tenure. The tenures for women are longer in Malaysia and the Philippines than in Australia, the Netherlands, the United Kingdom and the United States. For men, Malaysia and the Philippines have longer tenures than Australia and the United States. Malaysia and the Netherlands have the same tenures. However, Malaysia has shorter tenure than the United Kingdom. Germany has the longest tenure for all persons and men. Korea has the shortest tenure for men and women. Although these comparisons call for some caution, tenure in the developing countries seems higher than tenure in the developed countries. Of course, Germany (as a developed country) and Korea (as a developing country) are opposite cases. I do not consider the trend of these two countries in order to focus on the main trend mentioned above.

To the extent that tenure is positively associated with the level of loyalty, and that tenure in the developed countries tends to be shorter than that in the developing countries, it is possible to assume that the loyalty in the North is lower than that in the South: $\eta < \eta^*$. This assumption says that the marginal disutility of effort in the North, $(2-\eta)$, is higher than that in the South, $(2-\eta^*)$.

Assume that the maximal effort levels in the North and the South are equal, because the upper limit on the effort that employees are able to provide physically is the same: $e^{\max} = e^{*\max}$. Then the effort of the South relative to the North, $\frac{(25)}{(24)}$, is

$$\frac{e^*(m)}{e(m)} = \frac{w^*}{w}\left(\frac{2-\eta}{2-\eta^*}\right)^2, \quad \text{where } (2-\eta) > (2-\eta^*). \tag{26}$$

This equation does not include the number of total stages m. That is, a change in the total number of stages does not influence the relative effort since the Northern effort and the Southern effort are influenced by the same total number of stages. Since $(\frac{2-\eta}{2-\eta^*}) > 1$, (26) implies that

$$\frac{w}{e(m)} > \frac{w^*}{e^*(m)}.$$

This says that the wage per unit of effort in the North is higher than the wage per unit of effort in the South. Thus the Northern firm has an incentive to shift production stages to the South.

[2-2-2] Northern Firm's Cost under International Fragmentation

The Northern firm's cost under international fragmentation is first explained. Assume that the firm locates the stages with stage number n, $n \in (0, \hat{n}]$, in the North, and the stages with stage number n^*, $n^* \in (0^*, m-\hat{n}]$, in the South.[21] The firm faces the headquarters costs and production costs occurred in the North and the South, respectively. Let me explain the headquarters cost incurred by the $\hat{n}$ Northern stages. Since the headquarters remains in the North and employs Northern labor, the unit headquarters cost, hc^f, is obtained by substituting $\hat{n}$ for $\overline{n}$ in (10').

$$hc^f = \frac{\Psi w \hat{n}}{e^{\max}}, \quad \text{where } \Psi = \frac{\theta}{N}. \tag{27}$$

I explain production in the North. Under international fragmentation, effort and labor productivity are affected by the total number of stages m. The production function for the Northern stages $n \in (0, \hat{n}]$ is derived by substituting $e(m)$ for $e(\overline{n})$ in the production function (2) in national fragmentation, l_n^f for l_n, and $b(m)$ for $b(\overline{n})$:

$$z_n = \min\{z_{n-1}, \frac{e(m) l_n^f}{b(m)}\}. \tag{28}$$

Using (3'), the unit production cost at the $\hat{n}$th stage changes to

[21] I take $\hat{n}$ and m as exogenous variables for the moment. However, in the subsequent analysis, m and $\hat{n}$ will be determined endogenously by the firm's optimal choice.

$$pc_{\hat{n}} = \frac{\hat{n}b(m)w}{e(m)}. \tag{29}$$

From (27) and (29), the unit total cost at the $\hat{n}$th stage is

$$tc_{\hat{n}}^{f} = hc^{f} + pc_{\hat{n}} = \frac{\Psi w\hat{n}}{e^{\max}} + \frac{\hat{n}b(m)w}{e(m)}. \tag{30}$$

The headquarters in the North also manages the $(m-\hat{n})$ Southern stages. It employs Northern labor and faces the headquarters cost incurred by the Southern stages. The headquarters faces more difficulty operating the Southern stages than the Northern stages for a given number of stages, because of the smaller network size and border barrier in the South. This causes the unit headquarters cost to be higher for the Southern stages than for the Northern stages.

I evaluate the headquarters cost for the Southern stages for simplicity on the basis of $\Psi\ (=\frac{\theta}{N})$, which consists of the Northern network N and the input requirement of effective Northern labor θ for providing the headquarters services for the Northern stages. This cost is denoted as hc^{f*}. However, the small Southern network and the high Southern border barriers lower the firm's efficiency. The lower efficiency should be embedded in this unit headquarters cost. I formalize this by defining hc^{f*} as a two-part functional form. The first part is linear with respect to Ψ and the number of the Southern stages, respectively: $\frac{\Psi(m-\hat{n})w}{e^{\max}}$. The second part is expressed as an increasing marginal cost of the Southern stages, $\frac{\beta(m-\hat{n})^{\mu}w}{e^{\max}}$. The parameter β, $\beta > 0$, reflects an increase in the headquarters cost due to the small Southern network. Also, the parameter μ, $\mu > 1$, reflects an increase in the headquarters cost due to the high Southern border barriers.

$$hc^{f*} = \frac{\{\Psi(m-\hat{n}) + \beta(m-\hat{n})^{\mu}\}w}{e^{\max}}, \quad \text{where } \beta > 0,\ \mu > 1. \tag{31}$$

Before addressing the production cost for the Southern stages such that $n^{*} \in (0^{*},\ m-\hat{n}]$, I look at the production function for the Southern stages. The first stage in the South is denoted by $(\hat{n}+1^{*})$. This stage uses intermediate good $Z_{\hat{n}}$ which the $\hat{n}$th stage in the North produces, and employs Southern labor $l_{\hat{n}+1^{*}}^{f}$. Since the Southern stage is one part of the production process of m stages, the productivity of labor in the

South depends on m, and is the same as in the North. The production function for the $(\hat{n}+1^*)$ th stage is derived by substituting $z_{\hat{n}+1^*}$ for z_n in (28), $z_{\hat{n}}$ for z_{n-1}, $e^*(m)$ for $e(m)$ and $l^f_{\hat{n}+1^*}$ for l^f_n,

$$z_{\hat{n}+1^*} = \min\{z_{\hat{n}}, \frac{e^*(m) l^f_{\hat{n}+1^*}}{b(m)}\}. \tag{32}$$

Suppose that output of this stage is $\bar{z}_{\hat{n}+1^*}$. Then the $(\hat{n}+1^*)$ th stage uses $\bar{z}_{\hat{n}+1^*}$ units of the intermediate $Z_{\hat{n}}$ in the South. The $\hat{n}$ th stage produces $\bar{z}_{\hat{n}+1^*}$ units of $Z_{\hat{n}}$: $z_{\hat{n}} = \bar{z}_{\hat{n}+1^*}$. Since the production of one unit of $Z_{\hat{n}}$ uses $\frac{\hat{n}b(m)}{e(m)}$ units of Northern labor, $\bar{z}_{\hat{n}+1^*}$ units of $Z_{\hat{n}}$ embodies Northern labor of $\frac{\hat{n}b(m)\bar{z}_{\hat{n}+1^*}}{e(m)}$ units. The $(\hat{n}+1^*)$ th stage also uses Southern labor $l^f_{\hat{n}+1^*} = \frac{b(m)\bar{z}_{\hat{n}+1^*}}{e^*(m)}$ for the production of $\bar{z}_{\hat{n}+1^*}$ units of $Z_{\hat{n}+1^*}$. The total labor input is the sum of embodied labor and the directly used labor: { $\frac{\hat{n}b(m)}{e(m)}$ units of Northern labor + $\frac{b(m)}{e^*(m)}$ units of Southern labor} $\bar{z}_{\hat{n}+1^*}$. The expression in the curly brackets is the total labor input per unit of $Z_{\hat{n}+1^*}$.

Consider the second stage in the South. The production function at the $(\hat{n}+2^*)$ th stage is

$$z_{\hat{n}+2^*} = \min\{z_{\hat{n}+1^*}, \frac{e^*(m) l^f_{\hat{n}+2^*}}{b(m)}\}.$$

I apply here the same logic as applied for the $(\hat{n}+1^*)$ th stage. For the production of $\bar{z}_{\hat{n}+2^*}$ units of $Z_{\hat{n}+2^*}$, the $(\hat{n}+2^*)$ th stage uses $\bar{z}_{\hat{n}+2^*}$ units of $Z_{\hat{n}+1^*}$. And $\bar{z}_{\hat{n}+2^*}$ units of $Z_{\hat{n}+1^*}$ embodies labor by { $\frac{\hat{n}b(m)}{e(m)}$ units of Northern labor + $\frac{b(m)}{e^*(m)}$ units of Southern labor} $\bar{z}_{\hat{n}+2^*}$ since the total labor input per output of $Z_{\hat{n}+1^*}$ is { $\frac{\hat{n}b(m)}{e(m)}$ units of Northern labor + $\frac{b(m)}{e^*(m)}$ units of Southern labor}. The Southern labor used for the $(\hat{n}+2^*)$ th stage is $l^f_{\hat{n}+2^*} = \frac{b(m)\bar{z}_{\hat{n}+2^*}}{e^*(m)}$. The total labor used for the $(\hat{n}+2^*)$ th stage is the sum of the embodied labor and the labor used for this stage: { $\frac{\hat{n}b(m)}{e(m)}$ units of Northern labor + $\frac{2b(m)}{e^{*}(m)}$ units of Southern labor} $\bar{z}_{\hat{n}+2^{*}}$. The expression in the curly brackets is the total labor input per unit of $Z_{\hat{n}+2^{*}}$.

Applying this logic to the $(\hat{n}+n^{*})$ th stage, the total labor used for the production of $\bar{z}_{\hat{n}+n^{*}}$ units of $Z_{\hat{n}+n^{*}}$ is { $\frac{\hat{n}b(m)}{e(m)}$ units of Northern labor + $\frac{n^{*}b(m)}{e^{*}(m)}$ units of Southern labor} $\bar{z}_{\hat{n}+n^{*}}$. The unit production cost for the $(\hat{n}+n^{*})$ th stage is obtained from the product of wage and units of labor per unit of $Z_{\hat{n}+n^{*}}$.

$$pc^{f}_{\hat{n}+n^{*}}=\frac{\hat{n}b(m)w}{e(m)}+\frac{n^{*}b(m)w^{*}}{e^{*}(m)} . \tag{33}$$

The unit total cost at the $(\hat{n}+n^{*})$ th stage is the sum of the two unit headquarters costs in (27) and (31), plus the unit production cost in (33). For simplicity, assume that $\mu=2$. Rearranging,

$$tc^{f}_{\hat{n}+n^{*}}=\frac{\Psi\hat{n}w}{e^{\max}}+\frac{\hat{n}b(m)w}{e(m)}+\frac{\{\Psi(m-\hat{n})+\beta(m-\hat{n})^{2}\}w}{e^{\max}}+\frac{n^{*}b(m)w^{*}}{e^{*}(m)} . \tag{34}$$

The sum of the first term and the second term on the RHS of (34) is the cumulative unit total cost from the first stage to the $\hat{n}$ th stage. The sum of the third term and the fourth term is the cumulative unit total cost from the $(\hat{n}+1^{*})$ th stage to the $(\hat{n}+n^{*})$ th stage.

[2-2-3] National Fragmentation versus International Fragmentation

The Northern firm changes its mode of production from national fragmentation to international fragmentation if the unit total cost under international fragmentation is lower than the unit total cost under national fragmentation.

To compare these unit total costs, I consider two cases. The first case is that the Northern firm locates all $\tilde{\tilde{n}}$ stages in the North. Recall that $\tilde{\tilde{n}}$ is the optimal number of total stages under national fragmentation. In the second case, I consider a deviation from equilibrium of national fragmentation: the firm locates the first $(\tilde{\tilde{n}}-1)$ stages in the North and moves the $\tilde{\tilde{n}}$ th stage to the South.[22]

[22] This is suboptimal. Exclusively in this section, I consider this suboptimal state to confirm whether there is an incentive to split the stages across the borders. I will address in section [2-2-4] how the firm optimally splits the stages into the North and the South.

In the first case, since the $\tilde{\tilde{n}}$th stage is in the North, the production function for this stage is obtained by (2). The production function of $Z_{\tilde{\tilde{n}}}$ is $z_{\tilde{\tilde{n}}} = \min\{z_{\tilde{\tilde{n}}-1}, \frac{e(\tilde{\tilde{n}})l_{\tilde{\tilde{n}}}}{b(\tilde{\tilde{n}})}\}$. The unit production cost for the $\tilde{\tilde{n}}$th stage is $\frac{\tilde{\tilde{n}}b(\tilde{\tilde{n}})w}{e(\tilde{\tilde{n}})}$ from (3'). The unit headquarters cost for all $\tilde{\tilde{n}}$ stages is $\frac{\Psi w\tilde{\tilde{n}}}{e^{\max}}$ from (10'), where $\Psi = \frac{\theta}{N}$. The unit total cost for the $\tilde{\tilde{n}}$th stage is the sum of these two costs:

$$\frac{\Psi w\tilde{\tilde{n}}}{e^{\max}} + \frac{\tilde{\tilde{n}}b(\tilde{\tilde{n}})w}{e(\tilde{\tilde{n}})}. \tag{35}$$

Turning to the second case, the firm keeps $(\tilde{\tilde{n}}-1)$ stages in the North and shifts one stage to the South. First, I address the size of the employment pool since this may be affected by the shift of one stage. If the shift takes place, the cost for the final stage $\tilde{\tilde{n}}$ in the North under national fragmentation and the cost for the final stage $\tilde{\tilde{n}}$ in the South under international fragmentation will diverge, while the cost for the first $(\tilde{\tilde{n}}-1)$ stages have remained the same. However, the difference between the two costs is very small compared with the total cost covering all the stages under national fragmentation. The firm would not feel the necessity to adjust the size of the pool in order to internalize the difference between the two costs for the final stage. I assume that the sizes of the pool for national fragmentation and international fragmentation would be the same, l.

The production function for the Southern stage is obtained by replacing $\hat{n}$ in (32) with $(\tilde{\tilde{n}}-1)$: $z_{(\tilde{\tilde{n}}-1)+1^*} = \min\{z_{\tilde{\tilde{n}}-1}, \frac{e^*(\tilde{\tilde{n}})l^f_{(\tilde{\tilde{n}}-1)+1^*}}{b(\tilde{\tilde{n}})}\}$. One unit of $Z_{(\tilde{\tilde{n}}-1)+1^*}$ is produced with $\frac{(\tilde{\tilde{n}}-1)b(\tilde{\tilde{n}})}{e(\tilde{\tilde{n}})}$ units of Northern labor and $\frac{b(\tilde{\tilde{n}})}{e^*(\tilde{\tilde{n}})}$ units of Southern labor. Using the first term on the RHS of (33), the unit production cost for the $(\tilde{\tilde{n}}-1)$th stage in the North is $\frac{(\tilde{\tilde{n}}-1)b(\tilde{\tilde{n}})w}{e(\tilde{\tilde{n}})}$. The unit headquarters cost for the $(\tilde{\tilde{n}}-1)$ stages in the North is $\frac{\Psi(\tilde{\tilde{n}}-1)w}{e^{\max}}$ from (27). Using the second term on the RHS of (33), the unit production cost for the first stage in the South is $\frac{b(\tilde{\tilde{n}})w^*}{e^*(\tilde{\tilde{n}})}$. Substituting $\{\tilde{\tilde{n}}-(\tilde{\tilde{n}}-1)\}$ for $(m-\hat{n})$ in (31), the unit headquarters cost for only the first stage in the South is $\frac{(\Psi+\beta)w}{e^{\max}}$. Then the unit total cost for the $\tilde{\tilde{n}}$ th stage is the sum of these four costs:

$$\{\frac{\Psi(\tilde{\tilde{n}}-1)w}{e^{\max}}+\frac{(\Psi+\beta)w}{e^{\max}}\}+\{\frac{(\tilde{\tilde{n}}-1)b(\tilde{\tilde{n}})w}{e(\tilde{\tilde{n}})}+\frac{b(\tilde{\tilde{n}})w^{*}}{e^{*}(\tilde{\tilde{n}})}\}. \qquad (36)$$

In order for the Northern firm to shift the $\tilde{\tilde{n}}$ th stage to the South, the cost in (36) should be lower than the cost in (35). In both cases, the unit total costs up to the $(\tilde{\tilde{n}}-1)$ th stage are the same. Thus, the difference between the unit total cost for all the stages $\tilde{\tilde{n}}$ in the first case and that for all the stages $\tilde{\tilde{n}}$ in the second case comes from the difference between the unit total cost for the $\tilde{\tilde{n}}$ th stage in the North and the unit total cost for the $\tilde{\tilde{n}}$ th stage in the South. For simplicity, assuming that $\delta=1$, the condition for international fragmentation, $(36)-(35)<0$, is identical to the following condition:[23]

$$\beta<\Psi\{1-(\frac{2-\eta^{*}}{2-\eta})^{2}\}<1. \qquad (37)$$

I assume that the model satisfies this condition.

[2-2-4] Optimal Determination of International Fragmentation

The Northern firm determines the optimal number of total stages and split these optimally between the North and the South as a way of incurring the lowest total cost. For the moment, assume that there is a number of total stages across both regions, m, minimizing the unit total cost. For a given m, the firm locates $\hat{n}$ stages in the North and the other $(m-\hat{n})$ stages in the South. The unit total cost function for the final stage m is obtained by replacing n^{*} in (34) with $(m-\hat{n})$.

$$tc^{f}_{\hat{n}+(m-\hat{n})}=\{\frac{\Psi\hat{n}w}{e^{\max}}+\frac{\hat{n}b(m)w}{e(m)}\}+[\frac{\{\Psi(m-\hat{n})+\beta(m-\hat{n})^{2}\}w}{e^{\max}}+\frac{(m-\hat{n})b(m)w^{*}}{e^{*}(m)}],$$

where $\frac{b(m)}{e(m)}=\frac{dm^{-2}}{w}$, $\frac{b(m)}{e^{*}(m)}=\frac{d^{*}m^{-2}}{w^{*}}$,

$$d\equiv e^{\max}(2-\eta)^{2}l^{f},\ d^{*}\equiv e^{\max}(2-\eta^{*})^{2}l^{f},\ d>d^{*}\text{ since }\eta<\eta^{*}. \qquad (38)$$

This is simplified as

[23] See Appendix 2.

$$tc_m^f = \frac{w}{e^{\max}}\{\Psi m + \beta(m-\hat{n})^2\} + \{d\hat{n} + d^*(m-\hat{n})\}m^{-2}\,. \tag{39}$$

For the minimization of this cost with respect to $\hat{n}$ for a given m, the first derivative with respect to $\hat{n}$ is zero and the second derivative is positive.

$$\frac{\partial tc_m^f}{\partial \hat{n}} = \frac{-2\beta w(m-\hat{n})}{e^{\max}} + (d-d^*)m^{-2} = 0\,,\ \frac{\partial^2 tc_m^f}{\partial \hat{n}^2} = \frac{2\beta w}{e^{\max}} > 0\,. \tag{40}$$

The number of Northern stages, $\hat{n}$, is obtained from (40).

$$\hat{n}(m) = m - (d-d^*)(\frac{e^{\max}}{2\beta w})m^{-2}\,. \tag{41}$$

Since $\hat{n}$ is a function of m, (39) is re-expressed as

$$tc_m^f = \frac{w}{e^{\max}}[\Psi m + \beta\{m-\hat{n}(m)\}^2] + [d\hat{n}(m) + d^*\{m-\hat{n}(m)\}]m^{-2}\,. \tag{39'}$$

The number of total stages is obtained by minimizing tc_m^f (39') with respect to m. For the cost minimization, the first derivative is zero and the second derivative is positive. For a given w, the first derivative is

$$\frac{\partial tc_m^f}{\partial m} = \frac{w\Psi}{e^{\max}} - dm^{-2} + \{\frac{e^{\max}(d-d^*)^2}{\beta w}\}m^{-5} = 0\,. \tag{42}$$ [24]

The second derivative is positive:

$$\frac{\partial^2 tc_m^f}{\partial m^2} = 2dm^{-3} - \{\frac{5e^{\max}(d-d^*)^2}{\beta w}\}m^{-6} > 0\,,\quad \text{where } m > (\frac{5e^{\max}(d-d^*)^2}{2d\beta w})^{\frac{1}{3}}\,. \tag{43}$$ [25]

The optimal number of total stages is determined by (42). However, for graphical analysis, I rewrite (42) as a new expression with A and B that are defined as follows.

$$A - B = 0\,,\quad \text{where } A \equiv \frac{w\Psi}{e^{\max}} + \{\frac{e^{\max}(d-d^*)^2}{\beta w}\}m^{-5}\,,\ B \equiv dm^{-2}\,. \tag{44}$$

The optimal number of total stages $\tilde{m}$ is determined where $A = B$. The shapes of A and B are monotonic in m. The slope of A is flatter than that of B.[26] As m becomes small, the value of A becomes smaller than

---

[24] See Appendix 3.

[25] For the second derivative in (43) to be positive, $\{2dm^3 - \frac{5e^{\max}(d-d^*)^2}{\beta w}\}m^{-6} > 0$. Since the final stage m is a positive number, $m^{-6} > 0$. $\{2dm^3 - \frac{5e^{\max}(d-d^*)^2}{\beta w}\}$ has to be positive. Thus, $m > \{\frac{5e^{\max}(d-d^*)^2}{2d\beta w}\}^{\frac{1}{3}}$.

that of B.[27] As m becomes large, the value of A becomes larger than that of B.[28] Thus A and B are drawn as in Figure 7. They intersect at c where $\tilde{m}$ is determined.

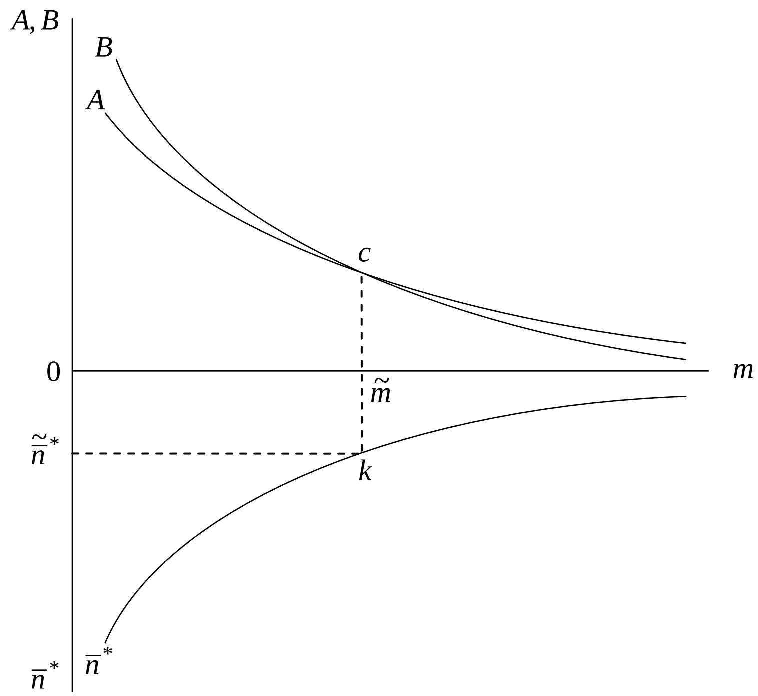

Figure 7

The number of the stages located in the South is obtained by using (41).

$$\bar{n}^*(m) = m - \hat{n}(m) = (d - d^*)(\frac{e^{\max}}{2\beta w})m^{-2}, \tag{45}$$

$$\text{where } \frac{\partial \bar{n}^*(m)}{\partial m} = \{-\frac{e^{\max}(d-d^*)}{\beta w}\}m^{-3} < 0, \ \frac{\partial^2 \bar{n}^*(m)}{\partial m^2} = \{\frac{3e^{\max}(d-d^*)}{\beta w}\}m^{-4} > 0. \tag{46}$$

The curve of $\bar{n}^*(m)$ is drawn in the space $(m, \bar{n}^*)$ below the horizontal axis of Figure 7. The vertical axis below the origin represents a positive number of Southern stages $\bar{n}^*$. This number becomes larger moving downward from the origin along this axis. The signs in (46) indicate that the curve of $\bar{n}^*(m)$ is convex toward

---

[26] $\frac{\partial A}{\partial m} = \{\frac{-5e^{\max}(d-d^*)^2}{\beta w}\}m^{-6}$. $\frac{\partial B}{\partial m} = -2dm^{-3}$. From $\frac{\partial^2 tc_m^f}{\partial \hat{n}^2} > 0$ in (43), $\left|\frac{\partial A}{\partial m}\right| < \left|\frac{\partial B}{\partial m}\right|$.

[27] See Appendix 4.

[28] $\lim_{m\to\infty} A = \frac{w\Psi}{e^{\max}}$ and $\lim_{m\to\infty} B = 0$. Then $\lim_{m\to\infty} A > \lim_{m\to\infty} B$.

the origin in the space $(m, \bar{n}^*)$. This can be understood as follows. As the total number of stages m increases, the effect of the increasing returns to the labor division increases, and thus the production cost falls; the firm has a smaller incentive for shifting the stages to the South to reduce the production cost; and this means that the number of Southern stages $\bar{n}^*$ decreases with the total number of stages m. The optimal number of Southern stage, $\tilde{\tilde{n}}^*$, is determined at k where the vertical line extended downward from the point c touches the curve $\bar{n}^*(m)$.

I have to mention the possibility that the headquarters remains in the North and all production stages shift to the South. The possibility of this case is very low in the model, but this case can also be ruled out if η^* satisfies $\eta^* > \frac{\eta}{\sqrt{2}} + (2 - \sqrt{2})$.[29]

[2-2-5] Effect of International Fragmentation on the Total Number of Stages

I address what happens to the total number of stages when the firm changes its mode of production from national fragmentation to international fragmentation. I compare the total number of stages in national fragmentation $\tilde{\tilde{n}}$ to that in international fragmentation $\tilde{m}$.

The curves A and B in Figure 7 are used for this purpose. First, consider a point on the horizontal axis which is located at $\tilde{m}$. Where m is equal to $\tilde{m}$, the values of A and B are represented by c, and are the same.

$$\frac{A(m)}{B(m)} = 1, \quad \text{where } m = \tilde{m}. \tag{47}$$

Second, consider a point on the left of $\tilde{m}$ on the horizontal axis. Where $0 < m < \tilde{m}$, the value on the curve A corresponding to this point is smaller than the value on the curve B in Figure 7. Using (47),

$$\frac{A(m)}{B(m)} < \frac{A(\tilde{m})}{B(\tilde{m})} = 1, \quad \text{where } 0 < m < \tilde{m}. \tag{48}$$

Third, consider a point on the right of $\tilde{m}$, that is, $m > \tilde{m}$. The value on A corresponding to this point is larger than the value on B in Figure 7. By (47),

$$\frac{A(m)}{B(m)} > \frac{A(\tilde{m})}{B(\tilde{m})} = 1, \quad \text{where } m > \tilde{m}. \tag{49}$$

---

[29] See Appendix 5.

Using $A(m)$ and $B(m)$ in (44), $\frac{A(m)}{B(m)}$ is obtained as follows.

$$\frac{A(m)}{B(m)} = \frac{1}{d}\{\frac{w\Psi m^2}{e^{\max}} + \frac{e^{\max}(d-d^*)^2 m^{-3}}{\beta w}\}.^{30} \tag{50}$$

This ratio is used to compare $\tilde{m}$ and $\tilde{\tilde{n}}$. If m has the same number as the equilibrium number of total stages in national fragmentation $\tilde{\tilde{n}}$, this ratio is expressed as $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})}$. If $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})}$ is less than or equal to one, this implies that $\tilde{\tilde{n}} \le \tilde{m}$ from (47) and (48):

$$\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})} \le \frac{A(\tilde{m})}{B(\tilde{m})} = 1 \quad \Rightarrow \quad \tilde{\tilde{n}} \le \tilde{m}. \tag{51}$$

If $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})}$ is larger than one, this implies that $\tilde{\tilde{n}} > \tilde{m}$ from (49):

$$\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})} > \frac{A(\tilde{m})}{B(\tilde{m})} = 1 \quad \Rightarrow \quad \tilde{\tilde{n}} > \tilde{m}. \tag{52}$$

To find the value of $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})}$, substitute $\tilde{\tilde{n}}$ for m in (50). Using $\tilde{\tilde{n}}$ in (23), $\delta = 1$ and $\varepsilon = 1$,

$$\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})} = \frac{l}{l^f}[1 + \frac{\{(2-\eta)^2 - (2-\eta^*)^2\}^2 \theta^{\frac{3}{2}} w^{\frac{1}{2}}}{e^{\max}\beta(2-\eta)^5 l^{\frac{1}{2}} (\frac{l}{l^f})^2 N^{\frac{3}{2}}}],^{31} \quad \text{where } 0 < \eta < \eta^* < 1. \tag{53}$$

The second term in the square brackets on the RHS in (53) is positive. The value in the square brackets is larger than one. Thus, whether the value of $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})}$ is larger than one or not depends on the size of the employment pool in national fragmentation relative to that in international fragmentation, $\frac{l}{l^f}$. The following cases are considered.

(i) case: $l^f \le l$

If the size of the employment pool with international fragmentation is less than or equal to that with national fragmentation, the ratio of employment is $\frac{l}{l^f} \ge 1$. The RHS in (53) is the product of the ratio $\frac{l}{l^f}$ and

[30] See [A6-1] in Appendix 6.
[31] See [A6-5] in Appendix 6.

the value in the square brackets. The product is larger than one, so that the LHS is larger than one: $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})} > 1$.

This occurs, as explained in (52), where $\tilde{\tilde{n}} > \tilde{m}$.

This can be explained as follows. The internationally fragmented firm faces a low Southern wage as well as a high Northern wage, but the nationally fragmented firm faces only a high Northern wage. The number of workers hired by the internationally fragmented firm is smaller than or equal to that hired by the nationally fragmented firm. Thus the former firm faces a lower production cost than the latter firm.[32] The internationally fragmented firm feels less pressure to reduce the production cost compared with the nationally fragmented firm. The former firm has a smaller incentive to increase the total number of stages so as to raise their productivity than the latter firm. This says that the total number of stages should be smaller in international fragmentation than in national fragmentation.

(ii) case: $l^f > l$

If the size of the employment pool is larger with international fragmentation than with national fragmentation, the ratio of employment is $\frac{l}{l^f} < 1$. The RHS in (53) is the product of the ratio $\frac{l}{l^f}$ and the value in the square brackets, where the former is smaller than one and the latter is larger than one. According to which one dominates, the value of the product is greater than, equal to or less than one. Thus, the value of the LHS in (53) is $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})} \gtreqless 1$. This implies that $\tilde{\tilde{n}} \gtreqless \tilde{m}$ from (51) and (52).

Intuitively, the internationally fragmented firm in comparison to the nationally fragmented firm faces a lower wage, but hires a larger number of workers. This could cause the production cost with international fragmentation to be larger than, equal to or smaller than the production cost with national fragmentation. Thus, it is difficult to say which firm has a greater incentive to reduce its production cost. Therefore, the total number of stages with international fragmentation could be larger than, equal to or smaller than that with national fragmentation.

---

[32] The total cost – the sum of the headquarters cost and the production cost – is affected by the network size, loyalty, border barrier, size of employment pool and wage. The degree of segmentation of the production process is determined by the minimization of the total cost. These factors are reflected on the RHS in (53). Since the value in the square brackets on the RHS is larger than one, whether or not $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})}$ is larger than one is determined by the ratio of the numbers of workers hired by the two firms. That is, the second term in the square brackets, which includes the factors mentioned, does not play a key role in determining whether the value of the ratio in (53) is larger than one. Therefore, I focus on the wage cost of production.

[2-3] Comparative Statics

I focus on the internationally fragmented firm. I examine how the networks, loyalty and the wages affect the total number of stages, the number of outsourced stages and the effort level.

[2-3-1] Networks

Suppose the network size of the South becomes larger and the network size of the North does not change. The firm connected to the Southern network can communicate information faster and easily with the Southern stages. This improves efficiency of the headquarters service. Since β reflects the increment in the headquarters cost due to inefficiency of the small Southern network, an increase in the Southern network size brings on a decrease in β. To examine how a change in β affects the total number of stages and the number of Southern stages, I use (44) and (45).

From (44),

$$\frac{\partial A}{\partial \beta} = \frac{-e^{\max}(d-d^*)^2 m^{-5}}{w\beta^2} < 0\,,\ \frac{\partial B}{\partial \beta} = 0\,.$$

A decrease in β shifts the curve A upward for a given m in Figure 8. The extent of the shift of A is decreasing in m, so that the new curve is steeper. The curve B does not change. The point c moves to c' along the curve B. The total number of stages m falls.

Using (45),

$$\frac{\partial \bar{n}^*}{\partial \beta} = \frac{-e^{\max}(d-d^*)m^{-2}}{2w\beta^2} < 0\,.$$

An increase in the Southern network size (i.e., a decrease in β) leads to an increase in the number of Southern stages. This can be illustrated graphically by a shift of the curve $\bar{n}^*$ in Figure 8. Suppose that the curve $\bar{n}^*$ shifts downward to the curve $\bar{n}^{*'}$. The number of Southern stages increases for a given m. Therefore, the fact that a decrease in β increases the number of Southern stages is represented by the downward shift of the curve $\bar{n}^*$. The extent of the shift of $\bar{n}^*$ is decreasing in m, so that the new curve becomes steeper. The point k, prior to the shift, becomes k' after the shift. The number of Southern stages $\bar{n}^*$ increases.

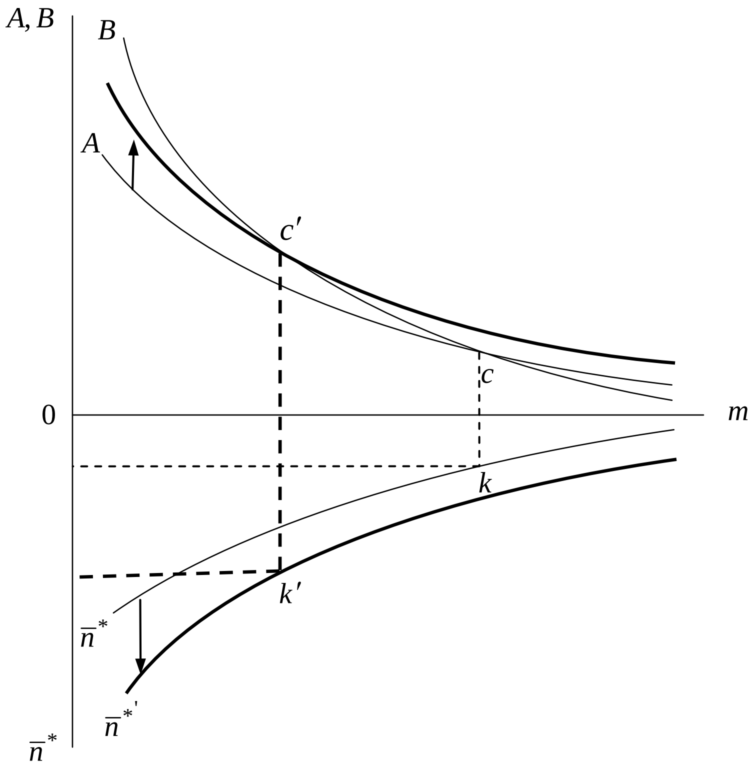

Figure 8

Intuitively, when the Southern network size increases, the headquarters efficiently manage the Southern stages. The headquarters cost decreases. The firm increases the number of Southern stages. On the other hand, the firm decreases the total number of stages. The reason is as follows. The lowered headquarters cost and the increase in the number of Southern stages using low wage labor decrease the firm's total cost; this reduces the necessity for dividing production into more stages for improving the efficiency.

The headquarters supervises all employees in the North and the South with the same monitoring ability, which is measured by the total number of stages in the North and the South. All employees in both countries face the same detection probability. This, along with the decrease in the total number of stages, decreases effort in each country in (24) and (25) by the same proportion. Thus, the relative effort in (26) is not affected.

[2-3-2] Loyalty of Employees

The new competitive environment through globalization makes the employment relationship more unstable in the South compared to that in the North, since the South faces increased international competition. This decreases loyalty in the South compared to the North, thus η^* falls for a given η .

To investigate how a decrease in η^* influences the total number of stages, (44) is used.

$$\frac{\partial A}{\partial \eta^*} = \{-\frac{2e^{\max}(d-d^*)m^{-5}}{\beta w}\}(\frac{\partial d^*}{\partial \eta^*}) > 0\,,\ \frac{\partial B}{\partial \eta^*} = m^{-2}(\frac{\partial d}{\partial \eta^*}) = 0\,,$$

$$\text{where } \frac{\partial d^*}{\partial \eta^*} = -2e^{\max}(2-\eta^*)l^f < 0\,,\ \frac{\partial d}{\partial \eta^*} = 0 \text{ from (38).} \tag{54}$$

This says that only the curve A shifts downward as η^* falls for a given m as in Figure 9. Since the extent of the shift of A is decreasing in m, the new curve becomes less steep. The curve B does not change. The point c moves to c' along the curve B. The total number of stages m increases.

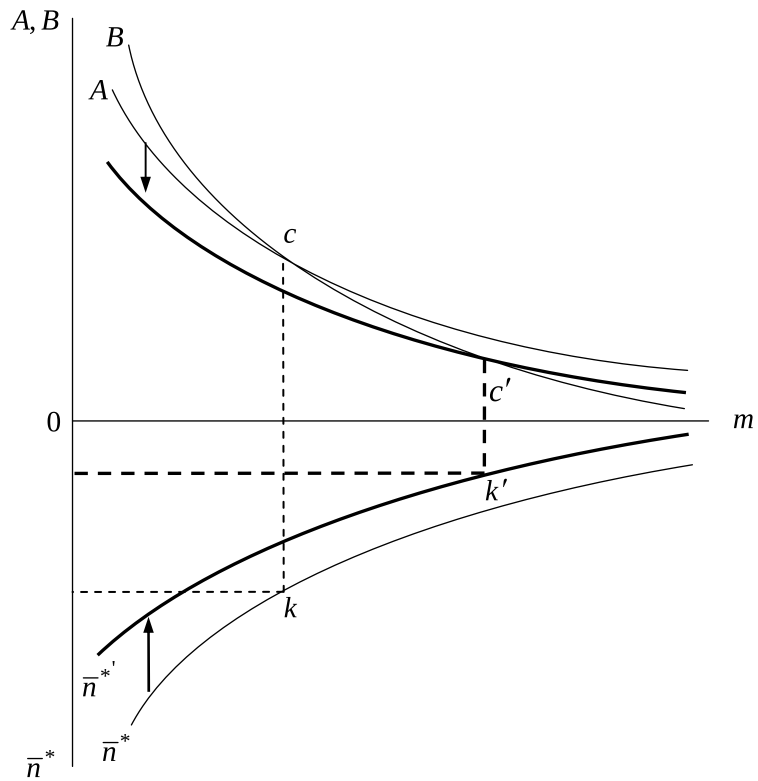

Figure 9

The effect of a decrease in η^* on the number of Southern stages is obtained from (45).

$$\frac{\partial \bar{n}^*}{\partial \eta^*} = (-\frac{e^{\max}m^{-2}}{2\beta w})(\frac{\partial d^*}{\partial \eta^*}) > 0\,,\quad \text{where } \frac{\partial d^*}{\partial \eta^*} < 0 \text{ in (54).}$$

A fall in Southern loyalty leads to a decrease in the number of Southern stages. This can be illustrated by a upward shift of the curve $\bar{n}^*$ since the number of Southern stages decreases for a given m. The extent of the shift of $\bar{n}^*$ is decreasing in m, so that the new curve becomes less steep. The point k shifts to k' after the curve $\bar{n}^*$ shifts upward to the curve $\bar{n}^{*'}$. The number of Southern stages $\bar{n}^*$ decreases.

If Southern loyalty falls, its effect on Northern effort is from (24):

$$\frac{\partial e(m)}{\partial \eta^*} = \frac{w}{e^{\max}(2-\eta)^2 l^f} \frac{\partial m}{\partial \eta^*} < 0, \quad \text{where } \frac{\partial m}{\partial \eta^*} < 0. \tag{55}$$

The negative signs in (55) can be explained as follows. The fallen Southern loyalty increases the total number of stages (i.e., $\frac{\partial m}{\partial \eta^*} < 0$); this increases the firm's detection probability and thus Northern effort rises.

The effect on Southern effort of lower Southern loyalty is from (25):

$$\frac{\partial e^*(m)}{\partial \eta^*} = \{\frac{2w^* m}{e^{\max}(2-\eta^*)^3 l^f} + \frac{w^*}{e^{\max}(2-\eta^*)^2 l^f} \frac{\partial m}{\partial \eta^*}\} \gtrless 0, \quad \text{where } e^{*\max} = e^{\max}.$$

The first term in the curly brackets is positive. This says that lowered Southern loyalty leads to a fall in Southern effort, given that m does not change. The second term in the curly brackets is negative; the increased total number of stages increases the detection probability and thus Southern effort. These opposite movements make the effect on Southern effort ambiguous.

I now examine the relationship between Southern loyalty and relative effort. From (26),

$$\frac{\partial\{\frac{e^*(m)}{e(m)}\}}{\partial \eta^*} = [\frac{\partial\{\frac{e^*(m)}{e(m)}\}}{\partial(\frac{2-\eta}{2-\eta^*})}]\{\frac{\partial(\frac{2-\eta}{2-\eta^*})}{\partial \eta^*}\} > 0, \tag{56}$$

$$\text{where } [\frac{\partial\{\frac{e^*(m)}{e(m)}\}}{\partial(\frac{2-\eta}{2-\eta^*})}] = 2(\frac{w^*}{w})(\frac{2-\eta}{2-\eta^*}) > 0, \ [\frac{\partial(\frac{2-\eta}{2-\eta^*})}{\partial \eta^*}] = \frac{(2-\eta)}{(2-\eta^*)^2} > 0.$$

The positive sign in (56) implies that a fall in η^* causes a fall in $\frac{2-\eta}{2-\eta^*}$ and thus a fall in $\frac{e^*(m)}{e(m)}$. Thus, lowered Southern loyalty reduces Southern effort relative to Northern effort.

These results are explained intuitively. If loyalty in the South falls relative to the North, the South's relative disutility of effort rises. The employees in the South provide less effort relative to the employees in the

North. This increases the production cost of outsourced stages. The firm decreases outsourcing to the South and brings back some of the Southern stages to the North. The increase in Northern stages using high wage labor, and the lowered Southern effort, increase the total cost. To improve the efficiency of production, the firm increases the total number of stages.

[2-3-3] Wages

Suppose that the Southern wage rises for a given Northern wage.[33] The change in the relative wage does not affect the position of the curves A and B, since A and B in (44) have no Southern wage term: $\frac{\partial A}{\partial w^*} = \frac{\partial B}{\partial w^*} = 0$. This says that the total number of stages m is not changed by the Southern wage. The effect on the number of Southern stages $\overline{n}^*$ is from (45): $\frac{\partial \overline{n}^*}{\partial w^*} = 0$. The number of Southern stages is not affected by the Southern wage.

If the Southern wage rises, Southern effort increases and Northern effort does not change.

$$\frac{\partial e^*(m)}{\partial w^*} = \{\frac{1}{e^{\max}(2-\eta^*)^2}\frac{m}{l^f}\} > 0\,,\ \frac{\partial e(m)}{\partial w^*} = 0\,.$$[34]

Also, a rise in the wage of the South relative to the North increases the effort of the South relative to the North. From (26),

$$\frac{\partial\{\frac{e^*(m)}{e(m)}\}}{\partial(\frac{w^*}{w})} = (\frac{2-\eta}{2-\eta^*})^2 > 0\,. \tag{57}$$

These results can be understood as follows. The rise in the Southern wage increases Southern employees' effort. This relatively decreases the production cost in the South. At the same time, the increased wage raises the production cost in the South. These opposite effects are traded off, and the degree of outsourcing is not changed. Also, the total number of stages is not changed.

---

[33] I treat the wages as exogenously given. However, it has also become an important issue how wages are affected by international fragmentation. In the single sector H-O model with skilled and unskilled labor of Deardorff (2005), the services outsourcing of skilled-labor from the developed country (the North) causes the wage of unskilled labor in the North to fall below that in the developing country (the South), if the North continues to diversify. However, the high-skilled and low-skilled labor in the North gains if specialization occurs due to largely different factor endowments. Also, Grossman and Rossi-Hansberg (2009) examine how falling costs of offshoring affect high-skill and low-skill wages in the source country in a H-O model with two sectors. They show that offshoring benefits the factors whose tasks are moved more easily overseas and generates shared gains for all domestic factors.

[34] See Appendix 7.

I consider another case of a rise in the Northern wage for a given Southern wage. The direction of shift of the curve A depends on the values of parameters, while the curve B does not change:

$$\frac{\partial A}{\partial w} = \frac{\Psi}{e^{\max}} - \frac{e^{\max}(d-d^*)^2 m^{-5}}{\beta w^2} \gtrless 0\,, \quad \frac{\partial B}{\partial w} = 0\,.$$

The direction of movement of m is ambiguous, so that the change in the total number of stages is ambiguous. The effect on the number of Southern stages $\bar{n}^*$ is

$$\frac{\partial \bar{n}^*}{\partial w} = \frac{-e^{\max}(d-d^*)}{2\beta m^2 w}\left(\frac{1}{w} + \frac{2}{m}\frac{\partial m}{\partial w}\right) \gtrless 0\,, \quad \text{where } \frac{\partial m}{\partial w} \lesseqgtr 0.$$

The number of Southern stages can decrease, increase or not change, since $\bar{n}^*$ depends on the change in m with respect to w.

When the Northern wage rises, the changes in effort in both regions are ambiguous due to the ambiguity of the change in the total number of stages. This is shown by using (24) and (25).

To examine how the rise in the Northern wage affects Northern effort relative to Southern effort, I use (57). This indicates that a decrease in the ratio of the Southern wage to the Northern wage $\frac{w^*}{w}$ leads to a decrease in the ratio of Southern effort to Northern effort $\frac{e^*(m)}{e(m)}$. Since a decrease in $\frac{w^*}{w}$ means an increase in $\frac{w}{w^*}$, and since a decrease in $\frac{e^*(m)}{e(m)}$ means an increase in $\frac{e(m)}{e^*(m)}$, an increase in $\frac{w}{w^*}$ leads to an increase in $\frac{e(m)}{e^*(m)}$. In other words, a rise in the Northern wage for a given Southern wage causes a rise in Northern effort relative to Southern effort.

Now I explain intuitively why the changes in the total number of stages and outsourcing with respect to a change in the Northern wage are ambiguous. A rise in the Northern wage increases Northern employees' effort relative to Southern employees' effort. This relatively decreases the production cost in the North. At the same time, the increased wage raises the headquarters cost and the production cost in the North. These work in opposite directions, so that the total cost can increase, decrease or not change. Thus, the effect of the increased Northern wage on the total number of stages is ambiguous. Since this ambiguity influences production costs of both Northern stages and Southern stages, the comparison of production costs between the North and the South is also difficult. This causes the effect on outsourcing to be ambiguous.

[3] Conclusion

This paper addresses how physical networks, loyalty of workers and wages affect international fragmentation and the work effort, in a world consisting of a developed country and a developing country. Splitting a single production process into multiple-stages increases a headquarters' monitoring ability to find where shirking occurs. This increases the employees' work effort. The firm has an incentive to fragment production. However, since the headquarters cost for coordinating and monitoring the stages increases with the number of production stages, this limits the number of total stages. The international differences in the network sizes, border barriers, loyalty and wages cause the stages to be separated into the North and the South. The Northern firm determines the optimal number of total stages as a way of incurring the lowest total cost.

When the mode of production changes from national fragmentation to international fragmentation, if international fragmentation brings on a decrease in employment, the internationally fragmented firm compared with the nationally fragmented firm faces low wage and decreased employment. This weakens the internationally fragmented firm's incentive to improve efficiency by increasing the total number of stages, thus leading to a decrease in the total number of stages. However, if international fragmentation brings on an increase in employment, this raises the incentive to improve efficiency by increasing the total number of stages. On the contrary, the low wage weakens this incentive. These opposite effects make the effect of international fragmentation on the total number of stages ambiguous.

An increase in the Southern network decreases the headquarters cost for the Southern stages compared to the headquarters cost for the Northern stages. This increases outsourcing from the North to the South. Since the increased Southern network reduces the firm's cost, the necessity for the division of production is diminished. The firm decreases the total number of stages. Since all employees in both regions face a lowered and identical detection probability, this decreases the level of effort in the North and the South, and keeps the effort of the South relative to the North unchanged.

The new competitive environment through globalization makes employment relationships more unstable in the South compared to that in the North, since the South is under more protection that limits free trade. As a result, Southern employees' loyalty falls and their effort relative to the Northern employees' effort decreases. This increases the production cost in the South and reduces outsourcing. However, the firm increases the total number of stages to improve productivity.

If the Southern wage rises, Southern effort increases, but Northern effort does not change. The increasing effort in the South decreases the production cost in the South. However, a rise in the Southern wage could also increase the production cost in the South. These opposite effects on the production cost are traded off, so outsourcing does not change. Also, the total number of stages is not changed.

If the Northern wage rises, the headquarters cost and the production cost in the North increase. At the same time, the risen Northern wage increases Northern effort relative to Southern effort. This decreases the production cost in the North. The effect of the risen wage and the effect of the risen effort work in opposite directions, so outsourcing is determined by which of these effects dominates. The effect of the risen Northern wage on outsourcing is ambiguous. The effect of the risen Northern wage on the total stages is also ambiguous.

[References]

Asian Development Bank, 1992. Key Indicators of Developing Member Countries of ADB.

Bai, Moo Ki and Cho, Woo Hyun, 1995. Employment Structures in Manufacturing of Four Surveyed Countries and terms of Female Employment in Women's Wages and Employment in Korea. Seoul: Seoul National University Press.

Becker, G. and Murphy, K., 1992. The Division of Labor, Co-ordination costs, and Knowledge. Quarterly Journal of Economics, 107.

Brecher, Richard A., 1992. An Efficiency-Wage Model with Explicit Monitoring: Unemployment and Welfare in an Open Economy. Journal of International Economics, 32.

Calvo, Guillermo A. and Wellisz Stanislaw, 1978. Supervision, Loss of Control, and the Optimum Size of the Firm. The Journal of Political Economy, 86.

Copeland, Brian R., 1989. Efficiency Wages in a Ricardian Model of International Trade. Journal of International Economics, 27.

Deardorff, Alan V., 2001 a. Trade and Welfare Implication of Networks. Paper prepared for the Murray S. Johnson Conference on International Economics, University of Texas.

Deardorff, Alan V., 2001 b. Fragmentation in Simple Trade Models. North American Journal of Economics and Finance, 12.

Deardorff, Alan V., 2005. A Trade Theorist's Take on Skilled-Labor Outsourcing. International Review of Economics and Finance, 14.

Directorate-General of Budget, 2005. Accounting and Statistics, Executive Yuan, Report on the Manpower Utilization Survey in Taiwan Area. Republic of China.

Edwards, Brian K. and Starr, Ross M., 1987. A Note on Indivisibilities, Specialization, and Economies of Scale. The American Economic Review, 77.

Either, Wilfred J., 1979. Internationally Decreasing Costs and World Trade. Journal of International Economics, 9.

Freund, Caroline and Weinhold, Diana, 2002. The Internet and International Trade in Services. The American Economic Review, v.92 no.2.

Grossman, Gene M. and Rossi-Hansberg Esteban, 2007. The Rise of Offshoring: It's Not Wine for Cloth Anymore. Federal Reserve Bank of Kansas City on "The New Economic Geography: Effects and Policy Implications," Jackson Hole Symposium.

Grossman, Gene M. and Rossi-Hansberg Esteban, 2009. Trading Tasks: A Simple Theory of Offshoring. The American Economic Review, forthcoming.

Harris, Richard G., 2001. A Communication-Based Model of Global Production Fragmentation, in Arndt, Sven W. and Henryk Kierzkowski (eds.) Fragmentation: New Production Patterns in the World Economy. Oxford: Oxford University Press.

Leamer, Edward E., 1999. Effort, Wages, and the International Division of Labor. The Journal of Political Economy, 107.

Lee, Kiong Hock and Sivananthiran, A., 1991. Report on Employment, Occupational Mobility and Earnings in the Kuala Lumpur Urban Labor Market with Special Reference to Women in the Manufacturing Sector: Draft.

Matusz, Steven J., 1996. International Trade, the Division of Labor, and Unemployment. International Economic Review, 37.

Minkler, Lanse, 2004. Shirking and Motivations in Firms: Survey Evidence on Worker Attitudes. International Journal of Industrial Organization, 22.

National Bureau of Statistics, Economic Planning Board, Report on Mining and Manufacturing Census 1990, Republic of Korea.

OECD, Employment Outlook 1993. Paris.

Reichheld, Frederick F. and Teal, Thomas, 1996. The Loyalty Effect: The Hidden Force behind Growth, Profits and Lasting Value. Bain & Company, Inc. Harvard Business School Press.

Sennett, Richard, 1998. Corrosion of Character: The Personal Consequences of Work in the New Capitalism. New York: W.W. Norton and Company, Inc.

Shapiro, Carl and Stiglitz, Joseph E., 1984. Equilibrium Unemployment as a Worker Discipline Device. The American Economic Review, 74.

Smith, Adam, 1937. The Wealth of Nations. New York: Modern Library; Orig. pub. 1776.

[Appendices]

[Appendix 1]

The maximal expected utility is obtained by solving the constrained maximization problem, which maximizes EU subject to a single constraint such that $\varphi(e) = e \le e^{\max}$. To solve this problem, it should first be checked whether the constraint qualification is satisfied. The condition of constraint qualification is as follows. If the expected utility is maximized at an effort level $\bar{e}$ such that $\bar{e} \le e^{\max}$, and if the constraint is binding at $\bar{e}$, then $\frac{\partial \varphi(e)}{\partial e}$ should not be zero at $\bar{e}$ on the boundary of the constraint: $\frac{\partial \varphi(\bar{e})}{\partial e} \neq 0$. When I turn to the maximization problem at hand, the constraint has the feature that $\frac{\partial \varphi(e)}{\partial e} = 1$ for all e. This means that $\frac{\partial \varphi(e)}{\partial e} \neq 0$ at a value of e on the boundary, so the constraint, $\varphi(e) = e \le e^{\max}$, satisfies the constraint qualification. Then we are able to form the Lagrangian function and to calculate the optimal effort.

$$L(e,\ \lambda) = EU - \lambda(e - e^{\max}) .$$

The first order conditions are

$$\frac{\partial L}{\partial e} = \frac{\partial EU}{\partial e} - \lambda = 0 , \qquad \text{[A1-1]}$$

$$\lambda(e - e^{\max}) = 0 , \qquad \text{[A1-2]}$$

$$\lambda \ge 0 , \qquad \text{[A1-3]}$$

$$e \le e^{\max} . \qquad \text{[A1-4]}$$

If $\lambda > 0$ in [A1-3], $\frac{\partial EU}{\partial e} > 0$ from [A1-1] and $(e - e^{\max}) = 0$ from [A1-2]. This means that $e = e^{\max}$.

However, if $\lambda = 0$ in [A1-3], $\frac{\partial EU}{\partial e} = 0$ from [A1-1]. Then $\frac{\partial EU}{\partial e}$ in (15) should be zero. The effort level is determined at $e = \frac{wP^2}{e^{\max}(2-\eta)^2}$.

The two effort levels, $e^{\max}$ and $e = \frac{wP^2}{e^{\max}(2-\eta)^2}$, are candidates that can be considered in the range of $e \le e^{\max}$ in [A1-4]. The first candidate represents case (iii) in Figure 2. The second candidate represents the case (ii) and $e = \frac{wP^2}{e^{\max}(2-\eta)^2} < e^{\max}$ by [A1-4].

Compare the levels of expected utility at these candidates. First, substituting $e = \frac{wP^2}{e^{\max}(2-\eta)^2}$ into the expected utility function in (14),

$$EU = \{\frac{wP^2}{e^{\max}(2-\eta)} + 2(1-P)\sqrt{w}\} > 0, \quad \text{where } 1 \le (2-\eta) < 2 ,\ 0 < (1-P) < 1 . \qquad \text{[A1-5]}$$

Second, substituting $e = e^{\max}$ into (14),

$$EU = 2\sqrt{w} - e^{\max}(2-\eta) \lesseqgtr 0 . \qquad \text{[A1-6]}$$

If [A1-6] is negative or zero, employees would provide effort $e = \dfrac{wP^2}{e^{\max}(2-\eta)^2}$ since [A1-5] is positive.

However, if [A1-6] is positive, the difference between [A1-5] and [A1-6] is positive:

$$\frac{1}{e^{\max}(2-\eta)}[P\sqrt{w} - \{e^{\max}(2-\eta)\}]^2 > 0 .$$

This says that [A1-5] is larger than [A1-6]. Thus employees choose [A1-5] that yields the higher expected utility. They provide effort $e = \dfrac{wP^2}{e^{\max}(2-\eta)^2}$. Regardless of whether [A1-6] is negative, zero or positive, [A1-5] is always larger than [A1-6]. Thus, case (ii) has higher expected utility than case (iii).

[Appendix 2]

The condition that international fragmentation occurs is that $(36)-(35)<0$:

$$[\{\frac{\Psi(\tilde{\tilde{n}}-1)w}{e^{\max}} + \frac{(\Psi+\beta)w}{e^{\max}}\} + \{\frac{(\tilde{\tilde{n}}-1)b(\tilde{\tilde{n}})w}{e(\tilde{\tilde{n}})} + \frac{b(\tilde{\tilde{n}})w^*}{e^*(\tilde{\tilde{n}})}\}] - \{\frac{\Psi w\tilde{\tilde{n}}}{e^{\max}} + \frac{\tilde{\tilde{n}}b(\tilde{\tilde{n}})w}{e(\tilde{\tilde{n}})}\} < 0 .$$

This is rearranged as

$$\frac{\beta w}{e^{\max}} + b(\tilde{\tilde{n}})\{\frac{-w}{e(\tilde{\tilde{n}})} + \frac{w^*}{e^*(\tilde{\tilde{n}})}\} < 0 . \qquad \text{[A2-1]}$$

Using (18), where $\bar{n} = \tilde{\tilde{n}}$,

$$\frac{w}{e(\tilde{\tilde{n}})} = \frac{e^{\max}(2-\eta)^2 l}{\tilde{\tilde{n}}} . \qquad \text{[A2-2]}$$

Recall that the sizes of the pool for national fragmentation and international fragmentation are assumed to be the same, l. Using [A2-2],

$$\frac{w^*}{e^*(\tilde{\tilde{n}})} = \frac{e^{\max}(2-\eta^*)^2 l}{\tilde{\tilde{n}}} . \qquad \text{[A2-3]}$$

For simplicity, I assume that $\delta = 1$. From (1),

$$b(\tilde{\tilde{n}}) = \frac{1}{\tilde{\tilde{n}}} . \qquad \text{[A2-4]}$$

Substituting [A2-2], [A2-3] and [A2-4] into [A2-1],

$$\frac{\beta w}{e^{\max}} + \frac{e^{\max} l}{\tilde{\tilde{n}}^2}\{-(2-\eta)^2 + (2-\eta^*)^2\} < 0 . \qquad \text{[A2-5]}$$

From (23), where $\delta = 1$, $\varepsilon = 1$ and $\Psi = \dfrac{\theta}{N}$, $\tilde{\tilde{n}} = [\dfrac{\{e^{\max}(2-\eta)\}^2 l}{\Psi w}]^{\frac{1}{2}}$. Substituting $\tilde{\tilde{n}}$ into [A2-5],

$\dfrac{w}{e^{\max}}[\beta + \Psi\{-1 + (\dfrac{2-\eta^*}{2-\eta})^2\}] < 0$. Since $\dfrac{w}{e^{\max}} > 0$, $[\beta + \Psi\{-1 + (\dfrac{2-\eta^*}{2-\eta})^2\}] < 0$. This means $\beta < \Psi\{1 - (\dfrac{2-\eta^*}{2-\eta})^2\}$. Since $\Psi = \dfrac{\theta}{N} < 1$ in (9) and $0 < (\dfrac{2-\eta^*}{2-\eta}) < 1$ from (26), $\beta < \Psi\{1 - (\dfrac{2-\eta^*}{2-\eta})^2\} < 1$.

[Appendix 3]

The first derivative of tc_m^f in (39') with respect to m is

$$\frac{\partial tc_m^f}{\partial m} = \frac{w}{e^{\max}}[\Psi + 2\beta\{m - \hat{n}(m)\}\{1 - \frac{\partial \hat{n}(m)}{\partial m}\}] + [d\{\frac{\partial \hat{n}(m)}{\partial m}\} + d^*\{1 - \frac{\partial \hat{n}(m)}{\partial m}\}]m^{-2}$$

$$-2[d\hat{n}(m) + d^*\{m - \hat{n}(m)\}]m^{-3} = 0 .$$

This is rearranged as

$$\frac{w}{e^{\max}}[\Psi + 2\beta\{m - \hat{n}(m)\}\{1 - \frac{\partial \hat{n}(m)}{\partial m}\}] + \{(d - d^*)\frac{\partial \hat{n}(m)}{\partial m} + d^*\}m^{-2}$$

$$-2\{(d - d^*)\hat{n}(m) + d^* m\}m^{-3} = 0 . \qquad \text{[A3-1]}$$

Rearranging (41),

$$\{m - \hat{n}(m)\} = (d - d^*)(\frac{e^{\max}}{2\beta w})m^{-2} . \qquad \text{[A3-2]}$$

The derivative of $\hat{n}(m)$ in (41) with respect to m is

$$\frac{\partial \hat{n}(m)}{\partial m} = 1 + (d - d^*)(\frac{e^{\max}}{\beta w})m^{-3} . \qquad \text{[A3-3]}$$

Substituting (41), [A3-2] and [A3-3] into [A3-1],

$$\frac{\partial tc_m^f}{\partial m} = \frac{w}{e^{\max}}[\Psi + 2\beta\{(d - d^*)(\frac{e^{\max}}{2\beta w})m^{-2}\}\{-(d - d^*)(\frac{e^{\max}}{\beta w})m^{-3}\}]$$

$$+[(d - d^*)\{1 + (d - d^*)(\frac{e^{\max}}{\beta w})m^{-3}\} + d^*]m^{-2}$$

$$-2[(d - d^*)\{m - (d - d^*)(\frac{e^{\max}}{2\beta w})m^{-2}\} + d^* m]m^{-3} = 0 . \qquad \text{[A3-4]}$$

[A3-4] is rearranged as

$$\{\frac{w\Psi}{e^{\max}} - (d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5}\} + \{dm^{-2} + (d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5}\}$$

$$-2\{dm^{-2} - (d-d^*)^2(\frac{e^{\max}}{2\beta w})m^{-5}\} = 0 .$$

Then $\frac{\partial tc_m^f}{\partial m} = \frac{w\Psi}{e^{\max}} - dm^{-2} + \{\frac{e^{\max}(d-d^*)^2}{\beta w}\}m^{-5} = 0$.

[Appendix 4]

I need to prove that $(B-A) > 0$ for a small value of m. Rearrange the second derivative in (43),

$\{2dm^{-3} - 5(d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-6}\} > 0$. Factoring out,

$$2m^{-1}\{dm^{-2} - (2.5)(d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5}\} > 0 .$$

Since $2m^{-1} > 0$, $\{dm^{-2} - (2.5)(d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5}\} > 0$. Rewriting this,

$$\{dm^{-2} - (d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5} - (1.5)(d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5}\} > 0 . \quad \text{[A4-1]}$$

From (44), $(B-A) = dm^{-2} - \{\frac{w\Psi}{e^{\max}} + (d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5}\}$. Substituting this equation into [A4-1],

$[\{(B-A) + \frac{w\Psi}{e^{\max}}\} - (1.5)(d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5}] > 0$. Rearranging this,

$$(B-A) > -\frac{w\Psi}{e^{\max}} + (1.5)(d-d^*)^2(\frac{e^{\max}}{\beta w})m^{-5} . \quad \text{[A4-2]}$$

We want to know the sign of $(B-A)$ in the range left of $\tilde{m}$ on the horizontal axis in Figure 7. This sign is determined when the value on the RHS of [A4-2] is obtained. Thus, I calculate this value where m has a value smaller than $\tilde{m}$, such that $m = 1$ as an example. That is

$$-\frac{w\Psi}{e^{\max}} + (1.5)(d-d^*)^2(\frac{e^{\max}}{\beta w}) . \quad \text{[A4-3]}$$

Using the notations for d and d^* in (38),

$$(d-d^*)^2 = [e^{\max} l^f \{(2-\eta)^2 - (2-\eta^*)^2\}]^2 = [e^{\max} l^f \{(\eta^* - \eta)(4-\eta-\eta^*)\}]^2. \quad \text{[A4-4]}$$

Substituting [A4-4] and $\Psi = \dfrac{\theta}{N}$ into [A4-3], [A4-3] is rewritten as

$$\frac{w}{e^{\max}} [-\frac{\theta}{N} + \frac{(1.5)e^{\max 4}\{(\eta^* - \eta)(4-\eta^*-\eta)\}^2 l^{f\,2}}{\beta w^2}]. \quad \text{[A4-5]}$$

I check the magnitudes of the two terms in the square brackets in [A4-5]. First, consider the numerators. The parameter θ is the input requirement of effective labor per unit of headquarters service. l^f is the total labor input for the production process. Conceptually, θ should be smaller than l^f : $\theta < l^f$. Assume that $e^{\max} \geq 1$. The conditions that $0 < \eta \leq 1$, $0 < \eta^* \leq 1$ and $\eta^* > \eta$ make $0 < (\eta^* - \eta) < 1$ and $2 < (4 - \eta^* - \eta) < 4$. I assume that $(\eta^* - \eta)$ ensures that $(1.5)e^{\max 4}\{(\eta^* - \eta)(4-\eta^*-\eta)\}^2 l^f \geq 1$. Since $\theta < l^f$, the relation between the two numerators in the square bracket is

$$\theta < (1.5)e^{\max 4}\{(\eta^* - \eta)(4-\eta^*-\eta)\}^2 l^{f\,2}.$$

Next, consider the denominators in the square brackets in [A4-5]. If the Northern network size N is represented by the value of the network infrastructure, N is larger than w^2. Note that w is the wage per worker. The condition in (37), $\beta < 1$, makes that $\beta w^2 < w^2$. This yields that $N > \beta w^2$. Then the absolute value of the first term in the square brackets in [A4-5] is smaller than the value of the second term:

$$\left| -\frac{\theta}{N} \right| < \frac{(1.5)e^{\max 4}\{(\eta^* - \eta)(4-\eta^*-\eta)\}^2 l^{f\,2}}{\beta w^2}.$$

This makes [A4-5] positive and thus [A4-3] is positive. This says that $(B - A) > 0$ from [A4-2]. In other words, $B > A$ in the left range of $\tilde{m}$.

[Appendix 5]

I have to check whether there is a case that the headquarters remains in the North and all production stages shift to the South at the values of m which are determined by (42), in other words, at the values of m that are determined by the intersection of the curves A and B. This case means that there is no production stage in the North: $\hat{n}(m) = 0$. Substituting $\hat{n}(m) = 0$ into (41), m is obtained. When I define it as m',

$$m' = \{\frac{e^{\max}(d - d^*)}{2\beta w}\}^{\frac{1}{3}}. \quad \text{[A5-1]}$$

If this extreme case exists, the condition in (42), $\frac{\partial tc_m^f}{\partial m} = 0$, should also be satisfied where $m = m'$.

$$\underbrace{\frac{w\Psi}{e^{\max}} - dm'^{-2} + \{\frac{e^{\max}(d-d^*)^2}{\beta w}\}m'^{-5}}_{H} = 0 .$$

If H is not zero, this extreme case does not exist. If H is zero, this extreme case can exist.

To check whether H is zero, rewrite H as follows.

$$\frac{1}{m'^5}[\frac{w\Psi}{e^{\max}} m'^5 - dm'^3 + \{\frac{e^{\max}(d-d^*)^2}{\beta w}\}] . \qquad \text{[A5-2]}$$

Since $m'^5 > 0$, I should check whether the value in the square brackets is zero. Using [A5-1], the expression in the square brackets of [A5-2] is rewritten:

$$[\frac{w\Psi}{e^{\max}}\{\frac{e^{\max}(d-d^*)}{2\beta w}\}^{\frac{2}{3}} + (d-2d^*)]\{\frac{e^{\max}(d-d^*)}{2\beta w}\} . \qquad \text{[A5-3]}$$

Since $d > d^*$ in (38), $\frac{w\Psi}{e^{\max}}\{\frac{e^{\max}(d-d^*)}{2\beta w}\}^{\frac{2}{3}} > 0$ and $\{\frac{e^{\max}(d-d^*)}{2\beta w}\} > 0$. Then, from [A5-3], three cases are considered.

(i) If $d - 2d^* \geq 0$, [A5-3] is non zero.

(ii) If $d - 2d^* < 0$ and if $\frac{w\Psi}{e^{\max}}\{\frac{e^{\max}(d-d^*)}{2\beta w}\}^{\frac{2}{3}} \neq \left|d - 2d^*\right|$, [A5-3] is non zero.

(iii) If $d - 2d^* < 0$ and if $\frac{w\Psi}{e^{\max}}\{\frac{e^{\max}(d-d^*)}{2\beta w}\}^{\frac{2}{3}} = \left|d - 2d^*\right|$, [A5-3] is zero.

The extreme case exists if case (iii) occurs.

I now explain the likelihood that case (iii) occurs, compared with the cases (i) and (ii). The case (iii) has more restrictive conditions than the case (ii). The reason is that the situations satisfying the second condition in (iii), $\frac{w\Psi}{e^{\max}}\{\frac{e^{\max}(d-d^*)}{2\beta w}\}^{\frac{2}{3}} = \left|d - 2d^*\right|$, occur less frequently than the situations satisfying the second condition in (ii). Also, case (iii) has more restrictive conditions than case (i). Then the likelihood that case (iii) can occur is lower than the cases (i) and (ii). In fact, since the second condition in (iii)

consists of many parameters, it is very hard to find a combination of the values of the parameters that satisfy the second condition in (iii). Thus, the likelihood of case (iii) is very low.

Another way to avoid the extreme case is to assume that $d - 2d^{*} \geq 0$, like the condition in (i). This condition makes that $\eta^{*} > \frac{\eta}{\sqrt{2}} + (2 - \sqrt{2})$, when $d = e^{\max}(2-\eta)^2 l^f$ and $d^{*} = e^{\max}(2-\eta^{*})^2 l^f$ in (38) are substituted into $d - 2d^{*} \geq 0$. Also this inequality ensures that $0 < \eta < 1$ and $0 < \eta^{*} < 1$.

[Appendix 6]

Using $A(m)$ and $B(m)$ in (44),

$$\frac{A(m)}{B(m)} = \frac{\frac{w\Psi}{e^{\max}} + \{\frac{e^{\max}(d-d^{*})^2}{\beta w}\}m^{-5}}{dm^{-2}} = \frac{1}{d}\{\frac{w\Psi m^2}{e^{\max}} + \frac{e^{\max}(d-d^{*})^2 m^{-3}}{\beta w}\}. \quad \text{[A6-1]}$$

The ratio where $m = \tilde{\tilde{n}}$, $\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})}$, is obtained as follows. Substituting $\tilde{\tilde{n}}$ for m in [A6-1],

$$\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})} = \frac{1}{d}\{\frac{w\Psi\tilde{\tilde{n}}^2}{e^{\max}} + \frac{e^{\max}(d-d^{*})^2\tilde{\tilde{n}}^{-3}}{\beta w}\}. \quad \text{[A6-2]}$$

Using $\tilde{\tilde{n}}$ in (23), $\delta = 1$ and $\varepsilon = 1$,

$$\tilde{\tilde{n}} = [\frac{\{e^{\max}(2-\eta)\}^2 lN}{\theta w}]^{\frac{1}{2}}. \quad \text{[A6-3]}$$

From the definitions of d and d^{*} in (38),

$$(d - d^{*}) = e^{\max} l^f \{(2-\eta)^2 - (2-\eta^{*})^2\}. \quad \text{[A6-4]}$$

Substituting [A6-3] and [A6-4] into [A6-2],

$$\frac{A(\tilde{\tilde{n}})}{B(\tilde{\tilde{n}})} = \{\frac{1}{e^{\max}(2-\eta)^2 l^f}\}[\frac{\Psi e^{\max}(2-\eta)^2 lN}{\theta} + \frac{l^{f\,2}\{(2-\eta)^2 - (2-\eta^{*})^2\}^2 \theta^{\frac{3}{2}} w^{\frac{1}{2}}}{\beta (lN)^{\frac{3}{2}}(2-\eta)^3}]$$

$$= \frac{l}{l^f}[1 + \frac{\{(2-\eta)^2 - (2-\eta^{*})^2\}^2 \theta^{\frac{3}{2}} w^{\frac{1}{2}}}{e^{\max}\beta(2-\eta)^5 l^{\frac{1}{2}} (\frac{l}{l^f})^2 N^{\frac{3}{2}}}]. \quad \text{[A6-5]}$$

[Appendix 7]

Using (25), $\frac{\partial e^*(m)}{\partial w^*} = \{\frac{1}{e^{*\max}(2-\eta^*)^2}\frac{m}{l^f} + \frac{w^*}{e^{*\max}(2-\eta^*)^2}\frac{1}{l^f}\frac{\partial m}{\partial w^*}\}$. Since $e^{*\max} = e^{\max}$ and $\frac{\partial m}{\partial w^*} = 0$, $\frac{\partial e^*(m)}{\partial w^*} = \{\frac{1}{e^{\max}(2-\eta^*)^2}\frac{m}{l^f}\} > 0$.

Using (24), $\frac{\partial e(m)}{\partial w^*} = \{\frac{w}{e^{\max}(2-\eta)^2 l^f}\}\frac{\partial m}{\partial w^*}$. Since $\frac{\partial m}{\partial w^*} = 0$, $\frac{\partial e(m)}{\partial w^*} = 0$.