The theory of irrelevance of the size of the firm

Pak-Wai Liu\textsuperscript{a}, Xiaokai Yang\textsuperscript{b,*,1}

\textsuperscript{a} Department of Economics, Chinese University of Hong Kong, Hong Kong, PR China
\textsuperscript{b} Center for International Development, Harvard University, One Eliot Street Cambridge, MA 02138, USA

Received 8 August 1997; received in revised form 10 November 1999; accepted 22 November 1999

Abstract

This paper formalizes Cheung, Coase, Stigler, and Young's theory of irrelevance of the size of the firm. This theory states that if division of labor develops within the firm, the average size of the firm and productivity go up side by side. If division of labor develops between firms, the average size of firms decreases as productivity goes up. Inframarginal analysis of the trade off between positive network effects of division of labor on aggregate productivity and transaction costs can predict recently popular business practices of down sizing, outsourcing, contracting out, focusing on core competence, and disintegration. We present evidence for a negative correlation between average employment of labor by firms and productivity. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: D23; L11; L22; L23

Keywords: Size of the firm; Division of labor; Theory of the firm

1. Introduction

Many scholars have observed that average employment of labor by firms (average firm size) in many industries in both advanced countries and newly industrialized countries has declined. Data clearly show that in numerous developed countries, such as the US, UK, France, Austria and Belgium, employment share of small and medium enterprises in their manufacturing industries exhibits a U-shape pattern. It declined initially in the 1960s and early 1970s but started to increase in the late 1970s and early 1980s\textsuperscript{2}. Among the newly industrialized countries, South Korea’s manufacturing industries exhibit the same pattern as the developed countries, with a turning point in the early 1980s\textsuperscript{3}. In Taiwan, average
firm size in manufacturing measured in terms of average employment increased from 7.3 persons per firm in 1954 to a peak of 28.2 in 1971 after which it followed a decline trend, reaching 24.2 in 1984. Similarly in the 1980s, manufacturing firms in Singapore were getting smaller in size on the average.

This inverted U-shaped trend is not confined to manufacturing industries. Liu (1992) shows that in Hong Kong, besides manufacturing, industries such as business services, import/export and restaurant exhibit a trend of declining average firm size. Table 1 shows the change in firm size of these four industries measured in two ways, the average number of employees per firm and the percent of employees engaged in firms with employees larger than 50. The trend since the late 1970s when data became available is generally downward. The decline in the average size of firms is associated with an increase in per capita real income and total factor productivity.

This trend of declining average firm size is contrary to the common perception that due to technological change and economies of scale, firm size should get larger over time. For instance, Kim (1989) shows from his model that because there are economies of scale arising from increasing returns due to specialization, firm size gets larger with specialization and the extent of the market. Our study provides an explanation of why firm size may get smaller over time on the basis of an analysis of the division of labor and transaction cost. We will show that the institution of the firm will emerge if the transaction efficiency for labor is higher than that for intermediate goods. By transaction efficiency we measure the fraction of the purchased good which disappears in transit due to transaction cost. Transaction competes for time and management. For instance, a producer of an intermediate good can either pay the search cost to find a supplier of the primary good used in its production or the managerial monitoring cost of in-house production of the primary good. We will show that given the emergence of firms, their size (measured by employment) decreases

---

### Table 1
Declining firm size in Hong Kong

<table>
<thead>
<tr>
<th>Year</th>
<th>Manufacturing</th>
<th>Import/export</th>
<th>Business services</th>
<th>Restaurants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>23.72</td>
<td>6.35</td>
<td>10.67</td>
<td>25.27</td>
</tr>
<tr>
<td>1981</td>
<td>20.61</td>
<td>6.74</td>
<td>11.30</td>
<td>27.42</td>
</tr>
<tr>
<td>1984</td>
<td>19.10</td>
<td>6.07</td>
<td>10.42</td>
<td>28.31</td>
</tr>
<tr>
<td>1987</td>
<td>18.60</td>
<td>6.21</td>
<td>9.78</td>
<td>23.49</td>
</tr>
<tr>
<td>1990</td>
<td>14.47</td>
<td>5.45</td>
<td>8.50</td>
<td>21.30</td>
</tr>
<tr>
<td>1993</td>
<td>13.13</td>
<td>5.10</td>
<td>7.38</td>
<td>21.47</td>
</tr>
</tbody>
</table>

Percent of employees in industry engaged in firms with employees larger than 50

<table>
<thead>
<tr>
<th>Year</th>
<th>Manufacturing</th>
<th>Import/export</th>
<th>Business services</th>
<th>Restaurants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>58.72</td>
<td>23.45</td>
<td>52.85</td>
<td>59.03</td>
</tr>
<tr>
<td>1984</td>
<td>59.07</td>
<td>20.20</td>
<td>51.20</td>
<td>57.72</td>
</tr>
<tr>
<td>1987</td>
<td>57.77</td>
<td>23.14</td>
<td>50.47</td>
<td>56.25</td>
</tr>
<tr>
<td>1990</td>
<td>53.36</td>
<td>21.15</td>
<td>52.31</td>
<td>51.62</td>
</tr>
<tr>
<td>1993</td>
<td>52.73</td>
<td>18.98</td>
<td>45.45</td>
<td>42.74</td>
</tr>
</tbody>
</table>

* Source: Census and Statistics Department, Annual Digest of Statistics, Hong Kong Government, various years.

---

if the transaction efficiency for intermediate goods becomes higher than that for labor. The decrease in the size of firms is driven by two forces. First, each firm becomes more specialized if division of labor develops among firms, so that each firm’s scope of activities is reduced. However, if division of labor and specialization are developed within each firm, the level of specialization of each worker in a firm and the scope of the firm may increase hand in hand. But if the transaction efficiency for intermediate goods is higher than that for labor, then organizing division of labor among more specialized firms will be more efficient than organizing division of labor within a firm since the former involves more transactions of intermediate goods while the latter involves more trade in labor, thus generating a decline in the size of firms.5

Several recent studies including Grossman and Hart (1986); Hart and Moore (1990); Milgrom and Roberts (1990); Kreps (1990), and Lewis and Sappington (1991) explain the size of the firm in various ways. None of these models, however, can predict the negative relationship between an increase in per capita real income and the average size of firms, nor can they explore the interaction between that negative relationship and the level of specialization.

Our model extends the Camacho’s (Camacho, 1991, 1996) analysis of the trade-off between benefits of division of labor and co-ordination costs as an explanation of the optimal size of the firm. Echoing Jones’ empirical works (Jones, 1995a, b) which conclusively reject a scale effect, our theory explains economic development on the basis of the division of labor instead of economies of scale. It in effect formalizes Allyn Young’s (Young, 1928) criticism of the notion of economies of scale. It is also in keeping with the arguments of Young, Coase (1937); Stigler (1951); Cheung (1983), and Langlois (1988) which imply that the size of the firm may increase or decrease, depending on whether division of labor is organized through the labor market or through the market for intermediate goods.

In Section 2, we use an example to illustrate the story behind the formal model. In Section 3 a model of consumer-producers with economies of specialization in production and transaction costs in trading is presented. Section 4 characterizes eight market structures that may emerge from the development in division of labor and identifies the general equilibrium by comparing per capita real incomes in different structures. The main theoretical results of the paper on the emergence of the institution of the firm and average employment are contained in Section 5. The paper concludes in Section 6.

2. The theories of the firm and the story behind the formal model

There are three strands of the literature on the institution of the firm and transaction costs that closely relate to our paper. Recent literature of general equilibrium models with economies of scale and endogenous number of goods, represented by Dixit and Stiglitz (1977); Krugman (1980, 1981); Ethier (1982); Judd (1985), Romer (1990); Grossman and

---

5 A firm’s decision to increase its level of specialization which leads to the equilibrium phenomenon of disintegration and increasing division of labor between firms is called ‘outsourcing’, ‘downsizing’, ‘returning to core competencies’, and ‘diverting unrelated businesses’ in the literature of management. See for instance, Milgrom and Roberts (1992).
Helpman (1989, 1990); Fujita and Krugman (1995), and Krugman and Venables (1995), generates some important empirical implication, which is referred to as scale effects.

There are several types of scale effects. Type I scale effect exists if productivity or growth performance is positively correlated with the average size of firms. Type II scale effect exists if the growth rate of per capita income or per capita consumption goes up as the size or growth rate of population increases. Type III scale effect exists if the growth rate of per capita income goes up as the size of the research and development sector increases. Type IV scale effect exists if the growth rate of per capita income goes up as the investment rate increases. All four types of scale effects are wildly at odds with empirical evidence. The AK model and R&D based model have generated scale effects of Types II, III, and IV. The empirical work of Jones (Jones, 1995a, b), and National Research Council (1986) has rejected the three types of scale effect. Jones (Jones, 1995a, b), Alwyn Young (Young, 1998), and Segerstrom (1998) suggest several ways to avoid Type III scale effect in the R&D based model. But the modified models still have a Type-II scale effect. As Jones has shown, endogenous growth cannot be preserved in the neo-classical endogenous growth models in the absence of scale effect. On the ground of the empirical evidence, it is concluded that the neo-classical endogenous growth models have not provided a convincing explanation of the driving mechanism of economic growth (Jones, 1995a, pp. 508–509).

This class of models with economies of scale generates other types of scale effects too. For instance, the Fujita and Krugman model (1995) predicts a positive correlation between the degree of urbanization and the average size of firms. The Krugman and Venables model (1995) predicts a positive correlation between the degree of industrialization and the average size of firms. A preliminary examination of data does not support all of the scale effects. Coase, Stigler, Young, and Cheung have proposed a theory of irrelevance of the size of the firm. According to them, if division of labor develops within the firm, then the average size of firms goes up as productivity increases. But if division of labor develops between firms, then as productivity rises the average size of firms falls. This paper will use a general equilibrium model to formalize this idea and provide a theoretical foundation for empirical work that rejects various scale effects.

The second strand of the literature relates to the notion of endogenous transaction costs. Many economists of information economics claim that the principal-agent model with moral hazard, the model of adverse selection, and other models with information asymmetry provide microeconomic theories of the firm. However, as Hart (1991) suggests, the principal-agent model has nothing to do with the theory of the firm. The contingent contract in the principal-agent model is not labor contract because of the assumption that labor effort is not observable or is non-verifiable in the court when dispute arises. Also, both parties of the contingent contract claim part of residual returns. In addition, decision authority is symmetrically distributed, so that asymmetric distribution of residual returns and residual control rights, which are two essential components of the firm, does not occur in the model. All models of information asymmetry do not explain why and how firms emerge, though they study endogenous transaction costs associated with information problems. In this paper, we consider endogenous transaction costs as a departure of the equilibrium from the Pareto optimum. But recent dynamic game models with information asymmetry show that information asymmetry may reduce endogenous transaction costs. For instance, the Milgrom and Roberts model (1982) shows that information asymmetry can reduce the
distortions caused by monopoly power. From the game models we can learn that information asymmetry may not be a source of endogenous transaction cost. At least the endogenous transaction costs caused by information asymmetry are not as serious as the naïve model of adverse selection, which ignores complicated interactions between information and strategies, predicts. As Hurwicz (1973) points out, the distortion in the Walrasian equilibrium with extreme information asymmetry (each player does not know others’ utility and production functions and endowment and has no information about the distribution functions of their characteristics) is smaller than any other possible mechanisms and also involves lower information cost. Hence, information asymmetry and related endogenous transaction costs are not essential for the theory of the firm. This claim has been supported by two research lines.

First, Cheung (1983) argues that the institution of the firm is to replace trade in goods with trade in labor. According to this argument, if trading of goods involves endogenous transaction costs caused by moral hazard or other types of information asymmetry, trading in labor may involve endogenous transaction costs as well. Great measurement costs of quantity and quality of goods may cause information problem, so that endogenous transaction costs result. But quantity and quality of labor, in particular, labor allocated to research and entrepreneurial activities, involve prohibitively high measurement costs. This may cause information problem too. Hence, endogenous transaction cost per se is not enough for explaining the emergence of the firm. The key point is to explain why trading of goods is replaced with trading in labor since this replacement is essential for explaining the emergence of the firm.

Second, the recent literature of incomplete contract was believed to point to a hopeful research direction that may provide a theory of the firm. However, Maskin and Tirole (1997) poses a convincing challenge against the theory of incomplete contract. According to them, this theory does not have sound theoretical foundation. It involves logical inconsistency, so that it cannot provide a theoretical foundation for the theory of the firm. Also, Holmstrom and Milgrom (1995) and Holmstrom and Roberts (1998) point out that labor trade, which is essential for the existence of the firm, is not essential in the model of incomplete contract. Hence, the theory of incomplete contract is a theory of optimum ownership structure rather than a theory of the firm.

Yang and Ng (1995) have developed a general equilibrium model to formalize Coase and Cheung’s theory of the firm. Since many readers are not familiar with the theory and its technical substance, we briefly outline the story behind their model in this section.

According to Yang and Ng, the institution of the firm is a structure of transactions based on the division of labor that satisfy the following three conditions.

(i) There are two types of trade partners who are associated with a firm: the employer (she) and the employee (he). There is an asymmetric distribution of control rights or authority. The employer has control rights to use the employee’s labor. He must do what he is told to do, if she so demands. That is, the ultimate rights to use the employee’s labor are owned by the employer, subject to the employee’s freedom to quit the job and to other terms of the labor contract.

(ii) The contract that is associated with the asymmetric relationship between the employer and the employee never specifies how much the employer receives from the relationship. It only specifies how much the employee receives from the relationship. The employer
claims the residual returns, defined as the revenue of the firm net of the wage bill and other expenses. The employer is referred to as the owner of the firm. One of the most important components of the ownership of a modern firm is the entitlement to the business name of the firm. The exclusive rights to the business name are enforced through a business name search process when the firm is registered with the government and through recognition of the name in legal cases in the judicial process. The exclusive rights are also enforced through the laws of brand.

(iii) A firm must involve a process that transforms labor of the employee into something that is sold in the market by the owner of the firm. In the process, what is produced by the employee is owned by the employer (residual returns).

The relationship between a professor and his housekeeper does not involve the institution of the firm since it does not involve such a resale process of services provided by the housekeeper. The professor directly consumes what the housekeeper produces and never resells it to the market.

Yang and Ng’s story of the firm runs as follows. There are many ex ante identical consumer-producers. Each individual as a consumer must consume a final good, called cloth, the production of which requires an intermediate good, called management service, as an input. There is a trade-off between economies of specialization and transaction costs. If a transaction cost coefficient for a unit of goods or labor traded is small, then division of labor will occur at equilibrium. Otherwise autarky will be chosen as the equilibrium. There are two ways to interpret the transaction cost. First, the transaction cost can be considered as resource allocated to measure quantity and quality of goods and labor. Following common practice in the literature of equilibrium models of geography and monopolistic competition, we can use an iceberg transaction cost coefficient to represent such transaction cost. Second, the distortions caused by moral hazard and information asymmetry can be considered as endogenous transaction cost. Since simultaneous endogenization of transaction costs, specialization and the emergence of the firm involves formidable algebra, we use the iceberg transaction cost represents the anticipated endogenous transaction cost. For trade in goods, the anticipated endogenous transaction cost may be associated with the distortion caused by information asymmetry about the quality of goods, as in the model of lemons (Akerlof, 1970). For trade in labor, the anticipated endogenous transaction cost may be associated with moral hazard and shirking in the model of efficiency wage (Shapiro and Stiglitz, 1984).

There are three different structures of residual rights which can be used to organize transactions required by the division of labor. Structure D, shown in Fig. 1b, comprises markets for cloth and management services. Specialist producers of cloth exchange that product for the specialist services of management. For this market structure, residual rights to returns and authority are symmetrically distributed between the trade partners, and no firms or labor market exist. Structure E, shown in Fig. 1c, comprises the market for cloth and the market for labor hired to produce the management service within a firm. The producer of cloth is the owner of the firm and specialist producers of management services are employees. Control rights over employees’ labor and rights to the firm’s residual returns are asymmetrically distributed between the employer and her employees. The employer claims the difference between revenue and the wage bill, has control rights over her employees’ labor, and sells goods that are produced from employees’ labor. Structure F, shown in Fig. 1d comprises the market for cloth and the market for labor hired to produce cloth
within a firm. The professional manager is the owner of the firm and specialist producers of
cloth are employees. For the final two structures of residual rights, the firm emerges from
the division of labor. Compared with structure D, these involve a labor market but not a
market for management services. As Cheung (1983) argues, the firm replaces the market
for intermediate goods with the market for labor. Although both structures E and F involve
a firm and an asymmetric structure of residual rights, they entail different firm ownership
structures.

Suppose that the transaction cost coefficient is much larger for management service than
for labor. This is very likely to be the case in the real world, since the quality and quantity
of the intangible entrepreneurial ideas are prohibitively expensive to measure. Potential
buyers of the intellectual property in entrepreneurial ideas may refuse to pay by claiming
that these are worthless as soon as they are acquired from their specialist producer. Under
this circumstance, the institution of the firm can be used to organize the division of labor
more efficiently because it avoids trade in intangible intellectual property.

Suppose further that the transaction cost coefficient for labor hired to produce manage-
ment services is much larger than for labor hired to produce cloth because it is prohibitively
expensive to measure the efforts exerted in producing intangible management services (can
you tell if a manager sitting in the office is pondering business management or his girl-
friend?). Then the division of labor can be more efficiently organized in structure F than in
structure E. This is because structure F involves trade in cloth and in labor hired to produce
cloth but not trade in management services nor in labor hired to produce management ser-
vices, while structure E involves trade in cloth and in labor hired to produce management
services. Hence, structure F will occur at equilibrium if the transaction cost coefficients for
labor hired to produce cloth and for cloth are sufficiently small. The claim to the residual
return of the firm by the manager is the indirect price of management services. Therefore,
the function of the asymmetric structure of residual rights is to get the activity with the
lowest trading efficiency involved in the division of labor while avoiding direct pricing and
marketing of the output and input of that activity, thus promoting the division of labor and
productivity. In a sense, the function of the asymmetric structure of residual rights is similar
to that of a patent law which enforces rights to intangible intellectual property, thereby
promoting the division of labor and productivity in producing the intangible. However, the
asymmetric structure of residual rights to returns and control can indirectly price those intangible intellectual properties which are prohibitively expensive to price even through a patent law.

Intuitively, there are two ways to do business if an individual has an idea for making money. The first is to sell the entrepreneurial idea in the market. This is very likely to create a situation in which everybody can steal the idea and refuse to pay for its use. The second way of proceeding is for the entrepreneur to hire workers to realize the idea, while keeping the idea to herself as a business secret. Then she can claim as her reward the residual returns to the firm, which represent the difference between revenue and the wage bill. If the idea is a good one, then the entrepreneur will make a fortune. If the idea is a bad one, she will become bankrupt. The residual return is the precise price of the idea and the entrepreneur gets what she deserves, so that stealing and over or under pricing of intellectual property is avoided. To understand this, you may imagine that Bill Gates did not set up a company to hire others to realize his entrepreneurial ideas about how to make money from computer software. Instead he sold his ideas as consultant services. Would anybody pay billion dollars (which is the real market value of the ideas indicated by Bill Gates residual returns) to buy the ideas?

This theory is referred to as the theory of indirect pricing. The Yang–Ng model (1995) formalizes the Cheung–Coase theory of the firm (Coase, 1937; Cheung, 1983) that the institution of the firm can save on transaction costs by replacing trade in intermediate goods with trade in labor if the former involves higher transaction costs than the latter. The model does not endogenize transaction costs if these are defined as a departure from the Pareto optimum. But as discussed before, endogenous transaction cost is not essential for telling story of the firm though we can use some transaction cost coefficient to represent anticipated endogenous transaction cost. The essence of the story is that why endogenous or exogenous transaction costs of goods are replaced with endogenous or exogenous transaction costs of labor as the institution of the firm emerges.

The essence of the theory of indirect pricing can be summarized as follows. If individuals engage in division of labor in producing two goods \( x \) and \( y \), trade in two of the four elements comprising outputs \( x, y \), and inputs \( l_x, l_y \) can be used to organize the division of labor. Hence, there are six possible structures of transactions (two combinations of four elements) that can be used to organize the division of labor: \( x \) and \( y \) (structure D), \( y \) and \( l_x \) (structure E), \( y \) and \( l_y \) (structure F), \( x \) and \( l_y \) (infeasible since a specialist of \( x \) who must consume \( y \) cannot buy \( y \)), \( x \) and \( l_x \) (infeasible too), \( l_x \) and \( l_y \) (a structure that may be seen in some collective organization with direct exchange of labor but is rarely seen in the real world). Hence, individuals can at least consider three of the structures and choose one of them to avoid trade that involves the lowest trading efficiency. (Yang and Ng, 1993, Chapter 13) have shown that as the number of traded goods and the level of division of labor increase, the number of possible structures of transactions increases more than proportionally. This implies that choice of transaction structure is increasingly more important than choice of production structure and resource allocation for a given structure of production for economic development as division of labor develops. This is why in a highly commercialized society (high level of division of labor) there are more opportunities for entrepreneurs to make fortunes from playing around with structures of transactions. In general, the number of possible pricing structures of goods and factors increases more than proportionately as the network size of division of labor is
enlarged. For instance, one of recent five patented internet trading practices pay reviewers of advertisements for clicking an ad or provide free email services. The generous pricing policies can make a lot of money using positive network effects that pay the companies via other channels. Hence, conventional marginal analysis of optimum price of a single good is out-of-date. In a large and highly interconnected network of division of labor and communication, players can price a subset of the set of all goods and services. It is a very complicated optimization decision problem to identify which subset of all goods and services should be priced and which can be provided free of charge. The function of the market in sorting out the efficient pricing structure is even much more complicated than a player’s problem of optimum pricing structure.

The Yang–Ng model cannot predict a negative correlation between productivity and the average size of firms since it involves only two goods. In the next section we introduce three goods in to the model to formalize the theory of irrelevance of the size of the firm.

3. A model with consumer-producers, economies of specialization and transaction costs

We consider an economy with M ex ante identical consumer-producers (where M is assumed large). There is a single consumer good \( z \), the amount of this good that an individual self-provides is denoted by \( z^s \) and the amounts sold and purchased are \( z^d \) and \( z^d \), respectively. A fraction \( 1-k \) of \( z^d \) disappears in transit because of transaction costs. Hence, the quantity that a person receives from the purchase of \( z^d \) is \( k z^d \). Total consumption is thus \( z + k z^d \). \( 1-k \) is referred to as the transaction cost coefficient for food and \( k \) is referred to as the transaction efficiency coefficient or simply transaction efficiency for the consumer good. As discussed in Section 2, the transaction cost can be considered as the measurement cost of quantity and quality of goods or anticipated distortion caused by moral hazard or other opportunistic behavior. An alternative formulation is to introduce an explicit time constraint on the consumer-producers in such a way that producing the good takes a certain amount of time but purchasing it takes more or less time\(^6\). The individual maximizes the objective function \( u \) (which can be interpreted as utility) given by total consumption:

\[
    u = z + k z^d
\]  

(1)

The consumer good is produced using labor \( L_z \) and an intermediate good \( x \), i.e.

\[
    z + z^s = (x + r x^d)^\alpha L_z z^{(1-\alpha)\alpha}, \quad \alpha \in (0, 1) \text{ and } \alpha > 1
\]  

(2)

where \( x \) and \( x^d \) are the respective quantities of the intermediate good self-provided and purchased, \( r \) is the transaction efficiency of the intermediate good, \( 1-r \) is the transaction cost coefficient for a unit of the intermediate good purchased and, therefore, \( r x^d \) is the quantity received by a person who buys \( x^d \). The parameter \( \alpha \) is the elasticity of output with respect to input of the intermediate good and can be interpreted as the relative importance of the

\(^6\) The alternative formulation was suggested by the Editor which we gratefully acknowledge. Our specification of iceberg transaction technology avoids formidable indices of origins and destinations of deliveries and other technical difficulty.
roundabout productive sector compared to labor. The degree of economies of specialization is represented by $a$.

Production of the intermediate good requires labor $L_x$ and a primary good $y$.

$$x + x^s = (y + ty^d)^a L_x (1-a)^a, \quad \alpha \in (0, 1) \text{ and } a > 1$$

where $x^s$ is the quantity of the intermediate good sold and $y$ and $y^d$ are, respectively, the quantities of the primary good self-provided and purchased. $1-t$ is the transaction cost coefficient for the primary good ($t$ is the transaction efficiency) and $ty^d$ is the amount that a person receives from the purchase of $y^d$. It is also assumed that fraction $1-s$ of goods produced by an employee within a firm disappears when it is delivered to the employer because of transaction cost of labor. The production function for the primary good is

$$y_s + y^s = L_y^a, \quad a > 1$$

where $y^s$ is the quantity of the primary good sold. The labor endowment constraint for each individual is

$$L_z + L_x + L_y = 1, \quad L_i \in [0, 1], \quad i = z, x, y.$$  

where $L_i$ is a person’s level of specialization in producing good $i$.

This system of production exhibits economies of specialization. The total factor productivity of the consumer good and the labor productivity of the primary good increase with the level of specialization in producing a good concerned. But economies of specialization are individual specific and do not extend beyond the scale of an individual’s working time. Also, they are related to diseconomies of an individual’s scope of activities.

There are $2^9=512$ combinations of zero and non-zero values of $z, z^s, z^d, x, x^s, x^d, y, y^s$ and $y^d$. Following the approach of Borland and Yang (1995), we can exclude many combinations of zero and nonzero variables from the list of candidates for an individual’s optimal decision.

4. Configurations, market structures and general equilibrium

A combination of zero and nonzero variables that is compatible with the Kuhn–Tucker condition and other conditions for the optimum decision is called a configuration. For each configuration, an individual can solve for a corner solution for a given set of relative prices of traded goods. A combination of configurations that satisfies the market clearing conditions is called a market structure, or simply structure. There are eight market structures if the institution of the firm is allowed. We define the firm as contractual arrangements which involve asymmetric relationship between an employer and his employees and the production of intermediate goods using labor. The employer has control rights of his employees’ labor and claims the difference between the firm’s revenue and wage bill.

The eight market structures include autarky, two structures involving partial division of labor with no firm and two with firm, one structure involving complete division of
labor with no firm and two with firms. The eight market structures and the combination of configurations that make up these structures are categorized and characterized as follows:

**Autarky**

1. **Structure A**
   Structure A consists of a configuration with \( z^s = x^s = y^s = z^d = x^d = y^d = 0 \). In this structure each individual self-provides all goods he needs.

**Partial division of labor with no firm**

2. **Structure P1**
   Configuration 1: an individual sells the consumer good, self-provides the intermediate good, and buys the primary good
   Configuration 2: an individual sells the primary good, and buys the consumer good

3. **Structure P2**
   Configuration 1: an individual sells the intermediate good, self-provides the primary good, and buys the consumer good
   Configuration 2: an individual sells the consumer good, and buys the intermediate good

**Partial division of labor with firm**

4. **Structure PF1**
   Configuration 1: an individual sells the consumer good, self-provides the intermediate good, buys labor, and directs workers to produce the primary good in his firm
   Configuration 2: an individual sells labor which is hired to produce the primary good by a firm, and buys the consumer good

5. **Structure PF2**
   Configuration 1: an individual sells labor which is hired by a firm to produce the intermediate good and the primary good in the firm, and buys the consumer good
   Configuration 2: an individual sells the consumer good, buys labor, and directs each worker to produce both the primary and the intermediate goods.

**Complete division of labor with no firm**

6. **Structure C**
   Configuration 1: an individual sells the consumer good, and buys the intermediate good
   Configuration 2: an individual sells the primary good, and buys the consumer good
   Configuration 3: an individual sells the intermediate good, and buys the primary good and the consumer good

**Complete division of labor with firm**

7. **Structure CF1**
   Configuration 1: an individual sells the consumer good, buys the primary good, hires workers and directs them to produce the intermediate good using the primary good in his firm
   Configuration 2: an individual sells the primary good, and buys the consumer good
   Configuration 3: an individual sells labor which is hired by a firm to produce the intermediate good, and buys the consumer good

8. **Structure CF2**
   Configuration 1: an individual sells the consumer good, buys labor, and directs some of the workers to produce the primary good and others to produce the intermediate good using the primary good
Configuration 2: an individual sells labor which is hired by a firm to produce the intermediate good, and buys the consumer good.

Configuration 3: an individual sells labor which is hired by a firm to produce the primary good, and buys the consumer good.

A corner equilibrium for a certain structure is defined by a set of relative prices of traded goods and a set of relative numbers of individuals choosing different configurations that equilibrate total corner-demand to total corner-supply of each traded goods and equalize all individuals’ consumption levels. Here, competition among ex ante identical consumption-maximizing individuals, combined with free entry into each configuration, implies that all configurations in a structure will be chosen only if consumption is equalized across the configurations. In this section, the corner equilibria for the eight structures are first solved.

A two-step approach developed in Yang and Ng (1993) is used to solve for the corner equilibrium for each of the eight market structures. As an illustration, the algebra for solving for the corner equilibrium for structure PF1 is shown as follows.

Solving for the corner equilibrium for structure PF1

Structure PF1 consists of configurations 1 and 2 as summarised earlier. The individual decision problem for configuration 1 is as follows:

\[
\text{Max} : \quad u_z = z \quad \text{(objective function)}
\]

\[
s.t. \quad z + z^s = x^a L_z^{(1-a)}, \quad x = (s y^d)^a L_x^{(1-a)} \quad \text{(production function)}
\]

\[
y^d = n y^s \quad \text{(material balance in a firm)}
\]

\[
L_x + L_z = 1 \quad \text{(endowment constraint for the employer’s labor)}
\]

\[
y^s = L_y^a \quad \text{(production function for an employee)}
\]

\[
L_y = 1 \quad \text{(endowment constraint for an employee)}
\]

\[
p z^s = wn \quad \text{(budget constraint for the firm)}
\]

where \(L_x\) and \(L_z\) are the employer’s labor allocation to the production of \(x\) and \(z\), respectively, \(n\) is the number of workers hired by an employer, \(L_y\) employee’s labor which is subject to the employer’s control, \(y^s\) each employee’s output of \(y\) within the firm, \(y^d\) the total output level of \(y\) in the firm or input requirement for the production of \(x\) in the firm, \(w\) is the wage rate, and \(s\) is the transaction efficiency for labor or \(1-s\) is the transaction cost coefficient of labor in terms of the loss of goods produced out of labor in transit. If \(1-s\) is assumed to be in terms of the loss of labor in transit, the essence of the results will not change, but the algebra will be more complicated. The solution of the problem is

\[
L_x = \frac{\alpha}{(1 + \alpha)}, \quad L_z = \frac{1}{(1 + \alpha)}, \quad L_y = 1,
\]

\[
n = \left( \frac{\alpha^2 A L_z W}{W} \right)^{1/(1-\alpha)(1+\alpha)}, \quad A \equiv [s^a \alpha^{(1-\alpha)a}]^a (1 + \alpha)^{-(1-\alpha)(1+\alpha)a}
\]
\[
\frac{z^b}{p_z} = \frac{wn}{p_z}, \quad u_z = (1 - \alpha^2)A \left( \frac{\alpha^2Ap_z}{w} \right)^{\alpha^2/(1-\alpha^2)}
\]

where \(n\) is demand for labor by the firm. \(z^b\) is supply of \(z\) by the firm, and \(u_z\) is the indirect objective (utility) function for configuration 1. The decision for configuration 2 is fixed as follows

\[
u_y = k\xi^d = \frac{kw}{p_z} \quad \text{(objective function)}
\]

\[wL_y = p_z\xi^d, \quad L_y = 1 \quad \text{(budget and endowment constraints)}
\]

where \(u_y\) is the indirect objective function (utility) of an employee.

The consumption equalization condition \(u_z = u_y\) yields the corner equilibrium relative price \(w/p_z\) and the market clearing condition for good \(z\), \(M_z\xi^d = M_y\xi^d\), yields the relative number of individuals choosing the two configurations. Here, \(M_y/M_z = n\) is the corner equilibrium relative number of individuals choosing to be employees to those choosing to be employers, which happens to be the same as the corner equilibrium value of \(n\), a decision variable for an employer. The solution for the corner equilibrium is summarized in Table 2.

Following this procedure, the corner equilibria in all market structures can be solved. Table 2 summarises these solutions. A general equilibrium is defined as a fixed point that satisfies the following two conditions: (i) each individual uses inframarginal analysis to maximize his consumption respect to configurations and quantities of each good produced, consumed, and traded for a given set of relative prices of traded goods and a given set of the numbers of individuals selling different goods; and (ii) the set of relative prices of traded goods and the set of individuals selling different goods clear the markets for traded goods and equalize consumption for all individuals selling different goods. (Yang and Ng, 1993, Chapter 6) have proved that in this kind of models, the general equilibrium is the corner equilibrium that generates the highest per capital real income.\(^7\) The other corner equilibria are not general equilibria.

A revealed preference argument can be used to show that each corner equilibrium is locally Pareto optimal for a given structure (see Yang and Ng, 1993). Hence, the general equilibrium is globally Pareto optimal. This implies that the competitive market can efficiently co-ordinate division of labor to fully utilize network division of labor. This result critically depends on the assumption that economies of specialization are individual specific and different from economies of scale.

5. Development in division of labor and average employment

A comparison between per capita real incomes generated by the eight market structures yields

**Proposition 1.** As transaction efficiency is improved, the general equilibrium shifts from structure A first to structure P1 or PF1, followed by P2 or PF2, and finally to structure C.

\(^7\) The proof of this statement is available from the authors.
Table 2
Solutions to corner equilibria of market structures$^a$

<table>
<thead>
<tr>
<th>Market structure</th>
<th>Relative price</th>
<th>Labor allocation</th>
<th>Real per capita income</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>$L_1=\alpha(1-\alpha)$</td>
<td>$u(A)=\beta^u(1-\alpha)a$</td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>$L_2=\alpha^2L_2=1-\alpha$</td>
<td>$u(P1)=(krt)\nu a^\rho[2\nu(1-\alpha)^u]$</td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>number of sellers of primary good</td>
<td>$<a href="1-%5Calpha">(1-\alpha)(1+\alpha)^u</a>$</td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>relative number of sellers of consumer good</td>
<td>$=\alpha^u(1-\alpha)$</td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>$P_2$ = $\alpha^u(1-\alpha)$</td>
<td>$u(P2) = (krt)^u \beta^{1+ua}$</td>
</tr>
<tr>
<td>P1</td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>$P_1$ = $\alpha^u(1-\alpha)$</td>
<td>$u(PF1) = (krt)^u \beta^{1+ua}$</td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>number of workers hired by a firm</td>
<td>$=\alpha^u(1-\alpha)$</td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>$u(PF2) = (krt)^u \beta^{1+ua}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>number of Workers hired by a firm</td>
<td>$=\alpha^u(1-\alpha)$</td>
</tr>
<tr>
<td>CF1</td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>$u(CF1) = (krt)^u \beta^{1+ua}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>number of workers hired by a firm</td>
<td>$=\alpha^u(1-\alpha)$</td>
</tr>
<tr>
<td></td>
<td>$-\alpha^u(1-\alpha)$</td>
<td>$u(CF2) = (krt)^u \beta^{1+ua}$</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Note $\beta \equiv \alpha^u(1-\alpha)^{1-ua}$.

CF1, or CF2. This process increases per capita real income, productivity, trade dependence, individuals’ level of specialization, and the number of specialized activities.

(2) The institution of the firm will emerge from the development in division of labor if transaction efficiency is higher for labor than for intermediate goods.

(3) Given the emergence of the institution of the firm from the development in division of labor, the average employment increases if the transaction efficiency for labor is higher than that for intermediate goods. It decreases if transaction efficiency is improved in such a way that the transaction efficiency for intermediate goods becomes higher than that for labor.

Proposition 1 is proved in the Appendix. The intuition of Proposition 1 is as follows. There is a trade-off between economies of specialization and transaction costs. Individuals will choose a low level of specialization and a large scope of production activities if transaction
efficiency is low because economies of specialization are outweighed by transaction costs generated by a high level of division of labor. As transaction efficiency is improved, the equilibrium level of specialization increases and the equilibrium scope of each individual’s production activities becomes narrower. As the level of division of labor increases with improvements in transaction efficiency, two types of market structure can be used to organize a higher level of division of labor. The first involves markets for final and intermediate goods and the second involves markets for final goods and for labor. If the transaction efficiency for labor hired to produce an intermediate good is higher than that for the intermediate good, the institution of the firm will emerge from division of labor. If the institution of the firm emerges, but the transaction efficiency for intermediate goods is improved more quickly than that for labor, then the equilibrium size and scope of firms will decrease as each individual increases his level of specialization and narrows down his scope of production activities.

The decrease in the average employment is driven by two forces. First, each firm becomes more specialized if division of labor between firms and individuals develops, so that each firm’s scope of activities is reduced. Second, if the transaction efficiency is improved in such a way that the transaction efficiency for intermediate goods becomes higher than that for labor, then organizing division of labor between more specialized firms and individuals will be more efficient than organizing division of labor within a firm since the former involves more transactions of intermediate goods while the latter involves more trade in labor. Hence, an increase in the level of specialization of firms and an increase in division of labor between small scale firms and individuals will occur, generating a decline in average employment of firms.

To illustrate the shifting of market structures. Suppose the initial values of parameters are such that the transaction efficiencies for final and intermediate goods are very low but the transaction efficiency for the intermediate good is higher than for labor. Then the general equilibrium will be autarky, as shown in the Appendix. If the transaction efficiencies for final and intermediate goods are improved while the transaction efficiency for labor is still lower than for intermediate goods, then the general equilibrium will be structure P1 where firms and labor market do not exist and completely specialized producers of primary good exchange it for final good with non-specialized producers of intermediate and final goods in the market. Suppose the transaction efficiencies for final goods and labor are improved relative to the transaction efficiency of the intermediate good, then the general equilibrium will shift to structure PF2 where a specialist producer of final good hires $k\alpha/(1-\alpha)$ workers and directs each of them to produce primary and intermediate goods within a firm. In PF2, each firm produces three goods and the level of specialization of workers is not high although it is higher than in autarky. If the transaction efficiency for primary good is sufficiently improved such that it is higher than the transaction efficiency for labor which in turn is higher than that for the intermediate good, then the general equilibrium will shift from PF2 to structure CF1 where a specialist producer of final good hires $ak$ workers, buys the primary good, and directs the workers to specialize in producing the intermediate good using the primary good within his firm.

For this particular exogenous changing pattern of values of parameters, the general equilibrium will shift along the path represented by the structure sequence $A \Rightarrow P1 \Rightarrow PF2 \Rightarrow CF1$ in Fig. 2.
The development of specialization in output associated with a decrease in the scope of a firm’s activities, an increase in specialization of labor assignment within firms and replacement of self-provision of intermediate goods with market procurement, are common in the economy. For instance, law firms may give up providing a full range of legal services to specialize in only one or two aspects of the legal practice such as conveyancing or family law. Specialization in labor assignment is also common in manufacturing assembling and processing. A manufacturing firm may purchase intermediate services such as advertising, storage and distribution in the market instead of advertising its own products and operating its own warehouses and fleet of trucks.

To summarise, all patterns of changes in division of labor that involve a decrease in average employment of firms are listed in (6a), and those patterns that involve a monotonic increase in average employment are listed in (6b). The values of \( S \) are the average employment of firms in the relevant structures.

\[
A \Rightarrow P1 \Rightarrow PF2 \Rightarrow \begin{cases} 
C \\
CPF1 \\
PF2 \\
CF2 
\end{cases} \quad S = \alpha k / (1 - \alpha) \quad (6a)
\]

\[
A \Rightarrow PF1 \Rightarrow PF2 \Rightarrow \begin{cases} 
C \\
CPF1 \\
PF2 \\
CF2 
\end{cases} \quad S = \alpha k / (1 - \alpha^2) \quad (6b)
\]

where the average employment of firms is \( \alpha k / (1 - \alpha) \) in structure PF2, \( \alpha k \) in CF1, \( \alpha^2 k / (1 - \alpha^2) \) in PF1, and \( \alpha k / (1 - \alpha) \) in CF2, and \( \alpha k / (1 - \alpha) > \alpha^2 k / (1 - \alpha^2) \), \( \alpha k \). (6) is proved in the Appendix.
6. Conclusions

In this paper we use inframarginal comparative static equilibrium analysis to show, after the manner of Coase and North, a decline in the transaction cost coefficient leads to organizational changes from self-sufficient production to specialization in production\(^8\). If we assume that each individual maximizes total discounted utility and transaction efficiency changes over time, our model in this paper will become a dynamic general equilibrium model which can predict evolution of firm size over time. Another way to extend the static model is to assume bounded rationality and interactions between dynamic decisions and information of prices. Zhao (1998) has developed such a Walrasian sequential equilibrium model that can predict spontaneous evolution of the institution of the firm and of society’s knowledge of institution.

Our model predicts that the average employment of firms will decrease as division of labor develop if transaction efficiency is improved in such a way that the transaction efficiency for intermediate goods becomes higher than that for labor. To test the model one needs to have good measures of the relative transaction conditions of goods to labor. However, these data are difficult to obtain. Nevertheless, empirical work of Murakami et al. (1996), and Yang et al. (1992) have provided indirect evidence for this theory. Murakami, Liu, and Otsuka show that in China’s machine tool industry the average size of firms declined while division of labor and productivity rose in the 1980s. Yang, Wang, and Wills show that an institution and policy index (a weighted average of 12 subindices of transaction conditions for four types of properties and three components of property rights) that affects transaction efficiency in China increased in the 1980s. The subindex for transaction efficiency of goods was higher than that of labor during that period. These two pieces of empirical work indirectly support the hypothesis based on our theory. More empirical work needs be done to test our theory.

Appendix A. Proof of Proposition 1

Since the corner equilibrium that generates the highest per capita real income is the general equilibrium, the general equilibrium can be identified by comparisons among per capita real incomes in all structures.

A comparison between \(u(P1)\) and \(u(PF1)\), or between \(u(P2)\) and \(u(PF2)\) yields

\[
u(PF1) > u(P1) \text{ if } s > t \text{ and } u(PF2) > u(P2) \text{ if } s > r.\]

(7)

This implies that for a given pattern of partial division of labor, the institution of the firm will be used in the general equilibrium to organize the division of labor if the transaction efficiency for labor \(s\) is higher than the transaction efficiency for intermediate goods \((r\) or \(t\)). Comparisons among \(u(C)\), \(u(CF1)\) and \(u(CF2)\) yield the following results.

\[
u(C) > u(CF1) \text{ and } u(CF2) \text{ if } s < \min\{r, (rt)^{1/(1+\alpha)}\}\]

(8a)

\(^8\) Development in division of labor in this paper is based on comparative statics. It is exogenous and cannot take place in the absence of exogenous change of transaction efficiency parameters. A dynamic version of the model is developed by Borland and Yang (1995) to generate endogenous evolution of the firm.
\( u(CF1) > u(C) \) and \( u(CF2) \) if \( s \in (r, t) \) \& \( t > r \). \tag{8b} \\
\( u(CF2) > u(C) \) and \( u(CF1) \) if \( s > \text{Max}(t, (rt^{a})^{1/(1+a)}) \) \tag{8c}

(8) together with Table 1 imply that the larger the transaction efficiency for labor compared to that for intermediate goods, the more likely the general equilibrium is a structure involving the institution of the firm, and that the larger the transaction efficiency for labor compared with that for intermediate goods, the larger is the employment of a firm.

Since the general equilibrium, which is the corner equilibrium with the highest per capita real income, will shift between corner equilibria as transaction efficiency parameters reach some critical values, there are many possible paths of such development in the division of labor, indicated by the arrows with numbers in Fig. 2. The critical values can be identified by comparing per capita real incomes in all structures. Such comparisons yield the following results.

Shift 1, \( A \Rightarrow P1 \), will take place if \( kt \) increases to the critical value \( kt > \beta_0 \)

Shift 2, \( A \Rightarrow PF1 \), will take place if
\( ks \) increases to the critical value \( ks > \beta_0 \)

Shift 3, \( P1 \Rightarrow P2 \), will take place if \( k \) and 
\( r \) increases to the critical value \( 
\frac{k^{(1-a)r}}{r^{a}} > \beta_1
\)

Shift 4, \( P1 \Rightarrow PF2 \), will take place if \( k \) and 
\( s \) increases to the critical value \( 
\frac{k^{(1-a)s}}{s^{a}} > \beta_1
\)

Shift 5, \( PF1 \Rightarrow P2 \), will take place if \( k \) and 
\( r \) increases to the critical value \( 
\frac{k^{(1-a)r}}{s^{a}} > \beta_1
\)

Shift 6, \( PF1 \Rightarrow PF2 \), will take place if \( ks \) increases to the critical value \( (ks)^{1-a} \) \( \beta_1 \)

Shift 7, \( P2 \Rightarrow C \), will take place if \( k \) increases to the critical value \( 
\frac{t}{r^{(1-a)}} > \beta_2
\)

Shift 8, \( P2 \Rightarrow CF1 \), will take place if \( k \) and \( t \) increases to the critical value \( 
\frac{st^{a}}{r} > \beta_2
\)
Shift 9, $\text{P2} \Rightarrow \text{CF2}$, will take place if $s$ and $r$ increases to the critical value $s^{1+\alpha} \frac{1}{r} > \beta_2$

Shift 10, $\text{PF2} \Rightarrow \text{C}$, will take place if $r$ and $t$ increases to the critical value $r \frac{t^\alpha}{s} > \beta_2$

Shift 11, $\text{PF2} \Rightarrow \text{CF1}$, will take place if $t$ increases to the critical value $t^\alpha > \beta_2$

Shift 12, $\text{PF2} \Rightarrow \text{CF2}$, will take place if $s$ increases to the critical value $s^\alpha > \beta_2$

where $\beta_0 = \alpha^{-2\alpha^2}(1-\alpha)(1-\alpha^2)(\alpha-1)$, $\beta_1 = [\alpha^{(1-2\alpha^2)}(1-\alpha)\alpha^{-1}]^{1-1}(1+\alpha)^{(1-\alpha^2)(\alpha+1)/\alpha}$, and $\beta_2 = [\alpha^{\alpha}(1-\alpha)^{1-\alpha}]^{a-1}$.

It can be shown that the size and scope of firms will decrease if the development in division of labor follows the structure sequences: $A \Rightarrow \text{P1} \Rightarrow \text{PF2} \Rightarrow \text{C}$ (Paths 1, 4, 10 in Fig. 2), $A \Rightarrow \text{P1} \Rightarrow \text{PF2} \Rightarrow \text{CF1}$ (Paths 1, 4, 11 in Fig. 2), $A \Rightarrow \text{PF1} \Rightarrow \text{PF2} \Rightarrow \text{C}$ (Paths 1, 6, 10 in Fig. 2), and $A \Rightarrow \text{PF1} \Rightarrow \text{PF2} \Rightarrow \text{CF1}$ (Paths 2, 6, 11 in Fig. 2). According to (7–9), this will take place if transaction efficiencies for goods and labor increase but the transaction efficiency for labor does not increase too fast.

On the other hand, the size of firms will increase if the development in division of labor follows the structure sequences $A \Rightarrow \text{P1} \Rightarrow \text{P2} \Rightarrow \text{CF1}$ (paths 1, 3, 8 in Fig. 2), $A \Rightarrow \text{P1} \Rightarrow \text{P2} \Rightarrow \text{CF2}$ (paths 1, 3, 9, in Fig. 2), and $A \Rightarrow \text{PF1} \Rightarrow \text{PF2} \Rightarrow \text{CF1}$ (Paths 2, 6, 11 in Fig. 2). According to (7–9) in the Appendix, this will take place if transaction efficiencies for goods and labor increase or the transaction efficiency for labor increases more quickly than the transaction efficiency for intermediate goods. The results are summarized in (6). Fig. 2 or (9) in the Appendix, combined with Table 1 imply that as transaction efficiencies for labor, and intermediate and final goods are improved, the level of division of labor increases. Individuals’ level of specialization, trade dependence, per capita real income, and the number of specialized activities increase. This change will not involve the institution of the firm if the transaction efficiency for labor is lower than that for intermediate goods. The institution of the firm will emerge from the development in division of labor if the transaction efficiency is higher for labor than for intermediate goods.

(6), (7), (8), and (9) are sufficient to establish Proposition 1.

References

Borland, J., Yang, X., 1995. Specialization, product development, evolution of the institution of the firm, and
and Organization 15, 137–149.
Camacho, A., 1996. Division of Labor, Variability, Co-ordination, and the Theory of Firms and Markets, Boston, 
67, 297–308.
Ethier, W., 1982. National and international returns to scale in the modern theory of international trade. American 
1261–1283.
796–815.
73–94.
National Taiwan University, Taipei, (in press).
the Far Eastern Meeting of Econometric Society, Taipei.
70, 950–959.
973.
110, 857–880.
economics 144, 635–657.
81, 965–982.
Liu, P.W., 1992. Extent of the market, specialization, transaction efficiency and firm size in Hong Kong. Working  
Paper No. 20, Department of Economics, Chinese University of Hong Kong.
of Economic Research Discussion Paper No. 1796, Harvard University, USA.
57–89.