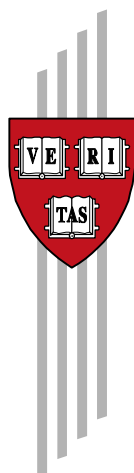


**Incomplete Contingent Labor Contract,
Asymmetric Residual Rights and Authority,
and the Theory of the Firm**

Xiaokai Yang

CID Working Paper No. 45
May 2000

© Copyright 2000 Xiaokai Yang and the President and Fellows of
Harvard College



Working Papers

Center for International Development
at Harvard University

**Incomplete Contingent Labor Contract,
Asymmetric Residual Rights and Authority,
and the Theory of the Firm**

Xiaokai Yang

Abstract

In the paper the trade-offs among endogenous transaction costs caused by two-sided moral hazard, exogenous monitoring cost, and economies of specialization are specified in a Grossman, Hart and Moore (GHM) model to absorb Maskin and Tirole's recent critique and Holmstrom and Milgrom's criticism of the model of incomplete contract. The extended GHM model allowing incomplete contingent labor contract as well complete contingent contract of goods trade is used to explore the implications of structure of ownership and residual rights for the equilibrium network size of division of implications of structure of ownership and residual rights for the equilibrium network size of division of labor and productivity.

Keywords: theory of the firm, incomplete labor contract, asymmetric authority, two-sided moral hazard, transaction cost, asymmetric residual rights, division of labor, specialization

JEL Classification Codes: D23, L11, L22, L23

Xiaokai Yang is a Research Fellow at the Center for International Development at Harvard University and Personal Chair Professor, Department of Economics, Monash University.

I. Introduction

The purpose of the paper is to develop a general equilibrium version of the models of Grossman and Hart (1986) and Hart and Moore (1990) (GHM model). The motivation of the extension is twofold. First, it extends the GHM model to absorb the recent criticisms of the theory of incomplete contract (Hart and Moore, 1999, Maskin and Moore, 1999, Maskin and Tirole, 1999a, b, Segal, 1999, Tirole, 1999, Holmstrom and Milgrom, 1991, and Holmstrom and Robert, 1998). Second, the general equilibrium model with endogenous specialization as well as endogenous structure of property rights allows us to explore the implications of two-sided moral hazard and structure of residual rights for the equilibrium network size of division of labor and productivity progress. Let us motivate the two tasks one by one.

In the GHM models, ex ante indescribability of contingent states creates scope for ex post renegotiation. An incomplete contract that specifies the ownership structure of assets will affect ex post bargaining power of players and therefore generate important implications of structure residual rights for reducing distortions caused by double moral hazard.

Maskin and Tirole (1999a, see Tirole, 1999, p. 760 as well) have shown that with the assumption that the equilibrium monetary transfers between the parties resulting from the renegotiation are the same whenever the states are payoff equivalent and the assumption of strict risk aversion (together with innocuous assumptions), the possibility of renegotiation and the indescribability of contingencies (even combined together) do not restrict the payoffs that can be attained through contracting.¹ In response to the indescribability irrelevance theorem, Hart and Moore (1999) and Segal (1999) show that as the number of contingent states tends to infinity, the optimal complete contract is a null contract.

In the present paper, it will be shown that neither an infinite number of contingent states nor possibility for renegotiation is essential for exploring the implications of structure of residual rights predicted by the GHM model. I follow the idea of Cheung (1983) and Coase (1937) that the difference in transaction costs between labor and goods is enough for generating the implications of structure of residual rights in the GHM model. Our story runs as follows.

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. Each individual can choose self-provision of the two goods to avoid moral hazard caused by transactions. But autarkic production implies a low productivity since economies of specialization are not fully exploited. If individuals choose the division of labor between specialist producers of two goods, necessary transactions involve two-sided moral hazard. Unobservable effort of each type of specialist in reducing a transaction risk affects not only this person's probability to receive good that she buys, but also the probability for her trade partner receiving good purchased. It is assumed that there is the trade-off between measurement cost of goods and labor and moral hazard caused by imprecise measurement (Holmstrom and Milgrom, 1991). The measurement cost coefficient is different between two goods and labor employed to produce different goods.

There is a tradeoff between economies of specialization and transaction costs (including moral hazard and measurement cost). If measurement efficiency is high, then individuals have a greater scope for trading off among economies of specialization, moral hazard, and measurement costs, thereby division of labor occurring at equilibrium. Otherwise, autarky, shown in Fig. 1(a) where ellipse with A denotes a person choosing autarky, x is self-provided quantity of management service and y is self-provided quantity of cloth, will be chosen as the equilibrium.

¹ Maskin and Xu (1999) provide a survey of commitment game models (sequential equilibrium models with information asymmetry and possible renegotiation) which provide a solid theoretical

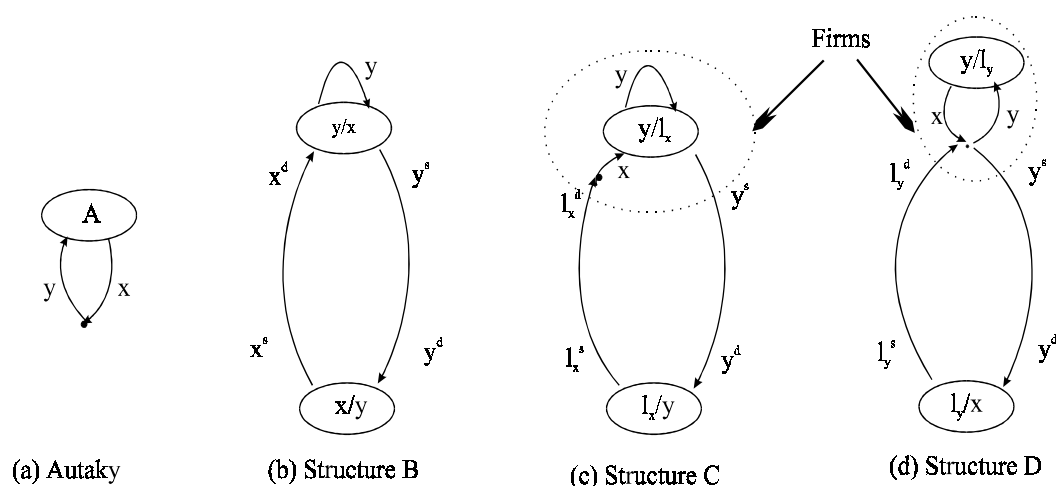


Figure 1: Various Structures of Transactions and Residual Rights

There are three different structures of residual rights, which can be used to organize transactions required by the division of labor. Structure B, shown in Fig. 1(b) where the ellipse with y/x represents a person selling cloth and buying management services and the ellipse with x/y represents a person selling management service and buying cloth and arrowed curves represent goods flows, 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 C, shown in Fig. 1(c) where the ellipse with y/l_x represents the employer who hires labor to produce management service and then uses it to produce cloth and ellipse with l_x/y represents the employee selling labor and buying cloth, comprises the market for cloth and the market for labor hired to produce the management service within a firm. The dashed ellipse represents the firm, the dot represents a process transforming labor to management service within the firm, and arrowed curves represent flows of goods and labor. 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 decision rights over her employees' labor, and sells goods that are produced from employees' labor. Structure D, shown in Fig. 1(d), 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 B, 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 C and D involve a firm and an asymmetric structure of residual rights, they entail different firm ownership structures.

Suppose that measurement efficiency is much lower 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. Hence, risk for delivery failure increases even for the same effort level in reducing the transaction risk. 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 measurement efficiency for labor hired to produce management services is much lower 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?) Therefore, the efficient balance of the trade-off between moral hazard and monitoring cost generates more serious moral

hazard in labor trade in structure C than in structure D. This is because structure D involves trade in cloth and in labor hired to produce cloth, but not trade in management services nor in labor hired to produce management services, while structure C involves trade in cloth and in labor hired to produce management services. Then the division of labor can be more efficiently organized in structure D than in structure C. Hence, structure D will occur at equilibrium if the measurement efficiencies for labor hired to produce cloth and for cloth are sufficiently high. 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 measurement 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 that 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.²

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 is 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

² According to Mokyr (1990, p. 268, 1993), the legal protection of the residual rights to the firm were more important than patent laws for technical and managerial inventions and innovations in

the precise price of the idea and the entrepreneur gets what she deserves, so that stealing and over-or underpricing 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 a billion dollars (which is the real market value of the ideas indicated by Bill Gates residual returns) to buy the ideas?³

In this story, infinite contingent states, indescribability, ex post renegotiation are not essential. Incomplete labor contract and associated institution of the firm can be defined as follows. The employer has exclusive rights to a specific asset: the name of the firm via name search in the registration process of the company. Labor contract between her and employees does not specify what the employer receives from the relationship, but specifies what the employee receives. Hence, the employer claims the residual return which is the difference between revenue and wage bill. This residual right also implies that whatever produced by the employees within the firm is owned by the employer. In addition the employer has decision rights to allocating employees' labor. This structure of asymmetric authority implies that the employee must do whatever she is told to do. This is consistent with Hart's idea about asymmetric control rights, but different from own consumption of asset which is considered by Maskin and Tirole (1999b, p. 139) as the essential feature of asset ownership in the GHM model. The function of the asymmetric decision rights is to avoid directly pricing the goods with the high measurement cost, which is different from the function of asymmetric authority in the literature of complete contract (Aghion and Tirole, 1997). Such labor contract can be considered as incomplete contract since it requires the employee to do

British industrialization.

³ If more layers of principal-agent relationships are introduced, the division of labor between production management and portfolio management may be associated with shareholders being the

whatever she is asked to do rather than specifying what exactly the employee must do. However, the labor contract has a feature of complete contingent contract: the employee receives a higher wage when the business goes well than when business is bad.

Hence, our model is to formalize Cheung and Coase's theory of the firm that emphasizes the implications of the difference in measurement efficiency among different goods and factors rather than Williamson's (1975) idea about holding up and related opportunism. Our model absorbs Holmstrom and Roberts (1998) and Holmstrom and Milgrom's (1991, 1995) criticism of the model of incomplete contract that labor contract is inessential in the GHM model. Also, their idea on the trade-off between moral hazard and monitoring or measurement cost is absorbed.

In the GHM models, the pattern of division of labor between players is exogenously given, so that the implications of structure of residual rights for the network size of division of labor cannot be explored. In a general equilibrium environment, individuals' decisions in choosing their levels of specialization are dependent on endogenous transaction costs caused by moral hazard and on structure of residual rights, while the level of the endogenous transaction cost itself is determined by individuals' levels of specialization. For instance, if all individuals choose autarky, endogenous transaction cost is zero because of the absence of transactions in autarky. Hence, we need a general equilibrium model to figure out a mechanism that simultaneously determines the interdependent level of endogenous transaction cost and network size of division of labor.

The simultaneous endogenization of transaction cost and network size of division of labor can be motivated by Houthakker and George Stigler's following statements. Houthakker claims (1956): "there is hardly any part of economics that would not be advanced by a further analysis of specialization." This claim implies

owners of the firm. A recent survey of the literature of multiple layers of principal-agent relationships can be found from Gibbons (1998).

that studies of specialization is not merely a field of economics, but rather it is the core of mainstream economics. However, Stigler noted (1976):

“... almost no one used or now uses the theory of division of labor, for the excellent reason that there is scarcely such a theory. ... there is no standard, operable theory to describe what Smith argued to be the mainspring of economic progress. Smith gave the division of labor an immensely convincing presentation – it seems to me as persuasive a case for the power of specialization today as it appeared to Smith. Yet there is no evidence, so far as I know, of any serious advance in the theory of the subject since his time, and specialization is not an integral part of the modern theory of production.”

Since the late 1970s, many works have been developed to fill the gap noted by Stigler (see a survey on the literature on specialization and division of labor by Yang and S. Ng, 1998 and references there). But in the literature, transaction costs are exogenous. This paper will apply inframarginal analysis developed in the literature of endogenous specialization and in the literature of incomplete contract to simultaneously endogenize individuals' levels of specialization and transaction costs.

As Allyn Young (1928) noted, economies of division of labor are typical network effect rather than economies of scale. Hence, partial equilibrium model is not appropriate for an investigation of equilibrium network size of division of labor. In order to explore the implications of endogenous transaction costs for the equilibrium network size of division of labor and related productivity, we need to develop a general equilibrium version of the GHM model.⁴ The extended GHM

⁴ Previous studies of general equilibrium models of moral hazard can be found from, for instance, Helpman and Laffont (1975), Kihlstrom and Laffont (1979), and Legros and Newman (1996), which do not focus on implications of structure of residual rights for the equilibrium network size of division of labor.

model predicts Pareto improving increases in moral hazard as a result of improved monitoring efficiency, which enlarges the scope for trading off one against the others among positive network effect of the division of labor on aggregate productivity, moral hazard, and monitoring costs. Hence, the general equilibrium discontinuously jumps from autarky where there is no transaction and related moral hazard to the division of labor where aggregate productivity and welfare are higher and moral hazard occurs. The Pareto improving increase in moral hazard provides a rationale for concurrent increases in per capita real income and income share of transaction cost, observed by North (1985). The model in this paper also formalizes Cheung (1983), and Coase (1937), and Stigler's (1953) theory of irrelevance of the size of the firm, which is supported by empirical evidences provided by Liu and Yang (2000) and Murakami, Liu, and Otsuka (1996). This theory states that the average size of firms increases if division of labor develops within firms and it decreases if division of labor develops among firms.

Section II specifies a general equilibrium model with endogenous specialization and endogenous structure of residual rights. All possible structures of transactions and residual rights are enumerated and local equilibrium in each structure is solved in section III. Then the general equilibrium and its comparative statics are solved in section IV. The final section summarizes the findings of the paper.

II. A General Equilibrium Version of the GHM Model

Consider a model with a continuum of *ex ante* identical consumer-producers of mass M , consumption good y and intermediate good x , which is used to produce y . Each person's utility function is:

$$(1) \quad u = y + k_y y^d$$

where y is the amount of the final good self-provided, y^d is the amount of the good purchased from the market. The fraction $1-k_i \in (0, 1)$ of good i purchased disappears in transit because of transaction risk. Hence, k_i is the fraction of one unit of good i that is received by the buyer. Or k_i can be considered as a trading efficiency coefficient of good i . k_i is a contingent variable because of transaction risk. In the absence of labor trade, the trading efficiency coefficient for good i is:

$$(2a) \quad k_i = \begin{cases} \theta_i & \text{with probability } s_x s_y \\ 0 & \text{with probability } 1 - s_x s_y \end{cases}$$

where s_i is the effort level of a buyer of good i in avoiding transaction risk and $i=x,y$. (2a) implies that the trading efficiency for an exchange between x and y is determined by two trade partners' efforts in avoiding the transaction risk. Hence, each partner's effort not only affects her trading efficiency, but also affects the other party's trading efficiency. But if the effort is not observable or not verifiable, then each player may ignore the effect of her effort on the other party, so that two-sided moral hazard arises.

θ_i can be considered as negatively dependent on exogenous transaction cost in measuring quantity and quality of goods i . In other words, for the same effort levels s_x and s_y , as the difficulty in measuring quantity and quality of good i decreases, the trading efficiency coefficient $k_i = \theta_i$, which occurs with probability $s_x s_y$, increases. This is a simple way to formalize the trade-off between moral hazard and measurement cost. As the measurement cost decreases, moral hazard may decrease even if the effort level for avoiding transaction risk is fixed. A more complicated way to specify the trade-off is to specify labor cost in measuring goods. Other sophisticated ways to specify the trade-off between moral hazard and monitoring cost can be found from, for instance, Holmstrom and Milgrom (1991, 1995), Milgrom and Roberts (1992), and Aghion and Tirole (1997).

Because of a great structure variety of transactions and the distinction between ex ante and ex post production conditions (drawn by Rosen, 1978, see also, Yang,

2000, ch. 8) in the model of endogenous specialization, structure of production functions depends on structure of transactions. Hence, I first specify the production conditions in the absence of labor market. The ex post production conditions after individuals choose a structure with labor market will be specified later on. Each person is equipped with the same production functions for the two goods and the endowment constraint for working time.

$$(3) \quad x^p \equiv x+x^s = \text{Max}\{L_x-\alpha, 0\}, \quad y^p \equiv y+y^s = \text{Max}\{(x+k_x x^d)^\beta (L_y-\alpha), 0\}$$

$$L_x+L_y+s_x+s_y=1,$$

where s_i is the effort level of a buyer of good i exerted in avoiding transaction risk of good i . It is trivial to show that s_i is zero iff $t^d = 0$ or iff the individual does not buy good i . x is the amount of the intermediate good self-provided and x^d is the amount of the intermediate good purchased from the market, and $k_x \in (0, 1)$, a contingent variable, is the trading efficiency coefficient for the intermediate goods. x^s and y^s are respective quantities of the two goods sold. x^p and y^p are respective output levels of the two goods. L_i is an individual's quantity of labor allocated to the production of good i . α is a fixed learning cost in producing each good. The system of production functions and endowment constraint (1b) displays economies of specialization which exist if the labor productivity or total factor productivity of an individual in producing a good increases as her amount of labor allocated to the production of that good increases. Economies of specialization differ from economies of scale for it is associated with diseconomies of an individual's scope of production activities. It is individual specific and does not extend beyond the scale of an individual's working time and it cannot be obtained by simply pooling labor together in the absence of workers' specialization within the firm. Aggregation of individuals' economies of specialization generate economies of division of labor for society as a whole. The difference between economies of division of labor and economies of scale was drawn by Allyn Young (1928) and

Yang (1994).⁵ The difference and connection between individual specific economies of specialization and economies of division of labor for society as a whole is drawn by Yang and Ng (1993) and Yang (1994).

There are $2^6 = 64$ profiles of positive and zero values of decision variables x, x^s, x^d, y, y^s, y^d ($x^p, y^p, L_x, L_y, s_x, s_y$ are not independent of the other six decision variables). Hence, there are 64 possible corner and interior solutions for each individual in making a plan of production, trade, and consumption. According to theorem 1 in Borland and Yang (1995, see also Sun, 1999), for a Walrasian equilibrium model with consumer-producers, economies of specialization, and transaction costs each consumer-producer's optimum decision can be achieved by not simultaneously buying and selling or producing the same good and by selling at most one good. Consider the budget constraint and positive utility constraint in addition to this theorem, each individual needs to consider only three *configurations* of production, trade, and consumption in the absence of labor trade: autarky, denoted by A, which implies $x, y, L_x, L_y > 0, x^s = x^d = y^s = y^d = s_x = s_y = 0$; specialization in x, denoted by (x/y), which implies $x^s, y^d, s_x, L_x > 0, x = x^d = y = y^s = L_y = s_y = 0$; and specialization in y, denoted by (y/x), which implies $y, y^s, x^d, s_y, L_y > 0, y^d = x = x^s = L_x = s_x = 0$. Configurations (x/y) and (y/x) are shown in Fig. 1(b).

Combinations of those configurations that are compatible with the market clearing conditions constitute several *market structures* or *structures* for short. Structure autarky (A), shown in Fig. 1(a), occurs if M individuals choose configuration A. Structure B (division of labor), shown in Fig. 1(b), occurs if the population is divided between configurations (x/y) and (y/x). I assume that measure of individuals selling good i is M_i , which are endogenously determined by the utility equalization condition in a competitive equilibrium process.

⁵ Young (1928, p. 539) stated "The mechanism of increasing returns is not to be discerned adequately by observing the effects of variations in the size of an individual firm or of a particular industry, for the progressive division of labor and specialization of industries is an essential part of the process by

In structure B, complete contingent contracts of goods are used to organize the division of labor and reciprocal and bilateral principal-agent relationships (an x specialist is an agent of y specialist in producing x and is a principal of an y specialist in producing y). Players sign complete contingent contracts and each player chooses her allocation of labor between production and transaction activities to maximize expected utility before a state of trading efficiency coefficient k is realized. After a state of k is realized, players deliver goods according to contracts. Note that with probability $s_x s_y$ the two types of specialists trade x for y at a positive relative price and that with probability $1 - s_x s_y$ the specialist producer of y fails to receive the intermediate good x , so that the production of y cannot take place, that is, each individual receives zero utility. The transaction condition is given in (2a). The assumption that there are only two contingent states of k_i and one of the states generates zero utility significantly simplifies the algebra. In the future research, the assumption may be relaxed to allow for four contingent states of k which is determined by the two specialists' decisions in choosing s_i . For instance, $k = \theta_{hh}$ with probability $s_x s_y$, $k = \theta_{hl}$ with probability $s_x (1 - s_y)$, $k = \theta_{lh}$ with probability $(1 - s_x) s_y$, and $k = \theta_{ll}$ with probability $(1 - s_x)(1 - s_y)$.

In structure C, shown in Fig. 1(c), a specialist producer of y is the owner of a firm producing and selling the final good y and hiring labor to produce good x within the firm. In this structure, the trading efficiency coefficient for the final good and the trading efficiency coefficient for labor employed to produce the intermediate good are

$$(2b) \quad k_y = \begin{cases} \theta_y & \text{with probability } r_x s_y \\ 0 & \text{with probability } 1 - r_x s_y \end{cases} \quad \text{and} \quad t_x = \begin{cases} \mu_x & \text{with probability } r_x s_y \\ 0 & \text{with probability } 1 - r_x s_y \end{cases}$$

respectively, where s_y is the employee's (x specialist's) effort level in reducing transaction risk in buying good y and r_x is the employer's (y specialist's) effort level

which increasing returns are realized." He argued (p. 533) that the use of the notion of large-scale-production misses the phenomenon of economies of division of labor.

in reducing transaction risk in buying labor, μ_x represents monitoring and measuring efficiency of the labor used to produce x . The labor contract between the x specialist and the y specialist is an incomplete contract in the following sense. It does not specify exactly what the employee must produce. Instead, it specifies that the employee must do whatever the employer asks her to do and the employer owns goods produced by her and her employees before selling them in the market. Also, the labor contract specifies contingent wage rates that the employee receives, but does not specify what the employer receives from the relationship. In other words, the employer has control rights to the employee's labor allocation and claims residual returns of the firm. But this incomplete contract with asymmetric authority and asymmetric residual returns is mixed with complete contingent contract in terms of Maskin and Tirole's definition of the term. If business goes well (with probability $s_y r_x$), the employee receives a positive wage. If business is not good (with probability $1-s_y r_x$), the employee receives 0 wage rate. Hence, the labor contract in this structure is a mixture of Hart's incomplete contract which specifies ownership structure of assets and conventional complete contingent contract.

The two features of the incomplete labor contract here deserve particular attention. First, when the ex ante ownership structure of x and y is specified in the incomplete contract, x and y are yet to be produced. This, together with the fact that the employee does not know what exactly she is supposed to produce ex ante, has a flavor of the GHM model with unforeseen contingent states. However, the infinite number of contingent states is not needed for justifying the function of the labor contract in this model.

The second feature relates to the definition of the firm. Here, institution of the firm is defined as a structure of transactions that satisfies the following conditions.

(i) There are two types of trade partners who are associated with a firm: employers and employees. There is an asymmetric distribution of decision rights or

authority. The employer has decision rights to use the employee's labor. "Decision rights" here implies that ultimate rights to use the employee's labor is owned by the employer subject to the employee's freedom to quit the job and to other terms of labor contract. A labor contract is an incomplete contract as discussed above. This incompleteness of labor contract is purposely created in order to protect the employer's rights to intangible intellectual properties and prevent them from being stolen by the employee. When an employer-entrepreneur directs the employee to take some action, the latter actually uses the former's idea as an input to produce some final goods. But different employees just get a very small part of the whole idea. The incompleteness of the labor contract prevents the employee to get the whole idea. Hence, inefficient direct trade and pricing of intangible ideas are avoided. Here, bilateral monopoly between the employer and the employee, discussed by Maskin and Tirole (1999a, b) and Tirole (1999), is not essential. Each employer can turn to other potential employees in labor market and each employee can turn to other potential employers.

(ii) The employer claims the *residual returns* that is the revenue net of wage bill and other cost of the firm. 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 is enforced through the business name search process and recognition of the name in the legal cases in the judicial process. The exclusive rights are also enforced through laws of brand. This has a flavor of Hart and Williamson's notion of ownership of specific asset. However, the ownership of other tangible assets here is not essential as long as all production factors other than intangible intellectual properties (entrepreneurial ideas, reputation, client base, and so on) can be precisely priced (via renting and hiring). Here, monopolistic possession of specific physical assets is inessential. For instance, if an entrepreneur founds a copying shop, she uses the institution of the firm to protect her rights to the good

will of doing business, client base, and her idea to align location of shop, machines, and clients in a right way. But she may not have any ownership of physical assets (copy machines may be hired from another company and real estate may be rented too). But exclusive rights to the business name can be considered as essential ownership of specific assets for the firm.

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

The relationship between a professor and her housekeeper does not involve the institution of the firm although it has features (i) and (ii) since it does not have feature (iii). The professor directly consumes what the housekeeper produces and she does not resell it to the market. If an individual hires a broker to conduct stock exchanges, the relationship involves asymmetric rights to control and to residual returns. But the relationship does not involve the firm since it does not satisfy condition (iii). This definition of institution of the firm is consistent with the spirit of the GHM model that incomplete contract that specifies ex ante ownership structure of goods makes a difference on the one hand. It has absorbed Holmstrom and Milgrom's criticism of the model of incomplete contract by well defining labor contract and institution of the firm, on the other.

In structure C, incomplete contingent labor contracts are used to organize the division of labor. Players sign incomplete labor contracts before a state of trading efficiency coefficient is realized. If state with $k_y = \theta_y$ and $t_x = \mu_x$ is realized, a positive relative price of labor in terms of the consumption good is implemented. If state with $k_y = 0$ or $t_x = 0$ is realized, prices are irrelevant and players' utilities are zero. After the incomplete contingent labor contract is signed, an employee allocates her labor between sale in the labor market and reduction of transaction risk. Then the x specialist (the owner of a firm) directs labor that she receives to the production of good x, and allocates her own labor between production and

transaction activities to maximize her expected utility. After states of k_y and t_x are realized, players deliver goods according to the contracts. In this economy with a continuum of individuals, we may assume that in an equilibrium the fraction $s_y r_x$ of total population end up with a positive wage in terms of y and the rest of population end up with 0 utility.

Compared with structure B, the moral hazard in the market for intermediate good x in structure B is replaced by the moral hazard in labor market in structure C. This formalizes Cheung's (1983) refinement of Coase's theory of the firm that the institution of the firm is to replace transaction costs of intermediate good with transaction costs of labor rather than replacing market with non-market institution. The structure of ownership and residual rights makes a difference because the parameter that relates to measuring or monitoring efficiency of intermediate good θ_x (relevant only to structure B) differs from that related to monitoring efficiency of labor, μ_x (relevant only to structure C).

In structure D, shown in Fig. 1(d), a specialist producer of x is the owner of a firm selling y and hiring labor to produce y . She owns goods produced by her and her employees before selling them. Specialist producers of x exchange y for labor used to produce y . In this structure, the trading efficiencies for the final good and for labor used to produce y are

$$(2c) \quad k_y = \begin{cases} \theta_y & \text{with probability } r_y s_y \\ 0 & \text{with probability } 1-r_y s_y \end{cases} \quad \text{and} \quad t_y = \begin{cases} \mu_y & \text{with probability } r_y s_y \\ 0 & \text{with probability } 1-r_y s_y \end{cases}$$

respectively, where μ_y relates to the monitoring and measuring efficiency of labor used to produce y , which is different from μ_x in (2b) that relates to the monitoring and measuring efficiency of labor used to produce x . Note that good y and labor used to produce y must be measured and priced in structure D.

A Walrasian regime is assumed since economies of specialization are individual specific or increasing returns are localized, all individuals are *ex ante* identical in all aspects, and they are allowed to choose between different patterns of

specialization and different contractual structures. As shown by Sun (1999), a general equilibrium may not exist if the set of consumer-producers is finite, and a general equilibrium exists if the set of consumer-producers is a continuum and the following assumptions hold in such a model. Preferences are rational, increasing, and continuous, production functions of intermediate goods are continuous and weakly increasing, production functions of consumption goods may exhibit at most local increasing returns, and transaction functions, which are $k_x x^d$ and $k_y y^d$ in this paper, are continuous and weakly increasing.

Since the interior solution is never optimal and the optimum values of decision variables are not continuous across structures, marginal analysis does not work for this model. A two-step inframarginal analysis is then used to solve for the general equilibrium. In the first step (section III), each individual uses marginal analysis to solve for a corner solution in a given configuration, which sorts out the efficient resource allocation for a given configuration of specialization. The corner solution generates indirect expected utility function for each configuration. Then for each structure, the market clearing conditions for traded goods and utility equalization condition between different constituent configurations generate a set of contingent relative prices of traded goods and a set of numbers (measures) of individuals choosing different configurations within a structure. We call them, together with quantities produced, traded, and consumed of goods under the contingent contract, the *corner (partial or local) equilibrium* for the given structure. Plugging the corner equilibrium relative prices and number of different specialists into the individual optimum plans of production, trade, and consumption, the corner equilibrium resource allocation for a given structure of division of labor can be found.

In the second step (section IV), the corner equilibrium terms of contract in a given structure are plugged into the expected utility functions. Expected utility of each constituent configuration in this structure is compared with all other

alternative configurations to identify the condition under which each individual has no incentive to deviate from her chosen configuration in this structure. This so-called total cost-benefit analysis can then be conducted for each corner equilibrium to partition the parameter space into several subspaces within each of which a corner equilibrium is the general equilibrium.

III. Corner Solution for a Configuration, Corner Equilibrium for a Structure

We first solve for the corner equilibria in structures A (autarky), B, C, and D.

(A) The decision problem in structure A is

$$\text{Max: } u = y = x^\beta (L_y - \alpha) = (L_x - \alpha)^\beta (1 - L_x - \alpha)$$

Its solution is $L_x = [(1-\alpha)\beta + \alpha]/(1+\beta)$ and $U(A) = \beta^{2\beta}/[(1-2\alpha)/(1+\beta)]^{1+2\beta}$, where $U(A)$ is the maximum per capita real income in structure A.

(B) The decision problem for a specialist of x in structure B is

$$\text{Max: } U_x \equiv Eu_x = E(k_y y^d)$$

$$\text{s.t. } y^d = p x^s \quad (\text{budget constraint})$$

$$x^s = L_x - \alpha \quad (\text{production function of x})$$

$$L_x + s_y = 1 \quad (\text{endowment constraint})$$

$$k_y = \begin{cases} \theta_y & \text{with probability } s_x s_y \\ 0 & \text{with probability } 1 - s_x s_y \end{cases} \quad (\text{trading efficiency})$$

where L_x , x^s , y^d , s_y are decision variables and $p \equiv p_x/p_y$ is the price of good x in terms of good y. s_x is given by the decision of the buyer of x who is the y specialist. The solution of this decision problem yields optimum values of L_x , x^s , y^d , s_y , and expected indirect utility function of this configuration for a given positive realized value of $p \equiv p_x/p_y$.

The decision problem for a y specialist in structure B is

$$\begin{aligned}
 & \text{Max: } U_y \equiv Eu_y = y && \text{(utility function)} \\
 \text{s.t. } & y+y^s = (k_x x^d)^\beta (L_y - \alpha) && \text{(production function of } y) \\
 & y^s = px^d && \text{(budget constraint)} \\
 & L_y + s_x = 1 && \text{(endowment constraint)} \\
 & k_x = \begin{cases} \theta_x & \text{with probability } s_x s_y \\ 0 & \text{with probability } 1 - s_x s_y \end{cases} && \text{(trading efficiency)}
 \end{aligned}$$

where L_y , x^d , y^s , y , s_x are decision variables. s_y is given by the decision of the buyer of good y who is the x specialist. The solution of this decision problem yields optimum values of L_y , x^d , y^s , y , s_x and expected indirect utility function for this configuration as functions of contingent p .

For a continuum of individuals, the law of large number implies that in equilibrium, $s_y s_x M$ individuals end up with realized state of $k_x = \theta_x$ and $k_y = \theta_y$, while $(1 - s_y s_x)M$ individuals end up with realized state of $k_x = 0$ or $k_y = 0$, which implies that the deal fails to be substantiated and therefore the relative price p is irrelevant, utility levels are 0. For the realized state of $k_x = \theta_x$ and $k_y = \theta_y$, the market clearing conditions $M_y y^s = M_x y^d$, the population equation $M_x + M_y = M$, and utility equalization condition $U_x = U_y$ yield the corner equilibrium in structure B

$$\begin{aligned}
 p &= [\beta \theta_x (1 - \alpha)]^\beta [2(1 - \beta)/\theta_y]^{1 - \beta} / (2 - \beta), \\
 M_x &= [2/(1 - \alpha)]^\beta [\theta_y / (1 - \beta)]^{1 - \beta} \beta^{1 - 2\beta} M / \{1 + [2/(1 - \alpha)]^\beta [\theta_y / (1 - \beta)]^{1 - \beta} \beta^{1 - 2\beta}\}, \\
 M_y &= M / \{1 + [2/(1 - \alpha)]^\beta [\theta_y / (1 - \beta)]^{1 - \beta} \beta^{1 - 2\beta}\}, \\
 s_y &= (1 - \alpha)/2, \quad s_x = (1 - \alpha)(1 - \beta)/(2 - \beta), \quad x^s = (1 - \alpha)/2, \\
 y^s &= p x^d, \quad x^d = [(1 - \alpha)\beta \theta_x^\beta / (2 - \beta)p]^{1/(1 - \beta)}, \quad y^d = (1 - \alpha)/2p, \\
 U(B) &= (1 - \alpha)^{3 + \beta} (1 - \beta)^{2 - \beta} (\beta \theta_x \theta_y)^\beta 2^{-\beta - 1} (2 - \beta)^{-2},
 \end{aligned}$$

where $U(B)$ is the expected per capita real income in structure B.

Before solving for the corner equilibria in other two structures, let us establish the claim that the corner equilibrium in structure B is not locally Pareto optimal. To see this, an allocation that is locally Pareto superior to the corner equilibrium in B

can be found by maximizing expected utility in configuration (x/y) subject to the constraints that expected utility in (x/y) equals that in (y/x) and total consumption equals total production. I assume that a benevolent dictator uses the utility equalization condition to express p as a function of other decision variables and uses material balance constraints to eliminate as many variables as possible. Then she maximizes expected utility of a configurations with respect to all remaining decision variables in structure B.

It is not difficult to show that the optimum expected per capita real income solved by the dictator is

$$U^* = (1-\alpha)^{3+\beta}(1-\beta)^{1-\beta}(\beta^2\theta_x\theta_y)^\beta 2^{-2}(1+\beta)^{-1-\beta},$$

which is larger than the corner equilibrium expected utility in structure B, $U(B)$, iff $\beta^\beta(2-\beta)^2-(1-\beta)(1+\beta)^{1+\beta}2^{1-\beta} > 0$, which always holds for $\alpha, \beta \in (0, 1)$. In this model with two-sided moral hazard, each specialist's trading efficiency depends not only on her own effort level, but also on her partner's effort level in reducing transaction risk. Because of moral hazard, each player ignores the interdependence of the decisions, so that effort level in avoiding transaction risk is not Pareto optimal. The difference between U^* and $U(B)$ is the endogenous transaction costs caused by two-sided moral hazard in structure B.

(C) In structure C, an y specialist is the owner of a firm. She hires N workers and directs them to specialize in producing y in the firm. She owns all goods produced by her and her employees and claims residual returns. She pays a contingent market wage rate to her employees and has decision rights to use her employees' labor. The decision problem for her is

$$\begin{aligned} \text{Max: } U_y \equiv E u_y &= E y && \text{(utility function)} \\ \text{s.t. } y^s &= wN && \text{(budget constraint)} \\ x^s &= t_x(1-s_y) - \alpha && \text{(production condition of employee-specialist of x)} \\ y + y^s &= (x^d)^\beta (1-r_x - \alpha) && \text{(production condition of employer-specialist of y)} \\ x^d &= N x^s && \text{(input-output material balance within the firm)} \end{aligned}$$

$$t_x = \begin{cases} \mu_x & \text{with probability } r_x s_y \\ 0 & \text{with probability } 1 - r_x s_y \end{cases} \quad (\text{trading efficiency})$$

where y , x^s , x^d , y^s , N , r_x are the employer's decision variables, though x^s is the consequence of employees' actions under the employer's direction. s_y is given by the decision of the buyer of y who is an employee (x specialist), w is the wage rate in terms of the final good. The employer buys $1-s_y$ unit of labor from each employee, but only a fraction t_x of that labor is actually employed by the employer. The employer directs each employee to allocate $t_x(1-s_y)$ of labor to the production of x . The optimum decision for this configuration yields demand function for labor and supply function of y , and expected indirect utility function. Since, there are only two production activities in this model, the employer's control rights in using the employee's labor seems trivial. If there are more goods and the employee engages in more than one production activity in the model, the employer's control rights of the employee's labor will be nontrivial.

Note that the production function for the firm is a combination of the employer and employees' individual specific production functions. In other words, the employer hires employees' production functions and combines them with her own production function to generate a production function for the firm. This differentiates the model in this paper from other neoclassical model of production and differentiates the notion of economies of division of labor from the neoclassical concept of economies of scale.

Table 1: Corner Equilibria in 4 Structures

| Structure i | Relative prices | Numbers of Specialists | Demand, supply, effort level | Expected per capita real income $U(i)$ |
|-------------|--|--|--|---|
| A | | | | $=\beta^{2\beta}/[(1-2\alpha)/(1+\beta)]^{1+2\beta}$ |
| B | $p = \frac{[\beta\theta_x(1-\alpha)]^\beta}{[2(1-\beta)/\theta_y]^{1-\beta}(2-\beta)^{-1}}$ | $M_x = \frac{[2/(1-\alpha)]^\beta}{[\theta_y/(1-\beta)]^{1-\beta}\beta^{1-2\beta}M} \div \{1+[2/(1-\alpha)]^\beta [\theta_y/(1-\beta)]^{1-\beta}\beta^{1-2\beta}\}$, $M_y = M/\{1+[2/(1-\alpha)]^\beta [\theta_y/(1-\beta)]^{1-\beta}\beta^{1-2\beta}\}$ | $s_y = (1-\alpha)/2$, $s_x = (1-\alpha)(1-\beta)/(2-\beta)$, $x^s = (1-\alpha)/2$, $y^s = px^d$, $x^d = [(1-\alpha)\beta\theta_x^\beta/(2-\beta)p]^{1/(1-\beta)}$, $y^d = (1-\alpha)/2p$ | $(1-\alpha)^{3+\beta}(1-\beta)^{2-\beta}(\beta\theta_x\theta_y)^\beta 2^{-\beta-1}(2-\beta)^{-2}$ |
| C | $w = \frac{[2\beta(0.5\mu_x-\alpha)]^\beta}{[(1-\beta)/\theta_y]^{1-\beta}(1-\alpha)/(2-\beta)}$ | $M_x/M_y = N = 2\beta\theta_y/(1-\beta)$ | $s_y = L_x = 0.5$, $r_x = (1-\alpha)(1-\beta)/(2-\beta)$, $x^s = 0.5(1-\alpha)(1-\beta)/(2-\beta)-\alpha$, $y^s = 0.5wN$, $x^d = Nx^s$, $y^d = w/2$ | $[(1-\alpha)/(2-\beta)]^2 (1-\beta)^{2-\beta} [\beta\theta_y(0.5\mu_x-\alpha)]^\beta 2^{\beta-1}$ |
| D | $w = \frac{[\beta(1-\alpha)/\theta_y]^\beta}{[2(1-\beta)]^{1-\beta}(0.5\mu_y-\alpha)}$ | $M_y/M_x = N = (1-\beta)\theta_y/\beta$ | $s_y = L_y = 0.5$, $r_y = 0.5(1-\alpha)$, $x^s = 0.5(1-\alpha)$, $y^s = 0.5wN$, $x^d = x^s/N$, $y^d = w/2$ | $(1-\alpha)^{1+\beta} [(1-\beta)\theta_y]^{1-\beta} \beta^\beta (0.5\mu_y-\alpha) 2^{-\beta-2}$ |

The decision problem for an employee is

$$\text{Max: } U_x \equiv E u_x = E(k_y y^d)$$

$$\text{s.t. } y^d = w(1-s_y) \quad (\text{budget constraint})$$

$$k_y = \begin{cases} \theta_y & \text{with probability } r_x s_y \\ 0 & \text{with probability } 1 - r_x s_y \end{cases} \quad (\text{trading efficiency of } y)$$

The optimum decision is: $s_y = (1-\alpha)/2$ and $U_x = \theta_y w r_x / 2$. The market clearing and utility equalization conditions yield the corner equilibrium in structure C, summarized in Table 1. It should be noted that structure C with the firm replaces moral hazard in the market for good x with moral hazard in the market for labor.

(D) Structure D is symmetric to structure C. The only difference is that specialist producers of x rather than specialist producers of y are owners of firms in structure D. The corner equilibrium in structure D is reported in Table 1.

Following the method to show that the corner equilibrium in structure B is not locally Pareto optimal, we can show that the corner equilibria in structures C and D are not locally Pareto optimal. The efficient trade off between exogenous

monitoring costs, which relate to θ , μ , and endogenous transaction costs caused by moral hazard in different structures of transactions and residual rights determines the general equilibrium and its comparative statics.

IV. General Equilibrium and Its Inframarginal Comparative Statics

I now consider the second step of the inframarginal analysis. Inserting corner equilibrium contractual terms in a structure into the expected utility function, each individual can compare her expected utility in a constituent configuration in this structure with all other alternative configurations. Letting expected utility in each constituent configuration be not smaller than in any alternative configurations yields a system of inequalities that specifies a parameter subspace within which the corner equilibrium in this structure is a general equilibrium (or individuals have no incentive to deviate from the corner equilibrium). Taking structure B as an example. The condition under which an individual has no incentive to deviate from configuration (x/y) in structure B requires

$$(4a) \quad U_x(B) \geq U_y(B), U_x(C), U_x(D), U_y(C) U_y(D), U(A).$$

where $U_i(j)$ is the expected utility function for a specialist producing good i in structure j under the corner equilibrium contractual terms in structure B. The condition under which an individual has no incentive to deviate from configuration (y/x) in structure B requires

$$(4b) \quad U_y(B) \geq U_x(B), U_x(C), U_x(D), U_y(C) U_y(D), U(A).$$

$U_x(B) \geq U_y(B)$ in (4a) and $U_y(B) \geq U_x(B)$ in (4b) imply $U_y(B) = U_x(B) = U(B)$, where per capita real income in structure B, $U(B)$, is given in Table 1. This, together with $U_x(B) \geq U_x(C)$, $U_y(C)$, imply

$$(5a) \quad U(B) \geq 0.5\theta_y(1-\alpha)(1-\beta)w/(2-\beta) \text{ and}$$

$$U(B) \geq 0.5[2(1-\alpha)(0.5\mu_x-\alpha)\beta/w(2-\beta)]^{\beta/(1-\beta)}[(1-\alpha)(1-\beta)/(2-\beta)]^2,$$

(5a) can be rearranged as follows.

$$(5b) \quad 2U(B)(2-\beta)/\theta_y(1-\alpha)(1-\beta) \geq w \\ \geq 2\{[(1-\alpha)(1-\beta)/(2-\beta)]^2/2U(B)\}^{(1-\beta)/\beta}(1-\alpha)(0.5\mu_x-\alpha)\beta/(2-\beta)$$

where w is the corner equilibrium wage rate in terms of good y in structure B. For any finite positive value of w , (5b) holds only if $U(B) > U(C)$ where $U(C)$ is the per capita real income in structure C, given in Table 1. Similarly, it can be shown that $U_x(B) = U(B) \geq U_x(D)$, $U_y(D)$ in (4) hold, only if $U(B) > U(D)$. Hence, (4) holds only if $U(B) > U(C)$, $U(D)$, $U(A)$, or only if

$$(6) \quad \theta_x/\theta_y(0.5\mu_x-\alpha) > 4/(1-\alpha)^{(1+\beta)/\beta}, \quad \theta_x^\beta/(0.5\mu_y-\alpha) > 0.5[(2-\beta)(1-\beta)^{1-\beta}/(1-\alpha)]^2, \\ \text{and } \theta_x\theta_y > \beta[(1-2\alpha)/(1+\beta)]^{1+2\beta}[2^{1+\beta}(2-\beta)^2/(1-\beta)^{2-\beta}(1-\alpha)^{3+\beta}]^{1/\beta}.$$

(6) defines a parameter subspace within which the corner equilibrium in structure B is the general equilibrium. The second step is referred to as total cost-benefit analysis. Inframarginal analysis consists of marginal analysis of each corner equilibrium in the first step and total cost-benefit analysis in the second step. Following this procedure, we can identify parameter subspaces within which structures A, C, and D, respectively, occur in equilibrium. The inframarginal comparative statics of general equilibrium are summarized in Table 2 and proposition 1.

Table 2: General Equilibrium and Inframarginal Comparative Statics

| | | | | | | |
|-----------------------|--|--------------------------|--|-----------------------------------|---|---|
| Parameter subspaces | $\theta_x/(0.5\mu_x-\alpha) > b_1$ and $\theta_x^\beta\theta_y^{2\beta-1}/(0.5\mu_y-\alpha) > b_2$ | | $\theta_x/(0.5\mu_x-\alpha) < b_1$ and $(0.5\mu_x-\alpha)^\beta(0.5\mu_y-\alpha)\theta_y^{2\beta-1} > b_3$ | | $\theta_x^\beta\theta_y^{2\beta-1}/(0.5\mu_y-\alpha) < b_2$ and $(0.5\mu_x-\alpha)^\beta(0.5\mu_y-\alpha)\theta_y^{2\beta-1} < b_3$ | |
| | $\theta_x\theta_y < c_1$ | $\theta_x\theta_y > c_1$ | $\theta_y(0.5\mu_x-\alpha) < c_2$ | $\theta_y(0.5\mu_x-\alpha) > c_2$ | $\theta_y^{1-\beta}(0.5\mu_y-\alpha) < c_3$ | $\theta_y^{1-\beta}(0.5\mu_y-\alpha) > c_3$ |
| Equilibrium structure | A | B | A | C | A | D |

where $b_1 \equiv 4/(1-\alpha)^{(1+\beta)/\beta}$, $b_2 \equiv 0.5[(2-\beta)/(1-\alpha)]^2/(1-\beta)$,

$$b_3 \equiv 2^{-1-2\beta}(2-\beta)^2/(1-\beta)(1-\alpha)^{1-\beta},$$

$$c_1 \equiv \beta[(1-2\alpha)/(1+\beta)]^{1+2\beta} [2^{1+\beta}(2-\beta)^2/(1-\beta)^{2-\beta}(1-\alpha)^{3+\beta}]^{1/\beta},$$

$$c_2 \equiv \beta \{ [(1-2\alpha)/(1+\beta)]^{1+2\beta} 2^{1-\beta} [(2-\beta)/(1-\alpha)]^2 / (1-\beta)^{2-\beta} \}^{1/\beta}, \text{ and}$$

$$c_3 \equiv \beta^\beta \{ [(1-2\alpha)/(1+\beta)]^{1+2\beta} 2^{2+\beta} / (1-\alpha)^{1+\beta} (1-\beta)^{1-\beta} \}.$$

Proposition 1:

- (1) The corner equilibrium in structure A is the general equilibrium if θ_x , θ_y , μ_x , and μ_y are sufficiently small.
- (2) The corner equilibrium in structure B is the general equilibrium if θ_x and θ_y are sufficiently large, compared to μ_x and μ_y .
- (3) The corner equilibrium in structure C is the general equilibrium if μ_x and θ_y are sufficiently large, compared to θ_x and μ_y .
- (4) The corner equilibrium in structure D is the general equilibrium if μ_y and θ_y are sufficiently large, compared to θ_x and μ_x .

$1-\theta_i$ and $1-\mu_i$ relate to exogenous monitoring or measurement costs of good i and labor employed to produce good i , respectively. Consider, for instance, $1-k_i = 1-\theta_i$ with probability $1-s_x s_y$ in structure B. This implies that with fixed effort levels of the two types of specialists s_x, s_y in reducing transaction risk, the transaction cost $1-\theta_i$ is lower as parameter θ_i increases. The increase in θ_i may be due to an increase in monitoring and measuring efficiency of quantity and quality of goods. Hence θ_i can be considered as an exogenous monitoring or measurement efficiency coefficient for good i and μ_i can be interpreted as that for labor employed to produce good i . But an increase in θ_i or in μ_i will affect the decisions in choosing s_x, s_y , thereby affecting the endogenous transaction cost (or degree of moral hazard) in a structure which is defined by the difference between the Pareto optimum and a corner equilibrium in the structure.

Proposition 1 indicates that monitoring or measurement efficiency determines if autarky or division of labor is the general equilibrium. Relative measurement efficiency between goods and labor determines if the division of labor is organized by the complete contingent contracts in the market for goods or the incomplete contingent contract in labor market. If labor is cheaper to measure (thereby labor contract involving a lower degree of moral hazard) than goods, then the institution of the firm emerges from the division of labor. Otherwise, labor trade and asymmetric structure of residual rights to returns and decision authority will not occur in equilibrium. The essence of our story of the firm and residual right is that any two elements of outputs of two goods (x and y) and labor employed to produce the two goods (L_x and L_y) can be used to organize the division of labor between the production of x and the production of y . If the division of labor occurs in equilibrium, the one out of the three possible structures of transactions that involves the lowest moral hazard will occur in the general equilibrium, so that division of labor will be promoted. As the number of production activities involved in the division of labor increases, the number of possible structures of transactions will increase more than proportionally. Then choice of structure of residual rights and ownership becomes increasingly more important for achieving a higher level of division of labor and related greater extent of the market and productivity.

Proposition 1 suggests that as measurement efficiency is improved within a certain parameter sub-space, the general equilibrium discontinuously jumps from locally Pareto efficient corner equilibrium in A to a locally Pareto inefficient corner equilibrium in structure B, C, or D. This increase in equilibrium degree of moral hazard is Pareto improving since the expansion of the network of division of labor increases aggregate productivity, which outweighs increased moral hazard. This again highlights the trade-offs between economies of specialization and endogenous and exogenous transaction costs.

This theory has an interesting empirical implication. It predicts that the equilibrium size of the firm may decrease when division of labor develops between firms as a result of improvements in measurement (or pricing) efficiency of goods which is faster than that for labor. Empirical evidence to support the hypothesis can be found in Liu and Yang (2000) and Murakami, Liu, and Otsuka (1996), who show that productivity, level of division of labor, and per capita real income increase and the average size of firms declines concurrently as transaction conditions are improved more rapidly for goods than that for labor.⁶

V. Concluding Remarks

The model in this paper supports the spirit of the GHM model by absorbing Maskin, Tirole, Holmstrom, and Milgrom's critiques of the theory of incomplete contract. It shows that the spirit of the GHM model can be kept in the absence of indescribability of infinite contingent states and of renegotiation if the trade-off between moral hazard and monitoring or measurement costs of goods and labor is introduced in a general equilibrium version of the GHM model. In the extended GHM model, it is relative measurement cost of different goods and factors rather than prohibitively high absolute contract writing cost that matters. It shows that contingent incomplete labor contracts can be used to avoid direct pricing and marketing of output and input of the activity that involves prohibitively high measurement cost and related moral hazard, meanwhile getting the activity involved in the division of labor. The labor contract is incomplete since it may not specify exactly what the employee has to do and it requires the employee to do whatever she is directed to do. Ex ante specified ownership of whatever produced by the

⁶ Increasingly more common business practice of down-sizing, disintegration, outsourcing, and focusing on core competency also supports the model in this paper (see Hart, 1995).

employee and asymmetric claims to residual returns and decision rights, together with exclusive rights to the company name (firm specific asset), is a very sophisticated way to indirectly pricing those intellectual properties, such as management know-how, which are too intangible to be directly priced even via a patent law. Hence, contingent incomplete labor contract can increase aggregate productivity by promoting the division of labor between the production of intangible intellectual properties and the production tangible goods.

References

- Aghion, Philippe and Tirole, Jean (1997), "Formal and Real Authority in Organization," *Journal of Political Economy*, 105, 1-29.
- Borland, J. and Yang, X. (1995): "Specialization, Product Development, Evolution of the Institution of the Firm, and Economic Growth", *Journal of Evolutionary Economics*, 5, 19-42.
- Cheung, S. (1983), "The Contractual Nature of the Firm", *Journal of Law & Economics*, 26, 1-21.
- Coase, R. (1937): "The Nature of the Firm", *Economica*, 4: 386-405
- Coase, R. (1960), "The Problem of Social Cost," *Journal of Law and Economics*, 3, 1-44.
- Gibbons, Robert (1998), "Incentives in Organization," *Journal of Economic Perspectives*, 12, 115-32.
- Grossman, S. and Hart, O. (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, 94: 691-719.
- Hart, O. (1995): *Firms, Contracts, and Financial Structure*, Oxford, Clarendon Press.
- Hart, O. and Moore, B. (1990): "Property Rights and the Nature of the Firm", *Journal of Political Economy*, 98: 1119-1158
- Hart, O. and Moore, B. (1999): "Foundations of Incomplete Contracts", *Review of Economic Studies*, 66: 115-38.
- Helpman, Elhanan and Laffont, Jean-Jacques (1975), "On Moral Hazard in General Equilibrium Theory," *Journal of Economic Theory*, 10, 8-23.
- Holmstrom, Bengt and John Roberts (1998), "The Boundaries of the Firm Revisited," *Journal of Economic Perspectives*, 12, 73-94.
- Holmstrom, Bengt, and Paul Milgrom (1991), "Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 7, 24-51.
- Holmstrom, B. and Milgrom, P. (1995): "The Firm as an Incentive System", *American Economic Review*, 84, 972-91.
- Kihlstrom, Richard and Laffont, Jean-Jacques (1979), "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," *Journal of Political Economy*, 87, 719-
- Legros, Patrick and Newman, Andrew (1996), "Wealth Effects, Distribution, and the Theory of Organization," *Journal of Economic Theory*, 70, 312-41
- Liu, P-W. and X. Yang (2000) "The Theory of Irrelevance of the Size of the Firm," *Journal of Economic Behavior and Organization*. Forthcoming.
- Maskin, E. and Moore, J. (1999), "Implementation and Renegotiation," *Review of Economic Studies*, 66: 39-56.
- Maskin, E. and Tirole, J. (1999a), "Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, 66: 83-114.
- Maskin, E. and Tirole, J. (1999b), "Two Remarks on the Property Rights Literature," *Review of Economic Studies*, 66: 139-49.
- Maskin, Eric and Xu, Chenggang (1999), "Soft Budget Constraint Theories: From Centralization to the Market", Working Paper, Department of Economics, Harvard University.
- Milgrom, P. and Robert, J. (1992), *Economics, Organization, and Management*, Englewood Cliffs, Prentice-Hall.
- Mokyr, Joel (1990) *The Lever of Richs: Technological Creativity and Economic Progress*, New York, Oxford University Press.
- Mokyr, Joel, (1993), "The New Economic History and the Industrial Revolution," in Mokyr, J. ed. *The British Industrial Revolution: An economic perspective*, Boulder and Oxford, Westview Press.
- Murakami, N., Liu, D., and Otsuka, K. (1996), "Market Reform, Division of Labor, and Increasing Advantage of Small-Scale Enterprises: The Case of the Machine Tool Industry in China", *Journal of Comparative Economics*, 23, 256-277.
- North, D. (1986), "Measuring the Transaction Sector in the American Economy", S. Eugerman and R. Gallman, eds., *Long Term Trends in the American Economy*, Chicago, University of Chicago Press.
- Rosen, S. (1978), "Substitution and the Division of Labor," *Economica*, 45, 235-50.
- Segal, I. (1999), "A Theory of Incomplete Contracts," *Review of Economic Studies*, 66: 57-82.
- Sun, Guangzhen (1999), *Increasing Returns, Roundabout Production and Urbanization: A General Equilibrium Analysis of the Division of Labor*. Ph.D. Dissertation, Department of Economics, Monash University.
- Tirole, J. (1999), "Incomplete Contracts: Where Do We Stand?" *Econometrica*, 67: 741-81.

- Yang, X. (1994), "Endogenous vs. Exogenous Comparative Advantages and Economies of Specialization vs. Economies of Scale", *Journal of Economics*, 60, 29-54.
- Yang, X. (2000): *Economics: New Classical Versus Neoclassical Framework*, Cambridge, MA, Blackwell.
- Yang, X. and Ng, S. (1998): "Specialization and Division of Labor: A Survey," in K. Arrow, et al, (ed.), *Increasing Returns and Economic Analysis*, London, Macmillan.
- Yang, X. and Ng, Y-K. (1993): *Specialization and Economic Organization: A New Classical Microeconomic Framework*, Amsterdam, North-Holland.
- Yang, X. and Ng, Y-K. (1995): "Theory of the Firm and Structure of Residual Rights", *Journal of Economic Behavior and Organization*, 26, 107-28.