SEMINAR PAPER 95-06

SPECIALIZATION AND DIVISION OF LABOR: A SURVEY

Xiaokai Yang and Siang Ng

Department of Economics
Monash University
Victoria, Australia 3168

and

Department of Banking and Finance
and Department of Economics
Monash University
Victoria, Australia 3168

April 1995

The Centre was established in 1983 by the Economics Department of the University of Adelaide to strengthen teaching and research in the field of international economics and closely related disciplines. Its specific objectives are:

- to promote individual and group research in the area by staff and students within the University of Adelaide and between Adelaide and other research institutions;
- to strengthen undergraduate and postgraduate education in this field;
- to provide specialized consulting services;
- to conduct seminars, workshops and conferences for academics and for the wider community;
- to publish and promote research results; and
- to improve public understanding of international economic issues.

Both theoretical and applied, policy-oriented studies are emphasised, with a particular focus on developments within, or of relevance to, the Asia-Pacific region. Further details and a list of publications are available from:

The Director, Professor Kym Anderson
Centre for International Economics Studies
University of Adelaide
Adelaide, S.A. 5005, AUSTRALIA

Telephone: (08) 303 7712, Facsimile (08) 228 4460
[International prefix (061) 8]

ISBN 0 86956 346 3
ABSTRACT

Specialization and Division of Labor: A Survey

The paper first briefly reviews the literature of specialization from Petty up to the 1980s. The classical research line of specialization that is followed by Young, Stigler, and Houthakker is distinguished from neoclassical economics structured by Marshall. The modern literature of formal models of specialization and division of labor is then surveyed in detail. Three lines of research are identified. One is associated with neoclassical trade theory which assumes constant returns to scale and explains the pattern of specialization and division of labor by exogenously comparative advantage between countries. The second line is associated with new trade and growth models which endogenize one aspect of the division of labor, the number of goods, by formulating the tradeoff between economies of scale and economies of variety of consumption or producer goods. The third line is associated with the models that endogenize all aspects of the division of labor: individuals’ level of specialization, the length of a roundabout production chain, and the number of goods in each link of the chain. In particular, the implications of new classical economics and inframarginal analysis for the resurrection of the spirit of classical economics in a modern body are explored.

JEL Classifications: B20, O00, F10, F43

Xiaokai Yang
Department of Economics
Monash University
Victoria
Australia 3168

and

Slang Ng
Department of Banking and Finance and
Department of Economics
Monash University
Victoria, Australia 3168

1. Classical Literature of Specialization and Division of Labor

The purpose of the paper is to survey the literature of formal models that endogenize individuals' level of specialization and the level of division of labor for society as a whole.

In the introductory section the classical literature of specialization is briefly reviewed. Neoclassical literature of specialization and new trade theory based on marginal analysis will be reviewed in section 2. Formal models of specialization based on inframarginal analysis will be surveyed in section 3. New classical equilibrium models of specialization based on corner solutions will be surveyed in section 4.

Houthakker [1956, p. 182] expressed the belief that “Most economists have probably regarded the division of labor, in Schumpeter’s words, as an ‘external common place,’ yet there is hardly any part of economics that would not be advanced by a further analysis of specialization.” This implies that the analysis of specialization and division of labor is not a subfield of economics, but rather it is at the central place of economics. The focus of classical economics was on the implications of specialization and division of labor for economic growth and welfare. William Petty [1690, I, pp. 260-61] noted that specialization contributes to skilful clothmaking and pointed out that Dutch could convey goods cheaply because they specialize each ship for a specific function. In another place, Petty gave a more striking example of the division of labor in the manufacture of a watch. Joseph Harris [1757] and Josiah Tucker [1756, 1774] referred directly to the productivity of the division of labor, the possibility for the subdivision of labor, and the intimate relationship between a greater variety
of goods, production roundaboutness, and a higher level of division of labor. Before Adam Smith, three advantages of the division of labor (improving the skill, or human capital in modern terms, of individual workers, saving the time and effort involved in having to switch from one operation to another, and facilitating the invention of machinery) and the role of the market and population size in permitting specialization were spelt out by French Encyclopedia, Henry Martin [1699, p. 591], Henry Maxwell [1721, p. 33], and Tucker [1774].

Smith [1776] called public attention to the central role of specialization and division of labor in economic analysis by systematically investigating their implications for economic growth and prosperity. Among other contributions, he proposed the conjecture that the extent of the market is determined by transportation efficiency [1776, pp. 31-32] in addition to the proposition that the division of labor is limited by the extent of the market [1776, chapter 3 of book I]. He proposed a theory of capital that takes capital to be a vehicle for increasing division of labor in roundabout productive activities (p. 371). He proposed what is now referred to as the concept of endogenous comparative advantage which implies that economies of specialization and division of labor may exist even if all individuals are ex ante identical (p. 28).

He explained the relative difference of productivity in the agricultural sector vs. the industrial sector between rich and poor countries as determined by the relative difference of benefit of specialization compared to seasonal adjustment cost caused by specialization

between the two sectors. This theory explains economic structure by the different balance points between economies of division of labor and coordination cost of the division of labor, instead of by tastes, or exogenous technical conditions. An extension of the theory implies that a decline in income share of the agricultural sector is not because of a change in tastes, in income, or in exogenous technical condition, but because the agricultural sector has a higher coordination cost of the division of labor compared to its benefit and it can improve productivity only by importing increasingly more industrial goods that are produced by a high level of division of labor in the manufacturing sector where transaction costs are relatively easier to be outweighed by economies of division of labor.

David Ricardo [1817] pursued an alternative line of studies of specialization and division of labor. He emphasized exogenous comparative advantage and what is referred to by Rosen [1978] as superadditivity which implies a type of interpersonal or social complementarity of production that generates a higher transformation curve for the division of labor than that for non-specialized production. Here, exogenous comparative advantage is defined as ex ante differences in tastes, technology, and endowments between individuals or countries that generate gains to trade.

Charles Babbage [1835, pp. 170-74] indicated that the division of labor can be used to save on fixed learning cost by avoiding duplicated learning and training. Karl Marx [1867, Vol. I, p. 368] drew the distinctions between simple cooperation without the division of labor and cooperation based on the division of labor, between the division of labor in production

---


2. According to Rashid [1986, p. 295], Cannan [1937], and Campbell and Skinner [1973], "not only does the example of pin-making appear to have been taken (by Smith) from the French Encyclopedia but also the three advantages of the division of labor are distinctly stated there."
and the division of labor in transacting activities, and between the division of labor within a firm and the division of labor in the market. Amasa Walker [1874, p. 39] called people’s attention to the implications of the division of labor in schools, lectures, churches, and journals and stated that division of labor facilitates the invention of tools, machines, and technology [1874, pp. 36-37]. He discussed the tradeoff between benefits of division of labor and related costs caused by a tension between specialization and seasonal changes in demand (p. 41), a sort of coordination costs for division of labor. He conjectured that division of labor will increase when the benefits outweigh the costs.

It is interesting to note that all the classical economists did not use the concept of economies of scale or increasing returns to scale. The concepts that they used are specialization, division of labor, and related benefits and costs. A careful reading of Alfred Marshall [1890] and Allyn Young [1928] indicates that the subtle distinction between the concepts of economies of specialization and economies of scale was crucial for subsequent development of economics.

Marshall [1890] tried to formalize classical economic thinking within a mathematical framework. His principles textbook consists of two parts. One (chapters 8-12) is full of classical insights into the economic implications of specialization and division of labor without mathematical formalization. The other is the marginal analysis of demand and supply within an internally consistent mathematical framework. In the first part the most important part of classical economic thinking about specialization and division of labor was not formalized within a mathematical structure. The second part successfully formalized the relatively unimportant part of classical economic thinking: the problem of resource allocation. Here, the problem of resource allocation is defined as a problem to find the efficient relative quantities of different goods and the efficient relative quantities of factors allocated to produce different goods for a given degree of scarcity (or a given transformation function) and a given pattern and level of division of labor. The problem of organization is to find the efficient level and pattern of division of labor to expand the production possibility frontier (or to reduce scarcity) against transaction costs for a given relative quantities of different goods consumed and produced. As Yang and Ng [1993] show, this was because the formalization of the classical economic thinking of problems of economic organization must involve corner solutions, but the technique to handling corner solutions and related inframarginal analysis was not available until the 1950s. Here, inframarginal analysis is defined as total benefit-cost analysis across corner solutions in addition to marginal analysis of each corner solution.

The success of the second part of Marshall’s principles textbook is based on an unrealistic dichotomy between pure consumers and pure producers (firms) which is essential for avoiding corner solutions. This dichotomy makes Marshall’s marginal analysis incapable of explaining the emergence of firms, business cycles, cities, money, middlemen, and a hierarchical structure of transactions from division of labor and of explaining the evolution of the extent of the market, productivity, comparative advantages, trade dependence, productivity, and many other interesting and important economic phenomena.

However, Marshall’s formalization of the resource allocation problem established the mainstream of economics in the following sense. Marshall’s mathematical structure of marginalism gives teaching a well organized structure. Within this structure, not only different generations of economists and students share a common dictionary, but also teachers can give good questions and exercises in classrooms and in examinations to which unique correct

---

6 Buchanan and Stubblebine [1962] propose this concept. The application of inframarginal analysis to a decision problem can be found from Kendrick [1978], Little and Mirrlees [1980], and Rosen [1983]. The application of inframarginal analysis to general equilibrium models can be found from Yang [1990], Yang and Wills [1990], Yang and Borland [1991], Yang and Shi [1992], and Yang and Ng [1993].
answer is expected. What teachers teach on the blackboard can be exactly duplicated by many students. This common mainstream facilitates the division of labor between different generations of economists and between different fields of economics and for continuous accumulation of economic knowledge. Due to the success of Marshall, what several generations of economics professors have been teaching in classrooms share the core part with Marshall's economics. We are all grand students of Marshall. Unfortunately, the mainstream does not carry the core part of classical economic thinking: problems of specialization and division of labor. As Buchanan [1994, p. 6] observes, "with one part of his mind always in classical teachings, Marshall recognized that this genuinely marvellous neoclassical construction requires that the Smithian proposition on labor specialization be abandoned." As an unexpected consequence of Marshall's success in formalizing problems of resource allocation, the core part of classical economics concerning specialization and division of labor has been forgotten.

Samuelson's principles textbook [1948] was another dividing line. This textbook consists of microeconomics that is Marshall's marginal analysis of demand and supply and macroeconomics that incorporates Keynesian economics which tries to explain many economic phenomena that Marshall's economics cannot predict. Since the 1950s when Samuelson's textbook became a prototype of principles textbook of economics, there have not been a place for problems of specialization and division of labor in principles textbooks. Each principles textbook just spends one paragraph to pay symbolic respect to classical economic thinking concerning specialization and division of labor. No formal models of Smith's endogenous comparative advantage are developed to endogenize individuals' level of specialization in the textbooks although they cover formal models of Ricardo's exogenous comparative advantage.

Marshall's neoclassical framework is characterized by the dichotomy between pure consumers and firms, the replacement of the concept of economies of specialization with the concept of economies scale, and marginal analysis of demand and supply. The debate on external vs. internal economies of scale and on other issues within the framework clarified some confusion. But as Buchanan [1994, p. 7] indicates, "Allyn Young sensed that the focus of economists' attention was shifting too readily and too rapidly toward clarification of analysis within neoclassical structure and away from classical emphasis." Young's paper [1928] is most cited by the modern literature of specialization and is regarded by Rosen [1983, p. 44] as "the zenith of the analysis of the connection between specialization and economic development" before the 1950s. Young emphasized the concepts of specialization, roundaboutness, and division of labor, and criticized the concept of economies of scale or increasing returns to scale which had been already very popular in economic teaching and research due to the success of Marshall's principles textbook of neoclassical economics. He argued [1928, 531] that "the view of the nature of the processes of industrial progress which is implied in the distinction between internal and external economies is necessarily a partial view. Certain aspects of those processes are illuminated, while, for that very reason, certain other aspects, important in relation to other problems, are obscured." Hence, it seemed to Young that the concept of external economies of scale is a misrepresentation of classical concept of economies of specialization and division of labor. Since Young and Marshall, the research of specialization has been developing along two lines. One is associated with Marshall's concept of (external or internal) economies of scale and with his marginal analysis, and the other follows Young's concept of economies of specialization and division of labor.

\[\text{The modern Arrow-Debreu model of general equilibrium, which is featured with the first two properties of Marshall's framework, has generalized and consolidated Marshall's framework.}\]
Young's concept of "social increasing returns" is very similar to Buchanan's [1994] concept of "generalized increasing returns" and to Rosen's [1978] concept of "superadditivity". Young stated several times that his increasing returns are not the ones that relate to the scale of a firm or an industry. Young's increasing returns are generated by specialization and the division of labor rather than by economies of scale. He used three concepts to describe the division of labor. The first is individuals' specialization. An individual's level of specialization increases as he narrows down his scope of activities. The second is the length of a roundabout production chain, or so-called roundaboutness. The third is the number of intermediate goods in each link of the chain. Certainly, the three concepts are related to and distinct from the concept of economies of scale. Indeed, Young's concept of social increasing returns based on specialization and division of labor is equivalent to the modern concept of a positive network effect of division of labor. The Young theorem [1928, p. 539, p. 534] consists of the following three statements. "The securing of increasing returns depends on the progressive division of labor"; "Not only the division of labor depends upon the extent of the market, but the extent of the market also depends upon the division of labor"; and "Demand and supply are two sides of the division of labor". The Young theorem represents the view that takes economies of division of labor as a network effect.

He suggested that the extent of the market is determined not only by population size, but also by purchasing power, which is determined by productivity, which is in turn dependent on the extent of division of labor. He then went on to argue that the circle that the division of labor depends upon the extent of the division of labor implies that a dynamic mechanism generates progressively increasing division of labor and the extent of the market. On the other hand, this circle implies that the size of the market network and the degree of division of labor are simultaneously determined. Suppose there are three ex ante identical individuals who prefer diverse consumption and specialized production in producing each of three goods x, y, and z. If an individual chooses to completely specialize in producing x, then he will demand y and z. If he chooses partial specialization in producing x and y, then he has no demand for x and y from the market and he will demand z. But if two individuals choose self-sufficiency, then the other cannot choose specialization. This implies that each person's decision of his level of specialization not only determines his productivity, but also determines the extent of the market for others' produce, thereby setting constraint for others' decisions of their level of specialization and for their productivity. Hence, the Young theorem explores a typical feature of network effects of the division of labor and related market. He implicitly, therefore, set up a research agenda to use economic models to explain how the size of the market network based on specialization and division of labor is determined in a decentralized system. Another more explicit target set by Young is to formalize the concept of economies of division of labor which includes economies of individual specialization, economies of

---

8 Young (p. 533) even argued that the use of the notion of large-scale-production misses the phenomenon of economies of division of labor.

9 Young [1928, p. 539] spells this out as follows. "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 which increasing returns are realized. What is required is that industrial operations be seen as an interrelated whole."

10 Roumasset and Smith [1984] provide evidence for the proposition that individuals' level of specialization determines the extent of the market.

11 This relates to Say's law. Yang and Ng [1993, chap. 18] show that a new classical dynamic equilibrium model may generate efficient business cycles and unemployment even if Say's law holds.

12 The concept of network effect here is consistent with Katz and Shapiro's definition [1983, 1986].
roundaboutness, and economies of variety of producer goods. On the basis of the formalization, a dynamic equilibrium model may be able to simultaneously explain the three aspects of the division of labor.

Although Young represented a more promising research line than Marshall’s neoclassical framework, he did not, unfortunately, formalize his ideas within a well-organized mathematical structure. However, he was sure that neoclassical marginal analysis is not suitable for his research agenda (p. 534). Certainly, Marshall’s concepts of demand and supply are substantially different from Young’s. The essence behind Marshall’s concepts of demand and supply is a tradeoff between quantities of different goods in raising utility and a tradeoff between quantities of factors allocated to produce a certain level of output. The market trades off one against the other to achieve an equilibrium. The equilibrium relative prices and relative quantities of goods and factors are explained by relative tastes, relative technologies, and relative quantities of different factor endowments. This story has nothing to do with the aggregate or absolute level of demand (extent of the market) which intimately relates to the level of division of labor and productivity.

Due to the dichotomy between pure consumers and firms, Marshall’s neoclassical model does not have any tradeoff that can be used to endogenize the level of specialization for individuals and the level of division of labor for society that determine the extent of the market and productivity. In other words, each pure consumer in Marshall’s model has to buy all goods that he consumes and he cannot survive in the absence of exogenously given firms. Hence, this model cannot explain why and how an economy evolves from an autarkical state where each individual self-provides all goods he consumes to a completely commercialized state, and why and how firms emerge and develop as division of labor evolves. Young inquires how individuals choose among different patterns and levels of division of labor to reduce scarcity or improve productivity. Marshall’s marginal analysis is devised to address the question how the market sorts out the relative quantities of goods consumed and produced for a given degree of scarcity, a given transformation curve, and a given pattern of organization within firms and between pure consumers and firms. Hence, Young’s economics is referred to as economics of organization and Marshall’s economics is referred to as economics of resource allocation by Yang and Ng (1993).

Houthakker (1956, p. 182) develops Smith’s proposition that the extent of the market is determined by transportation condition to suggest that the tradeoff between economics of specialization and transaction costs can be used to explain the level of division of labor. If the transaction cost coefficient is very large, then economies of specialization are outweighed by transaction costs caused by the division of labor, the equilibrium level of division of labor will be very low, so that the extent of the market is small and market demand and supply are zero. As the transaction cost coefficient falls, the efficient level of division of labor and productivity will increase, so that the extent of the market and demand and supply in the market place will increase. This is a promising direction toward the formalization of Young’s concepts of demand and supply. However, Houthakker argues (p. 182) “such an analysis of specialization involves the use of methods that are rather unlike those by which the classical questions of economics are discussed. These classical questions are treated with the aid of traditional calculus methods, but the latter are not suited to deal with indivisibility. It is in fact from indivisibility that the division of labor takes its start, and the basic indivisibility is that of the individual.”

He draws the distinction between economies of specialization and economies of scale when he discusses the necessity of new analytical approach. “We have increasing returns to the extent that if several activities are replaced by a single one, there is less need for (internal)
coordination and switching time and more scope for acquiring experience. The output of the single activity may thus be raised above the combined outputs of the several activities.

![Diagram of Individual and Aggregate PPF](image)

**Figure 1: Individual and Aggregate PPF**

He uses a graph to illustrate the distinction between Smith's concept of economies of division of labor and Ricardo's concept of comparative advantage. Suppose there are two ex ante identical individuals with the same production functions for two goods and the same endowment for labor: \( x_i = L_i - A, x_i = L_2 - A, \) and \( L_i + L_2 = 1 \), where \( x_i \) is the quantity of good \( i \) produced, \( L_i \) is the quantity of labor employed to produce good \( i \), and total amount of labor is one for each person. Then, an individual's transformation curve is ecdf in Fig 1. The aggregate transformation curve for the two persons when each of them produces two goods is segment gh. The aggregate transformation curve for the division of labor which implies at least a person producing only one good is point D and segments GBC and EFH. It is obvious that the aggregate transformation curve for the division of labor is higher than the aggregate transformation curve for autarky even if the two persons are ex ante identical or even if no Ricardian comparative advantage exists. This shows that Smith's concept of economies of division of labor may be more general than Ricado's concept of comparative advantage. Yang and Borland [1991] refer to Smith's comparative (and absolute) advantage as endogenous comparative advantage and refer to Ricardo's comparative advantage as exogenous comparative advantage.\(^{12}\) The following two points from Houthakker's graph deserve particular attention.\(^{14}\) (1) Specialization and diversification are two sides of the division of labor. Point G represents that two persons specialize in producing good y, but no division of labor exists. Also, point H is specialization but not division of labor. Point D, segment BC excluding point B, and segment EF excluding point F are associated with division of labor as well as specialization.\(^{15}\) (2) An ex post difference in productivity between sellers and buyers emerges from ex ante identical individuals only if they choose different levels of specialization in producing a good. If each of the two persons spends the same labor in producing each good, then ex post productivity is the same for them, so that no endogenous comparative advantage exists. This implies that the existence of endogenous comparative advantage depends upon individuals' decisions on their level of specialization.

Houthakker complained that evolution of specialization and division of labor in an economic system seems more significant and important than the evolution of species, but

\(^{12}\) Also, Arrow [1979, p. 154, p. 162] has spelt out the distinction between Smith's endogenous comparative advantage and Ricardo's exogenous comparative advantage and the tradeoff between economies of specialization and communication costs.

\(^{14}\) A similar graph for non-linear production functions can be found from Yang (1994a) or Yang and Ng (1993, chap. 2).

\(^{15}\) Many economists (for instance, Arrow [1979]) regard specialization as equivalent to division of labor. Though the two go together in many situations, they are not equivalent.
research of the former in economics is far behind studies of the latter in biology. This complaint still has very important implication for current economic research.

Another important paper that followed up Young's research line was George Stigler [1931]. Like Houthisker, Stigler used a graph to emphasize the distinctive nature of specialization that a firm's productivity increases as it narrows down its range of production activities. He demonstrated that cost function will be endogenously and discontinuously changed by a firm's decision on its level of specialization. The discontinuous change of cost function that is caused by a change in a firm's level of specialization is similar to the inframarginal analysis developed by Rosen [1983] and Yang [1990]. However, Stigler still followed Marshall's approach of separating the analysis of demand from the analysis of decision making regarding the level of specialization. He emphasized internal increasing returns to specialization against Marshall's concept of external economies of scale. When discussing the problem of vertical integration, he mixed up the concept of economies of specialization with the concept of economies of scale, departing from Young's research line.

The literature of specialization has been developing along three lines since Stigler and Houthakker. The first is to develop formal models based on Ricardo's concept of exogenous comparative advantage and on Marshall's framework with the dichotomy between pure consumers and firms and marginal analysis of demand and supply, focusing on the division of labor between countries rather than on the endogenization of individuals' level of specialization. This turned out to be a field called the theory of international trade. Dixit and Norman's textbook [1980] is representative of this research line. The second is to develop formal models using the concept of economies of scale and Marshall's framework to endogenize one aspect of Young's concept of division of labor: the number of consumption and producer goods. Recent works along this line are Dixit-Stigler [1977], Ethier [1982],

Krugman [1979], Romer [1986], Grossman-Helperman [1989], and others. The third is to endogenize individuals' level of specialization and the level of division of labor for society as a whole, following Smith, Young, and Houthakker's original ideas. Some of the formal models apply inframarginal analysis to endogenize the level of division of labor and explain all other economic phenomena by the level and pattern of division of labor. Section 2 briefly reviews the major works along the first two research lines. Then sections 3 and 4 survey the third line of research.

2. Exogenous Comparative Advantage and Economies of Scale

2.1. Neoclassical Trade Theory

Ricardo's model is a razor edge between increasing returns and diminishing returns from which either Young and Houthakker's ideas can be formalized or Marshall's line can be pursued. In his model production technology has constant average as well as marginal productivity of factor(s). Hence, it is not easy to apply marginal analysis to Ricardo's model. For instance, if two agents (individuals or countries) have the following production functions, and endowment constraint: \( x_i = 3L_{ij}, y_i = 2L_{ij}, L_{ij} = L_{ij} = 1 \), \( x_i = 2L_{ij}, y_i = L_{ij}, L_{ij} = L_{ij} = 1 \), \( u_i = X_i^a Y_i^b \), where \( x_i \) and \( y_i \) are respective output levels of the two goods for agent \( i \), \( L_{ij} \) is agent \( i \)'s amount of labor allocated to produce good \( j \), \( u_i \) is agent \( i \)'s utility level, and \( X \) and \( Y \) are respective quantities of the two goods consumed by agent \( i \). In this model, agent 2 has no absolute advantage, but has a comparative advantage in producing good \( x \). Equilibrium may be one combination of several corner and/or interior solutions. It can be shown that the Walrasian equilibrium based on zero profit conditions cannot support trade between the two agents.
Of course, a Nash bargaining game can generate a well defined trade equilibrium. However, the Nash bargaining solution is determined by a polynomial equation, which may not yield the analytical solution of the terms of trade. However, an analytical formula for each agent’s utility is needed for comparisons across several corner solutions. Certainly, marginal analysis does not work here. Each agent must do inframarginal analysis. Trade economists do not like inframarginal analysis. As Dixit and Norman [1980, p. 38] observe, “The Ricardo model is unsuitable for comparative statics. The phenomenon of multiple output choices with non-differentiable revenue functions makes it difficult to apply most standard techniques of analysis. For analyses which need single valued supply choices, therefore, attention has shifted to a post-Ricadian model. We have several goods produced using only one variable factor (the Ricardo-Viner model) or several factors (neoclassical trade model), but the factor(s) has diminishing returns in each use. Price change then cause a smooth shift of the factor from one use to another.” Actually, the comparative statics can be analyzed if a Nash bargaining equilibrium based on inframarginal analysis is solved though the algebra is not as simple as the one for marginal analysis based on interior solutions. Trade economists missed an opportunity for formalizing classical thought about specialization and division of labor when they were only one step away from inframarginal analysis which is essential for that exciting intellectual adventure.

From here, neoclassical trade theory goes on to focus on the models with constant returns to scale (diminishing returns to each factor) which cannot explain international trade from individuals’ decisions of their level and patterns of specialization. The dichotomy between pure consumers and firms in neoclassical trade theory implies that the rationale for international trade differs from the one for domestic trade. Domestic trade is essential even in the absence of exogenous comparative advantage (the differences in technology, endowments, and tastes) between individuals, and economies of scale because pure consumers cannot survive in the absence of domestic trade. But the driving forces of international trade are these elements. Also, international trade cannot be endogenized in two senses. First, trade is generated by exogenous comparative advantage which decision makers cannot change. Second, international trade is always better than autarky if any of these elements is there. Various degrees of involvement of countries in international trade cannot be explained by agents’ decisions. Wong and Yang [1994a, b] show that even if transaction costs are introduced into a neoclassical trade model, the degree of involvement in trade cannot be explained by the transaction cost coefficient for one unit of traded goods, as long as the coefficient is between 0 and 1 and the differential in the coefficient between international trade and domestic trade is not very large.

The neoclassical trade theory explains patterns of specialization and division of labor between countries by relative tastes and technical conditions compared to relative endowments between them. Suppose, for instance, countries 1 and 2 have the same tastes for all goods, the same endowment of labor, which is the sole factor for production, and the same technology in producing good y, but country 1’s labor productivity of x is higher than that of country 2. All consumers like y much more than x. Then country 2 will specialize in producing y and country 1 produces both x and y. If consumers like two goods to the same degree, then country 1 will completely specialize in producing x and country 2 in producing y. The determination of the pattern of division of labor is a problem of resource allocation concerning the efficient relative quantities of goods and efficient relative quantities of factors.

---

16 Buchanan and Stubblebine [1962], Coase [1946, 1960], Llewellyn and Margolis [1994] and other economists argue that many so called problems of coordination, multiple equilibria, and network externalities are generated by the assumption that individuals cannot do total benefit-cost analysis across corner solutions and they naively stick to marginal analysis of a globally non-optimum corner solution.
allocated to produce different goods. It has nothing to do with total factor productivity implication of specialization (due to the assumption of constant returns to scale), which is the focus of classical economic thinking.

2.2 New Trade and Growth Theory
Since the end of the 1970s, a new trade and growth theory has been developed to formalize the concept of gains to trade that are generated by economies of scale. In Dixit and Stiglitz's classical paper [1977], a tradeoff between distortions caused by global economies of scale and per capita real income, and the number of goods by the size of an economy. The CES utility function in their model implies that each good is not a necessity individually and utility increases with the number of consumption goods. A larger size of an economy generates a greater scope for trading off economies of scale against consumption variety, so that productivity and per capita real income increase as the size of the economy increases. The specification of a monopolistically competitive regime is essential for the tractability of the general equilibrium model with increasing returns to scale because the zero profit conditions can be used to get around the difficulty caused by the problem of profit distribution. The implications of the model for trade theory, which are explored by Krugman [1979], are straightforward. Opening up of international trade will improve productivity, raise per capita real income, and reduce the distortion caused by the monopoly power because the size of the pooling economy in an integrated world market is always larger than that for individual countries.18

As Krugman shows [1979 1980], gains to trade exist even if all countries are ex ante identical. Trade between identical countries might be more than trade between differentiated countries. This was an important step towards the formalization of Smith's concept of endogenous comparative advantage along the line suggested by Houthakker. Ethier [1982] extends the Dixit-Stiglitz model to the case with the CES production function. The Ethier model formulates a tradeoff between productivity gains of a variety of producer goods in producing final goods and global economies of scale in producing the producer goods. As the size of an economy increases, the scope for the market to trade off one against the other is enlarged, so that total factor productivity and per capita real income increase and distortions are reduced.

Following Judd [1985] who develops a dynamic version of the Dixit-Stiglitz model, Romer [1986, 1990] develops a dynamic version of the Ethier model, by replacing the concept of internal economies of scale with the concept of external economies of scale and by specifying perfect competition instead of monopolistic competition. This model formalizes one aspect of Young's concept of division of labor, that is the number of producer goods. Also, this model formalizes Young's conjecture that a dynamic mechanism generates spontaneous evolution of the number of producer goods and productivity. However, none of the new trade and growth models have endogenized the level of specialization for individuals and for firms.

In the models, each firm completely specializes in producing one good and an intermediate level of specialization never occurs in equilibrium, and each consumer has to buy all goods

---

17 Ng's model [1977, 1980] is similar to the Dixit-Stiglitz model in specifying a monopolistically competitive regime, but is used for macroeconomic analysis.

18 However, Krugman uses the Dixit-Stiglitz formula for the own price elasticity of demand which implies that opening up of international trade will not reduce distortions. Yang and Heijdra [1992] show that if a correct formula for the elasticity is used, the implication of international trade for reducing distortions can be explored.
that he consumes, so that the level of specialization for each individual and for each firm cannot be explained. Romer's story runs as follows.

There is a tradeoff between economies of scale and economies of complementarity between different intermediate goods in raising productivity of final goods. In addition another tradeoff exists between current consumption which can be increased by reducing saving and future consumption which can be increased by a higher saving level that increases the quantity of a primary factor which can be employed to increase the number of intermediate goods and thereby the productivity of final goods. A competitive market trades off one against others among the conflicting forces to achieve a dynamic equilibrium that generates long-run growth and spontaneous evolution of the number of intermediate goods. The dynamic equilibrium is not Pareto optimal due to myopic decisions of firms which do not take into account external economies of scale.

Grossman and Helpman [1989, 1990] develop a dynamic version of the Dixit-Stiglitz model and a dynamic version of the Ethier model, keeping the original flavor of internal economies of scale and monopolistic competition. Their story is similar to Romer's one, but distortions are caused by monopoly power instead of external economies of scale. The model can generate endogenous evolution of productivity and the number of goods. All of the new trade and growth models are featured with (static or dynamic) marginal analysis, economies of scale, and dichotomy between pure consumers and firms. Hence, this is a development within Marshall's neoclassical framework. This research line does not really follow Smith, Young, and Houthakker's original thoughts.

3. Specialization and Division of Labor

3.1. The Return to the Endogenization of Individuals' Level of Specialization

The first effort to bring the economic research based on formal models back to the original ideas of Smith, Young, and Houthakker might be attributed to Rosen [1978] and Becker [1981]. Rosen extends the Ricardo model to the case with many goods and many individuals. He applies linear programming rather than marginal analysis to handle the problem of corner solutions. Different from neoclassical economists, he concentrates attention on the implications of corner solutions for endogenization of individuals' level and pattern of specialization instead of getting around them. He uses a managerial decision model to clarify several important problems which are essential for the endogenization of individuals' level of specialization and the level of division of labor for society as a whole. He shows that economies of division of labor that are endogenously determined by individuals' decisions of their level and pattern of specialization look like external economies of scale, but they may exist in the absence of economies of scale. Economies of division of labor is not a technical concept because the linear programming model can be used to show that the elasticity of substitution between factors in an ex post production function generated by individuals' decisions of their level and pattern of specialization is different from that for the

---

19 Here, dynamic marginal analysis refers to the calculus of variations. The first order condition for the interior solution, the Euler equation, in a dynamic model implies that accumulated discounted marginal benefit equals instant marginal cost of investment. This is an analogue of the equalization condition between marginal benefit and marginal cost in static models. The control theory is sufficient but not necessary for managing the new growth models. The dynamic inframarginal analysis involves dynamic corner solutions (bang-bang control), the control theory, or dynamic programming (see Yang and Borland [1991] and Borland and Yang [1994]).

20 As soon as economists have realized the implications of economies of scale and CES function for the endogenization of aggregate demand, the Dixit-Stiglitz model is applied to macroeconomic analysis. See Blanchard and Kiyotaki [1987] and Rotemberg [1987].
corresponding ex ante production function. Hence, economies of division of labor is a social concept. Suppose there are exogenous comparative advantages among many individuals, then there are many possible transformation curves. As individuals choose different levels and patterns of specialization, resource allocation may jump from one transformation curve to another, generating changes in productivity (see Fig. 1 of Rosen [1983, p. 236]). This means more interaction and interdependence among individuals yields a larger scope for productivity improvement which is an interesting type of social and interpersonal complementarity. Rosen calls such social complementarity "superadditivity" which differs from economies of scale and from technical complementarity.

There are two types of technical complementarity. One is studied by Milgrom and Roberts [1990] who consider to what degree a factor complements another in producing goods for a given number of factors and goods. The second is studied by Dixit-Süglicht, Edhier, Judd, Romer, and Grossman and Helpman who investigate the implications of the complementarity between goods when the number of goods is endogenized. Rosen's superadditivity differs from both types of technical complementarity that are independent of individuals' decisions of their level and pattern of specialization. The distinction between Rosen's interpersonal superadditivity and technical complementarity can be used to draw the distinction between economies of division of labor, which relates to Rosen's superadditivity, and economies of scope which is equivalent to technical complementarity. Since various combinations of corner solutions generate many possible patterns of specialization and division of labor, each of the patterns is a network of exchanges. Hence, Rosen's superadditivity is a typical and important network effect. The essence of the Smith-Young theorem about the interdependence between the extent of the market and the division of labor relates to the nature of the network effects of the division of labor.

Indeed, if transaction costs are introduced into the Rosen model, a tradeoff between exogenous comparative advantage and transaction costs might be used to endogenize not only the pattern of division of labor, but also the level of division of labor and the related size of the network of exchanges. However, the algebra for this kind model is too complicated to manage because exogenous comparative advantages are based on the differences between each pair of individuals. Also, this model is difficult to extend to a general equilibrium model since the Walrasian equilibrium may not be well defined, while the Nash bargaining solution may be too complicated to analyze. However, Rosen's work is an important step toward shifting economists' attention from the marginal analysis of interior solutions to the endogenization of individuals' level and pattern of specialization and related inframarginal analysis of corner solutions.

Becker [1981] develops a model to endogenize individuals' decisions of specialization within a family. This model is solved implicitly using inframarginal analysis of many corner and interior solutions. The positive interactions between labor and human capital allocated to produce a certain good generates a pattern of complete specialization for each member of the family except one who might not be completely specialized because of a need for smoothing resource allocation between different consumption goods or services. Although this model is not explicitly specified as a dynamic decision model, and human capital plays a role similar to the one of the difference in endowment between agents in neoclassical models, it focuses on the endogenization of individuals' pattern of specialization and emphasizes the role of endogenous comparative advantage. This might be taken to be a starting point for formalizing Smith and Babbage's idea that the divisions of labor can be used to avoid duplicated fixed
Rosen was aware of the limitation of exogenous comparative advantages in his 1978 paper. He develops another model based on endogenous comparative advantage to explain individuals' level of specialization [1983]. Following Becker, he formalizes Arrow [1979, p. 154] and Barzel and Yu's ideas [1981] that the division of labor can increase utilization rate of a fixed learning and training cost. In this decision model, an agent maximizes the difference between benefit and cost of learning. 

\[ V = w_i k_i (1-t) + w_k (1-t) - C(k_i, k) \]

with respect to \( t \), which is the time allocated to produce good 1 and \( k \), which is learning and training level in activity \( i \), where \( C \) is the total learning and training cost, \( 1-t \) is the amount of time allocated to produce good 2, and \( w_i \) is a given benefit coefficient for activity \( i \). Since \( V \) is linear in \( t \), the optimum value of \( t \) may involve corner solution. Rosen uses marginal analysis to solve for the two corner solutions which represent specialization in different activities, and the interior solution which represents non-specialization, then compares total benefit-cost across all possible solutions. The result is that non-specialization takes place iff economies of technical complementarity between two learning activities outweigh economies of specialization generated by a higher utilization level of a particular learning and training investment. Rosen uses this model to emphasize again that interpersonal complementarity may exist in the absence of technical complementarity. If \( \frac{\partial C}{\partial k_i, \partial k} < 0 \), then technical complementarity does not exist, but two individuals can be better off by completely specializing in different activities and using the social complementarity.

The economies of specialization in this model are individual specific, so that increasing returns are localized. This implies substantial economies of specialization may be compatible with a competitive market. Many economists claim that Smith's economies of scale are incompatible with the invisible hand. However, Smith never used the concept of economies of scale which is imposed on him by others. Smith's concept of benefit from specialization and division of labor can be more accurately represented by the concept of economies of specialization or economies of specialized learning by doing that are individual specific. We cannot see why such localized increasing returns are necessarily incompatible with a competitive market. When two individuals engage in the division of labor, economies of specialized learning by doing will not go beyond the scale of each individual's working time. This implies that pooling two persons' labor together will not generate so great economies of scale that correspond to the size of the pooling labor although the social and interpersonal superadditivity (or complementarity) can be exploited by the division of labor between the two persons. Hence, economies of scale by pooling labor together within a firm and economies of division of labor based on specialization and interpersonal superadditivity are two distinctive concepts. The distinction is very subtle, but very important too. As Young argued, the replacement of Smith's concept of economies of division of labor with Marshall's concept of economies of scale obscures the distinction, thereby misleading economics.

The first model that endogenizes individuals' level of specialization and the level of division of labor for society as a whole by abandoning the dichotomy between pure consumers and firms was Yang's model [1984, 1985]. In this model, each individual is a consumer-producer who prefers diverse consumption and specialized production because of economies of specialization. A tension between specialized production and diverse consumption for each consumer-producer generates a tradeoff between economies of specialization and transaction costs. A central planner may trade off economies of

---

21 Also, Schultz [1986] has explored the intrinsic connection between human capital and specialization.

22 See Kim [1989] and Sah and Stiglitz [1986].
specialization against transaction cost to achieve the efficient level of division of labor by equalizing marginal benefit of the division of labor and marginal transaction cost. Since this is a planning model, marginal analysis can be used for decision making, and corner solutions which may emerge from a decentralized market are avoided. In terms of mathematics, the Becker and Murphy model [1992] is very similar to this model if the coordination cost in the Becker-Murphy model is interpreted as transaction costs in the Yang model. Both of them are decision models where the optimum level of division of labor is determined by the equalization condition between marginal benefit and marginal costs of the division of labor. Also, endogenous comparative advantages are driving forces in both models. If the marginal transaction or coordinate cost increases more rapidly than economies of division of labor as the number of different specialties increases, then the comparative statics of the optimum implies that division of labor will evolve if the degree of economies of specialization (in the Yang model) or a human capital parameter (in the Becker-Murphy model) and/or the transaction or coordination cost coefficient change.

In Becker and Murphy's model, however, each person is always completely specialized and each person's level of self-sufficiency is not endogenized. Hence, the extent of the market in Young's sense (each person's decision of self-sufficiency determines the extent of the market) is endogenized by Yang, but not by Becker and Murphy. But the Becker-Murphy model is a very important step toward formalizing ideas of Smith, Young, and Houthakker. This model shows that the efficient level of division of labor is determined not only by the population size which is usually considered as the extent of the market, but also by the efficient balance between economies of division of labor and coordination or transaction costs.

Also, Becker and Murphy recognize the implications of the Smith theorem that transportation condition determines the extent of the market which determines the level of division of labor. They discuss the implication of coordination costs for urban economics and explore the possibility for developing a dynamic model to explain evolution of division of labor by interactions between human capital and economies of specialization.

The Becker-Murphy model reminds us that the concept of extent of the market needs to be refined. There are three aspects of the extent of the market: population size; the number of goods; and the number of traded goods compared to the number of all goods. Certainly, the population size is a determinant of the extent of the market. If there is only one person in the world, the extent of the market is zero and the division of labor is impossible. If each individual buys all goods that he consumes, then his number of consumption goods determines his trade volume which in turn affects the extent of the market. All new trade and growth models (Dixit-Stiglitz, Krugman, Eber, Romer, Grossman-Helpman) have endogenized the extent of the market by endogenizing the number of goods. However, the new trade and growth models have not endogenized the third aspect of the extent of the market which is determined by individuals' level of specialization or their degree of self-sufficiency. If individuals' level of self-sufficiency is not appropriately endogenized, then the ultimate driving force of productivity, the number of goods, and division of labor is the population size. This result is inconsistent with the evidence for a negative correlation between the population size and productivity in some less developed economies, as documented in Wong and Yang [1994]. The Becker-Murphy model shows that a tradeoff between economies of specialization and coordination costs can endogenize the level of division of labor, while the driving force of the division of labor can be a falling coordination or transaction cost coefficient even if the population size is fixed.
Since the middle of the 1980s, the endogenization of individuals' level of specialization has been developing along two distinctive lines. One line is to endogenize individuals' level of specialization on the basis of the neoclassical dichotomy between pure consumers and pure producers or to endogenize individuals' level of specialization by assuming economies of scale. Baumgardner [1988], Kim [1989], and Locay [1990] are representative of the studies. Another line is to follow Yang's 1985 model and Becker and Murphy's 1992 model. The rest of this section is to survey the models of the first type. The second type of models will be surveyed in the next section.

3.2. Equilibrium Models of Specialization

The Baumgardner model [1988] represents a desperate effort to endogenize individuals' level of specialization within Marshall's framework. This is a partial equilibrium model with a given demand function for a continuum of service types. The analysis of demand is separated from the endogenization of individuals' level of specialization, a typical Marshallian way of thinking. The dichotomy between pure consumers and producers implies that marginal analysis is a major analytical instrument and inframarginal analysis is not needed. Global economies of scale in producing each type of service entails monopoly in the market for the range of service provided by each producer who can choose his range of service which differ from all others' services. The market will trade off the distortions caused by the monopoly power against economies of scale to determine the equilibrium range of services provided by each monopolist producer. As the population size increases, the scope for balancing the tradeoff is enlarged, so that more economies of scale will be exploited by narrowing down each producer's range of activities. Here, the population size is interpreted as the extent of the market. The model is used to explain why physicians in a large city are more specialized than those in small towns. Becker and Murphy [1992] point out that a higher level of specialization is generated by a lower coordination cost in a large city rather than by a larger population size. Young's idea that demand and supply are two sides of the division of labor and that the extent of the market and division of labor are two sides of the same coin is yet to be formalized.

Kim [1989], following Lancaster [1980] and Grossman and Shapiro [1982], specifies a model with many consumer-workers. A point on the circumference of a circle represents the characteristics of a worker which differ from that of each of other workers. Hence, there are exogenous absolute advantages between workers. But a matching cost between specialist workers and firms counteracts economies of division of labor generated by the exogenous absolute advantages. Each worker can invest in two kinds of human capital with a certain cost. One can be used to reduce the matching cost with a potential employer by expanding his characteristics, and the other can improve productivity. Each worker will trade off the benefit from a higher productivity against the matching cost to choose a pattern of two kinds of human capital. For a given pattern of each worker's human capital, free entry in the market will trade off the productivity gains from a narrower range of activities of each firm against the matching cost caused by the narrower range of activities of the firm, thereby determining a structure of firms. In this structure, all firms are evenly located on the circle, so that the average distance between each worker's range of characteristics and his employer's range of activities is minimized. Also, this structure of firms is characterized by the range of each firm's activities and the number of firms. As population size increases, the scope for the market to trade off economies of specialization against matching costs is enlarged, so that the equilibrium level of specialization of each firm, the number of various firms, and wage rate increase. As Weitzman [1994] shows, the models of this kind are dual to the Dixit-Stiglitz model [1977]. Hence, another interpretation of the Kim model is that a tradeoff between
economies of specialization and productivity gains of input variety in the market determines
the equilibrium number of various firms and their output level. However, Kim assumes a Nash
bargaining game rather than the regime of monopolistic competition. Since each consumer-
worker's welfare depends on only his wage, the tradeoff between diverse consumption and
increasing returns does not exist. This implies each person's level of self-sufficiency and the
extent of the market in the sense of Young's definition is not endogenized.

Locay's model of specialization [1990] is closer to classical thought of specialization
and division of labor than both Baumgardner's and Kim's models. Locay assumes that there
are many consumer-producers and that endogenous comparative advantages exist in producing
goods. Each person consumes only one particular good that is different from each of other
goods preferred by others. In producing the consumption goods, a long roundabout production
chain is needed. The consumption goods and producer goods are structured in a hierarchy. On
the top of the hierarchy is a producer good that is essential for the production of all producer
goods at the second layer. Each of the producer goods at the second layer is connected by a
span to the several producer goods at the third layer. Each producer good is essential for the
production of the producer goods that are in its span at a lower layer. For instance, steel can
be used to produce all kinds of machines, but a particular harvester can be only employed to
produce wheat. Hence, the demand for the upstream producer goods at the top of the structure
is greatest and the demand for the downstream goods at the bottom of the structure is smallest
due to specialized preferences of consumers. The extent of the market for producer goods at
different layers decreases as the production process descends from upstream to downstream.

The production of each good can be organized within a firm or within a household.
Economies of scale rather than economies of specialization are assumed. That is, as labor is
pooled together in a firm, productivity is increased even if the level of specialization of all
workers within the firm is fixed. The disadvantage of the production within a firm is a higher
monitoring cost than household production. This tradeoff between economies of scale and
monitoring cost implies that the production within a firm is superior to household production
only if the extent of the market for a good is sufficiently great such that economies of scale
outweigh increasing monitoring cost within a firm. When this tradeoff is balanced by a
competitive market within the hierarchical structure of goods, those upstream goods are more
likely to be produced within firms due to a greater extent of the market for upstream goods,
while those downstream goods are more likely to be produced by households due to a smaller
extent of the market for downstream goods. A larger population size will enlarge the scope
for the market to trade off economies of scale against monitoring costs within a firm. Hence,
the equilibrium dividing line between household production and the production within firms
will move towards the one that is in favor of firms as the population size increases.

This model mixes the concept of economies of scale with the concept of economies
of specialization. The hierarchical structure that is based on preferences for specialized
consumption is quite artificial. As in all models with the tradeoffs involving economies of
scale, the driving force of specialization in this model is the population size. This model
successfuly formalizes Stigler's thought about specialization and the extent of the market. It
endogenizes the extent of the market as determined by households' level of specialization. In
a sense, the monitoring cost within a firm can be considered as a kind of transaction costs.
Hence, this model touches on the Smith theorem that transaction efficiency determines the
extent of the market which determines the level of division of labor. If a monitoring cost
coefficient is specified, the Locay model may explain the extent of the market and households' level of self-sufficiency by the coefficient in the absence of any changes in the population
size.
4. New Classical Economics and Inframarginal Analysis

The first general equilibrium model that endogenizes individuals' level of specialization and the level of division of labor for society as a whole within a framework with consumer-producers, economies of specialization, and transaction costs was developed by Yang [1988]. For brevity, we call the framework "new classical framework." Inframarginal analysis and the concept of general equilibrium based on this analysis shall be outlined in subsection 4.1. The other subsections shall survey the application of the inframarginal analysis in different fields.

4.1 Basic Approach

We use the Yang model [1990] to illustrate the concept of equilibrium based on inframarginal analysis. In order to endogenize each individual's level of specialization, each consumer-producer must be allowed to choose any range of production activities. Hence, each decision variable can take on zero and positive values. In order to capture Smith's concept of endogenous comparative advantage, all consumer-producers are assumed as ex ante identical, equipped with the same production functions for each and every goods and the same quasi-concave utility function. An individual's labor share in producing a good is defined as his level of specialization in producing that good. A production function for a good is said to display economies of specialization if labor or total factor productivity of that good increases with a person's level of specialization in producing the good. In addition to the specification of production functions for each individual, an endowment constraint for individual specific labor is specified to capture the fact that economies of specialization are individual specific and increasing returns are localized. Hence, simply pooling labor together without an increase in individuals' level of specialization cannot increase their productivity. Each individual's self-interested behavior is represented by a non-linear programming problem that maximizes a person's utility with respect to his quantities of goods consumed, produced, and traded and level and pattern of specialization, subject to the production functions, endowment constraint, and the budget constraint.

For the decision problem, the first technical problem that must be tackled is a formidable large number of corner solutions. In each person's decision problem there are 3 decision variables for each good: quantity self-provided, quantity sold, and quantity purchased. Hence, for a model with m goods, there are 2^m combinations of zero and positive values of 3m decision variables, and therefore 2^m-1 possible corner solutions and one interior solution for each consumer-producer. In real world, individual decision makers can solve for the non-linear programming problem numerically for given parameters. But how can we economists solve for the problem analytically to identify demand and supply functions? This problem has been solved by Yang. He applies the Kuhn-Tucker theorem to narrow down the set of candidates for the optimum decision. According to proposition 1 in Yang [1988] or in Yang [1990], if there are economies of specialization and transaction costs, a consumer-producer never simultaneously sells and buys the same good, he never simultaneously buys and produces the same good, and he sells at most one good although he may produce several goods. For brevity, we call this proposition the Wen theorem since Wen [1994] proves this proposition for general specification of quasi-concave utility function and separable production functions with economies of specialization, and non-increasing transaction cost coefficient functions. The Wen theorem implies that the number of elements of the set of candidates for the optimum decision is much smaller than the number of elements of the set of all corner and interior solutions. For instance, the number of elements of the set of candidates is 10, but the number of all possible corner and interior solutions is 2^8 = 512 if there are 3 goods. The difference between the two numbers increases with the number of goods. Also, this theorem
implies that the interior solution can never be optimal, so that marginal analysis for interior solution does not work for the new classical framework.

A profile of zero and non-zero variables that is compatible with the Wen theorem is referred to as a configuration. The corner solution for each configuration can be solved using marginal analysis. Each corner solution determines the efficient resource allocation for a given level and pattern of specialization. For instance, corner demand and supply functions for traded goods and corner consumption and production plans of non-traded goods are part of each corner solution. The optimum corner solution determines the optimum level and pattern of specialization. It can be identified by a total benefit-cost analysis across configurations. An important feature of the inframarginal analysis is that the demand and supply functions and the indirect utility function discontinuously jump, respectively, across corner demand, corner supply, and corner indirect utility functions when a decision maker shifts between configurations.

Despite the fact that the Wen theorem significantly narrows down the set of candidates for the optimum decision, each individual still needs inframarginal analysis for identifying the optimum one out of many corner solutions. Hence, the second problem arises when we try to define and solve for a general equilibrium based on one of many combinations of corner solutions using inframarginal analysis.

A combination of several configurations that is compatible with the market clearing conditions for traded goods is referred to by Yang as a market structure or simply a structure. For each structure, a market clearing condition can be established for each traded good by specifying the numbers of individuals selling different goods and by equalizing total corner market demand and supply. Also, utility equalization conditions can be established by competition for a higher income between specialties (configurations). Hence, for each structure, there may exist a set of relative prices of traded goods and a set of numbers of individuals selling different goods that satisfy the neoclassical market clearing condition. Yang refers to the set of relative prices and the set of numbers of individuals selling different goods in each structure as a corner equilibrium. Each corner equilibrium is associated with a certain network of the market. Different corner equilibria are associated with different numbers of traded goods for society, differing degree of interdependence between different specialists, and different productivity levels. A Walrasian regime is assumed because the number of ex ante identical individuals is large and economies of specialization are individual specific.

General equilibrium is defined as a fixed point that satisfies the following conditions:

(i) Each individual uses inframarginal analysis to maximize his utility with 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; (ii) The set of relative prices of traded goods and the set of numbers of individuals selling different goods clear the markets for traded goods and equalize utility for all individuals selling different goods. There are two steps in solving for general equilibrium. First, a corner equilibrium is solved for each structure. Then, the general equilibrium is identified as the corner equilibrium that generates the highest utility level since it satisfies the two conditions for the definition of general equilibrium. The other corner equilibria are not general equilibrium since they do not satisfy condition (i) for general equilibrium. A rigorous proof of the statement that individuals have an incentive to deviate from these inefficient corner equilibria can be found from Yang and Ng [1993, chap. 6].

There are two types of comparative statics of the general equilibrium. The first type of comparative statics implies that the general equilibria, demand and supply functions, and indirect utility function will discontinuously shift between corner equilibria as parameters of
transaction cost and production have reached some critical values. The discontinuous jump of supply function is consistent with Stigler’s conjecture [1951] that a change in the level of division of labor will discontinuously shift the cost function which can be interpreted as endogenous technical progress. Another type of comparative statics of general equilibrium imply that the equilibrium relative prices, quantities of goods, and number of individuals selling different goods will continuously change in response to continuous changes of the parameters within the range defined by the critical values given by the comparative statics of the first type. The second type of comparative statics are analogous to neoclassical comparative statics of equilibrium based on marginal analysis. It generates the implications for resource allocation for a given level and pattern of division of labor. But there is no neoclassical counterpart of the first type of comparative statics based on inframarginal analysis.

It is easy to see that the efficient extent of the market and efficient level of labor productivity, scarcity, and per capita real income are different aspects of the level of division of labor. The efficient level of division of labor is determined by the tradeoff between economies of division of labor and transaction costs. Different levels of division of labor are associated with different transformation curves, as shown in Fig 1. But a very high transformation curve may not be efficient because it may be associated with a very high level of transaction cost. Hence, the conventional coincidence between the PPF and the utility frontier may not hold in this framework. But as the transaction cost coefficient falls, the utility frontier in general equilibrium will be closer to the PPF.

The first type of comparative statics substantially enhance the power of general equilibrium models in explaining changes in patterns of market network. The Yang model shows that the invisible hand can efficiently sort out the problem of network effects. Which network of the market and related division of labor is efficient depends on the transaction efficiency coefficient. If transaction efficiency is low, then the positive network effect of the market is outweighed by transaction costs, so that autarky or a low level of division of labor, which is associated with a small size of the network of the market, is efficient and will be chosen by the invisible hand. If transaction efficiency is improved, the efficient and equilibrium level of division of labor and related efficient size of market network will increase. Hence, whether the positive network effects can be utilized all depends on where is the efficient tradeoff between the positive network effects and transaction costs.

![Diagram](image)

**Figure 2: Configurations, Structures, and Evolution of Division of Labor**

Fig 2 gives an intuitive illustration of the first type of comparative statics of general equilibrium where the number of goods and the population are assumed to be 4 in an symmetric version of the Yang model [1990]. The lines in Fig 2 denote goods flows. The small arrows indicate direction of goods flows. The numbers beside the lines signify goods
involved. A circle with number \( i \) denotes a person selling good \( i \). Panel (a) denotes autarky where each person self-provides 4 goods, due to an extremely low transaction efficiency. Panel (b) denotes partial specialization where each person sells one good, buys one good, trades two goods, and self-provides three goods, as an improvement in transaction efficiency generates a partial division of labor. Panel (c) denotes extreme specialization where each person sells and self-provides one good, buys three goods, and trades four goods, due to a very high transaction efficiency.

It is interesting to see that marginal cost pricing rule no longer holds in the equilibrium involving specialization. In the Yang model [1990], for instance, the marginal opportunity cost of good \( x \) in terms of good \( y \) is 0 for a specialist of \( x \), but is infinite for a specialist of \( y \), but the equilibrium price of good \( x \) in terms of good \( y \) is a finite positive number in the general equilibrium with complete division of labor. This is a formal substantiation of Coase's argument [1946] against marginal cost pricing rule. He shows that total benefit-cost analysis is necessary and marginal analysis is inappropriate to the pricing of those goods with increasing returns in production. Moreover, emergence of professional middlemen from a high level of division of labor is predicted by the Yang model [1990] where economies of specialization in transacting activities are assumed.

The Yang model [1990] simultaneously formalizes many of the original ideas of Smith, Young, and Houthisker. This model shows that demand and supply are two sides of the division of labor and that the extent of the market (absolute level of aggregate demand) can be endogenized as one aspect of the level of division of labor. The most important function of the market is to choose the efficient size of the market network based on the division of labor. The concept of endogenous comparative advantage is formalized. This result has a significant implication for economics of transaction costs and institution because here it is transaction efficiency, which is affected by the legal system and other institutional arrangements, that determines the equilibrium (and efficient) productivity rather than that technology determines institution. From Fig 2, it can be seen that the absolute level of aggregate demand of each person, which is one aspect of the extent of the market is determined by each person's level of specialization. Yang and Ng [1994, chap. 3] specify a Rubinsteinian alternating-offer bargaining game in the Yang model. This subgame perfect equilibrium model shows that the pricing process in a bilateral bargaining game will generate inefficient resource allocation despite the Pareto efficient level of division of labor in the equilibrium. A multilateral bargaining equilibrium will eliminate the distortion of resource allocation.

Yang and Ng [1993, chap. 4] introduce Kreps and Wilson's concept of sequential equilibrium into the bargaining model to explore the dual functions of the market in exploiting economies of information asymmetry that generate economies of specialization and in restricting ex post transaction costs caused by information asymmetry. As Smythe [1994] comments, Yang and Ng's book [1993] reorients economics from problems of resource allocation to problems of economic organization. He realizes that the new classical framework and the concept of equilibrium based on inframarginal analysis provides a unifying core part for trade and growth theory, the theory of the firm, the theory of transaction costs and property rights, and macroeconomics. Many works have been done to apply the new classical

---

\(^{24}\) Simple versions of the Yang model that are easy to duplicate and teach can be found from Yang [1994a, d] and Yang and Ng [1993, chap. 2].
framework to the fields. The rest of the section shall survey the applications.

4.2 New Classical Trade Theory

Although the general equilibrium models based on inframarginal analysis is not easy to manage, its symmetric version with identical tastes and identical conditions of production and transactions for all goods is quite easy to handle since each consumer-producer's choice among different configurations can be simplified as a choice of the number of traded goods. Yang and Shi [1992] introduce the CES utility function with the number of goods as a variable into such a symmetric version of the Yang model [1990] to endogenize the two aspects of the division of labor: individuals' level of specialization and the number of all goods. A tradeoff between preference for consumption variety and each individual's management cost of the consumption variety is specified in addition to the tradeoff between economies of specialization and transaction costs. If the transaction cost coefficient is large, then economies of specialization generated by a high level of division of labor are outweighed by transaction costs. Hence, each individual will choose autarky (self-sufficiency) where there is no demand and supply in the market place and the extent of the market is zero. Autarky implies a very narrow scope for trading off economies of specialization against consumption variety because of each person's limited time. Hence, a small number of goods are produced and consumed by each person. As the transaction cost coefficient falls, each person will choose a higher level of specialization, so that a greater scope for trading off one against others among economies of specialization, economies of complementarity between different goods in raising utility level, transaction costs, and management costs of consumption variety emerges from the higher level of division of labor between different specialists. The equilibrium level of specialization, which determines the number of traded goods, and equilibrium number of all goods increase side by side as transaction efficiency is improved.

The extent of the market, absolute levels of aggregate demand and supply, productivity, trade dependence, consumption variety, the degree of market integration, the degree of production concentration, diversity of economic structure, variety of professions, each person's level of specialization, and the extent of endogenous comparative advantage which is defined by Yang and Borland [1991] as the difference in productivity between sellers and buyers of a good, all increase concurrently as division of labor develops.26

The new classical trade model is similar to the new trade models of Dixit-Stiglitz and Krugman because both of them endogenize the number of all consumption goods, productivity, and trade dependence (ratio of trade volume to income). However, the differences between the two types of models are important. Smythe [1994] draws the distinction by addressing the new classical trade model as "endogenous trade theory" in the sense that the degree of involvement of a country in international trade is endogenized by the new classical trade model, but not by the new trade models of Dixit-Stiglitz and Krugman.

Because of the feature of the D-S model and its variants, each consumer must buy goods from each of the monopolist producers. This, combined with the productivity implication of the population size, implies that the integrated world market is always better than any other patterns of organization and that separate local markets cannot occur in

26 The Herfindahl index of specialization cannot reflect the level of division of labor and thereby may be misleading because it does not reflect another side of the division of labor: diversity of various professions. According to the index, the level of specialization of Los Angeles, Chicago, San Francisco, and New York City are lower than Albany, Gary, and Norfolk (see Diamond and Simon [1990, pp. 180-183]). But from casual observation, we can perceive that the four large cities have certainly a much higher level of division of labor than the three small cities because of a higher degree of diversity of professions in the four large cities than in the three small cities. Chandler [1990] has documented that full exploitation of economies of scale and scope was a condition for rapid economic growth in the US in the end of the 19th century and early 20th century. In essence, the economies of scale and scope are two sides of economies of division of labor which can be fully exploited only if transaction efficiency is sufficiently great.
equilibrium. In other words, the degree of market integration is not endogenized in the D-S model. But in Fig 2, it is clear that the degree of market integration is endogenized if we assume that individuals trade first with those closest. Suppose the transaction cost coefficient is slightly larger for international trade than that for domestic trade, then international trade can be endogenized. If transaction efficiency is extremely low in a symmetric new classical trade model, autarky is the general equilibrium where no domestic and international trade exists and the economy is divided as M (population size) separate "local communities". As transaction efficiency is improved, the number of traded goods for each person as well as for the economy, n, increases and the number of separate local business communities, M/n, decreases. With continuous improvements in transaction efficiency, an integrated national market emerges from separate local business communities, followed by international trade with several separate international trade blocks, and finally ending up with the integrated world market. The story is explicitly spelt out by S. Ng [1995], using a two country model.

Here, the rationale for domestic and international trade is the same: the tradeoff between economies of specialization and transaction costs.

4.3. New Classical Theory of the Firm

It is easy to see that the new classical framework has very important implications for the theory of the firm since firms are not ex ante actors and they may emerge as ex post actors endogenously from the division of labor if trade in goods is replaced with trade in labor. Yang and Ng [1994] introduce intermediate goods and a differential in the transaction cost coefficients between trade in goods and trade in labor into the Yang model to develop a new classical theory of the firm.

Yang and Ng's story of the firm 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. There is a tradeoff between economies of specialization and transaction costs. If transaction efficiency is high, then division of labor will occur in equilibrium. Otherwise autarky will be chosen as the equilibrium. There are three different structures of residual rights which can be used to organize transactions required by the division of labor. Structure 1 is comprised of markets for cloth and management services. Specialist producers of cloth exchange cloth for management consultant service with specialist producers of management services. For this market structure, residual rights to returns and control are symmetrically distributed between trade partners and no firms and labor market exist. Structure 2 is comprised of the market for cloth and the market for labor hired to produce 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 of employees' labor and rights to the firm's residual returns are asymmetrically distributed between the employer and his employees. The employer claims the difference between revenue and wage bill, has control rights of his employees' labor, and sells goods that are produced from employees' labor. Structure 3 is comprised of 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 1, these two structures 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 hired to produce the intermediate goods. Although both structures 2 and 3 involve the firm and asymmetric structure of residual rights, they have different structures of ownership of a firm.

Assuming that transaction efficiency is much lower for management service than for
labor, the institution of the firm can be used to organize the division of labor more efficiently because it avoids trade in management services. Suppose further that transaction 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 efforts exerted producing intangible management (a sort of intellectual property) and to measure output level (quality and quantity) of management services. Then the division of labor can be more efficiently organized in structure 3 than in structure 2 because structure 3 involves trade in cloth and in labor hired to produce cloth but not trade in management services and in labor hired to produce management services, while structure 2 involves trade in cloth and in labor hired to produce management services. Hence, structure 3 will occur in equilibrium if the transaction efficiencies for labor hired to produce cloth and for cloth are sufficiently high. The claim to the residual 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 transaction efficiency involved in the division of labor while avoiding direct pricing and marketing of the activity, such that the division of labor and productivity are promoted.

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 in research and development. However, the asymmetric structure of residual rights to returns and control can indirectly price those intangible intellectual properties which are prohibitively expensive to enforce even through a patent law.

The Yang-Ng model, which formalizes Coase's [1937] theory of the firm, can explain the emergence of the firm from the division of labor and other endogenously complicated stories in the absence of uncertainties, exogenous comparative advantages, incomplete contracts, and other exogenous complications. The model does not endogenize ex post transaction costs in the sense of S. Grossman and Hart's definition (1986). However, if ex post transaction cost is defined as the transaction cost whose value is endogenously determined by individuals' decisions and the equilibrium process, then the Yang-Ng model has endogenized ex post transaction costs. This is because the number of transactions for individuals and for the economy as a whole is endogenized in this model due to the tradeoff between economies of specialization and transaction costs. It is not difficult to endogenize distortions generated by opportunistic behavior in an extended version of this model with the time dimension because economies of specialized learning by doing generate a scope for bargaining power and therefore for such opportunistic behavior. Borland and Yang [1994] have developed a dynamic version with the CES production function of the Yang-Ng model.

This theory has an interesting empirical implication. It predicts that the equilibrium size of the firm may decrease as division of labor evolves if the relative transaction cost coefficient of labor to goods rises or if increasingly more intangible intellectual properties need to be protected by different types of firms as they are involved in the division of labor.

Liu and Yang [1994] have tested the hypothesis against Hong Kong's data set.

4.4. New Classical Urban Economics and New Classical Theories of Industrialization and Hierarchy

Yang and Rice [1994] introduce a differential of the transaction cost coefficient between the manufacturing and agricultural sectors into the Yang model to show how and why the dual
structure between urban and rural sectors emerges from the evolution of division of labor. A
dual structure in terms of a differential in level of specialization and productivity may also
emerge in a transitional period from a low level to a high level of division of labor. Further,
Yang [1993] uses the new classical framework to explore the economic implications of
topological properties of economic organisms. Shi and Yang [1994] have extended the Yang-
Shi model to incorporate producer goods and CES production function. The Shi-Yang model
is a synthesis between the Yang model and Edhier model. It endogenizes the three aspects of
the division of labor described by Smith [1776] and Young [1928]: the level of specialization
of individuals, the length of a roundabout production chain, and the number of producer goods
in each link of the chain. Shi and Yang's story runs as follows.

Where there are economies of specialization, economies of complementarity between
producer goods in producing the final good, economies of roundaboutness, and transaction
costs, a tradeoff exists. A greater degree of horizontal division of labor in producing
downstream producer goods may generate more opportunities for the vertical division of labor
in producing upstream and downstream goods, which is associated with a larger number of
layers of the hierarchy of goods. This implies higher productivity, generated by a greater
variety of sophisticated professional equipment and machines, but, at the same time, greater
transaction costs.

If transaction efficiency is extremely low, then the gains to introducing more layers
of the hierarchy and further horizontal and vertical division of labor are outweighed by
transaction costs. In this case, each individual will choose autarky, that is, he will self-
provide all producer goods and consumer goods. A tradeoff still exists in autarky between
economies of specialization and increasing returns to a variety of producer goods. If a large
number of producer goods are produced in autarky, a person's level of specialization in
producing each good must be low. Thus, in autarky the foregone economies of specialization
due to the production of many producer goods at many layers of the hierarchy of goods
outweigh the gains to a variety of producer goods. Therefore, in autarky, each individual will
choose a hierarchy of goods with a small number of layers and a small number of producer
goods at each layer, so that he can capture more economies of specialization by concentrating
his limited labor in a few activities directly related to final consumption.

If transaction efficiency is extremely high, then people may choose a greater degree of
horizontal as well as vertical division of labor and in the meantime maintain each
individual's level of specialization at a high level through the division of labor between many
different specialists. Therefore, a high transaction efficiency may bring out some new layers
in the hierarchy of goods and new producer goods at each layer in the hierarchy. The
emergence of the new layers and new producer goods implies new technology and new
industries that are associated with an industrialization process. Hence, the general equilibrium
model predicts concurrent increases in the number of producer goods at each layer, in the
level of specialization, and in the number of layers of the hierarchy of goods. Since the supply
functions will discontinuously jump across structures as improvements in transaction
efficiency drive the evolution of division of labor, productivity progress and emergence of
new producer goods and new layer of roundabout production generate endogenous technical
progress.

Yang and Ng [1993, chap. 14] explore the intrinsic relationship between the level of
division of labor and the number of layers of the efficient hierarchy of transactions. Yang
[1994c] develops a general equilibrium model of hierarchy. This model shows that
improvements in transaction efficiency for transaction services will generate simultaneous
evolution of division of labor in producing various goods and different kinds of transaction
services and a hierarchical structure of wholesale and retail network. Professional wholesale and specialist retail middlemen will emerge from the evolution.28

4.5. New Classical Growth Models

All new classical models of specialization so far reviewed are static models. Although the comparative statics of the general equilibrium based on inframarginal analysis can generate shifts in the level of division of labor and productivity, the evolution of division of labor and productivity is exogenous evolution driven by exogenous improvements in transaction efficiency. Hence, Young's conjecture that dynamic equilibrium (the original phrase used by Young was "moving equilibrium") may generate spontaneous (endogenous) evolution of division of labor and extent of the market had not been formalized until Yang and Borland's dynamic general equilibrium model of specialization [1991].

The Yang and Borland model with consumer-producers, economies of specialized learning by doing, and transaction costs shows that the distinguishing feature of learning by doing in an economic system is specialized learning by doing through the division of labor. In their model, a mechanism of spontaneous evolution of the division of labor can generate endogenous growth. The evolution increases all individuals' productivity and decreases the ability of a single individual to survive independently of society at the same time. The feature distinguishes the learning by doing and knowledge accumulation that are associated with the evolution of the division of labor from the learning by doing that does not depend on the division of labor. Learning by doing based on the division of labor is associated with increases in the degree of organization and in the degree of interdependence between individuals, while learning by doing without division of labor is independent of organization.

Learning by doing in Arrow's [1962] and Alwyn Young's [1991] models is independent of the evolution of division of labor.

There are three patterns of learning by doing in the Yang-Borland (Y-B) model. The first is learning by doing in autarky in the absence of the division of labor and its evolution. The growth rate of percapita real income generated by this kind of learning by doing declines over time. The second pattern is learning by doing based on the evolution of division of labor. The compounded effect of individuals' specialized learning and an increase in division of labor accelerates the learning speed of society and increases the accumulation rate of human capital for society as a whole. As shown in the Y-B model, this pattern of learning by doing based on evolution of division of labor can generate takeoff (accelerated growth or increasing growth rate). This pattern of takeoff can be interpreted as economic miracle referred to by Lucas [1993]. Intuitively, a person's experience in an activity built up through the learning over a long period of time when he disperses his time among many activities at each point in time is equivalent to his experience in this activity through specialized learning by doing in this activity over a short period of time. If learning by doing occurs in the absence of division of labor, learning speed declines. If learning over time and specialized learning through increasing division of labor take place simultaneously, then speed of learning for society as a whole increases. The third pattern of learning by doing is based on a high level of division of labor in the absence of the evolution of division of labor. The growth rate of percapita real income generated by this kind of learning by doing declines over time although it is always positive.

The intuition behind the model is quite straightforward despite the technical sophistication of the control theory and dynamic general equilibrium based on corner solutions. Suppose there are productivity gains from specialized learning by doing,

28 This model formalizes Marshall's idea (p. 256, p. 264) that the division of labor in production creates more scope for the division of labor in management.
transaction costs, and consumer-producers who prefer diverse consumption, then each person does not have much experience in producing each and every goods at t=0, so that his productivity is low and he cannot afford transaction costs caused by specialization and division of labor. Autarky is thus chosen. As time goes by, each person builds up some experience (or so called human capital) in producing each and every goods, so that his productivity goes up slightly and he can afford a slightly larger transaction cost and therefore will choose a slightly higher level of specialization. The specialized learning by doing will speed up the accumulation of professional experience, so that each person's productivity in his professional activity increases further and therefore he can afford an even larger transaction cost and will choose an even higher level of specialization, and so on, until the potential for further evolution of division of labor has been exhausted. In the process, the growth rate of per capita real income declines in autarky, then increases (takeoff) as the division of labor evolves, and finally declines again (but is always positive) as the potential for further evolution of division of labor has been exhausted. The evolution of division of labor will increase the extent of the market (per capita effective demand times population size), production concentration of each traded good, the diversity of different professions, the extent of endogenous comparative advantages, the degree of market integration, each person's level of specialization, income share of transaction cost, each person's productivity in his profession, and so on. Fig. 2 provides illustration of the spontaneous evolution of division of labor if the comparative statics are interpreted as dynamics in the Yang-Borland model. The calculus of variations does not work and the control theory is essential for managing the model.

This model not only formalizes Young's insights into the relationship between the level of division of labor and the extent of the market and his point that demand and supply are two sides of the division of labor, but also can accommodate convergence and divergence phenomena of growth rates between developed and less developed economies. In the Y-B model, if the transaction cost coefficient falls, the evolution of division of labor will be speeded up although the change in the parameter is not necessary for the spontaneous evolution of division of labor. Suppose an economy started the evolution process earlier than other economies because of a lower transaction cost parameter (due to geographical position, say, this economy is in an island with a lower transaction cost coefficient for shipment) and it entered the takeoff stage when other economies were still in autarky (or at a low level of division of labor), then the growth rates between the two kinds of economies will diverge. But as the developed economy has finally exhausted the potential for further evolution of division of labor, and other economies have eventually entered the takeoff stage, then the growth rates between the two types of economies will converge. The model is consistent with Barro and Sala-i-Martin [1991, 1992], and Tamura's [1991] argument and evidence for convergence phenomenon, on the one hand, and is consistent with Romer [1986, 1990] and Lucas' [1988] argument and evidence for divergence phenomenon, on the other.

In the Yang-Borland model, as division of labor evolves to a very high level, a few specialists in producing each traded good will gain monopoly power from accumulated specialized learning by doing. Hence, long term contracts are essential for eliminating ex post transaction costs caused by economies of specialized learning by doing even if uncertainty is absent. However, casual observation indicates that a high level of division of labor may intensify competition between a few specialists through magnification effects of specialization. A high level of specialization implies that a small differential in performance will generate a large difference in the market share. This may more than offset the monopoly power resulting from specialized learning by doing. This idea needs another dynamic general
equilibrium model to formalize. Wang [1993] touches on this problem. He uses a new classical general equilibrium model with individuals' preference for relative position to show that division of labor may be promoted by individuals' behavior of pursuing relative position.

Since the Yang-Borland model [1991] is a dynamic version of the Yang model [1990] and the Borland-Yang model [1994] is a dynamic version of the Shi-Yang model [1994], on the one hand, and Romer and Grossman and Helpman's new growth models are dynamic versions of new trade model of Ethier, on the other hand, the differences and similarities between the new classical growth models of Borland and Yang and the new growth models of Romer, Grossman, and Helpman are analogous to those between the new classical trade models (Yang and Yang-Shi) and new trade models (Dixit-Stiglitz and Ethier).

Ng and Yang [1994] incorporate the concept of Kreps and Willson's concept of sequential equilibrium into the Yang model [1990] to investigate the function of the market in experimenting with various patterns of division of labor. They specify a tradeoff between information gains from experimentation with various patterns of division of labor and pricing costs incurred in the information acquisition process. If there is a fixed bargaining cost or a fixed cost for the communication between the Walrasian auctioneer and individuals, then individuals must trade off the pricing cost against expected information gains before they acquire all information about economic organization. Suppose individuals can experiment with one pattern of division of labor at each period. Each individual applies dynamic programming to maximize expected total discounted utility according to updated information at each period. The sequential dynamic equilibrium depends upon the four parameters: the pricing cost coefficient, the transportation cost coefficient, the degree of economies of specialization, and the discount rate. The smaller the two types of transaction cost coefficients and the discount rate and/or the greater the degree of economies of specialization, the faster is the evolution of division of labor and productivity.

4.6. New Classical Theory of Contract and Property Rights

Since the size of the market network is endogenized in the Yang model [1991], if uncertainties are introduced into the new classical framework, a tradeoff between economies of specialization and coordination reliability of the market network can be formulated to tell interesting stories about contracts, transaction costs, and property rights. The first of this kind of models was developed by Yang and Wills [1990]. This general equilibrium model explains the equilibrium level of division of labor and equilibrium degrees of competition, reliability, and vagueness of contractual terms by the tradeoffs among economies of specialization, coordination reliability, a risk of losing property rights, and transaction costs in specifying and enforcing property rights. The story runs as follows. A risk is specified for each transactions in the Yang-Wills model with m goods. This risk is caused by anticipated opportunistic behavior or uncertainties in transactions. As the level of division of labor increases, the chain of many different specialties connected in series becomes longer, so that the compounded risk for the coordination failure in the complicated division of labor increases more than proportionally, or total reliability decreases more rapidly. There are two ways to reduce the risk for the coordination failure.

One is that each buyer keeps in touch with many potential specialist supplies connected in parallel to put pressure on the incumbent one. A large number of incumbent and potential partners connected in parallel will raise total reliability for a given risk for each purchasing contract. But a larger number of the same type of specialists implies a smaller number of different types of specialties and thereby a lower level of division of labor for a

---

29 Rosen's statement [1981, p. 48] that "greater division of labor is itself a manifestation of increasing competition in the labor market" is consistent with our conjecture.
given population size. Also, an increase in the number of each person's incumbent and potential partners for each good will increase ex ante costs in keeping many potential relations (see Yang and Ng [1993, chap. 11]). The other way is to increase resources allocated for the specification and enforcement of each incumbent contract, so that the risk of coordination failure for each contract can be reduced. This will reduce welfare loss caused by the compounded risk of coordination failure. This welfare loss can be interpreted as ex post transaction costs. However, the decrease of the ex post transaction costs is associated with the increase in the ex ante transaction costs for specification and enforcement of contracts. Hence, there are tradeoffs among economies of specialization, ex ante transaction costs in widening potential relations, ex ante specification and enforcement costs of incumbent contracts, and ex post transaction costs.

The comparative statics of equilibrium suggests that as a transaction cost coefficient in specifying and enforcing property rights falls, the equilibrium level of division of labor, equilibrium extent of the market network, and productivity increase and the equilibrium risk for coordination failure which relates to vagueness of contract terms may either increase or decrease, depending upon relative values of the parameters representing the degree of economies of specialization, transportation efficiency, and conditions for specifying and enforcing property rights which relate to technical and institutional environment. Yang and Ng [1993, chap. 11] specify two cost coefficients to characterize the transaction costs, respectively, for deepening a relation and for widening potential relations. For a large value of the cost coefficient for deepening the incumbent contract relative to the cost coefficient for widening potential contracts, "classical contracts" with many potential trading partners (similar to perfect competition) may turn up in equilibrium. For a small value of the deepening cost coefficient, "relational contracts" without potentially alternative trading partners may turn up in equilibrium (Williamson [1975]).

This is the first model that formalizes Cheung's [1970, 1983] insights into the problem of "externality." According to the model, as transaction efficiency is improved, the equilibrium level of division of labor and equilibrium level of the risk for coordination failure may increase side by side. The equilibrium degree of unreliability and related ex post transaction costs can be considered as the equilibrium extent of externalities. Eliminating all the "externalities" is certainly not efficient. Another interpretation of the equilibrium degree of unreliability of each contract (specified as an endogenously determined probability) is the degree of softness of budget constraint. With this interpretation, the model implies that a perfectly hard budget constraint is not efficient if all the complicated tradeoffs are taken into account. The theory is tested against China's data by Yang, Wang, and Wills [1992]. The equilibrium degree of risk for coordination failure is a risk for mass unemployment since individuals will be forced to choose autarky and the corresponding low productivity when the coordination of complicated network of division of labor breaks down. However, this equilibrium and efficient degree of the risk for mass unemployment will increase as the transaction cost coefficient falls provided the economies of division of labor outweigh the expected welfare loss caused by the risk.

4.7. New Classical Theory of Money

Since aggregate demand is endogenized in the new classical equilibrium models and discontinuous jumps of the general equilibrium across many corner equilibria enhance

---

Cheung [1970, 1983] is the first economist who argues that the essence behind the concept of externality is a tradeoff between two kinds of transaction costs. Ex post transaction costs caused by distortions can be reduced by an increase in ex ante transaction cost for specifying and enforcing property rights. If the tradeoff is taken into account, according to Cheung, many seemingly inefficient contractual arrangements are the outcomes of the efficient balance of the tradeoff.
predictive power, explaining many so-called macroeconomic phenomena is the most interesting application of new classical economics. Subsections 4.7, 4.8, and 4.9 review three new classical equilibrium models which explain the emergence of money, unemployment, and business cycles from the division of labor and explore the intrinsic relationship between capital and the evolution of division of labor. This intellectual adventure will show that the dichotomy between microeconomics and macroeconomics is an unfortunate consequence of the fatal flaw of neoclassical microeconomics which was called economics and was supposed to be able to explain all micro and macro economic phenomena.

As Ostroyn and Starr (1990, p. 29) note "... a general formal treatment of the Smithian view of the interaction of money and specialization is still absent". Although Smith's suggestion (1776, chapter 4) that the driving force behind the use of money is specialization is generally accepted, Borland and Yang (1991) have shown that specialization is in fact a necessary but not sufficient condition for the emergence of money. They introduce a sufficiently long chain of roundabout production activities into the new classical framework and assume that transactions take time to complete. Hence, simultaneous implementation of all transactions essential for a high level of division of labor in a sufficiently long chain of roundabout production activities is infeasible. As transaction efficiency is sufficiently improved, the division of labor will involve a long chain of roundabout production activities, so that double coincidence of demand and supply between some pair of specialist trade partners is not satisfied. Hence, commodity money or fiat money become essential for the high level of division of labor in roundabout production. Their story runs as follows.

In the economy there are many ex ante identical consumer-producers who derive utility from food, which is produced using as inputs labor and hoes. Hoes are produced using labor and steel, which is itself produced using labor. There are transaction costs and economies of specialization in producing each good. Three main types of feasible market structure based on several configurations may occur: (i) autarky where each individual self-provides the intermediate and consumption goods; (ii) partial division of labor where, for example, some individuals specialize in the production of food, and other individuals produce both types of intermediate goods; and (iii) complete division of labor where each good is produced by an individual who specializes in that production activity.

Their model shows that given a sufficiently high degree of returns to specialization and a sufficiently small transaction cost coefficient, the general equilibrium market structure will involve the complete division of labor; if returns to specialization and the transaction cost coefficient are at intermediate levels, the general equilibrium will involve the partial division of labor; and if returns to specialization are insignificant and the transaction cost coefficient is large, individuals will remain in autarky. For the market structure with the complete division of labor, where an enforceable credit system exists, the equilibrium will involve trade which occurs through that credit system. In the absence of an enforceable credit system, the medium of exchange will be commodity money which is associated with the lowest transaction cost coefficient. An advantage of the endogenization of both the division of labor and emergence of money is that it allows the productivity implications of the emergence of money to be investigated.31

31 In other studies of the transaction role of money (for example, Kiyotaki and Wright (1989)) money is not necessary if all individuals are ex ante identical and the transition of an economy without money to an economy with money cannot be predicted. Kiyotaki and Wright (1993) claim, by interpreting a parameter of acceptability of commodity money as the reciprocal of level of specialization, that their model has endogenized the level of specialization. According to Yang and Ng's definition of an individual's level of specialization, the Kiyotaki and Wright model (1993) has not endogenized individuals' level of specialization.
division of labor in the inventing activities [1920, p. 256]. Edison's experience is another evidence for the implication of the division of labor for successful inventions. Not only Edison himself specialized in inventing electrical machines for most of his whole life, but also, he organized the first professional research institution with more than one hundred employees who specialized in different inventing activities (Josephson [1959]).

The observation suggests that investment in physical capital goods, in education, or in research would not automatically increase productivity in the future if the investment were not used to develop the right level and pattern of division of labor. Hence, the essential question around the notion of capital is not so much as to how much we invest and save, but rather as to what level and pattern of division of labor are used to invest in machines, education, and research.

The Yang model [1994b] has formalized the classical theory of capital. The story of capital runs as follows. There are many ex ante identical consumer-producers in an economy where food can be produced out of labor alone or out of labor and tractor. In producing each good, there are economies of specialized learning by doing. A fixed cost is incurred in the period when an individual engages in a job first time or when job shifting takes place. Each individual can choose between specialization and self-sufficiency. The advantage of specialization is to exploit economies of specialized learning by doing and to avoid job shifting cost. However, it increases productivity in the future at the expense of current consumption because of an increase in transaction cost.

Moreover, in producing a tractor, there is a significant fixed learning cost. The production of a tractor cannot be completed until the learning cost has reached a threshold level. Hence, there are tradeoffs among economies of specialized learning by doing, economies of roundaboutness, transaction costs, and fixed learning costs. If the transaction

---

22 Smith stated [1776, p. 371] "when the division of labor has once been thoroughly introduced, the produce of a man's own labor can supply but a very small part of his occasional wants. The far greater part of them are supplied by the produce of other men's labor, which he purchases with the produce, ... of his own. But this purchase cannot be made till such time as the produce of his own labor has not only been completed, but sold. A stock of goods of different kinds, therefore, must be stored up somewhere sufficient to maintain him, and to supply him with the materials and tools of his work, till such time, at least, as both these events can be brought about."
cost coefficient is sufficiently great, the economy is in autarky in all periods. If the transaction cost coefficient is sufficiently small and economies of specialized learning by doing and of roundaboutness are significant, the dynamic equilibrium involves division of labor. For the division of labor there are two patterns of investment and saving. If the fixed learning cost in producing tractor is not large, each individual will sacrifice consumption in period 1 to pay transaction costs in order to increase the level of division of labor, so that productivity in period 2 can be increased. This is a self-saving mechanism which does not involve the transfer of saving fund from an individual to another. The story of saving and investment for increasing division of labor is the same as in the Yang-Borland model [1991]. If the fixed learning cost in producing tractor is so large that the production of a tractor cannot be completed until time for specialized learning by producing tractor is longer than one period, then an explicit saving arrangement which involves a loan from a specialist producer of food to a specialist producer of tractor in period 1 is necessary for the specialization in producing tractors.

Under the assumptions of a great fixed learning cost in producing tractors, a small transaction cost coefficient, and significant economies of specialized learning by doing and roundaboutness, dynamic general equilibrium yields the following picture. A specialist producer of food produces food using his labor only and makes a loan in terms of food to a specialist producer of tractor in period 1 when the production of a tractor is yet to complete. In period 2, a specialist producer of tractor sells tractors to a specialist farmer in excess of the value of his purchase of food in period 2. The difference is his repayment of the loan received in period 1. Per capita consumption of food in period 1 is lower than in an alternative autarkical pattern of organization. But in period 2, tractors are employed to improve productivity of food. The discounted gains will exceed the reduction in the level of per capita consumption in period 1 if the transaction efficiency coefficient and economies of specialized learning by doing and roundaboutness are great. Economic growth takes place not only in the sense of an increase in per capita real income between periods, but also in the sense that total discounted real income is higher than in alternative autarkical patterns of organization.

The dynamic general equilibrium model based on corner solutions shows that investment does not necessarily increase future productivity. Productivity in the future can be increased by an investment that is used to create a higher level of division of labor which can speed up accumulation of professional experience (human capital) through specialized learning by producing roundabout productive equipment or services. If the transaction cost coefficient is large due to a deficient legal system or to a protectionist tariff, such opportunity for lucrative investment for increasing division of labor does not exist, so that investment may not increase real income. A decrease in the degree of economies of specialization and roundaboutness, an increase in the transaction cost coefficient, and/or exhaustion of the potential for further evolution of division of labor will reduce real return rates on investment and reduce opportunity for lucrative investment. This new classical theory of capital and interest rate is substantially different from Keynes' theory of capital [1937], which explains a sudden decline of interest rates by pure consumers' preference for liquidity. The new classical theory of capital is supported by the success of Hong Kong and Taiwan governments' liberalization and internalization policies towards investment and capital flow.

4.9 New Classical Theory of Business Cycles and Unemployment

The intimate relationship between the division of labor in producing durable goods, unemployment, business cycles, and economic growth is explored by Yang and Ng [1993, chap. 18] using a new classical dynamic equilibrium model. The story of efficient endogenous business cycles with unemployment runs as follows.
Suppose each individual can produce a non-durable consumer good called food and a durable producer good called tractor. A tractor is indivisible and can be used for 2 years. Each driver can drive one and only one tractor to produce food at any point in time. There are economies of specialized learning by doing in producing any good and two kinds of costs will be incurred if an individual shifts between productive activities. An individual will forget his experience built up in an activity if he shifts to another activity from this activity. Also there is an entry cost, such as a threshold learning cost, into any activity. Assume further, a tractor can be used for two years. Each consumer derives utility from food and maximizes his total discounted utility. For this simple economy, there are at least three possible organizational structures of production and consumption. The first is autarky where each person self-provides each good himself. He spends some time producing a tractor and the rest of time driving the tractor to produce food in the first year and produces only food using the tractor in the second year. Therefore, no business cycle and no unemployment exists for this structure. The second structure is that the population is divided between the production of tractors and the production of food. Professional farmers drive tractors to produce food in two years. Professional producers of tractors produce tractors in the first year and are unemployed in the second year. Aggregate output level is higher in the first year than in the second. This is a business cycle of two years with unemployment in the second year. The second cycle occurs over the third and fourth years, and so forth. The third structure is the same as the second except that the producers of tractors shift to the production of food in the second year. The final two structures involve the division of labor and specialization, but the second structure generates business cycles and unemployment and the third does not. However, autarky and the division of labor without business cycles and unemployment involves shifting costs of the producers of tractors between activities but the structure with business cycles and unemployment does not. If such shifting costs and economies of specialized learning by doing are sufficiently great, the division of labor with business cycles and unemployment is Pareto superior to autarky and the division of labor without business cycles and unemployment. The market mechanism (the invisible hand) may choose the efficient structure with business cycles and unemployment.  

5. Where Are We Going from Here

A principles textbook of new classical economics that is substantially different from Samuelson's principles textbook is badly needed. From several readers of Yang and Ng's book (1993), we received requests for new classical economics teaching materials. A series of textbooks of new classical economics will resurrect classical economic spirit in the modern body of new classical framework. A broad range of research can be further pursued. For instance, application of topology and graph theory in the analysis of economic organisms may generate some ground-breaking results, analogous to those which emerged from the application of topology and graph theory in the analysis of bio-chemical organisms. Introduction of exogenous comparative advantage into the new classical models with endogenous comparative advantage may generate much richer and more realistic stories. More general specification of functional forms may make the framework more applicable to the analysis of different cases.

If uncertainties are introduced into Yang's general equilibrium model of hierarchy

---

62

31 Weitzman's micro-equilibrium model (1982) and the models of monopolistic competition and sticky prices (see, for example, Mankiw (1983) and Ball, Mankiw, and Romer (1988)) explain business cycles by economies of scale and market failure. However, they cannot generate persistent efficient business cycles.
of labor may be explored. The blend between the Yang and Wills model of property rights [1990] and Borland and Yang's dynamic model of the firm [1994] may be able to predict a spontaneous evolution of sophisticated structure of property rights and contractual arrangements. If the distinction is drawn between learning from others' experience (learning by learning as Sah and Stiglitz [1986] call it), learning by teaching, learning by thinking, learning by experimentation (research), and learning by producing, then emergence and evolution of professional education and research may be explained by a new classical model. It is quite easy to propose the ideas, but much more difficult to substantiate them by striking an efficient tradeoff between generality, tractability, realism, and predictive power of models.

A blend of Borland and Yang's model of money [1991] and Yang and Ng's model of the firm [1994] may yield a new classical theory of professional banking firms. A combination of Ng and Yang's model of organization experimentation [1994] and Yang and Ng's model of the firm may yield a new classical theory of stock market which is a powerful vehicle for society to share risk in experimenting with various patterns of division of labor. If information asymmetry is introduced into Ng and Yang's sequential Walrasian equilibrium model [1994], the role of entrepreneurship in experimenting with various patterns of division

34 Krugman's economics of geography [1991] based on economies of scale relates to and differs from implications of specialization and division of labor for economics of geography.
References


Little, I. and Mirrless, J. (1980), Manual of Industrial Project Analysis in Developing Countries, Paris, OECD.


Marin, Henry (1699), Consideration on the East-India Trade.


Maxwell, Henry (1721), Reasons offered for erecting a Bank in Ireland, the second edition, Dublin.

Meek, R. and Skinner, A. (1973), "The Development of Adam Smith's Ideas on the Division


Petyt, William (1690), Political Arithmetic, in Hull (1899).


CIES SEMINAR PAPER SERIES

This Seminar Paper series provides a means for circulating, prior to publication, research results by staff or visitors associated with the Centre for International Economic Studies (CIES) at the University of Adelaide. Its purpose is to stimulate discussion and critical comment prior to the publication of papers. To facilitate prompt distribution, papers are screened but not formally refereed. Once published, papers will become available as part of the Centre's Reprint series.

Copies of CIES Seminar Papers are available by contacting the Director of the Centre (see inside front cover for the address). Single copies to individual researchers are free; the cost to institutions is $5 to each individual researcher and $25 to institutions, including postage charge every 4 months (payment by cheque or Visa/Mastercard/ Bankcard).


95-05 Andersen, Kym and Mari Pungesta, "Agriculture and Rural Development in Indonesia into the 21st Century", April 1995. (Forthcoming as an FAO Report, Rome.)

95-06 Yang, Xiaokai and Siang Ng, "Specialization and Division of Labor: A Survey", April 1995.
ACADEMIC ADVISORY BOARD OF THE
CENTRE FOR INTERNATIONAL ECONOMIC STUDIES

The Centre benefits from and is very grateful to members of its Academic
Advisory Board comprising professors both in Australia and overseas.
The following professors are currently serving a two-year term:

Australian Members

Jeff Carmichael, Bond University, Queensland
Kenneth W. Clements, University of Western Australia
Ross Garnaut, Australian National University, Canberra
Murray C. Kemp, University of New South Wales
Peter J. Lloyd, University of Melbourne, Victoria
D.T.C. (Tom) Nguyen, Griffith University, Queensland
Jonathan J. Pincus, University of Adelaide
Richard H. Snape, Monash University, Victoria
Alan D. Woodland, University of Sydney, New South Wales

Overseas Members

Robert E. Baldwin, University of Wisconsin and NBER
W. Max Corden, Johns Hopkins University
Steffi Hamilton, Stockholm School of Economics
Eric G. Johnson, University of Chicago
Heiner-Klirzkowski, Graduate Institute for International
Economics, Geneva
Paul R. Krugman, Stanford University
Richard D. Portes, London Business School
Jean-Waelbroeck, Free University of Brussels
John Whalley, University of Western Ontario