On the Emergence and Evolution of Mark-up Middlemen: An Inframarginal Model

Chun Pang\textsuperscript{1} and He-ling Shi\textsuperscript{2}

Abstract
This paper is aimed to provide an economic interpretation on the emergence and evolution of the specialised middlemen whose duty is to facilitate the transactions of goods and services in an economy. In a general equilibrium framework, the emergence and evolution of the specialised middlemen conforms to Adam Smith’s insight of deepening specialisation and the division of labour with the improvement in institutions and/or transaction technologies. Consequently, the emergence and the growth of the intermediation sector in both absolute and relative terms, the expansion of the network which provides transaction services, the evolution of market structure from autarky towards division of labour, the improvement in productivity, the reduction in wholesaling-retailing price dispersion, will be realised in concurrency

JEL Classification: L14, M21, O43

Keywords: middlemen, transaction efficiency, inframarginal economics

\textsuperscript{1} Research Centre of Division of Labour and Inframarginal Economics, Nanjing Audit University, Nanjing, Jiangsu, Province, P.R. China
\textsuperscript{2} Department of Economics, Monash University, Australia
Correspondence to: He-ling Shi, Department of Economics, Monash University, Clayton, Vic 3800, Australia.
Email: he-ling.shi@buseco.monash.edu.au.

© 2010 Chun Pang and He-ling Shi
All rights reserved. No part of this paper may be reproduced in any form, or stored in a retrieval system, without the prior written permission of the author
1. Introduction

The importance of market intermediation has been long recognized in literature. Jones (1936) investigated retail stores in the United States in the years between 1800 and 1860 and found that as transportation and communication facilities became better, the number of retailing outlets increased and the competition intensified. Attack and Passell (1994, p. 523) provided a recent incarnation of a similar story. Their data is reproduced in Table 1 which lists the employment data in the US labour market during 1840-1990. The whole economy is divided into three sectors - primary (agriculture), secondary (manufacture), and tertiary (services) sectors, respectively. The workforces working in the agriculture sector (column A and B of Table 1) increased from 1840, peaked in 1910 and then declined. The workforces employed in the manufacture sector (column C and D of Table 1) continued to increase, but with a declining rate. In contrary, workforces in the tertiary sector (column E and F of Table 1) continuously increased during the whole period of time. More impressively, the proportion of labour working in the service sector constantly increased - the ratio between the labour in service sector and the total workforce in agriculture and manufactures combined increased from 0.08 to over 1 in the last 150 years.

<table>
<thead>
<tr>
<th>Year</th>
<th>Agriculture</th>
<th>Industry</th>
<th>Service</th>
<th>Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>1840</td>
<td>3570</td>
<td>3594</td>
<td>500</td>
<td>918</td>
</tr>
<tr>
<td>1850</td>
<td>4520</td>
<td>4550</td>
<td>1200</td>
<td>1839</td>
</tr>
<tr>
<td>1860</td>
<td>5880</td>
<td>5911</td>
<td>1530</td>
<td>2391</td>
</tr>
<tr>
<td>1870</td>
<td>6790</td>
<td>6818</td>
<td>2470</td>
<td>3643</td>
</tr>
<tr>
<td>1880</td>
<td>8920</td>
<td>8961</td>
<td>3290</td>
<td>4775</td>
</tr>
<tr>
<td>1890</td>
<td>9960</td>
<td>10020</td>
<td>4390</td>
<td>6711</td>
</tr>
<tr>
<td>1900</td>
<td>11680</td>
<td>11749</td>
<td>5895</td>
<td>8722</td>
</tr>
<tr>
<td>1910</td>
<td>11770</td>
<td>11838</td>
<td>8332</td>
<td>12025</td>
</tr>
<tr>
<td>1920</td>
<td>10790</td>
<td>10843</td>
<td>11190</td>
<td>14513</td>
</tr>
<tr>
<td>1930</td>
<td>10560</td>
<td>10633</td>
<td>9684</td>
<td>13628</td>
</tr>
<tr>
<td>1940</td>
<td>9575</td>
<td>9635</td>
<td>11309</td>
<td>14995</td>
</tr>
<tr>
<td>1950</td>
<td>7870</td>
<td>7947</td>
<td>13648</td>
<td>20478</td>
</tr>
<tr>
<td>1960</td>
<td>5970</td>
<td>6015</td>
<td>17145</td>
<td>23234</td>
</tr>
<tr>
<td>1970</td>
<td>3463</td>
<td>3463</td>
<td>20746</td>
<td>26080</td>
</tr>
<tr>
<td>1980</td>
<td>3364</td>
<td>3364</td>
<td>21942</td>
<td>29136</td>
</tr>
<tr>
<td>1990</td>
<td>3186</td>
<td>3186</td>
<td>21184</td>
<td>29610</td>
</tr>
</tbody>
</table>

Source: Attack et al. (1994), “A New Economic View of American History: From Colonial Times to 1940” (2nd Ed.), New York: W. W. Norton & Company, p. 523. A – agriculture (not include fishing); B – agriculture and fishing; C – Total MFG (not include mining, construction, cotton textiles, iron & steel); D – all the manufacturing sectors; E – trade (not include ocean shipping, railroads, teaching and domestic); F – all the service sectors.

As a snapshot, Spulber (1996a) elaborated the role of intermediaries in the U.S. economy and estimated the intermediation accounted for 25% of value-added to GDP in 1993 in the U.S. by taking into account of retail trade (9.33%), wholesale trade (6.51%), finance and insurance (7.28%), and selected services (1.89%).

Although the emergence and continuous development of the service sector has been extensively discussed and its importance has been fully recognized in management literature – there is little economic study to tackle this issue.

Two lines of economic research are relevant to the current study

---

2 See example, Braudel (1982, 1984), Chandler (1977, 1990), Hall (1950), Hicks (1969), Porter and Livesay (1971), Shaw (1912), and Westerfield (1915),
North (1991) and more recently Kohn (2001a, 2001b, 2003a, 2003b, 2003c) provides historical insights on the continuous development of the service sector. Based on historical evidences, they stressed the importance of institutions in driving this trend. They argued that specialization in trading – the emergence of a specialized service sector which produces no tangible goods but facilitates transactions, is driven by the increasing transaction efficiency, which is further based on the improvement in institutions and/or transaction technologies. They asserted that the separation of trading activity from production of physical goods is a remarkable progress for economic growth and development.

Along with the transaction cost literature initiated by Ronald Coase (1937) and Oliver Williamson (1975) who have identified the importance of transaction costs in shaping the modes of market transaction, Spulber (1996b) initiated a search-equilibrium model which endogenized the microstructure of trade – that is, directing trading versus through market intermediation, in commodity and financial markets. The choice of intermediaries is dependent on the relative transaction costs and the nature of information asymmetry which leads to costly search. His model was followed and extended by including more variety of market intermediaries and endogenizing the equilibrium types and numbers of various market intermediaries. For example, Rust and Hall (2003) endogenized the choice between middlemen and market makers, and Ju, Linn, and Zhu (2010) modelled the choice between middlemen and oligopolistic market makers and its impacts on price dispersion.

While the first research line provides insights on the historical evolution of market intermediaries with increasing transaction efficiency, the second research line offers rigorous analysis on information asymmetry which drives the emergence of market intermediaries to facilitate market exchange via reducing search costs.

This paper will construct a general equilibrium model to address both the emergence and the historical evolution of market intermediaries in the wake of transaction efficiency improvement.

Yang and Ng (1993), Yang (2000, 2003) developed a research line – inframarginal economics, which formalizes Adam Smith’s insight of specialisation and division of labour. In short, with the increasing returns to specialisation, the division of labour will only be hampered by the extent of market which in turn is limited by transaction costs. Inframarginal economics have been used to investigate the microstructure of trade by endogenizing the number of tradable goods, number of total goods, and the trading patterns among market participates (for example, Yang and Shi, 1992, Shi and Yang, 1995).

Based on inframarginal economics, this paper will endogenize the level of specialization and the division of labour to analyse the mechanism for the emergence and evolution of intermediaries. In addition, in a general equilibrium framework, we will predict the increasing share of employment devoted to the intermediation/services sector (as is evident in Table 1, for example) along with economic growth measured by the increasing per capita income.

The intuition is as follows. As in standard inframarginal analysis, everyone could choose to be self-sufficient or rely on market for sell and buy. Self-sufficiency is lower in efficiency due to the inherited costs in producing multiple consumer goods – for example, multiple learning costs, switching costs, and so on. Market exchange, on the other hand, enjoys
economies of specialization in Smithian fashion, but will suffer from market transaction costs. Therefore, the choice of self-sufficiency or market transaction depends on the balance of the economies of specialization and market transaction costs. Any improvement in market transaction efficiency will lead to the increase in specialization and division of labour – consequently, specialists emerge, exchange network expands, and number of transactions increases.

Service, although intangible, is a type of measurable product. Banking services facilitate payments from buyers to sellers, and wholesale and retail services reduce search costs for buyers to find products. In fact, any product which ends up in a consumer’s hand, is a combination of the physical “core” product and all sorts of services attached to the “core” product. In this sense, a Lenovo notebook is NOT a notebook by itself. It is a combination of a physically tangible notebook and all sorts of attached services - including financial services, logistical services, wholesaling and retailing, branding and other marketing services, and after-sale quality warranty services.

In a self-sufficient economy, we might not need these services at all. Robinson Crusoe cares for himself without the need and possibility for market exchange. With the improvement of market transaction efficiency, specialisation and the division of labour start to shape up - including the emergence of some specialists who supply aforementioned intangible services. Professional bankers emerge from some producers who used to occasionally lend or borrow production capitals. Professional wholesalers emerge from some producers who used to store their products by themselves. Retailers might be a sort of people who used to go-between.

This illustrates the emergence of professional middlemen and dedicated service sectors. With further improvement in transaction efficiency, more tangible goods will be traded in the market – which exponentially increases the overall size of trading network. For instance, if two goods are traded in the market, only two trading connection are required. If three goods are traded, six trading connection are needed. Trading four goods requires a further12 trading connections. The expansion of trading network will therefore increase the demand for transaction services. As a result, the number of service specialists as well as the added value of the service sector will increase, more than proportionally than the sectors which produce tangible “core” products – such as the primary and secondary sectors.

Our model differs from Spulber-type models in the following three aspects: (a) A Walrasian regime is maintained in this paper - which allows the free entry in the intermediaries market; consequently, the equilibrium number of middlemen could be endogenized in the current model to capture the theme presented in Table 1. (b) For the tractability of the model which has a general equilibrium setting, a generic “iceberg” type of transaction cost is used to capture all costs associated with market transaction; rather than the specific and delicate search costs specified in Spulber-type models. (c) The inframarginal approach enables us to study comparative statics across different market structures; consequently, some quasi-dynamic stories could be demonstrated from our model.

In this paper, we mark-up middlemen who buy from producers and sell to end users. The rest of this paper is organized as follows. In Section 2, we will setup the economic environment which specifies production, market transaction, consumption and utility. In Section 3, individual’s optimum choices will be identified and solved – which will be combined in such a way to formulate a variety of market structures. In structures with market transactions, relative price and relative employment will then be investigated. In Section 4, we conduct
general equilibrium analysis by assuming the evolution of transaction efficiency; consequently, comparative static analysis will be conducted to draw some propositions which will be used to shed some lights on the explanation of relevant historical evidences. The final section contains concluding remarks.

2. The Model

Consider an economy, with a continuum of ex ante identical individuals of mass $M$, in which two consumer goods (both are necessities), $x$ and $y$, are produced and consumed. Transaction services for the trading of $x$ and $y$ are denoted as $r_x$ and $r_y$ respectively. Potentially, there are three types of economic agent – consumer, producer, and middleman. An individual is free to choose any combination of these roles in a specific time – for example, an individual can produce both $x$ and $y$, or just $x$ but need to purchase $y$ from others, or act as a middleman who produce $r_x$ but needs to purchase $x$ and $y$ for survival. Note the combination of roles will evolve over time. All individuals have diverse consumption preference for the two consumer goods. If individuals self-produce all the two consumer goods only for self-consumption, then there are not market transactions. Accordingly, there is no need for mark-up middlemen whose sole function is to provide transaction services during market transactions. On the other hand, when individuals purchase consumer goods from the market, trading activities arise. Trading activities must involve transaction services, including packaging, labelling, bar coding, product lot tracking, inventory controls, delivery and other after-sale services.

Apparently, to carry out trading activities, sellers have two options to handle transaction services. One is to self-provide transaction services when they independently sell goods to consumers. The second option is to sell “core” products to mark-up middlemen for resale. The second option implies the middlemen will produce transaction services and attach these services to the “core” products and then resell the package to end users. It can be seen that in the former setting without specialized mark-up middlemen, producers undertake two types of economic activities: the production of physical goods as well as the provision of transaction services. They play a dual role as producers and independent sellers. This is an economy with partial division of labour. By contrast, in the latter setting, trading activities separate from production activities. Such a separation implies that specialized mark-up middlemen come upon and the economy exhibits the division of labour. In this specialized economy, producers specialize in the production activities of physical goods; concurrently, middlemen specialize in transaction services. Correspondent to these two ways of providing transaction services, there are two approaches to purchase consumer goods. One is that consumers buy directly from producers; the other is that consumers buy from mark-up middlemen.

Before building up the model, let’s first summarize the specialized mark-up middleman’s most essential attributes - that is, (a) not producing physical goods but providing transaction services; (b) selling goods bundled together with transaction services rather than simply selling transaction services; (c) making a profit on price differences through resale with added values; and (d) playing an independent distribution or wholesale-retail role. For simplicity, it is assumed in this paper that mark-up middlemen perform a dual role as wholesalers and retailers - so-called independent distributors, directly linking consumers and producers.
With the economies of specialisation, a specialised producer and middleman will always outperform someone who produces both “core” product and transaction services. The counter-balance comes from the factor that there exist Coase-Williamson type transaction costs in the market – searching, bargaining, opportunistic behaviour, and etc. A produce could directly sell to a consumer through one market transaction, or sell to a middleman who then resell to a consumer through two market transactions. These two alternative transaction methods will certainly have different implications on the magnitude of transaction costs.

For the tractability of the model, we use a generic iceberg type of transaction costs to represent all costs associated with information asymmetry and leave the disaggregation of transaction costs to future study. More specifically, we assume $0 \leq k \leq 1$ as the transaction efficiency coefficient in the market which sells goods to the end users, and $0 \leq \mu \leq 1$ as the transaction efficiency coefficient in the market which sells goods to middlemen.

It is worthwhile distinguishing between transaction services ($r_x$ and $r_y$) and transaction costs ($1-k$, and $1- \mu$) in the context of this paper. In short, transaction services are intangible services which can be produced (endogenized by the model); while transaction costs are these which will be lost in transit (exogenous to the model).

Let’s now specify the model.

We use $x^p$ and $y^p$ respectively to denote the output levels of consumer goods $x$ and $y$. The production functions of the two types of goods are written as below

$$x^p = m(0|l_x - a) \text{ and } y^p = \max\{0, l_y - a\}$$

where $a \in (0,1)$ denotes fixed learning costs of producing goods; $l_x \in [0,1]$ and $l_y \in [0,1]$ are respectively an individual’s level of labour force devoted to producing goods $x$ or $y$.

Producers may have two approaches to supply their goods. One is to directly sell and dispatch goods to all the consumers. This approach means that producers self-provide transaction services for selling. In contrast, the second approach is that producers sell goods to middlemen and then the middlemen resell the goods to the consumers. The second approach means that middlemen provide transaction services for selling to consumers. In essence, these two ways all involve the allocation of output of producers. For producers of good $x$, the output allocation equations can be written as below

$$x^p = \begin{cases} 
  x + x_{self}^E & \text{if producers directly sell goods to all the consumers} \\
  x + x' & \text{if producers sell goods to middlemen who resell to consumers}
\end{cases}$$

Similarly, for producers of good $y$, the output allocation equations can be written as below

$$y^p = \begin{cases} 
  y + y_{self}^E & \text{if producers directly sell goods to all the consumers} \\
  y + y' & \text{if producers sell goods to middlemen who resell to consumers}
\end{cases}$$

where $x$ (or $y$) (here, $x, y>0$) denotes the self-consumed amount of producers; $x_{self}^E$ (or $y_{self}^E$) denotes the dispatched amount of producers who directly supply to the consumers; $x'$ (or $y'$)
denotes the amount of producers who sell to middlemen. In an autarkic economy, 
\[ x_{self}^E = x^i = y_{self}^E = y^i = 0. \]

As it is known, mark-up middlemen provide resale services with added values. As a consequence, the process of “buying in and reselling out” must involve the transactions of the same “core” good.

In trading activities, transaction costs will be incurred. A \( 1-k \ (0 \leq k \leq 1) \) amount will be lost in transit in the market which sells goods to the end users. So, the actually supplied amount of goods for the market must be less than the dispatched amount of sellers. For good \( x \), the actually supplied amount is \( kx_{self}^E \) (if producers self-provide transaction services for direct selling) or \( kx_{rx}^E \) (if middlemen specializing in transaction services \( r_x \) act as resellers). Here, \( x_{rx}^E \) denotes the dispatched amount of middlemen to consumers. It is assumed that trading activities must require transaction services \( r_x \), provided by sellers or middlemen. Thus, \( kx_{self}^E \) (or \( kx_{wx}^E \)) and \( r_x \) can be regarded as two intermediate inputs, both of which need to be combined, for outputs of sellers. In this model, we employ the Leontief technology to describe transaction process of consumer goods. Accordingly, the actually supplied amount of good \( x \) is written as below

\[
x_{self}^s = \min \{kx_{self}^E, r_x\} \text{ if the market is supplied directly by producers}
\]

\[
or \quad x_{rx}^s = \min \{kx_{rx}^E, r_x\} \text{ if the market is supplied by middlemen}
\]

where \( x_{self}^s \) (or \( x_{rx}^s \)) represents the actual supplied amount of good \( x \). The operator “\( \min \)” means that \( x_{self}^s \) (or \( x_{rx}^s \)) is given by the smaller of the two values in parentheses. In particular, suppose that \( kx_{self}^E \) (or \( kx_{rx}^E \)) \( < r_x \), then \( kx_{self}^E \) (or \( kx_{rx}^E \)) is the binding constraint in this transaction process. The employment of more transaction services \( r_x \), cannot raise the actually supplied amount of goods, and hence the marginal product of \( r_x \) is zero; that is, additional \( r_x \) is superfluous in this case. Similarly, if \( kx_{self}^E \) (or \( kx_{rx}^E \)) \( > r_x \), \( r_x \) is the binding constraint on the actually supplied amount of goods and additional \( kx_{self}^E \) (or \( kx_{rx}^E \)) is superfluous. Therefore, when \( kx_{self}^E \) (or \( kx_{rx}^E \)) \( = r_x \), both inputs are fully utilized. The economic intuition behind the Leontief transaction function employed here is reasonable. Consider a computer trader sells a given amount of computers in one transaction. If ten trucks are enough to meet the need for transporting these computers, then more trucks are not able to yield any more output at all.

Similarly, the transaction function of sellers of good \( y \) is written as below

\[
y_{self}^s = \min \{ky_{self}^E, r_y\} \text{ if the market is supplied directly by producers}
\]

\[
or \quad y_{ry}^s = \min \{kx_{ry}^E, r_y\} \text{ if the market is supplied by middlemen}
\]
where \( y_{self}^{x} \) (or \( y_{ry}^{x} \)) is the actual supplied amount of good \( y \) for the market; \( y_{self}^{E} \) is the dispatched amount of producers; \( y_{ry}^{E} \) denotes the dispatched amount of middlemen and \( r_{y} \) is transaction service for goods \( y \).

When mark-up middlemen, as re-sellers, mediate in the distribution and supply of consumer goods, they first purchase some amount of goods (for goods \( x \), for example) from producers, \( x_{rx}^{d} \) (namely, received amount). In the process of purchase from producers, there may be some losses. \( 1-\mu \) (here \( 0<\mu \leq 1 \)) denotes the loss ratio of the purchase in this market. Here, we define \( \mu \) as purchasing efficiency of middlemen (correspondingly, \( k \) is their selling efficiency). So, the actually received amount of good \( x \) is \( \mu x_{rx}^{d} \). Mark-up middlemen then dispatch a portion of \( \mu x_{rx}^{d} \) - that is, \( x_{rx}^{E} \) which is dispatched amount, bundle together with transaction services, for the supply. Of course, middlemen need to keep another portion, \( x_{rx}^{d} \), for their own consumption for survival. Therefore, the trading balance equation of mark-up middlemen for good \( x \) is established as below

\[
\mu x_{rx}^{d} = x_{rx}^{d} + x_{rx}^{E}
\]

Similarly, the trading balance equation of middlemen for good \( y \) is \( \rho y_{ry}^{d} = y_{ry}^{d} + y_{ry}^{E} \). These two balance equations can also be understood as “actually received amount = self-consumption + dispatched amount”. Here, \( \mu x_{rx}^{d} \) (or \( \rho y_{ry}^{d} \)) is the actually purchased (received) amount (or inbound amount); \( x_{rx}^{E} \) (or \( y_{ry}^{E} \)) is dispatched amount (or outbound amount, or stock amount). Obviously, for middlemen, \( \mu x_{rx}^{d} > 0 \), \( x_{rx}^{d} > 0 \) and \( x_{rx}^{E} > 0 \) as well as \( \rho y_{ry}^{d} > 0 \), \( y_{ry}^{d} > 0 \) and \( y_{ry}^{E} > 0 \).

We use \( r_{x} \) and \( r_{y} \) respectively denote two different transaction services for marketing goods \( x \) and \( y \). This differentiation is also reasonable. A retailer of daily consumer goods and a trader of electronic products have different professional know-how in marketing their items. The production functions of the two transaction services are written as below

\[
r_{x} = \max \{0, t_{d}(l_{rx} - b)\} \quad \text{and} \quad r_{y} = \max \{0, t_{d}(l_{ry} - b)\}
\]

where \( b \in (0,1) \) denotes fixed learning costs of producing transaction services; \( t_{d} > 0 \) is a parameter of production technology for transaction services, which may here called the distribution service technique coefficient. \( l_{rx} \in [0,1] \) and \( l_{ry} \in [0,1] \) respectively represents an individual’s level of specialization in producing the two different transaction services.

Apparently, there are four “production” activities in the economy, including the production of good \( x \), good \( y \), transaction services for good \( x \) and transaction services for good \( y \). Thus, the endowment constraint of the specific labour for an individual is

\[
l_{x} + l_{y} + l_{rx} + l_{ry} = 1
\]

where \( l_{x}, l_{y}, l_{rx}, l_{ry} \in [0,1] \).
The budget constraint of an individual is generally written as below

\[ p_x x^d + p_x x^d_{self} + p_y y^d + p_y y^d_{self} = p_x x^d + p_y y^d + p_y y^d + p_y y^d \]

where \( p_x \) and \( p_y \) respectively represent unit selling price of good \( x \) and \( y \) when producers self-supply these goods; \( p'_x \) and \( p'_y \) are respectively unit selling prices of goods \( x \) and \( y \) when middlemen supply these goods. Apparently, selling prices of producers are equal to buying prices of middlemen; selling prices of middlemen are equal to buying prices of consumers through a middleman. We will show later that \( p'_x > p_x \) and \( p'_y > p_y \). This means that, for a core good, there are two different prices in the case of mark-up middlemen. Here, \( x^d \) (or \( y^d \)) is the demand for \( x \) and \( y \) by a consumer who produce goods \( y \) (or \( x \)); and \( x^d_{rx} \) (or \( y^d_{ry} \)) is the demand for \( x \) and \( y \) by a mark-up middlemen (including quasi-specialized mark-up middlemen) who resell goods \( x \) (or \( y \)).

The general utility function of individuals is written as below

\[ U = (x + x^d + x^d_{rx})(y + y^d + y^d_{ry}) \]

Cobb-Douglas utility function is used to capture the necessity of two consumer goods. Take good \( x \) for example, an individual can end up with the total consumption of \((x + x^d + x^d_{rx})\) where \( x \) the self-produced and self-consumed amount, \( x^d \) is the demand for \( x \) if the consumer produces and/or resells goods \( y \), and \( x^d_{rx} \) is the demand for \( x \) if this individual also acts as a middleman who resell goods \( x \).

In this paper, the Walrasian regime is assumed\(^3\) - which implies an individual can freely choose to take up any combination of activity without barriers to entry. Without the loss of generality, it is assumed that transaction costs are paid by sellers (producers and middlemen) rather than buyers.

Here, the model can be re-stated as:

Max. \[ U = (x + x^d + x^d_{rx})(y + y^d + y^d_{ry}) \]

---

\(^3\) For the proof on the existence of general equilibrium in a Walrasian regime with increasing returns to specialisation and transaction costs, see Sun, Yang, Zhou (2004)
s.t. 
\[ x^p = \max \{ 0, l_x - a \}, \quad y^p = \max \{ 0, l_y - a \} \]
(production functions of goods \( x \) and \( y \))

\[ x^p = \begin{cases} x + x^E_{self} & \text{if producers sell to all the consumers} \\ x + x^s & \text{if producers sell to middlemen} \end{cases} \]
(output allocation equations of good \( x \))

\[ y^p = \begin{cases} y + y^E_{self} \\ y + y^s \end{cases} \]
(output allocation equations of good \( y \))

\begin{align*}
    x^s_{self} &= \min \{ kx^E_{self}, r_s \} & \text{if producers self-supply the market} \\
    x^s_{rs} &= \min \{ kx^E_{rs}, r_s \} & \text{if middlemen supply the market} \\
    y^s_{self} &= \min \{ ky^E_{self}, r_y \} \\
    y^s_{ry} &= \min \{ ky^E_{ry}, r_y \}
\end{align*}
(transaction functions of good \( x \))

\[ r_s = \max \{ 0, t_d(l_{rs} - b) \}, \quad r_y = \max \{ 0, t_d(l_{ry} - b) \} \]
(production functions of transaction services for goods \( x \) and \( y \))

\[ \mu^x_{rs} = x^d_{rs} + x^s_{rs}, \quad \mu^y_{ry} = y^d_{ry} + y^E_{ry} \]
(trade balance equations of middlemen for goods \( x \) and \( y \))

\[ l_s + l_y + l_{rs} + l_{ry} = 1 \]
(endowment constraint for time)

\[ (p_x x^d + p_s x^s_{self} + p'_s x^s_{rs}) + (p_y y^d + p_y y^s_{self} + p'_y y^s_{ry}) \]
\[ = (p_x x^d + p'_x x^s + p_s x^s_{rs}) + (p_y y^d + p'_y y^d + p_y y^s_{ry}) \]
(budget constraint)

3. Individual Decisions, Structures and Corner Equilibrium

Since every individual could take any combination of four occupations, so there is \( C_4^4 + C_4^2 + C_4^2 + C_4^4 \) corner solutions. The combination of these corner solutions could generate a myriad amount of structures which make the model intractable. Fortunately, the Optimal Configuration Theorem (see Yang, 2001, pp. 134-136) generates the following two claims which can be used to eliminate large number of structures which are not consistent with the economies of specialisation and the existence of transaction costs in market exchange.

Claim 1: the optimum decision does not involve selling the same “core” good to the middlemen and market concurrently\(^4\)

Claim 1 implies: if \( x^E_{self}, y^E_{self} > 0 \), then \( x^s = y^s = 0 \); if \( x^s > 0 \), then \( x^E_{self} = y^E_{self} = 0 \).

Claim 2: the optimum decision does not involve buying and producing the same “core” good\(^5\)

\(^4\) Intuitively, an individual would not engage in the production of transaction services \((r_s, r_y)\) to self-market proportion of its output while leaves the remaining proportion to the specialised middlemen in a Walrasian regime.

\(^5\)
Claim 2 implies: if \( x^d, x_{rx} > 0 \), then \( x^s, x_{self} = 0 \); if \( x^s, x_{self} > 0 \), then \( x^d, x_{rx} = 0 \). The same is applicable to good \( y \).

Those structures that do not conform to the last two claims of Optimal Configuration Theorem are eliminated from the subsequent analysis and therefore four structures - \( A, P1, P2 \) and \( C \) (see Figure 1), are remained to be examined. Of the four structures, \( A \) is autarkic structure; \( P1 \) and \( P2 \) are partial division of labour; and \( C \) is complete division of labour.

Now, we investigate each structure in depth.

---

\(^5\) Intuitively, an individual would not involve in the production of a product which involves a positive learning cost while buy the same good from market which attracts transaction costs.
**Figure 1: Mark-up Middlemen, Specialization and the Division of Labour**

**Structures A**

Structure A is self-sufficient (autarky), in which each individual produces all the two consumer goods $x$ and $y$ only for self-consumptions. The trade of goods does not occur. Obviously, transaction services for trading activities are not necessary; there are no mark-up middlemen. Although other services for the production of goods are necessary to autarkic producers, it is reasonable to consider these services as part of the production of goods. For this structure, some variables are equal to zero. That is, $x^d = x^d_s = x^E = x^E_s = x^s = x^E_r = x^E_r_s = y^d = y^d_s = y^E = y^E_s = y^r = y^r_s = y^E_r = y^E_r_s = r_s = r_s = l_s = l_r = 0$

Thus, the decision of each individual is as below

Max. $U^A = xy$
S.t. $x^p = l_x - a = x$
$y^p = l_y - a = y$
$l_x + l_y = 1$

Solving for this problem, we get

$$x = y = \frac{1 - 2a}{2}, \quad l_x = l_y = \frac{1}{2}, \quad U^A = \left(\frac{1 - 2a}{2}\right)^2,$$

for $0 < a < 1/2$.

**Structure P1**

In this structure with partial division of labour, some people produce good $x$ and self-provide transaction services $r_s$, for the supply of good $x$. Other people produce good $y$ and self-provide transaction services $r_y$, for the supply of good $y$. These two groups of producers
are mutually trading partners and carry out direct exchange for consumption. In this structure, these non-specialized producers have not reselling activities.

For producers of goods \( x \), their decision configuration is denoted as \( x_{rx} \), \( y_{ry} \) (producing and directly selling \( x \) and \( r_x \), but buy \( y \)). Some variables are equal to zero. That is, \( x^d = x_{rx} = x^s = x_{rx} = x^E = x_{rx} = y = y^d = y^f = y_{self}^s = y_{self}^r = y_{self}^E = y_{self}^f = y_{self}^r = r_y = l_{rx} = l_y = 0 \)

Their maximization problem is as below:

\[
\begin{align*}
\text{Max. } U^p_x & = xy^d \\
\text{S.t. } & x^p = x + x_{self}^E = l_x - a \\
& x_{self}^s = \min\{kx_{self}^E, r_x\} \\
& r_x = t_d(l_{rx} - b) \\
& l_x + l_{rx} = 1 \\
& p_x x_{self}^s = p_y y^d 
\end{align*}
\]

Solving for this problem, we get

\[
\begin{align*}
y^d & = \frac{(1 - a - b)t_d kp_s}{2(t_d + k)p_y}, \quad x^p = \frac{(1 - a - b)(2t_d + k)}{2(t_d + k)}, \quad x = \frac{1 - a - b}{2}, \quad x_{self}^E = \frac{(1 - a - b)t_d}{2(t_d + k)}, \\
x_{self}^s & = r_x = \frac{(1 - a - b)t_d k}{2(t_d + k)}, \quad l_x = \frac{2(1 - b)t_d + (1 + a - b)k}{2(t_d + k)}, \quad l_{rx} = \frac{2bt_d + (1 + a - b)k}{2(t_d + k)}, \\
U^p_x & = \left(\frac{1 - a - b}{2}\right)^2 \frac{t_d kp_s}{(t_d + k)p_y}, \quad \text{for } 0 < a + b < 1, a > 0, b > 0.
\end{align*}
\]

The trading partners have the configuration of \( y_{ry} \), \( x_{rx} \) that is, producing and directly selling \( y \) and \( r_y \), but buy \( x \). Some variables are equal to zero. That is, \( x = x^d = x^f = y_{self}^s = x^E_s = x_{rx}^E = x_{rx}^f = y^d = y^r = y_{self}^r = y_{self}^E = y_{self}^f = r_x = l_{rx} = l_y = 0 \).

Their decision is as below.

\[
\begin{align*}
\text{Max. } U^p_y & = x^d y \\
\text{S.t. } & y^p = y + y_{self}^E = l_y - a \\
& y_{self}^s = \min\{ky_{self}^E, r_y\} \\
& r_y = t_d(l_{ry} - b) \\
& l_y + l_{ry} = 1 \\
& p_y y_{self}^s = p_x x^d 
\end{align*}
\]

Solving for this problem, we get

\[
\begin{align*}
x^d & = \frac{(1 - a - b)t_d kp_s}{2(t_d + k)p_x}, \quad y^p = \frac{(1 - a - b)(2t_d + k)}{2(t_d + k)}, \quad y = \frac{1 - a - b}{2}, \quad y_{self}^E = \frac{(1 - a - b)t_d}{2(t_d + k)}, \\
U^p_y & = \left(\frac{1 - a - b}{2}\right)^2 \frac{t_d kp_s}{(t_d + k)p_y}.
\end{align*}
\]
For quasi-middlemen of \( r_x \), who produce non-traded goods \( y \) for their own consumption (configuration \( r_y/y/r_x \)), \( x = x^d = x^s = x^r_y = y^d = y^s = y^l = y^l_y = y^l_y \) = 
\[ y^r_y, \quad y^r_y = y^r_y = r_y = l_{ry} = l_{x} = 0. \] Their decision is as below.

Given a Walrasian regime, equilibrium is attainable when (a) utility equilibrium equalisation across these two types of economic agents, and (b) market clearing for both goods \( x \) and \( y \).

For \( U_x^{p_1} = U_y^{p_1} \), we have the corner equilibrium relative price of goods \( x \) to goods \( y \) and the corner equilibrium utility, respectively, \( \frac{p_x}{p_y} = 1 \) and
\[ U_x^{p_1} = \left( \frac{1 - a - b}{2} \right)^2 \frac{t_d k p_y}{(t_d + k) p_x}, \text{ for } 0 < a + b < 1, \]
a > 0, b > 0. Under the market clearing conditions \( M_x x^r_y = M_y x^d \) and \( M_y y^r_y = M_x y^d \), where \( M_x, M_y \) respectively represents the number of producers of goods \( x \) and goods \( y \), we have the relative number of the two types of producers \( M_x / M_y = 1 \).

**Structure P2**

In this structure, producers of goods \( x \) and \( y \) both are fully specialized. This means that both of them do not self-provide transaction services for the supply of goods. Notably, producers of good \( x \) need to purchase good \( y \) for consumption; similarly, producers of goods \( y \) need to purchase goods \( x \) for consumption. Hence, this must involve the exchange for different goods. But the exchange between the two groups of specialized producers must not directly occur; otherwise, it will violate the Claim 1 of the Optimal Configuration Theorem. Accordingly, this exchange needs to be carried out by providers of different transaction services. These providers of different transaction services can be categorized into two groups: one group buys and resells goods \( x \) for the supply of the market through providing transaction services, \( r_x \); the other group buys and resells goods \( y \) for the supply of the market through providing transaction services, \( r_y \). These providers of transaction services may be quasi-specialized. This means that they self-produce one non-traded good for their own consumption. In accordance to Claim 2, the providers of \( r_x \) transaction services will only self-supply good \( y \) for self-consumption. They may be called quasi-middlemen. These four groups of individuals comprise structure \( P2 \), which is partial division of labour. On the other hand, if these service providers are fully specialized, then specialized middlemen of \( r_x \) and \( r_y \) as well as specialized producers of goods \( x \) and \( y \) will make up another structure \( C \). This structure will be investigated later. Now, let’s examine structure \( P2 \).
Max. $U_{rx}^{p_2} = x_{rx}^d y$

S.t. $y^p = y = l_y - a$

$x_{rx}^d = \min\{k y_{rx}^E, r_s\}$

$\mu x_{rx}^d = x_{rx}^{d'} + x_{rx}^E$

$r_s = t_d (l_{rx} - b)$

$l_y + l_{rx} = 1$

$p_x x_{rx}^d = p_x x_{rx}^l$

Solving for this problem, we have

$y^p = y = \frac{1 - a - b}{2}$, $x_{rx}^d = r_s = \frac{(1 - a - b)t_d}{2}$, $x_{rx}^{d'} = \frac{(1 - a - b)t_d}{2k}$, $x_{rx}^l = \frac{(1 - a - b)t_d p_x^l}{2p_x}$

$x_{rx}^d = \frac{1 - a - b}{2} \left( \frac{p_x^l}{p_x} - \frac{1}{\mu k} \right) t_d \mu$, $l_{rx} = \frac{1 - a + b}{2}$, $l_y = \frac{1 + a - b}{2}$

$U_{rx}^{p_2} = \left( \frac{1 - a - b}{2} \right)^2 \left( \frac{p_x^l}{p_x} - \frac{1}{\mu k} \right) t_d \mu$, for $0 < a + b < 1, a > 0, b > 0$.

For quasi-middlemen of $r_y$, who produce non-traded goods $x$ for their own consumption (configuration $r_yx/yr_y$), $x^d = x_{rx}^d = x^s = x_{rx}^{s'} = x_{rx}^E = x_{rx}^l = x_{rx}^d = y = y^d = y^s = y_{self}^E = y_{self}^s = y_{self}^l = r_s = l_{rx} = l_y = 0$. Their decision is as below.

Max. $U_{ry}^{p_2} = xy_{ry}^d$

S.t. $x^p = l_k - a = x$

$y_{ry}^l = \min\{k y_{ry}^E, r_y\}$

$\mu y_{ry}^l = y_{ry}^d + y_{ry}^E$

$r_y = t_d (l_y - b)$

$l_x + l_y = 1$

$p_y y_{ry}^l = p_y y_{ry}^l$

Solving for this problem, we have

$x^p = x = \frac{1 - a - b}{2}$, $y_{ry}^s = r_y = \frac{(1 - a - b)t_d}{2}$, $y_{ry}^{E} = \frac{(1 - a - b)t_d}{2k}$, $y_{ry}^l = \frac{(1 - a - b)t_d p_y^l}{2p_y}$

$y_{ry}^{d'} = \frac{1 - a - b}{2} \left( \frac{p_y^l}{p_y} - \frac{1}{\mu k} \right) t_d \mu$, $l_y = \frac{1 - a + b}{2}$, $l_x = \frac{1 + a - b}{2}$

$U_{ry}^{p_2} = \left( \frac{1 - a - b}{2} \right)^2 \left( \frac{p_y^l}{p_y} - \frac{1}{\mu k} \right) t_d \mu$, for $0 < a + b < 1, a > 0, b > 0$.

For specialized producers of goods $x$ (configuration $x/yr_x, r_y$), $x^d = x_{rx}^d = x_{rx}^{s'} = x_{rx}^l = x_{self}^E = x_{self}^s = x_{self}^l = y = y_{ry}^d = y_{ry}^s = y_{ry}^l = y_{ry}^E = y_{ry}^l = r_s = r_y = l_{rx} = l_y = 0$. Their decision is as below.
Max. $U^p_2 = xy^d$
S.t. $x^p = l_x-a = x + x^i$
\[ l_x = 1 \]
\[ p_x y^d = p'_y y^d \]

Solving for this problem, we have
\[ y^d = \frac{(1-a)p_x}{2p'_y}, \quad x^p = 1-a, \quad x = \frac{1-a}{2}, \quad x^i = \frac{1-a}{2}, \quad U^p_2 = \left(\frac{1-a}{2}\right)^2 \frac{p_x}{p'_y}, \text{ for } 0 < a < 1. \]

For specialized producers of goods $y$ (configuration $y/x/r_y$), $x = x^d = x^i = x^s = x^r = x^l_y = y^d = y^s = y^l = r_x = r_y = l_x = l_y = l_x = 0$. Their decision is as below
Max. $U^p_2 = x^d y$
S.t. $y^p = l_y-a = y + y^s$
\[ l_y = 1 \]
\[ p_y y^s = p'_y x^d \]

Solving for this problem, we have
\[ x^d = \frac{(1-a)p_x}{2p'_y}, \quad y^p = 1-a, \quad y = \frac{1-a}{2}, \quad y^s = \frac{1-a}{2}, \quad U^p_2 = \left(\frac{1-a}{2}\right)^2 \frac{p_y}{p'_y}, \text{ for } 0 < a < 1. \]

Once again, a Walrasian regime is assumed. Under utility equalization conditions between these four configurations $U^p_2 = U^p_2 = U^p_2 = U^p_2$, the corner equilibrium relative buying price of goods $x$ to goods $y$, corner equilibrium relative selling price of $x$ goods to $y$ goods, and corner equilibrium relative selling price to buying price for a same “core” good, as well as corner equilibrium utility are obtained
\[ \frac{p_x}{p_y} = \frac{p'_x}{p'_y} = 1, \quad \frac{p'_x}{p_x} = \frac{p'_y}{p_y} = \left(\frac{1-a}{1-a-b}\right)^2 \frac{1}{t_d \mu} + \frac{1}{4 \mu^2 k^2} + \frac{1}{2 \mu k}, \]
\[ U^p_2 = \left(\frac{1-a-b}{2}\right)^2 t_d \mu \left(\frac{1-a}{1-a-b}\right)^2 \frac{1}{t_d \mu} + \frac{1}{4 \mu^2 k^2} - \frac{1}{2 \mu k} \right), \text{ for } 0 < a + b < 1, a > 0, b > 0. \]

Simultaneously, under the market clearing conditions $M_x x^i = M_{rx} x^d$, $M_y y^s = M_{ry} y^d$, $M_{rx} x^s = M_{rx} x^d$ and $M_{ry} y^r = M_{ry} y^d$, where $M_{rx}, M_{ry}$ respectively represents the number of quasi-middlemen of goods $x$ or goods $y$, with the above obtained corner equilibrium relative prices, we get the relative number of specialized producers of goods $x$ to that of specialized producers of goods $y$ as well as the relative number of quasi-middlemen of goods $x$ to that of quasi-middlemen of goods $y$, $\frac{M_x}{M_y} = \frac{M_{rx}}{M_{ry}} = 1$. We simultaneously get the relative number of quasi-middlemen of goods $x$ (or $y$) to that of specialized producers of goods $x$ (or $y$) as below
\[ \frac{M_{rx}}{M_x} = \frac{M_{ry}}{M_y} = \left(\frac{1-a}{1-a-b}\right)^2 \frac{1}{t_d \mu} + \frac{1}{4 \mu^2 k^2} + \frac{1}{2 \mu k} \right), \text{ for } 0 < a + b < 1, a > 0, b > 0. \]
**Structure C**

In this structure with complete division of labour, all the individuals are divided into four types of specialists, respectively specializing in producing goods \( x \) and \( y \) as well as proving transaction services, \( r_s \) and \( r_y \). This structure mirrors a fully specialized economy, in which trading activities are entirely mediated by specialized mark-up middlemen. In this structure, the degree of trade dependence achieves the highest level. This structure also reflects distribution trade pattern.

For specialized middlemen of \( r_s \) (configuration \( r_s xy r_y \)),

\[
\begin{align*}
    x &= x^d = x^s = x^E_{self} = y = y^d = y^s = y^E_{self} = y^E_{ry} = y^I_{ry} = r_s = l_{ry} = l_y = l_s = 0. \\
    \text{Max.} \ U^C_{rx} &= x^d_{rx} y^d_{rx} \\
    \text{S.t.} \quad x^d_{rx} &= \min \{k x^E_{rx}, r_s \} \\
    \mu x^d_{rx} &= x^d_{rx} + x^E_{rx} \\
    r_s &= t_d (l_{rx} - b) \\
    l_{rx} &= 1 \\
    p'_s x^d_{rx} &= p_s x^d_{rx} + p'_s y^d_{rx} \\
\end{align*}
\]

Solving for this problem, we have

\[
\begin{align*}
    x^d_{rx} &= r_s = (1-b)t_d, \quad x^E_{rx} = \frac{(1-b)t_d}{k}, \quad x^d_{rx} = \frac{1-b}{2} \left( \frac{p'_s}{p_s} + \frac{1}{\mu k} \right) t_d, \quad x^d_{rx} = \frac{1-b}{2} \left( \frac{p'_s}{p_s} - \frac{1}{\mu k} \right) t_d \mu, \\
    y^d &= \frac{1-b}{2} \left( \frac{p'_s}{p_s} - \frac{p_s}{\mu k p'_s} \right) t_d, \quad U^C_{rx} = \left( \frac{1-b}{2} \right)^2 t_d^2 \left( \frac{\mu k p'_s - p_s}{p_s p'_s \mu k^2} \right), \quad \text{for } 0 < b < 1. \\
\end{align*}
\]

For specialized middlemen of \( r_y \) (configuration \( r_y xy r_s \)),

\[
\begin{align*}
    x &= x^d = x^s = x^s_{self} = x^E_{rx} = x^I_{rx} = y = y^d = y^s = y^E_{self} = y^E_{ry} = r_s = l_{rx} = l_y = l_s = 0. \\
    \text{Max.} \ U^C_{ry} &= x^d_{ry} y^r_{ry} \\
    \text{S.t.} \quad y^r_{ry} &= \min \{k y^E_{ry}, r_y \} \\
    \mu y^r_{ry} &= y^r_{ry} + y^E_{ry} \\
    r_y &= t_d (l_{ry} - b) \\
    l_{ry} &= 1 \\
    p'_y y^r_{ry} &= p'_y x^d_{ry} + p_y y^r_{ry} \\
\end{align*}
\]

Solving for this problem, we have

\[
\begin{align*}
    y^r_{ry} &= r_y = (1-b)t_d, \quad y^E_{ry} = \frac{(1-b)t_d}{k}, \quad y^E_{ry} = \frac{1-b}{2} \left( \frac{p'_y}{p_y} + \frac{1}{\mu k} \right) t_d, \quad y^r_{ry} = \frac{1-b}{2} \left( \frac{p'_y}{p_y} - \frac{1}{\mu k} \right) t_d \mu, \\
    x^d &= \frac{1-b}{2} \left( \frac{p'_y}{p_y} - \frac{p_y}{\mu k p'_y} \right) t_d, \quad U^C_{ry} = \left( \frac{1-b}{2} \right)^2 t_d^2 \left( \frac{\mu k p'_y - p_y}{p'_y p_y \mu k^2} \right), \quad \text{for } 0 < b < 1. \\
\end{align*}
\]
For specialized producers of goods $x$ (configuration $x/yr, r$), $x^d = x^d_{self} = x^d_{rx} = x^d_E = x^{d}_{self} = x^{d}_{rx} = y = y^s = y_{self}^s = y_{self}^E = y^d_{ry} = y^d_{ry} = r_x = r_y = l_x = l_y = l_y = 0$. Their decision is as below:

$$\text{Max. } U^C_x = xy^d$$

S.t. $x^p = l_x - a = x + x^s$

$l_x = 1$

$p_x, x^s = p'_y y^d$

Solving for this problem, we have

$$y^d = \frac{(1-a)p_x}{2p'_y}, \quad x^p = 1-a, \quad x = \frac{1-a}{2}, \quad x^s = \frac{1-a}{2}, \quad U^C_x = \left(\frac{1-a}{2}\right)^2 p_x / p'_y, \quad \text{for } 0 < a < 1.$$

For specialized producers of goods $y$ (configuration $y/rx, r$), $x = x^d = x^d_{self} = x^d_{rx} = x^d_E = x^{d}_{self} = x^{d}_{rx} = y = y^s = y_{self}^s = y_{self}^E = y^d_{ry} = y^d_{ry} = r_x = r_y = l_x = l_y = l_y = 0$. Their decision is as below:

$$\text{Max. } U^C_y = x^d y$$

S.t. $y^p = l_y - a = y + y^s$

$l_y = 1$

$p_y, y^s = p'_x x^d$

Solving for this problem, we have

$$x^d = \frac{(1-a)p_y}{2p'_x}, \quad y^p = 1-a, \quad y = \frac{1-a}{2}, \quad y^s = \frac{1-a}{2}, \quad U^C_y = \left(\frac{1-a}{2}\right)^2 p_y / p'_x, \quad \text{for } 0 < a < 1.$$

In equilibrium, under the utility equalization conditions between these four configurations $U^C_x = U^C_y = U^C_{rx} = U^C_{ry}$, we get the corner equilibrium relative buying price of goods $x$ to goods $y$, corner equilibrium relative selling price of goods $x$ to goods $y$, and corner equilibrium relative selling price to buying price for a same “core” good, as well as corner equilibrium utility,

$$\frac{p_x}{p_y} = \frac{p'_x}{p'_y} = 1, \quad \frac{p'_x}{p_x} = \frac{p'_y}{p_y} = \frac{1-a}{(1-b)t_d \sqrt{\mu}} + \frac{1}{\mu k}$$

$$U^C = \left(\frac{1-a}{2}\right)^2 \frac{1-a}{(1-b)t_d \sqrt{\mu}} + \frac{1}{\mu k}^{-1}, \quad \text{for } 0 < a < 1, 0 < b < 1.$$

Simultaneously, under the market clearing conditions $M_x x^s = M_{rx} x^d_{rx}$, $M_y y^s = M_{ry} y^d_{ry}$, $M_{rx} x^d = M_y x^d + M_{ry} x^d_{ry}$ and $M_{ry} y^d = M_x y^d + M_{rx} y^d_{rx}$ with the above obtained corner equilibrium relative prices, we get the relative number of specialized producers of goods $x$ to that of specialized producers of goods $y$ as well as the relative number of specialized middlemen of goods $x$ to that of specialized middlemen of goods $y$, $\frac{M_x}{M_y} = \frac{M_{rx}}{M_{ry}} = 1$. We simultaneously get the relative number of specialized middlemen of goods $x$ (or $y$) to that of specialized producers of goods $x$ (or $y$) as below.
\[ M_{ss} = M_{rv} = \frac{(1-a)\mu k}{2(1-b)t_d + (1-a)\sqrt{\mu k}}, \quad \text{for } 0<a<1, 0<b<1. \]

4. General Equilibrium and Inframarginal Analysis

4.1 Emergence and evolution of mark-up middlemen

In the previous section, corner equilibrium of each of four structures has been solved. Note Structure A does not involve middlemen, Structure P1 and P2 have quasi-middlemen, and Structure C contains specialised middlemen. Now, we conduct general equilibrium and inframarginal analysis. First, let’s restate the corner equilibrium utilities of the four structures as below

\[ U^A = \left(\frac{1-2a}{2}\right)^2, \quad \text{for } 0<a<0.5; \]
\[ U^{P1} = \left(\frac{1-a-b}{2}\right)^2 \frac{t_d k}{t_d + k}, \quad \text{for } 0<a+b<1, a>0, b>0; \]
\[ U^{P2} = \left(\frac{1-a-b}{2}\right)^2 t_d \mu \left(\left(\frac{1-a}{1-a-b}\right)^2 \frac{1}{t_d \mu} + \frac{1}{4 \mu^2 k^2} - \frac{1}{2 \mu k}\right), \quad \text{for } 0<a+b<1, a>0, b>0; \]
\[ U^C = \left(\frac{1-a}{2}\right)^2 \left[\frac{1-a}{(1-b)t_d \sqrt{\mu} + 1}\right]^{-1}, \quad \text{for } 0<a<1, 0<b<1. \]

A general equilibrium is defined as the structure which generates the highest per capita utility (see Yang, 2001). Consequently, we now conduct the comparison of the utilities across structures A, P1, P2 and C. Recall five parameters – \( a, b, k, \mu, \) and \( t_d \) represent fixed learning costs in the production of goods and transaction services, transaction efficiency in the market selling to the end users \( k \) and selling to middlemen \( \mu \), and the technology in the production of transaction services \( t_d \). Since the main focus of this paper is to study the impacts of the specialisation and transaction costs on the emergence and evolution of middlemen, we concentrate on three parameters, \( a, b, \) and \( k \). Some interesting results are summarised in Table 2.

<table>
<thead>
<tr>
<th>( a+b )</th>
<th>( a+b&lt;1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>( a&lt;0.5 )</td>
</tr>
<tr>
<td>( b )</td>
<td>( b&lt;\alpha )</td>
</tr>
<tr>
<td>( k )</td>
<td>( k&lt;k_1 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equilibrium Structures</th>
<th>A</th>
<th>P1</th>
<th>P2</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a+b )</td>
<td>( a+b&lt;1 )</td>
<td>( a+b&gt;1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a )</td>
<td>( a&gt;0.5 )</td>
<td>( a&gt;0.5 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( b )</td>
<td>( b&lt;\alpha )</td>
<td>( 1-\alpha&lt;\beta&lt;1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( k )</td>
<td>( k&lt;k_2 )</td>
<td>( k_2&lt;k&lt;k_3 )</td>
<td>( k&gt;k_3 )</td>
<td></td>
</tr>
</tbody>
</table>

| Equilibrium Structures | P1 | P2 | C |

Where:
\[
k_1 = \frac{(1-2a)t_d}{(1-a-b)^2 t_d - (1-2a)^2}
\]
\[
k_2 = \left[ \frac{1}{t_d} \frac{1}{\mu} - \frac{2(1-a)^2}{(1-a-b)^2} \right] + \left[ \frac{1}{t_d \mu} \left( \frac{1}{t_d} \frac{4(1-a)^2}{(1-a-b)^2} - \frac{4}{\mu} \right) \right] \left[ \frac{2}{t_d} \left( \frac{(1-a)^2}{t_d (1-a-b)^2} - \frac{1}{\mu} \right) \right]^{-1}
\]
\[
k_3' = \frac{(1-2a)^2 (1-a-b)^2 t_d}{(1-a)^2 (1-a-b)^2 t_d \mu - (1-2a)^2}, k_3 = \frac{(1-a-b)(1-b)t_d}{(1-a)((1-b)^2 t_d - (1-a-b)^2)\sqrt{\mu}}.
\]
0 < \mu < 1. t_d > 0.

Let first exam the impacts of \(k\) – the transaction efficiency in the market selling to the end users, \(ceteris paribus\). It can be seen that autarky (structure A) is the equilibrium when transaction efficiency is very low. Conversely, complete division of labour (structure C) is the equilibrium when transaction efficiency is sufficiently high. Between autarky and complete division of labour are structures P1 and P2 with partial division of labour, each of which emerges as the equilibrium when transaction efficiency is moderate. As compared with structure P1, structure P2 occurs

Now let us turn to other two parameters – \(a\) and \(b\), the respective learning costs in the production of goods and transaction services and their relative magnitude. Recall the significance of the fixed learning costs represents the degree of the economies of specialisation – a higher \(a\) and \(b\) indicates a significant economies of specialisation. In particular, when \(a+b<1\), \(a<0.5\), and \(b<a\). \(A \rightarrow P1 \rightarrow P2 \rightarrow C\) represent a smooth evolution path with the improvement in transaction efficiency \(k\). When \(a<0.5\) and \(a\leq b<1-a\), \(A \rightarrow P2 \rightarrow C\) with P1 does not appears. When \(a\geq 0.5\) and \(b<1-a\), \(P1 \rightarrow P2 \rightarrow C\) with A does not appear. When \(a<0.5\) and \(1-a\leq b<1\), \(A \rightarrow C\) and P1 and P2 do not appear. Especially when \(a\geq 0.5\) and \(1-a\leq b<1\), only structure C may be the equilibrium. Intuitively, for these four structures to appear sequentially, the value of \(a\) and \(b\) could not be too higher – otherwise, the benefits from self-providing transaction services for carrying out trading are not enough to cover or even make up for the total learning costs of production and transaction activities.

By the nature of the definition of general equilibrium, per capita utility will increase along with the evolution path.

Figure 2 indicates diverse evolving pathways of specialization and the division of labour with increasing transaction efficiency.

![Figure 2](image-url)
Figure 2: The Emergence and Evolution of Mark-up Middlemen from the Division of Labour

Herein, we formalize the insights by Shaw (1912), Weld (1917), and Kohn (2001a,b and 2003a,b,c) – the merchant appears from the division of labour as an organizer of the market, accompanying the expansion of the market.

Let us exam these structures in details of their respective topological perspectives, structure $P1$ represents the direct trade. This trade pattern is a dispersed transaction form, since it is transacted directly by non-specialized producers. Hence, its degrees of trade interdependence and economics integration are lower. In contrast, structure $P2$ reflects higher degrees of trade interdependence and economics integration, as compared with structure $P1$. $P2$ exhibits the embryonic form of the distribution trade, in which the exchange between different specialized producers is mediated and transacted by quasi-specialized mark-up middlemen, who simultaneously act as non-specialized producers to self-provide a non-traded good. However, it is noted that in structure $P2$ there is no trade between quasi-specialized mark-up middlemen of different goods. Conversely, the trade takes places between specialized mark-up middlemen of different goods in structure $C$. Using graph theory, structure $C$ as a network displays a higher clustering coefficient of $5/6$ which is near 1, than structure $P2$ with $2/3^6$. So, structure $C$ reflects not only fully specialized distribution trade pattern, but also the highest degrees of trade interdependence and economic integration in all the structures. Actually, the distinctions among these structures stem from different levels of division of labour. In addition, it can be recognized from structure $P2$, if as an equilibrium structure, that the presence of specialized producers does not necessarily imply the appearance of specialized mark-up middlemen. Nevertheless, when transaction efficiency achieves an enough high level to ensure complete division of labour between production and trade and between production activities of different goods, specialized producers and specialized middlemen concomitantly arise in structure $C$ and the economy exhibits the distribution trade pattern.

Furthermore, if we define the extent of market as the aggregation of all goods traded in the market, we can see a clear evolution path on the expansion of market. In Structure A, because there are no transactions for goods $x$ and $y$ in autarkic economy, i.e. no market, the extent of the market is $E^A = 0$. In structure $P1$, individuals are both producers and traders, playing a twofold role in the direct exchange economy. The extent of the market – which is a summation of all goods traded in the market, $E^{P1}$ is

$$E^{P1} = M_x y^d + M_y x^d = \frac{M(1-a-b)t_y k}{4(t_d + k)}$$

In structure $P2$, specialized producers emerge, and the extent of the market is

$^6$ In graph theory, a clustering coefficient is a measure of degree to which nodes in a graph tend to cluster together (see Watts, 2003). A higher clustering coefficient indicates a higher degree of connectedness (integration) in a network. See Barabási (2002) and Watts (2003) for some applications of this concept in social sciences.
\[ E^{P2} = M_{rx}x^d + M_{ry}x^d + M_{xy}y^d + M_{yx}y^d = \frac{M(1-a)((1-a)^2 + (1-a-b)\gamma_1 \gamma_2)}{2(1-a-b)(1-a+(1-a-b)\gamma_1 \gamma_2)} \]

where \( \gamma_1 = \sqrt{\frac{(1-a)^2}{1-a-b}} \cdot \frac{1}{\mu k} + \frac{1}{2\mu k} \cdot \frac{p_x'}{p_x} = \frac{p_y'}{p_y} \); \( M \) is the total size of population.

In structures C, both producers and traders are fully specialized and the extent of the market is

\[ E^C = M_{rx}x^d + M_{ry}x^d + M_{xy}y^d + M_{yx}y^d = \frac{M(1-a)[(1-b)t_d\sqrt{\mu + (1-a)\mu}]}{2(1-b)t_d\sqrt{\mu \gamma_2 + (1-a)(\sqrt{\mu - 1})}} \]

where \( \gamma_2 = \frac{1-a}{(1-b)t_d\sqrt{\mu}} \cdot \frac{1}{\mu k} \cdot \frac{p_x'}{p_x} = \frac{p_y'}{p_y} \).

It can be seen that

\[ \partial E^C / \partial k > 0, \partial E^{P2} / \partial k > 0, \partial E^{P1} / \partial k > 0; \]

In summary, we can establish the following proposition:

**Proposition 1:** Given the existence of economies of specialisation, an improvement in transaction efficiency\(^7\) will transform the economy from autarky to partial division of labour and/or to complete division of labour. Such an improvement in transaction efficiency not only leads to the increase in the levels of specialization in both trading and production activities, but also results in the evolution of the trade pattern from non-trade to direct trade between producers and/or to distribution trade mediated by mark-up middlemen. Along with this evolution path, per capital utility will increase which represents the expansion of the market and economic growth. The emergence of specialized mark-up middlemen increases the degrees of trade interdependence, economic integration and commercialization.

### 4.2 Dispersion of wholesale/retail price

Structure P2 and C can be used to demonstrate the impacts of transaction efficiency on wholesale/retail price dispersion. Recall in Structure P2,

\[ \left( \frac{p_x'}{p_x} \right)^{P2} = \left( \frac{p_y'}{p_y} \right)^{P2} = \sqrt{\frac{(1-a)^2}{1-a-b} \cdot \frac{1}{4\mu^2k^2} + \frac{1}{2\mu k}} \]

\(^7\) We use \( k \) to represent transaction efficiency, it should be noted that the above inframarginal comparative statics is also subject to the existence of transaction costs in selling goods to middlemen - namely, \( \mu < 1 \). When \( \mu = 1 \), Structure C will always dominate Structure P1 – therefore, P1 will not appear.
And in Structure C:

\[
\left( \frac{p'_x}{p_x} \right)^C = \left( \frac{p'_y}{p_y} \right)^C = \frac{1-a}{(1-b)t_d\sqrt{\mu}} + \frac{1}{\mu k}
\]

It can be shown that for Structure P2 and Structure C:

\[
\begin{align*}
\partial(p'_x / p_x) / \partial k &= \partial(p'_y / p_y) / \partial k < 0 \\
\partial(p'_x / p_x) / \partial \mu &= \partial(p'_y / p_y) / \partial \mu < 0,
\end{align*}
\]

More interesting, it can be shown that: when \( a+b<1 \), \( a<0.5 \), and \( b<a \),

\[
k > k_3 \equiv \frac{(1-a-b)(1-b)t_d}{(1-a)(1-b)^2t_d - (1-a-b)^2}\sqrt{\mu}
\]

\[
\left( \frac{p'_x}{p_x} \right)^C < \left( \frac{p'_y}{p_y} \right)^{p2}
\]

This leads to the following proposition

**Proposition 2:** For a same “core” good, an improvement in transaction efficiency will result in a decrease in the dispersion between wholesale and retail prices, both within a structure and along with the evolution of market structures.

The decrease in the dispersion of wholesale and retail prices is interesting but hardly unexpected. Recall we assume a Walrasian regime and therefore in equilibrium price reflects total costs. With the improvement of transaction efficiency, there are two driving forces which increase the productivity and reduce the costs in supplying transaction services. The first is the reduction in the costs (lower \( 1-k \) and \( 1-\mu \) in transit. More interestingly, the improvement in transaction efficiency also promotes the specialisation and division of labour – which is the second force to increase the productivity in the provision of transaction services.

Proposition 2 indicates that with the improvement of transaction efficiency, such as the employment of advanced communication and sophisticated transportation system, the price differences between wholesaling and retailing will shrink. This finding is consistent with the work by Nakamura (1999) who shows that the retail revolution facilitated by the rapid computerization of retail transactions is the source of the decrease of retail prices compared with wholesale prices. Atack et al (1994), through citing statistic evidence of North (1966) and Taylor (1951), show that increasing productivity in trading, stimulated by the transportation revolution in roads, canals, steamboats and railways during the period between 1815 and 1860 in America, lowered the price differences of interregional trade. For instance, the use of steamboats shifted agriculture from pioneer self-sufficiency to market oriented production in the western states during this period. The wholesaling price difference of mess pork between New Orleans and Cincinnati slumped from $7.50/Bbl in 1818, to $2.40/Bbl in 1828, and to only $1.25/Bbl by the late 1850s. Over the same period, the price difference of
wheat flour between these two cities was cut about 70%. The opening of the Erie Canal in 1825 had a significant effect in moving western bulk goods to the eastern market. Prior to the construction of this canal, the goods from Cincinnati to New York had to be transported over the Appalachian Mountains, so that the price difference of mess pork between these two places was as much as $9.53/Bbl in 1820. After the use of the waterway, however, this price difference precipitated to $3.48/Bbl in the middle 1830s and over the same period the price difference of flour was cut roughly in half (also see North, 1966, p. 261).

4.3 Relative Number of Mark-up Middlemen

Let us now turn to employment. It can be seen from structures $P_2$ and $C$ that \[ \frac{\partial (M_x / M_y)}{\partial k} > 0, \frac{\partial (M_{x_1} / M_{y_1})}{\partial \mu} > 0, \frac{\partial (M_{x_2} / M_{y_2})}{\partial \mu} > 0, \text{ for } k \in (0,1) \text{ and } \mu \in (0,1). \] Thus, the following proposition can be established:

**Proposition 3:** An improvement in transaction efficiency and/or distribution service technology will lead to a higher proportion of labour working in distribution services sectors.

Proposition 3 is also anticipated. However, this conclusion is based on the expansion of overall market transaction in the wake of transaction efficiency improvement, rather than the assumption that consumer’s preference towards services will increase overtime (for example, Betancourt, 2004)

Proposition 3 is consistent with the research by Jones (1936) and Atack and Passell (1994, p. 523) as cited in the beginning of this paper. Other notable studies include Hartwell (1973) and Czinkota et al (2003). As Hartwell (1973) notes, with advanced transport facilities, the share of distribution sectors in total economic activity rises with economic growth. Distribution trade has become one of the largest sectors of any developed countries. In 1955, for example, there were 14.5 million workforce in eighteen countries of Western Europe working in the distribution sectors, representing about 11% of the total European labour force. Importantly, these countries with highest per capita real incomes (like UK, France, Germany and the Netherlands) had a relatively high proportion of labour in trade, compared with countries of lower average income (like Greece and Spain). In addition, Czinkota et al (2003), by statistic data between the years 1840-2000, also substantiate the increasing role of the service sector in the U.S. economy.

5. Concluding Remarks

In the research we have explored the interplay among transaction efficiency, distribution technique, the relative productivity between trading and production, the relative price between wholesales and retails, the relative number of mark-up middlemen to producers, the number of transaction contacts, the extent of the market, the real income per capita and economic growth. By conducting general equilibrium analysis with inframarginal analytical apparatus of the division of labour, we have developed several propositions which are
consistent with the historical evidences documented in economics and management literature. We show that the mark-up middleman is the product of escalating specialization and evolving division of labour between production activities and trading activities, propelled by the increasing transaction efficiency. The improvement in transaction efficiency transforms the economy from autarky to partial division of labour and to complete division of labour. This transformation essentially results in the evolution of the trade pattern from non-trade to direct trade between non-specialized producers and to distribution trade mediated by mark-up middlemen. This transformation, eventually, characterizes the emergence of specialized mark-up middlemen who play a hub role in effectively and efficiently organizing and coordinating trading activities in the market. The emergence of specialized mark-up middlemen increases the degrees of trade interdependence and economic integration, and enlarges the extent of the market. We verify that economic growth is correlated positively with the extent of the market.

The results of this research not only show the consistence with some existing theoretical, empirical and historical studies relevant to this theme, within the neoclassical framework, but also elucidate the relationship between endogenous variables for economic development and growth. This research can also provide useful implications to marketing strategies. Based on this research, we can develop more models to probe into the world of commerce. For instance, we can build up the model of the commercial firm to explore the interplay between organization hierarchy and market hierarchy, and to investigate the trade-off between specialization and diversification and to explain the nature of commercial contracts. We can also construct a model to explore the underlying relationship between urbanization and commercialization.

For the tractability of the model, we treat transaction costs as an iceberg type exogenous variable - which is not as elegant as the Spulber-type models that explicitly model the game-theoretic search equilibrium. For the same reason, we have incorporated only one type of middlemen rather than the existence of a variety of middlemen observed in the real world – for example, commission-base middlemen, market makers, and etc. Incorporating these features into the current model could shed further lights on the emergence and evolution of not only the quantity of middlemen but also the composition of middlemen. This is highly in our research agenda.
References


