

ISSN 0075-6407

KOBE
ECONOMIC & BUSINESS
REVIEW

49th

ANNUAL VOLUME



RESEARCH INSTITUTE FOR ECONOMICS
AND BUSINESS ADMINISTRATION
KOBE UNIVERSITY

2004

KOBE ECONOMIC & BUSINESS REVIEW

Editor in Chief

Kenji Kojima

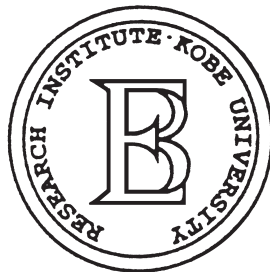
Co-Editors

Junichi Goto; Takashi Kamihigashi; Nobuaki Hamaguchi

ISSN 0075-6407

KOBE
ECONOMIC & BUSINESS
REVIEW

49th
ANNUAL VOLUME



RESEARCH INSTITUTE FOR ECONOMICS
AND BUSINESS ADMINISTRATION
KOBE UNIVERSITY

2004

CONTENTS

	Page
An Economic Analysis of Labor Market Integration in East AsiaJunichi GOTO	1
Implicit Collusion Models of Export Pricing: An Econometric Application to the Japanese CaseEiichi TOMIURA	51
National Money of Account, with a Second National Money or Local Monies as Means of Payment: A Way of Finessing the Zero Interest Rate Bound.Stephen J. DAVIES	69
Product Competitive Advantage and Product Architecture – Value Creation and Value Capture in the Digital Camera Industry –Munehiko ITOH	93

AN ECONOMIC ANALYSIS OF LABOR MARKET INTEGRATION IN EAST ASIA

JUNICHI GOTO Kobe University

Abstract

The purpose of this paper is to present an objective analysis, both theoretical and empirical, of the impact of the labor market integration in East Asia on both host countries and home countries. In addition to the economic impact of migration, the social effect is also discussed.

While international migration generally gives negative impact on the host country, it can give economic benefits at least to sending countries, because it can increase income and employment opportunities in (poorer) sending countries. However, the same economic benefits could be achieved by other measures. Especially, trade liberalization seems to be hopeful. Instead of moving “people” internationally, the international movement of “goods” could increase the income and employment opportunities of sending countries with smaller adjustment cost. The important point is that trade liberalization benefits both host countries and home countries.

JEL Classification: F22

Keywords: International migration, Guest workers, Japan and Asia, Immigration policy

I. Introduction

Asia has been increasingly integrated in recent years. Countries are more and more interconnected with each other in terms of goods, capital, and labor. As many studies suggest, the increase in the amount of intra-Asian trade is dramatic. Indeed, there have been various initiatives toward economic integration in Asia. As early as in 1967, Indonesia, Malaysia, the Philippines, Singapore, and Thailand were united into the Association of South East Asian

Corresponding Address:

Junichi Goto

Research Institute for Economics and Business Administration (RIEB), Kobe University

2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501 Japan

E-mail: jgoto@rieb.kobe-u.ac.jp

Phone / Fax: 81-78-803-7007

Nations (ASEAN). The movement toward trade integration in Asia and the Pacific is intensified in the 1990s. As manifested in the Bogor Declaration of 1994, members of Asia Pacific Economic Cooperation (APEC) agreed to achieve free and open trade (and investment) in the region by 2010 for industrialized countries and by 2020 for developing countries. As for the international movement of capital, the increasing amount of the Japanese foreign direct investment (FDI) is flowing into East Asian countries in recent years, and the FDI from Asian NIEs to other countries in East Asia is becoming more and more important. While it is less dramatic than the integration of movements of goods and capital, the intra-regional integration of labor has also become increasingly important. Although the share of foreign population in total population in East Asia remains a little more than one percent, which is substantially lower than that in North America (8.6 percent) and that in Europe (5.0 percent), the number of intra-Asian migration has increased from one million in the beginning of 1980s to 6.5 million in 1997. Increasing number of people are flowing from less developed Asian countries, such as Indonesia and the Philippines, to more developed Asian countries, such as Japan and Singapore. As discussed in detail below, the increase in migration gives myriad of economic and social effects to both receiving countries and sending countries. For example, the increase in the number of migration has some beneficial impact, e.g., the increase in remittances to sending countries, it also brings about detrimental impact, such as brain drain and the abuse of migrant workers by the employers in the host country.

In view of the above, the purpose of this paper is to present an objective analysis, both theoretical and empirical, of the impact of the labor market integration in East Asia on both host countries and home countries. In addition to the economic impact of migration, the social effect is also discussed.

As become clearer in the discussions below, migration can give economic benefits at least to sending countries, because it can increase income and employment opportunities in (poorer) sending countries. However, it should be noted that the same economic benefits could be achieved by other measures. Especially, trade liberalization seems to be hopeful. Instead of moving "people" internationally, the international movement of "goods" could increase the income and employment opportunities of sending countries.

In Section II, the salient features of migration in East Asia in recent years are examined to obtain some background facts for the theoretical and empirical analysis in the later sections. In view of the fact that many receiving Asian countries, such as Malaysia and Korea, have been trying to repatriate migrant workers in order to save jobs for native workers after the Asian Financial Crisis, the impact of the recent Crisis is also discussed there.

In Section III, the impact of migration on the host country (Japan) is analyzed, because there exist heated, and often emotional, debates on the desirable policies toward immigration. Although traditional economic theorists are generally in favor of migration, at least as far as the economic effect is concerned, because it involves the movement of labor from a labor

abundant country to a labor scarce country, it is demonstrated that such a rosy picture depends on fairly restrictive assumptions, such as perfect competition and no trade barriers, and that, when we relax some assumptions of the traditional theory, the admission of migrant workers could give (or is likely to give) adverse economic effects to the host country. In Section III, a simple trade model, which incorporates trade barriers and non-traded goods, is presented. Using the model, the economic impact of accepting migrant workers on the Japanese economy is analyzed. In addition to the economic effect, some social effects of migration are also discussed there, because migration involves the movement of workers as human being, rather than the movement of labor as a production factor.

Section IV is deals with the economic effect of migration on sending countries. In Section IV, an empirical analysis of migration is made, taking the relationship between Japan and seven countries / area in East Asia (China, Indonesia, Korea, Malaysia, the Philippines, Thailand, and Taiwan province of China) as an example. A simple $2 \times 2 \times 2$ CGE model is developed first, which incorporates migration, FDI, and international trade. Using the model, welfare impacts of migration and trade liberalization are calibrated and compared with each other. The simulation result suggests that, generally speaking, migration is inferior to trade liberalization as a means of increasing the welfare of the sending country.

Section V summarizes major findings of the paper and discusses some agenda for future research.

II. Overview of Asian Migration

1. Salient Features of Asian Migration

Let us briefly examine the recent data on migration in East Asian countries in order to obtain some background facts for the theoretical and empirical analyses of the impact of migration in later sections. First, while the degree of labor market integration in East Asia is not as large as that in North America or in Europe, it has been rapidly increasing since East Asian LDCs started “take-off” of economic development. Table 1 lists the share of foreign population in total population in selected countries in the world. As the table shows, the foreign population share in total East / Southeast Asian countries is only 1.2 percent, which is substantially lower than that in North America (8.6 percent) and that in Europe (5.0 percent). Some internationalized countries (areas) such as Hong Kong (40.0 percent) and Singapore (15.5 percent) are notable exceptions. Recently, however, the degree of labor market integration in Asia has been dramatically increasing, at least until the Asian Financial Crisis. According to the ILO (1998), intra-Asian migration has increased from one million in the beginning of 1980s to 6.5 million in 1997. Major host countries include Japan and the NIES such as Hong Kong, Singapore, and Taiwan, while Indonesia and the Philippines are major

exporters of migrant workers. Korea, Malaysia, and Thailand are both exporters and importers of migrant workers, i.e., they are receiving some types of foreign workers and sending out other types of workers. The increasing trend of Asian migration can be observed from the data both in receiving and sending countries. Table 2 lists the number of registered foreigners in Japan, a typical receiving country, by sending countries. As the table shows, the number of registered foreign residents originating in Asia has increased by 69 percent from 734,476 in 1980 to 1,244,629 in 2000. Although the Asian share in the total number of registered foreign residents is decreased due to the dramatic increase in the foreign residents from South America, it is still as high as 74 percent in 2000. In addition to the legal residents, there are many illegal overstayers, too. According to the estimate by the Japanese Ministry of Justice, the number of illegal overstayers is 251,697 in 2000, almost all of which are from neighboring East Asian countries, such as Korea, the Philippines, China, and Thailand. The increasing trend of intra-Asian migration can also be observed from the data reported by sending countries. Table 3, Table 4, and Table 5 show the number of emigration of workers by destination from Indonesia, the Philippines, and Thailand, respectively. The number of deployed workers from these countries into Asian countries has doubled in a few years. The

Table 1 : Foreign Population in Selected Countries [1990]

	Total population (thousand)	Foreign population (thousand)	Foreigner ratio (%)
East / Southeast Asia	652,927	7,594	1.2
Japan	123,267	1,075	0.9
Korea	42,663	900	2.1
Malaysia	17,670	745	4.2
Singapore	2,690	418	15.5
Taiwan	19,080	1,508	7.9
Hong Kong	5,680	2,271	40.0
Thailand	55,138	314	0.6
North America	276,384	23,868	8.6
Canada	27,606	4,266	15.5
United States	248,778	19,603	7.9
Europe	498,740	24,908	5.0
France	56,563	5,897	10.4
Germany	79,195	5,037	6.4
Switzerland	6,804	1,092	16.0
United Kingdom	57,332	3,718	6.5
Latin America	402,285	6,550	1.6
Argentina	32,325	1,675	5.2
Brazil	147,134	1,138	0.8
World	5,926,830	120,000	2.0

Source: United Nations, Japanese Ministry of Labor, U.S. Census Bureau

increase in Indonesian emigrant workers deployed in Asia is dramatic, i.e., it shows more than a five-fold increase in just two years (from 68,436 in 1995 to 375,383 in 1997). Asia and the Middle East are the two major destinations for these workers, and the importance of destinations in Asia is increasing. For example, as shown in Table 4, while the number of Filipino workers deployed in the Middle East has hardly increased in the 1990s, those

Table 2 : Registered Foreigners in Japan, by Origin

(number)	1970	1980	1990	2000
Asia	672,280	734,476	924,560	1,244,629
North America	1,290	2,719	71,495	58,100
South America	20,892	24,743	44,643	312,921
North Europe	11,902	15,897	25,563	47,730
Oceania	1,044	1,561	5,440	12,839
Africa	232	795	2,140	8,214
Total	707,640	780,191	1,073,841	1,686,444
(share, %)	1970	1980	1990	2000
Asia	95.00	94.14	86.10	73.80
North America	0.18	0.35	6.66	3.45
South America	2.95	3.17	4.16	18.56
North Europe	1.68	2.04	2.38	2.83
Oceania	0.15	0.20	0.51	0.76
Africa	0.03	0.10	0.20	0.49
Total	100.00	100.00	100.00	100.00

Source: Japanese Ministry of Justice

Table 3 : Indonesian Workers Abroad by Destination

	1995		1997	
	number	share (%)	number	share (%)
Asia Pacific	68,436	56.7	375,383	74.6
Brunei	832	0.7	2,426	0.5
Hong Kong	4,205	3.5	2,019	0.4
Japan	1,438	1.2	3,245	0.6
Korea	6,732	5.6	8,390	1.7
Malaysia	29,712	24.6	317,685	63.2
Singapore	20,975	17.4	31,928	6.3
Taiwan	4,106	3.4	9,445	1.9
Other	436	0.4	245	0.0
America	3,483	2.9	736	0.1
Europe	64	0.1	577	0.1
Middle East & Africa	47,524	39.4	116,844	23.2
Other	1,160	1.0	10,014	2.0
Total	120,603	100.0	502,977	100.0

Source: Scalabrini Migration Center

deployed in Asian countries has doubled during the same period. For Thailand, the Asian share in the total number of deployed workers is as high as eighty percent.

Most of the foreign workers in Asia are temporary migrants, who go to the foreign country for a short period of time to make money, leaving their family in their home country. In order to support their family left in the home country, these emigrated workers remit substantial part of their incomes to the home country. So, the remittances from those workers are very important not only for their family but also for the balance of payments of the home country. As Table 6 shows, for major sending countries, remittances from emigrated workers play a very important role in foreign-exchange receipts of the home country. For example, in

Table 4 : Philippines: Land-based Workers, by Destination

	1900		1996	
	number	share (%)	number	share (%)
Asia	90,768	27.1	174,308	36.0
Hong Kong	34,412	10.3	43,861	9.0
Japan	41,558	12.4	20,183	4.2
Malaysia	4,397	1.3	12,340	2.5
Singapore	4,698	1.4	15,087	3.1
Taiwan	54	0.0	65,464	13.5
Other Asia	5,649	1.7	17,373	3.6
Europe	6,853	2.0	11,409	2.4
Americas	9,557	2.9	8,378	1.7
Middle East	218,110	65.1	221,224	45.6
Other	9,595	2.9	69,334	14.3
Total	334,883	100.0	484,653	100.0

Source: Philippines Overseas Employment Administration

Table 5 : Thailand: Deployed Overseas Workers, by Destination

	1993		1996	
	number	share (%)	number	share (%)
Asia 5	94,038	82.4	148,831	80.3
Singapore	1,664	1.5	17,601	9.5
Brunei	14,403	12.6	20,714	11.2
Hong Kong	5,398	4.7	4,301	2.3
Japan	5,682	5.0	10,118	5.5
Taiwan	66,891	58.6	96,097	51.8
Middle East	16,906	14.8	22,425	12.1
Other	3,118	2.7	14,180	7.6
Total	114,062	100.0	185,436	100.0

Source: Thai Government

Pakistan, the amount of remittances is almost equal to the total value of exports in the 1980s. As discussed in the next section, however, the amount of remittances is unstable, and it tends to decrease in recession years, when sending countries need foreign exchange most.

It seems that most migrant workers are unskilled workers. Table 7 lists the deployment of Filipino migrants by skill category. As the table shows, the majority is unskilled workers, such as maid and production workers. The share of professional workers is only 23.1 percent in 1997, half of which are “entertainers.” It is believed, at least in Japan, that many, perhaps most, of such “entertainers” are not typical entertainers like singers or actors but workers in the notorious sectors of the economy. The share of professional workers other than entertainers was mere

Table 6 : Share of Remittances to GDP and Merchandise Exports (%)

	Year	GDP	Merchandise Export
Philippines	1980	1.9	10.6
	1985	2.6	17.4
	1990	3.3	17.8
	1993	4.7	22.3
Bangladesh	1980	2.2	36.1
	1985	4.0	50.2
	1990	3.7	46.6
	1993	4.2	44.1
India	1980	1.6	32.7
	1985	1.1	25.6
	1990	0.7	12.4
	1993		
Pakistan	1980	8.9	82.1
	1985	8.7	97.2
	1990	5.5	40.4
	1993	3.3	23.7
Sri Lanka	1980	3.5	13.1
	1985	3.9	17.7
	1990	4.6	19.9
	1993	5.4	19.8

Source: Asian Development Bank

Table 7 : Deployment of Filipino Migrants by Skill Category (New Hires)

	1992		1997	
	number	share (%)	number	share (%)
Professional	72,848	28.0	51,228	23.1
Entertainers	49,996	19.2	25,636	11.6
Other	22,852	8.8	25,592	11.6
Administrative	495	0.2	555	0.3
Clerical	4,943	1.9	3,534	1.6
Sales	2,725	1.0	2,560	1.2
Services	82,440	31.6	76,402	34.5
Maid	58,700	22.5	47,544	21.5
Caretakers	11,399	4.4	19,225	8.7
Other	12,341	4.7	9,633	4.4
Agricultural	1,920	0.7	538	0.2
Production	94,525	36.3	83,560	37.7
Other	698	0.3	3,027	1.4
Total	260,594	100.0	221,404	100.0

Source: Scalabrini Migration Center

Table 8 : Number of Migrant Workers in Selected Country in East Asia

	legal	illegal	total	illegal share (%)
Korea	149	148	297	49.8
Taiwan	268	20	288	6.9
Thailand	316	845	1,161	72.8
Malaysia	1,121	800	1,921	41.6
Total of above	1,854	1,813	3,667	49.4

Source: Compiled from data in Battistella and Asia, and The Scalabrini Migration Center

11.6 percent in 1997 (8.8 percent in 1992). While data by skill categories are limited, we can observe similar trends in other countries, too. For example, out of two million foreign workers in Malaysia, the number of skilled workers is only 36,498 (Battistella and Asis, p. 22), or less than two percent.

Since the international flow of legal migrants is severely restricted by receiving countries, the number of illegal migration is substantial. Although it is difficult to accurately determine the number of illegal migrant workers, existing data suggest that there are many illegal workers in East Asia. Table 8 lists the numbers of legal migrants and illegal migrants in selected Asian countries. The illegal shares in total migration in Thailand, Korea, and Malaysia are as high as 72.8 percent, 49.8 percent, and 41.6 percent, respectively. Due to the illegal status, their working conditions are often worse than those of native workers and legal migrant workers.

2. Migrant Workers in Japan

(1) Three Categories of Migrant Workers

(a) Legal Skilled Workers

Table 9 summarizes the number of migrant workers by visa categories in Japan. The number of legal and skilled workers is very small at 154,748 or about 0.2 percent of total labor force in Japan. Even when unskilled workers, mostly from Latin America, are included, the number of legal migrant workers is 392,355, or just 0.58 percent of total labor force. This is partly due to the strict Japanese immigration law, which severely restricts jobs that foreign workers can take in Japan. Therefore, most legal foreign workers, except for the Latin Americans of Japanese origin called Nikkei (see below for the detailed discussions of Nikkei), are professional workers, such as professors, researchers, lawyers, accountants etc. The share of legal foreign workers in total labor force in Japan is far smaller than those in European countries. The share of migrant workers, including illegal migrants, in the total labor force in Japan is less than one percent, while the share of foreign workers in total labor force is

around seven percent in France and Germany and as high as seventeen percent in Switzerland.

While the number of real skilled and professional migrant workers in Japan is small and there are no rapidly increasing trend in recent years, there are two groups of migrant workers that have dramatically increased in recent years: (i) illegal unskilled workers from neighboring Asian countries and (ii) legal Latin American workers of Japanese origin, who are often called *Nikkei (or Nikkeijin)* workers. In what follows, these two groups of migrant workers will be discussed in detail.

(b) Illegal Unskilled Workers

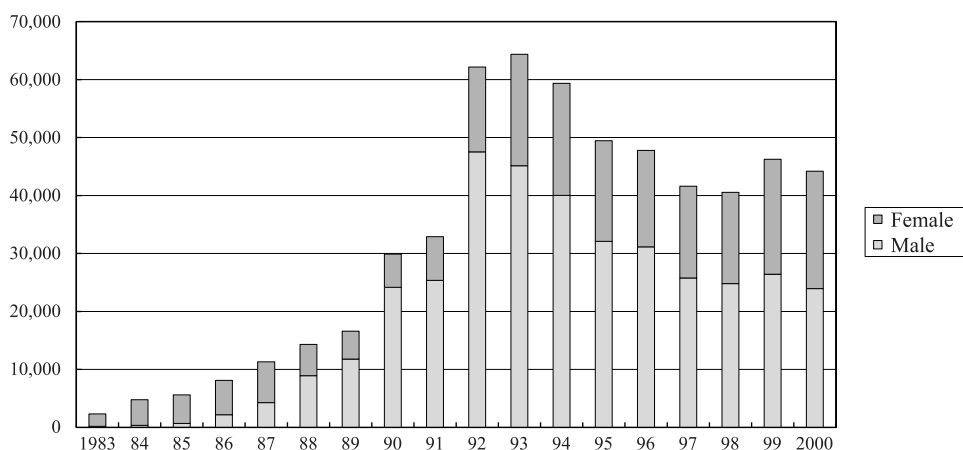
Although the number of migrant workers (both legal and illegal) in Japan is less than a million, or less than one percent of her labor force (see Table 9), the rate of increase in the number of illegal foreign workers has been dramatic since the mid-1980s, at least until the severe economic recessions in the 1990s. As Figure 1 shows, the number of illegal foreign workers apprehended by the authorities has sharply increased from 2,339 in 1983 to 64,341 in 1993. Although the number declined a little after that due to the severe recession of the Japanese economy, the number of illegal migrant workers are far greater than the level before mid-1980s. Needless to say, these

Table 9 : Foreign Workers in Japan (2000)

Legal foreign workers	392,355
Skilled workers	(154,748)
Migrant workers of Japanese origin	(237,607)
Illegal foreign workers	232,121
Total	624,476
Total Labor Force In Japan	67,660,000

Source: Japanese Ministry of Labor

Fig. 1 ILLEGAL FOREIGN WORKERS APPREHENDED



Source: The Japanese Ministry of Justice

numbers represent only a small part of the total illegal foreign workers in Japan. According to the Ministry of Justice, the total number of illegal foreign workers in Japan is estimated at 200, 000-300, 000.

Just as most illegal aliens in the United States come from Mexico and other neighboring countries in Latin America and the Caribbean, most illegal foreign workers in Japan come from neighboring Asian countries (See Table 10). Since the wage rate in their home countries is extremely low, even a discriminatory low wage by the Japanese standard means a lot to these workers.

The recent influx of Asian workers is markedly different from earlier migrations. Until the middle of the 1980s, most of the illegal foreign workers were women who worked as bar hostesses (so-called “*Japayuki San* (Miss Japan-going)”). In 1983-84, for example, more than

Table 10 : Illegal Foreign Workers by Country of Origin (2000)

	Number			Share (%)		
	Total	Male	Female	Total	Male	Female
Korea	13,211	4,802	8,409	100.0	36.3	63.7
China	9,437	6,018	3,419	100.0	63.8	36.2
Taiwan	700	167	533	100.0	23.9	76.1
Hong Kong	52	24	28	100.0	46.2	53.8
Philippines	8,974	3,129	5,845	100.0	34.9	65.1
Thailand	4,459	1,633	2,826	100.0	36.6	63.4
Indonesia	1,498	1,071	427	100.0	71.5	28.5
Iran	1,872	1,829	43	100.0	97.7	2.3
Malaysia	1,239	832	407	100.0	67.2	32.8
Bangladesh	1,147	1,109	38	100.0	96.7	3.3
Peru	1,534	909	625	100.0	59.3	40.7
Pakistan	1,207	1,171	36	100.0	97.0	3.0
Other	6,129	4,462	1,667	100.0	72.8	27.2
Total	51,459	27,156	24,303	100.0	52.8	47.2

Table 11 : Illegal Foreign Workers by Activities (2000)

Male			Female		
	Number	Share (%)		Number	Share (%)
Construction worker*	7,290	30.4	Bar hostess*	9,056	44.7
Factory worker	6,956	29.0	Factory worker	2,510	12.4
Unskilled help*	1,830	7.6	Waitress*	2,495	12.3
Cook*	1,650	6.9	Other service*	1,373	6.8
Bartender*	1,561	6.5	Dish washer*	1,214	6.0
Other service*	954	4.0	Cook*	896	4.4
Other	3,708	15.5	Other	2,697	13.3
(Nontraded)	13,285	55.5	(Nontraded)	15,034	74.3
Total	23,949	100.0	Total	20,241	100.0

(Note) “Nontraded” is a summation of items with an asterisk.

90 percent of the illegal foreign workers were female. But, the number of male workers dramatically increased to about 50-80 percent of the total illegal immigrants by 1990. (See Figure 1)

As seen in Table 11, in 2000, about 30 percent of the illegal male workers were construction workers, and 45 percent of female workers are bar hostess. Most of the illegal aliens are doing work that few Japanese want to do because of unfavorable working conditions. It should be noted that about two-thirds of the illegal migrants are working in the nontraded goods sector, such as construction and service industries.

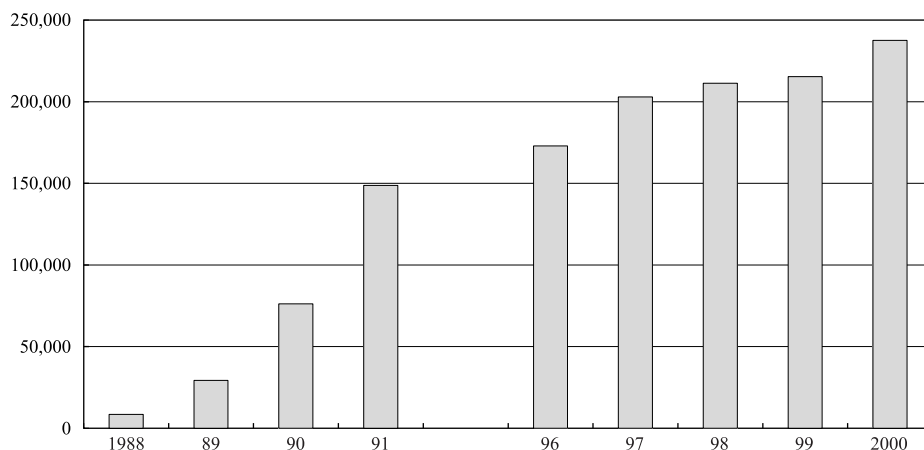
(c) Legal Unskilled Workers — Migrant Workers of Japanese Origin (Nikkei)

(i) The impact of the revision of the Japanese immigration law in 1990

In addition to the illegal foreign workers discussed above, there has been a dramatic increase in the number of Latin American workers of Japanese origin (*Nikkei*) since late 1980s. The influx of these workers is mainly due to the revised immigration law in Japan, which was enacted in 1989 and was put into effect in June 1990. While the Japanese immigration law does not allow foreigners to take an unskilled job in principle, the revised law made it possible for “a foreign citizen whose parent or grandparent was a Japanese citizen” to do whatever activities (including unskilled work) in Japan. Further, these people are allowed to stay in Japan for three years (instead of three months for visitors).

The revision of the immigration law resulted in the dramatic increase of workers from

Fig. 2 Migrant Workers of Japanese Origin



Source: The Japanese Ministry of Justice

Latin American countries such as Brazil, Argentina, Peru, and Bolivia, to which many Japanese citizens had emigrated long time ago. Since the wage rate in Japan is much higher (and Japanese society is much safer) than that in Latin America, a host of Latin American people of the Japanese origin (*Nikkei*) were attracted to Japan. According to the newspaper reports, some people fell into huge debt to pay for their travel cost, and other people without Japanese origin were arrested for forgery of their birth certificate or using other person's identification. Figure 2 shows the number of these *Nikkei* workers since 1988. While the number of *Nikkei* workers in Japan was only eight thousand at the end of 1988, the year before the revised immigration law was enacted, after that the number of *Nikkei* workers doubled every year to become around 150,000 in June 1991. In 2000, the number of *Nikkei* staying in Japan is 237,607.

(ii) Single male in car parts factory

Then, what are salient features of these *Nikkei* workers coming from Latin America to Japan? The Ministry of Labor in Japan published a result of the survey on *Nikkei* workers in Japan, and the following discussion is mainly based on the survey result.

According to the survey, most of them are young male: About two thirds of them are male, and the majority of them are under thirty years old. About two thirds are coming to Japan by himself, and only thirteen percent of them bring their entire family to Japan.

Almost all (more than ninety percent) of these males are working as production workers in manufacturing sector. About one third of them are employed in transport equipment production sector (most of them are car parts factory). This constitutes a striking contrast with illegal unskilled workers from Asia, who are employed mainly in the nontraded good sectors such as construction and services.

(iii) Conditions of *Nikkei* workers in Japan

What are the working conditions of these *Nikkei* workers in Japan? When we compare hourly wage of the *Nikkei* workers with that of the Japanese counterpart, there are no big difference between the two. Since they are legally employed, employers tend to pay them by regular hourly wages. However, the annual income of these *Nikkei* workers is much smaller than that of Japanese workers, because most of the *Nikkei* workers are employed on daily basis and paid by hourly wage. Note that, in terms of wage structure, blue-collar production workers in Japan are similar to white-collar workers in the United States. Most of the blue-collar production workers in Japan receive monthly salary (instead of hourly wages). Further, these salaried workers in Japan receive bonus payments twice a year, and enjoy various fringe benefits. The bonus payment in Japan constitutes a substantial part of their

annual income, and the amount of average bonus is equal to five months salary. Therefore, even though hourly wage rate of *Nikkei* workers is similar to that of the Japanese counterpart, their annual income is much smaller than Japanese workers because most of *Nikkei* workers do not receive bonus payment and fringe benefit.

Further, the *Nikkei* workers are often exploited by brokers or mediators. According to the survey result, less than half of *Nikkei* workers are directly employed by the firm where they actually work. More than half of *Nikkei* workers are employed by mediator agencies and are sent by them to the factories. Hence, the amount of income of *Nikkei* workers is usually smaller than what the factories are paying for their work. In some cases, exploitation by gangsters was also reported.

Further, the above survey reveals that most of the *Nikkei* workers are working without medical insurance and unemployment insurance. In Japan, almost hundred percent of workers (and their family) are covered by very generous government sponsored medical insurance program. Workers contribute to the medical insurance program according to their income. And, the deductibles are nominal. No matter how high the actual medical cost is, the payment out of the patient pocket does not exceed about five hundred dollars a month and the balance is paid by the insurance program. Very poor people can enjoy the same benefit as the others without contributing anything.

However, according to the above survey, only twenty-three percent of the *Nikkei* workers are covered by the generous medical insurance program, probably because both employers and *Nikkei* workers themselves do not want to pay their contribution to the program. But, in case of illness or accident, these *Nikkei* workers have to pay 100 percent of their medical expenses from their pocket.

(2) Reasons for the Sharp Increase — Push and Pull

Why did many unskilled foreign workers suddenly come to Japan after the middle of the 1980s? While it is clear that the most important reason for the increase in the *Nikkei* workers from Latin America is the revision of the Japanese immigration law in 1990, the reasons for the influx of illegal foreign workers from neighboring Asian countries are not so obvious.

One of the most important reasons is that a push-force in neighboring Asian countries coincided with a pull-force in the Japanese economy in the 1980s. Probably, one of the most important reasons on the supply side for the sharp increase is that the destination of Asian migrant workers has shifted from the Middle East to Japan. In the 1970s, an increasing number of Asians had been recruited to work at construction sites in the oil producing Middle Eastern countries. When the price of crude oil quadrupled after the First Oil Crisis in 1973, a construction boom occurred in the oil-rich countries because their oil revenue dramatically increased. But the population size of these countries in the Middle East is

relatively small. Therefore, these rich countries recruited a large number of temporary immigrants mostly from southern Europe and Asia. As a result, the number of migrant workers from eight Asian countries (Bangladesh, India, Indonesia, Pakistan, Sri Lanka, the Philippines, Thailand, and South Korea) to the Middle East grew from a little more than 0.1 million in 1976 to more than 1.2 million in 1982. But, as the price of crude oil went down in the 1980s, the construction boom in the Middle East subsided, and some 400 thousand Asian migrant workers lost their jobs and had to return to their home countries.

The return of these workers was a serious blow to the Asian sending countries, because remittance from them was an important source of foreign exchange receipt. In Pakistan and Bangladesh, for example, remittance from migrant workers was almost as large as the total value of their exports. Moreover, the dependence of the migrant workers from these Asian countries on the Middle East for their destination was extremely heavy. Therefore, the decline in labor demand in the Middle East created a large pool of Asian workers who lost jobs in the Middle East and were eager to find new jobs in some other countries. Probably, to these unemployed workers, one of their rich neighbors, Japan, must have looked like a new land of opportunity.

The increase in the supply pressure coincided with the increased demand of Japanese businesses for the migrant workers. Due to the strong performance of the Japanese economy, the labor market in Japan became very tight since the mid-1980s. The labor shortage was especially keen in the construction and service industries. Moreover, an important source of the domestic supply of marginal workers (i.e., a group of seasonal workers called "*dekasegi*") has shrunk, and therefore, the demand for migrant workers to fill the gap in this marginal labor market increased.

The performance of the Japanese economy after the middle of the 1980s was dramatic: the annual growth rates of the real GNP in 1988, 1989, and 1990 were 6.2%, 4.7%, and 5.6%, respectively; those of industrial production in 1988, 1989, and 1990 were 9.5%, 6.1%, and 4.6%, respectively. Consequently, the labor market became very tight, and the job-opening / job-seeker ratio (one of the most commonly used indicators of the labor market condition in Japan) sharply increased. While the ratio nose-dived in 1975 (the First Oil Recession) and stayed at around 0.6 (i.e., only six jobs were available for every 10 job seekers), it began to increase after 1987. In 1988, the ratio exceeded one for the first time since 1974, and it went as high as 1.40 in 1990.

A typical practice of Japanese firms in boom years has been to increase the number of marginal workers, such as seasonal and temporary workers, because firms had at least a moral obligation to keep their regular employees on the payroll even in a recession. However, the construction industry is losing an important source of its supply of domestic nonregular workers. Until the end of the 1970s, the labor shortage in the construction industry in boom years had been largely filled by seasonal workers, *dekasegi*, who were mostly farmers in the

northern part of Japan who came to metropolitan areas like Tokyo and Osaka to take temporary jobs in an attempt to supplement their farm incomes in the farmers' slack season. In the early 1970s, the number of dekasegi amounted to about 600,000. But, because of increased job opportunities in their home towns, that number has been steadily diminishing: only 142,200 dekasegi were reported in 1993. The decline in the supply of dekasegi, along with the recent construction boom, created a serious labor shortage in the construction industry. The strong demand for marginal workers in the Japanese construction industry attracted an increasing number of foreign workers whose supply pressure had been increased by the decreasing demand in the Middle East.

Faced with the strong push-force and pull-force, illegal mediators between Japanese employers and Asian migrants (like the "coyote" figure for Mexican illegal aliens in the United States) have become prevalent. Although the details of their illegal activities are unknown, involvement of gangsters was often reported. According to an estimate by the Japanese Ministry of Justice, in 1990 about 70 percent of illegal migrant workers entered Japan with the help of such illegal mediators.

3. Impact of Asian Financial Crisis

The increasing trend of intra-Asian migration discussed above seems to be suspended, at least temporarily, by the Asian Financial Crisis started in July 1997. In order to cope with the dramatic depreciation of their currencies and inflationary pressures, many countries were obliged to take deflationary monetary and fiscal policies. As a result, many countries, perhaps with the exception of Taiwan, have been suffering from negative economic growth and high unemployment. Due to the severe recession, unemployment rate more than doubled in many countries, including Hong Kong, Indonesia, Korea, Malaysia, and Thailand in one year from 1997 to 1998. In Indonesia, the unemployment rate in 1998 (20.0%) is five times higher than that in 1997 (4.7%). Due to the severe unemployment problem, there are some changes in government policies towards migration in both receiving and sending Asian countries.

Receiving countries, such as Korea, Malaysia, Singapore, and Thailand,¹ took various measures to reduce the number of immigrant workers in order to save employment for native workers. For example, at the end of 1997, the governments of Malaysia and Thailand announced the plan to repatriate one million immigrant workers to save domestic jobs. In order to reduce the number of immigrant workers, especially illegal immigrants, various measures were taken by the governments of receiving countries in East Asia. First, the enforcement of immigration laws was strengthened in most countries. Second, amnesty programs taken by

¹ Thailand is both receiving and sending countries, and the Thai government took restrictive policies toward immigration and encouraged outflow of Thai workers in order to relieve the unemployment problem in the country.

Korea and Malaysia, coupled with the threat of stiff punishment after the amnesty period, brought about a large number of exodus of illegal migrant workers. Third, in the hope of encouraging employers to shift from foreign workers to native workers, some countries, including Malaysia and Singapore, imposed fees (or increased existing fees) against migrant workers. Through various restrictive measures, a large number of migrant workers were forced out of many East Asian countries. As shown in Table 12, from 1997 to mid-1998, the number of migrant workers decreased sharply in Korea (43.8% decline), Thailand (36.5% decline), Malaysia (23.5% decline), and Singapore (11.1% decline) in less than a year. In addition to the decline in the number of migrant workers, worsening of working conditions has been reported.²

While receiving countries took various restrictive measures to reduce immigration, sending countries, such as Indonesia, the Philippines, and Thailand, strengthened their efforts to increase emigration to mitigate unemployment problems in their countries. For example, in the beginning of 1998, Thailand announced the objective of sending out 210,000 Thai workers abroad. Similarly, the Filipino government suspended their policy to reduce outflow of unskilled workers, and, instead, encouraged the outflow of workers. Due to such efforts, the total number of emigration does not seem to have decreased at all after the Crisis, in spite of the decline in the emigrated workers destined to East Asia. Table 13 lists the outflow of Filipino workers by destination in 1997 and 1998. As shown in the table, although the number of deployment (flow data) of Filipino workers in Asia decreased by 13,872 (or six percent decline), the deployment in all other regions increased. In other words, in spite of the substantially decreased job opportunities in East Asia, the total number of emigration from the Philippines actually increased from 1997 to 1998. However, the increase in deployed overseas workers did not bring about the increase of remittances from these workers. As Table 14

Table 12 : Migrant Workers before and after the Asian Financial Crisis

	Pre-crisis 1997 (thousand)	Post-crisis mid-1998 (thousand)	Change	
			number (thousand)	percent (%)
Hong Kong	300	300	0	0
Japan	700	680	-20	-2.9
Korea	267	150	-117	-43.8
Malaysia	1,700	1,300	-400	-23.5
Singapore	450	400	-50	-11.1
Thailand	1,260	800	-460	-36.5
Total of above	4,677	3,630	-1,047	-22.4

Source: ILO (1999), "Toward Full Employment: Prospects and problems in Asia and the Pacific"

² For example, for maids in Hong Kong, 20 percent wage reduction and 16 hours daily work were proposed in 1998.

shows, the amount of remittances decreased by 14.2 percent from 5,742 million dollars in 1997 to 4,925 million dollars in 1998, in spite of the increase in the number of deployed overseas workers. This suggests that wages (and other working conditions) of workers deployed overseas have been adversely affected by the Asian Financial Crisis. In fact, it is reported that some governments allowed the deployment of migrant workers in high-risk countries in an effort to increase the number of total deployment after the Crisis.

III. The Impact of Migration on the Host Country

In this section, the impact of migration is theoretically (and empirically to some extent) analyzed. While the theoretical model here is fairly general and applicable to many situations in many countries, I relate the theoretical analysis with the real world example, taking the Japanese experience in recent years.

1. Conventional Wisdom on the Economic Effect of Migration

Traditional economic theorists usually consider that the overall effect of international migration is favorable to both home and host countries, because it involves a movement of labor from a labor abundant (and capital scarce) country to a labor scarce (and capital abundant) country, and therefore, it will increase productivity (and economic welfare) in both

Table 13 : Filipino Overseas Worers by Region

	1997	1998	Change	
			number (thousand)	percent (%)
Africa	3,517	5,548	2,031	57.7
Asia	235,129	221,257	-13,872	-5.9
Americas	7,058	8,210	1,152	16.3
Europe	12,626	15,682	3,056	24.2
Middle East	221,047	226,803	5,756	2.6
Oceania	1,970	2,061	91	4.6
Other	77,880	82,823	4,943	6.3
Total	559,227	562,384	3,157	0.6

Source: Philippines Overseas Employment Administration

Table 14 : Overseas Filipino Workers' Remittances

(US \$ million)

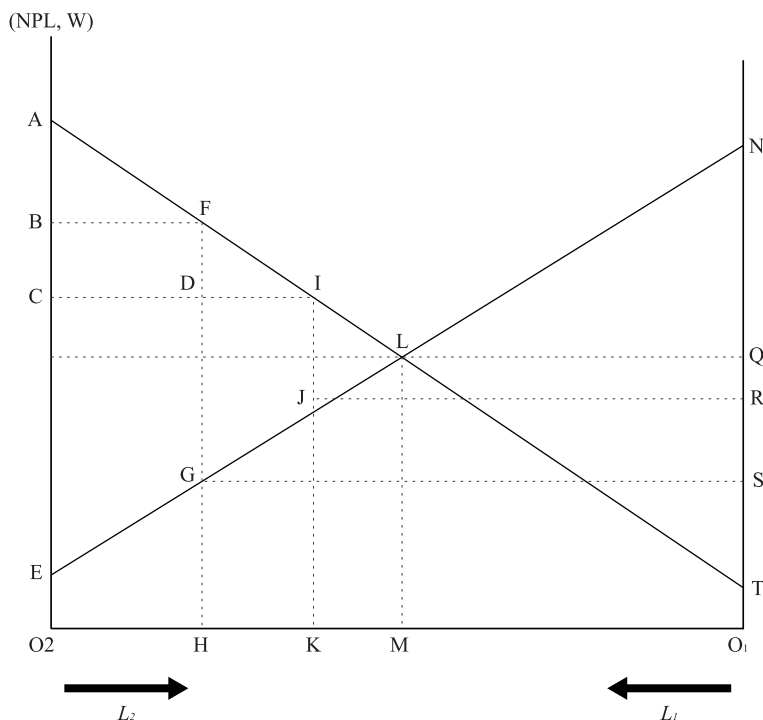
	Landbased	Seabased	Total	% change
1997	5,484	258	5,742	
1998	4,650	275	4,925	-14.2

Source: Internet site of Scalabrini Institute

countries. For example, when some workers move from the Philippines (a labor abundant country) to Japan (a labor scarce country), the Japanese employers who have been suffering from unfilled vacancy can gain from hiring these workers and the Filipino workers can earn more than what they could earn in the Philippines. If these workers remit some part of their income earned in Japan to their home country, people left behind in the Philippines are also benefited from the migration of their fellow Filipinos indirectly. Of course, there could be some conflict of interests among various economic agents in each country. For example, an inflow of Filipino workers may give dampening effect on Japanese wages, and thereby the income of Japanese workers could decrease while income of employers in Japan increases. But, overall effect is positive in both countries. Thus, the movement of workers (or unemployed persons) from the home country to the host country would increase national incomes (and economic welfare) in both countries.

The economic reasoning for their argument of the economic gain is summarized in Figure 3. In the figure, the horizontal axis plots the amount of labor supply, where the amounts of labor supply in country 1 (home country) and that in country 2 (host country) are measured from O_1 and O_2 , respectively. Vertical axis plots marginal productivity of labor

Fig. 3 ECONOMIC EFFECT OF MIGRATION



(MPL), which is equal to wage rate in the competitive equilibrium. MPL of labor in country 1 (country 2) is expressed by line NE (by line AT).

Suppose that at the initial stage before migration the labor endowment in country 1 (home country) is O_1H and that in country 2 (host country) is O_2H , and therefore labor supply in the two countries as a whole is O_1O_2 . At this stage, the value of total production (i.e., national income) of the sending home country is the area of trapezoid $NGHO_1$, and the value of the national income of the receiving host country is the area of trapezoid $AFHO_2$. In this pre-migration situation, the wage rate in the host country is BO_2 , which is higher than that in the home country (SO_1). Such wage gap between the two countries constitutes an incentive for the workers in country 1 to migrate to country 2.

Now, suppose that some workers in country 1 (the number of workers expressed by HK) migrate to country 2 in order to seek higher wage there. In this post-migration situation, the amount of labor that can be mobilized for the production in country 2 is increased to O_2K , and that in country 1 is decreased to O_1K , because the labor HK moves from country 1 to country 2. Now, the value of goods produced in country 2 (Gross Domestic Product, GDP) increases to the area of trapezoid $AIKO_2$. But, the area of rectangle $DIKH$ is paid as wage to the workers from country 1, and the net gain of income of host country's citizen is equal to the area of triangle FDI . While the GDP in country 1 is decreased to the area of trapezoid $NJKO_1$, the national income of the citizen of country 1 (the Gross National Product, GNP), which includes the income earned by the workers who are migrating to country 2, is increased to the area $NJDHO_1$. So, the net gain to Country 1 is equal to the area of trapezoid $DIJG$. Needless to say, the national incomes in both countries keep increasing until the number of workers expressed by HM migrate to country 2.

Therefore, according to the traditional economic theory, international migration increases the national income (and economic welfare) of both sending and receiving countries.

2. Economic Effects under the New Framework

However, the above rosy picture of international migration hinges upon various restrictive assumptions, such as perfect competition and constant returns to scale etc. There are several studies demonstrating that the conclusion of the traditional framework breaks down when new realistic assumptions are introduced. For example, Schiff (1999) introduced the role of social capital, and Goto (1998) analyzed the impact of migration when trade barriers and nontraded goods exist. Both authors cast some doubts on the orthodox conclusions based on Figure 3, and pointed out that migration can (is likely to) give adverse economic effect on receiving countries. In this sub-section, I summarize the argument of Goto (1998) on the economic impact of migration of the host country.

(1) Basic Characteristics of the Model

The formal model developed in Goto (1998) has three main characteristics, which are somewhat different from the traditional theory of international factor mobility. The following three features have been added in order to capture more realities for the analysis of the migrant workers in Japan: (1) a possible change in factor prices (e.g. a decline in wage rate) by admission of migrant workers even when the prices of tradable goods are constant; (2) an existence of non-traded goods; (3) an existence of trade barriers.

(a) Possible Change in Factor Prices

Under the standard Heckscher-Ohlin-Samuelson (H-O-S) framework, factor prices are completely determined by the prices of goods without regard to relative factor endowment. Thus, for a small open economy where prices of goods are given, the factor prices are unchanged by the admission of foreign workers. In reality, however, many empirical studies, including Morgan and Gardner (1982), found that the influx of foreign unskilled labor results in a depressing effect on the domestic wage rate. In order to overcome the discrepancies between the theoretical prediction and the empirical findings, a Jones-type specific factor model is used in this paper. More specifically, capital is assumed to be fixed to each sector while labor is mobile between sectors. Under the specific factor model, the change in relative factor endowment due to an influx of foreign labor plays an important role in determining factor prices, as discussed in Sapir (1983). As will become clear in the following discussion, the influx of labor causes a decline in wage rate, which agrees with the empirical findings.

(b) Non-Traded Goods

In the real world, the share of non-traded goods sector such as construction and services in the total economy is high, although less attention has been paid to this by traditional trade theories. Further, as seen in Table 11 above, two-thirds of the foreign unskilled workers apprehended by the Japanese authorities are employed in the non-traded goods sector. Moreover, in Europe and the United States, it is often pointed out that migrant workers contribute to the lower prices of non-traded goods (e.g., maid service and street-cleaning). While most formal studies of international labor mobility has been based on the two sector model (i.e., exportables and importables), the economy in the model here is assumed to consist of three sectors producing three kinds of goods: exportables, importables, and non-tradables. As seen below, the inclusion of non-traded goods yields additional insights into the economic effect of migration.

(c) Trade Barriers

While most studies of international factor mobility have assumed free trade, in the model below international trade is restricted by tariff and / or non-tariff barriers (NTBs), as is often the case in the real world. This is an application of the framework developed by Brecher and Diaz-Alejandro (1977), which was used to analyze international capital mobility. As shown in their study, the economic impact of factor inflow under trade barriers is very different from that under free trade.

(2) Specification of the Model

In the model, consumers are characterized by the following Cobb-Douglas social utility function.

$$U = C_1^\alpha C_2^\beta C_3^\gamma, \quad \alpha + \beta + \gamma = 1 \quad (3-1)$$

where C_1 , C_2 , and C_3 are the amount of consumption of exportables (good 1), importables (good 2), and non-tradables (good 3), respectively, and U is social utility. Consumers maximize the social utility function (3-1) subject to the budget constraint (3-2).

$$P_1 C_1 + (1+t)C_2 + P_3 C_3 = Y \quad (3-2)$$

where P_1 and P_3 are the prices of exportables and non-tradables, respectively, and Y is the national income. The world price of importables, which is considered to be the numeraire goods here, is set to unity. And t is the rate of domestic price markup of importables due to tariffs and non-tariff barriers. In order to avoid further complication, it is assumed that the world prices of tradables are given to the economy (i.e., the “small country” assumption). From the above utility maximization problem, the following three demand functions are obtained.

$$C_1 = \alpha Y / P_1 \quad (3-3)$$

$$C_2 = \beta Y / (1+t) \quad (3-4)$$

$$C_3 = \gamma Y / P_3 \quad (3-5)$$

The producers in the three sectors are characterized by the following Cobb-Douglas production function.

$$Q_1 = K_1^a l_1^{1-a} \quad (3-6)$$

$$Q_2 = K_2^b l_2^{1-b} \quad (3-7)$$

$$Q_3 = K_3^c l_3^{1-c} \quad (3-8)$$

where $a > b > c$

Q_i , l_i and K_i are, respectively, production, labor input and capital input in the production sector of the i -th good ($i = 1, 2, 3$). Note that capital is assumed to be fixed to each sector in the equilibrium after foreign workers are admitted, although it was mobile before the initial long run equilibrium was reached.

Producers maximize the following profit function:

$$\pi_i = P_i Q_i - (r_i K_i + w l_i) \quad (3-9)$$

where π_i and r_i are, respectively, the profit and rental rate of the i -th production sector, and w is wage rate. Solving the profit maximization problem, the following equilibrium conditions are obtained.

$$a K_1^{a-1} l_1^{1-a} P_1 = r_1 \quad (3-10)$$

$$(1-a) K_1^a l_1^{-a} P_1 = w \quad (3-11)$$

$$b K_2^{b-1} l_2^{1-b} (1+t) = r_2 \quad (3-12)$$

$$(1-b) K_2^b l_2^{-b} (1+t) = w \quad (3-13)$$

$$c K_3^{c-1} l_3^{1-c} P_3 = r_3 \quad (3-14)$$

$$(1-c) K_3^c l_3^{-c} P_3 = w \quad (3-15)$$

Note that equations (3-10) to (3-15) show that factor prices are equal to their marginal value product in equilibrium.

Domestic labor supply is assumed to be given, i.e., there is no wage leisure trade-off. Therefore, the sum of labor input in the three sectors is equal to the sum of the domestic labor endowment (L) plus the number of admitted foreign workers (L_f).

$$l_1 + l_2 + l_3 = L + L_f \quad (3-16)$$

The domestic supply of non-tradables must be equal to their domestic demand because, by definition, no international trade is allowed for them.

$$C_3 = Q_3 \quad (3-17)$$

Since the tariff revenue accrued to the government is assumed to be distributed to domestic consumers in a lump-sum fashion, and since there is no profit in equilibrium, the national income (GNP rather than GDP), which does not include the income accrued to migrant workers, consists of factor payments and tariff revenue.

$$r_1 K_1 + r_2 K_2 + r_3 K_3 + wL + t(C_2 - Q_2) = Y \quad (3-18)$$

Note that, in the calibration in the next sub-section, t includes the rate of price markup by non-tariff barriers (NTB) as well as by tariffs. Therefore, equation (3-18) might not be valid in some situations, because the government does not collect tariffs when import restrictions are given by NTBs. But, as long as the economic rent from the import restriction is accrued to domestic people, equation (3-18) is still valid. In fact, many studies have shown that, under the NTBs such as import quotas, the economic rent often accrued to the importers. Therefore, equation (18) seems to be justified even when a part of import restrictions are imposed in the form of NTBs. By substitution, equation (3-18) can be expressed by (3-19).

$$P_1 Q_1 + (1+t) Q_2 + P_3 Q_3 - wL_f + t(C_2 - Q_2) = Y \quad (3-19)$$

Note that temporary guest workers are assumed to be paid in terms of exportables here. In other words, the consumption pattern of guest workers is different from that of domestic workers, in that the former is assumed to spend their entire income on the consumption of exportable goods of the host country. This assumption seems reasonable because the temporary guest workers stay in Japan for a short period of time and they are remitting a large portion of their incomes to the home country.

(3) Quadruple Effects of Immigration on the Welfare of the Host Country

Because the model in the above section is complete, the economic impact of the admission of migrant workers on the Japanese economy can be *quantitatively* analyzed by calibrating the model. But, before calibration, let us theoretically (or *qualitatively*) analyze the model

in order to obtain some insights into the welfare effect of labor inflow under the three realistic assumptions included in this model.

The following is an examination of the impact of the admission of migrant workers on welfare that is represented by the social utility (U) in equation (3-1) above. Substituting (3-3), (3-4), (3-5) into (3-1), (3-20) is obtained.

$$U = (\alpha/P_1)^\alpha (\beta/(1+t))^\beta \gamma^\gamma Y/P_3^\gamma \tag{3-20}$$

Taking natural logarithm of both sides of (3-20), (3-21) is obtained.

$$\ln U = \ln(\alpha/P_1)^\alpha (\beta/(1+t))^\beta \gamma^\gamma + \ln Y - \gamma \ln P_3 \tag{3-21}$$

Differentiating equation (3-21) with respect to L_f , (3-22) is obtained.

$$(\ln U)' = (\ln Y)' - \gamma (\ln P_3)' \tag{3-22}$$

where the variables with the prime notations are defined as the derivative of those variables with respect to L_f . Similar short-hand notations are used throughout this paper.

From (3-22), (3-23) is obtained.

$$(\ln U)' = Y'/Y - \gamma P_3'/P_3 \tag{3-23}$$

Equation (3-23) shows that the total welfare effect of admission of migrant workers can be broken down into the effect of the change in income and the effect of the change in the price of non-traded goods.

Tedious but straightforward substitution using equilibrium conditions yields the basic equation (3-24), which divides the total welfare effect into four sub-effects.

$$\begin{aligned} (\ln U)' = & B(-L_f w') \dots\dots \text{Effect 1 (Cheaper foreign labor effect)} \\ & + B(-tQ_2') \dots\dots \text{Effect 2 (Trade barrier effect)} \\ & + B(Q_3 P_3') \dots\dots \text{Effect 3 (Nontradables income effect)} \\ & - (C_3 P_3') \dots\dots \text{Effect 4 (Nontradables consumption effect)} \end{aligned} \tag{3-24}$$

Where $B \equiv (1+t)/(1+t-\beta t)$, Note that $B > 1$

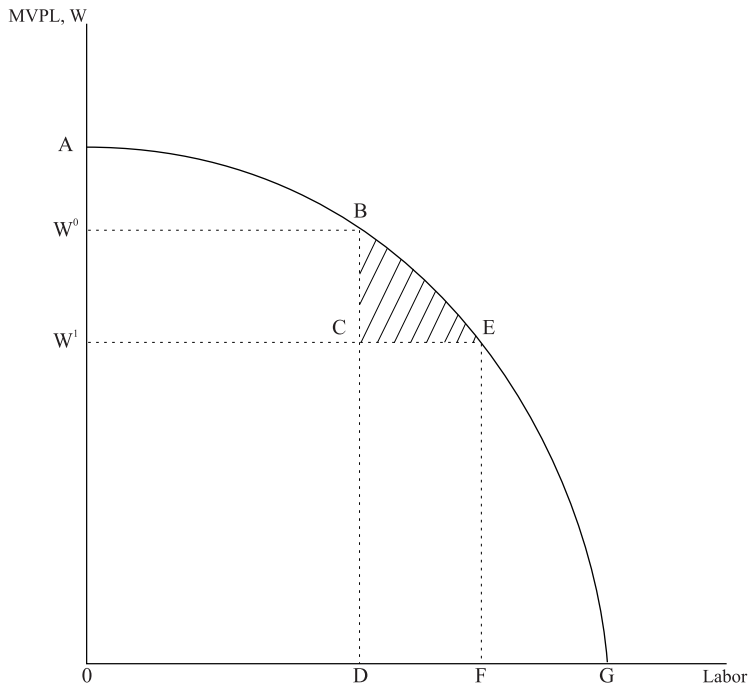
Note that the first three effects results from the effect on income. Following is a discussion of the above four sub-effects of the admission of migrant workers on the welfare of a host

country such as Japan.

(a) Effect 1: Cheaper foreign labor effect (positive)

Since it can be shown that w' is negative, Effect 1 has a positive impact on the welfare of the host country. In other words, the admission of foreign labor results in a positive sub-effect through the decline in wage rates. Figure 4 demonstrates an intuitive reasoning for this effect. In the figure, ABEG shows the marginal value product of labor (MVPL) curve. Since wage rate is equated with the MVPL in equilibrium, the equilibrium before the admission of foreign labor is B , where total domestic labor (OD) is employed with the wage rate of w^0 . In this case, total labor income is w^0ODB and total capital income is Aw^0B . If the foreign labor of DF is admitted to the country, then, the new equilibrium point moves to E , and the wage rate decreases to w^1 . In this case, capital income increases to Aw^1E , and total labor income accrued to the native workers and the income accrued to migrant workers become w^1ODC and $CDFE$, respectively. Thus, total income of domestic factors (capital and labor) is increased by the hatched area BCE . Note that the traditional $H - O - S$ model overlooks this effect, while labor economists often emphasize it. Further,

Fig. 4 CHEAPER FOREIGN LABOR EFFECT



note that the magnitude of the (positive) cheaper foreign labor effect increases, *ceteris paribus*, as the scale of the admission of migrant workers becomes larger.

(b) Effect 2: Trade barrier effect (Brecher = Diaz-Alejandro effect) (negative)

Since it can be rigorously shown that Q'_2 is positive (i.e., the amount of production of labor-intensive goods increases after foreign labor is imported), Effect 2 has a negative impact on the welfare of the host country. This effect stems from the existence of trade barriers which was first pointed out by Uzawa (1969) and rigorously discussed by Brecher and Diaz-Alejandro (1977). The intuition behind Effect 2 (trade barrier effect) is as follows: Although the international price of the importable is set to unity, its domestic price is higher ($l+t$) due to trade barriers. Therefore, the MVPL of a migrant worker under the domestic price is $(l+t)dQ_2/dl_2$, which is higher than the MVPL under the international price (dQ_2/dl_2). Since wage rate w is equated with domestic MVPL instead of international MVPL, the admitted foreign labor is in some sense overpaid. Hence, the admission of one migrant worker results in the overpayment of lQ'_2 , which tends to decrease the national income of the host country. Note that the magnitude of the (negative) trade barrier effect decreases, *ceteris paribus*, as t decreases (i.e., as trade is liberalized). In the extreme, under free trade (i.e., $t = 0$) Effect 2 disappears.

(c) Effect 3: Nontradables income effect (negative)

Since it can be shown that P'_3 is negative under a reasonable range of parameter values, Effect 3 is also negative. Since the national income consists of values of the three goods ($P_1Q_1 + (l+t)Q_2 + P_3Q_3$) plus tariff revenue less payment to foreign labor, the decline in the price of non-traded goods (P_3) results in the depressing sub-effect on the national income, and therefore on the welfare of the host country. Note that the traditional 2 goods (exportables and importables) model cannot capture Effect 3 which results from the change in the price of non-traded goods.

(d) Effect 4: Nontradables consumption effect (positive)

Since P'_3 is negative, Effect 4 is positive. In some sense, Effect 4 is the other side of the coin of Effect 3. From the consumers' viewpoint, the decline in the price of non-tradable goods is favorable because the same amount of income allows them to consume more goods. For example, the inflow of migrant workers enables consumers in the host country to enjoy cheaper maid service or street cleaning. However, note that the net nontradables effect (Effect 3 plus Effect 4) is negative. The proof of this proposition is straightforward. From (3-5) and

(3-24), (3-25) is obtained.

$$\begin{aligned} \text{Effect 3 + Effect 4} &= (B/Y)(Q_3 P_3') - C_3 P_3' \\ &= \gamma/P_3(B-l)P_3' \end{aligned} \quad (3-25)$$

Since B is greater than 1, and P_3' is negative, the net effect is negative. In other words, while consumers enjoy gains from cheaper maid services and street cleaning, the adverse effect resulting from the decline in income of native maids and native street-cleaners should be kept in mind.

(4) Trickle vs. Flood

The next important question is whether the net effect of the above four sub-effects has any systematic relationship to the level of admitted migrant workers (L_f) and the magnitude of trade barriers (t). The answer to this question is "yes." After some tedious algebra, it can be shown that the following two propositions hold under reasonable parameter values:

- (i) The welfare declines by the initial inflow of migrant workers, but after a certain number of admitted foreign workers (L_f^1) the welfare increases;
- (ii) The smaller the value of t , the smaller the value of L_f^1 . In other words, the less severe the trade barriers are, the more likely it is that the admission of a certain number of migrant workers is welfare improving.

Figure 5 summarizes the above two propositions. In the figure, the welfare level of host country (U) is plotted on the vertical axis, while the number of admitted foreign workers is plotted on the horizontal axis. Curve I plots the welfare level as a function of admitted migrant workers when the magnitude of trade barriers is t_1 . The admission of migrant workers decreases the welfare level of host country first, but when the number of admitted foreign workers reaches L_f^1 , the welfare level begins to increase, and exceeds the initial level when the number of admitted foreign workers exceeds L_f^2 . In other words, the admission of a small number (or *trickle*) of migrant workers produces a negative effect on the host country while a large number (or *flood*) produces a positive impact on the host country. This finding implies that when migrant workers are admitted, the admission quota should be large if it is to produce a positive welfare impact. Curve II plots the welfare level when the magnitude of trade barriers decreases to t_2 due to, for example, a successful implementation of the Uruguay Round agreement. The curve shifts upward and leftward, and therefore the trough of the curve also shifts leftward. In other words, a smaller number of migrant workers can be welfare improving.

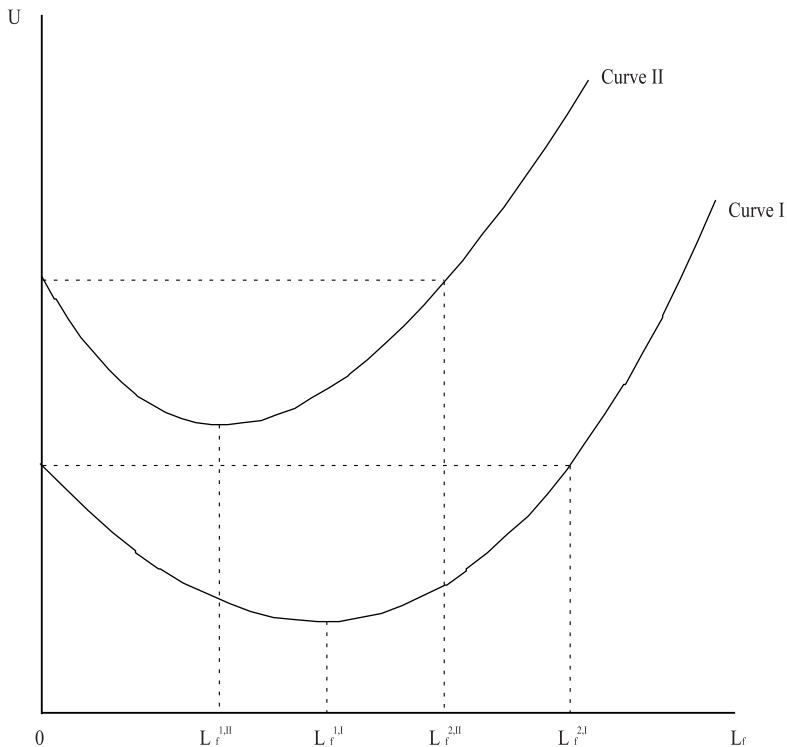
Although the proof of the above propositions requires a cumbersome manipulation of the equilibrium conditions of the model, the intuition behind these propositions is obtained from

the following, somewhat rough, argument: First of all, note that equation (24) can be rearranged to equation (3-26).

$$(\ln U)' = (1/Y)(-L_f w' + tM') \tag{3-26}$$

where $M \equiv C_2 - Q_2$ (i.e., the amount of imports). We already know that $-L_f w'$ (the cheaper foreign labor effect) is positive. It can be shown that M' is negative, i.e., the amount of the import of labor-intensive goods declines as the number of migrant workers (or the import of labor) increases, and therefore the second effect in equation (3-26) is negative. In equation (3-26), first note that, when L_f is zero, the first effect $(-L_f w')$ is also zero. Thus, the economic effect of the beginning influx of migrant workers is always welfare worsening. Second, note that the cheaper foreign labor effect is the change in wage rate (w') multiplied by the number of foreign workers (L_f), and therefore the magnitude of the first effect tends to increase more rapidly than the second effect, as the number of admitted foreign workers increases. Therefore, at a certain level of admitted foreign workers ($L_f^{1,1}$ in Figure 5), the first effect begins to dominate the second effect, which means that an additional admission of

Fig. 5 MIGRANT WORKERS AND HOST COUNTRY'S WELFARE



migrant workers is welfare-improving. Third, the second negative effect (tM') seems to decrease as t decreases because it is M' multiplied by t . Therefore, the admission of the same number of migrant workers can produce a positive impact on the welfare of the host country when the magnitude of trade barriers is smaller. The above analysis has the following policy implication for Japan: (1) while small scale admission of foreign workers has a negative impact on Japan, a large-scale admission is beneficial; (2) the liberalization of trade barriers increases Japan's chance to benefit from the admission of foreign labor.

3. The Effect of Migrant Workers on the Japanese Economy — Calibration

(1) Data and Method of Calibration

The impact of the *legal* admission of foreign unskilled workers on the Japanese economy is calibrated using the formal model developed above. The Japanese industries are classified into three categories to conform to the specification of the model. The classification was made according to the following criteria: (1) when both export share (the share of the value of exports in the value of total supply) and import share (the share of the value of imports in the value of total demand) are less than a threshold value (five percent), the industry is classified as a non-tradable sector; (2) the rest of the industries are classified in either an exportable sector or an importable sector, depending on which share (i.e., export share or import share) is larger.

The next task in the calibration is to identify the values of parameters in the model ($\alpha, \beta, \gamma, a, b, c, t, K_1, K_2, K_3, L, P_1$). First, capital stock in the three industries (K_i) and domestic labor supply (L) must be identified. Fortunately, actual data exist for these. Second, a, b , and c in the production function (3-6)-(3-8) must be determined. Since it is very difficult to directly estimate these values, an indirect method has been taken: The actual situation in the base year (in 1986) was assumed to be the long run equilibrium generated by the model, and the values of a, b , and c were calculated using actual (observed) values of endogenous variables (r_i, w , and l_i). I obtained $a = 0.4242$, $b = 0.3785$, and $c = 0.2234$. Third, the values of P_1 and t must be determined. By using equations (3-6), (3-7), (3-19), and observed values of Q_i, K_i , and l_i , I obtained $P_1 = 2.40$. Furthermore, using the data on the average tariff rate and the tariff equivalency of the NTBs, $t = 0.1329$ was obtained. Fourth, the values of α, s , and γ in the social utility function (3-1) must be identified. Since the utility function is Cobb-Douglas, α is equal to the expenditure share of exportables, and similar arguments can be made for s and γ . From this relationship, $\alpha = 0.2232$, $\beta = 0.1971$, and $\gamma = 0.5797$ were obtained. In other words, more than half of the income is spent on non-tradables. In view of the large expenditure share of non-tradables in total consumption, the inclusion of the third sector in the model seems to be all the more important.

(2) Results of Calibration

With values for all parameters (obtained as above), the model can be solved for the sixteen endogenous variables. The model was solved for various values of

L_f . This paper concentrates on the overall welfare effect (U'). In order to answer the question of how many foreign workers must be admitted if Japan is to derive a favorable economic impact, the simulation is repeated for different numbers of admitted foreign workers, using the model and parameter values discussed above. Figure 5 is also the result of the numerous simulations. In the figure, the value of welfare (U) is plotted for various values of admitted migrant workers (L_f). According to the simulations, Japanese welfare continues to decline until L_f reaches 1.66 million (or about three percent of the Japanese labor force), and it is not until L_f reaches 3.43 million (or about five percent of the Japanese labor force) that the welfare recovers to the initial level without migrant workers. In view of the fact that when the legalization of unskilled foreign labor is discussed in Japan, the proposed number is about one percent of the labor force (about 650,000), such a small-scale admission is very likely to have a negative impact on the Japanese economy.

Table 15 shows the effect of trade liberalization on the impact on Japanese welfare through the admission of foreign workers. While the base calibration uses $t = 0.1329$, the simulation is repeated for the reduced values of t in order to incorporate further trade liberalization. As Table 15 shows, if Japan succeeds in reducing trade barriers by half (i.e., $t = 0.066$), Japanese welfare turns to increase when 0.79 million foreign workers, which is a little over one percent of the Japanese labor force, are admitted. When 1.61 million foreign workers (about three percent of the Japanese labor force) are admitted, the Japanese welfare level is even higher than in the case of no admission of foreign workers. Thus, further attempts for trade liberalization seem to be very important.

4. The Social Effect of Migration in the Host Country

(1) Diversification and Internationalization

Since international migration involves the international movement of human beings as a whole, it brings about various social effects in addition to the economic effects discussed above. For example, suppose that a Japanese university in Tokyo hires a Korean mathematical professor. Although the job description of the Korean professor is probably to teach and make research into mathematics, his contribution to the university, and perhaps to the

Table 15: Trade Barriers and Welfare Effect of Migrant Workers

t (percent)	L_f^1 (thousand)	L_f^2 (thousand)
7.0	910	1,830
11.0	1,480	2,960
21.7	3,160	6,410

Source: Author's estimate. See main text for details.

Japanese society, is much more than that. His colleagues in the university in Japan can learn from him the economic and social situation in Korea as well as mathematics, and listening to the Korean fellow professor will widen their perspective. These interactions between Japanese and Korea would enhance mutual understanding between the two countries. In economic jargon, such an effect is called (positive) externality.

(2) Possible Burden on the Fiscal Expenditure in the Host Country

Since migrant workers pay taxes and receive various social services from the government of the host country, they give various effects on public finance in the host country. On the one hand, the existence of these migrant workers increases the revenue of the government, because they pay income taxes, consumption tax, and property taxes etc, and if they are enrolled in the social security system in the host country, they contribute to the social security system, too. On the other hand, it increases the expenditure of the government, because they receive various social services from the government of the host country, e.g., education for their children, medical services, and pension if they are enrolled in.

Recently, the Japanese government published an estimate of fiscal cost and benefit to the host government (both central and local government) for three different stages of admission of migrant workers (See Table 16). While the host government is benefited from migrant

Table 16 : Social Cost of Migrant Workers (billion yen)

	Stage 1	Stage 2	Stage 3
Receipt			
Central government	181.1	93.1	77.5
Local government	0.0	38.3	28.2
Social security	145.5	180.2	193.2
Total	326.6	311.6	298.9
Expenditure			
Central government	12.8	77.8	353.8
Local government	15.5	486.1	901.5
Social security	52.3	89.1	158.1
Total	80.6	653.0	1,413.4
Net			
Central government	168.3	15.3	-276.3
Local government	-15.5	-447.8	-873.3
Social security	93.2	91.1	35.1
Total	246.0	-341.4	-1,114.5

Note: Stage 1 : Single
 Stage 2 : Married
 Stage 3 : Married, two children

Source: Japanese Ministry of Labor

workers in Stage 1 (only single youth is admitted) because their tax payment exceeds social services they receive, in Stage 2 (with spouse) and Stage 3 (with spouse and two children), the fiscal cost for the social expenditure far exceeds tax revenues. When half a million migrant workers are admitted, the net cost to the government in Stage 3 exceeds one trillion yen (or about 8 billion dollars).

(3) Possible Increase in Crime

It is sometimes argued that the admission of migrant workers may increase crimes in the host country. Since Japan is a relatively homogeneous society and enjoys very low crime rate, some Japanese often argue that the migrant workers may bring about crimes into the safe country. Their logic behind such fear is very simplistic and goes something like the following: most large U.S. cities like Los Angeles and New York are filled with migrant workers and crime situation there is extremely serious, and therefore, there must be some correlation (and causality) between migration and crime. Therefore, the increase in migrant workers in Japan would make Japan more dangerous country.

However, in my view, such claim is not substantiated at all. There is no scientific evidence that connects high crime rate with migrant workers. On the contrary, some studies suggest that, as far as the first generation is concerned, migrant workers commit fewer crimes than natives in the host country. That is probably because newly migrated workers have higher motivation toward success in the new country, and because the effective penalty for committing a crime is far severer to migrant workers than to the native workers (i.e., only migrant workers face possible deportation from rich host country to poor home country!)

(4) Continuation of "3-D" jobs

In many cases, migrant workers are employed in the job whose working conditions are less favorable than other jobs. In Japan, such jobs are often referred to as "3-D job", i.e., "dangerous", "dirty", and "demanding" jobs. Since nobody prefers such 3-D jobs to other jobs, employers of 3-D jobs are often experiencing unfilled vacancies. The existence of such unfilled vacancies would encourage employers to make working conditions there more favorable (e.g., better wage, safer workplace etc.) However, if such 3-D vacancies are easily filled by migrant workers, the incentive for employers to achieve better working conditions would disappear. In other words, due to the hiring of migrant workers, who are willing to take jobs under the 3-D working condition because even the 3-D jobs there are better than regular jobs in their home country, the 3-D jobs in the host country may persist even in the long run.

IV. The Impact of Migration on the Sending Country

In this section, I will analyze the impact of migration from a perspective of sending countries. It should be noted that various economic effects of migration mentioned below, e.g., income creation and employment creation can be brought about by other measures such as foreign direct investment and international trade. For example, a Tokyo-based Japanese company that is currently hires migrant workers from Thailand can move the production plant to Thailand and hire workers there. Similarly, instead of importing labor from countries in East Asia, Japan can import labor-intensive products from them through international trade.

In view of the above, a simply CGE model will be developed below to empirically analyze the impact of migration in wider perspective, which incorporates migration, foreign direct investment, and international trade.

1. Various Effects of Emigration on the Sending Country

Before going into a rigorous analysis, various benefits and costs of emigration will be briefly discussed.

(1) Remittance and Income Creation

Needless to say, most migrant workers emigrate in an attempt to make more money in foreign countries than they can make in their home countries. It is often argued that such higher income benefits not only individual migrant workers but also their sending countries as a whole. Since poor sending countries are often suffering from prolonged balance of payment problems, the remittance from emigrant workers is an important source of foreign exchange receipt to these countries. In fact, the sizes of remittances for some of Asian developing countries are remarkable. As shown in Table 6 above, remittance is very important part of the national income in poor countries. But, as OECD (1987) pointed out, the flow of remittances is often unstable. When the host country is in recession, migrant workers are often the first to be fired, and therefore, in recession years, when sending countries particularly need foreign exchange receipts, the remittances tend to decrease. Further, it is often the case that, as more and more migrant workers decide to stay in the host country longer than initially expected and decide to invite families to join them, the remittances begin to decline. Moreover, the remittance may not necessarily help sending countries improve their balance of payment. As Stahl (1982) reported, the marginal propensity to consumption (especially consumption of imported luxury goods) out of remittance income is very high.

(2) Alleviation of Domestic Unemployment

Can emigration of labor alleviate serious unemployment problems in sending countries? In many Asian sending countries, there exists enormous unemployment (and underemployment) problem in rural agricultural sector and urban informal sector. Some argues that Japan should admit migrant workers from Asian LDCs in order to relieve unemployment problems in these countries. But, it seems questionable whether emigration to Japan relieves unemployment problems in Asian LDCs to a greater extent. As Todaro (1986) rigorously showed using his three sector model, while labor emigration may contribute to the relief of overall domestic unemployment, this favorable effect may be offset by a costly rise in urban unemployment caused by increased rural-urban migration. In other words, urban unemployment gets worse because more and more rural population move to an urban sector in the hope of further emigrating to Japan. Furthermore, the number of unemployment in Asian sending countries is too large to be relieved by the emigration to Japan. Probably, for a relief of unemployment, a creation of employment opportunities in these sending countries through foreign direct investment and / or increase of exports would be more important than emigration.

(3) Transfer of Knowledge and Skills

As well known, technology transfer has been one of the most important issues in economic development. In similar vein, it is often argued that emigration contributes to economic development and modernization of sending LDCs through the introduction of new knowledge and skills brought back by returned migrant workers. But, a survey by the Philippine Government suggests that this alleged benefit is questionable. According to the survey, two-thirds of the migrant workers, which include both skilled and unskilled, said that they acquired no skills. Mere 13.6 percent of migrants said that they acquired skills through employment in the host country. In view of this, it seems unlikely that unskilled migrant workers in Japan acquire skills through their employment. Provably, a formal training program would be much more useful than migration, in order to introduce new skills and knowledge to developing countries.

(4) Brain Drain

Economists have long pointed out that emigration causes "brain drain" from sending LDCs. Since good skilled labor emigrates in the hope of receiving higher pay in the developed country, sending LDCs often suffer from the lack of essential professional workers, such as doctors and nurses. It should be noted that even unskilled migrant workers in Japan often have very high level of education. Since the wage rate for an unskilled worker like a

construction worker in Japan is often much higher than the wage rate for a skilled worker like doctor or nurse in LDCs, people of higher education also come to Japan to take an unskilled job. According to a survey, more than 90 percent of migrants workers have high school education (42 percent have college education), although almost all of them are taking unskilled jobs in Japan. In other words, the persons who would work as an skilled workers like teachers and doctors in their home country come to work in Japan as unskilled workers, simply because the wage rate of unskilled workers in Japan is generally higher than that of skilled workers in their home country. Although it may economically benefit the migrant workers themselves, it is clearly a misallocation of labor, and gives a substantial loss to the persons left behind in the sending country.

2. The Model

In the general model, the world consists of two countries (or two groups of countries). Country A (B) is relatively capital (labor) abundant, and therefore, Country A (B) is a potential importer of labor (capital) and exporter of capital (labor). Also, country A (B) tends to export capital-intensive (labor-intensive) goods, while importing labor-intensive (capital-intensive) goods. The outline of the $2 \times 2 \times 2$ model is given below:

(1) Country A — capital-abundant

Consumers in country A are characterized by the following social utility function.

$$U_A = C_{1A}^\alpha C_{2A}^{1-\alpha}, \quad 0 < \alpha < 1 \quad (4-1)$$

where U_A is social utility of consumers in country A and C_{1A} and C_{2A} are the amount of consumption of good 1 and good 2, respectively, in country A. Good 1 is more capital-intensive than good 2. Hence, in a regular situation, country A exports good 1 and imports good 2. Consumers maximize social utility function (4-1) subject to the budget constraint (4-2).

$$P_{1A}C_{1A} + P_{2A}C_{2A} = Y_A \quad (4-2)$$

where P_{1A} and P_{2A} are, respectively, domestic (i.e., tariff-inclusive) prices of good 1 and good 2 in country A, and Y_A is the national income of country A. From the above utility maximization problem, the following two demand functions are obtained.

$$C_{1A} = \alpha Y_A / P_{1A} \quad (4-3)$$

and

$$C_{2A} = (1-\alpha) Y_A / P_{2A} \quad (4-4)$$

The producers of good 1 sector maximize the profit function (4-5) subject to the production function (4-6).

$$\pi_{1A} = P_{1A} Q_{1A} - r_A k_{1A} - w_A l_{1A} \quad (4-5)$$

$$Q_{1A} = k_{1A}^\alpha l_{1A}^{1-\alpha} \quad (4-6)$$

where π_{1A} , Q_{1A} , k_{1A} , and l_{1A} are, respectively, profit, amount of production, capital input, and labor input, in the good 1 sector of country A. r_A and w_A are rental rate of capital and wage rate of labor, respectively.

Solving the profit maximization problem, the following equilibrium conditions are obtained.

$$r_A = P_{1A} \alpha k_{1A}^{\alpha-1} l_{1A}^{1-\alpha} \quad (4-7)$$

$$w_A = P_{1A} (1-\alpha) k_{1A}^\alpha l_{1A}^{-\alpha} \quad (4-8)$$

Similarly, the production of good 2 in country A is characterized by the following four equations, which correspond to equations (4-5), (4-6), (4-7), and (4-8) for good 1 production.

$$\pi_{2A} = P_{2A} Q_{2A} - r_A k_{2A} - w_A l_{2A} \quad (4-9)$$

$$Q_{2A} = k_{2A}^b l_{2A}^{1-b} \quad (4-10)$$

$$r_A = P_{2A} b k_{2A}^{b-1} l_{2A}^{1-b} \quad (4-11)$$

$$w_A = P_{2A} (1-b) k_{2A}^b l_{2A}^{-b} \quad (4-12)$$

Note that we are assuming that good 1 is more capital-intensive than good 2, or assuming $a > b$.

Domestic labor supply is assumed to be given, i.e., there is no wage leisure tradeoff. So, the sum of labor input in two sectors is equal to the sum of the domestic labor supply (L_A) and the number of migrant workers (L_F) in country A.

$$l_{1A} + l_{2A} = L_A + L_F \quad (4-13)$$

Similarly, the amount of capital input in two sectors is equal to the domestic capital supply (K_A) less the amount of outgoing FDI from country A into country B (K_F).

$$k_{1A} + k_{2A} = K_A - K_F \quad (4-14)$$

Since the tariff revenue accrued to the government is assumed to be distributed to domestic consumers in a lump-sum fashion, and since there is no profit in equilibrium, the national income (GNP rather than GDP), which does not include the income accrued to the imported factor but does include the income accrued to the exported factor which is employed in the foreign country, is as follows.

$$r_A(K_A - K_F) + r_B K_F + w_A L_A + P_{2A}(t_A / (1 + t_A))(C_{2A} - Q_{2A}) = Y_A \quad (4-15)$$

where t_A is the tariff rate imposed on the imported goods by country A. Note that, since we are dealing with an interior solution here, capital-abundant (or labor-scarce) country A is importing labor-intensive good 2.

(2) Country B — labor-abundant

Consumers in country B are characterized by the following social utility function.

$$U_B = C_{1B}^\alpha C_{2B}^{1-\alpha}, \quad 0 < \alpha < 1 \quad (4-16)$$

where U_B is social utility of consumers in country B and C_{1B} and C_{2B} are the amount of consumption of good 1 and good 2, respectively, in country B. Consumers maximize social utility function (4-16) subject to the budget constraint (4-17).

$$P_{1B} C_{1B} + P_{2B} C_{2B} = Y_B \quad (4-17)$$

where P_{1B} and P_{2B} are, respectively, domestic (i.e., tariff-inclusive) prices of good 1 and good 2 in country B, and Y_B is the national income of country B.

From the above utility maximization problem, the following two demand functions are obtained.

$$C_{1B} = \alpha Y_B / P_{1B} \quad (4-18)$$

$$C_{2B} = (1 - \alpha) Y_B / P_{2B} \quad (4-19)$$

The producers of good 1 sector in country B maximize the profit function (4-20) subject to the production function (4-21).

$$\pi_{1B} = P_{1B}Q_{1B} - r_B k_{1B} - w_B l_{1B} \quad (4-20)$$

$$Q_{1B} = k_{1B}^\alpha l_{1B}^{1-\alpha} \quad (4-21)$$

where π_{1B} , Q_{1B} , k_{1B} , and l_{1B} are, respectively, profit, amount of production, capital input, and labor input, in the good 1 sector of country B. r_B and w_B are rental rate of capital and wage rate of labor, respectively, in country B.

Solving the profit maximization problem, the following equilibrium conditions are obtained.

$$r_B = P_{1B} \alpha k_{1B}^{\alpha-1} l_{1B}^{1-\alpha} \quad (4-22)$$

$$w_B = P_{1B} (1-\alpha) k_{1B}^\alpha l_{1B}^{-\alpha} \quad (4-23)$$

Similarly, the production of good 2 in country B is characterized by the following four equations, which correspond to equation (4-20), (4-21), (4-22), and (4-23) for good 1 production.

$$\pi_{2B} = P_{2B}Q_{2B} - r_B k_{2B} - w_B l_{2B} \quad (4-24)$$

$$Q_{2B} = k_{2B}^b l_{2B}^{1-b} \quad (4-25)$$

$$r_B = P_{2B} b k_{2B}^{b-1} l_{2B}^{1-b} \quad (4-26)$$

$$w_B = P_{2B} (1-b) k_{2B}^b l_{2B}^{-b} \quad (4-27)$$

Since, domestic labor supply is assumed to be given, the sum of labor input in two sectors is equal to the sum of the domestic labor supply (L_B) minus the number of emigrated workers (L_F) from country B to country A.

$$l_{1B} + l_{2B} = L_B - L_F \quad (4-28)$$

Similarly, the amount of capital input in two sectors is equal to the domestic capital supply (K_B) plus the amount of receiving FDI from country A into country B.

$$k_{1B} + k_{2B} = K_B + K_F \quad (4-29)$$

As in (4-15), the national income of country B is as follows:

$$r_B K_B + w_B (L_B - L_F) + w_A L_F + P_{1B} (t_B / (1 + t_B)) (C_{1B} - Q_{1B}) = Y_B \quad (4-30)$$

where t_B is the tariff rate imposed on the imported goods (good 1).

(3) Tariff wedge under no price discrimination by the firm

In the model, firms are assumed to have no monopolistic power to exercise price discrimination among the different markets. Therefore, the domestic price of good 2 in country A (P_{2A}) is higher by t_A than the domestic price of good 2 in country B (P_{2B}).

$$P_{2A} = (1 + t_A) P_{2B} \quad (4-31)$$

Similarly, the following relationship holds for the price of good 1.

$$P_{1B} = (1 + t_B) P_{1A} \quad (4-32)$$

(4) World market clearance

Since the amount of total production of each good must be equal to the amount of total consumption of the good in equilibrium, we have the following two equations.

$$C_{1A} + C_{1B} = Q_{1A} + Q_{1B} \quad (4-33)$$

$$C_{2A} + C_{2B} = Q_{2A} + Q_{2B} \quad (4-34)$$

The above model is complete, and the system of 28 equations ((4-1), (4-3), (4-4), (4-6), (4-7), (4-8), (4-10), (4-11), (4-12), (4-13), (4-14), (4-15), (4-16), (4-18), (4-19), (4-21), (4-22), (4-23), (4-25), (4-26), (4-27), (4-28), (4-29), (4-30), (4-31), (4-32), (4-33), and (4-34)) determines the equilibrium values of 28 endogenous variables in the model ($U_A, U_B, C_{1A}, C_{2A}, C_{2B}, Q_{1A}, Q_{2A}, Q_{1B}, Q_{2B}, k_{1A}, k_{2A}, k_{1B}, k_{2B}, l_{1A}, l_{2A}, l_{1B}, l_{2B}, P_{1A}, P_{2A}, P_{1B}, P_{2B}, r_A, w_A, r_B, w_B, Y_A, Y_B$). Note that one of the 28 equations is redundant by Walrus's law and that one of the prices (e.g., w_B) can be set to unity as the numeraire. Hence, the above model consists of 27 independent equations that can be solved for 27 endogenous variables. Therefore, as soon as the values of parameters ($\alpha, a, b, K_A, L_A, K_B, L_B, K_F, L_F, t_A, t_B$) are identified, the model can be solved to obtain the equilibrium values of endogenous variables for the parameter values.

In order to evaluate the impact of migration (FDI), all we have to do is to solve the

model for the different values of L_F (K_F) and compare the values of endogenous variables in each situation. Similarly, for evaluation of the impact of trade liberalization, we can solve the model for reduced values of $t_A t_B$ and t_B .

3. Impact of Asian Migration — Illustrative Simulation

(1) Method for Simulation

(a) Basic Strategy of Simulation

In what follows, we will calibrate the model developed above in order to present an integrated analysis of migration, FDI, and trade liberalization. While in many cases the effect of migration has been analyzed separately from the effects of FDI and of trade liberalization, we will analyze the impact of migration under the situation where FDI and international trade also exist, and compare the effects of migration with that of trade liberalization. Therefore our model is a little more complicated and sophisticated than traditional frameworks for the analysis of migration, where 'no trade' or 'free trade' has often been assumed and FDI has been assumed away. The motivation for our taking a more complicated approach is twofold:

First, contrary to the assumption of the traditional framework, migration, FDI, and trade barriers coexist in the real world. For example, while East Asian countries are exporting workers to Japan and other developed countries, they are also admitting many multinational corporations from Japan and other countries (e.g., SONY, Panasonic, Toyota, etc. to name a few). In addition to the movement of production factors (labor and capital), countries are integrated with each other through international trade that is subject to tariffs and nontariff barriers. As discussed in Section III above, the effect of migration in the economy with international trade subject to trade barriers, is very much different from that in the economy without trade. Therefore, in order to evaluate the impact of migration more precisely, we have to simultaneously incorporate the impact of trade barriers and capital movement.

Second, as often argued, there are various alternatives to migration. Instead of moving workers from East Asia to Japan, the Japanese firms may be able to move their plants to East Asia and employ workers there. Also, instead of importing labor, Japan can import labor-intensive goods such as textiles and clothing from neighboring countries. In other words, Japan can *indirectly* utilize labor in East Asia through international trade. In view of the fact that there are alternatives to migration, for policy discussions at least, benefits and costs of migration should be compared with those of alternatives such as trade liberalization or foreign direct investment, because, if all the benefits of migration can be achieved by an alternative measure with fewer costs, the alternative is superior to migration. As will become clearer below, trade liberalization is generally superior to migration as a means of enhancing welfare

in sending countries as well as receiving countries.

In the simulation below, I will analyze the economic relationship between Japan and East Asia as an aggregate. So, Country A in the model means Japan, and country B is the aggregated entity of seven countries in East Asia, i.e., China, Indonesia, Korea, Malaysia, Philippines, Thailand, and Taiwan province of China. These seven countries are chosen because they are major exporters of migrant workers to Japan. Also, the group of seven countries corresponds to the group of ASEAN and Asian NIEs plus China, minus Hong Kong and Singapore. As is well known, per capita income in Hong Kong and Singapore is so high that they are importers of migrant workers rather than exporters. In fact, according to the statistics compiled by the Japanese Government, there is very little immigration to Japan from Hong Kong or Singapore.

Basic strategy of simulation is as follows: The model is solved for the base year using parameter values in 1997, and we obtain equilibrium values of endogenous variables. Next, the model is solved again using different (increased) numbers of migrant workers (L_F). While the actual number of migrant workers, both legal and illegal, from East Asia to Japan is estimated at 319,000 in 1997, the model is again solved for the three other values of L_F , i.e. 419,000 (319,000 plus 100,000), 500,000, and one million, in order to evaluate the welfare impact of migration for both Japan and East Asia. Similarly, the model is solved using different values of t_A and t_B (K_F) in order to evaluate the impact of trade liberalization (foreign direct investment). Note that the simulation result using different values of L_F is also compared with those using different values of t_A and t_B , in order to obtain some policy implications on the relative importance of migration and trade liberalization as a means of improving economic welfare.

(b) Identification of parameter values of the model

As discussed above, since the model consist of 27 independent equations and 27 endogenous variables (plus the price of the numeraire good), it can be solved if values of parameters ($\alpha, a, b, K_A, L_A, K_B, L_B, K_F, L_F, t_A, t_B$) are identified. I borrow some of the parameter values from the simulation in Section III, which is based on Goto (1998).

First, a and b in the production functions (4-21) and (4-25) must be identified. Borrowing the values used in Goto, I obtained $a = 0.4242$ and $b = 0.3785$. As assumed in the above model, good 1 is more capital intensive than good 2 because $a > b$.

Second, α in social utility function (4-1) (or (4-16)) must be identified. With the proportional adjustment of the values of superscripts in the utility function in Goto's three good model, I obtained $\alpha = 0.4509$. Namely, the share of consumption of capital-intensive goods in total expenditure is equal to 45 percent.

Third, the magnitude of current trade barriers of both countries (t_A and t_B) must be

identified. Note that t_A and t_B include both tariffs and nontariff barriers. Although the data for tariffs exist, it is very hard to obtain the data for tariff equivalency of nontariff barriers. Borrowing from Goto, I obtained $t_A = 0.1329$. In other words, the domestic price of imports in Japan is on average 13 percent higher than the international price due to various trade barriers. In view of the fact that, generally speaking, trade restrictions imposed on capital-intensive imports in developing countries are much higher than trade barriers by industrialized countries,³ I set, somewhat arbitrary, t_B as double of t_A (i.e., $t_B = 0.2658$).

Fourth, amounts of endowments of capital and labor in both countries (K_A , L_A , K_B , L_B) must be identified. While Goto presented the data for Japan (i.e., K_A and L_A), it is very hard to obtain the corresponding data for seven countries in East Asia, and, even if they existed, they do not seem very reliable. Moreover, when we deal with a multi-country model consisting of countries that are very different from each other in terms of education, training, technology, etc., a simple head count of the number of workers may be misleading. For example, the productivity of well-trained workers in Country A is likely to be higher than that of less-trained workers in Country B. Therefore, some measures of effective workforce, which incorporates difference in skills and education of workers and a difference in technology among countries, seems to be more desirable than a simple head count of workers for the simulation here. In order to cope with such difficulties, I indirectly estimated K_A , L_A , K_B , and L_B , using the observed values of some endogenous variables. I used observed values of national incomes (Y_A and Y_B), the value of Japanese exports to East Asia ($P_{1A}(Q_{1A} - C_{1A})$), and the value of Japanese imports from East Asia ($P_{2B}(Q_{2B} - C_{2B})$). In other words, in this preliminary calibration, K_A , L_A , K_B , and L_B are treated as if they were endogenous variables. By this preliminary calibration, I obtained $K_A = 5055$, $K_B = 371$, $L_A = 10$, and $L_B = 15.1$. The values for K 's do not show the dollar amounts of capital endowment due to the different choice of unit. Also, values for L 's do not show the head count of workers due to the choice of unit and, more importantly, due to the fact that they show the effective unit of labor endowment which embodies difference in education, training, etc.

Finally, by using the units for K 's and L 's, I proportionally adjusted the data of migration in Japan from East Asia (318,600) and the stock of FDI from Japan to East Asia (\$ 60.45 billion), and obtained $K_F = 39.1$ and $K_F = 0.0052$.

(2) Results of Calibration

With values for all the parameters (obtained as above), the model can be solved for the 28 endogenous variables. The model was solved using the above parameter values for the

³ It is often the case that the tariff rate imposed by developing countries on capital-intensive "luxury" goods, such as cars, is as high as 100 percent.

base simulation first. Then, the model is repeatedly solved using different values of L_F , t_A and t_B , and K_F , in order to evaluate the impact of migration, trade liberalization, and FDI.

(a) Effect of Increased Migration from East Asia to Japan

First of all, let us examine how welfare in Japan (U_A) and East Asia (U_B) changes when the number of migration from East Asia to Japan is increased either by Japan's open-door policy or increased efforts for emigration by sending countries. In addition to the base simulation for the actual situation in 1997, I repeated the simulation using different values of L_F . In Table 17, three simulation results are reported: Case A-1 where $L_F = 419,000$ (i.e., 100,000 increase from the current number of 319,000), Case A-2 where $L_F = 500,000$, and Case A-3 where $L_F = 1,000,000$. Other parameters are kept constant for these three cases.

As shown in the table, the economic welfare of East Asia (country B) increases as the number of emigration to Japan increases because these emigrated workers earn higher wages than they could in their home country. They are more productive in the labor-scarce country than in the labor-abundant country. However, the economic welfare of Japan (country A) decreases as the number of migration from East Asia increases due to the Brecher = Diaz-Alejandro effect (or Uzawa effect), which is discussed in detail in Section III. However, note that the magnitude of the increase in the welfare of East Asian countries is larger than the magnitude of the decrease in the welfare of Japan. For example, when one million of migrant workers are admitted to Japan, the welfare index of East Asia increases by 176 while the Japanese welfare index decreases only by 20.⁴

Table 17

	Base	Case A-1 (small)	Case B-1 (middle)	Case C-1 (large)
Migration (LF) (thousand)	319	419	500	1,000
Japanese Trade Barrier (TA) (%)	13.29	13.29	13.29	13.29
East Asian Trade Barrier (TB) (%)	26.58	26.58	26.58	26.58
Stock of FDI (KF) (\$ billion)	60.45	60.45	60.45	60.45
Japanese Welfare (UA)	100,000	99,997	99,994	99,980
change	0	-3	-6	-20
East Asian Welfare (UB)	100,000	100,029	100,049	100,176
change	0	29	49	176

Source: The author's estimate. See main text for detail.

4 Although, at the first glance, the changes in welfare by 20 or by 176 (from 100,000) may look very small and essentially zero, we should note that the number of migrant workers here (one million) is also very small in comparison with the total population of the seven countries in East Asia (1.7 billion). Therefore, the important thing is the relative number rather than absolute number. Also, note that the utility is expressed as an ordinary number, rather than cardinal number.

(b) Effect of Trade Liberalization

Let us examine how welfare of Japan (U_A) and East Asia (U_B) changes when both countries decrease their trade barriers by reduction of tariffs and/or lifting some nontariff barriers. As well known, tariffs have been decreased by a series of world trade negotiations such as the Tokyo Round and the Uruguay Round. Recently, there are various initiatives to realize trade liberalization in Asia and the Pacific. In order to evaluate the impact of such efforts toward trade liberalization, I repeated the simulation using different values of t_A and t_B . In Table 18, three simulation results as well as the base simulation result are reported: Case B-1 where $t_A = 12.63\%$ and $t_B = 25.25\%$ (i.e, five percent decline in tariff rates in both countries from the base simulation), Case B-2 where $t_A = 11.96\%$ and $t_B = 23.92\%$ (i.e, ten percent decline from the base simulation), Case B-3 where $t_A = 9.03\%$ and $t_B = 18.60\%$ (i.e, thirty percent decline from the base simulation), to evaluate the welfare impact of trade liberalization. Other parameters are kept constant for these three cases.

As shown in the table, the economic welfare of East Asia (country B) increases as trade liberalization continues, because after the liberalization consumers benefit from cheaper foreign goods and because producers in the exporting good (good 2) sector benefit from increased demand abroad. In comparison with the effect of migration, the magnitude of welfare increase by trade liberalization is very large. When five percent, ten percent, and thirty-percent trade liberalization are achieved, the East Asian welfare increases by 780, 2088, and 5937, respectively.

Note that through trade liberalization, the economic welfare of Japan also increases and magnitude of such an increase is large. When five percent, ten percent, and thirty percent trade liberalization are achieved, the Japanese welfare increases by 87, 185, and 543, respectively. Thus, trade liberalization substantially benefits not only labor-exporting East Asia but also labor-importing Japan. The result conforms to the theoretical analysis of Goto (1998).

Table 18 : Result of Simulation B: Impact of Trade Liberalisation

	Base	Case B-1 (small)	Case B-2 (middle)	Case C-3 (large)
Migration (LF) (thousand)	319	319	319	319
Japanese Trade Barrier (TA) (%)	13.29	12.63	11.96	9.3
East Asian Trade Barrier (TB) (%)	26.58	25.25	23.92	18.61
Stock of FDI (KF) (\$ billion)	60.45	60.45	60.45	60.45
Japanese Welfare (U_A)	100,000	100,087	100,085	100,543
change	0	87	85	543
East Asian Welfare (U_B)	100,000	100,780	102,088	105,937
change	0	780	2,088	5,937

Source: The author's estimate. See main text for detail.

(c) Effect of the Increase in Japanese FDI to East Asia

Next, let us examine how welfare of Japan (U_A) and East Asia (U_B) changes when the amount of Japanese FDI to East Asia increases. In order to evaluate the impact of such efforts to increase the FDI, I repeated the simulation using different values of K_F . In Table 19 three simulation results as well as the base simulation result are reported: Case C-1 where $K_F = \$61.45$ billion (i.e., \$1 billion increase from the base value of \$60.45 billion), Case C-2 where $K_F = \$65.45$ billion (i.e., \$5 billion increase from the base value), and Case C-3 where $K_F = \$70.45$ billion (i.e., \$10 billion increase from the base value of \$60.45 billion). Other parameters are kept constant for these three cases.

As shown in the table, the economic welfare of Japan (country A) increases as the amount of the Japanese FDI increases because outgoing capital earns higher returns than it would in its home country. Capital is more productive in the capital-scarce (and labor-abundant) country. When the stock of the Japanese FDI in East Asia increases by \$1 billion, \$5 billion, and \$10 billion, the economic welfare of Japan increases by 45,229, and 457, respectively. However, the economic welfare of East Asia (country B) decreases as the amount of Japanese FDI increases. Brecher=Diaz-Alejandro effect (or Uzawa effect) is working here. When the stock of Japanese FDI in East Asia increases by \$1 billion, \$5 billion, and \$10 billion, the economic welfare of East Asia decreases by 105,530, and 1057, respectively. Namely, unless negative Brecher = Diaz-Alejandro effect is compensated by some other external benefit, such as technology transfer, the increase in FDI tends to give negative effects on the receiving country under the tariff-distorted world economy.

4. Migration vs. Trade Liberalization — Some Policy Implications

So far, we have examined the impact of migration, trade liberalization, and FDI sepa-

Table 19 : Result of Simulation C: Impact of FDI

	Base	Case C-1 (small)	Case C-2 (middle)	Case C-3 (large)
Migration (LF) (thousand)	319	319	319	319
Japanese Trade Barrier (TA) (%)	13.29	13.29	13.29	13.29
East Asian Trade Barrier (TB) (%)	26.58	26.58	26.58	26.58
Stock of FDI (KF) (\$ billion)	60.45	61.45	60.45	70.45
Japanese Welfare (U_A)	100,000	100,045	100,229	100,457
change	0	45	229	457
East Asian Welfare (U_B)	100,000	99,895	99,470	99,943
change	0	-105	-530	-1,057

Source: The author's estimate. See main text for detail.

rately. At the end of the calibration section of the paper, let us discuss the relative benefits of migration and other alternatives, especially trade liberalization. First of all, if we are interested in the welfare improvement of labor-importing developed country such as Japan, migration is the worst measure of three alternatives. Migration tends to reduce the welfare in Japan at least under such range of the amount of migration,⁵ while other two alternatives give a positive impact on Japanese welfare. However, we may well argue that the welfare improvement of East Asia is more important than the (small) decline in Japanese welfare, because, in view of the huge income gap between Japan and East Asian countries, Japan has some international obligation to contribute to the economic development of her neighbors.

In what follows, I will examine the relative effectiveness of migration and trade liberalization as a means of improving the welfare of East Asia (U_B). For that purpose, I repeat the simulation exercise using many different values of L_F , to find out how many migrant workers (L_F) have to be admitted to achieve the same values of U_B as that achieved by trade liberalization in Case B-1, Case B-2, and Case B-3, respectively. In other words, through repeated simulations, I tried to find how many migrant workers must be admitted to Japan, if we want to achieve the same welfare improvement in East Asia as the improvement achieved by various degree of trade liberalization. Table 20, which is probably the most important table in this paper, summarizes the result of simulations. As the table shows, when ten percent reduction in trade barriers is achieved for both t_A (from 13.29% to 11.96%) and t_B (from 26.58% to 23.92%), the welfare of East Asian improves by 2,088. If we achieve this 2,088 improvement of the East Asian welfare through migration to Japan, the number of East Asian migrant workers into Japan has to be increased to 8.7 million from current 0.3 million. While ten percent reduction in trade barriers does not seem so difficult to achieve, the admission of East Asian migrant workers to Japan by 8.7 million is far from being realistic (at least in the near future). Similarly, if we want to achieve the improvement of East Asian welfare realized by 30 percent reduction in trade barriers (i.e. t_A from 13.29% to 9.30% and t_B from 26.58% to 18.61%), the number of East Asian migrant workers has to be increased to 24.9 million from current 0.3 million. In view of the fact that current number of employees in Japan is 53.7 million, the admission of 24.9 million (or about half of total number of Japanese workers) is far from realistic.

Thus, it seems to me that trade liberalization is more effective (and realistic) way to improve the welfare in labor-exporting countries in East Asia than migration. Moreover, it should be noted that trade liberalization also improves the welfare of Japan, a labor-importing country. However, some caveats may be necessary. The above conclusion, based on the simulation results reported in Table 20, does not incorporate (positive or negative) externalities

⁵ As discussed earlier, very large scale of admission of migrant workers may increase the welfare of recipient country. See Goto (1998) for detailed discussions on this point.

nor does it include adjustment costs associated with international migration. Migration is not a mere movement of labor as a production factor but the movement of human beings as a whole across the border. Migration involves various externalities. For example, when a Japanese university hires a chemistry professor from the Philippines, the professor may increase the productivity of fellow Japanese professors of economics by familiarizing them with the economic development of the Philippines and other East Asian countries. Similarly, when the Filipino professor goes back to his country, he may bring back some advanced technology from Japan. The degree of positive externality would depend on the type of work performed by the migrant worker. If he (she) works as a professional worker, such as professor, doctor, and lawyer, positive externalities are easily expected. However, if the

Table 20 : Migration vs. Trade Liberalization

	Base	Trade Liberalization	Migration
Comparison with 5% reduction			
Migration (LF) (thousand)	319	319	3398
Japanese Trade Barrier (TA) (%)	13.29	12.63	13.29
East Asian Trade Barrier (TB) (%)	26.58	25.25	26.58
Stock of FDI (KF) (\$ billion)	60.5	60.5	60.5
Japanese Welfare (UA)	100,000	100,087	99,913
change	0	87	-87
East Asian Welfare (UB)	100,000	100,780	100,780
change	0	780	780
Comparison with 10% reduction			
Migration (LF) (thousand)	319	319	8662
Japanese Trade Barrier (TA) (%)	13.29	11.96	13.29
East Asian Trade Barrier (TB) (%)	26.58	23.92	26.58
Stock of FDI (KF) (\$ billion)	60.5	60.5	60.5
Japanese Welfare (UA)	100,000	100,185	99,770
change	0	185	-230
East Asian Welfare (UB)	100,000	102,088	102,088
change	0	2,088	2,088
Comparison with 30% reduction			
Migration (LF) (thousand)	319	319	24875
Japanese Trade Barrier (TA) (%)	13.29	9.30	13.29
East Asian Trade Barrier (TB) (%)	26.58	18.61	26.58
Stock of FDI (KF) (\$ billion)	60.5	60.5	60.5
Japanese Welfare (UA)	100,000	100,543	99,371
change	0	543	-629
East Asian Welfare (UB)	100,000	105,937	105,937
change	0	5,937	5,937

Source: The author's estimate. See main text for detail.

migrant worker is employed as an unskilled worker, which is the case for most migrant workers in East Asia, it is hard to expect for him (her) to bring about large degree of positive externalities to the Japanese economy, and to bring home advanced technology from Japan. On the other hand, the movement of workers may involve larger adjustment cost than the movement of goods. If this is the case, the case for trade liberalization over migration is further strengthened.

V. Concluding Remarks

After presenting some background facts of intra-Asian migration, we have theoretically and empirically examined the impact of migration on both the host country and the home country in East Asia. As demonstrated in Section III, the economic effect of immigration is not necessarily beneficial to the host country, because the negative trade barrier effect (Brecher-Diaz Alejandro effect) dominates other positive effects in many cases. This implies that, at least in the long run, the admission of guest workers may not be a desirable policy, unless the above negative effect is compensated by some positive externalities. As argued above, the migration by professional and skilled workers is more likely to have positive externalities than that by unskilled workers. Unfortunately, however, the majority of the current intra-Asian migration is the movement of unskilled workers, such as maids and construction workers. In addition to the economic impact on the host country, the impact on the home country and the social effects have been also examined in Section III.

In Section IV, after developing a simple $2 \times 2 \times 2$ CGE model, which incorporates international migration, FDI, and trade liberalization, the economic effects of migration and trade liberalization are compared with each other, taking the economic relationship between Japan and seven East Asian countries as an example. Through a series of simulation exercises, we found that migration tends to give negative welfare effects to the receiving country (Japan), although it has positive welfare impact on sending countries in East Asia, such as Indonesia and the Philippines. It was demonstrated that, in order to achieve the positive economic impact on the home country, such as income creation and alleviation of unemployment, there are *alternative measures* such as the increased flow of goods through trade liberalization by both parties. The simulation results suggest that trade liberalization is far more effective policy than international migration to enhance the economic welfare of the sending country. Moreover, it should be noted that the trade liberalization is beneficial to both the host country and home countries in East Asia, while migration tends to benefit only sending home countries. Thus, even from a pure economic viewpoint, trade liberalization seems to be a superior policy to migration.

In this paper, I have presented the theoretical and empirical analysis of the potential effect of migration. By presenting an objective analysis of migration, and by comparing the

impact of migration with that of trade liberalization, I hope that this paper gives some insights into desirable policies toward migration in Asia and the Pacific. I focused my discussion on what is the desirable policy towards immigration from the macroeconomic viewpoint. Needless to say, it is also very important to consider how to protect the welfare of currently existing migrant workers as individual human beings, without regard to the effect of migration as a whole to the overall economic welfare of the countries concerned.

Bibliography

- Abubakar, S. Y., "Migrant Labour in Malaysia: Impacts and Implications of the Asian Economic Crisis" (mimeo), 1999.
- Battistella, G. and M.M.B. Asis, "The Crisis and Migration in Asia," *mimeo*, 1999.
- Borning, W. R., "The Impact of Asian Crisis on Filipino Employment Prospect Abroad," *International Labor Office*, 1998.
- Brecher, R. and C. F. Diaz-Alejandoro, "Tariffs, Foreign Capital, and Immiserizing Growth," *Journal of International Economics*, vol. 7, pp. 317-322, 1977.
- Goto, Junichi, "The Impact of Migrant Workers on the Japanese Economy: Trickle vs. Flood," *Japan and the World Economy*, vol. 10, pp. 63-83, 1998.
- Goto, Junichi, *Labor in International Trade Theory — A New Perspective on Japanese-American Issues*, The Johns Hopkins University Press, Baltimore MD, June 1990.
- International Labor Organization, "Towards Full Employment: Prospect and Problem in Asia and the Pacific," 1999.
- Schiff, Maurice, "Trade, Migration, and Welfare," *Policy Research Working Paper (#2044)*, The World Bank, 1999.

IMPLICIT COLLUSION MODELS OF EXPORT PRICING: AN ECONOMETRIC APPLICATION TO THE JAPANESE CASE

EIICHI TOMIURA *Kobe University*

Abstract

This paper examines export competition by interpreting observed export price variations over time as a result of dynamic changes in the sustainability of implicit collusion among exporters. One model focuses on unpredictable negative shocks on demand as in Green and Porter (1984), while exchange rate fluctuation is emphasized in the alternative model as in Rotemberg and Saloner (1986). The regime classification dummy, which follows a Markov transition process, is estimated endogenously. Switches in export pricing are detected in some of the Japanese industries, especially in textiles.

JEL Classifications: F12, F14, D43, L16

Keywords: Implicit collusion, Export price, Switching regression

1. Introduction

To analyze the export price competition, this paper interprets observed export price variations over time as a result of dynamic changes in the sustainability of implicit collusion among exporters. This approach is motivated by the increasing importance of oligopoly in world trade and by the development of dynamic game analyses in industrial organization.¹

Two alternative models will be introduced, by applying the supergame theory, to formalize the individual firm's incentive to deviate from the collusive equilibrium.

The first model, or the "demand uncertainty model," which originates in Green and Porter (1984), considers the case where decisions by other firms are unobservable. Given the

Corresponding Address:

Eiichi Tomiura

Research Institute for Economics and Business Administration (RIEB), Kobe University

2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501 Japan

E-mail: tomiura@rieb.kobe-u.ac.jp

Phone/Fax: 81-78-803-7008

¹ The imperfect pass-through of exchange rate changes into export/import prices was previously analyzed mostly in microeconomic industrial organization context as in Dornbusch (1987), but is recently discussed also in macroeconomic perspectives related with propagation of policy effects. See Campa and Goldberg (2002) for the relation between these two views.

imperfect information, a rational firm gets suspicious of rivals' secret price cuts when the export demand for own product decreases unexpectedly. Firms sometimes shift to a punishment phase by cutting their prices even if all the firms keep the implicit collusion because firms cannot distinguish negative shocks to industry demand from a rival's deviation. Since the demand in foreign countries is often affected by so many factors that individual firms cannot correctly take account of, this model helps explain how unpredictable negative demand shocks affect export price changes in the real world.

The second model, or the "fluctuation model," which was originally developed by Rotemberg and Saloner (1986), emphasizes the cyclical aspect of export demand movements. If the level of export demand is currently higher than that expected in the future, the present time offers a good opportunity for deviation because the current gain from cheating during booms dominates the future loss from being punished in unfavorable market conditions.² Hence, this theory predicts that implicit collusion among exporters tends to destabilize during exchange rate depreciation.

To test the relevancy of these two alternative models, this paper estimates the export pricing equation by the econometric technique that has been established in the empirical industrial organization literature. The regime classification dummy, which is allowed to follow a Markov transition process, will be determined endogenously by the Bayes' rule based on maximum-likelihood estimates. Thus estimated regime probability could be interpreted as a measure of the "relative sustainability" of implicit collusion. The Japanese export is a good case to test implicit collusion models both because of their low pass-through in export prices and of the strong misperception of aggressive export behavior among Japanese corporations, once hotly debated in the 1980s. This paper estimates export pricing equations for four Japanese industries during that period. Although the industries are broadly defined, my estimation will hopefully serve as a first-step experiment for the empirical test of these industrial-organization based hypotheses.

This paper is structured as follows. Section 2 introduces the two models theoretically formalizing export price decisions. Section 3 constructs the econometric model and introduces approaches to test the theoretical predictions. Section 4 reports the estimation results from Japanese industries. Section 5 concludes.

2. Theory

As oligopoly industries generally increase their importance in world trade, the gains from applying the supergame theory to export pricing behaviors must be greater. This section introduces two alternative supergame models, adapted from industrial organization literature,

² Bagwell and Staiger (1997) theoretically generalizes the Rotemberg-Saloner model in the Markov growth model. They find the countercyclical pricing when the intertemporal correlation is negative.

to draw the predictions on export prices.

2.1 Demand uncertainty model

Suppose that firms compete in prices and that each firm cannot directly observe the prices chosen by competitors. Facing such uncertainty, a rational firm may infer others' price from the level of demand for his own product and detect a rival's secret price cut by a unanticipated decrease in his export demand. Any slight deviation from the collusive price is thus perfectly detected and drives firms into the competitive phase to punish the deviating firm. As the repeated game models show, given the discount rate of the firms, there exist the price level and the minimal punishment length that sustain the collusion without inducing a deviation.

In the real world, however, it is sometimes hard to detect a rival's deviation by indirect inference from changes in demand for each firm. If uncertain factors which cannot be known by the individual firm at the time of pricing decisions play a crucial role in export demand determination, the enforcement mechanism described above sometimes breaks down even if no firm deviates from the collusive pricing because each firm cannot distinguish a rival's secret price cut from an exogenous negative shock to the export market. The demand uncertainty model formalizes this intuition.

Before introducing formal models, we should clarify the information structure surrounding the export pricing because the timing issue is critical in dynamic games. This paper considers the following timing:

1. The exchange rate becomes known.³
2. Each firm chooses its export price.⁴
3. Aggregate export demand is revealed to the individual firm.

Consider an industry where domestic firms supply products to a foreign market in a simple two-country framework. Let me call the exporters as Japanese firms and the export market as U.S., just to make things concrete.

Suppose that firms are risk neutral and maximize the expected present value of future export profits:

$$W = E_t[\sum_t \delta^t \pi_t] \tag{1}$$

3 Since I assume no menu costs or no adjustment process for changing prices, there exists no uncertainty issue in the *current* exchange rate in this framework.

4 Here, I assume that the firm's decision variable is the yen price, even if export contracts are concluded in terms of the dollar in many cases of exports by Japanese firms. One of the supports for this assumption can be found in that many Japanese firms adopt the "in-house" exchange rate to evaluate each contract in terms of the yen.

where δ is the discount factor ($0 < \delta < 1$, assumed to be constant), π is the export profit per period and E_t is the expectation operator conditional on the all information available at the time t . Then, π is given by

$$\pi_{it} = (p_{it} - c)D\left(\frac{p_{it}}{e_t}, \frac{p_{-it}}{e_t}, P_t, Y_t\right) + u_t \quad (2)$$

where p_i denotes the export price charged by firm i ($i = 1, 2, \dots, I$) and p_{-i} is the price chosen by other exporting firms. Marginal cost (c) is assumed to be constant. Both p and c are expressed in terms of yen, and e is the exchange rate (yen per dollar). The export demand D also depends on the price offered by competing foreign firms and the income level of the importer country (P, Y , both expressed in dollars). As usual, $D_1 < 0, D_i > 0$ ($i = 2, 3, 4$) (where D_i is the partial derivative of D with respect to the i -th argument). And u denotes the disturbance term which cannot be known at the time of pricing decision and assume $E(u_{it}) = 0$ for all time t .

Without collusion, no firm can earn positive profit because the noncooperative Bertrand price competition is repeated every period.

Consider the following trigger strategy; keep choosing the collusive price as long as the collusion is sustained and charge the competitive price during the punishment phase once the collusion breaks down.

Now, suppose that export demand becomes lower than anticipated in spite of collusive pricing kept by all firms. Then, since the realized u in this case is negative,

$$\pi = (p^c - c)[D(p^c/e, p^c/e, P, Y) + u] < (p^c - c)E[D(p^c/e, p^c/e, P, Y)]$$

where p^c denotes the collusive price.

Since each firm cannot know the exact price charged by competitors, no firm can distinguish a rival's secret price cut from exogenous negative export demand shocks, even if all the firms know the export demand function correctly. Thus, an unanticipated export demand fall triggers a breakdown of collusion. Let me call this "uncertainty effect."

2.2. Exchange rate fluctuation model

The first theory, explained in the previous section, focuses on the uncertainty in export demand. On the other hand, we can sometimes safely assume that export demand is determined by observable factors such as current exchange rate. Then, there is no issue of uncertainty in export pricing and the issue left is the timing of deviation from collusive pricing. The second theory introduced below examines this side of export price competition.

Let the export demand for firm i now be a function of prices charged by firms:

$$D_{it} = D\left(\frac{p_{it}}{e_t}, \frac{p_{-i,t}}{e_t}\right). \quad (2')$$

There is no uncertainty in export demand determination in this model, since the current exchange rate is observable before the pricing decision.

Following Rotemberg and Woodford (1991), consider the case where the incentive compatibility constraint is binding:

$$\pi^m = \pi^c + x \quad (3)$$

where $\pi^m(\pi^c)$ denotes the per-period profit of the deviating firm (the per-period profit of a firm under collusion, respectively), and x is the future profit from the next period on ($x_t = W_t - \pi_t$ in (1)).

(3) can be rewritten as

$$(p^m - c)D\left(\frac{p^m}{e}, \frac{p^c}{e}\right) = (p^c - c)D\left(\frac{p^c}{e}, \frac{p^c}{e}\right) + x \quad (4)$$

where p^c is the collusive price which depends on the exchange rate and future profit and p^m is the price chosen by the deviating firm:

$$p^c = p^c(e, x)$$

$$p^m = \arg \max_p (p - c)D\left(\frac{p}{e}, \frac{p^c}{e}\right)$$

The effect of current exchange rate change on the export price can be examined by differentiating both sides of (4) with respect to e as follows:

$$\begin{aligned} (p^m - c) \left[D_1^m \left(-\frac{p^m}{e^2} + \frac{1}{e} \frac{\partial p^m}{\partial e} \right) + D_2^m \left(-\frac{p^c}{e^2} + \frac{1}{e} \frac{\partial p^c}{\partial e} \right) \right] + D^m \frac{\partial p^m}{\partial e} \\ = (p^c - c) (D_1^c + D_2^c) \left(-\frac{p^c}{e^2} + \frac{1}{e} \frac{\partial p^c}{\partial e} \right) + D^c \frac{\partial p^c}{\partial e} \end{aligned} \quad (5)$$

where D^m and D^c represent the export demand level when this firm deviates to the monopoly pricing and when all the firms keep collusive pricing, respectively, as follows:

$$D^m = D(p^m/e, p^c/e)$$

$$D^c = D(p^c/e, p^c/e)$$

and D_i denotes the partial derivative of the export demand function with respect to the i -th argument ($i = 1, 2$).

Let the own- and cross-elasticity of export demand be ε_1 and ε_2 ($\varepsilon_2 > 0 > \varepsilon_1$). Then, (5) can be rearranged to

$$\partial p^c / \partial e = (A - B)(C - D) \quad (6)$$

where

$$\begin{aligned} A &= (\varepsilon_1 + \varepsilon_2)(\pi^m - \pi^c) / e \\ B &= \left(D^m + \frac{\varepsilon_1 \pi^m}{p^m} \right) \frac{\partial p^m}{\partial e} \\ C &= \varepsilon_2 (\pi^m - \pi^c) / p^c \\ D &= D^c + \frac{\varepsilon_1 \pi^c}{p^c} \end{aligned}$$

Here, the following points help us determine the sign of each argument on the right-hand side of this equality. First, the deviating firm's profit is higher than collusive profit ($\pi^m > \pi^c$). Second, since the own price effect is normally stronger than the cross-price effect, we can safely assume $\varepsilon_1 + \varepsilon_2 < 0$.⁵ The combination of these two points requires that A must be negative and that C must be positive. Third, as the deviating firm chooses monopoly pricing, the price chosen by the deviating firm must satisfy the Lerner equality ($1 - c/p^m = -1/\varepsilon_1$).⁶ This relation means that B equals to zero and that D must be negative.

Therefore, we obtain

$$\partial p^c / \partial e < 0 \quad (7)$$

which indicates low export price under exchange rate depreciation.

This result sounds contradictory to that from the previous model because this asserts that the competitive phases emerge more often under higher level of export demand. Explicit contrast of basic assumptions, however, is important. The previous model focuses on the comparison of anticipated current export demand with realized one, while this model compares the current export demand with the future one. The difference in the assumption about the timing is crucial. Export price is chosen before total export demand becomes known in the first model. In the second model, however, the decision to undercut the price is made after knowing the actual exchange rate.⁷ Let me call this driving force in the second model

5 I also assume that the elasticity is constant for all levels of export demand. This assumption of constant-elasticity demand function is consistent with the log-linear econometric model, which will be introduced in the next section.

6 The Envelope Theorem also directly suggests that B is equal to zero.

7 Another important difference is that the uncertainty model predicts periodic switches between collusion and competition even if no firm deviates, while no switch is observed in equilibrium in the fluctuation model.

the “fluctuation effect.”

Next, we can study the effect of *future* changes on the current export price within the same model. Differentiating (4) with respect to x yields

$$(p^m - c)D_2^m \frac{1}{e} \frac{\partial p^c}{\partial x} = D^c \frac{\partial p^c}{\partial x} + (p^c - c)(D_1^c + D_2^c) \frac{1}{e} \frac{\partial p^c}{\partial x} + 1. \quad (8)$$

By rearranging, as in the previous differentiation,

$$\partial p^c / \partial x = 1 / (C - D) > 0. \quad (9)$$

Therefore, export price is high, for example, when the exchange rate depreciation is expected in the future. In other words, the collusion is more easily sustained at the same current exchange rate level when expected future export demand is on the expanding trend than on the declining trend because current gain from deviation is smaller than future loss from being punished. Let me call it “expected fluctuation effect” in contrast to the previously discussed “current fluctuation effect.”

3. Econometric procedures

This section explains the econometric approaches to test the theoretical predictions from the previous two supergame models.

3.1. Econometric model

To test the predictions provided by the two models, we need an econometric model, which flexibly allows various behavioral assumptions. First, consider the following pricing equation:

$$\ln p_t = \beta_0 + \beta_1 \ln q_t + \beta_2 \ln w_t + \beta_3 I_t + \sum_i \gamma_i DUM_{it} + v_t. \quad (10)$$

Here, I is the key indicator variable, of which the definition will be explained in detail later, taking zero in the collusive phase and one in the competitive phase for negative β_3 , and w denotes the input cost, DUM_i ($i = 2, \dots, 12$) are seasonal dummies for monthly data, and v is the i.i.d. error term with mean zero.⁸ When export quantity and seasonal dummies are not significant, the pricing equation will be simplified to

⁸ The pricing equation in the collusive phase is assumed to have the identical functional form to that in the competitive phase, except only for the constant term.

$$\ln p_t = \beta_1 + \beta_2 \ln w_t + \beta_3 I_t + v_t \quad (11)$$

Next, as in Section 2.1, consider the corresponding export demand function:

$$\ln q_t = \alpha_0 + \alpha_1 \ln p_t + \alpha_2 \ln e_t + \alpha_3 \ln P_t + \alpha_4 \ln Y_t + \sum_i \delta_i DUM_{it} + u_t \quad (12)$$

Thus, a system of export market consists of (10) (or (11)) and (12).

Since the goal of this paper is to check the significance of the switching dummy, we have to find a series of $\{I_t \mid t = 1, \dots, T\}$. Exploiting the results in empirical industrial organization literature, I will take the following two approaches:

- (A) Define $\{I_t\}$ exogenously from the economic theory (as in Baker (1989)).
- (B) Detect $\{I_t\}$ endogenously from data (as in Porter (1983), and Lee and Porter (1984)).

3.2. Estimation

3.2.1. Estimation with the exogenous dummy

The approach (A) in the previous section is suitable to test the “uncertainty effect.” The economic reasoning underlying this approach is straightforward. The demand uncertainty model suggests the definition of the switching dummy based on the unanticipated change in export demand by

$$\begin{aligned} I_t &= 1, \text{ if } u_t < 0 \\ &= 0, \text{ otherwise} \end{aligned} \quad (13)$$

where u_t is the estimation error of the export demand function.

Suppose that a firm forms its estimate about the export demand by the function (12). Then, it is rational for a firm to switch its pricing strategy depending on the sign of the estimated residual (u) of this export demand function; i.e., negative u implies the possibility of some firms’ defection and drives firms into a competitive period. This leads us to the definition of the switching dummy given above.⁹

After estimating the export demand function (12), the pricing equation (10) or (11) is estimated with a thus defined dummy. Under this definition, the essence of the uncertainty effect is boiled down to the one-tailed test of the null hypothesis $H_0: \beta_3 < 0$.

⁹ A slightly different definition is given in Baker (1989). He considers the case where the threshold level does not necessarily coincide with zero.

3.2.2. Estimation with the endogenous dummy

One of the most serious problems of the previous approach derives from the unobservability of the “true” sequence of the $\{I_t\}$. In other words, the sequence of dummies exogenously defined by the unanticipated demand changes does not necessarily coincide with the true switches even if the theoretical model is valid because each firm may have a different estimate of demand from ours and/or because some factors other than demand changes may have an impact on pricing. The exogenous dummy is a measure of the true switching factor with some errors. As in the usual cases of errors-in-variables models in econometrics, least squares estimation yields biased estimators.

Instead of using the imperfect sample separation information derived from the export demand function, we can estimate the probability of being in the “competitive” period (λ) by maximum likelihood and can classify each period into either “collusive” or “competitive” phase based on the estimation result.¹⁰ This consists of the essence of Approach (B), which I will explain, in what follows.

Suppose that the i.i.d. sequence of $\{I_t \mid t = 1, \dots, T\}$ follows the Bernoulli distribution¹¹

$$\begin{aligned} I_t = 1 & \quad \text{with probability } \lambda \\ & = 0 \quad \text{with probability } (1-\lambda). \end{aligned} \tag{14}$$

Then, the log likelihood function is given by

$$L = \prod_t \ln f(p_t)$$

with the probability density function

$$f(p_t) = \lambda g(p_t \mid I_t = 1) + (1-\lambda)g(p_t \mid I_t = 0)$$

where g is the conditional density given I_t .

By assuming normal distribution $N(0, \sigma^2)$ for the error term v , we can estimate the probability λ and coefficients β of the pricing equation by maximum likelihood. These estimates, in turn, determine the probability of $\{I_t = 1\}$ for each period by the following Bayes' rule:

¹⁰ Here, I follow the procedure taken by Porter (1983). But, on the other hand, Lee and Porter (1984) examines the method where imperfect information (in our case, unanticipated export demand changes) is also taken account of in the regime classification.

¹¹ As Porter (1983) pointed out, if the enforcement mechanism analyzed in the previous section actually works, then, the switches are not independent over time but rather serially correlated. I will soon relax the i.i.d. assumption and allow the correlation by introducing a Markov transition process.

$$Pr\{I_t = 1\} = \frac{\lambda g(p_t | I_t = 1)}{\lambda g(p_t | I_t = 1) + (1-\lambda)g(p_t | I_t = 0)} \quad (15)$$

This probability classifies the whole sample period into the “competitive” phase and the “collusive” phase by

$$I_t = 1 \quad \text{if } \text{Prob}\{I_t = 1\} > 0.5 \\ = 0 \quad \text{otherwise.}$$

Thus, the switching dummy I_t is endogenously determined from the data.

Another interpretation can be given to this probability, though it deviates from the original econometric procedure. Rather than focusing on the rigid threshold level of 0.5, we can interpret this probability as a continuous measure of “relative vulnerability/sustainability” of implicit collusion; i.e., lower probability of $I = 1$ means that the collusion is more sustainable.

The original implication of implicit collusion models, however, may not be captured by thus simple i.i.d. assumption about the switching dummy that I have just introduced. The competitive and collusive phases may persist more than one period because the underlying theory predicts the long enough punishment periods after a deviation to sustain the implicit collusion. Hence, the $\text{Prob}\{I_t = 1\}$ may be serially correlated rather than i.i.d. Consider the estimation including the dummy variable whose process is a Markov chain with the following transition probability:

$$\lambda_{ij} = \text{Prob}\{I_t = j | I_{t-1} = i\} \quad (16)$$

where i and j takes zero or one. After obtaining estimates of λ_{ij} , we can calculate the regime probability of the competitive or collusive phase by

$$\lambda_i = \lambda_{ij} / (\lambda_{ij} + \lambda_{ji}) \quad (17)$$

assuming stationarity. Further, the estimated mean duration of the phase i will be given by $1/\lambda_{ij}$.¹²

3.2.3. Comparison of two approaches

The performances of these two approaches should be compared to reach the final conclu-

¹² I employ the simple estimation procedure by Goldfeld and Quandt (1973), which yields consistent estimators. A more efficient estimator is proposed by Cosslett and Lee (1985). Recently, Ellison (1995) simultaneously estimates pricing/demand parameters and regime shift dynamics.

sion. The Hausman's specification test is applicable to the cases like ours, as Hajivassiliou (1986) pointed out. The estimator with the endogenous dummy does not depend on the information derived from export demand function, while the estimator with the exogenous dummy fully makes use of it. Then, if the switching dummy is really determined by the unanticipated export demand changes, the former is inefficient. The consistency of the latter depends on the correctness of the information from export demand function, while the former is always consistent. Therefore, under the null hypothesis that the export-pricing regime is actually classified by the unanticipated export demand changes, the test statistics defined below follows the χ^2 distribution

$$H = (\hat{\beta}_1 - \hat{\beta}_2)' [V(\hat{\beta}_1) - V(\hat{\beta}_2)]^{-1} (\hat{\beta}_1 - \hat{\beta}_2) \quad (18)$$

where β and V are the coefficient vector and variance-covariance matrix of the maximum-likelihood estimates of the export pricing equation and the subscript 1 and 2 means that the dummy is determined without (with, respectively) employing export demand residuals.

4. The case of Japanese exports

4.1. Data description

The data employed in estimation are briefly described in this section. The appendix will provide additional explanations of them in detail.

Out of the export price indices (p), reported monthly in yen terms by the Bank of Japan, I use the following four classifications: machinery, metals, chemicals and textiles. Though the main reason for my choice of this industrial classification is first-hand data availability, we must note here that there remains a problem in assuming collusion at this level of disaggregation. As will be reiterated in the concluding remarks, the results from this paper must be supplemented in the future by other studies on industries defined narrowly enough to safely assume collusion/punishment mechanism.

To avoid troubles caused by the conventionally used unit-value index, I define export "quantity" (q) as the export value which is reported in trade statistics divided by the corresponding export price " p "

Included as explanatory variables in the regressions are (a) the input cost of each industry, (b) the exchange rate, (c) the producers price indices as a proxy for the price offered by foreign competitors, (d) the industrial production as a proxy for importers' income, since GDP data are not available on a monthly basis, and (e) monthly dummies for seasonal adjustment. Both (c) and (d) are for U.S. and OECD European countries.

The sample period is from January 1976 to December 1988, with 156 monthly

Table 1 ESTIMATION WITH EXOGENOUS DUMMY

$$\ln P_t = \beta_0 + \beta_1 \ln W_t + \beta_2 I_t + v_t$$

$$I_t = 1 \text{ if } u_t < 0, u_t = Q_t - D(P_t)$$

	MACHINERY	TEXTILES	METALS	CHEMICALS
Constant	2.599 (0.421)	-0.956 (0.159)	-0.877 (0.277)	0.844 (0.241)
W	0.431 (0.093)	1.015 (0.354)	1.194 (0.062)	0.810 (0.055)
I	-0.008 (0.013)	-0.006 (0.006)	-0.009 (0.01)	0.003 (0.02)
\bar{R}^2	0.117	0.843	0.708	0.585

(NOTES)

1. \bar{R}^2 is the coefficient of determination after degree of freedom adjustment. Estimated standard errors are in parentheses.
2. Regressors in the export demand function are constant term, export price, exchange rate, total industrial production in U.S. and OECD European countries, producers price index in U.S. and OECD European countries, and monthly dummies. Instruments for two-stage least squares are input cost, and all the regressors excluding the export price.

observations.¹³ This choice enables us to test the theoretical predictions because drastic exchange rate fluctuations / uncertainty during this period provide us with rich information about pricing behaviors.

4.2. Estimation results

4.2.1. Estimates with the exogenous dummy

Table 1 reports the results of the approach employing the exogenous dummy based on the unanticipated changes in export demand.¹⁴ The export demand function of which the estimated residuals form the dummy is estimated by the two-stage least-squares (2SLS). Since export quantity and seasonal dummies are not significant in the export pricing equation, they are omitted from the final estimates. As a result, pricing equations can be estimated by the ordinary least squares method (OLS) because export demand and pricing equations, (11) and (12), form a recursive system.

All the four regressions of the export demand function record high fit and each estimated coefficient has the correct sign as predicted by the theory ($\alpha_1 < 0$, $\alpha_i > 0$ ($i = 2, 3, 4$) in (12)) (not shown in the table to save space).

As Table 1 shows, the coefficient of the switching dummy (β_3) is negative in three

13 While the choice of this sample period was motivated by the data availability at the time of the original research, the results reported here will be informative as an overview for the entire period from the oil shocks to the bubble period, when the aggressive exporting by Japanese manufacturers was hotly debated.

14 All the variables employed in estimation are in the logarithm of the original data. Export demand functions are estimated after taking the first-difference to deal with serial correlation. Export pricing equation is not estimated in the first-difference form. Ellison (1995) refers to some reasons for this, while Cosslett and Lee (1985) consider the case of serial correlation of order one for both equations.

industries. In the chemical industry, the sign of the coefficient is positive but statistically insignificant, as its t-value is the lowest of the four. This seems to be consistent with the demand uncertainty model. The regime dummy coefficient, however, is significantly different from zero only for the textile industry only if we choose a generous 20% significance level. Hence, the relative weakness of the statistical results from the exogenous dummy approach makes us wait for the results of endogenous-dummy estimates to draw the conclusions.

4.2.2. Estimates with the endogenous dummy

The estimation results of the maximum-likelihood without employing information from export demand residuals are shown in Table 2. The iterative calculations converged for all the four equations. The initial value of λ has been set at 0.5, which implies the uniform prior probability about switches. Naturally, the coefficients on the input cost are reasonably positive and significant in all industries. The estimated coefficient on the regime dummy is significantly different from zero in the textile industry at any conventional significance level and also in the machinery industry at the 10% significance level. This suggests that there exist switches in export pricing in these industries. The estimates imply that the export price level in the competitive phase is lower than that in the collusive phase by five percent in the textiles. The quantitative magnitude of price changes across competition phases in this case is roughly comparable with that in Baker (1989) in the case of steel prices, although it is quite moderate compared with that by Porter (1983), which reports price variation of more than 60%.¹⁵

**Table 2 ESTIMATION WITH ENDOGENOUS DUMMY
(SWITCHING REGRESSION)**

$\ln P_t = \beta_0 + \beta_1 \ln W_t + \beta_2 I_t + v_t$ $\text{Prob}\{I_t = j \mid I_{t-1} = i\} = \lambda_{ij} (i, j = 0, 1), v_t \sim N(0, \sigma^2)$				
	MACHINERY	TEXTILES	METALS	CHEMICALS
Constant	2.599 (0.006)	-0.074 (0.009)	-0.869 (0.180)	0.845 (0.142)
W	0.430 (0.001)	0.020 (0.002)	0.194 (0.011)	0.810 (0.270)
I	-0.009 (0.006)	-0.051 (0.020)	-0.018 (0.292)	-0.001 (0.088)
σ	0.077 (0.004)	0.0365 (0.002)	0.080 (0.004)	0.130 (0.034)
λ_1	0.293	0.915	0.815	0.667
λ_0	0.707	0.085	0.185	0.333
λ_{01}	0.369	—	0.759	0.314
λ_{11}	0.110	—	0.827	0.843
λ_{10}	0.890	—	0.173	0.157
λ_{00}	0.631	—	0.241	0.686

15 I have to note that the significance in the machinery industry is relatively marginal in that the likelihood ratio test rejects the null hypothesis of no switches only in the textiles industry. This results in a moderate price change across competition phases. The export price in the competitive period is lower than that in the collusive period by 0.94% (machinery), 5.24% (textiles), 1.84% (metals), and 0.08% (chemicals).

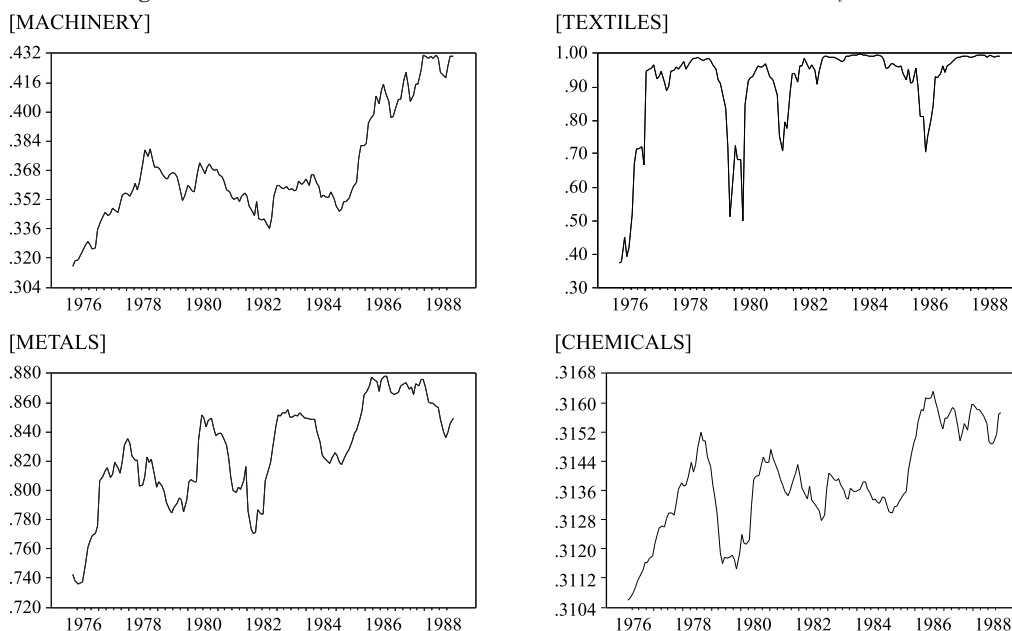
The probability of being in the “competitive” phase (λ_1) varies from industry to industry: lowest in the machinery industry ($\lambda_1 = 0.29$) and highest in the textile industry ($\lambda_1 = 0.92$), which may imply the industrial difference in the difficulty of coordinating price decisions.

By introducing the Markov process into the switching dummy series, richer results become available.¹⁶

For example, the probability of collusion breakdown in the machinery industry ($\lambda_{01} = 0.37$) is approximate to that in the chemical industry (0.31), but the regime probability of a competitive period (λ_1) is much lower in the machinery industry ($0.29 < 0.67$), because price wars in the chemical industry are more persistent ($\lambda_{11} = 0.84 > 0.11$). In other words, estimated mean duration of price war ($1/\lambda_{10}$) is much longer in the chemical industry than in the machinery industry (6.39 months $>$ 1.12 months), although the collusive phase also persists longer ($1/\lambda_{01} = 3.19$ months $>$ 2.71 months).

In the metal industry, the probability of continuing a price war ($\lambda_{11} = 0.83$) is almost the same as that in the chemical industry (0.84), but the collusion is much more volatile ($\lambda_{01} = 0.76 > 0.31$). As a result, the expected duration of the collusive phase is quite short ($1/\lambda_{01} = 1.32$ months), and price wars persist long ($1/\lambda_{10} = 5.79$ months).

Fig. 1 RELATIVE VULNERABILITY OF COLLUSION (Prob $\{I_t = 1\}$)



16 As for the textile industry, only the result of i.i.d. Bernoulli case is reported in Table 2 because the calculation allowing Markov transition did not converge within reasonable number of iterations.

The probability computed by (15), or the “relative vulnerability” of implicit collusion is drawn in Figure 1 for the four industries. A glimpse of the figures gives us some clear features as follows; First, the collusion in the machinery industry becomes significantly vulnerable after the drastic yen appreciation process since 1985, although the probability of $I_t = 1$ is always smaller than 0.5, which means no all-out price wars in the sample period. Second, in the textiles industry, firms charge competitive prices almost all the time in the sample period with some exceptional periods of short-lived collusion. Third, the sustainability of collusion in the metal industry decreases almost monotonically over time rather than being characterized by periodical switches. Lastly, the variation of collusion vulnerability in the chemical industry over time is quite small, $\text{prob} \{I = 1\}$ moving from 0.310 to 0.317, with a relatively big rise around 1978.

These industrial features summarized above may find intuitive interpretations in the following way; First, industries with substantial market power in export markets such as the Japanese machinery industry and oligopolistic industries such as the chemicals tend to have a lower probability of $I = 1$. Second, industries with less market power and/or with larger numbers of firms are likely to be in the competitive phase for a longer time as in the case of the textiles industry. Third, the sustainability of collusion may decline over time as the market power of the member firms diminishes in the global market, as in the case of the Japanese metal industry. Lastly, the sustainability of collusion in industries where cost changes dominate the pricing tends to be strongly affected by such episodes as the “oil shock” (e.g. the chemicals in 1978).

Thus, although none of them should be exaggerated without industrial case studies that investigate background information including historical and institutional factors surrounding export pricing decisions,¹⁷ we have obtained the estimated sustainability of implicit collusion which is consistent with reasonable industrial features.¹⁸

4.2.3. Specification test

Combining the results of both approaches suggests that unanticipated export demand changes affect the sustainability of collusion in the textiles industry. Therefore, it is useful to conduct the specification test discussed in Section 3.2.3 for the textiles industry to determine

17 Since a substantial share of Japanese exports was subject to trade-protection-related measures (for example, automobiles, steels and textiles), the results obtained in this section might be distorted.

18 To study determinants of the phase of competition (I_t), I conduct probit regressions with I_t as the dependent variable. Main results are as following. (a) The negative residuals in export demand raise the probability of collusion breakdown significantly in the textile industry. (b) In all industries, current yen appreciation significantly triggers the emergence of competitive pricing. (c) The results from regressions on current or future expected export size vary across industries and are not decisive. The analysis of expected exchange rates is worth independent work.

whether the unanticipated export demand change is *the* determinant of pricing regime switches.

The test statistics H defined in (18) is calculated as $H = 6.1811$. Under the 2.5% significance level, the null hypothesis that unanticipated export demand change is the correct regime classification information cannot be rejected.¹⁹ Therefore, beyond other factors, the demand uncertainty effect well characterizes the dynamic export price competition in the textile industry.²⁰ However, this result should be interpreted with caution because we observe only three significant episodes of regime shift in the sample period.

5. Concluding remarks

The econometric study in this paper suggests, at least as a first approximation to some of the industries examined, the relevance of dynamic oligopoly model in export pricing.²¹ Switches between collusion and punishment periods are observed in the textile export price, supporting the model that predicts that unanticipated negative shocks to industry demand triggers a breakdown of collusion. In other industries, although implied price changes are too moderate for us to characterize export pricing as switches between states, dynamic changes in the relative sustainability of collusion are consistent with reasonable industrial characteristics.

Divergence of the results across industries is not embarrassing but rather reasonable because the game theory from which we draw predictions on export pricing behaviors heavily depends on the industry-specific factors. This, however, implies the needs to further excavate the detailed information surrounding the pricing decisions in each industry. In this sense, my choice of aggregation level, which is larger than the minimum necessary to achieve collusion, might have clouded the results in this paper. Since the theory of implicit collusion assumes that exporting firms have high market power in the world market of narrowly defined products, we need to check the plausibility of this assumption. Actually, the success of researches triggered by Porter (1983) in empirical industrial organization critically depends on the database, which are sufficiently information-rich to investigate the regional railroad cartel.

19 The maximum-likelihood estimates with the exogenous dummy which is used in calculating H are not shown in the tables, but are quite similar to those by two-stage least squares in Table 1.

20 One of the possible reasons for this industrial difference could be found in the differing role of trading companies. The export of Japanese textiles has been often handled by trading companies, while manufacturers are directly involved in foreign trade of their own products in other industries. Hence, even if many manufactures compete, oligopolistic pricing can be observed in the textile industry if a small number of trading companies are the price setters.

21 This does not at all imply that other alternative explanations are rejected. Dynamic effects of sunk entry costs, for example, cannot be neglected for interpreting real-world exporting behaviors. See Krugman (1987) for an essence of various models.

Therefore, what must be required next after my preliminary attempt is an econometric study using more disaggregated data of narrowly defined industries.²²

Acknowledgement

This paper is a revised version of the first chapter in my Ph.D. thesis at M.I.T. The author acknowledges Paul R. Krugman, Julio Rotemberg and Rudiger Dornbusch for their insightful suggestions to the original thesis. The author also appreciates valuable comments from seminar participants at Kobe University, especially Seiichi Katayama, in revising the earlier version published as MITI-RI Discussion Paper, for which Masato Hisatake gave useful comments. All remaining errors are mine.

Appendix

[Export price] (p)

FOB price in terms of yen (Laspeyres index).

Source: Bank of Japan, *Economic Statistics Annual*, various issues.

Classifications: (1) Textiles, (2) Metals and related products, (3) Chemicals, and (4) Machinery (including transport equipment).

[Export quantity] (q)

Export “quantity” is defined as export value divided by export price (p).

Source: Ministry of Finance, Government of Japan, *Custom Clearance Statistics*.

Classifications: same as those of export price.

[Input cost] (w)

Input price based on Input-Output table (wholesale price for domestic input and import CIF price for imported input)

Source: Bank of Japan *Economic Statistics Annual*, various issues.

Classifications: (1) Textiles, (2) Basic metal products and metal products, (3) Chemicals, and (4) Machinery and equipment.

[Exchange rate] (e)

The yen-US dollar rate is the monthly closing rate (yen per dollar).

Source: Bank of Japan, *Economic Statistics Annual*, various issues.

The nominal effective exchange rate of the yen is from International Monetary Fund, *International Financial Statistics*.

[Foreign producer price and industrial production] (PUS, PEU; IUS, IEU)

²² The application of the implicit collusion model to international trade contexts remains limited. For example, even in the recent empirical study on micro-level strategic export pricing by Kadiyali (1997), no supergame model is considered.

Producer price and total industrial production in the US (PUS, IUS) and in OECD European countries (PEU, IEU).

Source: OECD (1990) *Main Economic Indicators: Historical Statistics 1969-1988*.

References

- Bagwell, K. and R. Staiger [1997], "Collusion over the Business Cycle," *RAND Journal of Economics* 28, Spring, pp. 82-106.
- Baker, J. [1989], "Identifying Cartel Policing under Uncertainty: the U.S. Steel Industry, 1933-1939," *Journal of Law & Economics* XXXII S, pp. 47-76.
- Campa, J. M. and L. Goldberg [2002], "Exchange Rate Pass-through into Import Prices: A Macro or Micro Phenomenon?" NBER Working Paper No. 8934.
- Cosslett, S. R. and L-F. Lee [1985], "Serial Correlation in Latent Discrete Variable Models," *Journal of Econometrics* 27, pp. 79-97.
- Dornbusch, R. [1987], "Exchange Rates and Prices," *American Economic Review* 77, pp. 93-106.
- Ellison, G. [1995], "Theories of Cartel Stability and the Joint Executive Committee," *RAND Journal of Economics* Vol. 25, No. 1, Spring, pp. 37-57.
- Goldfeld, S. M. and R. E. Quandt [1973], "A Markov Model of Switching Regressions," *Journal of Econometrics* 1, pp. 3-15.
- Green, E and R. Porter [1984], "Noncooperative Collusion under Imperfect Price Information," *Econometrica* 52, pp. 87-100.
- Hajivassiliou, V. A. [1986], "Two Misspecification Tests for the Simple Switching Regressions Disequilibrium Model," *Economics Letters* 22, pp. 343-348.
- Kadiyali, V. [1997], "Exchange rate pass-through for strategic pricing and advertising: An empirical analysis of the U.S. photographic film industry," *Journal of International Economics* 43, pp. 437-461.
- Krugman, P. [1987], "Pricing to Market when the Exchange Rate Changes," in S.W. Arndt and J. D. Richardson eds. *Real-Financial Linkages among Open Economies*, pp. 49-70, M.I.T. Press.
- Lee, L.F. and R. Porter [1984], "Switching Regression Model with Imperfect Sample Separation Information: with an Application on Cartel Stability," *Econometrica* 52, pp. 391-418.
- Marston, R. [1990], "Pricing to Market in Japanese Manufacturing," *Journal of International Economics* 29, pp. 217-236.
- Porter, R. [1983], "A Study of Cartel Stability: the Joint Executive Committee, 1880-1886," *The Bell Journal of Economics* 14, pp. 301-314.
- Rotemberg, J. and G. Saloner [1986], "A Supergame-theoretic Model of Business Cycles and Price Wars during Booms," *American Economic Review* 76, pp. 390-407.
- Rotemberg, J. and M. Woodford [1991], "Markups and the Business Cycles," in O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual 1991*, pp. 63-129, Cambridge, MA: MIT Press.
- Tomiura, E. [1992] "Export Pricing and Sustainability of Noncooperative Collusions," Chapter One, *Three Essays on Dynamic Export Competition*, pp. 12-58, Ph. D. Dissertation, M.I.T.

**NATIONAL MONEY OF ACCOUNT, WITH A SECOND NATIONAL
MONEY OR LOCAL MONIES AS MEANS OF PAYMENT:
A WAY OF FINESSING THE ZERO INTEREST RATE BOUND.**

STEPHEN J. DAVIES Kobe University

Abstract

This paper shows how negative interest rates can be implemented — at much less cost than under alternative proposals — by separation of the unit of account and means of payment roles of money — i. e. by having these roles performed by different moneys. Such separation has been far from unusual in history, however unnatural it may seem to those whose experience is of modern non-separated monetary systems. One possibility discussed is that the means of payment role of cash could be provided by local currencies rather than a national currency.

JEL Classification: B59, E31, E42, E43, E50, E52, E59

Keywords: Negative interest rates, Money of account, Local currencies

“If, other things remaining the same, the leading banks of the world were to lower their interest rate, say 1 per cent below its ordinary level, and keep it so for some years, then the prices of all commodities would rise and rise and rise without any limit whatever; on the contrary, if the leading banks were to raise their interest rate, say 1 per cent above its normal level, and keep it so for some years, then all prices would fall and fall and fall without any limit except Zero.”¹

“The government need not produce dollars in order to define the dollar, any more than it has to produce yardsticks in order to define the yard.”²

Corresponding Address:

Stephen J. Davies

Research Institute for Economics and Business Administration (RIEB), Kobe University

2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501 Japan

E-mail: sdavies@rieb.kobe-u.ac.jp

Phone/Fax: 81-78-803-7030

1 Wicksell (1907).

2 Hall (1999).

The motivation for this paper³ starts from the observation that, having reached a state of technology in which an advanced country could, if necessary, conduct its business without using cash, Japan has allowed the continued existence of cash — through the zero lower bound on interest rates that cash as we know it apparently entails — to cause it to “lose a decade”.

We can certainly envisage a world entirely without cash; indeed many eminent monetary economists have, as a matter of relatively urgent practical relevance, been debating in recent years the nature of such a world and in particular whether the monetary authorities will be able to control a meaningful interest rate in such a world.⁴ For many types of transactions, people in advanced economies can much of the time do perfectly well without cash already. In the UK, non-cash payments account for 58 per cent of retail payments by value and 25 per cent by volume in 2002: the volume share is forecast to rise to 38 per cent by 2009.⁵

Those who have most need of cash are drug-dealers and tax-evaders.⁶ (Also, apparently, drug-users: a 1999 survey showed that 99 per cent of Bank of England notes were contaminated with cocaine, and most euro notes were also contaminated with cocaine within a few months of the introduction of euro notes and coin.) Yet, as discussed more fully later in this paper, governments everywhere continue to allow the existence of cash to exercise a constraint on national monetary policy, in the form of the zero bound on short-term interest rates.

Although a cashless economy is imaginable, cash certainly still has some advantages even in certain legitimate contexts, and it is not my intention here to advocate the abolition of cash. The purpose of this paper is to show how we can both have cash and at the same time remove the zero bound on interest rates.

The zero bound on interest rates is not just of theoretical interest. Various estimates of interest rate rules for Japan show that interest rates needed to be negative for substantial periods of time from the mid 1990s.⁷ The Bank of Japan has held overnight interest rates close to zero from early 1999 up to the time of writing (September 2004). It is inconceivable that they would not have cut interest rates below zero for at least some of this time if they had thought this option open to them. As the quote above from Wicksell suggests, if the rate

3 The proposal in this paper is essentially that put forward in an earlier paper: Davies (2002). However, much of that earlier paper was devoted to an inevitably over-simplified discussion of a range of alternative analyses of Japan’s “lost decade”, and this will have distracted some readers from the technical proposal it made. This new paper contains an account of historical precursors to and parallels with my suggestion, of which I was not fully aware when I wrote the earlier paper. I have also benefited from the comments of Professor Toshiki Jinushi of Kobe University on the presentation of the argument in an intermediate paper.

4 See, for example, Friedman (1999).

5 Source: Association for Payment Clearing Services (APACS): <http://www.apacs.org.uk/>

6 Humphrey et al. (2004) estimate that illegal use of cash and hoarding of cash accounted for 67 per cent of total demand for cash in Norway in 2000.

7 E.g. McCallum (2002), Okina and Shiratsuka (2001).

of interest is too high (e.g. zero when it should be negative) the economy may simply move further and further away from equilibrium. There is no presumption that the economy will be stable if interest rates themselves cannot adjust in the required direction (though, of course, the dynamics depend on the way expectations are formed).

The central point made in this paper that the zero bound on interest rates is actually quite easy to get round. All that is needed is to escape from the frame of mind that takes it as inevitable that the monetary unit of account is identical to the means of payment. That it need not be is the point of the quote from Robert Hall above.

Separation of the means of payment and unit of account in medieval Europe

Most of us so take for granted the identity of cash with the unit of account that some historical digression is helpful both to explain what is meant by separation of the means of payment and unit of account and to demonstrate that separation is perfectly feasible.

The monetary systems that prevailed for hundreds of years in much of continental Europe were based on the monetary units established by Charlemagne: pound, shilling and penny.⁸ In these systems there certainly was material money in circulation: gold and silver coins, as well as token coins of lead; but these were not in the same denomination as the unit of account. As described by Kaye, “money of account functioned as an idealized monetary scale of artificially fixed ratios of named coins that were often no longer in circulation, against which the actual value of the coin in circulation was measured”.⁹

Spufford puts it thus: “In most parts of late medieval Europe...a dichotomy existed in the functions of money. On the one hand, money of account was the *measure of value*, whilst on the other, the actual coin was the *medium of exchange* and the *store of wealth*...”

Most financial transactions were first determined and expressed in money of account, although payments were naturally made subsequently in coin, or surprisingly often in other goods. Coin itself was valued as a commodity in terms of money of account, and, like any other commodity, its value frequently varied.”¹⁰

As Spufford’s reference to payment in other goods implies, a money of account does not necessarily entail the use of monetary payments. A monetary unit of account is convenient even in a barter economy, to summarise debts and assets in a “favour bank”. (So that someone can be said to owe 10 pounds, rather than 3 chickens plus one hammer plus two massage sessions, etc.)¹¹

8 In French: livre, sou and denier. A pound was equal to twenty shillings; a shilling was equal to twelve pennies.

9 Kaye (1998), p.190.

10 Spufford (1986), p. xx.

11 Kocherlakota (1998) argues that “fiat money’s only technological role in an economy is to act as *societal memory*: money allows people to credibly record some aspects of their transactions and make that record accessible to other people”. Clearly money of account can fulfil this role, without need for fiat money in material form.

In parts of continental Europe the separation of means of payment and unit of account continued well beyond the middle ages. Thus in France, for example, right up to the Revolution, coins in specie had a “price” in terms of the unit of account, a price that was changed from time to time by royal decree. Announcing a five per cent increase in the price of specie (so that a particular gold coin would exchange in future, say, for one pound one shilling instead of for one pound) meant that the price (in pounds) of all commodities would rise sooner or later by around one twentieth, and would reduce the real value of debts (including, of course, royal debts) denominated in pounds. (This was the standard mechanism, rather than changing the fineness or weight of coins, by which French kings used the inflation tax to improve their finances.) Under this system, inflation (measured in terms of the unit of account) is roughly the sum of any change in the price of specie relative to other prices (caused by discoveries of new mines, or exhaustion of old ones, say) and the change in the unit account price of specie.¹²

The fullest discussion of the history in Europe of this separation of the unit of account from money used in payment is in a classic paper by Einaudi (who uses the term “imaginary money” for money of account that has no cash form):¹³

“Today each country has only one monetary unit: the lira, the franc, mark, pound sterling, or dollar. This is the system established by the French assemblies at the end of the eighteenth century...Prior to the French Revolution, the monetary system of most European countries was based on altogether different principles....

In the time of Malestroit, a customer could pay a shopkeeper for an ell of velvet, priced at 10 pounds tournois by giving him 4 ecus de soleil rated at 2L 10s each. Similarly, the buyer of a barrel of wine, costing 12 pounds, could give in payment 20 testoons current at 12 sous per testoon. A shoemaker would be satisfied if he received 15 douzains in black money, at 12d. per douzain, for a pair of shoes selling at 15s.... If there was a change in the ratio between real and imaginary money...the number of coins to be paid in discharge of a debt would vary inversely.”

The “New Monetary Economics”

This account of historical experience has established the feasibility of the separation of unit of account and means of payment. In raising the possibility of separation, and also in questioning the essential role of cash, I am echoing some of the points made by the school of thought known as the “new monetary economics” (NME). The three essential ideas of this

12 A monetary reform of 1577, under which the ecu gold coin became the unit of account, was abandoned in 1602, when separation between the unit of account (the livre) and the means of payment was reinstated. See Sargent and Velde (2002), pp. 208-212.

13 Einaudi (1953).

school are, according to the relevant entry¹⁴ in the New Palgrave:

- Monetary and monetary institutions derive their special status from regulations imposed by governments
- Different regulations would produce radically different financial and monetary arrangements
- **The function of money as the unit of account may be separated from the function of money as the means of payment** (my emphasis).

Although the NME school was only identified as such in the 1980s, some of the measures it advocated (including the separation of the unit of account from the means of payment) can be traced back to the writings of nineteenth century economists.¹⁵ The school also tends to regard cash as non-essential: Fischer Black¹⁶ and Eugene Fama,¹⁷ in two of the best-known contributions to the NME literature, both discussed the possibility of economies in which cash had ceased to exist.

The contribution to this literature of most interest to me, however, is that of Greenfield and Yeager. They specifically referred to the historical evidence of separation of unit of account and means of payment when, in an article published in 1983, they proposed a radical monetary reform under which the national monetary authorities would define a unit of account (in terms of a bundle of commodities) but would not issue any money.¹⁸ (Note that there is not merely an absence of cash in their proposal, there is an absence of bank reserves as well; there is no monetary base. This is crucial for the feasibility of anything like traditional banking in the world they describe.)

The authorities would **define** a pound, say, as equal to one apple plus one banana plus one cherry. But they would not issue any notes or coins corresponding to pounds, or multiples or fractions of pounds; indeed, they would not issue any notes and coins. They called this monetary system the “BFH” system, after Fischer Black, Eugene Fama and Robert Hall, who had all at that stage contributed to the NME literature.

This BFH system “would get rid of any distinct money existing in a definite quantity. The government would be forbidden to issue obligations fixed in value in the unit of account and especially suitable as media of exchange”. In the absence of money, problems of monetary policy would disappear. “No longer, then, could there be too much of it, causing price inflation, or too little, causing depression, or a sequence of imbalances, causing stagflation. A

14 Harper and Coleman (1992).

15 As shown by Cowen and Kroszner (1987).

16 Black (1970).

17 Fama (1980).

18 Greenfield and Yeager (1983).

wrong quantity of money could no longer cause problems because money would not exist.”

Greenfield and Yeager were at pains to distinguish their proposed system from a commodity standard in which money is issued and is convertible into a bundle of commodities (or a single commodity, typically gold) at a fixed ratio. In the BFH system, a pound would be **defined** as a bunch of commodities; but neither the central bank nor any other official body would have any obligation to buy or sell any commodities in exchange for pounds. They would thus not need to store commodities. The pound would have a relatively stable value, and thus be an attractive unit of account, not because the authorities take any action to ensure its stability, but because the aggregate price of the bunch of commodities it represents — if broadly enough defined — is likely to be relatively stable in relation to the average price of other goods and services bought and sold in the economy.

The two main criticisms of Greenfield and Yeager in subsequent literature related to the sustainability of the unit of account (which I discuss later) and to the nature of the means of payment in the BFH system. By assumption, there is in the system no outside money; no cash that is the liability of the government or central bank. This lack of outside money also means that banks cannot create inside money in the conventional way. In the world we are familiar with, banks normally come into being as a result of customers depositing outside money. The deposits are then a claim to receive back outside money in due course. Banks cannot develop in this way in the BFH system, because there is no outside money. Even in the absence of cash, if the central bank bought private sector securities by crediting bank accounts at the central bank, normal banking could develop on the base of these reserves. But as noted already, there are no reserves of this sort created by government; there is no monetary base at all.

Nevertheless, Greenfield and Yeager insist that transactions in their economy would not involve “*crude barter*” (their italics); money market funds offering payment services would develop. (They were writing at a time when, in the US, regulation Q — the restriction on interest rates paid on bank current accounts — had led to the widespread availability of money market funds offering transaction services.)

One can think of mutual funds developing in the following way. A company might grant a mutual fund ownership of part of the firm, in return for which the company would receive shares in the mutual fund. The company would then use the shares in the mutual fund to remunerate its employees who would thereby become shareholders of the mutual fund. The mutual fund would offer transaction services that would enable shareholders to use their shares to purchase goods and services from other companies.

One problem with this arrangement is the risks to which the workers would be exposed through movements in share prices. A transactions medium is needed that is suitable for low net worth individuals (who might be spending the whole of their shareholding between regular salary transfers, and so would be severely affected by a fall in its value) as well as

the more affluent (whose shareholding might largely represent saving rather than transaction balances).

A second problem is the need for settlement between different mutual funds (to offset net payments by one fund's share holders to another fund's shareholders). This settlement has apparently to be effected by a transfer of shares between funds: but that means a constant change in the risks that a particular fund is exposed to as a result of changes in shareholdings in the course of settlement. The fund would seem to have to accept these risks passively with no chance of managing them, and possibly without the information to assess them.¹⁹

The background to Greenfield and Yeager's proposal was of course the high inflation of the 1970s and a belief (widely held at the time, but seemingly mistaken, given subsequent developments) that fiat money controlled by governments or central banks inevitably led to inflation. If we fast-forward to the late 1990s, we come to a literature that does not advocate a cashless system as a way of avoiding inflation; instead it contemplates the possibility that technical change in the private sector will make cash redundant, and discusses the consequences for monetary policy.

One strand of the recent literature is about whether cash is actually likely to become extinct, and the generally accepted answer is "probably not" (and not just because of the demand for cash for illegitimate activities).²⁰ Another strand is whether the monetary authorities would be able to maintain meaningful control of interest rates in the case that cash did become extinct: a question that is now, following Woodford's influential analysis, generally answered in the affirmative.

The feasibility of negative interest rates

How do monetary authorities control interest rates? The precise institutional details vary from country to country, but generally they do so by forcing leading players in the money market to borrow from them at an interest rate of the authorities' own choosing. Woodford describes how, even in a cashless economy with no requirement on banks to hold reserves at the central bank, the central bank can nevertheless exert tight control on short-term money market interest rates through the interest rate it sets on deposits by and lending to its counterparties. If banks can get, say, a return of 5 per cent on deposits with the central bank, then they will not lend out money for less than this on the interbank market. And if they can

¹⁹ Greenfield and Yeager see the replacement of banks by mutual funds as a significant advantage in its own right: it would create a financial system that was immune to the financial instability that can be caused by runs on banks. This is however really a separate issue. The replacement of banks by money market mutual funds has been advocated by others (e.g. Goodhart (1993)), but within an economy where cash circulates in the normal way.

²⁰ See e.g. Goodhart and Krueger (2001).

borrow overnight from the central bank at 5.5 per cent, they will not pay more than 5.5 per cent to borrow from the interbank market. Thus, with these interest rates set by the central bank, overnight interbank rates will not move outside a 5 to 5.5 per cent “channel”.²¹

Does the central bank’s ability to set interest rates extend to setting negative interest rates? While the central bank may set a negative interest rate on deposits from and lending to its counter-parties, if it wants that interest rate to feed into the rest of the economy it cannot supply on demand cash that does not depreciate in terms of the unit of account. Either it has to ration the cash it supplies, or it has to supply a form of cash that depreciates in terms of the unit of account (the unit used, for example, in the accounts of the commercial banks at the central bank), and that depreciates at a rate at least equal to the negative interest rate set on lending by the central bank. Providing one of these conditions on the supply of cash is met, variations in the central bank’s (negative) deposit and lending rates will feed automatically into interest rates throughout the economy.

Obviously if neither of these conditions is met, the central bank will not succeed in pushing negative interest rates throughout the economy. Banks will simply “round-trip”: they will use their borrowings from the central bank to invest in cash (provided the borrowing rate is sufficiently negative to cover security, insurance and storage costs). There is no reason why they would lend at negative interest rates when they could invest in cash at a zero interest rate (on the reasonable assumption that the risk of losses on any lending are at least as great as the cost of storing and insuring cash holdings). The central bank would face potentially unlimited losses on the round-tripping (an unbounded balance sheet, with negative interest rates on its assets and a zero interest rate on its main liability).

This source of a lower bound on interest rates is not one that needs to be or can be settled by empirical studies. The stability and shape of the money demand function at very low positive interest rates is certainly of interest and obviously open to empirical investigation.²² Such investigation may in principle discover whether or not there are factors that prevent the interest rate falling to zero (for Japan, of course, we now have the practical experience of short-term interest rates virtually at zero). However, even at exactly zero interest rates there is no incentive for round-tripping by banks. Interest rates have to fall marginally below zero for this to happen at all. Empirical evidence on what happens as interest rates fall close to zero can therefore provide no evidence at all to confirm whether or not this round-tripping will occur. The argument that, with unlimited supplies of non-depreciating cash, interest rates cannot fall (more than marginally) below zero thus has to be based on the hypothesis that banks will not ignore the opportunity to make riskless profits rather than on empirical evidence.

21 Woodford (2003), p.33.

22 See for example Miyao (2002).

Proposals have in the past been made for taxing cash, by Gesell²³ in the early part of the last century and by Goodfriend²⁴ at its end, and a tax on cash would certainly make negative interest rates feasible. A 5 per cent per annum tax on cash, for example, would create room for a central bank to force short term money market rates down to around – 5 per cent. However, these tax proposals have generally been regarded as impractical.

Gesell's proposal involves bank notes being required to have stamps affixed at regular intervals to maintain validity. The cost of the stamp corresponds in effect to a tax on cash, or equivalently a negative interest rate on cash. The value of a note would be equal to its face value only immediately after the latest stamp was affixed; thereafter it would be subject to an increasing discount until the next stamp was due. Although Gesell's scheme has been implemented in **local** scrips (and was endorsed by Irving Fisher as an emergency measure in the 1930s depression in America) it would clearly be extremely cumbersome and costly if adopted for **national** currency. (That, however, was certainly Gesell's intention, and I am not aware that Gesell had any interest in what are now known as "local currencies"; although his name has become closely associated with the local currency movement, particularly in Japan and to a lesser extent in other countries.)

Goodfriend's proposal involves a magnetic strip on the back of notes that would record the date each time a note was issued at a bank or ATM etc. When next paid into a bank the face value of the note would be discounted by a sum proportional to the latest period of circulation. Whereas in Gesell's proposal all current bank notes would exchange at a discount that would vary over time, but that would at any time at least be the same for all bank notes, under Goodfriend's scheme bank notes would also be depreciating, but different bank notes with the same face value would be subject to different discounts, according to the length of time they had been in circulation. So to calculate the value of a bundle of bank notes tendered in payment one would have to check individually the date on which each had been issued, work out the discounted value of each one, and sum to obtain the total value (or acquire a machine to carry out the operation).

Neither of these proposals, incidentally, has anything to say about coins, which are not amenable to either form of taxation. If coins remained untaxed there would be obviously be some (inefficient) increase in their use, though the authorities could minimise that increase by replacing higher value coins by notes.

Some Japanese commentators and academics publicly advocate the "Gesell tax" as a solution to Japan's economic problems without fully confronting the practical problems of implementation. For example, Fukao in a paper presented at BIS in 2002, argued for a

23 S Gesell *The Natural Economic Order*, available at <http://www.systemfehler.de/en/neo>. See also Buiter and Panigirtzoglou (2003). However, while earlier versions of this paper discussed the implementation of an updated form of the Gesell tax, the authors have omitted all discussion of practicalities from the final version.

24 Goodfriend (2000).

“Gesell Tax on Government Guaranteed Assets” which included the proposal that “bank notes should be taxed. In order to tax cash the Bank of Japan has to print new bank notes and levy fees for exchange. Alternatively, the government can levy stamp duty on old bank notes.”²⁵ This technique for taxation of cash is certainly possible, but Fukao does not acknowledge that the conversion of bank notes could not be a one-off event. After one exchange of bank notes, a further exchange has to be expected, in order for the expected rate of return on cash to be below zero. Each conversion would, of course, involve not only the costs of issuing the new notes, but also the cost of altering every vending machine in the country so that it would accept the new notes and reject the notes that were no longer current.

A key — but generally unacknowledged — point about both Gesell’s original proposal and Fukao’s variant on it (and in a more complicated way Goodfriend’s proposal) is that they necessarily involve the separation of the unit of account from the means of payment. Immediately prior to one of Fukao’s proposed conversions, an old 10,000-yen note could obviously not be worth 10,000 yen (as it would be exchanged for a new note, worth at most 10,000 yen, only after some payment). Clearly also the value of an old note would not jump, so there must be an extended period in which the value of an old note would be at a discount — and an increasing discount — to its face value. In Fukao’s scheme, if there were regular but not too frequent conversions and the payment for conversion was below some critical level, there would be periods in which the notes traded at par followed by periods in which notes traded at an increasing discount to par. During the periods when the notes were trading at par, overnight and other very short-term interest rates could not be significantly below zero, although longer term interest rates could be below zero, reflecting expectations of future negative short-term interest rates.

Suppose the authorities announce that in two years time the present note issue will become worthless, prior to which it will be exchangeable with new notes at a rate of 9,000 yen for each old 10,000-yen note. Suppose for simplicity that it is expected that immediately on issue a new 10,000-yen note will be worth 10,000 yen. If interest rates on bank deposits are, say, zero, then this equates to a similar relative return on cash to that seen when the interest rate on bank deposits is 5 per cent and cash yields a zero return. That has not been an unusual situation in the UK in recent decades (indeed interest rates have often been considerably higher than 5 per cent); but people do not stop holding cash because of such interest rates, or indeed markedly reduce their cash holdings: in econometric work, it has not been straightforward to find a significant interest rate effect on holdings of cash at all.²⁶

Thus it is entirely conceivable that, with a 10,000-yen note legal tender for 10,000 yen, two years or even one year before the conversion, 10,000-yen notes could continue to be

25 Fukao (2003).

26 A recent Bank of England paper, by Grant et al. (2004), finds a stable interest rate effect on demand for cash over the ten years ending in 2002, but not for the previous decade.

exchanged at par. But as the date of exchange approached, willingness to respect the legal tender law would surely break down. The day before the conversion, the negative return on the old cash (valued at par) is close to 100 per cent per annum, so no one is going to want to hold cash overnight on the day before exchange, if they can deposit it at par. Can everyone just deposit cash with their bank, and the banks return it to the central bank, without penalty, the day before conversion? No doubt the economy can perfectly well survive one day without cash, but this arrangement would defeat the whole point of the exercise; it would mean that the effective interest rate on cash was zero for all but the last day before conversion, so that money market interest rates could not be significantly negative, except for that brief period.

But in the absence of some way to avoid the penalty on conversion and the almost — 100 per cent per annum interest rate on the last day, by recursion there would be an unwillingness to accept notes at par for some period prior to the conversion. Over this period there would be a growing reluctance to respect legal tender provisions — if these recognised cash as settling debts at par — and cash would be increasingly inconvenient as a means of payment.

The above discussion has left out the role of government transactions in affecting whether cash would be valued at par or otherwise in the periods between conversions. While it is possible, as an intellectual exercise, to think of the government and central bank as price takers, willing to accept cash at par when the economy as a whole accepts cash at par, and requiring the same discount as that generally prevailing in the market when cash trades below par, such an arrangement is hardly plausible. But if the government announces a rate at which it will accept cash, in payment of tax bills, for example, then that rate is likely to be highly influential, to say the least. I will develop this point further below.

What Fukao proposes (and the proposal is similar to one made by Buitter and Panigirtzoglou) in effect gets round the problem of the excessive administrative costs of very frequent small tax payments on currency in Gesell's original scheme with much more infrequent but much larger tax payments. The administrative expenses are indeed greatly reduced, but only at the cost of much greater disruption to the normal use of cash in the period running into the day when the tax is paid (the day of currency conversion).

Thus, while I am fully in sympathy with Professor Fukao's objectives (at least as they applied to the state of the Japanese economy when he presented his paper), I believe it would be desirable to achieve them in a less disruptive way. I now present a way of achieving the objectives, simply exploiting the possibility of separating the unit of account and means of payment, as discussed in the historical digression above, and not requiring any taxes or repeated conversions of currency notes.

Two schemes for separating the means of payment and unit of account

I present first a scheme — “Scheme A” — which is not actually my preferred scheme, but which is simpler to understand than my preferred scheme, and also facilitates the understanding of my preferred scheme. Under Scheme A, the monetary authorities would withdraw all existing cash while maintaining the existing monetary unit of account; they would use interest rate policy to achieve an inflation target expressed in terms of this unit of account, as discussed by Woodford.²⁷ Separately, to facilitate transactions, the authorities would make a new issue of cash, which could depreciate in terms of the unit of account. The authorities would, of course, have complete control over the “exchange rate” between cash and the unit of account.

To be more specific, taking the case of the UK as an example, the Bank of England and the Debt Management Office would continue to denominate operations with their counterparties in pounds sterling, the Government would continue to pay salaries into its employees’ bank accounts in pounds sterling, and to send out tax bills expressed in pounds sterling. But there would be no sterling notes and coins in circulation. Rather, a new currency with a different name (“drachma”, say) would be issued which could depreciate relative to sterling. The Bank of England would publish each day on its home page the rate of exchange between the pound and the drachma (rather as the Central Bank of Chile now publishes on its website the daily rate of exchange between the peso and the Unidad de Fomento, the indexed unit of account used for many transactions in Chile). With this arrangement, there would be absolutely no obstacle to negative sterling interest rates.

This “pound plus drachma” proposal would involve significant set-up costs (the replacement of sterling notes and coins with drachma notes and coins), and the Government might also want to consider issuing a basic electronic calculator to every household in the country (at the cost of roughly 0.01 per cent of annual GDP). The ongoing administrative costs, however, would be trivial in comparison with Gesell’s stamped currency.

Major set-up costs can be avoided, however, with my preferred scheme: “Scheme B”. It would not be essential for the Bank of England to issue drachmas or withdraw sterling notes and coin. Existing cash could continue to be used, but the value of a ten-pound note, say, would no longer necessarily be ten pounds. The Bank of England would publish daily the value of cash in terms of the sterling unit of account and would use this exchange rate in debiting the accounts of banks to which it supplied cash. When the government issued cash (for example, in payment of welfare benefits) or received cash (payments for licence fees, or taxes etc.) it would of course use the same exchange rate. Sterling cash would be legal tender to settle debts denominated in sterling at the daily exchange rate. That should be enough to

²⁷ Woodford (2003), pp. 63ff.

ensure that the unique daily exchange rate was used pervasively throughout the economy.

Eisler made a similar proposal more than seventy years ago²⁸: his idea was that the unit of account (expressed in pounds and decimals) — which he called “money *banco*” — would appreciate against cash (expressed in shillings and pence) — which he called “current money”. Restaurant menus, for example, would be priced in pounds and decimals: someone paying a restaurant bill by cheque (more likely, these days, by credit card) would simply write out the amount in pounds and decimals. In the case of cash payments, a cashier would consult the daily rate to work out the required cash sum in shillings and pence.

In the 1930s Eisler’s idea was regarded as unworkable; but with the current availability of cheap electronic calculators the need to use an exchange rate to convert prices expressed in the unit of account into cash prices would be a relatively minor inconvenience. It would no doubt lead to some reduction in the use of cash. But the main adverse impact of depreciating cash would be on tax-evaders and drug-dealers, and that hardly seems to be a conclusive reason for rejecting the idea.

The main question over my proposal, and one that Buitter has raised in a reply to my comment²⁹ on his paper with Panigirtzoglou, is not whether it would succeed in achieving negative interest rates, but whether that achievement would mean anything. Going back to my “sterling plus drachma” suggestion, the UK monetary authorities might succeed in stabilising prices measured in sterling terms, but would that mean anything if people were using drachma cash?

Buitter distinguishes between the “shoe-leather” and “menu” costs of inflation. “Shoe-leather” costs are of course the costs of economising on cash holdings — motivated by the lower return on cash than on alternative assets — and thus of having to make more trips to a bank (or ATM machine located in a supermarket or convenience store) to withdraw cash. This is a traditional name, incidentally, which would be better referred to these days as “petrol costs”. In the UK, with the closure of bank branches in rural areas, people in parts of the countryside have to drive perhaps 10 kilometres to their nearest bank or ATM, or a round trip of 20 kilometres. Thus the “petrol costs” of inflation are not as trivial as the traditional name “shoe-leather costs” makes them sound.

The “petrol costs” of inflation in my proposed “sterling plus drachma” economy are clearly related to inflation measured in drachmas not inflation measured in sterling. So by setting an inflation objective in terms of sterling prices the authorities are doing nothing to minimise these “petrol costs” of inflation. However, as these “petrol costs” are determined by the difference between the return on cash and the return on financial assets, the depreciation of cash does not cause these “petrol costs” to be high if that depreciation is used to imple-

28 Eisler (1931).

29 Davies (2005), Buitter (2004).

ment negative nominal interest rates. In any case, if the choice were between accepting high “petrol costs” or accepting a “lost decade” for the macro-economy, then presumably the former would be chosen as the lesser of two evils.

What then of the point about “menu costs”, broadly understood as not only the costs of changing price labels, but also the costs of renegotiating prices and wages? The point is that for these costs to be affected by the rate of sterling inflation, the prices have to be set in sterling in the first place. This choice of numeraire is obviously the outcome of private sector decisions.

In the familiar situation in which there is no separation between the unit of account used by the authorities and the cash in use for payments, in normal economic circumstances the unit of account used by the private sector is simply the same as the unit of account used by the authorities. This is not always the case, however, as in periods of high inflation and currency instability the private sector sometimes switches to a foreign currency. In periods of high inflation in recent decades, several Latin American countries saw the adoption of inflation-adjusted units of account. Even after years of low inflation, an inflation-adjusted unit of account has survived in Chile, where, as described by Shiller, “Chile’s *unidad de fomento* (UF, or unit of development) is the world’s first indexed unit of account. Created in 1967.. .. In Chile today, people buy and sell, and sign long-term contracts not only in terms of their currency (the peso) but also in terms of UFs....”

The UF is upheld by the government and by the legal system as a unit of account for transactions. If one defines a future payment in UFs, one must later, on the date the payment is made, calculate the payment in pesos by multiplying the UF amount by the number of pesos per UF shown in the newspaper on that day. People in Chile today will quote the prices of houses for sale and of apartments for rent in UFs, and they specify mortgage payments, tax payments, and even child support and alimony payments in UFs.”³⁰

Shiller, indeed, advocates the introduction of a similar inflation-adjusted unit of account in other countries. Niehans, in an interesting discussion of the choice of unit of account as the outcome of minimising accounting costs,³¹ had also considered the use of an inflation-adjusted unit of account, separate from the means of payment, as a possible measure to cope with inflation; though he saw the separation of means of payment and unit of account as a serious loss of efficiency: it was better to get rid of inflation and avoid the need for separation of the means of payment and unit of account. (Incidentally, Niehans prefers the term “medium of account” to “unit of account”. While his argument for this has some force, I follow the more general practice here.)

Greenfield and Yeager were also concerned about whether the money of account they proposed, as discussed above, would be adopted by the private sector. In their framework, the

30 Shiller (2003), pp. 204-205.

31 Niehans (1978) pp. 118-139.

government actually has a relatively limited role and this must limit its influence on the private sector's choice of unit of account. In addition there is, as I have already emphasised, not merely no cash but no base money at all in Greenfield and Yeager's proposed scheme. In my proposal base money does exist — denominated in the government's unit of account.

Both White³² and McCallum³³ argued in discussions of Greenfield and Yeager's proposal that the means of exchange would almost inevitably become the unit of account. McCallum makes no reference to — and was perhaps not fully aware of — the historical evidence of separation; while White makes only a brief and disparaging reference to “ghost currencies” — currencies which enjoyed an afterlife as a unit of account after the relevant coinage had ceased to be used as means of payment — whose significance, he claimed, had been overplayed. He likewise does not seem to be fully aware of the history of separation, which goes much wider than the case of “ghost currencies”.

Buiter argues that with my “sterling plus drachma” proposal I am in fact creating two possible focus points for the private sector's choice of numeraire: the public sector's unit of account and the unit of cash. As the private sector will be transacting in cash on a day-to-day basis, it may end up adopting the unit of cash as its numeraire, rather than the public sector's unit of account. Clearly there is room for argument here, and clearly Woodford is wrong, as Buiter notes, when he **defines** the unit of account as the unit which the government and monetary authorities use: “The special feature of central banks, then, is simply that they are entities whose liabilities happen to be used to define the unit of account in a wide range of contracts that other people exchange with one another.”³⁴ The unit of account is ultimately a private sector choice, not a central bank definition.

In normal economic circumstances, however, there is good reason to believe the private sector will simply follow the public sector. First, the government accounts for a large share of the economy: in the case of the UK, government spending and taxes account for roughly 40 per cent of GDP. By denominating payments to its suppliers and its own employees, as well as transfer payments and tax bills in terms of sterling, sterling would have a very privileged position compared with the drachma. This could be enhanced by the government requiring all shops to show prices in sterling rather than drachmas (just as they recently required shops to switch to metric measures of weight from the long-familiar imperial measures). Shops would in any case prefer to use sterling price tags, as sterling inflation would by assumption be lower than drachma inflation and therefore their “menu costs” would be lower if they used sterling rather than drachma price tags: they would have to change the price tags less frequently.

The authorities would also conduct all their transactions with commercial banks in

32 White (1984).

33 McCallum (1985).

34 Woodford (2003) p.37.

sterling. As the banks would have sterling-denominated borrowing from/deposits with the Bank of England and Debt Management Office, they would minimise market risk, other things being equal, by lending to their retail customers in sterling and taking retail deposits in sterling. They have no obvious reason to push their customers in the direction of drachma deposits, or borrowing in drachmas.

Sterling also has the competitive advantage that it is what people in the UK are used to. They have most of their debts and assets denominated in sterling already. Non-cash retail payments would presumably continue in sterling initially, as would wholesale payments: there is no reason why non-cash payments should be affected by the change from sterling to drachma cash. It thus is hard to see how the drachma would establish itself as the unit of account in preference to sterling. Why would people move away from the unit of account they are used to, and which they will continue to be required to use in all their dealings with government, to a unit of account that it is, after all, inferior (in the sense that it is subject to a higher, and possibly more variable, inflation rate)³⁵? This seems similar to abandoning a superior keyboard layout and adopting QWERTY.

Some anecdotal evidence of private sector inertia in use of unit of account (which, in the context of the argument above, would favour the continued use of sterling over the drachma as the private sector's unit of account) is available from 19th century France, following the Napoleonic reforms which introduced the franc both as cash and as the government's unit of account. Weber gives several examples of people using the ecu and/or pound as numeraire decades after their replacement by the franc; for example, a land-owner in northern France who "kept his accounts in ecus until 1837, in sols until 1849...and in livres and francs interchangeably, as if they were one and the same thing, all the way to 1877".³⁶

The concern over the private sector not adopting the government's unit of account would in any case be less in my (preferred) Scheme B. In this case there is no rival unit for cash (drachma) that could be adopted as unit of account.

National money of account with quasi-free banking

One possible arrangement would be to combine a national money of account with cash provided competitively by private banks. In principle, there is no reason why this could not take a form that might perhaps be described as "free banking", even though the outside money — the reserves of the banking system — would be different from those in past epi-

35 McCallum (1985, p.5) notes, in a discussion that is generally sceptical about the sustainability of a unit of account separate from the means of payment, that "there can be inducements — such as the presence of substantial inflation in terms of the MOE — for agents to make contracts in terms of something other than the MOE". (n.b. McCallum's MOE stands for "means of exchange" — what I am calling "means of payment".)

36 Weber (1976) p.33.

sodes of “free banking”. As Selgin and White put it:

“Most discussions of free banking assume that the monetary base does not include currency or deposits issued by an extant central bank. By itself, however, free banking does not uniquely specify the base money regime. Base money could be gold or silver.... Or it could be some fiat money.”³⁷

“Free banking” is of course normally taken to involve banking without a central bank. A nominal anchor is provided by the convertibility of bank notes into gold or silver. With fiat money, some authority like a central bank is needed to control the fiat currency to provide a nominal anchor: e.g. to set interest rates on it in accordance with an interest rate rule. But this authority would not supervise banks or involve itself in the banking system in any way. It might be help to differentiate this situation from classic free banking by calling it “quasi-free banking”.

If the main national banks all issued their own notes, they might compete by offering different rates of interest (positive or negative) on their notes; and the notes, though all denominated in the money of account, might trade at a premium or discount to that unit of account, and to each other. (The historical experience of free banking in the US confirms this possibility, though in Scotland bank notes normally traded at par in its free banking period.) Of course, the stability of free banking (both in theory and in the historical record) is open to argument. The stability of the Scottish system probably owed a great deal to the unlimited liability of equity holders, which should have helped restrain any temptation to “gamble for resurrection”. It is not clear that unlimited liability would be sustainable or credible today with the inevitable media focus on hard cases.³⁸

Local moneys as means of payment

The authorities would make it more difficult for an alternative unit of account to emerge if they did not issue national bank notes. One option would be to delegate the issuance of cash to local government bodies. Although central government might be reluctant to lose the seigniorage revenues from currency issuance (in the normal circumstance of positive interest rates on government bonds), this could be part of an overall reform of the structure of financial flows between central and local government and could be offset by a reduction in other central government transfers to local government.

³⁷ Selgin and White (1994).

³⁸ Shareholders of the City of Glasgow Bank which collapsed as a result of fraud in October 1878, included “bakers... builders, drapers, artists, portioners, grocers, clothiers, tobacconists...” according to Buchan (1997), p.212. All were exposed to unlimited liability. It is true that the UK government managed to resist pressure to bail out “names” in the Lloyd’s insurance market in the 1980s and 1990s, many of whom were impoverished or ruined by their unlimited liability for meeting claims, but they were mainly members of a much smaller social group who aroused limited public sympathy. By contrast, the Government has moved quickly to address a legal quirk that could have left Equitable Life policyholders exposed to unlimited liability.

How might this work? I should stress that I do not envisage local governments paying their employees and suppliers in local currency.³⁹ They would still generally pay them (by bank transfers etc.) in the national money of account; they would also raise local tax in the national money. The local cash would be supplied to fill a need mainly for small-scale local transactions, for which locally issued and locally familiar cash would be acceptable; indeed it might play a valuable role as a contribution to community solidarity. (Just as nation states generally see national currency as a way of building a sense of national identity.⁴⁰ One might also suggest that the EU has seen the abolition of national currencies as a way of reducing the sense of national identity within Europe.)

To obtain cash from their local authority, people would transfer national currency from a bank account into their local authority's bank account. In return the local authority would issue them with the local cash. The local cash could be denominated in a unit that might be fixed in terms of the national currency, though the local authorities might seek to increase their seigniorage revenue by making the local currency unit depreciate in terms of the national money of account. (And if there were negative interest rates on assets denominated in the national money of account, they would have to allow the local money to depreciate if they were not to make losses on the issuance of local money.)

An alternative would be to have private banks, rather than local authorities, issue cash. However, there is some advantage in local authorities rather than commercial banks doing this as their power to raise local taxation would help to back the currency. There should be less risk of a flight from cash occasioned by worries about the viability of the issuer.

There should be no need for local legal tender laws, requiring local businesses to accept the local cash.⁴¹ Local businesses are not likely to refuse the cash their customers have in their purses and wallets, provided they are assured of being able to convert it back into national money by depositing it at their bank. Local currency might be used outside the area of the authority that issued it, but one would imagine that it would be increasingly unfamiliar and less readily accepted at greater distance from the issuing area. (This was the English experience in the eighteenth and early nineteenth century, when the note issue was almost entirely supplied by local banks. These banks' notes were not generally accepted outside the area of issuance.) There could, however, be some competition between neighbouring local authorities over the rate of depreciation of their cash, if more than one local authorities' cash

39 It is true that in Argentina, for example, in a situation of financial and fiscal crisis, local governments paid their employees in local money; but, except in extreme circumstances, one cannot seriously expect that employees will accept payment in a currency that cannot be used outside a limited local area.

40 As emphasised by Helleiner (2003).

41 Note that within the UK, while Bank of England notes are legal tender in England and Wales, there are no legal tender bank notes in Scotland; but that forms no impediment at all to the acceptability of either Bank of England bank notes or notes issued by Scottish banks in discharge of debts in Scotland. Scottish bank notes can also normally be used without any problem in England, despite their lack of familiarity.

was accepted in border areas. A local authority would then face a revenue maximisation problem: a faster rate of depreciation would increase its seigniorage revenue for given cash issuance but reduce the demand for its cash.

I have suggested that one reason for issuing local cash rather than a single national cash would be to reduce the competition with the government's unit of account. A decision not to have a single national cash is, to be realistic, hardly likely to be taken for this reason alone. It must also be understood as an attempt to throw grit into the wheel of national commerce and mobility. It would create some inconvenience for those travelling from one part of the country to another that the cash they used in their own locality would be unfamiliar and might not be an acceptable means of payment in another locality. (Of course, credit card and bank transfer payments in the national currency would be available and there would presumably be a drastic reduction in cash payments for restaurant and hotel bills, for example.)

The case for such interference with market mechanisms could be sociological: that competition between successful growing regions and more backward regions was disrupting local communities in the more backward regions. Or it could be based on concerns about income distribution: that those with property rights in the more backward regions were losing money relative to those with property rights in the more successful regions, and there was no non-distortionary way to implement transfers to offset these losses. Or it could be based on externalities: that the population movements prompted by regional imbalances are inefficient because those moving are not affected by the additional costs imposed on the public sector by their moves (e.g. the need to build new schools and hospitals in the growing area).

One could make a (somewhat far-fetched) analogy with the *hansatsu*, the paper money (denominated normally in silver or gold, sometimes in rice) that many feudal domains (*hans*) issued in the Edo period in return for deposits of specie, which in theory backed the paper currency issue, and in payment for local services. Inhabitants of domains that issued such paper money were normally required to use the domain paper money for cash payments within the domain area, with spies sometimes employed to check that specie was not being used. However, I cannot see any analogous compulsion being acceptable in modern Japan: people would have to be free to use the electronic form of national money if they wanted to. And in any case the vast increase in mobility and the centralisation of the economy (not to mention globalisation) makes any analogy with the monetary arrangements of the Edo period (when there were strict border controls between the different "*kuni*") of limited relevance today.

Some Japanese advocates of local currencies may find my suggestion for the nature of local government issuance too limited: they would want to see local currencies playing a bigger role in local government finance. As I have argued elsewhere, however, it is not realistic to expect local government employees (other than those doing voluntary work) and suppliers to agree to be paid in a local currency that is not convertible into the national

currency, except in financial crises where the employees face the alternative of accepting a temporary local money or nothing (and where no alternative employment is available)⁴².

The arrangement I suggest seems to me the most that can be realistically hoped for in terms of the contribution of local money to local finance. It could of course be additional to local moneys used to remunerate volunteers (the role of which is more to provide a significant way of recording and expressing appreciation for the volunteers' activities than to provide a source of income); and also additional to private arrangements such as LETS (Local Exchange Trading Schemes) which finance transactions between scheme participants. The latter may perhaps be best thought of as attempts to alleviate the "institutional crowding out"⁴³ of trust-based transactions by the monetary economy; they have no direct relevance to local government finance.⁴⁴

What about "e-money"?

It is cash in material form that creates the problem of the zero interest rate bound. My assumption is that material cash will continue to have attractions for use in small-scale transactions, that its replacement entirely by new de-materialised forms of money (such as e-purses) would increase transactions costs, as demonstrated by Goodhart and Krueger.

In discussing e-money, one needs to make a distinction between "inside e-money" (what most people think of as e-money, and the only sort that currently exists anywhere, as far as I am aware), and "outside e-money", created by monetary authorities. At present, the authorities in Singapore are planning to replace cash by "digital legal tender",⁴⁵ but the details of the scheme have not been published. Conceptually, it may bear some relationship to Tobin's scheme for "deposited currency", under which cash would be abolished and individuals would hold deposit accounts in local branches of the central bank, perhaps located in post offices.⁴⁶ As Tobin notes, interest could be paid on such deposits (obviously he would have been thinking in terms of a positive interest rate, but if there were no currency but "deposited currency" then a negative interest rate would be equally feasible). In discussing their proposal, the Singapore authorities have stressed, as one of its advantages, the facility for making interest payments on digital legal tender. They clearly have in mind positive interest payments, but negative interest payments would be equally unproblematic.

The main relevance of "inside e-money" is that it could make it somewhat easier to fill

42 Davies (2003).

43 I have taken this term from Bowles (2004), p. 495. For a recent formal theoretical discussion of the replacement of norms of reciprocity by money see Araujo (2004).

44 I discussed LETS in detail in Davies (2004).

45 Kok (2002).

46 Tobin (1985).

the gap left in payments mechanisms if the monetary authorities decided to solve the problem of the zero interest rate bound by simply not issuing cash. If the authorities implemented negative interest rates, issuers of inside e-money — simply to avoid bankruptcy — would have to find some way of making e-money depreciate: e-purses would have to lose value if unspent. (Alternatively, they could have an expiry date. A bus or telephone card, with which bus fares and telephone calls are paid for in advance, might be valid for only three months after purchase, say. Some forms of pre-paid cards already embody such expiry dates, for other commercial reasons.) Pre-paid airtime on mobile phones (which will increasingly be usable like an e-purse to make a range of purchases) would also have to depreciate if not subject to an expiry date. Payment of items through surcharges to a post-paid mobile phone bill — another area of e-money with substantial growth potential — involves only unit of account money, not cash.⁴⁷

Concluding remarks

The method of achieving negative interest rates I have described in this paper would certainly work, but there are two possible lines of objection to it. The first is that it would be confusing for the citizens of a country operating it, and would increase the costs (broadly defined) of cash transactions. The second is that while one might achieve negative interest rates in terms of what one might term a “government-sponsored unit of account”, that would not have any significance if the private sector participants in the economy adopted an alternative unit of account.

In reply to the first objection, I would point out that what one has to compare are the costs of my proposal and the costs of prolonged macro-economic paralysis, if that is the alternative (as it has been in Japan). I would also point out that historically people have coped with much more confusing monetary regimes (with numerous coins from different countries with different weights and fineness, and different degrees of wear and clipping) all without the benefit of electronic calculators and without the option of simple electronic transfers of money. In reply to the second I have pointed out the pervasive influence of government transactions, and argued that the private sector has no obvious motive to abandon the government-sponsored unit of account and little reason to adopt a (higher inflation) alternative.

My proposal also ties in with the advocacy of local currencies, issued as means of payment, as these would be even less likely than a nationwide means of payment to threaten to supplant the role of a government sponsored national money of account.

⁴⁷ A concise summary of the use of the potential use of mobile phones for payments can be found in Allen (2003), p.430.

References

- Helen Allen "Innovations in retail payments: e-payments" Bank of England Quarterly Bulletin Winter 2003, pp. 428-438.
- Luis Araujo "Social Norms and Money" Journal of Monetary Economics 51 (2004) pp. 241-256.
- Fischer Black "Banking and interest rates in a world without money: The effects of uncontrolled banking" Journal of Banking Research (1970) 1, pp. 8-20.
- Samuel Bowles "Microeconomics" Russell Sage Foundation (2004).
- James Buchan "Frozen Desire" Picador (1997).
- Willem H. Buiter "Negative Interest Rates by Unbundling Numeraire and Medium of Exchange: Is Numeraireology the Future of Monetary Economics?" Mimeo, European Bank for Reconstruction and Development, October 2004.
- Willem H. Buiter and N. Panigirtzoglou "Overcoming the Zero Bound on Nominal Interest Rates with Negative Interest on Currency: Gesell's Solution" The Economic Journal 113 (October 2003), pp. 723-746.
- Tyler Cowen and Randall Kroszner "The Development of the New Monetary Economics" Journal of Political Economy 1987 vol. 95, no.3 pp. 567-590.
- Stephen J. Davies "Japan's Deflation and the Feasibility of Negative Interest Rates" Kobe Economic & Business Review 47 (2002), pp. 93-115.
- Stephen J. Davies "Local Currencies — Lessons from the World" in Yasuo Konishi (ed) "Local Currencies — Sono Genjou to Kadai" Kobe University 2003, pp. 1-38.
- Stephen J. Davies "Lessons from LETS" Kobe University Kokumin Keizai Zasshi vol. 189 No. 2 (February 2004) pp. 55-70.
- Stephen J. Davies Comment On "Overcoming The Zero Bound On Nominal Interest Rates With Negative Interest On Currency: Gesell's Solution" Submitted to The Economic Journal (2005).
- Robert Eisler "This money maze: a way out of the economic world crisis" London: Search Pub. Co. 1931.
- Luigi Einaudi "The Theory of Imaginary Money from Charlemagne to the French Revolution," chapter 14 of Frederic C. Lane and Jelle C. Riemersma (eds) "Enterprise and Secular Change" George Allen and Unwin 1953. (Translation of article which appeared originally in Rivista di storia economica 1936, I, 1-35).
- Eugene F. Fama "Banking in the Theory of Finance" Journal of Monetary Economics 6 (1980) pp. 39-57.
- Benjamin J. Friedman "The Future of Monetary Policy: The Central Bank as an Army with only a Signal Corps?" International Finance 2 (1999), pp. 321-338.
- Mitsuhiro Fukao "Financial strains and the zero lower bound: the Japanese experience" BIS Working Papers no. 141 (September 2003).
- M. Goodfriend "Overcoming the Zero Bound on Interest Rate Policy" Journal of Money Credit and Banking (November 2000) part 2, pp. 1007-1035.
- Charles A. E. Goodhart "Can We Improve the Structure of Financial Systems?" European Economic Review 37 (1993), pp. 269-291.
- Charles A.E. Goodhart and M. Krueger "The Impact of Technology on Cash Usage" LSE Financial Markets Group Discussion Paper 374 (April 2001).
- Kathryn Grant, Gertjan Vlieghe and Andrew Brigden "Assessing the stability of narrow money demand in the United Kingdom" Bank of England Quarterly Bulletin Summer 2004, pp. 131-141.

- Robert L. Greenfield and Leland B. Yeager “*A Laissez-Faire Approach to Monetary Stability*” *Journal of Money Credit and Banking* vol. 15 no. 3 (August 1983) pp. 302-315.
- Robert E. Hall “*Controlling the Price Level*” National Bureau of Economic Research Working Paper 6914 (January 1999).
- Ian R. Harper and Andrew Coleman “*New Monetary Economics*” in Peter Newman, Murray Milgate and John Eatwell (eds.) “*The New Palgrave dictionary of money and finance*” Macmillan 1993, vol. 3, pp. 28-31.
- Eric Helleiner “*The Making of National Money: Territorial Currencies in Historical Perspective*” Cornell University Press (2003).
- David Humphrey, Aris Kaloudis and Grete Owre “*The future of cash: falling legal use and implications for government policy*” *Journal of International Financial Markets, Financial Institutions and Money* 14 (2004) pp. 221-233.
- Joel Kaye “*Economy and Nature in the Fourteenth Century*” Cambridge University Press 1998.
- Narayana R. Kocherlakota “*The Technological Role of Fiat Money*” *Federal Reserve Bank of Minneapolis Quarterly Review* vol. 22 no. 3 (Summer 1998) pp. 2-10.
- Low Siang Kok “*Singapore Electronic Legal Tender (SELT) — A Proposed Concept,*” chapter 6, pp. 147-156 of “*The Future of Money*” OECD 2002.
- Bennett T. McCallum “*Bank Deregulation, Accounting Systems of Exchange, and the Unit of Account: A Critical Review*” NBER Working Paper no. 1572 (March 1985).
- Bennett T. McCallum “*Monetary Policy Rules and the Japanese Deflation*” Paper prepared for the March 20, 2002, workshop sponsored by the Economic and Social Research Institute of the Japanese Government.
- Ryuzo Miyao “*Liquidity Traps and the Stability of Money Demand: Is Japan Really Trapped at the Zero Bound?*” Kobe University RIEB Working Paper, October 2002.
- Jurg Niehans “*The Theory of Money*” The John Hopkins University Press 1978.
- Kunio Okina and Shigenori Shiratsuka “*Asset Price Bubbles, Price Stability, and Monetary Policy: Japan's Experience*” Bank of Japan Institute for Monetary and Economic Studies Discussion Paper 2001-E-16.
- Thomas J. Sargent and Francois R. Velde “*The Big Problem of Small Change*” Princeton University Press 2002.
- George A. Selgin and Lawrence H. White “*How Would the Invisible Hand Handle Money?*” *Journal of Economic Literature* vol. 32 (December 1994) pp. 1718-1749.
- Robert J. Shiller “*The New Financial Order*” Princeton University Press 2003.
- Peter Spufford “*Handbook of Medieval Exchange*” London: Royal Historical Society 1986.
- James Tobin “*Financial Innovation and Deregulation in Perspective*” *Bank of Japan Monetary and Economic Studies* vol. 3 no. 2 (September 1985) pp. 19-29.
- Eugen Weber “*Peasants into Frenchmen: the modernization of rural France, 1879-1914*” Stanford University Press 1976.
- Lawrence H. White “*Competitive Payments Systems and the Unit of Account*” *American Economic Review* vol. 74 no. 4 (September 1984) pp. 699-712.
- Knut Wicksell “*The Influence of the Rate of Interest on Prices*” *The Economic Journal* 17 (June 1907), pp. 213-220.
- Michael Woodford “*Interest and Prices*” Princeton University Press 2003.

**PRODUCT COMPETITIVE ADVANTAGE AND PRODUCT ARCHITECTURE
– VALUE CREATION AND VALUE CAPTURE
IN THE DIGITAL CAMERA INDUSTRY –**

MUNEHIKO ITOH Kobe University

Abstract

With a focus on the digital camera industry, which of all IT devices Japan has high industry competitive advantage, we discuss what product strategy did companies employ to create and capture value as the market developed, and how product architecture has changed. To meet the different needs of customers, the digital camera industry is made up of two different vectors of successive and destructive technology which each take either a modular or integrated approach to product architecture. It is an industry where the opposing industrial constructs of vertical integration and horizontal specialization are intermixed. Furthermore, the de-integration that has occurred for the supply of for core modules such as CCDs and the industry reliance on a few specialized companies shows how competitors are dividing into three camps adopting one of three strategies: vertical integration, horizontal specialization or core module supply.

JEL Classification: M11

Keywords: Product development, Product architecture, Horizontal architecture, Vertical integration, Product strategy

1. Introduction

Since the 1990s, the IT equipment industry has undergone rapid progress. Especially products such as mobile telephones, digital still cameras (from this point referred to as digital cameras), large scale displays, DVDs, personal computers, and car navigation systems have all seen growth on a global scale. There seems to be a common pattern in this IT device growth.

Corresponding Address:

Munehiko Itoh

Research Institute for Economics and Business Administration (RIEB), Kobe University

2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501 Japan

E-mail: itohm@rieb.kobe-u.ac.jp

Phone / Fax: 81-78-803-7009

Regardless of the perspective in regarding the similarity of device composition or focus on the relation between information technology and shared infrastructure in major social change, it is difficult to deny that development in the IT equipment marketplace is attributable to friendly rivalry between corporations in the marketplace, through cooperation and competition.

In this report, with a focus on the digital camera industry, which of all these IT equipment categories Japan has a relative competitive advantage, we analyze through this development what product strategy did corporations employ to create or capture value, and also what set apart companies that had accomplished successful product strategies from those that did not. Then, we will consider the origin of Japan's industry competitive advantage that is the driver behind this industry.

1.1 Preceding Research on Modularization

In new growth areas, what kind of product strategy is needed to quickly launch products and ensure success? When launching new products, to what extent should in-house technical innovations be used, and when should consideration be given to standardization? Before debating these topics, we shall review the Modularization that is this paper's key concepts.

Modules have an extremely long history to the extent that descriptions of modules are noted in which wheels, hubs, bearings, and a gasoline tank are standardized by the mass-production method developed by Henry Ford (Swan, 1914). A half-century later, the concept of "Modularity" was widely recognized through the discussion of product efficiency brought by modularization in the watchmaker analogy (Simon, 1962) and mass-customization advocated by Starr (Starr, 1965). In the 1990s, detailed research on modularization by Ulrich and others started, and the concept of an interface between modules was advocated (Ulrich and Tung, 1991). Numerous results were reported such as detailed research on shortening of product development lead time (Thomke and Reinertsen, 1998), on managing both mass customization and low cost (Pine, 1993; O'Grady, 1999), and on influencing product innovation (Baldwin and Clark, 2000). However, problems with modularization research have become obvious. First is that the interpretations, definitions, and viewpoints regarding modules are not uniform among the researchers. For example, researchers turning their viewpoints to product development management regarded modules as common, standardized, and compatible components and have pointed out the misdistribution of innovative activities among many organizations (Galsworth, 1994; Sanchez and Mahoney, 1996; Schilling, 2000; Baldwin and Clark, 2000). With technology theory, however, a module is understood as a design manual for product development and its product development management is treated as a given condition. For example, the module is defined as the kind of product feature, product demand, or material (Smith and Eppinger, 1997; Newcomb, Bras and Rosen, 1998; Stone and Wood, 2000). Thus, there have been a large number of attempts to define a module from a common

viewpoint, irrespective of management theory or technology theory (Brusoni and Prencipe, 1999; Baldwin and Clark, 2000; Fixson, 2003; Schilling, 2002; Sturgeon, 2002; Sako, 2003), but the current situation is one where most have not specified common definitions. Two factors for the existence of different viewpoints of management and technology mentioned above have been noted.

First is the industrial peculiarity in a module argument. The fact that the definition of a module differs substantially between products requiring assembly and small electric designed with a high density is unavoidable in the combination of hardware and software, or hardware itself. For example, numerous pieces of industry-specific research have been reported such as modularization of an automobile (Sako and Murray, 1999), modularization of an elevator as heavy machinery (Mikkola, 2001), a personal computer (the following, personal computer) (Fine, 1998; Baldwin and Clark, 2000), and an electronics product (Sturgeon, 2003), although definitions regarding modules are not uniform.

Another problem with the modular definition is the existence of several viewpoints for specification of a module (Fixson, 2003). The definition of a module can be broadly analyzed from three viewpoints of system, class, and product life cycle. A system viewpoint is for the purpose of analyzing the product architecture and understanding the connection between modules and their interfaces as a product structure element (Ulrich, 1995; Garud and Kumaraswamy, 1995; Schilling, 2000; Baldwin and Clark, 2000). Product architecture is a concept showing the relation between product feature and product structure, namely, the constituent elements of a product, (Ulrich, 1995; Baldwin and Clark, 2000; Fujimoto, Takeshi and Chingtao, 2001; Aoki and Ando, 2000). The class viewpoint indicates that if modularization is further advanced higher-class modules are formed commercially, which is expressed by the term "platform" in many papers. That is, skillful product development is defined as forming a platform by combining limited parts and united modules so as to form a product family that flexibly and quickly suits customer needs (Hyer and Wemmerlov, 1984; Nobeoka and Cusmano 1994; Meyer and Lehnerd, 1997; Robertson and Ulrich, 1998; Gonzalez-Zugasti, Javier, Otto and Baker, 2000; Gawer and Cusumano, 2001). Finally, the viewpoint of product life cycle is defined differently as three stages, development (modularity in design [MiD]), use (modularity in use [MiU]), and production (modularity in production [MiP]) (Sako and Murray, 1999; Baldwin and Clark, 2000). The current product structure is too complicated to be entirely designed by one person, and modularization can be promoted in relation to organization structure (Brusoni and Prencipe, 1999). In other words, a product development organizational structure is a view related to the modular design of a product (Henderson and Clark, 1990; von Hippel, 1990). However, important factors regarding modularization as seen from the point of view of the user are additional features, upgradability, and the diversity of options at the time of purchase (Pine, 1993; Sako and Murray, 1999; Yu, Javier, Gonzalez-Zugasti and Otto, 1999). Finally, the discussion of modularization includes the most practical

problem of formation of a supply chain with regard to product manufacturing (Sturgeon, 2002), and selection of commercially procurable modules becomes a requirement when procuring modules or utilizing outsourcing (Whilhelm, 1997).

In the above portion, the genealogy of research in modularization has been described. In this paper, the digital camera industry is examined by using the concept of "Modularization" to clarify how a company establishes modules and implements product differentiation in a limited range of selection to heighten competitiveness.

2. Framework for Analysis

We set 564 digital camera models from about 30 companies sold between 1995 and 2003 as subject and range for analysis. The digital camera industry was selected for two reasons. First, Japan is a global supplier and a strong international competitor, making this an optimal case study on competitive advantage in product development for the purpose of analyzing product strategy of market competitors. Second, the digital camera industry is relatively young, making it possible to acquire data for total and objective analysis of actual selling prices and specifications of all commercialized products. For these reasons we did not include other products with digital camera functions such as mobile phones or personal computers. Furthermore, for enhancements in the parts market and component technology such as operating system diversification,¹ improved microcomputer performance, increased memory and miniaturization of display components such as LCDs, digital cameras share many common traits with IT devices outside of the digital camera market. Therefore, it was thought that analysis of the digital camera industry had a high probability of giving clues for research on the IT device industry overall. In this section, we provide an overview of the digital camera industry and discuss indicators used to analyze the product strategy and competitive advantage of market competitors.

2.1 Overview of the Digital Still Camera Industry

In October 1988, Fuji Photo Film Co. Ltd. announced the prototype model of the first digital camera, called the DSP-1, and commercialized it three years later as the DSP-100. Its price at the time was a very costly 680,000 yen, and its sales volume was limited. At nearly the same time, Kodak began selling two models, the DC3 / 32 and DCS 200ci, which had internal hard disks and were developed together with Nikon. Costing over 3 million yen

¹ The overall computer system operating software which provides the basic features shared by most application software. Also called "base software." By using the core features provided by the OS, software developers can reduce development effort and standardize the application operation. There are operating systems such as those used in personal computers as well as the embedded operating systems, such as those found in IT devices.

apiece, they were adopted for use in specialized applications. Kodak originally saw digital cameras as an alternative to instant cameras for test shots in studio photography.² Then, in October 1993, the VC-1,000 from Olympus appeared on the market. These four models from these three companies were the pioneer models, and accordingly they functioned mainly as prototypes, and because they were built from existing technology such as hard disks and the bodies of optical cameras or video devices, they were very unrefined products. At the time, the immaturity of the market for personal computers, which were needed to store, display and print digital camera images, would be an obstacle to the spread of digital cameras. Also, since the optical camera industry was in competition with digital cameras, the product development of cameras using a cartridge film called APS continued, and the future for digital cameras remained unclear.³

The digital camera market took off in March 1995 when Casio released the affordable model QV-10 which had 250,000 pixels. The number of digital camera models has been steadily growing ever since. All of the major manufacturers had entered the market in the three years between 1995 and 1997. There have been no new domestic entrants to the market since 1997. Up to 2003, 564 different products have been commercialized and the rate of new product development has not lessened. The number of competitors remains unchanged while the number of new products continues to grow, raising the ferocity of product development competition.

2.2 Analytical Perspective

Here, we define the framework through which we will analyze the product changes in the overall digital camera industry. Here, the problem is how to gather design-level component, software framework, and OS information gathered through on-site surveys and apply this data to all 564 products targeted for product architecture analysis. Furthermore, in reality it is difficult to delve into the product cost structure data to determine the market value. Instead, in this research we gather product specifications and apply indicators where they can show changes in competitive advantage. Furthermore, we attempt to analyze competitor strategy by

2 This information gathered through interviews with members of the Kodak R&D department who were involved with digital camera research in the 1980s.

3 APS: Advanced Photo System is a unified standard developed by the 5 corporations: Kodak, FujiFilm, Canon, Minolta and Nikon; it is also known as new system camera. Differing from conventional cameras, the film was compact and employed a cartridge based loading system. The format's film was different from conventional films and could store data magnetically. In addition to date features, it could save data regarding print messages or specify the print type (C, P and H) for each frame. Additionally, the shooting conditions could be saved to each frame, allowing appropriate corrections to be made at the time of printing. Furthermore, it was also technically possible to record audio, but digital cameras appeared on the market before it was commercialized, and APS never was able to garner a large market share.

looking at the details of the actual selling price in the year of introduction and showing the relative value of all products for the industry overall. Since product value (factors related to cost) were greatly affected by corporate fixed costs, logistics routing, and corporate strategy, earlier researchers did not attempt to analyze them because the variable cost of the product itself could not be identified. However, recent advances in information technology allow easier access to pricing data. For this reason, we intentionally relied on this data for value analysis. Therefore, our analysis in this report is based on data acquired from product specifications, actual market prices, and industry association manufacturing statistics.⁴

First, the information on features that can be acquired from digital camera specification sheets is extremely limited. Table 1 shows the shift in some main digital camera features. Pixels,⁵ CCD sensor size,⁶ weight and price are continuous values and will be considered separately. Discussed here are items which can be judged in binary form. First, conversion method⁷ refers to the processing used on captured image data. The shift from YeMaCy to RGB conversion is prominent -the YeMaCy method that was most common when digital cameras first were marketed has slackened off from 1999, and as of 2003 all models use RGB conversion. Cameras with replaceable lenses are primary single lens reflex (SLR) digital cameras. Initially, product developers encountered technical problems with large format lenses and small CCDs which resulted in distortion around the edges of the image. However, from about 2001 technology for primarily optical device manufacturers to design large format

4 For recent models, Product specifications were gathered from various media such as catalogs, trade magazines, and manufacturer home pages. Actual selling price data was collected from the time of beginning to create this database in 1999 by frequently researching prices displayed at discount stores and online retailers and rounding to 1,000 yen. No adjustments to pricing has been made for inflation. The actual price fluctuates most as a model gets older. Whenever possible, but we referred to selling prices of new models in March and December when sales were most brisk. Data for product volumes and market share for each manufacturer was collected with the assistance of the Camera and Imaging Products Association (CIPA).

5 Pixel count for digital cameras refers to the number of pixels used in the CCD sensor. The number of sensors converting light into digital signals is an important element in measuring the performance of a digital camera. More pixels to collect and convert light means more image information can be recorded. When digital cameras first appeared on the market, most models had about 300,000 pixels. Recent "megapixel" models with CCDs exceeding 3 million pixels are now standard.

6 CCD: Charge Coupled Device — a device which electronically encodes images by converting light energy into electronic signals. A core component of digital cameras

7 The CCDs in digital cameras typically only record black and white images. For this reason, they are used together with filters to recreate color. For an RGB filter, a filter is applied to each sensor in the CCD, and by passing R (red), G (green) and B (blue) light through each of the sensors, the color data of the image is recreated. YeMaCy filters are filters which filter light complementary to RGB. Since using YeMaCy filters increases the ability of light to pass through, they have a benefit of having higher sensitivity, but they also have a drawback that complicated data processing is required to recreate the color data. On the other hand, RGB filters have the benefit of crisp representation of the original colors without complicated data processing, but they are less sensitive. Recently growing is the trend to combine large diameter lenses with RGB filters.

optical sensors became established, and from 2003 their market share has increased dramatically. External storage media specifications make up for nearly all storage media. Large pixel counts result in images which requiring large amounts of storage space; internal storage media could not keep up with the demand. Finders⁸ have become diversified. Here, this refers to optical viewfinders, but there are also many non-optical finders such as LCD viewfinders. As shown in the table, almost every model has an LCD monitor as a standard feature. For power, there are cameras which use conventional dry cell batteries or conventional rechargeable batteries, while there are also cameras which use completely proprietary battery designs. For instance, compact, thin digital cameras are unable to use conventional dry cells because they of their size, leading to the coexistence of standard type cameras and thin cameras following both specifications.

Finally, since 2000, USB has become the standard type of interface. As seen above, we found that specifications such as color conversion methods, external media, LCD monitor, and USB interface have become neutral standard features, while pixel quantity, lens interchangeability, viewfinder, power and weight have a tendency to be more differentiated. To more easily visualize this differentiation tendency, we attempted to categorize digital cameras according to characteristics such as lens interchangeability, CCD size, model name, weight, and maximum pixels. Resultingly, we split the cameras into four categories: SLR, compact, standard and

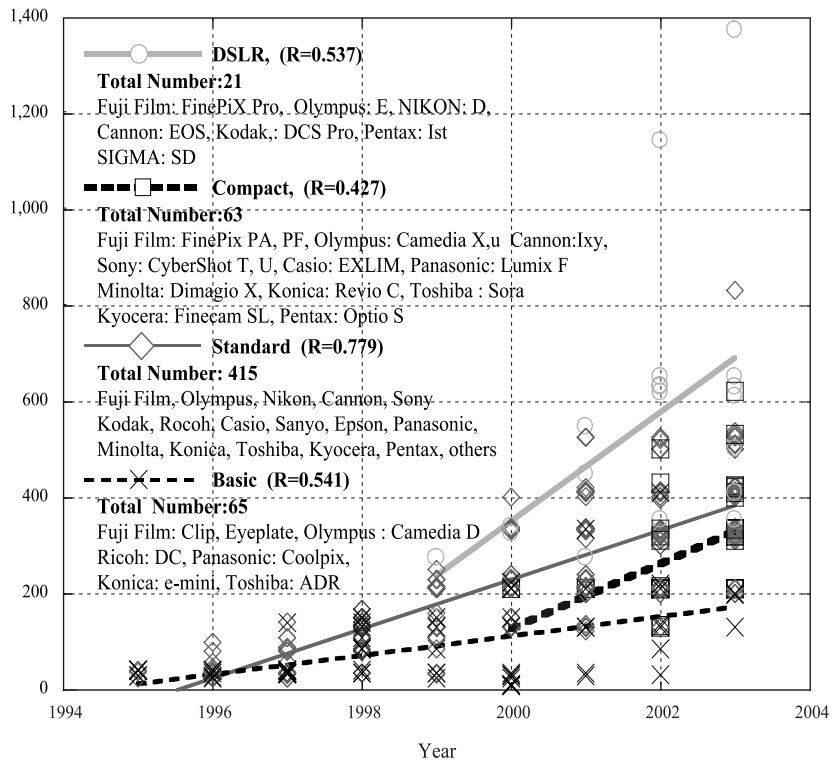
Table 1 : Shift in Digital Camera Core Specifications

Year	Conversion Method		Replaceable Lens		Storage Media		Finder		LCD Monitor		Power		Interface	
	Comp	Prim.	No	Yes	Internal	External	Other	Optical	None	Yes	Exclusive	AAA	Others	USB
1995	6	1	7	0	4	3	2	5	5	2	3	4	7	0
1996	21	7	27	1	14	14	9	19	10	18	6	22	28	0
1997	32	14	46	0	15	31	19	27	6	40	9	37	45	1
1998	29	18	47	0	2	45	12	35	2	45	13	34	47	0
1999	22	35	55	2	2	55	8	49	2	55	18	39	46	11
2000	28	51	77	2	9	70	17	62	10	69	29	50	16	63
2001	25	66	88	3	6	85	15	76	9	82	37	54	6	85
2002	14	84	92	6	5	93	21	77	1	97	55	43	2	96
2003	0	109	96	15	0	111	23	88	1	110	60	51	2	109
Total	177	385	535	29	57	507	126	438	46	518	230	334	199	365
Percentage (%)	31.5	68.5	94.9	5.1	10.1	89.9	22.3	77.7	8.2	518	40.8	59.2	35.3	64.7

⁸ The optical finder is the window which you look through in a camera to view the subject being photographed. In film cameras these are simply called “finders,” but since an LCD monitor can be used as the finder in digital cameras, the term optical has been added to differentiate the two. Here, filters that work together with the photographed image, in other words finders where the image is seen through a lens, are called optical finders. An example is a finder that zooms the image when the camera operator uses the zoom feature on the camera. There are also digital finders, for these, an LCD is attached to the finder and the actual image is displayed.

basic⁹ according to lens interchangeability and weight. The analysis below is based on this grouping. Figure. 1 shows the change in number of pixels for the four digital camera types broken down by year and having been undergone primary regression. Each company's model name is shown for each product category. The SLR type is easily recognized. SLRs can be recognized by their large lens, large format optical sensor with a high pixel count, and their lens replaceability. Products in the compact category weigh less than 200 grams, are thin and use proprietary 1batteries, and are easily recognized because each company has a model name especially for cameras belonging to this category. Cameras in the standard category also have easily recognized model names, recognition of standard cameras as being different compact type cameras by their weight and their conventional batteries (discussed below). The basic category is an affordable series with less than 2 megapixels. They have separate model designations to differentiate them from standard models. To further validate our method of classification, in Figure. 2 we created a boxplot based on the weight of products belonging to

Fig. 1 Trend in Pixels of Digital Camera Category



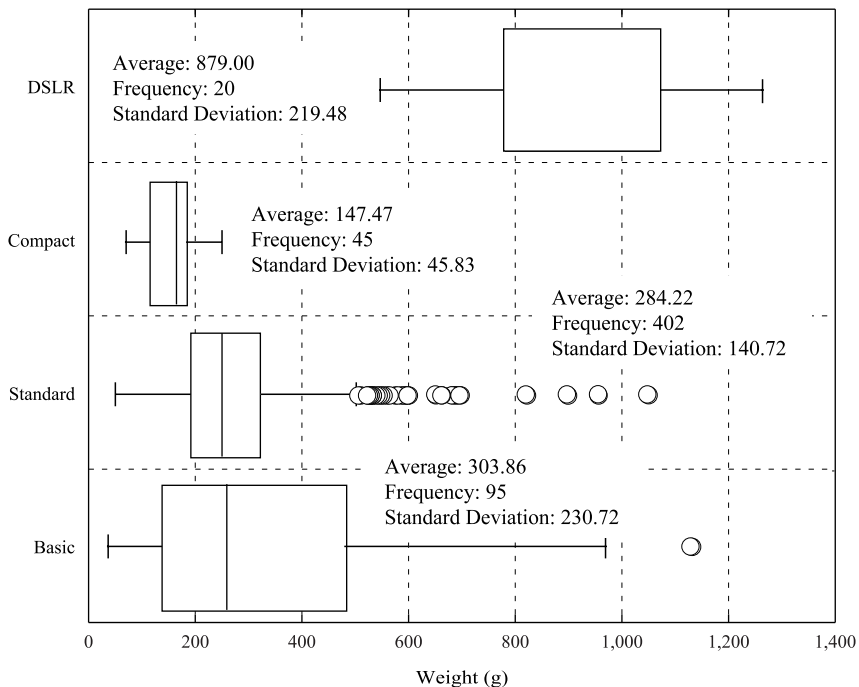
9 For weight and pixel count, when we performed a 2x2 analysis on each type, the asymptotic significance was significant at 0.5%. Therefore, this grouping can be considered to be statistically valid.

the various categories. As is clearly shown in the figure, for standard products excluding basic types, SLR and compact types the weight distribution is totally discriminate. Based on these results, we judged our grouping to also be statistically valid. This classification also implies the estimation of product makeup and the product strategy chosen by each manufacturer. Below, we proceed to analyze the data based on this grouping.

3. Analysis of Product Competitive Advantage

While product specifications for digital cameras are growing more generic, our analysis showed that competition was occurring for weight reduction and number of pixels. In this section we will consider how competitors implemented product strategy in this competitive environment and how did they form their competitive advantage. By performing a detailed analysis on how competitors performed product development, for each of the four previously defined product categories, the results will give clues as to how they dealt with successive improvement of existing technology and destructive technology for new areas, and how they changed the accompanying product architecture.

Fig. 2 Weight Distribution according to Digital Camera Type



3.1 Analysis of Competitor Strategy

What factor determines a digital camera's product competitive advantage? To answer this question, in the previous section we investigated the changes in specification. We found that while most features tended toward homogeneity, pixel quantity and weight underwent continuous and successive improvement. In this section, we will consider how each company formulated their product development strategy in this industry framework.

First, Table. 2 shows corporate product sales broken down by year and product name or series model for each category. First, the number of new products has grown every year. 111 products were released in 2003. From this table we can see that competition is especially strong in the digital camera industry. First, in the 2 years from 1995 to 1997, all the major players had achieved product commercialization. Accordingly, the unclear separation between old competitors and new entrants to the market is characteristic. One possible driver here is that digital cameras are products which are technically influenced by silver salt cameras¹⁰ and video cameras, and manufacturers of both of these have entered into the digital camera market. Since silver salt camera manufacturers and electronics manufacturers both entered the market, the result was sudden intense competition. This shift in the competitive landscape occurred twice. The first shift was in 1999 when digital camera production exceeded production of APS cameras, which had appeared on the market at roughly the same time. The second shift was in 2001 when digital camera production exceeded production of all silver salt cameras. Sony, Olympus, Fuji Film and Canon seized on these turning points. Sony and Olympus provided new value by drastically expanding their product lineup from 1999, and Canon did the same by creating the IXY series of compact cameras and the EOS line of SLR cameras. The actions of these firms which seized on these turning points enabled them to increase their sales volume.

Next, based on Table. 2 we attempt to divide the competitors into strategy groups. There are three levels of customer requirement in the digital camera market.¹¹ First, there is the level of requirement for less than 2 megapixels, mainly for use in web pages or printing stickers. This strategy group would go on to compete with camera-equipped mobile telephones. Next, there is the level which requires more than 2 megapixels for leisurely enjoyment of snapshots or other photos. Finally, there is the requirement for SLRs used by professional photographers or camera maniacs, requiring distortion-free images made possible

10 Refers to cameras which use silver salt film. Normally these are called optical cameras, but since digital cameras are also optical cameras, the Camera and Imaging Products Association (CIPA) has standardized on using this naming convention.

11 This was confirmed through interviews with the Camera and Imaging Product Association (CIPA). Industry statistics here are made from category of cameras with less than and more than 2 megapixels. According to survey results of CIPA, the requirement level for users of SLR cameras is 6 megapixels or more.

through the combination of resolutions over 6 megapixels and large diameter lenses, and additionally the requirement of advanced features such as high speed continuous shooting. As is evident from Table. 2, most companies are ramping down sales of basic products with less than 2 megapixels. This is likely a due to market share being eroded by camera-equipped mobile telephones.¹² To summarize the above discussion, the digital camera industry is in a running battle where optical device manufacturers compete on high end products like SLRs and both optical and electronics manufacturers compete on lightweight compact products in the compact category. On the other hand, all the competitors have released products in the standard category. In other words, the competitive makeup of the digital camera industry appears to be one where two vectors coexist: one vector that aims for volume expansion through the homogeneity of standard type devices while successively increasing pixels, and a different vector that aims to create new value in the compact and SLR categories. For optical device makers and electronics device makers possessing different core technologies and following different successive improvement and destructive technology vectors, we will consider what role did they each filled and how they related to each other.

Figure. 1 showed that the 4 types of cameras successively increased the number of pixels. By focusing on and investigating this pixel increase in detail, we find that for standard and compact types, there is an apparent tendency toward modularization primarily for three reasons. First, we find that after 2000 the CCDs were standardized on 3 types: 1/2.7 and 1/1.8 (both made by Sony) and 1/2.5 (made by Matsushita). the 69.1% concentration of these components in 2000 has reached as high as 91.1% in 2003. Next, for external media, other than Fuji Film and Olympus which use the xD Picture Card and Sony which uses the Memory Stick format, most manufacturers have gathered around the SD card standard. Finally, for interfaces, the USB port and color conversion methods have become standardized and software has become modularized. In other words, software reuse is progressing. This can be clearly seen from the year on year increase in models being released, as well as from the perspective of the reduced burden for new product development. As shown above, while indirect, these three reasons suggests the advancement of modularization for standard and compact digital cameras. On the other hand, SLR product development is handled completely differently. First, the sensor components are designed specifically for each product. In Canon's case, $22.7 \times 15.1\text{mm}$, $28.7 \times 19.1\text{mm}$ CMOS sensors and $35.8 \times 23.8\text{mm}$ CCD sensors are especially developed, and through pairing with aspheric lenses and specialized imaging engines, products are brought to market through proprietary specifications.

¹² According to digital camera production statistics from the Camera and Imaging Product Association (CIPA), 2 megapixel or less cameras in 1999 were 3.4 million units and 3.8 million units in 2000, 3 megapixel and higher models were 1.77 million units in 1999 but increased to 6.98 million units, at this point exceeding the production of 2 megapixel models. In the following 2002, production of models with less than 2 megapixels was 2.3 million units, and 2 megapixels unit production had grown to 21 million units, more a difference of more than tenfold.

Table 2 : Number of Released Products and Product Types

Company	Number of Products Released										Basic			Standard			Compact			SLR			
											Product Name	N	Period	Product Name	N	Period	Product Name	N	Period	Product Name	N	Period	
	1995	1996	1997	1998	1999	2000	2001	2002	2003	Total	DS, Clip, Eyeplate, Camedia, D, VC	18	1995-2000	Finepix P, PM, PS, Camedia C	29	1998-1996	Finepix PA, PF, Camedia X	12	2001-2002	Finepix Pro	2	2000-2000	
Fuji Film	1	2	4	7	7	9	11	12	8	61	DS, Clip, Eyeplate, Camedia, D, VC	18	1995-2000	Finepix P, PM, PS, Camedia C	29	1998-1996	Finepix PA, PF, Camedia X	12	2001-2002	Finepix Pro	2	2000-2000	
Olympus	1	3	5	5	7	13	11	8	12	65	DS, Clip, Eyeplate, Camedia, D, VC	5	1995-2000	Camedia C	48	1996	Camedia X	8	2002	E	4	2000-	
Nikon	—	—	2	3	4	2	6	7	6	30	—	—	—	COOLPIX	25	1997-	—	—	D	5	1999-		
Canon	—	1	2	3	2	4	9	11	11	43	—	—	—	Power Shot	28	1996-	IXY	9	2000-	EOS Digital	6	2000-	
Sony	—	2	4	6	6	11	13	14	15	71	—	—	—	MVC, CyberShot	63	1996-	CyberShot T, U	8	2002-	—	—		
Kodak	1	3	2	3	6	4	3	—	1	23	—	—	—	DC, DX	22	1995-2001	—	—	DCS Pro	1	2003-		
Ricoh	1	5	3	3	3	4	3	2	7	31	DC	12	1995-1998	RDC, Caplio	19	2001-	—	—	—	—	—		
Casio	1	4	4	3	3	5	7	7	6	40	—	—	—	QV, LV, XV	31	1995-	EXILIM	9	2002-	—	—		
Sanyo	—	—	2	3	3	3	3	3	4	21	—	—	—	XACTI	21	1997-	—	—	—	—	—		
Epson	—	2	1	2	2	1	2	—	2	12	CP	10	1996-2001	Calario	2	2003-	—	—	—	—	—		
Panasonic	—	—	5	1	1	3	2	4	6	22	COOLSHOT, LKR	10	1997-2000	Luminux FZ, LC	8	2001-	LUMIX F	4	2001-	—	—		
Minolta	—	—	1	1	3	2	5	5	7	24	—	—	—	DiMage	19	1997-	DiMageX	5	2002-	—	—		
Konica [i]	—	1	2	2	1	—	5	6	4	21	e-mini, Qmini, Q	8	1996-2001	Revio KD, DG	12	1998-	Revio C	1	2002-	—	—		
Toshiba	—	—	1	2	4	2	3	6	2	20	ADR, PDR	2	1997-1998	Alegro	14	1998-	Sora	4	2002-	—	—		
Kyocera	—	—	1	1	1	1	1	5	9	19	—	—	—	Finecam, Contax	17	1997-	Finecam SL	2	2003-	—	—		
Pentax	—	—	1	—	—	2	2	5	7	17	—	—	—	Optio	15	2001-	Optio S	—	—	istD	1	2003-	
Sigma [ii]	2	5	5	2	4	11	5	3	4	44	—	—	—	—	42	—	—	—	—	SIGMA SD	2	2002-	
Total Models	7	28	46	47	57	79	91	98	111	564	—	65	—	—	415	—	—	63	—	—	—	21	—

[i] Due to the merger of Konica and Minolta, two cameras were brought to market under the Konica-Minolta brand name in 2003. These have been added in the 2003 Konica totals.
 [ii] The "Other" category includes models from the following 19 companies: Apple Computer, Chinon, NEC, Polaroid, Sega, Plus, Sharp, Victor, Mitsubishi Electric, Tomy, Takagi Industries, Takara, Bandai, Nichimen, Hitachi, Muji, Maxell, Leica, and Sigma.

Other SLR manufactures also establish proprietary designs based on special sensors combined with optics technology. An SLR camera must be able to provide rapid continuous shooting functionality. Digital cameras read the image through the optic sensor and store this digitally converted data in memory. To swiftly process the large amounts of data inherent in large pixel sizes, this function is handled by the imaging engine each company has developed in-house. Furthermore, to meet the exceptionally high customer requirements, there are products which can provide the raw data before filter processing, which have special power circuit designs, and which have anti-dust features for large optical sensors. These are totally integrated designs.

In this way, the digital camera industry is an industry where opposing product architectures of modularization and integration co-exist, and the dichotomous industry frameworks of horizontal specialization and vertical integration are intermixed.

3.2 Considering Product Competitive Advantage

For each of the four SLR, compact, standard and basic product types, Table. 3 represents the shift in digital camera prices. Here, we calculated and plotted the unit cost (per pixel cost) for the market price¹³ per 1 million pixels in the year of introduction and performed primary regression. As the diagram shows, the regression reveals a strong correlation coefficient. For a compact product, the average pixel unit cost in 2003 was calculated at 14,140 yen / 1 million pixels, meaning a 3 megapixel camera would have sold for about 42,000 yen. As shown in the diagram, for the four years from 2000 to 2003, the pixel unit price has dropped by roughly 75%. As for type differences, standard and compact types have shown nearly identical price movement. The rate of price reduction is even larger for basic models, and as a result the per-pixel unit cost is the lowest. The rate of price reduction for the SLR type is the largest, but the price disparity between SLR and other types remains great.

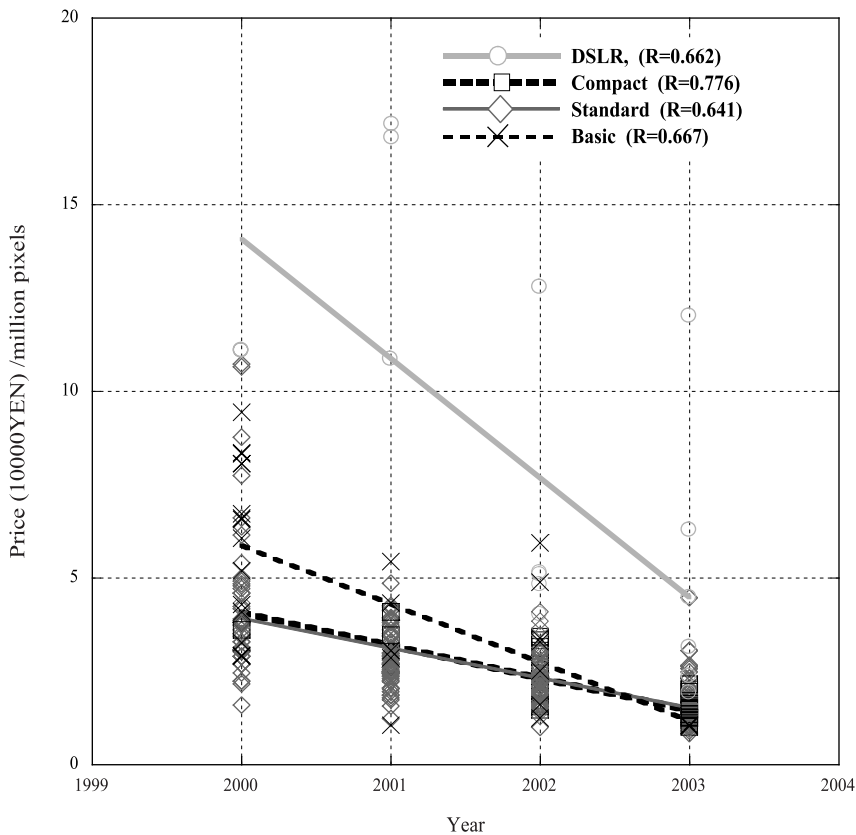
Notably, the results in Figure. 3 show no difference in pixel unit cost between basic, standard and compact models. In reality, there is little range in the sale price charged by different discount retailers; the only notable price difference was for models with different pixel quantities. In other words, the market is a competitive one with many entrants, none of whom are able to charge a reasonable premium. As possible factor could be that many optical device makers and electronics manufactures with brand value have entered the market, and due to high rate of market expansion it is difficult for any of them to have a runaway hit product. Furthermore, to cope with the rapid market expansion, horizontal specialization has accelerated, resulting in increasingly similar products appearing in the market. As mentioned previously, two companies supply nearly all optical sensors, the core component of digital

¹³ The market price information here was gathered from in store prices averaged to 1,000 yen intervals. In store price information was gathered from various price comparison websites. Price differences between discount stores were small with little distribution.

cameras, and for external memory, the market is evolving into one where product advantages for external memory are becoming difficult due to the standardization of formats such as the SD card. For optical device manufacturers, products in most of the categories are in fact OEM consignments from electronics manufacturers which results in ever-increasing product similarity.¹⁴ However, because of the fast market growth rate and market slack, all the competitors have been able to maintain their rate of growth.

On the other hand, the SLR market has become an oligopoly market based on product differentiation. There was the pre-existing silver salt film-based SLR market, but the technology to combine CCD or CMOS optical sensors and large diameter lenses, instantly capture large amounts of data from megapixel photosensors, convert to electronic data and save to memory, was not on the growth path of standard digital camera technology. It would not have been possible to address these requirements without the destructive merging of techno-

Fig. 3 Change in Pixel Unit Cost



14 For instance, Olympus, Nikon, Kodak, Pentax, Konika and Minolta have their standard model OEM supplied by Sanyo Electric and only produce SLR models in-house.

logical advances cultivated by optical device manufacturers and new electronics technologies. However, there is movement towards standardization in SLR arena as well. Through the Four Thirds Specification,¹⁵ and participation of companies having critical technology used in CCD and lenses, the aim is to provide higher quality SLR cameras at a lower price.

Let us think about this market from the perspective of two previous studies. Henderson and Clark performed research related to company products and their modularization. The invention of the digital SLR at first appears to be an evolutionary step up from a standard digital camera, but in reality, to make large diameter lenses a reality, it was necessary to develop large optical sensors with a completely different design than existing small CCDs. Innovations to process that data, resist dust, and stabilize images were also necessary, leading to this unavoidable “architectural shift.” Optical device makers dealt with this architectural change in various ways. For example, Canon separated digital SLR development from its existing standard digital camera development. It dealt with architectural change by separating its standard camera business by purchasing CCDs from Sony and Matsushita Electric and sending the manufacturing operations to Taiwan, but manufactured SLR cameras based on in-house technology at its Oita factory. On the other hand, Olympus entered into new agreements with corporations possessing core technologies, such as Kodak or FujiFilm, Matsushita Electric, Sanyo Electric, and Sigma, and developed its products based on the new Four Thirds specification. As pointed out by Henderson and Clark, corporations have to deal with architectural change by either internally dividing technology development organizations, or realizing change in coordination patterns by building new inter-company relationships.

Let us also examine this subject from the perspective of core competence. Successive improvement of standard and compact digital cameras was a core competence of digital camera manufacturers, who were able to increase pixel count and make their products more lightweight. Christian states that when companies assume they possess the capability to respond to demands of existing customers, they risk falling into the “competence trap.” By setting up new coordination patterns, corporations like Canon, Nikon or Olympus were able to avoid the competence trap and build structures for customers who were excited about the new digital SLR technology, mirroring Christensen’s assertions.

In this section, we have considered product competitive advantage for the digital camera industry. Standard, compact and basic products met customer demand through increased pixel counts gained through successive improvement, but due to the rapid increase in demand a large market slack emerged. Product development, modularization and manufacturing horizontal specialization continued, which led specification homogeneity, resulting in with flat prices

¹⁵ This suggests the interchangeability lens specifications for digital cameras using 4/3 inch size CCD units. Kodak, Olympus, FujiFilm, Sanyo Electric, Sigma and Matsushita Electric all participated and as of 2004, products meeting this specification are already appearing on the market. It is thought that this is a plan for Matsushita is repurchases for CCD sensors while Sigma provides the lenses, and OEM fulfillment by Sanyo Electric.

due to a competitive marketplace where charging a premium is impossible. On the other hand, the digital SLR required an architectural change which could not be achieved through the natural growth of existing digital camera technologies, and new value was created by establishing an integral product architecture which differed from standard digital cameras. The digital camera market has therefore segmented according the customer requirement levels, and it is a market which demands different product strategies for these different requirements.

4. Conclusion

The digital camera industry is an industry which typifies Japan's industry competitive power. Until the previous section, we discussed this competitive strength from the point of view of creation of a structure to build new market with both successive product improvement and destructive technological innovation. For the conclusion of this report, we consider how a company's choice of product development activity affects how they create and capture value.

4.1 Change of Architecture through Competition

We have analyzed the competitive makeup of the digital camera industry. Our analysis showed that this competition was a combination of successive product improvement, such as increased pixels and more lightweight and compact, and destructive technological innovation, such as digital SLRs.

Here, we show that integration and modularization occurs through differences in product architecture and customer requirement levels, and also that through this change in product architecture, companies separate their product development activities into the two strategy vectors of value creation and value capture. Figure. 4 shows the product development orientation and the makeup of value creation and value capture for corporate activities. Let us consider each of these. First, most companies were conducting technological research on digital cameras in the late 1980s, but at the time nobody had thought that they would replace silver salt cameras.¹⁶ In this kind of situation, where companies were unable to sufficiently

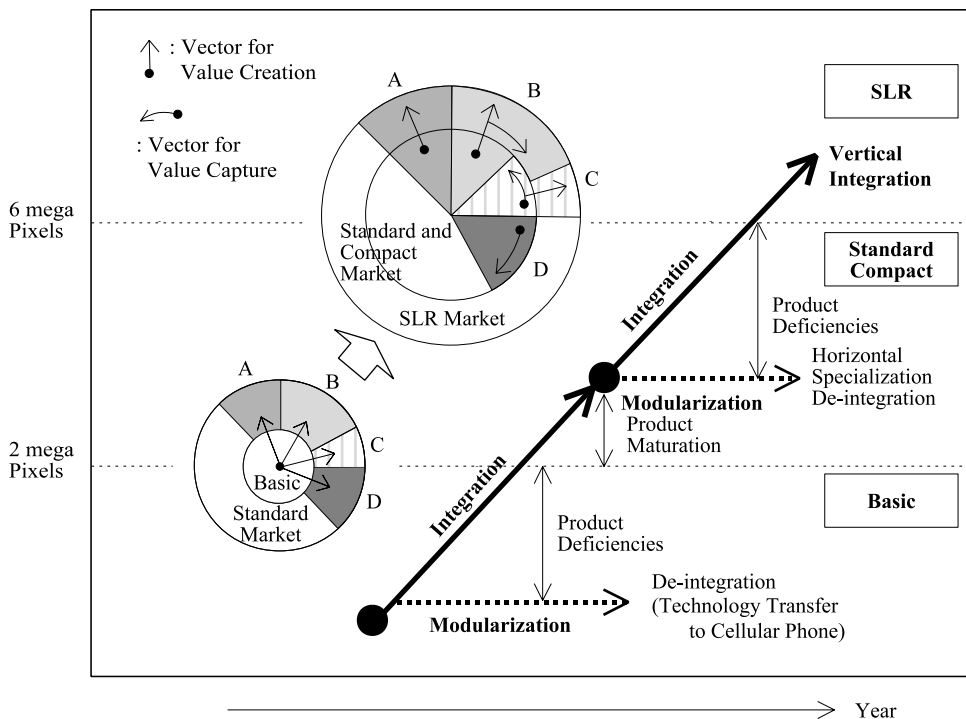
16 By the late 1980s companies like Fuji Film, Sony, and Kodak had already begun technological research aiming to make digital camera breakthroughs. We conducted interviews with Research Director Mr. Nanai who was responsible for Kodak's digital camera research and development from the late 1980s until the birth of the digital camera in 1995, "At the time, nobody even thought that digital camera quality would every reach the quality of silver salt film. We conducted our research whole heartedly assuming that it would be applied to only specialized applications such as test shooting in photo studios. Kodak developed a 1 megapixel CCD before anyone else, but even thought it had an advantage we were unable to take control of the market. The sales and profit of silver salt film at the time was so great and in comparison the sales of the first digital camera were so small, we thought that it was a miserable opportunity." In other words, at the time nobody thought that early digital cameras would replace silver salt cameras, and they were developed for specialized applications.

satisfy the customer, to compensate for product shortcomings, product development needed to increase product performance through trial and error while sticking to the integral product architecture. Then, when the product is mature and customers are satisfied, de-integration occurred through which product architecture was modularized and core modules were sold externally, and horizontal specialization continued to grow. Additionally, to resolve product shortcomings to meet the top customer requirements for products such as SLR cameras, again, product architecture was revisited, integration occurred and new market value was created. In this way, product architecture integration occurred to compensate for product deficiencies, and modularization occurred when the products were mature.

Product architecture modularization encourages horizontal specialization and de-integration. Actually for the standard market, the core CCD module is provided by oligopolistic firms, the SD card has become the dominant external memory format, and the number of LCD suppliers is limited.

Based on kind of product architecture change, the strategy division that occurs between competitors is as follows. The circle graph in the lower-left corner Figure. 4 shows proprietary specifications of digital camera manufactures, such as the realization of resolutions approaching that of silver salt cameras made possible through integration. Under these conditions, new

Fig. 4 Directed Vectors of Product Development



demand was created, and each company could freely expand its sales. Then, as the market took form, various types of similar products came to market, homogeneity took hold and the market became more competitive. The large circle graph in the middle of Figure. 4 shows this competitive situation. There are four strategy groups. First A, which represents corporations who performed vertical integration based on in-house technology for both standard and SLRs. In group B, in-house development of SLRs continues and standard models are outsourced, resulting in a shift toward horizontal specialization. C refers to de-integrated companies who focus on cost competition through standard model modularization and actively sell core technology as modules. Finally, D represents competitors who maximize their cost competitive advantage for standard models, and supply OEM products to other companies.

Why does this kind of strategic division of competitors occur? In a market like the digital camera market, where the customer requirement levels are diversified, corporations can choose from several strategic options. New markets are created when new technology fulfills customer requirements and insufficient performance concerns are resolved. In the diagram, this is the value creation vector which points from the center of the circle graph outwards. This sort of value creating activity is realized commonly through inter-company cooperative relations such as tie-ups and consortiums. Representative of this kind of consortium is the SLR camera Four Thirds specification or SD card of the early digital cameras. On the other hand, after value is created and the market matures, it evolves into fight over limited PIE.¹⁷ In the diagram, this is the value acquisition vector which curves around the the circle graph. These value capture activities are totally competitive in nature. As shown in the previous chapter, through horizontal specialization and de-integration, product homogeneity continues and companies compete on cost. In this way, companies' strategies are determined by the pull of collaboration and competition, value creation and value capture vectors on each other. In the digital camera industry there is a strong tendency for optical device manufacturers to work to create value and electronics device manufacturers to work to capture value. This is because a company's strategy is different depending on if it regards digital cameras as optical devices or as electronic devices.

4.2 Summary

We have looked at digital camera product development from the perspective of product architecture. By analyzing all the products in one industry, we attempted to define the intricate details of competitors' product development strategy based on the market trends. We close this report here by showing and summarizing the knowledge we gained through analysis.

15 PIE: Potential Industry Earning, A measure of the latent earning potential of a given market.

First, product architecture should be determined based on its relation to customer requirement levels. Integration is an architecture adopted to fill the customer expectations gap and aims to create value. Modularization, on the other hand, is an architecture which aims to capture value after customer requirement levels have been satisfied. When implementing integration to create value, inter-company collaboration is natural, and this kind of cooperation encourages modularization, but because modularization works to de-integrate the industry and encourages horizontal specialization, corporations try to capture value, and as a result, companies compete with each other on cost. In the digital camera industry, through this kind of collaboration and competitive structure, companies differentiate into strategic groups. When this kind of strategic group differentiation occurs, vertically integrated companies do not always have to adopt their in-house technologies which may be thought as weak, and coordination becomes possible where companies can choose and collaborate with the most appropriate partner company. The digital camera market in Japan is strong because all the conditions have been met which make full domestic coordination possible.

References

- Baldwin, C. Y. and K. B. Clark (2000). "Design Rules: The Power of Modularity." MIT Press, Cambridge, MA.
- Brusoni, S. and A. Prencipe (1999). "Modularity in Complex Product Systems: Managing the Knowledge Dimension." Working Paper, SPRU, Sussex University.
- Christensen, C. M. and M. E. Raynor (2003). "The Innovator's Solution: Creating and Sustaining Successful Growth." Harvard Business School Press, Boston.
- Fixson, S. (2001). "Three Perspectives on Modularity — A Literature Review of a Product Concept For Assembled Hardware Products." MIT Working Paper, ESD-WP-2001-06.
- Fixson, S. (2003). "The Multiple Faces of Modularity — A Literature Analysis of a Product Concept for assembled Hardware Products." Technical Report 03-05, Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI: 87.
- Fine, C. (1998), "Clockspeed — Winning Industry Control in the Age of Temporary Advantage." Perseus Books: New York.
- Galworth, G. D. (1994). "Smart, Simple Design." Oliver Wight Publications, Essex Junction, Vermont.
- Garud, R. and A. Kumaraswamy (1995). "Technological and Organizational Designs for Realizing Economies of Substitution." *Strategic Management Journal*, Vol. 16 (Summer Special Issue), pp. 93-109.
- Gawer, A. and M. Cusumano (2001), "Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation." Harvard Business School Press, Boston.
- Gonzalez-Zugasti, Javier P., Kevin N. Otto and J. D. Baker (2000), "A Method for Architecting Product Platforms." *Research in Engineering Design*, Vol. 12, pp. 61-72.
- Henderson, R. M. and K. B. Clark (1990). "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms." *Administrative Science Quarterly*, Vol. 35, pp.9-30.

- Hyer, N. L. and U. Wemmerlov (1984), "Group Technology and Productivity." *Harvard Business Review* (January / August, 1984), pp. 3-12.
- Meyer, M. H. and A. P. Lehnerd (1997) , "The Power of Product Platform." The Free Press, New York.
- Mikkola, J. H. (2001), "Modularity and Interface Management: The case of Schindler Elevators." Proceeding of the DRUID 2001 Winter Conference Korsor, Denmark, January 17-19, 2001.
- Newcomb, P. J., B. Bras and D. W. Rosen (1998). "Implications of Modularity on Product Design For The Life Cycle." *Journal of Mechanical Design* Vol. 120, pp. 483-491.
- Nobeoka, K. and M. A. Cusumano (1994), "Multi-Project Strategy and Market Share Growth: The Benefits of rapid Design Transfer in New Product Development." International Center for Research on the Management of Technology Working Paper # 105-94.
- O'Grady, P. (1999). "The Age of Modularity — Using The New World of Modular Products to Revolutionize Your Corporation." Adams and Steel Publishing
- Pine, B. Joseph II (1993). "Mass Customization: The New Frontier in Business Competition." Harvard Business School Press, Boston.
- Robertson, D. and K. T. Ulrich (1998), "Planning for Product Platforms." *Sloan Management Review*. Vol.39 (4), pp. 19-31.
- Sako, M. and F. Murray (1999). "Modules in design, Production and Use: Implications for the Global Automotive Industry." Proceeding of the International Motor Vehicle Program (IMVP) Annual Sponsors Meeting, October 1999, Cambridge Massachusetts.
- Sako, M. (2003), "Modularity and Outsourcing: The Nature of Co-evolution of Product Architecture in The Global Automotive Industry." Proceeding of Eleventh Gerpisa International Colloquium
- Sanchez, R. and J. T. Mahoney (1996). "Modularity, Flexibility, and Knowledge Management in Product and Organization Design." *Strategic Management Journal*, Vol. 17 (Winter Special Issue), pp.63-76.
- Schilling, M. A. (2000). "Towards a General Modular Systems Theory And Its Application to Interfirm Product Modularity." *Academy of Management Review*, Vol. 25, pp. 312-334.
- Schilling, M. A. (2002). "Modularity in Multiple Disciplines: Managing in the Modular Age: Architectures, Networks and Organizations. R. Grud, A. Kumaraswamy and R. N. Langlois." Blackwell Publishers, Oxford.
- Simon, H. A. (1962). "The Architecture of Complexity." Proceeding of the American Philosophical Society, Vol. 106 (6), pp. 467-482.
- Smith R. P. and S. D. Eppinger (1997). "A Predictive Model of Sequential Iteration in Engineering Design." *Management Science*, Vol. 43, No. 8, pp. 1104-1120.
- Starr, M. K. (1965). "Modular Production — A New Concept." *Harvard Business Review* Vol. 43 (November-December), pp. 131-142.
- Stone, R. B. and K. L. Wood (2000). "Development of a Functional Basis For Design." *Journal of Mechanical Design* Vol. 122, pp. 359-370.
- Sturgeon, J. S. (2002). "Modular Production Networks: A New American Model of Industrial Organization." *Industrial and Corporate Change* Vol. 11: 3, pp. 451-496.
- Sturgeon, J. S. (2003). "Exploring The Risks of Value Chain Modularity: Electronics Outsourcing During The Industry Cycle of 1992-2002." Working Paper, MIT-IPC-03-003.
- Swan, W. A. (1914). "Proposes Standardization of Car Sizes." *The Automobile* Vol. 31, pp. 76-77.
- Thomke, S. and D. Reinertsen (1998). "Agile Product Development: Managing Development Flexibility in Uncertain Environment." *California Management Review*, Vol. 41, pp. 8-30.

- Ulrich, K. T. and K. Tung (1991), "Fundamentals of Product Modularity." Working Paper WP #3335-91-MSA, MIT Sloan School of Management.
- Ulrich, K. T. (1995), "The Role of Product Architecture in The Manufacturing Firm." *Research Policy*, Vol. 24, pp. 419-440.
- Von Hippel, E. (1990). "Task Partitioning: An Innovation Process Variable." *Research Policy* Vol. 19 (5), pp. 407-418.
- Wilhelm, B. (1997). "Platform and Modular Concepts at Volkswagen — Their Effects on the Assembly Process." *Transforming Automobile Assembly — Experience in Automation and Work Organization*. K. Shimokawa, U. Juergens and T. Fujimoto. Berlin, Springer, pp.146-156.
- Yu, J. S., Javier P. Gonzalez-Zugasti and Kevin N. Otto (1999). "Product Architecture Definition Based Upon Customer Demands." *Journal of Mechanical Design*, Vol.121, pp.329-335

KOBE ECONOMIC & BUSINESS REVIEW

CONTENTS

No.48 (2003)

The Usefulness of Financial Statements Under Chinese-GAAP vs. IAS:

Evidence from the Shanghai Stock Exchange in PRC. *Hu Dan*

The Effects of Structural Changes in Financial System on Corporate Governance in Japan *Kenji Kojima*

A Dynamic Model of Export Adjustment with Deep-pocket Effect:

Evidence from Japanese Keiretsu Firms *Eiichi Tomiura*

Roscas and Credit Unions: Is Modern Japan Missing Something? *Stephen J. Davies*

No.47 (2002)

The Theory of International Trade *Ronald W. Jones*

The Fall of 'Companyism' in Japanese Corporate System *Kenji Kojima*

Creation of New Enterprises and Development of Existent Enterprises

in East Asia and Latin America *Akio Hosono*

Dynamic Export Pricing and Survey-based Exchange Rate Expectations *Eiichi Tomiura*

Technological Trajectories and the Origins of New Industries: A Theoretical Model *Jeffrey L. Funk*

Japan's Deflation and the Feasibility of Negative Interest Rates *Stephen J. Davies*

The Past and Present of "Local Currencies"; with Especial Reference to England and Japan *Stephen J. Davies*

No.46 (2001)

Economic Theory: Past and Future *Murray C. Kemp*

Long-Term Cooperation in Japanese Vertical Relationships *Kenji Kojima*

Experiences of Technological Partnership of Japan with Overseas:

The Case of Steel Industry *Akio Hosono / Nobuaki Hamaguchi*

Analysis of Adoption of Forest Certification in Japan

From Perspective of Organizational Theory *Akira Kajiwara*

Sensitivity of Domestic Production to Import Competition:

Evaluation at Different Levels of Aggregation *Eiichi Tomiura*

Japan-Korea Free Trade Area and Structural Reforms *Bonggil Kim / Kazuhiro Igawa*

No.45 (2000)

Imperfect Substitutes and Strategic Trade Policies under Cournot Duopoly:

- Mathematical Note*Seiichi Katayama / Hiroshi Ohta*
Shifts in the Japanese Corporate Governance*Kenji Kojima*
FSC Certification in Japan: An Uncertain Future*Akira Kajiwara*
The Long-Run Impact of APEC Trade Liberalization on Real GDP and Sectoral Adjustments
.....*Hiro Lee / David Roland-Holst / Dominique Van Der Mensbrugge*
Monetary Cooperation in East Asian Countries:
A Possibility from Macro Economic Indexes and Intra-Regional Trade Dependency
.....*Bonggil Kim / Kazuhiro Igawa*
Discretionary Accrual Models and the Accounting Process
.....*Xavier Garza-Gomez / Masashi Okumura / Michio Kunimura*

No.44 (1999)

A Note on Possibilities about Japan-Korea Free Trade Area:

- A Strategic Approach*Kazuhiro Igawa / Bonggil Kim*
Crisis Cambial Brasileira
-Fundamentos vs. Ataque auto-realizavel-*Shoji Nishijima / Eduardo Kiyoshi Tonooka*
Japanese Supplier Relations:
A Comparative Perspective*Kenji Kojima*
Development of the Forest Stewardship Council
Certification Program*Akira Kajiwara / Todd Malinick*

Published on February 2005
by
Research Institute for Economics & Business
Administration (RIEB)
Kobe University

2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501 Japan
Phone : 81-78-803-7036
F a x : 81-78-803-7059
Home Page at <http://www.rieb.kobe-u.ac.jp/>

Printed in Japan