RIETI Discussion Paper Series 23-E-001

Long-run Effect of a Horizontal Merger and Its Remedial Standards

FUKASAWA, Takeshi
Tokyo University

OHASHI, Hiroshi
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

1RIETI Discussion Paper Series 23-E-001
January 2023

Long-run Effect of a Horizontal Merger and its Remedial Standards

Takeshi Fukasawa (Tokyo University) * Hiroshi Ohashi (RIETI)†

Abstract

This paper estimates a dynamic oligopoly model with firms' continuous investment decisions to assess long-run consequences of a horizontal steel merger. It employs a novel simulation method to show that the merger improved social welfare. While the merger discouraged the merged firm from investing in capacity, it encouraged investment within non-merged firms, absent the efficiency gains of the merger. The paper also evaluates the remedial measure targeting asset divestiture that was endorsed by the competition authority. The paper finds that the effects of the merger remedy persisted for the 20 years after its implementation covered by this study, and the prescribed remedy differed considerably on the standpoint of either consumer or social welfare standards.

Keywords: Horizontal merger, Merger remedies, Long-run effect, Investment, Consumer Consumer welfare standard, Total surplus standard

JEL classification: D43, L13, L41

*Graduate School of Economics, University of Tokyo.
†Faculty of Economics, University of Tokyo, Tokyo 113-0033, Japan.
This study is conducted as a part of the Project "Globalization, Innovation, and Competition Policy" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). We are grateful to seminar participants at RIETI for their helpful comments.

1 Introduction

Horizontal mergers affect not only firms' pricing or output choices, but also firms' investment choices. Also, merger remedies, in which competitive authorities redistribute the firms' capital to mitigate the problems associated with the merger, might affect firms' investment decisions. Since firms' endogenous investment choices might affect the welfare consequences of the policies, we need to take account for the firms' investment choices in evaluating horizontal mergers and merger remedies.

This paper attempts to quantitatively assess the long-run consequences of a horizontal merger and merger remedies, allowing for firms' endogenous investment behavior. For this purpose, the paper uses a unique case of a horizontal merger and merger remedies that occurred in the Japanese steel industry. In 1970, Japan celebrated the birth of Nippon Steel, the world's second largest steelmaker. The new Japanese company came into being through the merger of Yawata and Fuji, the two largest Japanese steel producers at the time. The merger was approved by the Japan Fair Trade Commission (J-FTC) under the condition that the merging party would transfer a total of 1.8 % of its capital equipment to two smaller firms, namely, Nihon Kokan and Kobe. A notable aspect of the Japanese steel industry in the postwar period was its active investment in new plants and equipment. This paper quantifies the extent to which investment on capital affected steel production and explicitly accounts for the dynamics resulting from firm's capital investment behavior.

Although there does not appear to exist any opportunity to conduct controlled experiments on the 1970 merger and merger remedies, we can still perform counterfactual exercises by following two steps. The first is to use observed data along with an structural model of the steel industry, where oligopolists make optimal decisions regarding production and investment on the basis of their competitors' strategies, to estimate the parameters of the underlying economic primitives that were invariant in the horizontal merger and merger remedies. We then simulate changes in equilibrium outcomes on the basis of the counterfactual situation in which Yawata and Fuji do not engage in the merger. We also consider another counterfactual situation in which the merger took place in the absence of the merger remedies. For the simulation approach to be successful, the model used for the exercise must closely approximate the economic environment under study. We follow the research of Ericson and Pakes (1995) to compute a Markov-perfect equilibrium of the dynamic game in order to take account for the firms' capital investment processes.

Regarding the merger, the results show that the merger improved social welfare, though it lowered consumer welfare. We can attribute the improvement in social welfare mainly to the efficiency gains of the merged party. After the merger, merged party could lower its production cost, and it largely increased the producer surplus. Furthermore, though the merger discouraged the merged party's investment, it encouraged the non-merged party's investment, and firms' endogenous investment behaviors after the merger contributed positively to the improvement in social welfare. The reason is that lower investment of the merged party in a dominant position and non-merged party's larger investment make the firms more symmetric, and it makes the market more competitive. Also, marginal gains from investment is larger for smaller scale firms, and encouragement of non-merged firms' investments largely lower the production costs. The result implies we should also pay attention to the non-merged party's investment decisions in evaluating the effect of the merger, though lower merged party's investment seems to be the largest problem.

To interpret and clarify the mechanisms behind the effect of the merger on firms' investment, we further conduct a detailed decomposition of firms' investment incentives. The quantitative results show that the change in output largely affected the firms' investment strategies, which is classified as "margin expansion effect" in Jullien and Lefouili (2018). Merged party has less incentive to invest, because its optimal output is smaller after the merger than the case absent the merger. Non-merged firms have more incentive to invest, because its optimal output is larger after the merger than the case absent the merger. Nevertheless, other factors, including the scale economy in production cost and the range of the product affected by a marginal investment, also play large roles in the directions of effect of the merger on the firms' investment.

Regarding the merger remedies, our simulation results show that the merger remedies failed to fully offset the loss of consumer welfare, though it improved consumer welfare. Also, it lowered the total surplus. Regarding the long-term effect of the merger remedies, roughly half of the effect persisted even 10 years after the merger remedies, though the effect of the merger remedies had diminished over time. One reason for these long-run effect is the firms' endogenous investment decisions. In an extreme case where no adjustment cost in investment exists, firms decide their investment levels so as to equate the marginal revenue from investment and the user cost of capital, and the firms' capital stocks would go back to the no-remedies levels in the next period the merger remedies was implemented. If the adjutsment cost exists, it does not necessarily hold, and it takes time to go back to the no-remedies level. They imply diminishing, but long-run effect of merger remedies.

In this study, we further show the optimal merger remedies based on the consumer welfare standard / total surplus standard. First, we show that the size of optimal merger remedies based on the long-run consumer welfare standard should be larger than that based on the short-run consumer welfare standard. Under consumer welfare standard, we aim at offsetting the loss of consumer welfare due to the merger. Short-run consumer welfare standard only consider the consumer welfare at the period the policy was implemented, but long-run consumer welfare standard also take account of the consumer welfare in the future. As discussed above, firms would endogenously adjust their investment levels, and the effect of the merger remedies might disappear soon after the policy was implemented. Then, the loss of consumer welfare might reappear in the future, even if the merger remedies could offset the loss of consumer welfare at the period the merger remedies. Consequently, larger remedies are required to offset the loss of consumer welfare in the long run. Note that it does not necessarily hold if the competitive authorities can intervene the market even after the period of the merger and merger remedies.

Second, we also show that the optimal distributions of the merged party's divested assets based on the consumer welfare standard might be largely different from the one based on the total surplus standard. Though the consumer welfare standard has been widely applied by competitive authorities and studied by the previous literature, investigating the effect of merger remedies on producer surplus or total surplus is also important. Under consumer welfare standard, it is optimal to distribute the divested assets mainly to smaller firms to recover the symmetry of firms in the market (Vergé (2010)). In the context of the 1970 merger remedies, distributing mainly to the smallest firm Kobe is optimal. Nevertheless, it is not necessarily optimal when we take account of the producer surplus. The distribution to Kobe intensifies the competition in the market, and it lowers producer surplus. Even though this study only show the quantitative evidence in one market, competitive authorities should be careful about the impact on total surplus, even when applying the consumer welfare standard.

ex-post or ex-ante perspective.

1.1 Literature

This paper relates and contributes to three strands of literature.

1.1.1 Merger evaluations allowing investment

First, it adds to a growing body of both theoretical and empirical works on the evaluation of horizontal mergers allowing for firms' investment behaviors [1]. Regarding the effect of horizontal mergers on firms' investment, Mermelstein et al. (2020) considered the setting where a merger discourages merged party's incentive to invest based on a numerical model. Nishiwaki (2016) discussed that mergers promoted divestment based on a structural model of Japanese cement industry. Chen (2009) also showed that a merger reduces merged firm's investment based on a numerical model.

In contrast, Bennato et al. (2021) found a significant evidence of increased R&D after mergers in the hard disk drive industry using the reduced-form approach. Genakos et al. (2018), investigating telecommunication industry using reduced-form approach, showed that per-firm investment is higher in more concentrated markets, even though the impact on total investment is ambiguous.

Besides, Grajek et al. (2019), studying telecommunication industry using reduced form approach, showed that firms invested more in the markets where post-merger prices are higher than pre-merger levels, but invested less in the markets where post-merger prices are lower than pre-merger levels.

In the theoretical literature, Motta and Tarantino (2021) showed that merger always reduces total investments absent efficiency gains, but Jullien and Lefouili (2018) discussed that the effect of mergers on firms' investment is determined by several factors, and the claim by Motta and Tarantino (2021) does not hold in more general settings.

Because of its ambiguity on the effect of mergers on firms' investment, disentangling firms' incentives to invest after mergers helps us better understand the consequence of mergers and desirable merger policies. In our study, we fully incorporate dynamic incentives, such as preemptive motives which does not present in the theoretical studies applying static models, and quantify the contribution of each incentive based on the dynamic structural model on Japanese steel industry. We contribute to the literature by explicitly studying the effect of a merger on firms' investment decisions and its mechanism.

1.1.2 Merger remedies

Second, this paper also contributes to a growing literature on merger remedies. So far, most of the theoretical and empirical studies have investigated the effect of merger remedies and desirable policies by focusing on the short-run effect without accounting for firms' endogenous investment behaviors. Examples include Vergé (2010), Dertwinkel-Kalt and Wey (2016) in the theoretical studies, and Tenn and Yun (2011), Friberg and Romahn (2015), and Osinski and Sandford (2021) in the empirical studies. Nevertheless, as empirically shown in Bennato et al. (2021) based on a reduced-form approach, merger remedies altered firms' investment behavior. Merger remedies is the reallocation of merging firm's capital to non-merging firms, and firms' endogenous investment decisions after merger remedies is also important for assessing the long-run welfare consequence of merger remedies. In this study, we show that the optimal size of merger remedies allowing firms' investment decisions in the long-run should be much larger than the case focusing on short-run welfare ignoring investment decisions, when relying on the consumer welfare standard. Though not discussed in the previous literature, the analysis provides insight into the optimal merger remedies, and we contribute to the literature by stressing the role of firms' endogenous investment decisions in the long run in assessing merger remedies.

[1] See Igami and Uetake (2019) for empirical analysis of endogenous mergers allowing for firms' investment.

1.1.3 Quantitative method to solve dynamic investment competition model

Third, this paper contributes to the quantitative method for solving dynamic investment competition model. One large inhibitant in the application of dynamic competition model for policy analysis, including merger analysis, is the computational complexity of solving the equilibrium.Even though it is possible to estimate dynamic parameters using the two-step method proposed by Bajari et al. (2007) and others, we still have to solve the equilibrium to conduct counterfactual analysis[2]. In the case of continuous investment specification we employ, numerical methods previous studies have used are not so easy to handle or take much time to compute. In this study, we develop a novel algorithm that avoids solving optimal investment in each iteration, which is a nonlinear problem and computationally costly to solve. The idea is to change the structure of the algorithm, and make the optimal investment problem in a linear form. We combine the method with other computational methods, including Smolyak method (Smolyak (1963), Judd et al. (2014)) and spectral algorithm (La Cruz et al. (2006), Aguirregabiria and Marcoux (2021)), and solve the equilibrium with relatively small computational cost[3]. Our dynamic model with firm's investment decision is relatively simple, and the model and the method can be easily modified or extended to other industries.

The rest of this paper is organized as follows. Section 2 provides an overview of the Japanese steel industry in the postwar period, particularly the merger between Yawata and Fuji in 1970. This section also evaluates the merger from a static perspective, as a starting point of the further detailed analysis based on the structural model incorporating firms' endogenous investment decisions. Section 3 presents a dynamic model that explicitly accounts for the dynamics arising from investment behavior, which was one of the most important features of the Japanese steel industry in the 1960s and 1970s. Section 5 discusses the estimation results. Using these results, Section 6 performs policy experiments to evaluate the welfare consequences of the horizontal steel merger and assess the effectiveness of divestiture as a merger remedy. Section 7 concludes. Detailed procedure to conduct the simulation, additional discussion on the effect of the merger on firms' investment, further discussion on the optimal distribution of divested assets, and the details of the data are shown in Appendix A, B, C, and D, respectively.

2 Industry

This section begins with a historical overview of the Japanese steel market, mainly focusing on the study period 1960-1979. It illustrates that each firm's active investment was an essential characteristic of the Japanese steel industry. The features of the market described in this section lead us to develop a dynamic structural model, which is discussed in Section 3. Also, before introducing such dynamic decision making, Section 2.2 uses the static analytical framework without considering firms' endogenous investment choices to preliminarily assess the welfare impact of the merger and merger remedies. This static analysis is inadequate for the study of the steel merger, because it neglects the dynamic features of the firms' investment decision making; however, it

[2]Benkard et al. (2020) proposed a method to assess the long-run effect of mergers without explicitly solving the equilibrium, building on the two-step method proposed by Bajari et al. (2007). Even though convenient to use, they require some assumptions and it is hard to conduct flexible counterfactual analysis.

[3]In our setting, it takes less than 1 minute to run each counterfactual simulation even when incorporating the asymmetry in 6 firms

provides a useful starting point from which to consider the effect of the 1970 merger and merger remedies.

2.1 Overview of the industry

Yawata and Fuji were originally under the same ownership: this was dissolved in 1950 by the occupation forces, who were attempting to create a competitive environment for the Japanese steel industry. At the same time, the occupation forces established the J-FTC along with antitrust monopoly law. However, despite their efforts, only a handful of dominant major firms operated during the 1960s and 1970s. Indeed, over 80 % of Japanese steel production was accounted for by integrated steel manufacturers. These manufacturers transformed raw materials (iron ore and coking coal) into pig iron in a blast furnace. Pig iron was then transformed in a refining furnace into crude steel, the homogeneous product we focus on in this study. The following seven integrated companies enjoyed the largest shares in the market: Nippon Steel, Yawata, Fuji, Nihon Kokan, Kawasaki, Sumitomo, and Kobe (in order of average market share). Note that Nippon Steel was created in 1970 by the merger between Yawata and Fuji. Since then, Nippon Steel has remained the second largest steel producer in the world - after U.S. Steel at the time of the merger and now after Arcelor Mittal. This paper focuses on the above-mentioned seven Japanese integrated companies and characterizes the structure of the market. From 1960 to 1990, no entries and exits took place, except those associated with the Yawata-Fuji merger. Therefore, during the study period, the Japanese steel market was of little relevance to the merger waves observed in other markets, including the U.S. paper and pulp industry studied by Pesendorfer (2003).

Beginning in the 1960s, integrated steel makers faced increasing competitive pressure from a new type of steel producer, namely minimills. In contrast to the integrated steel makers, minimills own no blast furnaces but electric arc furnaces to use steel scrap and electricity as primary inputs to produce crude steel. In the 1960s, minimills appeared to catch up with integrated steel makers in terms of production capacity size and crude steel quality, as electric arc furnaces began producing on an increasingly larger scale with supplies of high-voltage electricity. Thus, the emergence of the alternative source of steel supplied by minimills should have increased the elasticity of steel demand faced by the integrated steel manufacturers.

Table 1 presents important statistics, classified according to the pre- and post-merger periods.

From 1960 to 1990, the Japanese steel industry was characterized by the firm's active investments in new plants and equipment. This feature is demonstrated by the index of the capital-labor ratio, which is defined as the ratio of the value of tangible fixed assets to the number of employees. The index indicates that the capital intensity of the steel industry was three times higher than the average of the manufacturing sector and twice that of the chemical industry. Note that the new production facilities utilized the latest technology, which further pushed firms' production frontiers and thus reduced the cost of steelmaking. Therefore, the active investment observed in the steel industry reflects the firm's incentives for efficient steel production in a market where only a handful of firms dominated. It is also worth noting that non-merging firms invested more in the post-merger period than in the pre-merger period: Table 1 shows that the investment share of a non-merging party became larger after the 1970 merger (54.5% $\rightarrow$ 71.9%), whereas that of a merging party became smaller (45.5% $\rightarrow$ 28.1%). In Section 3, we introduce the dynamic decision-making model to associate this finding with firms' endogenous investment behavior.

The rapid production growth indicated in Table 1 was accompanied by export expansion, and Japan's share of the world export market grew from less than 5% in 1955 to 9% in 1965. Most of Japan's steel had been shipped to Asian countries until the early 1960s, when an increasing proportion began to be exported to North America. Nevertheless, the steel export market was fairly competitive from 1955 to 1990, and there is little evidence that Japanese steelmakers had market power during that period. The Japan Iron and Steel Exporters' association (1974) observed that the Japanese Freight on Board (FOB) steel price was not significantly different from the price in Antwerp, Belgium, which was known as the center of the world steel trade at that time. It is thus reasonable to assume that the exported steel was competitively supplied in the world market[4].

Japan had an import tariff of 15% on steel until 1967 when it agreed to reduce the rate by half at the Kennedy Round of General Agreement of Tariffs and Trades (GATT). However, while the import tariff protected domestic steel makers from direct competition with foreign steel makers, it may have had little to do with the increase in Japanese steel production shown in Table 1, because Japan also exported steel during that period. Indeed, the share of steel imports accounted, on average, for a mere 0.2 % even after the tariff was reduced. We therefore assume that steel imports were not substituted by steel produced by Japanese companies and did not affect the domestic Japanese market.

Table 1: Summary Statistics for Important Variables: Japanese Steel from 1960 to 1990

		Pre-Merger 1960-1969	Post-Merger 1970-1990
Price (in Thousand JPY per tons)		42.55	41.72
		(2.81)	(2.49)
Output (in Million tons)		34.15	74.41
		(16.83)	(16.83)
Capital Investment (in billion JPY)	Merging Party	0.1	0.09
		(0.06)	(0.09)
	Non-merging Party	0.12	0.23
		(0.08)	(0.12)
Capital Stock (in billion JPY)	Merging Party	0.57	2.55
		(0.24)	(0.65)
	Non-merging Party	0.56	4.16
		(0.29)	(1.48)

Notes.
Average values and standard errors are shown.
Price, capital investment and stock are in terms of 1960 price.
The merging party is the sum of Yawata and Fuji in 1960-1969, and Nippon Steel in 1970-1990.
Values in non-merging party in the table is the sum of Nihon Kokan, Kawasaki, Sumitomo, and Kobe.

[4]The assumption is also consistent with the evidence presented in Ohashi (2005), which indicates that the export subsidy in Japanese steel was not based on profit shifting.

Table 2: Market Share of Steel Production: Impact of the 1970 Merger

1969		1970	
	% (Output)		% (Output)
Yawata (merging)	23.73	Nippon Steel (NSC)	45.33
Fuji (merging)	21.65		
Nihon Kokan (nonmerging)	17.5	Nihon Kokan	17.45
Kawasaki (nonmerging)	15.43	Kawasaki	15.05
Sumitomo (nonmerging)	15.34	Sumitomo	15.22
Kobe (nonmerging)	6.35	Kobe	6.94
HHI	1852	HHI	2866
Steel Production for the six firms (Million ton)	68.53	Steel Production for the five firms (Million ton)	72.75

Figure 1: Firms' outputs over time

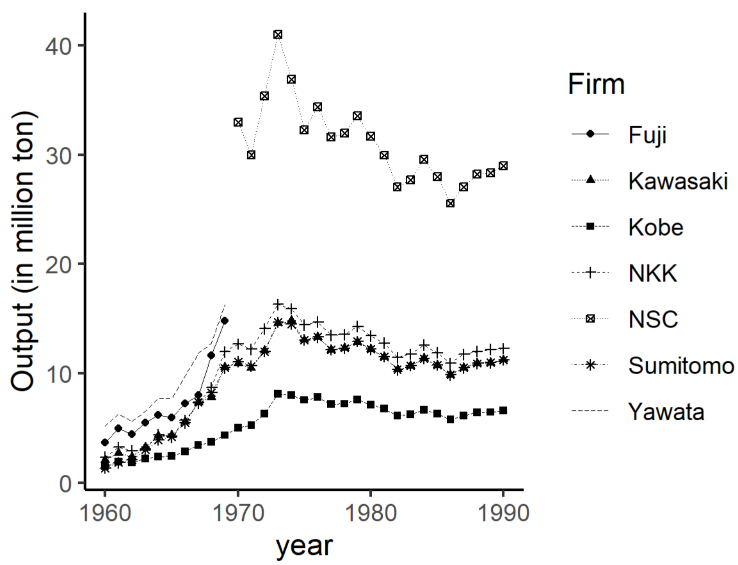
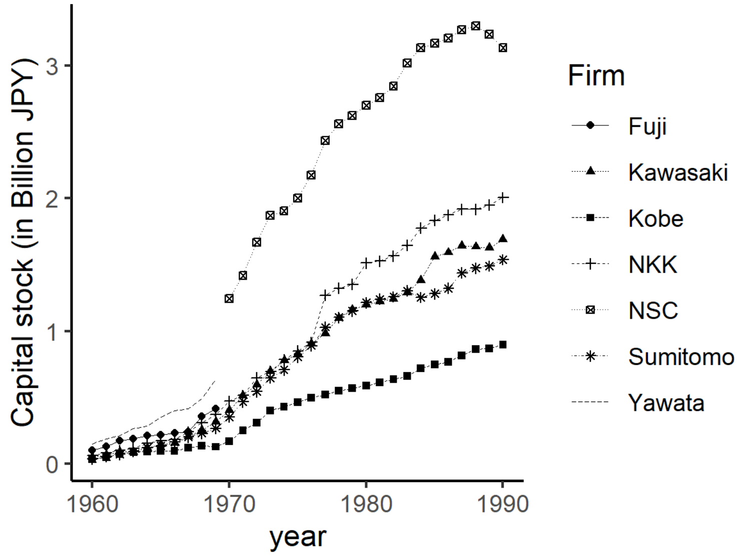


Figure 2: Firms' capital stocks over time



2.2 Preliminary Analysis

2.2.1 Merger

In evaluating a proposed merger, antitrust officials in the U.S. generally apply the rules summarized in the Department of Justice's Merger Guidelines (1992). Traditional merger analysis under the guidelines involves estimating the effect of a proposed merger on market concentration. Roughly speaking, the guidelines permit mergers that will result in either a low initial level of concentration in the industry or small predicted changes in concentration. In the guidelines, concentration is measured according to the HHI, which is defined as the sum of the squares of the firms' market (output) shares. In retrospect, the initial level of and change in HHI due to the merger between Yawata and Fuji exceeded 1800 and 100, respectively, as shown in Table 2. Thus, unless further analysis indicates that entry would be easy or that important efficiencies would be created, the Merger Guidelines would most likely recommend that the merger be challenged.

Careful assessment of horizontal mergers requires in-depth analysis of how they will affect equilibrium output and welfare; however, this analysis was lacking in the abovementioned traditional approach using the concentration index. In the previous literature, Nocke and Whinston (2022) proposed a convenient method to compute the minimum size of the merged party's efficiency gains needed not to raise the equilibrium price and not to harm consumers. In the context of Cournot competition in a homogeneous product market without firms' endogenous investment decisions, It is analytically given by:

$$\frac{c_M - \overline{c}_M}{c_M} = \frac{\Delta H}{s_M(\epsilon - s_M) + \Delta H} \tag{1}$$

where c_M denotes the output-weighted average marginal costs of the merging firms prior to the merger, and $\overline{c}_M$ denotes the merged firm's marginal cost after the merger. ΔH denotes the increment of the HHI index, s_M denotes the sum of merging firms' market shares, and $\epsilon > 0$ denotes the price elasticity of demand. Using the data in the industry, we find that the minimum size of the efficiency gains required is 27.51%. As we show later, the estimated efficiency gains are roughly 20%, implying that the realized efficiency gains could not fully offset the loss of consumer welfare.

Regarding the assessment of horizontal mergers based on the total surplus standard, Farrell and Shapiro (1990) derived sufficient conditions for a merger to increase total surplus. Suppose that firms in set I contemplate merging. Let q_j denote firm j's output and Q, the industry output. Under the presumption that the proposed merger is profitable for the merging firms and in the absence of firms' endogenous investment choices, a sufficient condition for a merger to increase aggregate surplus is given as (in Whinston et al. (2008)):

$$M_I \;<\; -\sum_{j \notin I} M_j \left(\frac{dq_j}{dQ}\right) \tag{2}$$

where M_j is firm j's premerger market share, M_I is the collective market share of the firms in set I, and $\frac{dq_j}{dQ}$ is the differential change in non-merging firm j's output when the industry output changes marginally. Using the data in the industry, the condition is violated: the value of the left hand side of the equation is 0.4538, but the right hand side is 0.3845[5]. It implies that there is no guarantee that the merger increased total welfare.

Although convenient, these simple methods ignore firms' endogenous investment decisions. The firms make not only output choices (reflected in HHI) but also investment choices. The merger might change firms' incentives for investment and consequently change the welfare consequence of the merger[6].

As described in Section 2.1, firms' investment is the nonnegligible feature in the Japanese steel industry. Since investment is a main strategic choice variable in the steel industry, careful analysis of the steel merger requires a more complete model that accounts for a dynamic environment in which firms make intertemporal decisions on investment.

2.2.2 Merger Remedies

In the discussion above, we evaluated the merger itself. In the evaluation, there are two alternatives: complete approval or complete rejection of the mergers. Nevertheless, in practice, mergers are sometimes approved, given the implementation of merger remedies. Many of the merger remedies are the redistribution of capitals, which is not directly reflected in HHIs or firms' output choices. Even though there is no simple method to assess merger remedies, we can think of the evaluation of merger remedies based on the HHIs using the firms' capital shares as a preliminary analysis. The value of HHI index based on the capital market shares absent the merger remedies is 2962, and the value after the merger remedies based on the capital market shares is 2959. Since the value based on the capital market shares absent the merger is 1977, the merger remedies actually implemented in 1970 seems to be too small to mitigate anticompetitive concerns, though the value of the HHI slightly gets lower due to the policy. As we discuss in Section 6.2 based on the structural model, this is true: the actual remedies were not sufficient to offset the loss of consumer welfare.

[5]To compute the values using the methods of Nocke and Whinston (2022) and Farrell and Shapiro (1990), we need the information on the elasticity of the demand function, and we specified the demand function described in the next section and used the estimated parameters.

[6]Regarding the method of Nocke and Whinston (2022), it is implicitly assumed that non-merged firms' marginal costs would not change before and after the merger. Nevertheless, under the existence of firms' endogenous investment decisions, non-merged firms' marginal costs would change in the long run, and the application of the method might lead to a different conclusion.

Note that capital shares do not have a direct connection with output shares in general, and the relationship between the analysis above and economic models is not necessarily clear. Also, the effect of merger remedies may not persist over time, when we allow for the firms' endogenous investment decisions. Competitive authorities can change the initial distribution of the firms' capital by implementing merger remedies, but firms can flexibly adjust capital levels in the long run. To deal with these issues and derive more realistic evaluations of the policy based on the economic discussion, we need to develop a structural model explicitly specifying capital and firms' endogenous investment decisions, such as the one we present in the next section.

3 Model

This section describes a model used to explain the Japanese steel market. We begin the section by providing an overview of the estimation model.

3.1 Setup

Our empirical goal is to evaluate the welfare effects of the 1970 merger and its remedies between Yawata and Fuji by explicitly accounting for the dynamics resulting from investment behavior. The merger may have lessened competition in the steel market and simultaneously yielded efficiency gains in production. To assess this tradeoff, which was originally identified by Williamson (1968), it is necessary to construct a theoretical model that captures the salient features of the Japanese steel industry in the postwar era.

Time is discrete, and at the beginning of each period, firms produce steel using blast furnaces considering their marginal cost and capital stock. As described in Section 2, the Japanese steel market is characterized by active investment in the capital; moreover, only a handful of major firms operated in the market under minimal international competitive pressure. Capital investment improves production efficiency in future periods, whereas an oligopolistic market structure generates concerns for strategic behavior. Since these market features contain essential implications for our assessment of the 1970 merger, we build a dynamic model of firm behavior that allows for strategic interactions between firms. Our data set, which comprises annual data, covers the period from 1960 to 1990. In our application, the number of firms, N, takes the value of 6 when t is before 1970 and 5 in the post-merger period. The industry state at each period is summarized by a vector consisting of the commonly observed variables, s_t. This state vector includes the variables affecting the demand and production cost, as discussed later in this section.

At the beginning of period t, and given the state s_t, firm j decides the level of investment, i_{jt} for $j = 1, \cdots, J$. In the model, investment enhances the efficiency of steel production through capital accumulation. Note we cannot disaggregate firm investment into furnace-level, and thus i_{jt} should be considered as net value. To align with the feature of the industry, we assume that the installation or scraping of steel furnaces was often delayed by a year. Thus, the transition of the capital stock is characterized by $k_{jt+1} = k_{jt} + i_{jt}$.

The crude steel product is considered to be homogeneous, and firms simultaneously choose their quantities for each period. Let q_{jt} denote firm j's quantity of crude steel chosen in period t, and Q_t^B denote the sum of q_{jt} over j. We assume that current demand or current output do not affect future sequences of shocks nor outputs. Therefore, the equilibrium quantities are obtained by use of a static decision-making problem. Further detail is discussed in Section 3.3.

As discussed in Section 2, since no firms entered or exited the market during the study period, the number of active firms is assumed to be exogenous by J_t.

3.2 Demand

Even though our main focus is on the investment competition among firms using blast furnace, we cannot miss the other type of steel existed in the market: steel produced in electric furnace. The steel produced in blast furnace and steel produced in electric furnace are somewhat substitutable.

We assume that a representative consumer maximizes his or her CES utility function U_t with respect to the consumptions of two types of steel given their income Y_t:

$$\max_{Q_t^B, Q_t^E} U_t = \left[\left(\xi_t Q_t^B\right)^{\frac{\sigma-1}{\sigma}} + \left((1-\xi_t)\, Q_t^E\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

$$s.t.\ P_t^B Q_t^B + P_t^E Q_t^E = Y_t$$

where P_t^B and P_t^E denotes the prices of steel produced in blast furnaces and electric furnaces. Q_t^B and Q_t^E denote the consumption of steel produced in blast furnaces and electric furnaces. $\xi_t \in (0,1)$ denotes the weight of the steel produced in blast furnaces, and we allow ξ_t to change over time.

Then, optimal amount of consumption Q_t^B and Q_t^E given the prices P_t^B and P_t^E are given by:

$$Q_t^B = \frac{Y_t}{P_t^B} \frac{\xi_t \left(P_t^B\right)^{1-\sigma}}{\xi_t \left(P_t^B\right)^{1-\sigma} + (1-\xi_t)^{\sigma} \left(P_t^E\right)^{1-\sigma}} \tag{3}$$

$$Q_t^E = \frac{Y_t}{P_t^E} \frac{(1-\xi_t)^{1-\sigma} \left(P_t^E\right)^{1-\sigma}}{\xi_t \left(P_t^B\right)^{1-\sigma} + (1-\xi_t)^{\sigma} \left(P_t^E\right)^{1-\sigma}} \tag{4}$$

We assume that the behaviors of the producers using electric furnaces are exogenous. Namely, we presume that P_t^E are exogenously given. Further, let $\Theta^d = \sigma$.

3.3 Output Choice

We begin with the model of steel production technology. The availability of firm-level factor input data is limited; therefore, we build a cost function that describes the steel-making process. We assume that an increase in the firm's capital reduces the marginal cost of production. This is a reasonable assumption because the firm's capital investments mostly take the form of augmenting new steel making furnaces, which utilize the latest cost-reducing technologies. Thus, it is likely that an increase in k_{jt} will improve productive efficiency. Since the firm's investment is capitalized at the end of the period, we model firm j's marginal cost at time t, mc_{jt}, as the following form:

$$\ln(mc_{jt}) = -\gamma \ln(k_{jt}) + c_t + c_j \tag{5}$$

Eq.(5) assumes no spillovers in that the benefits of making investments are fully appropriated within the firm. The characteristics of steel production mentioned above suggest that firm j's cost-reducing technology is not transferable to other firms because the technology is physically utilized in furnaces owned by the firm itself. Let $\Theta^c \equiv (\gamma, c_t, c_j)$ be the parameters to be estimated.

We allow marginal cost to have firm- and time-specific components, denoted by c_j and c_t, respectively. With this specification, we can introduce asymmetry between firms.

Since it is difficult to find accurate cost data to directly analyze Eq.(5), we estimate the price-cost margins by building a competition model and thereby obtain the cost parameters. In particular, we construct a steel maker's profit maximization problem and solve the first-order condition. We establish the following supply-side model. Suppose that firm j competes and chooses its output at time t in the domestic market for crude steel. In each period t, firms face the domestic demand function $Q_t^B(P_t^B; P_t^E, \xi_t, I_t, \Theta^d)$ in which Θ^d is a set of demand parameters to be estimated. The other variables have been defined in the previous subsection. In this paper, we treat the amount of export as exogenously given because exported steel is reasonably assumed to be competitively supplied in the world market, as discussed in Section 2.

In each period, after the choice of investment, firm j observes the shocks $\eta_t \equiv (\xi_t, P_t^E, Y_t)$, and simultaneously chooses the output quantity q_{jt} to maximize the following per-period payoff:

$$\left(P_t^B - mc_{jt}\right) q_{jt} \tag{6}$$

Under the assumption made in Section 3.1, steel output and price are determined in the static equilibrium conditional on the current state $s_t \equiv (k_{jt_j})$. Hence, the maximized per-period payoff for firm j is a function of the current state vector and denoted by $\pi_{jt}(s_t, \eta_t; \Theta)$, where $\Theta \equiv (\Theta^d, \Theta^c)$. The first-order condition derived from firm j's static profit maximization under Cournot competition takes the familiar form of the Lerner index, namely,

$$\frac{P_t^B - mc_{jt}}{P_t^B} = \frac{1}{|\epsilon_t|}\frac{q_{jt}}{Q_t^B} \tag{7}$$

where ϵ_t denotes the price elasticity of demand. We do not consider the possibility of capacity constraint in Eq.(7) because it is known to be difficult to define the maximum available production capacity in this industry. Note that the unit of measurement differs between q_{jt} and k_{jt}; the former is in terms of physical tonnage, while the latter is in terms of monetary value (at 1960 prices). Since a small-sized furnace with advanced technology was often more expensive than a conventional large-scale furnace, it is nearly impossible to determine the link between the two variables to infer the industry's utilization rate.

Using the demand estimates obtained in Section 5 and the data, we can derive mc_{jt} from the first-order condition in Eq.(7). In the next subsection, we construct the discounted future payoffs and introduce a set of parameters associated with firm investment decisions.

3.4 Firm's Investment

In this subsection, we describe the model of investment choice, or the decision that is made before the output choice. At the beginning of period t, and given the state s_t, each firm makes its decision on the investment amount. The investment decision is inherently dynamic because according to Eq.(5), a firm receives benefit from the investment in the future periods. In the investment choice, firm j is assumed to maximize the following expected future profit:

$$E_t \sum_{\tau=0}^{\infty} \beta^\tau \left[\pi_{jt+\tau}(k_{t+\tau}, \eta_{t+\tau}; \Theta) - \phi(k_{jt+\tau}, i_{jt+\tau}; \theta) | s_t\right] \tag{8}$$

where $k_t = \{k_{jt}\}_j$. The expectation is taken over other firms' investment choices in the current and future periods as well as over the current and future values of all the state variables. The maximized per-period payoff, $\pi_{jt}(k_t, \eta_t; \Theta)$, was defined in the previous subsection. Each firm discounts its future profits according to a common discount factor with a common information set. In the estimation, we set the discount factor equal to 0.9. Recall that we do not consider the issue of firm entry and exit in this study.

Investments incur costs. We assume that the investment cost $\phi(k_{jt}, i_{jt}; \theta)$ is in the following form:

$$\phi(k_{jt}, i_{jt}; \theta) \quad = \quad \theta_k i_{jt} + \theta_a (i_{jt}/k_{jt})^2 k_{jt} \tag{9}$$

where θ is the set of parameters, $\theta_k and \theta_a$, be estimated. The first term captures the direct cost of investment. The second term represents the quadratic adjustment cost, which has been intensively discussed in the macro literature (e.g. Hayashi (1982) and Cooper and Haltiwanger (2006))[7]. The quadratic term captures the investment cost relative to its capital stock (firm size). For smaller firms, a large investment is so costly, because of the relocation cost of workers and facilities, or the financial market imperfections (It is costly for smaller firms to borrow much money for investment).

As in Ericson and Pakes (1995), we consider a pure Markov perfect Nash equilibrium (hereafter MPNE). We restrict our attention to pure strategies and do not consider mixed strategies. The MPNE in this paper consists of a set of best-response strategies that govern investment decisions. An equilibrium is assumed to exist, and we actually confirm the existence by numerically solving the model. In our modeling assumption, a Markov strategy for firm j describes the firm's behavior at time t as a function of state variables $s_t = (k_t, \eta_t)$. We assume that after time T the environment surrounding the steel industry is constant, and firms do not make additional investments. This assumption reflects the observation in Figure 2, which illustrates steel firms' investment activities slowed down as time passed. We assume that year 1990 is the terminal period. Besides, we presume that firms form perfect foresight.

The value function V_{jt} is the discounted sum of firm j's profits at the beginning of time t. It can be decomposed into two components: the per-period payoff and the continuation value. Each firm uses the value function to determine its investment amount in the intertemporal optimization condition, where today's incremental cost incurred by making a unit of investment equals the present value of the benefit of cost reduction through capital accumulation. Firms base their investment strategy on the current state variables, and let $I_t \equiv \{I_{jt}\} \equiv \{i^*_{jt+\tau}(s_{t+\tau})\}_{j,\tau\geq 0}$ be a set of investment strategies for all firms after time t. Under the MPNE, we can rewrite the value function in Eq. (8) in the following recursive form:

$$V_{jt}(s_t; I_t) \quad = \quad \pi_{jt}(k_t, \eta_t; \Theta) - \phi(k_{jt}, i^*_{jt}(s_t); \theta) + \beta E_t V_{jt+1}\left(s_{t+1}(s_t, i^*_{jt}(s_t), i^*_{-jt}(s_t)) I_{t+1}\right) \tag{10}$$

where $s_t = (k_t, \eta_t)$. Since we assume that the market conditions will stay constant and firms do not invest anymore after time T, $V_{jT}(s_T)$ is in the following form:

$$V_{jT}(s_T) \quad = \quad \frac{1}{1-\beta} \pi_{jT}(k_T, \eta_T; \Theta)$$

[7]Even though it is possible to introduce other quadratic term i^2_{jt}, the introduction contribute little to the fit of the model and the results of counterfactual simulation. The estimation results are shown in Appendix E.

4 Estimation

4.1 Demand Function

By Eqs.(3) and (4), market shares of the steel produced in blast furnaces and electric furnaces measured by output is:

$$\begin{aligned} S_t^B &= \frac{\xi_t \left(P_t^B\right)^{-\sigma}}{\xi_t \left(P_t^B\right)^{1-\sigma} + (1-\xi_t)^{\sigma} \left(P_t^E\right)^{1-\sigma}} \\ S_t^E &= \frac{(1-\xi_t)^{\sigma} \left(P_t^E\right)^{-\sigma}}{\xi_t \left(P_t^B\right)^{1-\sigma} + (1-\xi_t)^{\sigma} \left(P_t^E\right)^{1-\sigma}} \end{aligned}$$

Then, the following equation holds:

$$\log S_t^B - \log S_t^E = -\sigma \log\left(\frac{P_t^B}{P_t^E}\right) + \sigma \log\left(\frac{\xi_t}{1-\xi_t}\right)$$

Let $\sigma \log\left(\frac{\xi_t}{1-\xi_t}\right) = c + u_t$ and $\widetilde{P}_t \equiv \frac{P_t^B}{P_t^E}$. $\widetilde{P}_t$ represents the price of steel produced in blast furnaces relative to the steel produced in electric furnaces. Then, we obtain:

$$\log S_t^B - \log S_t^E = -\sigma \log\left(\widetilde{P}_t\right) + c + u_t$$

To estimate parameter σ, we impose the following orthogonality condition:

$$E[u_t | z_t] = 0$$

Here, z_t denotes the instrumental variables which are correlated with the relative price $\widetilde{P}_t$, but orthogonal to the demand shock u_t. We use the prices of the major factor inputs used in steel making, iron ore, and heavy oil (both in terms of logarithms) as instrumental variables. We estimate parameters σ and c by the Generalized Method of Moments (GMM).

4.2 Marginal Cost

Using the results of demand estimates and the first order condition in Eq. (7), we calculate the marginal cost of steel production and estimate the following equation by OLS:

$$\ln(mc_{jt}) = -\gamma \ln(k_{jt}) + c_t + c_j + e_t \quad (11)$$

Since Eq. (5) does not include any stochastic terms, we add an error term e_t.

4.3 Investment Cost

Finally, having estimated the parameters of demand and marginal cost functions, we estimate the investment cost parameters θ using a full-solution approach.[8] As in Cooper and Haltiwanger (2006), Goettler and Gordon (2011), and Chen and Xu (2022), we apply Simulated Method of Moment (SMM), that minimizes the distance between the moments derived from the model and the ones based on the observed data. Let $\widehat{\Theta}$ be the estimated values of marginal cost and demand parameters $\Theta = (\Theta^d, \Theta^c)$. The procedures of the estimation are summarized as follows:

1. Set the initial values of θ.
2. Given $(\theta, \widehat{\Theta})$, solve the equilibrium and derive the policy function (optimal investment choice as a function of state variables) of each firm in each period (see Appendix A for the solution method).
3. Compute the firms' optimal investment decisions in each period, and compute the following objective function:

$$\arg\min_{\theta} \left[m_{model}(\theta, \widehat{\Theta}) - m_{data}\right]' W \left[m_{model}(\theta, \widehat{\Theta}) - m_{data}\right] \tag{12}$$

   where m_{data} is the moments from the observed data, and $m_{model}(\theta, \Theta)$ is the moments based on the model given parameter values (θ, Θ). W denotes a positive definite efficient weight matrix. The weight matrix is the inverse of the covariance matrix of actual data's moments, which we use the bootstrap procedure to estimate.

4. Obtain θ that minimizes the objective function (12)

We use three moments to estimate the parameters: average ratio of investment relative to current capital stock, average value of the merged firm's investment, and average value of the non-merged firms' investment.

Since the model is nonlinear, all the parameters influence all the moments. Nevertheless, primary connection is clear. Parameter θ_k (direct investment cost) is mainly identified by the variation in the values of investment in each period On the other hand, parameter θ_a is mainly identified by the ratio of investment to current capital stock.

5 Estimation Results

This section applies the estimation models described in the previous section to the annual frequency data set from 1960 to 1990. We chose to start the sample in 1960, when Kobe - the smallest company in the data set - had a fully operational blast furnace and became an integrated steel maker. Including Kobe in this study helps us expand the data set.

We first discuss the estimation of the demand and marginal cost functions. We then proceed to the estimation of the dynamic parameters associated with the investment cost. The summary

[8] Several methods have been proposed to estimate the parameters of dynamic games. The most prevalent method is the two-step method, such as Bajari et al. (2007). While the method is simple, many observations are often needed to precisely approximate a policy function, and it may suffer from finite sample biases. Unfortunately, the industry under focus does not have many markets, and the number of observations is not sufficiently large to employ the two-step method

statistics of the important variables used in the estimation appear in Table 1, and the data sources are presented in Appendix D. Section 6 uses the estimates reported in this section to assess the economic consequences of the 1970 merger.

5.1 Demand Estimates

Table 3 presents demand estimates. The first column is based on the ordinary least squared (OLS) method, and the second is based on the GMM method using instrumental variables. The upper portion of Table 3 reports the estimates of the regression coefficients. The implied demand elasticity with respect to price is also calculated in the lower portion of the table. The elasticity is relatively inelastic: its value is roughly 1.05. The comparison between the demand elasticities obtained from the OLS and GMM methods indicates the successful elimination of endogeneity from the positive correlation between steel output and demand shock: the mean value of the implied demand elasticity obtained from the GMM estimates is approximately 20% larger than those obtained from the OLS estimates.

It is known that the GMM method can produce severely biased estimates if the instruments are weak. We thus check the explanatory power of the instruments, conditional on the included exogenous variables in the first stage estimation. We obtain an F-statistic for the endogenous variables discussed above. Table 3 reports the value of the F-statistics, which equals 32. All the instruments used in this paper are not weak at the 99 % confidence level.

Table 3: Demand Estimates

	OLS	*GMM*
σ	0.580**	1.116***
	(0.266)	(0.248)
c	0.470*	0.084
	(0.246)	(0.226)
1st stage F-stat.	-	32.057
Elasticity w.r.t. price	0.8	1.055

Notes:
Standard Error of the estimate is shown in the parenthesis.
The number of samples used in the estimation is 31.
* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.
Elasticity w.r.t. price is the average value of annual demand elasticities of steel produced in blast furnaces with respect to price from 1960 to 1990.

5.2 Marginal cost Estimates

Using the demand estimates obtained in Table 3 and the first-order condition in Eq. (7), we calculate the marginal cost of steel production and estimate Eq. (11). The estimation results are presented in Table 4. Specification (1) only includes the year-specific components, yet specification (2) includes both the firm and year-specific components.

The value of γ represents the elasticity of marginal cost with respect to capital stock, it is estimated to be negative in both specifications. Further, the estimated firm-specific components in specification (2) indicate that the merger substantially improves production efficiency for the merging party. In comparison with the average estimates of the firm-specific effects in Yawata and Fuji, the table indicates that the merger enhanced the efficiency of Nippon Steel by roughly 20%.

In the remainder of this paper, we use specification (2) to more flexibly allow steel firms' asymmetry[9].

Table 4: Marginal cost Estimates

	(1)	(2)
γ	−0.241***	−0.109***
	(0.010)	(0.010)
Firm-specific component (c_j):		
NSC		0
Yawata		0.18
Fuji		0.21
NKK		0.25
Kawasaki		0.25
Sumitomo		0.24
Kobe		0.25
Year dummies	Yes	Yes
Firm dummies	No	Yes
Adjusted R^2	1.00	1.00

Notes:
Standard Error of the estimate is shown in the parenthesis.
The number of samples used in the estimation is 165.
*** Significant at the 1 percent level.

5.3 Investment Cost Estimates

The estimates for the dynamic parameters of investment cost are presented in Table 5, to avoid the problem associated with the firms' expectations on the realization of the 1970 merger before the merger period. We estimate them through the Simulated Method of Moment (SMM). Table 5 shows the coefficients of linear and quadratic terms (θ_k and θ_a). The estimation results imply that the existence of adjustment costs is not negligible. In the absence of adjustment costs in investment, firms optimize their capital stocks by equating the marginal revenues of capital and the rental price of the capital. Hence, even if firms have small capital stock in the current period, they can catch up with large capital stock firms in the next period by active investment. In contrast, under the existence of adjustment costs, a large investment is too costly for small capital stock firms and have weak incentives to catch up with larger firms. As shown in Figures 1 and 2, the order of the Japanese steel firms measured by capital stock / output is mostly stable over time

[9]Regarding the specification of marginal cost function, in principle we can consider the case where the marginal cost itself depends on the firm's output. Nevertheless, the estimated coefficient concerning the log of output $\log(q_{jt})$ was negative and insignificant, and we proceed with the baseline estimated parameters. The estimation results are available upon request.

in 1970-1990. Consequently, we can guess that this would be the reflection of the high adjustment costs in investment.

Table 5: Investment cost Estimates

	Est.
θ_k (i_{jt})	0.363
	(0.017)
θ_a (i_{jt}^2/k_{jt})	0.885
	(0.147)

Notes:
Standard Error of the estimate is shown inside parenthesis.
The number of samples used in the estimation is 100.

5.4 Model Prediction

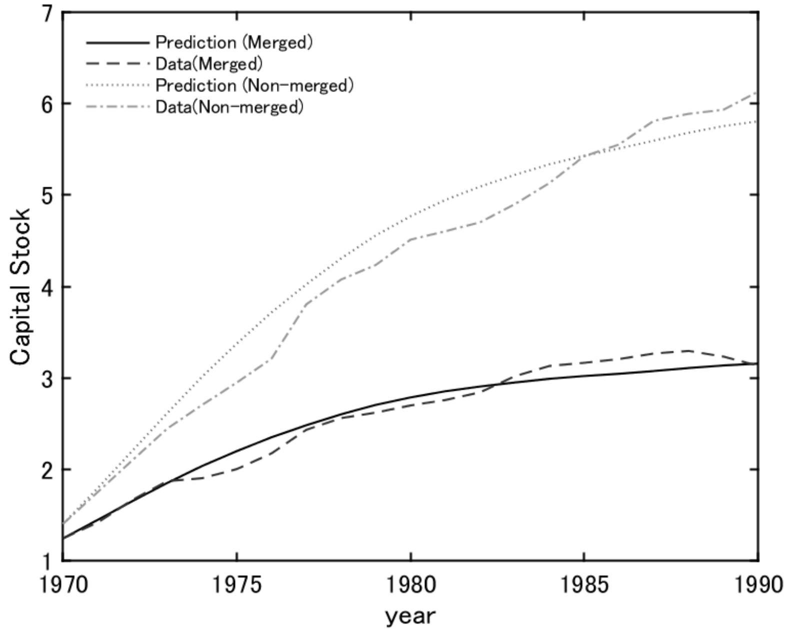
To understand how the model fits the data, we compare the actual and predicted industry prices, outputs, market shares, and capital stocks over the study period. We compare the averages and standard errors for each of the periods. Table 6 shows the results of this comparison. The LHS of the table presents the prediction based on the base model introduced in this section, while the RHS presents the actual data, some variables of which were already introduced in Table 1. To save space, we list the market shares and capital stock of the merged and non-merged parties. Further, Figure 3 shows the predicted path of capital stock over time compared to the actual data. The results show that the model explains the data well. All the variables in the table are predicted fairly accurately.

Table 6: Model Prediction

	Prediction	Actual
Industry Output	74.89	74.41
(in million ton)	(8.64)	(8.40)
Price	41.47	41.72
(in JPY per thousand ton)	(2.35)	(2.49)
Merged party Market share	41.38	41.72
(%)	(0.81)	(1.28)
Merged party Capital stock	2.55	2.55
(in billion JPY)	(0.59)	(0.65)
Non-merged party Market share	58.62	58.28
(%)	(0.81)	(1.28)
Non-merged party Capital stock	4.29	4.16
(in billion JPY)	(1.39)	(1.48)

Notes:
Standard Error of the estimate is shown inside parenthesis.
Merging party is Nippon Steel in 1970-1990. Non-merging party is the sum of Nihon Kokan, Kawasaki, Sumitomo, and Kobe. JPY is in terms of the 1960 price.

Figure 3: Predicted path of capital stock



Notes:
The values of capital stocks are represented in billion JPY in 1960.
In 1960, 1 USD is equivalent to 360 JPY.

6 Economic Consequence of the 1970 Merger

This section comprises three subsections and assesses the economic consequences of the horizontal merger between Yawata and Fuji in 1970. Based on the model and the estimates reported in the previous section, Section 6.1 evaluates the welfare tradeoff associated with the horizontal merger by comparing the counterfactual situation, in which no merger took place.

The section also assesses the effectiveness of the structural merger remedies endorsed in the negotiation process between the J-FTC and the companies involved in the merger. The merger was approved under the condition that the merged party would transfer its capital equipment to two smaller companies, Kobe and Nihon Kokan. Section 6.2 performs simulation exercises to assess the welfare consequence of this divestiture. In Section 6.3, we also consider the optimal merger remedies, where the merger remedies aim at maximizing consumer welfare or total surplus.

6.1 Economic Impacts of Merger

This subsection intends to assess the economic impacts of the merger. To conduct the assessment, we compare the merger and no-merger outcomes. In this subsection, we consider the consequences without merger remedies as the merger outcomes. The no-merger outcome is simulated by investigating what would have happened to the steel market had no merger occurred between Yawata and Fuji. Under this counterfactual scenario, we treat Yawata and Fuji as different business entities even in 1970-1990. Thus, we assume that Yawata and Fuji independently decided their outputs and investments based on their capital stocks. Thus, the number of firms in the market, N_t, is assumed to be six throughout the period in the counterfactual simulation procedure.

Figure 4 and Figure 5 show the effects of the merger on the economic variables and surpluses by year. Figure 4 shows the trend of steel prices and merged/non-merged firms' investment, and it indicates that the merger did increase steel prices, but by a small margin (0.8% on average. Regarding firms' investments, the merger had a large negative impact on the merged firm's investment. The merged firm lowered its investment, and its accumulated capital stock was 25% lower than the outcome absent the merger. In contrast, non-merged firms increased their investments, and the accumulated capital stocks are several percent larger than the case absent the merger in the 1980s. Regarding surpluses, Figure 5 shows that the loss of equivalent variation was relatively small. In contrast, the producer surplus largely increased, and In total, the total surplus was positive not only in the short run but also in the long run. It implies that the merger was socially desirable from the viewpoint of total surplus.

Figure 4: Effect of 1970 merger on economic variables (Comparison with the absence of Merger)

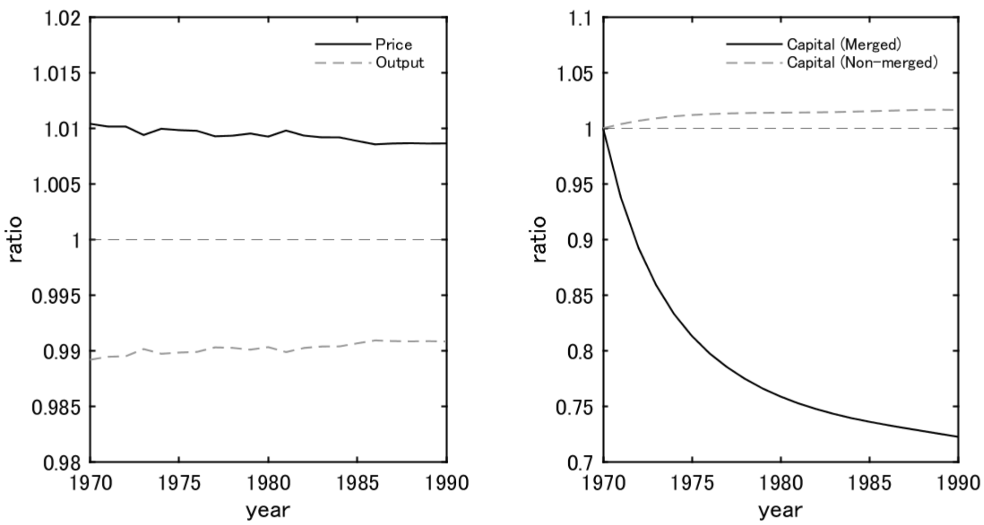
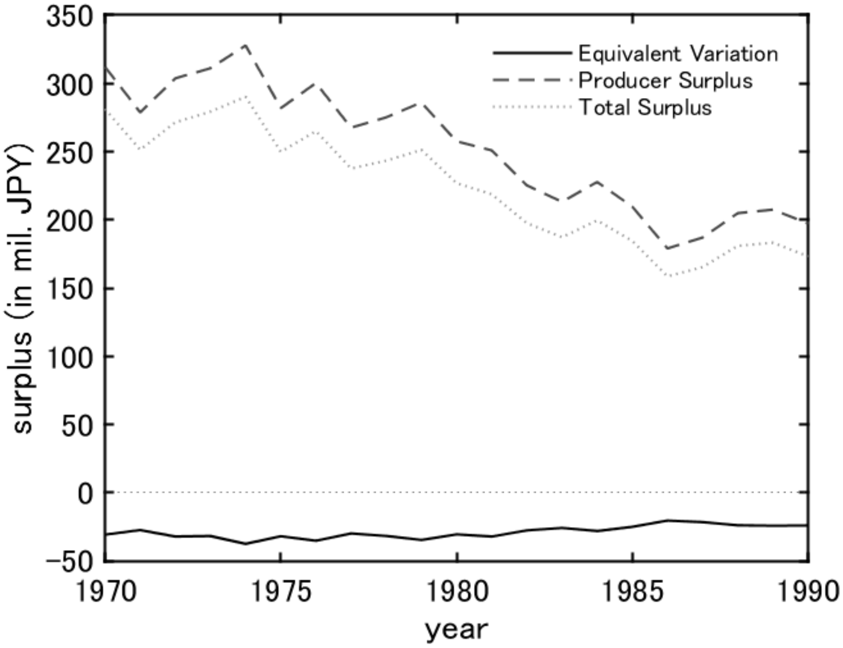


Figure 5: Impacts of Merger on Welfare (Comparison with the absence of Merger)



To see the background behind the results, we simulate the outcomes (economic variables and welfare) under the hypothetical cases where no efficiency gains in the merger exists, and the case where firms' investment decisions are exogenous, and compare the results. In the model, the merger affected the merged party's marginal cost through the merger-specific efficiency gains, where the firm-specific fixed effect term (c_j in Eq. (5)) gets lower. As discussed in Mermelstein et al. (2020), the merged party's marginal cost is also affected by the existence of scale economy, where larger capital stocks itself lowered merged party's marginal cost, as we can easily verify in Eq. (5). To isolate these effects, we also simulate the hypothetical case where there are no efficiency gains and the merged party's marginal cost remains the same level before the merger[10]. By comparing the cases with and without the efficiency gains, we can assess the contribution of the efficiency gains to the evaluation of the merger. Regarding the "exogenous investment", we presume that firms follow the equilibrium investment strategies absent the merger. Table 7 shows the results. The effects on economic variables (output, price, investment, merging party's profit; (1)-(9)) are measured as the ratios of corresponding economic outcomes under the merger relative to those in the absence of merger. The effects on surpluses((10)-(13)) are represented by million JPY.

[10]More specifically, we constructed the case without efficiency gains by searching for the value of c_{j_0} in Eq. (5) (merged firm's firm-specific fixed effect term in the marginal cost equation) such that $mc_{j_0,t=1970}(c_{j_0}) = \sum_{j \in j_1, j_2} w_{j,t=1970} mc_{j,t=1970}$. Here, j_0 denotes the merged firm and j_1, j_2 denote the merging firms. $w_{j,t=1970}$ denotes the capital share of firm j among the merging firms.

Table 7: The effect of the merger on economic variables and surplus

		No efficiency gain		With efficiency gain	
		(a)Exo. inv.	(b)Endo. inv.	(c)Exo. inv.	(d)Endo. inv.
(1)	Marginal cost (Merged)	0.993	1.053	0.762	0.783
(2)	Marginal cost (Non-merged)	1	0.978	1	0.997
(3)	Output (Merged)	0.571	0.466	0.931	0.896
(4)	Output (Non-merged)	1.267	1.368	1.044	1.07
(5)	Price	1.052	1.045	1.008	1.009
(6)	Investment (Merged)	1	0.229	1	0.61
(7)	Investment (Non-merged)	1	1.379	1	1.022
(8)	Merged firm's profit	0.607	0.614	2.011	2.026
(9)	Non-merged firms' profit	2.59	2.537	1.228	1.274
(10)	EV	-1590	-1402.3	-258.2	-299
(11)	PS (excluding Inv. cost)	787.2	567.9	2591	2343.5
(12)	PS (including Inv. cost)	789.9	750.4	2593.7	2678.8
(13)	TS (including Inv. cost)	-800.1	-651.9	2335.5	2379.7

Notes. The values in (1)-(9) represent the ratios of each economic variable under the merger compared to the absence of the merger. The values in (1)-(7) are represented as the average of the values in 1970-1990 and the values in (8) and (9) represent the merged / non-merged firms' discounted sum of profit. "EV", "PS", "TS" denote equivalent variation, producer surplus, and total surplus, and the values in (10)-(13) represent the discounted sum of the surpluses due to the merger.

First, the effect of the merger on consumer welfare (equivalent variation) was negative in all scenarios ((10)), because of the higher equilibrium prices compared to the case without the merger ((5)). In contrast, the effect on producer surplus was positive ((11),(12)). The effect on total surplus was positive under the existence of efficiency gains ((13-c),(13-d)), but was negative absent efficiency gains under the endogenous investment case ((13-b)). Consequently, we can argue that the large efficiency gains led to the positive assessment of the merger. Note that absent the efficiency gains, the merger was not necessarily profitable for the merged party. (8) shows the merged firm's profit compared to the sum of merging firms' profit absent the merger. The merger decreased the merged party's profit under the exogenous investment case ((8-a)), and the merger hardly affected the merged party's profit under the endogenous investment case ((8-b)). We can guess that the large efficiency gain would be the motivation to merge for the merging party.

Regarding firms' investment, the merged firm invested less compared to the no-merger scenario ((6-b,d)) [11]. Besides, by comparing the merged party's investment decisions without and with efficiency gains, we can also see that the merged firm invested more due to the efficiency gain((6-b,d)). One large factor affecting the merged party's investment is the effect of the merger on the output (See Appendix B for details). As discussed in Jullien and Lefouili (2018), lower post-merger output reduces incentives to invest in margin-enhancing innovation (margin expansion

[11] Mermelstein et al. (2020) argued that the typical models applying Ericson and Pakes (1995) type models are not "merger neutral" regarding merged firm's investment opportunities. By the construction of the model, it is implicitly assumed in the models that the merging party can add up to two units of capitals if they do not merge, but they can add only up to one unit of capital if they merger. Even though our model builds on the literature applying Ericson and Pakes (1995) type models, our model is merger neutral, because there are no constraints on the size of the capitals to be added.

effect). In the current setting, the merger absent efficiency gain lowered the merged firm's output (3-a), and consequently discouraged the firm's investment. In contrast, the efficiency gains after the merger raised the output level of the merged party ((3-a,c)), and encouraged more investment.

In contrast, the non-merged firms invested more in comparison with the absence of the merger, as we can verify in Table 7. This is mainly due to the non-merged party's larger output after the merger. Note that the merged party's efficiency gains negatively affected the non-merged party's incentive.

Finally, we consider the effect of the firms' endogenous investment behavior for merger assessment. Here, we focus on the case without efficiency gains[12]. By comparing (6-a) and (6-b), we can observe that the merged firm invested roughly 75% lower compared to the case without firms' endogenous investment behaviors. In contrast, non-merged firms invested more compared to the case absent endogenous investment ((7-a,b)), and the market environment got less concentrated ((3-a,b),(4-a,b)). Consequently, the equilibrium price got lower compared to the exogenous investment case ((5-a,b)), and the loss of equivalent variation got lower, even though the equivalent variation was negative in both cases ((10-a,10-b)). In contrast, producer surplus got lower due to the lower equilibrium price ((11,12-a,b)).

Overall, we obtained three main findings. First, we can attribute the positive evaluation of the 1970 merger largely to the efficiency gains due to the merger. Second, the merger lowered the merged party's investment, but it was determined by the tradeoffs between the lower incentives to invest due to the lower output after the merger without efficiency gains, and the higher incentives to invest due to the larger output due to the efficiency gains. Third, even though the merged firm lowered its investment, nonmerged firm invested more, and the loss of consumer welfare decreased due to the introduction of firms' endogenous investment behavior absent the efficiency gains.

In Appendix B, we show the detailed decomposition of the effect of the merger on firms' investment incentives. Since the discussion on the decomposition is lengthy, we summarize the results below. In the discussion, using firms' first-order conditions regarding their investment choices, we first decompose firms' investment incentives into five terms: Production cost reduction, Strategic use of investment to decrease other firms' investment, Investment cost reduction due to the scale economy of investment, Investment cost reduction due to lower future investment, and Preemption. Then, we show how the merger affected these incentives. The quantitative results show that the following four factors largely affected the merged party's investment:

1. Range of the product affected by a unit investment after the merger (+)

2. Efficiency gains(+)

3. Scale economy of production cost (−)

4. Lower merged party's output after the merger (−)

Regarding the first factor, merging firms can only affect the production cost of their own individual firms' output before the merger. In contrast, after the merger, they can affect the production cost of the sum of their output, which is larger than the case before the merger, and it encourages the merged party's investment. Even though this effect does not appear in the model of differentiated products, we show that the effect is quantitatively nonnegligible.

In terms of the third factor, the merger itself increases a firm's capital, lowering its marginal cost, as shown in Eq. (5). However, as the size of capital gets large, marginal reduction in marginal

[12] We can analogously interpret the results under the case with efficiency gains.

cost due to a unit investment decreases. Consequently, mergers discourage firms' investment. The results show that it largely affected the merged party's investment decisions.

Because of the contradicting directions of the effect of the merger on the merged party's investment, the overall effect was determined by the tradeoffs. The quantitative results show that the merger discouraged merged party's invesmetment.

In contrast, regarding the non-merged party's investment incentive, larger output due to the merger largely encouraged its investment, but the efficiency gains of the merged party discouraged the non-merged party's investment. Overall, the merger slightly increased the non-merged party's investment.

6.2 Evaluations of the 1970 merger remedies

Next, we consider the economic consequences of the merger remedies in the long run, using the example taken from the 1970 steel merger. The merger was approved under the condition that Yawata and Fuji would transfer 1.5% of their production facilities to Kobe and 0.3% to Nihon Kokan. First, Figure 6 and 7 show the impacts of 1970 merger remedies on economic variables from 1970 to 1990 in comparison with the absence of the remedies. Figure 6 is on the merging and non-merging firm's capital stocks and prices, and Figure 7 is on the surplus. Figure 6 shows that the merger remedies lowered steel price immediately after the 1970 remedies by roughly 0.25 %, but the effect gradually decreased over time. Half of the effect disappeared 10 years after the merger remedies, and approximately 80 % of the effect disappeared 20 years later. This is mainly due to the firm' investment behavior.

Figure 6: Effect of 1970 merger remedies on firms' capital stocks and prices

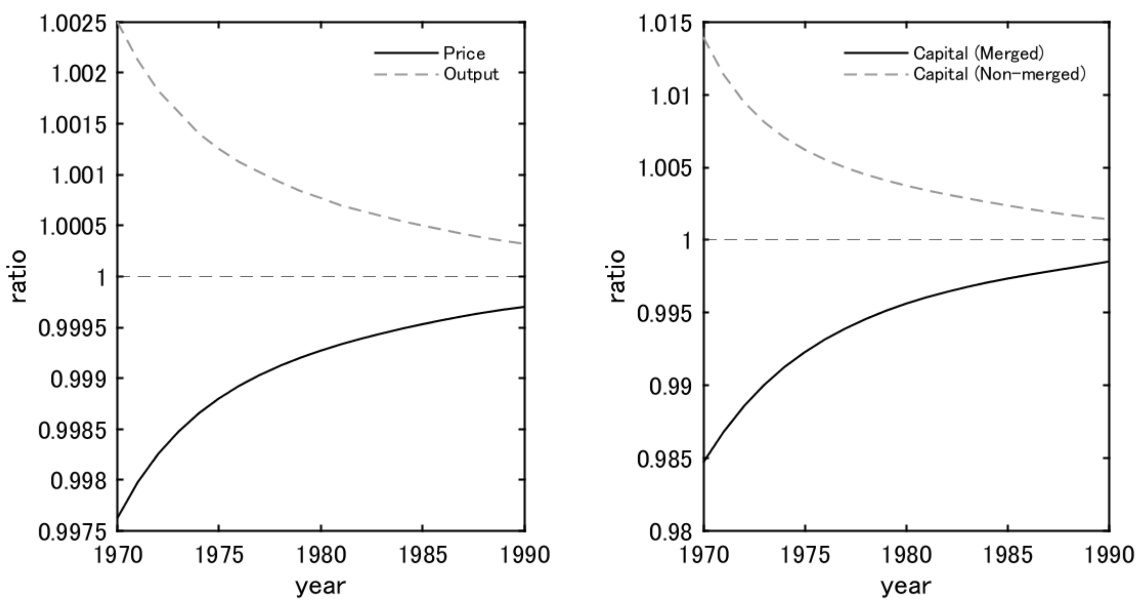
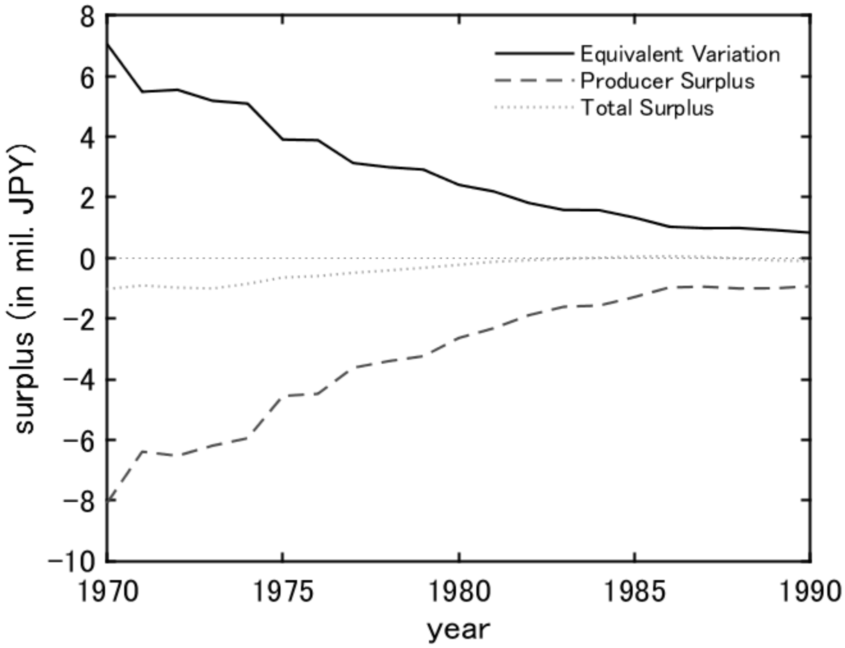


Figure 7: Effect of 1970 merger remedies on welfare



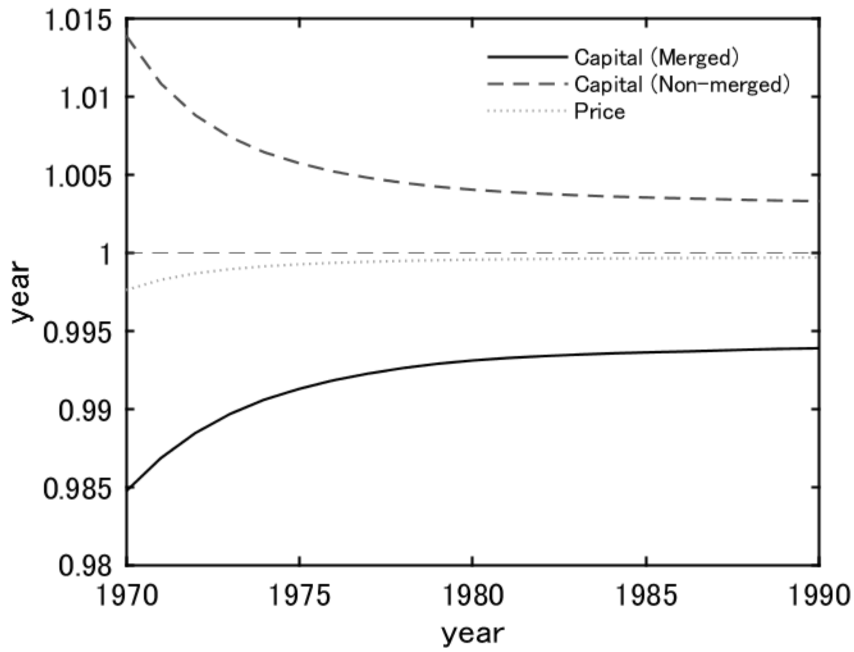
As the figure shows, the merged firm lost 1.5% of its assets following the merger remedies, but the firm invested more than in the case without merger remedies. Consequently, in the long run, the capital stock of the merging firm got close to the level without the merger remedies. Similarly, the non-merged firms obtained some assets from the merging firm following the merger remedies, but the firm invested less compared to the case without merger remedies. Consequently, in the long run, the capital stock of the merging firm got close to the level without the merger remedies.

Figure 7 shows that the merger remedies increased equivalent variation but decreased producer surplus. In total, the merger remedies decreased producer surplus. Nevertheless, the effects decreased over time: half of the increase in the equivalent variation disappeared ten years later. Note that the impacts of the remedies were so small compared to the impacts of the merger itself, as we can easily verify by comparing the values in Figure 5 and 7.

These figures show that the effects of the merger remedies diminish in the long run. Then, why did the phenomenon happen? We can attribute this to growing capital stocks and firms' endogenous investment behavior.

First, steel industries was characterized by firms' active investment behavior, and capital stocks grow over time, as depicted in Figure 2. Then, changes in the capital have smaller effect in the later periods, because capital stocks themselves get larger in the later periods. Figure 8 considers the case where firms' investments are exogenously given, and the figure shows that the effects of the remedies diminished especially in 1970s. In 1970s, the growth rates of the firms' capital stocks were large, and it led to the diminishing effect of merger remedies.

Figure 8: Effect of 1970 merger remedies on price and capital stocks (Under firms' exogenous investment)



Next, we discuss the effect of firms' endogenous investment behavior. The effects of firms' endogenous investment behavior on the persistency of the effect of merger remedies are represented in the difference between Figure 6 (Under firms' endogenous investment behavior) and Figure 8 (Under firms' exogenous investment behavior). As the figures show, the effect is not negligible, especially in the 1980s, in which the growth rate of capital stocks gets smaller. To consider the intuitions behind the phenomenon that the effects of the merger remedies diminish in the long run, consider the case where no adjustment cost in investment exists ($\theta_a = 0$). Then, we can easily show that (k_{jt}, k_{-jt}) satisfying the following equation is in the equilibrium[13]:

$$\underbrace{\beta \frac{\partial \pi_{jt}(k_{jt}, k_{-jt})}{\partial k_{jt}}}_{\text{Marginal revenue from investment}} = \underbrace{(1-\beta)\theta_k}_{\text{User cost of capital (rental price of capital)}} \tag{13}$$

[13]Firm j's first order condition with respect to its investment choice is:

$$\begin{aligned} 0 &= -\theta_k + \beta \frac{\partial k_{jt+1}}{\partial i_{jt}} \frac{\partial V_{jt+1}}{\partial k_{jt+1}} \\ &= -\theta_k + \beta \frac{\partial}{\partial k_{jt+1}} \left[\pi_{jt+1}(k_{jt+1}, k_{-jt+1}) - \theta_k(k_{jt+2} - k_{jt+1}) + \beta V_{jt+2}(k_{jt+2}, k_{-jt+2})\right] \end{aligned}$$

Since (k_{jt+2}, k_{-jt+2}) in the equilibrium satisfying Eq. (13) does not depend on the capital stocks in the previous period (k_{jt+1}, k_{-jt+1}), $\frac{\partial k_{jt+2}}{\partial k_{jt+1}} = \frac{\partial k_{-jt+2}}{\partial k_{jt+1}} = 0$ holds. Consequently, the first-order condition is equivalent to the following equation, and capital stocks satisfying Eq. (13) is in equilibrium.

$$0 = -\theta_k + \beta \frac{\partial \pi_{jt+1}(k_{jt+1}, k_{-jt+1})}{\partial k_{jt+1}} + \beta\theta_k$$

The equation implies that firms choose the levels of capital stocks to equate the marginal revenue of capital ($\beta\frac{\partial\pi_{jt}(k_{jt},k_{-jt})}{\partial k_{jt}}$) and unit capital cost $(1-\beta)\theta_k$. It implies that firms' future capital stocks k_{jt+1} do not depend on the current capital stocks k_{jt}, meaning that the effect of merger remedies at time t disappears at time $t+1$. Note that the strong results rely on the assumption of no adjustment cost in investment, and the effect of merger remedies at time t would not immediately disappear even at time $t+1$ under the existence of investment costs. Nevertheless, the model indicates that firms can flexibly adjust capital stocks even after merger remedies, and the effects may decrease in the long run.

6.3 Optimal merger remedies

In the previous subsection, we evaluated the merger remedies actually implemented in 1970. Nevertheless, there is no guarantee that the actual merger remedies were optimal. Next, we consider the optimal merger remedies. Since most of the merger remedies are closest to the consumer welfare standard (Whinston et al. (2008)), we first consider the optimal merger remedies under the consumer welfare standard in Section 6.3.1, which aims at offsetting the higher prices and the loss of consumer welfare from the merger discussed in Section 6.1. In Section 6.3.2, we consider the alternative scenario where competitive authorities implement merger remedies based on the total surplus standard, aiming to maximize the total surplus.

6.3.1 Consumer welfare standard

To design optimal merger remedies, we should solve the following optimization problem:

$$\begin{aligned} \max_{\{\Delta_j\}_j,D} \quad & CW(\{\Delta_j\}_j,D) \\ s.t. \quad & \sum_{j\in\mathcal{J}^{(nonmerged)}} \Delta_j = 1 \end{aligned}$$

Here, D denotes the total amount of divested assets by the merged firm, and Δ_j denotes the ratio of distributed assets to non-merged firm j. Note that $D\Delta_j$ represents the amount of divested assets to non-merged firm j. By definition, $\sum_{j\in\mathcal{J}^{(nonmerged)}}\Delta_j = 1$ must hold. $CW(\Delta,D)$ denotes the consumer welfare as a function of $\Delta\equiv(\{\Delta_j\}_{j\in\mathcal{J}^{(nonmerged)}})$ and D.

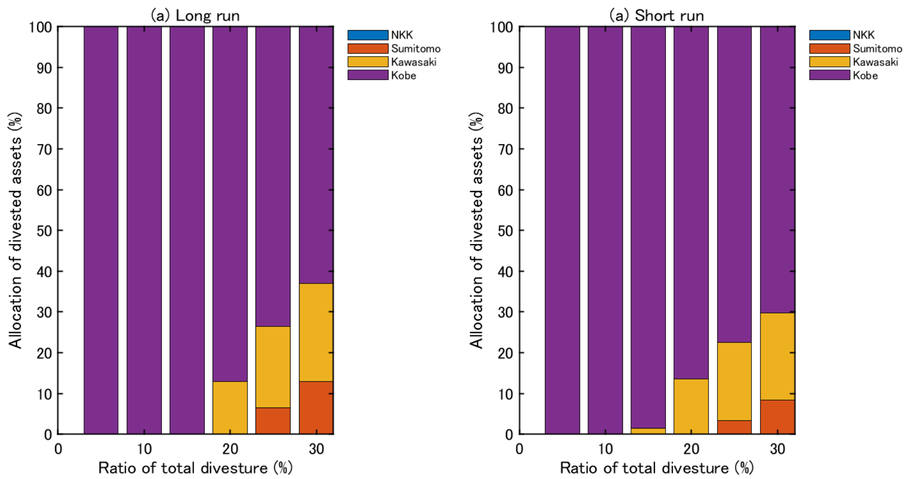
To solve the problem, we divide it into two problems:

$$\begin{aligned} \max_D \max_{\Delta(D)=\{\Delta_j(D)\}_{j\in\mathcal{J}^{(nonmerged)}}} \quad & CW(\Delta(D),D) \\ s.t. \quad & \sum_{j\in\mathcal{J}^{(nonmerged)}} \Delta_j = 1 \end{aligned}$$

Intuitively, we first maximize consumer welfare by changing the distribution of divested assets ($\Delta(D)$) given the size of the total divesture (D). Then, we maximize consumer welfare by searching for the value of D with the largest value of consumer welfare under the optimal distribution ($\max_{\Delta(D)=\{\Delta_j(D)\}_{j\in\mathcal{J}^{(nonmerged)}}} CW(\Delta(D),D)$).

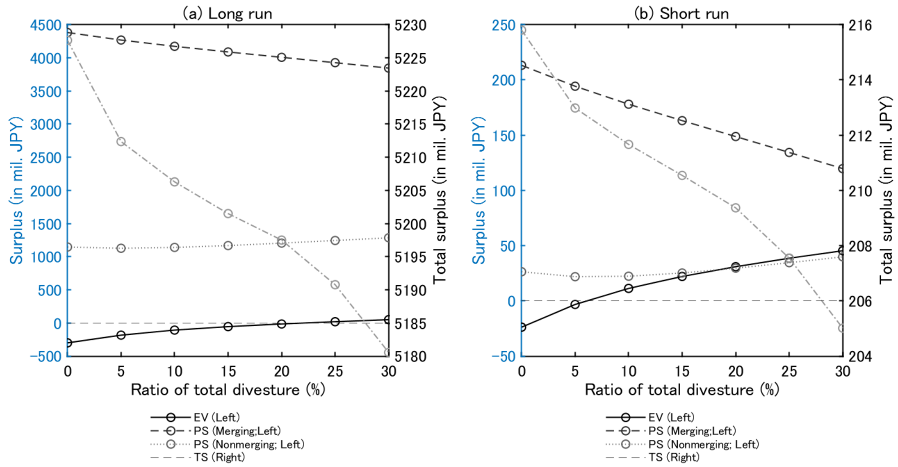
Figure 9 shows the solutions to the first problem, namely, the distribution of divested assets to non-merged firms given the size of total divesture. Figure 9(a) shows the distribution of the divested assets maximizing the discounted sum of the equivalent variation (consumer welfare) over time by endogenizing firms' investment decisions. Figure 9(b) shows the distribution of the divested assets maximizing the consumer welfare only in 1970. In both settings, it is optimal to distribute the assets mainly to Kobe, the smallest among non-merging firms. As discussed in Vergé (2010) based on the static framework, restoring symmetry mitigates the negative effect on consumer welfare. In the current setting, Kobe is the smallest, and letting Kobe larger is preferable from the viewpoint of the consumer welfare standard.

Figure 9: Optimal distribution of divested assets (Consumer welfare standard)



Even though short-run and long-run analyses yield similar results regarding the optimal distribution of the divested assets, we can find a large difference in terms of the optimal size of the divesture. Figure 10 shows the values of each surplus as functions of the amount of the divesture. As in the previous figures, the left figure (a) shows the case based on the long-run consumer welfare standard, and the right figure (b) shows the case based on the short-run consumer welfare standard. The left figure (a) shows that equivalent variation increases as the size of the divested assets increases, and the loss of consumer welfare would be offset by the merger remedies when the ratio of total divesture reaches 20%. In contrast, The right figure (b) shows that the loss of consumer welfare would be offset by the merger remedies even when the ratio is only 5%. The large difference in the required size of the divested assets comes from the diminishing effect of merger remedies in the long run discussed in the previous subsection. In the long run, the effect of merger remedies decreases over time, and the loss of consumer welfare might reappear in the future, even if competitive authorities implemented merger remedies which fully offset the loss of consumer welfare due to the merger at the period the policy was implemented. Consequently, considering consumer welfare in the long run, larger-scale remedies are required to offset the loss of the discounted sum of the consumer welfare over time even though not needed in the short run.

Figure 10: The size of divestitures and surplus (Consumer welfare standard)



Note that larger remedies would not be necessary if the competitive authorities could monitor the market and enforce additional remedies each year even after the merger. If taking such procedures is too costly, imposing stronger remedies is desirable from the viewpoint of long-run consumer welfare.

6.3.2 Total surplus standard

Next, we consider the optimal merger remedies based on the total surplus standard. Figure 11 shows the optimal distributions of the divested assets given the amount of the divested assets based on the total surplus standard. Unlike the results based on the consumer welfare standard, distributing the divested assets to the smallest firm Kobe is not preferable. Instead, distributing mainly to the second smallest firm Kawasaki is optimal. As discussed in Appendix C, marginal cost reduction through the increase in capital stock would induce a lower equilibrium price and might lead to a lower producer surplus. In the current empirical setting, distributing to Kobe largely affects the marginal cost of the firm, and it lowers the equilibrium price and producer surplus by a significant amount. Then, distributing to Kobe is not necessarily optimal based on the total surplus standard, even though it is optimal based on the consumer welfare standard. Consequently, we should be careful about the tradeoff between the producer surplus and consumer welfare. For details of the discussion, see Appendix C.

Figure 11: Optimal distribution of divested assets (Total surplus standard)

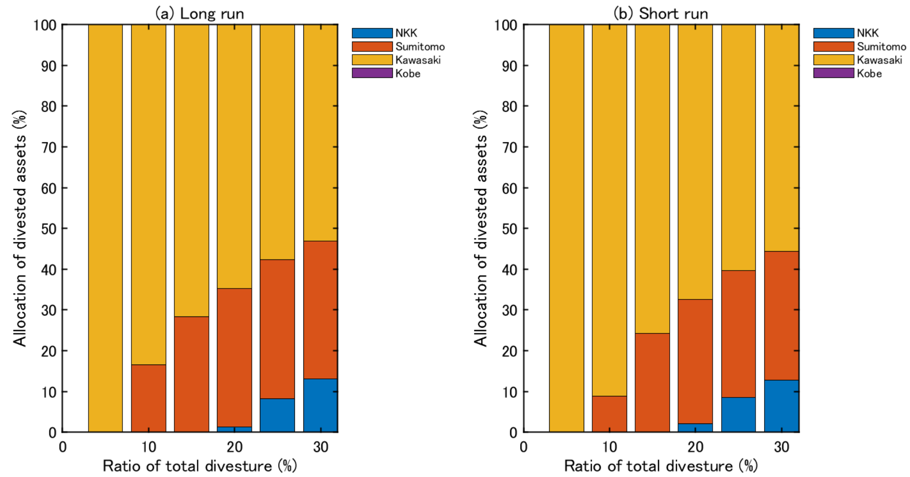
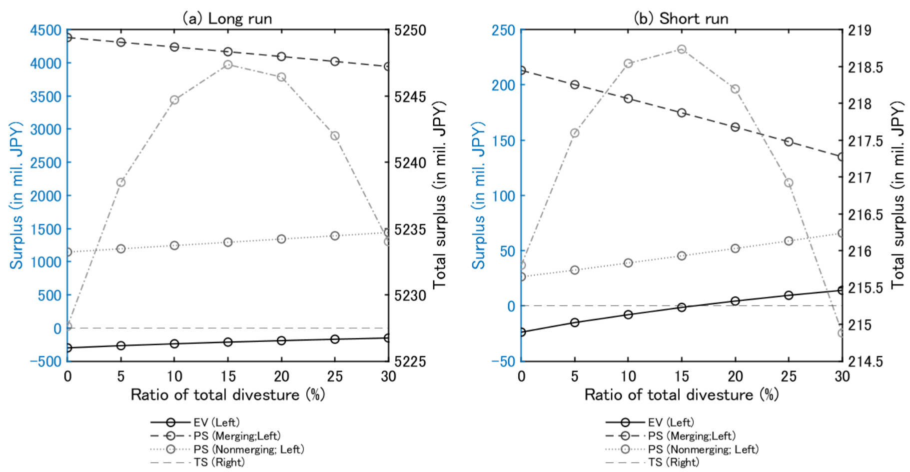


Figure 12: The size of divestitures and surplus (Total surplus standard)



Note that the ratio of the actual divesture of 1% is much lower than the optimal level, whether based on a consumer welfare standard (over 25%) or a total surplus standard (15%). Even though the merger remedies actually implemented in 1970 helped alleviate the loss of consumer welfare due to the merger between Fuji and Yawata, it was not enough to offset the effect. Besides, the remedies led to a lower total surplus.

7 Conclusion

This paper estimated a dynamic oligopoly model with firms' investment choices to evaluate the economic consequences of the horizontal merger that took place in 1970 between Yawata and Fuji. The simulation results show that even though the merger lowered consumer welfare, it increased the total surplus. Regarding the effect of the merger on firms' investment, it discouraged the merged party's investment and encouraged the non-merged party's investment. This firms' endogenous investment behavior contributed to positively to the increase in social welfare, absent the efficiency gains of the merged firm.

In this study, we further evaluated the merger remedies, and the results show that the effect of merger remedies persisted even ten years after the policy, and its long-run consequence is also important. Besides, the optimal size of merger remedies based on the short-run consumer welfare standard underestimated that based on the long-run consumer welfare standard. Furthermore, the optimal distribution of the merged party's divested assets based on the consumer welfare standard might be largely different from the one based on the total surplus standard.

A Solution method for solving the optimal investment decision

One large obstacle to implementing the estimation with the full solution approach and counterfactual simulation is that it is costly to compute equilibrium outcomes given parameter values. In this study, we solve the equilibrium model with continuous investment choice by applying Pakes and McGuire (1994) type algorithm combining the following tricks: a new algorithm that avoids the root-finding process, the Smolyak method, and the spectral algorithm. As pointed out in Hashmi and Biesebroeck (2016), there are three sources of computational burden for solving fixed point problems: (i)computation of the continuation values / optimal investment decision in each iteration, (ii)the size of the state space, and (iii)the number of iterations until convergence. The new algorithm that avoids the root-finding process mitigates the first source of the burden. The smolyak method mitigates the second source of the burden. The spectral algorithm mitigates the third source of the burden. The simplicity of the procedure is not lost even with these tricks.

Solution method for computing policy function

Optimal investment $i_{jt}(s_t)$ satisfies the following equation[14]:

$$-\frac{\partial\phi(k_{jt}, i_{jt}(s_t))}{\partial i_{jt}} + \beta\frac{\partial V_{jt+1}\left(k_{jt} + i_{jt}(s_t), k_{-jt} + i_{-jt}(s_t), \eta_{t+1}\right)}{\partial i_{jt}} = 0$$

Since $\phi(k_{jt}, i_{jt}) = \theta_k i_{jt} + \theta_a i_{jt}^2/k_{jt}$, we have:

$$-\theta_k - 2\theta_a i_{jt}(s_t)/k_{jt} + \beta\frac{\partial V_{jt+1}\left(k_{jt} + i_{jt}(s_t), k_{-jt} + i_{-jt}(s_t), \eta_{t+1}\right)}{\partial i_{jt}} = 0 \tag{14}$$

Similarly, value function $V_{jt}(s_t)$ satisfies the following equation:

$$V_{jt}(s_t) = \pi_{jt}(s_t) - \theta_1 i_{jt}(s_t) - \theta_a\frac{i_{jt}(s_t)^2}{k_{jt}} + \beta E_t V_{jt+1}\left(k_{jt} + i_{jt}(s_t), k_{-jt} + i_{-jt}(s_t), \eta_{t+1}\right) \tag{15}$$

In principle, we need to solve $V_{jt}(s_t)$ and $i_{jt}(s_t)$ satisfying equations (14) and (15) for all $j = 1, \cdots, J$, $t = 1, \cdots, T$ and for all s_t[15]. To solve these variables, we use the following algorithm:

[14]Hereinafter, we omit Θ and θ for notational simplicity.
[15]Note that we impose the assumption of $i_{jT} = 0$

Algorithm 1 Solution method for solving the equilibrium

Take arbitrary grid points of state variables $s_t = (k_{jt}, k_{-jt}, \eta_t)$ $(t = 1, \cdots, T)$. To accurately approximate the value function, the number of grid points should be sufficiently large.

1. Set initial values of $V_{jt}^{(0)}(s_t)$ (value function) for all $t = 2, \cdots, T$, $j = 1, \cdots, J$ and grid points s_t and $i_{jt}^{(0)}(s_t)$ for all $t = 1, \cdots, T$, $j = 1, \cdots, N$ and grid points s_t (optimal investment).

2. Iterate the following process until the convergence of $V_{jt}^{(n)}(s_t)$ and $i_{jt}^{(n)}(s_t)$:

   (a) At each grid point s_t, solve for $i_{jt}^{*}(s_t)$ satisfying the following equation, which is the counterpart of Eq. (14):

$$-\theta_k - 2\theta_a i_{jt}^{*}(s_t)/k_{jt} + \beta \frac{\partial V_{jt+1}^{(n)}\left(k_{jt} + i_{jt}^{(n)}(s_t), k_{-jt} + i_{-jt}^{(n)}(s_t), \eta_{t+1}\right)}{\partial i_{jt}} = 0 \tag{16}$$

   (b) Interpolate the values of $V_{jt+1}(k_{jt} + i_{jt}^{(n)}(s_t), k_{-jt} + i_{-jt}^{(n)}(s_t), \eta_{t+1})$ using the values of $V_{jt+1}^{(n)}(s_{t+1})$. Let $\overline{V}_{jt+1}$ be the interpolated values.

   (c) Find

$$V_{jt}^{*}(s_t) = \pi_{jt}(s_t) - \phi(k_{jt}, i_{jt}) + \beta \overline{V}_{jt+1}^{(n)}(k_{jt} + i_{jt}^{(n)}(s_t), k_{-jt} + i_{-jt}^{(n)}(s_t), \eta_{t+1}) \tag{17}$$

   , which is the counterpart of Eq.(15).

   (d) Update $i_{jt}(s_t)$ and $V_{jt}(s_t)$ by:

$$i_{jt}^{(n+1)}(s_t) = \rho_I^{(n)} \cdot i_{jt}^{*}(s_t) + (1 - \rho_I^{(n)}) \cdot i_{jt}^{(n)}(s_t) \tag{18}$$

$$V_{jt}^{(n+1)}(s_t) = \rho_V^{(n)} \cdot V_{jt}^{*}(s_t) + (1 - \rho_{V^{(n)}}) \cdot V_{jt}^{(n)}(s_t) \tag{19}$$

   where $\rho_I^{(n)}$ and $\rho_V^{(n)}$ represent tune parameters.

If the algorithm converges, we can easily find that computed V_{jt} and i_{jt} satisfy Eqs. (14) and (15). Hence, they are the solutions of the equilibrium.

One unique point in this algorithm is the step 2(a). In the Pakes and McGuire (1994) algorithm, which has been widely used in most of the literature, researchers compute optimal investment which satisfies the first order condition (14) given the value function $V_{jt}^{(n)}(s_t)$ and other firms' investment strategies $i_{-jt}^{(n)}(s_t)$ in each iteration. Namely, we compute $i_{jt}^{*}(s_t)$ by solving the following equation:

$$-\theta_k - 2\theta_a i_{jt}^{*}(s_t)/k_{jt} + \beta \frac{\partial V_{jt+1}^{(n)}\left(k_{jt} + i_{jt}^{*}(s_t), k_{-jt} + i_{-jt}^{*}(s_t), \eta_{t+1}\right)}{\partial i_{jt}} = 0 \tag{20}$$

The problem of solving this equation is that this is a nonlinear equation with respect to $i_{jt}^{*}(s_t)$. To find the solution of the equation, we need to numerically solve the nonlinear equation. Since this root finding process should be done in each iteration, computational burden gets so large.

In contrast to Eq.(20) in Pakes and McGuire (1994) algorithm, Eq.(16) in our algorithm is linear. We can compute the solution of equation (16) analytically:

$$i_{jt}^{*}(s_t) = \frac{-\theta_k + \beta \frac{\partial V_{jt+1}^{(n)}\left(k_{jt}+i_{jt}^{(n)}(s_t), k_{-jt}+i_{-jt}^{(n)}(s_t), \eta_{t+1}\right)}{\partial i_{jt}}}{2\theta_a / k_{jt}}$$

As in Pakes and McGuire (1994), there is no guarantee that the algorithm always converges. Nevertheless, in our case, we can successfully solve the equilibrium.

Spectral Algorithm

One problem with solving the equilibrium is that it takes much time to compute equilibrium. A large number of iterations are needed to attain convergence. To speed up the convergence process, I incorporate the spectral algorithm, developed in Barzilai and Borwein (1988) and improved in La Cruz et al. (2006). In Algorithm 1, in principle we can update the values of $i_{jt}^{(n+1)}$ and $V_{jt}^{(n+1)}$ by specifying $\rho_I^{(n)} = \rho_V^{(n)} = 1$, namely, $i_{jt}^{(n+1)} = i_{jt}^{*}$ and $V_{jt}^{(n+1)} = V_{jt}^{*}$. Nevertheless, the convergence process under the specification is not so fast. The essence of the spectral algorithm is adjusting the values of tune parameters ρ_I and ρ_V flexibly based on the convergence process[16].

More specifically, we update the values of $i_{jt}^{(n+1)}$ and $V_{jt}^{(n+1)}$ as follows. Let $i_{jt}^{*}(s_t) \equiv \psi_i(i_{jt}^{(n)}, V_{jt}^{(n)})$ and $V_{jt}^{*}(s_t) \equiv \psi_V(i_{jt}^{(n)}, V_{jt}^{(n)})$. Next, we define functions $\phi_i(i_{jt}^{(n)}, V_{jt}^{(n)}) \equiv i_{jt}^{(n)} - \psi_i(i_{jt}^{(n)}, V_{jt}^{(n)})$ and $\phi_V(i_{jt}^{(n)}, V_{jt}^{(n)}) \equiv V_{jt}^{(n)} - \psi_V(i_{jt}^{(n)}, V_{jt}^{(n)})$. Then, we employ the specification in Varadhan and Roland (2008), and update the values of $i_{jt}^{(n+1)}$ and $V_{jt}^{(n+1)}$ by $i_{jt}^{(n+1)} = i_{jt}^{(n)} - \rho_i^{(n)} \phi_i(i_{jt}^{(n)}, V_{jt}^{(n)})$ and $V_{jt}^{(n+1)} = V_{jt}^{(n)} - \rho_V^{(n)} \phi_V(i_{jt}^{(n)}, V_{jt}^{(n)})$, where $\rho_i^{(n)} = sgn(\Delta i^{(n)\prime} \Delta \phi_i^{(n)}) \frac{\left\|\Delta i^{(n)}\right\|}{\left\|\Delta \phi_i^{(n)}\right\|}$ and $\rho_V^{(n)} = sgn(\Delta V^{(n)\prime} \Delta \phi_V^{(n)}) \frac{\left\|\Delta V^{(n)}\right\|}{\left\|\Delta \phi_V^{(n)}\right\|}$. Note that we can easily verify the updates described above is in the form of Eq.(19).

Another problem of the full solution approach is the possibility of multiple equilibria. If several equilibria exist, in principle we need to find all the equilibria to yield valid inference. In the context of NPL estimator, Aguirregabiria and Marcoux (2021) has shown that the spectral algorithm can find unstable equilibria which cannot be found in the standard algorithm, and performs well even when multiple equilibria exist. Hence, applying the spectral algorithm would mitigate the problems associated with the possibility of the existence of multiple equilibria. Note that we obtained only one solution given parameter values even though we tried several initial values.

Smolyak method

In Algorithm 1, we need to take a large number of grid points to approximate the shape of the functions well $i_{jt}(s_t)$ and $V_{jt}(s_t)$. Nevertheless, it takes more time as the number of grid points increases. To reduce the number of grid points without losing numerical accuracy, we use the Smolyak method developed by Smolyak (1963) and improved by Judd et al. (2014).

[16] With the spectral algorithm, the speed of convergence is roughly two times faster than the case under $\rho_I^{(n)} = \rho_V^{(n)} = 1$ in our model.

B Further discussion on the effect of the merger on firms' investment

The discussion in Section 6.1 shows that the merger lowered the merged party's investment, but encouraged the non-merged party's investment. Then, why did the merger altered the firms' incentives for investment? To answer the question, in the following, we analytically decompose the firms' incentives to invest into several parts and quantify the contributions of each incentive using the estimated structural model.

First, we define a term $\widetilde{V_{jt}}(i_{jt}, i_{-jt}; I_{jt+1}, I_{-jt+1}; k_t) \equiv \pi_{jt}(k_t) + \beta V_{jt+1}(k_{jt+1}(k_{jt}, i_{jt}), k_{-jt+1}(k_{jt}, i_{-jt}); I_{jt+1}, I_{-jt+1})$. Here, I_{jt+1} represents the set of firm j's investment strategies ($\{i_{jt+\tau}\}_{\tau \geq 1}$) as defined in Section 6.1. Then, the first order condition of firm j's investment choice at time t is in the following form:

$$\frac{\partial \widetilde{V_{jt}}(i_{jt}, i_{-jt}; I_{jt+1}, I_{-jt+1}; k_t)}{\partial i_{jt}} = \frac{\partial \phi(k_{jt}, i_{jt})}{\partial i_{jt}}$$

Here, the left-hand side represents firm j's marginal gains from investment, and the right-hand side represents the marginal cost of investment. Under the functional form assumption that $\phi(k_{jt}, i_{jt}) = \theta_k k_{jt} + \theta_a i_{jt}^2 / k_{jt}$, firm j's equilibrium investment at time t satisfies:

$$i_{jt}^*(k_t) = \frac{\frac{\partial \widetilde{V_{jt}}}{\partial i_{jt}}(i_{jt}^*, i_{-jt}^*; I_{jt+1}^*, I_{-jt+1}^*; k_t) - \theta_k}{2\theta_a} k_{jt}$$

B.1 The effect of the merger on merged party's investment

In this subsection, we consider the effect of a merger on the merged party's investment.

Let j_1, j_2 be the merging firms, and j_0 be the merged firm. Further, let the subscript N be represents the outcomes or strategies under the no merger case. Similarly, let the subscript M represents those under the merger case. Then, the following claim holds regarding the effect of the merger on merged party's investment[17]:

Proposition 1. *The impact of the merger on merged party's investment at the period the merger has occured, $i_{j_0 t}^{*M} - (i_{j_1 t}^{*N} + i_{j_2}^{*N})$, has the same sign as $\kappa_{j_0} - (w_{j_1}\lambda_{j_1} + w_{j_2}\lambda_{j_2})$, where $\kappa_{j_0} \equiv \frac{\partial \widetilde{V_{j_0 t}^M}}{\partial i_{j_0 t}}(i_{j_1 t}^{*N} + i_{j_2 t}^{*N}, i_{-j_0 t}^{*M}; I_{j_0 t+1}^M, I_{-j_0 t+1}^M; k_t^M)$ and $\lambda_j \equiv \frac{\partial \widetilde{V_{j_1 t}^N}}{\partial i_{j_1 t}}(i_{jt}^{*N}, i_{-jt}^{*N}; I_{jt+1}^N, I_{-jt+1}^N; k_t^N)$ $(j = j_1, j_2)$, and $w_j \equiv \frac{k_j}{k_{j_1} + k_{j_2}}$ $(j = j_1, j_2)$.*

Proof. See Appendix B.3. □

Here, we introduced subscripts M and N in $\frac{\partial \widetilde{V_{jt}}}{\partial i_{jt}}$ to make the point clear and for later convenience. The values of $\frac{\partial \widetilde{V_{jt}}}{\partial i_{jt}}$ represent the marginal gain from investment, and the proposition implies that the direction of the effect of the merger on the merged party's investment is determined by the relative size of the weighted sum of merging firms' / merged firm's marginal gains from investment. Intuitively, if the investment is more profitable after the merger than the case before the merger, the merger would increase merged party's investment.

[17] Note that this intuitive representation largely relies on the functional form assumption of investment cost function $\phi(k_{jt}, i_{jt})$. If $\phi(k_{jt}, i_{jt}) = \theta_k i_{jt} + \theta_a i_{jt}^2$, $i_{j_0 t} - (i_{j_1 t} + i_{j_2 t}) = \frac{1}{2\theta_a}\left(\theta_k + \frac{\partial \widetilde{V_{j_0 t}^M}}{\partial i_{j_0 t}} - \left(\frac{\partial \widetilde{V_{j_1 t}^N}}{\partial i_{j_1 t}} + \frac{\partial \widetilde{V_{j_2 t}^N}}{\partial i_{j_2 t}}\right)\right)$ holds, and the sign of the effect also depends on the value of θ_k.

Further decompositions

In principle, we need to compare the values of κ_{j_0} and $w_{j_1}\lambda_{j_1} + w_{j_2}\lambda_{j_2}$ to judge the direction of the effect of the merger on merged party's investment. Nevertheless, the direct comparison between these two values is not necessarily easy to understand, because several factors have effects on the difference. Below, we disentangle the effect of the merger by conducting further decompositions.

Decomposition of the marginal gains from investment

We can decompose the marginal gain from investment $\frac{\partial \widetilde{V_{jt}}(i_{jt}, i_{-jt}; I_{jt+1}, I_{-jt+1}; k_t)}{\partial i_{jt}}$ into five terms:

$$
\begin{aligned}
&\frac{\partial \widetilde{V_{jt}}(i_{jt}, i_{-jt}; I_{jt+1}, I_{-jt+1}; k_t)}{\partial i_{jt}} \\
= &\sum_{\tau=1}^{\infty} \beta^{\tau} \frac{\partial \pi_{jt+\tau}(k_{t+\tau}) - \phi(k_{t+\tau}, i_{t+\tau})}{\partial k_{jt+1}} \\
= &\underbrace{-\sum_{\tau=1}^{\infty} \beta^{\tau} \left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{own} \frac{\partial mc_{jt+\tau}(k_{jt+\tau})}{\partial k_{jt+\tau}} q_{jt+\tau}}_{\text{Production cost reduction}} + \underbrace{\sum_{\tau=1}^{\infty} \beta^{\tau} \left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{own} \frac{\partial mc_{jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial q_{-jt+\tau}}{\partial mc_{jt+\tau}} \frac{\partial}{\partial q_{-jt+\tau}} \pi_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau})}_{\text{Strategic use of investment}} + \\
&\underbrace{-\sum_{\tau=1}^{\infty} \beta^{\tau} \left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{own} \frac{\partial \phi(k_{jt+\tau}, i_{jt+\tau})}{\partial k_{jt+\tau}}}_{\text{Investment cost reduction due to scale economy}} + \underbrace{(-1) \cdot \sum_{\tau=1}^{\infty} \beta^{\tau} \left.\frac{\partial k_{jt+\tau-1}}{\partial k_{jt+1}}\right|_{own} \frac{\partial i_{jt+1}}{\partial k_{jt+\tau-1}} \frac{\partial \phi(k_{jt+\tau}, i_{jt+\tau})}{\partial i_{jt+\tau}}}_{\text{Change in future investment cost}} + \\
&\underbrace{\sum_{\tau \geq s} \beta^{\tau} \sum_{s=2}^{\infty} \frac{\partial k_{-jt+s}}{\partial k_{jt+1}} \frac{\partial (\pi_{jt+\tau}(k_{t+\tau}) - \phi(k_{jt+\tau}, i_{jt+\tau}))}{\partial k_{-jt+s}}}_{\text{Preemptive motive}}
\end{aligned}
\tag{21}
$$

Here, we used the optimality condition of firm j's output choice $\frac{\partial}{\partial q_{jt+\tau}} \pi_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau}) = 0$ to derive the second equality. $\left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{own} \equiv \prod_{s=2}^{\tau} \frac{\partial k_{jt+s}}{\partial k_{jt+s-1}}$ denotes the derivative of $k_{jt+\tau}$ with respect to k_{jt+1} excluding the effect on other firms' capital stocks $k_{jt+\tau}$. In general, firm j's capital stock k_{jt+1} at time $t+1$ might affect firm $-j$'s investment i_{-jt+1} at time $t+1$. Then, firm $-j$'s capital stock k_{-jt+2} also changes, and it might also affects firm j's profit or investment strategies at time $t+2$. By defining the term $\left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{own}$, we exclude the intertemporal strategic interactions among firms regarding their investment.

The first term in the decomposition can be interpreted as the incentive associated with the reduction of production cost, because the marginal increase in capital lowers the firm's marginal cost. The second term can be regarded as the incentive associated with the strategic use of investment. A firm's larger capital stock lowers its own marginal cost, and it lowers other firms' equilibrium outputs. Under the investment cost function $\phi(k_{jt}, i_{jt}) = \theta_k k_{jt} + \theta_a i_{jt}^2/k_{jt}$, larger scale firms incur lower investment cost, and firms have incentives to be larger to lower the investment cost. This incentive is reflected in the third term. Also, a firm's current investment might affect its future investment and future investment cost. This is reflected in the fourth term. Finally, the fifth term can be interpreted as the preemptive motive, because a firm's current investment decision affects other firms' future investment decisions, and it affects the own firm's future profit. Even though there is no guarantee that the increase in own firm's increase increases future profits through the effect on other firms' investment, the values are positive in the quantitative results, as we show later.

Decomposition of the effect of the merger on merged party's investment incentives

Here, aside from the weighted sum of the terms $(0)\sum_{j\in\{j_1,j_2\}} w_j \frac{\partial \widetilde{V^N_{jt}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-jt}; I^N_{jt+1}, I^N_{-jt+1}; k^N_t)$ and $(3)\frac{\partial \widetilde{V^M_{j_0t}}}{\partial i_{j_0t}}(i^{*N}_{j_1t} + i^{*N}_{j_2t}, i^{*M}_{-j_0t}; I^M_{j_0t+1}, I^M_{-j_0t+1}; k^M_t)$ which appeared in Proposition 1, we introduce two additional terms: $(1)\sum_{j\in\{j_1,j_2\}} w_j \frac{\partial \widetilde{V^M_{j_0t}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-jt}; I^N_{jt+1}, I^N_{-jt+1}; k^N_t)$ and $(2)\sum_{j\in\{j_1,j_2\}} w_j \frac{\partial \widetilde{V^M_{j_0t}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-j_0t}; I^M_{j_0t+1}, I^N_{-j_0t+1}; k^N_t)$. In (1), merged party decides its output as a single firm, and its objective function is the profit of the merged entity. Nevertheless, two merging firms follow the investment strategies absent the merger, and decide their investment levels individually. (2) is almost the same as (1), except for the investment strategies of the merged party after time $t+1$. In (2), merged party switches to the investment strategy under the merger after time $t+1$ [18].

Note that the difference between the values of (0) and (1) comes not only from the change in merged party's output choice but also from the change in production cost structure. The term of production cost reduction is $\sum_{j\in\{j_1,j_2\}} w_j \cdot (-1) \sum_{\tau=1}^{\infty} \beta^\tau \left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{own} \frac{\partial mc_{jt+\tau}(k^N_{jt+\tau})}{\partial k^N_{jt+\tau}} q^N_{jt+\tau}$ in (0), but $\sum_{j\in\{j_1,j_2\}} w_j \cdot (-1) \sum_{\tau=1}^{\infty} \beta^\tau \left.\frac{\partial (k^N_{j_1t+\tau}+k^N_{j_2t+\tau})}{\partial k^N_{jt+1}}\right|_{own} \frac{\partial mc_{j_0t+\tau}(k^N_{j_1t+\tau}+k^N_{j_2t+\tau})}{\partial (k^N_{j_1t+\tau}+k^N_{j_2t+\tau})} q^M_{j_0t+\tau}$ in (1). Hence, the values of the derivative of the marginal cost with respect to the firm's capital $\frac{\partial mc_{jt+\tau}(k_{jt+\tau})}{\partial k_{jt+\tau}}$, and firm's output $q_{j_0t+\tau}$. To isolate the change in these terms, we further introduce the following terms for the term of production cost reduction[19]:

$$(1') \quad : \quad \sum_{j\in\{j_1,j_2\}} w_j \cdot (-1) \sum_{\tau=1}^{\infty} \beta^\tau \left.\frac{\partial (k^N_{j_1t+\tau}+k^N_{j_2t+\tau})}{\partial k^N_{jt+1}}\right|_{own} \frac{\partial mc_{jt+\tau}(k^N_{jt+\tau})}{\partial k^N_{jt+\tau}} (q^N_{j_1t+\tau} + q^N_{j_2t+\tau})$$

$$(1'') \quad : \quad \sum_{j\in\{j_1,j_2\}} w_j \cdot (-1) \sum_{\tau=1}^{\infty} \beta^\tau \left.\frac{\partial (k^N_{j_1t+\tau}+k^N_{j_2t+\tau})}{\partial k^N_{jt+1}}\right|_{own} \frac{\partial mc_{jt+\tau}(k^N_{jt+\tau})}{\partial k^N_{jt+\tau}} q_{j_0t+\tau}$$

$(1')$ is the alternative value of (1) where the values of the derivative of the marginal cost with respect to the firm's capital $\frac{\partial mc_{jt+\tau}(k_{jt+\tau})}{\partial k_{jt+\tau}}$ and output $q_{jt+\tau}$ are fixed at the no merger level. $(1'')$ is the alternative value of (1) where the values of the derivative of the marginal cost with respect to the firm's capital $\frac{\partial mc_{jt+\tau}(k_{jt+\tau})}{\partial k_{jt+\tau}}$ are fixed at the no-merger level.

As shown in the difference between (1) and $(1')$, the merger enhances the range of the product where its production cost gets lower as the merged party increases its capital. Before the merger, merging firm j_1 could only lower the production cost of the product with output $q^N_{j_1t+\tau}$. In contrast, after the merger, merging firm j_1's investment lowers the production cost of the product with output $q^N_{j_1t+\tau} + q^N_{j_2t+\tau}$. Consequently, the merger increases merged party's incentive for investment.

In summary, we compare the following values:

[18] Since the nonmerged firms follow the investment strategies absent the merger, merged party's investment strategy is not necessarily optimal. Besides, nonmerged firms following no-merger investment strategies make invest decisions based on the state space in the form of $(k_{j_1t+\tau}, k_{j_2t+\tau}, \{k_{jt+\tau}\}_{j\neq j_1,j_2})$. Since the values of $k_{j_1t+\tau}, k_{j_2t+\tau}$ are not available when the merged party make investment decisions as a single entity, we compute these value values by $k_{jt+\tau} = w_j k_{j_0t+\tau}, j = j_1, j_2$, where w_j denotes the weight of the capital stock of firm j among the merging firms at the time the merger occurred, as defined in Proposition 1.

[19] In $(1'')$, we also show the corresponding values of the strategic use of the investment.

$$(0)\sum_{j\in\{j_1,j_2\}} w_j \frac{\partial \widetilde{V^N_{jt}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-jt}; I^N_{jt+1}, I^N_{-jt+1}; k^N_t)$$

$$(1)\sum_{j\in\{j_1,j_2\}} w_j \frac{\partial \widetilde{V^M_{j_0t}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-jt}; I^N_{jt+1}, I^N_{-jt+1}; k^N_t)$$

$$(2)\sum_{j\in\{j_1,j_2\}} w_j \frac{\partial \widetilde{V^M_{j_0t}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-j_0t}; I^M_{j_0t+1}, I^N_{-j_0t+1}; k^N_t)$$

$$(3)\frac{\partial \widetilde{V^M_{j_0t}}}{\partial i_{j_0t}}(i^{*N}_{j_1t}+i^{*N}_{j_2t}, i^{*M}_{-j_0t}; I^M_{j_0t+1}, I^M_{-j_0t+1}; k^M_t)$$

$(1')$: alternative value of (1) where the marginal decrease in marginal cost and output are fixed at the no merger level

$(1'')$: alternative values of (1) where the marginal decrease in marginal cost are fixed at the no merger level.

The difference between the values of (1) and (2) comes from the merged party's investment strategies after time $t+1$, and that of (2) and (3) comes from the nonmerged party's investment strategies. The difference between (0) and $(1')$ comes from the range of the production cost of the output affected by the investment, that between $(1')$ and $(1'')$ comes from the change in merged party's output after the merger, The difference between $(1'')$ and (1) comes from the change in the values of the derivative of the marginal cost with respect to the capital stock. Note that marginal decrease in marginal cost gets smaller as the size of the capital stock gets larger. Since merger increases the capital stock held by one firm, it consequently leads to the lower absolute values of $\frac{\partial mc_{jt+\tau}(k_{jt+\tau})}{\partial k_{jt+\tau}}$, implying lower investment incentive for the merged party. This effect can be regarded as a scale economy in production cost.

Table B.1 shows the quantitative results of the decomposition of the merged party's investment incentives. In the table, we show the case of the merger without and with efficiency gains. In the table, larger values imply larger incentives for investment. Even though the decomposition in Eq (21) assumes that the future investment $i_{jt+\tau}(\tau \geq 1)$ can change as the party changes the value of i_{jt}, we also show the hypothetical case where the future investments do not change irrespective of the change in i_{jt}, namely, the case where $\left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{own} = 1\ \forall \tau \geq 2$. The purpose is to make it easier to compare the results with the static analysis widely applied in theoretical studies, where it is assumed that investment decisions at different periods are independent. Note that the values of $\left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{own}$ are equal to one for all $\tau \geq 1$, in the "no variable future investment" case.

Table 8: Marginal gain from investment (Merged firm)

	Variable future inv.	No Merger (0)	Merger (No eff. gain) $(1')$	$(1'')$	(1)	(2)	(3)	Merger (1)	(2)	(3)
Lower Production cost	No	400	796.5	454.3	227.9	259.9	294.5	284	397.9	331.8
	Yes	331.7	658.5	375.5	188.2	184.3	202.1	234	268.8	232.5
Strategic use	No	277.8	-	290.9	145.9	164.8	182.8	202.7	279.3	235.1
	Yes	230.5	-	240.6	120.6	117.3	126.4	167.1	189.6	165
Inv. cost (scale economy)	No	78.5	-	-	78.4	35.6	9.8	78.4	10.5	35.5
	Yes	76.9	-	-	76.4	32.8	9.1	76.4	9.7	32.7
Future Inv. cost	Yes	120.7	-	-	123.6	176.3	153.5	123.6	153.3	177.3
Preemptive	Yes	25.6	-	-	33.7	32.3	11.7	43.7	45.7	16.2
Total	No	756.3	-	-	452.1	460.3	487.1	565.1	687.7	602.5
	Yes	785.4	-		542.5	542.9	502.9	644.8	667.2	623.7

Regarding the decomposition of the marginal gains from investment, each factor contributes to the party's incentive positively. The investment motive to lower production cost is the largest, but strategic use of investment to lower other firms' outputs also plays a large role. Dynamic incentives without intertemporal strategic interaction, including the investment cost reduction due to scale economy and the change in future investment cost are not negligible, but the values of preemptive motive is relatively small: roughly 3% of the total incentive under the case absent the merger.

Next, by comparing the values in different settings, we can observe that the merger decreased merged party's incentive to invest in total, and following factors largely contributed to the change in merged party's investment incentives:

- larger output affected by a unit investment after the merger (+)
- Merged party's smaller output after the merger (-)
- Scale economy (-)
- Efficiency gains (+)

Note that the preemptive motive take positive values because of the nature of (intertemporal) strategic substitutability in investment. Table 9 and 10 show the effect of marginal increase in each firm's capital stock in the current period on firms' investment in the next period, and they imply that increase in one firm's current capital stock discourages other firm's future investment. Consequently, increase in the merged firm's investment increases its own profits through discouraging non-merged firms' investment.

Table 9: Elasticity of investment with respect to capital stock (with merger case)

$k_{jt} \backslash i_{jt}$	NSC	NKK	Sumitomo	Kawasaki	Kobe
NSC	-0.18	-0.03	-0.03	-0.03	-0.10
NKK	0.00	0.24	-0.05	-0.05	-0.17
Sumitomo	0.00	-0.04	0.29	-0.04	-0.18
Kawasaki	0.00	-0.05	-0.04	0.32	-0.17
Kobe	-0.03	-0.11	-0.12	-0.12	0.75

Note: The value is calculated by: $\frac{\partial \log i_{jt}}{\partial \log k_{mt}}$ $(t = 1970)$

Table 10: Elasticity of investment with respect to capital stock (with merger case)

$k_{jt} \backslash i_{jt}$	Yawata	Fuji	NKK	Sumitomo	Kawasaki	Kobe
Yawata	0.07	-0.01	-0.03	-0.02	-0.02	-0.17
Fuji	-0.01	0.17	-0.03	-0.03	-0.03	-0.17
NKK	-0.01	-0.02	0.27	-0.04	-0.04	-0.21
Sumitomo	-0.01	-0.02	-0.04	0.33	-0.03	-0.23
Kawasaki	-0.01	-0.02	-0.04	-0.03	0.36	-0.22
Kobe	-0.09	-0.08	-0.13	-0.17	-0.16	1.12

Note: The value is calculated by: $\frac{\partial \log i_{jt}}{\partial \log k_{mt}}$ $(t = 1970)$

B.2 The effect of the merger on nonmerged party's investment

Analogous to the case of the merged party's investment, we can also obtain an results regarding the non-merged party's investment:

Proposition 2. *The impact of the merger on non-merged party's investment at the period the merger has occurred,* $\sum_{j\in\mathcal{J}^{(nonmerged)}} i^{*M}_{jt} - \sum_{j\in\mathcal{J}^{(nonmerged)}} i^{*N}_{jt}$*, has the same sign as* $\sum_{j\in\mathcal{J}^{(nonmerged)}} \widetilde{w_j}\mu^M_j - \sum_{j\in\mathcal{J}^{(nonmerged)}} \widetilde{w_j}\nu^M_j$*, where*

$\mu_j \equiv \frac{\partial\widetilde{V^M_{jt}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*M}_{-jt}; I^M_{jt+1}, I^M_{-jt+1}; k^M_t)$,

$\nu_j \equiv \frac{\partial\widetilde{V^N_{jt}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-jt}; I^N_{jt+1}, I^N_{-jt+1}; k^N_t)$*, and* $\widetilde{w_j} \equiv \frac{k_{jt}}{\sum_{j':non\text{-}merged} k_{j't}}$.

Proof. See Appendix B.3. □

The proposition implies that it is sufficient to compare the weighted sums of each firm's marginal gain from investment to assess the impacts of the merger on firms' investment.

Table 11: Marginal gain from investment (Non-merged firm)

	Variable future inv.	No Merger (0)	Merger (No eff. gain) (1)	(2)	(3)	Merger (1)	(2)	(3)
Lower Production cost	No	426	540.2	540.7	499.9	446.3	444.8	450
	Yes	479.8	622.4	621.1	490.1	504.8	505.2	466.2
Strategic use	No	281.3	322.6	323.1	299.8	258.8	258	261.3
	Yes	315.5	369.3	368.7	293.3	291	291.1	269.8
Inv. cost (scale economy)	No	103	103	103.4	153.1	103	103.2	106.2
	Yes	114.7	114.7	115.1	163.8	114.7	114.6	115.5
Future Inv. cost	Yes	-106.5	-106.5	-103.8	5.4	-106.5	-110.9	-36.2
Preemptive	Yes	50.5	74.6	72.1	28.3	62.6	60.1	48.5
Total	No	810.3	965.8	967.1	952.8	808	806	817.5
	Yes	854	1074.6	1073.2	980.8	866.7	860.1	863.6

The results show that the incentive of lowering production cost largely increased due to the merger with no efficiency gains. Non-merged party increased its output as shown in Table 7, and it increased the incentive of lowering production cost. In contrast, merged party's efficiency gains largely discouraged non-merged party's incentive to lower production cost, because of the smaller output due to the efficiency gains of the merged party.

B.3 Proof

Proof of Proposition 1

Proof. First, because of the merging firms' optimality conditions in terms of their investment choices, the following equation holds:

$$\begin{aligned}\frac{\partial\widetilde{V^N_{j_0t}}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-jt}; I^N_{jt+1}, I^N_{-jt+1}; k^N_t) &= \frac{\partial\phi(k^N_{jt}, i^N_{jt})}{\partial i_{jt}}\ \ j = j_1, j_2 \\ &= \theta_k + 2\theta_a\frac{i^N_{jt}}{k^N_{jt}}\ \ j = j_1, j_2\end{aligned}$$

The merged party has an incentive to increase its investment in comparison with the investment level absent the merger, if the sign of $\frac{\partial \widetilde{V_{j_0t}^M}}{\partial i_{j_0t}}(i_{j_1t}^{*N} + i_{j_2t}^{*N}, i_{-j_0t}^{*M}; I_{j_0t+1}^M, I_{-j_0t+1}^M; k_t^M) - \frac{\partial \phi(k_{jt}^M, i_{j_1t}^N + i_{j_2t}^N)}{\partial i_{jt}}$ is positive. Here,

$$
\begin{aligned}
& \frac{\partial \widetilde{V_{j_0t}^M}}{\partial i_{j_0t}}(i_{j_1t}^{*N} + i_{j_2t}^{*N}, i_{-j_0t}^{*M}; I_{j_0t+1}^{*M}, I_{-j_0t+1}^{*M}; k_t^M) - \frac{\partial \phi(k_{jt}^M, i_{j_1t}^{*N} + i_{j_2t}^{*N})}{\partial i_{jt}} \\
= & \ \kappa_{j_0} - \left(\theta_k + 2\theta_a \frac{i_{j_1t}^{*N} + i_{j_2t}^{*N}}{k_{j_0t}^M}\right) \\
= & \ \kappa_{j_0} - \frac{k_{j_1t}^N}{k_{j_0t}^M}\left(\theta_k + \theta_a \frac{i_{j_1t}^{*N}}{k_{j_1t}^N}\right) - \frac{k_{j_2t}^N}{k_{j_0t}^M}\left(\theta_k + \theta_a \frac{i_{j_2t}^{*N}}{k_{j_2t}^N}\right) \ (\because k_{j_0t}^M = k_{j_1t}^N + k_{j_2t}^N) \\
= & \ \kappa_{j_0} - (w_{j_1}\lambda_{j_1} + w_{j_2}\lambda_{j_2})
\end{aligned}
$$

Consequently, $i_{j_0t}^{*M} - (i_{j_1t}^{*N} + i_{j_2}^{*N})$, has the same sign as $\kappa_{j_0} - (w_{j_1}\lambda_{j_1} + w_{j_2}\lambda_{j_2})$. □

Proof of Proposition 2

The strategy of the proof is mostly the same as the one in Proposition 1 (merged party's case).

C Optimal distribution of the divested assets

In this section, we discuss the factors affecting the optimal distribution of the divested assets based on the total surplus standard. If the amount of the divested assets is not so large, it is sufficient to compare the relative size of the marginal effect of each firm's capital stock increase on the total surplus, namely, the derivative of the total surplus with respect to each firm's capital stock. For instance, if the value of Kobe's derivative is larger than Kawasaki's, distributing to Kobe is more desirable from the viewpoint of total surplus standard.

Table 12 shows the effect of the marginal increase in each firm's capital stock on economic variables and surpluses in 1970. To assess the effect of the firms' endogenous investment behaviors on welfare, we show two scenarios: exogenous/ endogenous investment. Regarding producer surplus, we decompose the effect of the marginal increase in capital stock into three factors:
z (1). Effect on producer surplus (excluding investment cost) through marginal cost change, given fixed equilibrium price ($\sum_{\tau=1}^{\infty} \beta^\tau (-1) \sum_{j'} \frac{\partial k_{j't+\tau}}{\partial k_{jt+1}} \frac{\partial mc_{j't+\tau}(k_{jt+\tau})}{\partial k_{j't+\tau}} q_{j't+\tau}$)
(2). Effect on producer surplus (excluding investment cost) through the change in equilibrium price ($\sum_{\tau=1}^{\infty} \beta^\tau \sum_{j'} \left[\frac{\partial k_{j't+\tau}}{\partial k_{jt+1}} \frac{\partial mc_{j't+\tau}}{\partial k_{j't+\tau}} \frac{\partial P_{t+\tau}}{\partial mc_{j't+\tau}} q_{j't+\tau} + \frac{\partial k_{j't+\tau}}{\partial k_{jt+1}} \frac{\partial mc_{j't+\tau}}{\partial k_{j't+\tau}} \frac{\partial P_{t+\tau}}{\partial mc_{j't+\tau}} \sum_{j''} \frac{\partial q_{j''t+\tau}}{\partial P_{t+\tau}} (P_{t+\tau} - mc_{j''t+\tau})\right]$)
(3). Effect on investment cost ($\sum_{\tau=1}^{\infty} \beta^\tau (-1) \sum_{j'} \left[\frac{\partial k_{j't+\tau}}{\partial k_{jt+1}} \frac{\partial \phi(k_{j't+\tau}, i_{j't+\tau})}{\partial k_{j't+\tau}} + \frac{\partial i_{j't+\tau}}{\partial k_{jt+1}} \frac{\partial \phi(k_{j't+\tau}, i_{j't+\tau})}{\partial i_{j't+\tau}}\right]$)

The first factor represents the increase in firms' profit through cost reduction. The second factor represents the effect through equilibrium price change, and the third factor represents the effect on investment cost.

First, we can observe that distributing to Sumitomo, the second smallest firm, leads to the largest total surplus, by comparing the effect of the marginal increase in each firm's capital. Even though distributing to Kobe is the most desirable alternative when the objective function is consumer welfare, it is not desirable considering the loss of producer surplus. Rows (2) in producer surplus show that the marginal increase in Kobe's capital led to a large decrease in producer surplus through the change in equilibrium price.

Regarding firms' endogenous investment behaviors, we can see that the introduction of endogenous investment behaviors intensifies the effect of the marginal increase in Kobe's capital stocks on consumer welfare and producer surplus. As shown in Table 9, Kobe's larger capital stocks discouraged the other firms' investment. Then, the market got less concentrated, leading to a lower equilibrium price, which was desirable for consumers. In contrast, the producer surplus decreased due to the increase in Kobe's capital, and even the total surplus decreased. It implies that the application of the models without explicit specifications of firms' investment decisions may underestimate the magnitude of the effect of merger remedies.

Table 12: Effect of the marginal increase in each firm's capital stock on economic variables and surpluses

			NSC	NKK	Kawasaki	Sumitomo	Kobe
mc_{jt}			-2.1	-7.9	-9.3	-10.8	-26.9
P_t^B			-0.5	-1.9	-2.3	-2.7	-6.6
EV	Exo. inv.		239.3	759.5	845.4	912.2	1918.5
	Endo. inv.		130.8	539.5	634.1	722.1	2494.1
PS	Exo. inv.	(1)	408.6	527.2	555.9	587.4	626.3
		(2)	54	-596.5	-694.1	-760.7	-2203.9
		(3)	64.7	132	161.2	198.2	311.2
		Total	527.3	62.7	23.1	25	-1266.3
	Endo. inv.	(1)	281.9	425.4	466.5	514.7	622.3
		(2)	-38.2	-475.6	-562.3	-627.5	-3086.9
		(3)	285	299.1	307.3	316.1	334.8
		Total	655.4	350.6	300.9	276	-2125.8
TS	Exo. inv.		766.6	822.1	868.4	937.1	652.2
	Endo. inv.		786.1	890.1	935.1	998.1	368.3

Notes.
"EV", "PS", "TS" denote equivalent variation, producer surplus, and total surplus respectively. "Exo. inv." denotes the case where all the firms' investment are predetermined, and firms do not alter their future investment even when each firm's capital stocks change in each period. "Endo. inv." denotes the case where all the firms follow optimal Markov strategies, and adjust their investments based on the distributions of the capital stocks.

D Data

Annual data on the industry- and firm-level output data were obtained from Japan Iron and Steel Federation (1960-1990). The annual price data of blast furnace companies for domestically produced crude steel were taken directly from blast furnace companies' semiannual financial reports (1960-1990). We found that the price level did not vary widely across firms, and hence, Yawata's crude steel prices in the 1960s and Nippon Steel's prices in the 1970s were used for the estimation. On the other hand, the annual electric furnace companies were not directly available. Since wholesale price index (WPI) published by the Bank of Japan (BOJ) was constructed as the weighted sum of the two types of steel weighted by the sale, we used the data to construct the proxy for the steel price of electric companies. These prices were adjusted by the manufactured goods WPI to a constant 1960 Japanese yen.

Two input prices were used in the paper: data on iron ore and heavy oil were taken from the Bank of Japan (1960-1990). Data on the average seaborne shipping distance of iron ore, the variable that is used as an instrument for the demand estimation, were obtained came from Japan Iron and Steel Federation (1960-1990).

Our measure of firm-level capital stock is the firm-level physical fixed production asset, taken from companies semiannual financial reports (1960-1990). From the data on capital stock, we constructed the annual amount of firm-level investment. Both capital stock and investment data were converted from book value to market value by following the method proposed in Ogawa and Kitasaka (1998). We used the national wealth survey of 1960 to obtain the average age of each physical asset in the Japanese steel industry. Finally, investment and capital stock were adjusted by the manufactured goods WPI to a constant 1960 Japanese yen.

E Additional results

E.1 Alternative investment cost function

In the baseline model, we specified the investment cost function in the following form following macroeconomics literature: $\phi(k_{jt}, i_{jt}; \theta) = \theta_k i_{jt} + \theta_a i_{jt}^2/k_{jt}$. Nevertheless, the quadratic term $\theta_a i_{jt}^2/k_{jt}$ seems to be restrictive, and we also estimated the parameters in the following form: $\phi(k_{jt}, i_{jt}; \theta) = \theta_k i_{jt} + \widetilde{\theta}_a i_{jt}^2 + \theta_a i_{jt}^2/k_{jt}$. Table 13 shows the estimation result.

Table 13: Investment cost Estimates (Alternative specification)

	Est.
θ_k (i_{jt})	0.294
	(0.077)
$\widetilde{\theta}_a$ (i_{jt}^2)	0.296
	(0.308)
θ_a (i_{jt}^2/k_{jt})	1.161
	(0.347)

Notes:
Standard Error of the estimate is shown inside parenthesis.
The number of the samples used in the estimation is 100.

The result shows that the the estimated value of the parameter $\widetilde{\theta}_a$ is insignificant, and $\widetilde{\theta}_a i_{jt}^2$ does not a play important role in the investment cost function. Also, it implies that smaller firms incur more investment cost, which is consistent with the literature on the macroeconomics literature, and the discussion in Mermelstein et al. (2020). Note that the the results of the evaluations of the 1970 merger and merger remedies did not largely change irrespective of the inclusion of the term $\widetilde{\theta}_a i_{jt}^2$. The results are available upon request.

References

Aguirregabiria, V. and Marcoux, M. (2021). Imposing equilibrium restrictions in the estimation of dynamic discrete games. *Quantitative Economics*, 12(4):1223–1271.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Barzilai, J. and Borwein, J. M. (1988). Two-point step size gradient methods. *IMA journal of numerical analysis*, 8(1):141–148.

Benkard, C. L., Bodoh-Creed, A., and Lazarev, J. (2020). Simulating the dynamic effects of horizontal mergers: Us airlines.

Bennato, A. R., Davies, S., Mariuzzo, F., and Ormosi, P. (2021). Mergers and innovation: Evidence from the hard disk drive market. *International Journal of Industrial Organization*, 77:102755.

Chen, J. (2009). The effects of mergers with dynamic capacity accumulation. *International Journal of Industrial Organization*, 27(1):92–109.

Chen, Y. and Xu, D. (2022). A structural empirical model of r&d investment, firm heterogeneity, and industry evolution. Technical report, National Bureau of Economic Research.

Cooper, R. W. and Haltiwanger, J. C. (2006). On the nature of capital adjustment costs. *The Review of Economic Studies*, 73(3):611–633.

Dertwinkel-Kalt, M. and Wey, C. (2016). Merger remedies in oligopoly under a consumer welfare standard. *The Journal of Law, Economics, and Organization*, 32(1):150–179.

Ericson, R. and Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62(1):53–82.

Farrell, J. and Shapiro, C. (1990). Horizontal mergers: an equilibrium analysis. *The American Economic Review*, pages 107–126.

Friberg, R. and Romahn, A. (2015). Divestiture requirements as a tool for competition policy: A case from the swedish beer market. *International journal of industrial organization*, 42:1–18.

Genakos, C., Valletti, T., and Verboven, F. (2018). Evaluating market consolidation in mobile communications. *Economic Policy*, 33(93):45–100.

Goettler, R. L. and Gordon, B. R. (2011). Does amd spur intel to innovate more? *Journal of Political Economy*, 119(6):1141–1200.

Grajek, M., Gugler, K., Kretschmer, T., and Mişcişin, I. (2019). Static or dynamic efficiency: Horizontal merger effects in the wireless telecommunications industry. *Review of Industrial Organization*, 55(3):375–402.

Hashmi, A. R. and Biesebroeck, J. V. (2016). The relationship between market structure and innovation in industry equilibrium: a case study of the global automobile industry. *Review of Economics and Statistics*, 98(1):192–208.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica: Journal of the Econometric Society*, pages 213–224.

Igami, M. and Uetake, K. (2019). Mergers, Innovation, and Entry-Exit Dynamics: Consolidation of the Hard Disk Drive Industry, 1996-2016. *The Review of Economic Studies*, (0):1–31.

Judd, K. L., Maliar, L., Maliar, S., and Valero, R. (2014). Smolyak method for solving dynamic economic models: Lagrange interpolation, anisotropic grid and adaptive domain. *Journal of Economic Dynamics and Control*, 44:92–123.

Jullien, B. and Lefouili, Y. (2018). Horizontal mergers and innovation. *Journal of Competition Law & Economics*, 14(3):364–392.

La Cruz, W., Martínez, J., and Raydan, M. (2006). Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. *Mathematics of Computation*, 75(255):1429–1448.

Mermelstein, B., Nocke, V., Satterthwaite, M. A., and Whinston, M. D. (2020). Internal versus external growth in industries with scale economies: A computational model of optimal merger policy. *Journal of Political Economy*, 128(1):301–341.

Motta, M. and Tarantino, E. (2021). The effect of horizontal mergers, when firms compete in prices and investments. *International Journal of Industrial Organization*, 78:102774.

Nishiwaki, M. (2016). Horizontal mergers and divestment dynamics in a sunset industry. *RAND Journal of Economics*, 47(4):961–997.

Nocke, V. and Whinston, M. D. (2022). Concentration thresholds for horizontal mergers. *American Economic Review*, 112(6):1915–48.

Ohashi, H. (2005). Learning by doing, export subsidies, and industry growth: Japanese steel in the 1950s and 1960s. *Journal of International Economics*, 66(2):297–323.

Osinski, F. D. and Sandford, J. A. (2021). Evaluating mergers and divestitures: A casino case study. *The Journal of Law, Economics, and Organization*, 37(2):239–277.

Pakes, A. and McGuire, P. (1994). Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model. *The RAND Journal of Economics*, 25(4):555.

Pesendorfer, M. (2003). Horizontal Mergers in the Paper Industry. *The RAND Journal of Economics*, 34(3):495–515.

Smolyak, S. A. (1963). Quadrature and interpolation formulas for tensor products of certain classes of functions. In *Doklady Akademii Nauk*, volume 148, pages 1042–1045. Russian Academy of Sciences.

Tenn, S. and Yun, J. M. (2011). The success of divestitures in merger enforcement: Evidence from the j&j–pfizer transaction. *International journal of industrial organization*, 29(2):273–282.

Varadhan, R. and Roland, C. (2008). Simple and globally convergent methods for accelerating the convergence of any em algorithm. *Scandinavian Journal of Statistics*, 35(2):335–353.

Vergé, T. (2010). Horizontal mergers, structural remedies, and consumer welfare in a cournot oligopoly with assets. *The Journal of Industrial Economics*, 58(4):723–741.

Whinston, M. D. et al. (2008). Lectures on antitrust economics. *MIT Press Books*, 1.

Williamson, O. E. (1968). Economies as an antitrust defense: The welfare tradeoffs. *The American Economic Review*, 58(1):18–36.