CIRJE-F- 1264

Investment Dynamics and Merger Policy:
Long-run Effects of Horizontal Merger in Oligopolistic Market

Takeshi Fukasawa
Waseda University

Hiroshi Ohashi
The University Tokyo

December 2025

Investment Dynamics and Merger Policy: Long-run Effects of Horizontal Merger in Oligopolistic Market*

Takeshi Fukasawa† and Hiroshi Ohashi‡

December 3, 2025

This paper evaluates long-run effects of a horizontal merger and its remedies in a capital-intensive industry. It estimates a dynamic oligopoly model with continuous investment, enabled by a computationally efficient simulation algorithm. A decomposition of investment incentives shows that static forces dominated dynamic considerations. The merger, despite raising prices, increased social welfare, primarily through efficiency gains. The analysis also finds that divestiture remedies had persistent effects lasting nearly two decades, and that optimal remedy design differs markedly depending on whether consumer or social welfare is used as the evaluative standard.

Keywords: Horizontal Merger; Structural Remedies; Dynamic Oligopoly; Spectral Algorithm; Smolyak Method
JEL classification: C63; L13; L41; L61

*We thank Ryutaro Komiya, Yoshiro Miwa, Adrian Pagan, Ivan Png, and the conference and seminar participants at the DBJ, ES Meetings, EARIE, Hitotsubashi, Korean Association of Industrial Organization, Kyoto, NUS, RIETI, SHUFE, and UNSW for their helpful comments. We are grateful to Tsuyoshi Nakamura who generously provided us with the data set and patiently answered our questions on the Japanese steel industry. This is a substantially revised version of the earlier working paper titled "Evaluating the Merger and its Remedies in a Dynamic Environment: Revisiting the 1970 Yawata-Fuji steel merger."(Ohashi et al., 2010)

†Waseda Institute for Advanced Study, Waseda University, Japan.
‡Faculty of Economics, University of Tokyo, Tokyo 113-0033, Japan.

1 Introduction

Traditional antitrust analyses of horizontal mergers, involving firms operating within the same market, have often emphasized immediate competitive impacts, such as short-run changes in pricing and market shares. However, there is growing recognition among policymakers and competition authorities of the crucial importance of mergers' effects on firms' investment behavior. Investment decisions, particularly in capital-intensive industries, have significant impacts on long-term market structure, competitive dynamics, and firm performance, thereby becoming an essential area of focus in merger assessments.

Recent antitrust cases illustrate regulators' increasing attention to how mergers affect corporate investment incentives. For instance, competition authorities scrutinized the merger between telecom giants Sprint and T-Mobile, focusing on its potential implications for future investments in network infrastructure. Similarly, the recent merger case involving UK mobile network operators Vodafone and Three resulted in regulatory approval conditional upon substantial investment commitments throughout the UK over the next several years. These examples highlight a pivotal role of investment incentives in evaluating merger proposals and in formulating effective merger remedies.

One critical yet underexplored dimension of horizontal mergers is their long-term impact on firms' investment behavior, which are often endogenously determined and central to influencing future market dynamics. While the static effects of mergers have been extensively studied, our understanding of how mergers influence long-term investment incentives remains limited. Addressing this gap is crucial for both competition authorities and merging parties seeking to design remedies that preserve competition while enabling potential efficiency gains.

This paper quantitatively assesses the long-run effects of a horizontal merger and its associated remedies by explicitly modeling firms' endogenous investment decisions. We compute a Markov-perfect equilibrium of a dynamic oligopolistic investment game that captures firms' forward-looking and strategic behavior. Our model builds on the dynamic industry framework pioneered by Pakes (1994) and Ericson and Pakes (1995), with a modification inspired by Mermelstein et al. (2020). In their calibration exercise, Mermelstein et al. (2020) advance beyond unitary capital additions by allowing firms to undertake multiple discrete investments. We extend this approach further by modeling continuous investment choices, which provides a more realistic representation of the capital-intensive industry in our data. Unlike Mermelstein et al. (2020), whose framework is designed for calibration, our setting requires full structural estimation. To make this feasible, we develop a computationally efficient algorithm that avoids the heavy burden typically associated with solving optimal continuous investment decisions at each iteration.

We apply our framework to a historical case study of a horizontal merger in the Japanese steel industry, accompanied by structural remedies. During the post-war period, the Japanese steel industry was characterized by sustained demand growth and large-scale investment in production capacity. With this historical context, we quantify the long-run effects of the merger and the effectiveness of the remedies on major economic indicators, explicitly accounting for dynamic interactions between investment decisions and market competition. We furthermore decompose the investment incentives of both merging and non-merged firms to identify the mechanisms driving post-merger investment behavior. Finally, we conduct counterfactual simulations to evaluate optimal remedy designed under two standards of consumer welfare and social welfare.

Our analysis shows that, despite raising steel prices, the merger improved overall social welfare. This welfare gain is primarily attributable to efficiency gains realized by the merged firm, which significantly reduced its production costs and increased producer surplus. At the same time, the merger altered firms' dynamic investment incentives in asymmetric ways. While the merged firm reduced its investment following the merger, non-merging rivals expanded theirs in response. These endogenous reactions played a central role in determining long-run outcomes.

Specifically, the increased investment by non-merging firms partially offset the merged firm's decline in capital accumulation, fostering a more balanced distribution of productive capacity across firms. This convergence in firm size strengthened competition, as smaller firms – facing higher marginal returns to investment – expanded output and reduced costs over time. Although aggregate industry investment decreased slightly, the reallocation of investment activity toward smaller, more constrained rivals enhanced productive efficiency and competitive pressure in the long run.

To examine the mechanisms behind these effects, we further decompose firms' investment incentives. Our results indicate that changes in output decisions – captured by the "margin expansion effect" coined in Jullien and Lefouili (2018) – continue to influence investment incentives in our dynamic setting, However, additional forces that do not arise in static models, such as scale economies in capital, also influence marginal investment. Overall, we find that the merger reduce the merged firm's incentive to invest, while non-merging firms expand their investment in response.

Turning to the merger remedies, our simulation suggest that, while implemented remedies improved consumer welfare, they failed to fully offset the loss in producer surplus. Moreover, we find that roughly half of the remedies' effects remained even 10 years after implementation, although their impact gradually diminished over time. These long-run effects arise from the presence of investment adjustment costs. In a frictionless environment, firms would instantaneously adjust capital stocks to equate marginal returns with the user cost of capital, causing the remedies' impact to disappear after the initial period. In contrast, with the estimated adjustment costs, investment responses unfold gradually, leading to effects that persist – albeit declining magnitude – over time.

We also examine the design of optimal merger remedies under alternative welfare standards. First, we show that remedies based on long-run consumer welfare standard must be larger than those derived under a short-run standard. This is because firms endogenously adjust investment over time, eroding the initial effects of the remedies. Consequently, more extensive intervention is required to sustain consumer welfare over the long run – unless competition authorities intervene repeatedly after the merger.

Second, we demonstrate that the optimal allocation of divested assets differs substantially depending on whether the remedies aim to maximize consumer or total welfare. While the consumer welfare standard – commonly used by competition authorities – favors reallocating assets to smaller firms to enhance symmetry in firm size and thus competition (as in Vergé, 2010), this approach may not be optimal from a total welfare perspective. In the historical context of the 1970 merger, our simulations indicate that assigning assets primarily to the smallest firm, Kobe Steel, is optimal under the consumer welfare standard. However, doing so intensifies competition and reduces producer surplus, potentially lowering total surplus. Our findings highlight the importance of considering producer surplus and long-term dynamic effects when designing remedies, under total welfare standard.

1.1 Literature review

This paper contributes to three strands of literature: (i) merger evaluations that incorporate firms' investment decisions, (ii) merger remedies in the long-run investment, and (iii) quantitative methods for solving dynamic oligopolistic investment models.

1.1.1 Merger and Investment Incentives

A growing theoretical literature explores how mergers influence firms' investment incentives, though most analyses remain within static framework. Motta and Tarantino (2021) study cost-reducing (process) innovation in a differentiated-product Bertrand setting and show that, absence efficiency gains, mergers reduce the incentives of the merging firms to invest, since profit margins expand but the incremental returns to investment fall. Bourreau et al. (2024) extend the analysis to demand-enhancing (product) innovation, finding that the effect of mergers is theoretically ambiguous, depending on the balance between incentives to capture new demand and incentives to soften competition.

Empirical work, by contrast, has often relied on reduced-form approaches at the industry level. Bennato et al. (2021) document increased R&D following mergers in the hard disk drive industry. In telecommunications, Genakos et al. (2018) show that per-firm investment is higher in more concentrated markets, though the aggregate impact remains unclear, while Grajek et al. (2019) find that firms increase investment in markets where post-merger prices rise but reduce it when prices fall. These studies underscore that merger-investment relationships are highly context-dependent.

Structural or numerical approaches are comparatively rare. Pesendorfer (2003) provides a one-shot investment model linking with merger waves in the U.S. pulp and paper industry, finding that mergers primarily raised efficiency by reallocating production across plants. Mermelstein et al. (2020) develops a calibrated model to show that mergers can discourage the merged firm's investment. Nishiwaki (2016), in a structural study of the Japanese cement industry, finds that mergers accelerated divestment of service stations relative to non-merged rivals. Chen (2009) similarly reports negative merger effects on investment in a numerical framework.

This paper contributes to the literature in two ways. First, we develop a fully dynamic empirical model of the merger in the Japanese steel industry. Embedding the merger in a dynamic oligopolistic framework allows us to capture intertemporal strategic interactions that static models do not accommodate and to quantify the merger's effects on both merging and non-merging firms' investment incentives. This setup also enables an evaluation of the long-run implications of the merger and the associated remedies.

Second, on the investment side, we allow firms to choose continuous investment levels in a dynamic environment. This feature better reflects firms' decision spaces and permits a more detailed decomposition of merger-induced investment incentives than models that restrict firms to discrete, unitary investment increments (as in Mermelstein et al., 2020, Nishiwaki, 2016).

1.1.2 Merger Remedies and Long-Run Investment Dynamics

This paper also contributes to the literature on merger remedies. Existing theoretical and empirical studies have examined remedies with a short-run focus, often abstracting from firms' endogenous investment responses. Theoretical contributions include Vergé (2010), Dertwinkel-Kalt and Wey (2016), and Nocke and Rhodes (2025), while reduced-form empirical work such as Tenn and Yun (2011), Friberg and Romahn (2015), and Osinski and Sandford (2021) evaluates divestitures and other remedies. Empirical evidence implies that remedies can meaningfully alter firms' investment behavior.

Our study builds on this insight by explicitly modeling remedies as the reallocation of capital from merging to non-merging firms in a dynamic oligopolistic investment framework. Extending Nocke and Rhodes (2025), we evaluate optimal remedies and demonstrate that accounting for endogenous investment responses is crucial. Under the consumer welfare standard, the optimal long-run remedy is substantially larger than what static or short-run analyses would suggest. Ignoring investment dynamics therefore risks underestimating the remedy needed to protect competition. We also examine optimal remedies under a total welfare standard and find that the preferred allocation of divested assets differs markedly from the consumer-welfare case.

1.1.3 Estimating Dynamic Oligopoly Models of Mergers

Empirical studies that estimate dynamic oligopoly models incorporating mergers remain rare, primarily because solving multi-agent dynamic games is computationally demanding. Igami and Uetake (2019) estimate a dynamic model of entry and exit in the Japanese hard disk drive industry under a nonstationary environment, allowing for mergers to influence market structure over time. Jeziorski (2014) develops an empirically estimated dynamic merger model, applied to the U.S. radio industry, in which firms endogenously decide to merge, enter, or exit. Benkard et al. (2020) study the U.S. airline industry modifying a two-step method proposed by Bajari et al. (2007) that quantifies the long-run effects of mergers without requiring the explicit computation of equilibrium. These studies highlight the potential of dynamic frameworks to capture intertemporal efficiency and competition effects that static models cannot.

Our work extends this literature by estimating a dynamic oligopoly model with continuous actions (investment) and continuous state variables (capital stocks and productivity shocks) in a multi-agent environment. Most existing IO applications, including Goettler and Gordon (2011), Ryan (2012), Fowlie et al. (2016), and Barwick et al. (2025), rely on the Pakes and McGuire (1994) algorithm, where firms' actions are updated as a solution to nested nonlinear optimization problem given competitors' strategies and current value functions. Though widely used, this procedure is computationally intensive because each iteration requires solving high-dimensional optimization problems – a challenge amplified in models with continuous controls of multiple agents.

We propose an alternative computational strategy motivated by recent advances in reinforcement learning and numerical analysis, as proposed by Fukasawa (2025). The approach combines the policy-gradient method (e.g., Silver et al., 2014) with the spectral algorithm developed in applied mathematics and recently employed in economics by Aguirregabiria and Marcoux (2021). The algorithm iteratively updates each firm's investment decision in the direction of the gradient of its long-run (present-discounted) profit, updating value functions through the Bellman equation until convergence.[1] This approach avoids computing optimal investment given competitors' investment in each iteration, because each update does not involve nonlinear root-finding procedures – substantially reducing computational cost – while maintaining convergence speed and stability via the spectral algorithm. We further integrate the Smolyak sparse-grid method (Smolyak, 1963; see also Judd et al., 2014) to approximate value functions efficiently and mitigate the curse of dimensionality.

Recent methodological advances also point to related directions. Gowrisankaran and Schmidt-Dengler (2024) propose a computational method for dynamic oligopoly models with private information shocks to marginal costs; however, their framework requires discretizing investment choices – a limitation for our analysis, which requires continuous differentiation to decompose investment incentives. In macroeconomics, techniques such as the endogenous grid-point method (Carroll, 2006), the envelope condition method (Maliar and Maliar, 2013), and Euler-equation-based approaches (Judd, 1998) have improved computational efficiency in single-agent settings. Nevertheless, their applicability to multi-agent dynamic games with strategic interdependence, such as our model, is not straightforward.

By building on these developments, our framework provides a computationally tractable way to

[1] The existing literature on continuous-action, continuous-state multi-agent dynamic models typically relies on cubic spline approximations to represent firms' value of profit functions. However, such approaches are not directly applicable in our settings, as the required dense grid evaluation would be computationally prohibitive and incompatible with the continuous differentiation used in our decomposition analysis.

estimate dynamic merger models with continuous investment behavior. This allows us to analyze not only the equilibrium outcomes of mergers but also the dynamic adjustment paths of firms' investment, capacity, and welfare-dimensions that are crucial for understanding long-run policy implications and remedy design.

The rest of this paper is organized as follows. Section 2 provides an overview of the post-war Japanese steel industry, with a particular attention to the 1970 merger that forms the focus of our analysis. The section also presents a static evaluation of the merger as a preliminary step toward the dynamic structural analysis that follows, establishing the benchmark against which endogenous investment responses are later compared,

Section 3 introduces a structural model with capital accumulation. The framework consists of two components: (i) a static model of demand and marginal cost in steel production, and (ii) a dynamic investment model that characterizes firms' forward-looking behavior in an oligopolistic environment. Section 4 describes the computational algorithm used to solve the model's Markov-perfect equilibrium. We implement a policy-gradient-based procedure that iteratively updates firms' investment policies and value functions, enabling the computation of equilibrium dynamics. Section 5 presents the estimation results and discusses the empirical implications of the estimated parameters. Building on these estimates, Section 6 conducts counterfactual policy simulations to quantify the economic consequences of the 1970 horizontal merger. Section 7 examines the competitive and investment effects of the merger, while Section 8 assesses the effectiveness of divestiture as a merger remedy and evaluates the optimal remedy design under both consumer- and social-welfare standards. Section 9 concludes. The appendices document the data sources, describe the computational procedures, and provide additional analytical details.

2 Overview of the Industry

This section provides a historical overview of the Japanese steel market, with a particular focus on the period from 1960 to 1990. The merger under the focus of this study took place in 1970. During the study period, defining features of the industry was active and sustained investment of individual firms. These characteristics of the market described in Section 2.1 serve as our motivation for the development of the dynamic structural model discussed in Section 3.

Before introducing the dynamic framework, however, Section 2.2 employs a static analytical approach based on market concentration, abstracts from firms' endogenous investment behavior. This approach provides a simplified description of market outcomes by focusing on short-term competitive effects, without incorporating long-term strategic investment responses that are central to our dynamic analysis developed in later sections.

2.1 Brief Historical Background

The two merging companies examined in this study, Yawata and Fuji Steels, trace their origins back to a single entity named Japan Steel, which was dissolved by the occupation forces after World War II.[2] Over the 1960 - 1990 period, seven integrated steelmakers, producing crude steel from iron ore and coking coal – accounting for roughly 80% of domestic steel production. These firms form the primary focus of our analysis.[3] Between 1960 and 1990, there were no significant entries or exits in the industry, except for the Yawata-Fuji merger of our study. Three key characteristics define

[2]Japan Steel merged with U.S. Steel in June 2025.

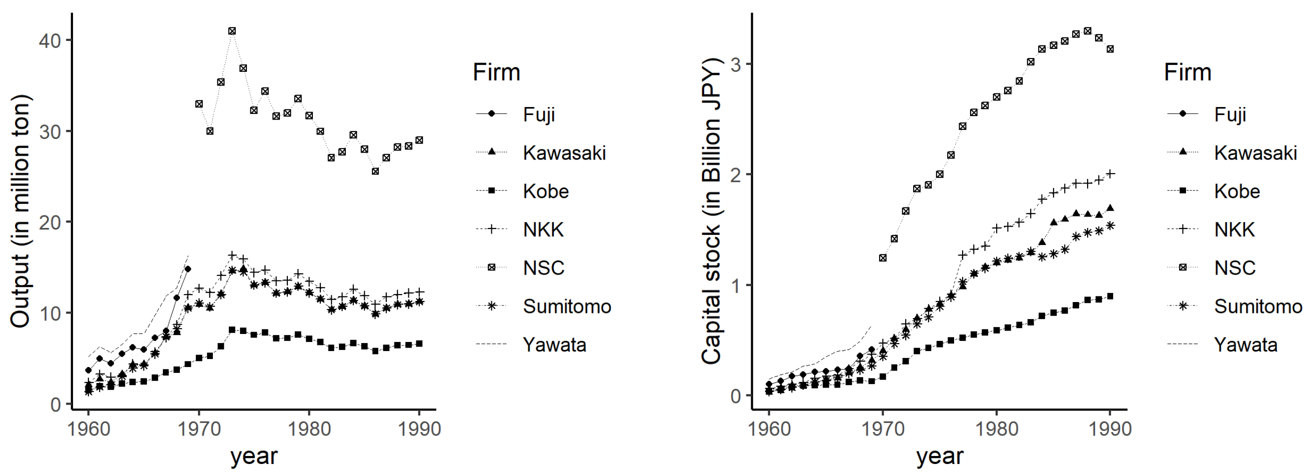
[3]The seven companies are in order of average market share: Nippon Steel, Yawata Steel, Fuji Steel, Nihon Kokan, Sumitomo Steel, Kawasaki Steel, and Kobe Steel.

the Japanese steel industry during this period; investment, minimills, and international trade. We discuss each below.

Investment The industry was marked by active investment (as shown in Figure 1), particularly in new plants and facilities to meet the growing demand in both domestic and international markets.[4] These new production often incorporated the latest technologies, allowing for more efficient production processes. Notably, Table 1 shows that non-merging firms tended to increase their investments in the post-merger period compared to the pre-merger period. In Section 3, we introduce a dynamic decision-making model to capture this pattern, incorporating firms' endogenous investment choices. We assume 1990 as the terminal period in the model, given that firms' investment activities plateaued around that time.

Figure 1: Firms' outputs and capital stocks: 1960–1990

(a) Output (b) Capital stock



Notes: The figure illustrates the evolution of output and capital stock for Japan's major integrated steelmakers from 1960 to 1990. Panel (a) shows annual crude steel output (in million tons), and Panel (b) depicts firms' capital stocks (in billion JPY, constant 1960 prices). The merged firm, Nippon Steel (NSC), was formed from the 1970 merger between Yawata and Fuji Steels. Other firms include Nihon Kokan, Sumitomo, Kawasaki, and Kobe Steels.

Minimills While integrated steel makers dominated the domestic market during the study period, minimills gradually increased their market share. Minimills produce crude steel from scrap using electric arc furnaces, which require a simpler production process and a smaller minimum efficient scale of production, compared to integrated steelmakers.

Crude steel serves as a fundamental input for manufacturing a wide range of steel products, including railway rails and construction materials. Although crude steel must undergo further processing (e.g., rolling) before it can be sold as a finished good, it remains the essential building block for thousands of steel items. Since our study focuses on investments directly related to expanding crude-steel production capacity, and given the availability of comprehensive production and price data for crude steel in Japan, we treat crude steel as the final good in our analysis.

Although both integrated steelmakers and minimills produce crude steel, the output from integrated producers is generally of higher quality and technological sophistication. High-grade

[4] 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, consistent with active investment in the Japanese steel sector at that time.

steel from integrated firms is widely used in advanced applications – such as light-weighted high-strength steel for automobile bodies – that significantly improve vehicle safety and fuel efficiency. Accordingly, in our demand model, we treat crude steel from integrated steelmakers and minimills as imperfect substitutes. Consistent with this modeling assumption, the estimation results in Section 5 confirm that the two types of crude steel are indeed differentiated products rather than perfect substitutes.

Table 1: Summary statistics for important variables: 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: Values are averaged over the study period. Industry output is measured in million tons. Prices are in JPY per thousand tons. Market shares are expressed in percentages. Capital stock is reported in billion JPY, adjusted to 1960 constant prices. Standard errors are shown in parentheses. The merging party refers to Nippon Steel, while the non-merged party includes Nihon Kokan, Sumitomo, Kawasaki, and Kobe.

International Trade The rapid growth in production, as highlighted in Table 1 was accompanied by a significant expansion of exports. Japan's share of the global steel export market rose from less than 5% in 1955 to 9% in 1965. Initially, most of Japan's steel exports were destined to Asian markets, but by the early 1960s, an increasing proportion was being shipped to North America. It is important to note that the international steel export market was highly competitive between 1955 and 1990, with little evidence that Japanese steelmakers possessed market power during this period.[5]

Japan maintained a 15% import tariff on steel until 1967, when it agreed to halve the rate during the Kennedy Round of the General Agreement on Tariffs and Trade (GATT). Although the import tariff provided some protection for domestic steelmakers against foreign competition, it likely had little impact on the growth in Japanese steel production, given that Japan was also a significant steel exporter during this period. In fact, steel imports accounted for an average of only 0.2% of the domestic market, even after the tariff reduction.

[5]For example, the Japan Iron and Steel Exporters' Association (1974) reported that the freight-on-board (FOB) price of Japanese steel was not significantly different from the price in Antwerp, a key hub of the global steel trade at the time. This finding aligns with Ohashi (2005), which indicates that Japanese steel export subsidies were not designed for profit-shifting purposes.

Table 2: Market share of steel production: impact of the 1970 merger

1969	% (Output)	1970	% (Output)
Yawata (merging)	23.73	Nippon Steel (NSC)	45.33
Fuji (merging)	21.65		
Nihon Kokan (nonmerging)	17.5	Nihon Kokan	17.45
Sumitomo (nonmerging)	15.34	Sumitomo	15.22
Kawasaki (nonmerging)	15.43	Kawasaki	15.05
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

Notes: The table reports market shares of crude steel production for major integrated steelmakers in 1969 (before the merger) and 1970 (after the merger). Market shares are calculated based on firms' annual crude steel output. The HHI is calculated using firm-level market shares expressed in percentages. The merged entity, Nippon Steel, represents the combination of Yawata and Fuji Steels in 1970.

2.2 Evaluating the 1970 Merger Using Static Concentration Measures

This subsection assesses the competitive effects of the 1970 merger using a static analytical framework based on market concentration. We begin with a conventional approach commonly employed by competition authorities, which evaluates mergers through concentration measures such as Herfindahl-Hirschman Index (HHI). According to the U.S. Department of Justice and the Federal Trade Commission's *Horizontal Merger Guidelines* (1992), a merger is presumed likely to enhance market power when the post-merger HHI exceeds 1800 and the increase in HHI (ΔHHI) exceeds 100 points.

Applying these criteria to the 1970 merger in the Japanese steel industry, we find that the post-merger HHI surpasses 1800, with an increase greater than 100 points (Table 2). Under the U.S. Merger Guidelines, such a merger would typically trigger antitrust concern and could be subject to challenge unless compelling evidence of substantial efficiency gains or low entry barriers exists.

In the related vein, Nocke and Whinston (2022) propose a more refined measure – derived from a homogeneous-good Cournot model – to quantify the minimum efficiency gains required for a merger not to increase equilibrium prices under a consumer welfare standard. The required cost reduction threshold is expressed as:

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

where c_M denotes the output-weighted average marginal cost of the merging firms before the merger, $\overline{c}_M$ represents the marginal cost of the merged firm, s_M is the combined market share of the merging firm, and ϵ is the (positive) price elasticity of demand. Using industry data, we calculate that the minimum efficiency gain necessary to prevent a price increase is approximately 27.5%. As will be discussed in Section 5, our structural estimates suggest that the actual efficiency gains realized from the 1970 merger were around 20%. This implies that, while the merger achieved notable cost reductions, the magnitude of these efficiencies was insufficient to fully offset the loss in consumer welfare predicted by the static framework, which is consistent with our simulation results discussed in Section 6.

From the perspective of total surplus, another practical measure is derived by Farrell and Shapiro (1990), providing sufficient conditions for a merger to increase total welfare. Suppose that the set of firms $\mathcal{M}$, representing the merging parties before the merger, are considering consolidation. Let q_j denote firm j's output, and Q represent total industry output. Assuming that the proposed merger is profitable for the merging firms, and abstracting from endogenous investment decisions, a sufficient condition for the merger to increase total surplus is expressed as (e.g., Whinston, 2008):

$$s_M \quad < \quad -\sum_{j \notin \mathcal{M}} s_j \left(\frac{dq_j}{dQ} \right), \tag{2}$$

where s_j is firm j's pre-merger market share, and $\frac{dq_j}{dQ}$ is the marginal change in output of non-merging firm j in response to a marginal change in total industry output. Based on industry data, Eq.(2) is violated in the case of the 1970 merger: the left-hand side of the equation is 0.45, while the right-hand side is 0.38. [6] This violation indicates that there is no guarantee the merger increased total welfare.

While these methods offer convenient and accessible tools for merger analysis, they overlook firms' endogenous investment decisions. Such decisions can influence not only the marginal costs of the merging firms but also those of non-merging rivals,[7] thereby determining both consumer and social welfare. Indeed, as the structural estimation and simulation results discussed in Section 6 below reveal, the merger reduced consumer welfare but increased total welfare – findings that align with the static analysis for consumer surplus, yet fall short of satisfying the sufficient conditions for overall welfare improvement.

Additionally, the static methods discussed in this section do not provide a clear framework for evaluating the effectiveness of merger remedies. This limitation arises because the merger and the implementation of remedies typically occur simultaneously, making it difficult to disentangle their separate effects. Identifying these impacts empirically would require counterfactual analysis to isolate the causal influence of remedies from that of the merger itself. In the subsequent sections, we develop a dynamic structural model that explicitly incorporates firms' investment decisions and enables a unified evaluation of both merger effects and the effectiveness of associated remedies.

3 The Model and Identification

This section develops a dynamic model designed to evaluate the welfare effects of the 1970 merger and its associated remedies, explicitly capturing the dynamics of firms' oligopolistic investment

[6] To calculate these values using the methods of Nocke and Whinston (2022) and Farrell and Shapiro (1990), we used demand elasticity estimates derived from Section 5.

[7] In the approach proposed by Nocke and Whinston (2022), it is implicitly assumed that non-merging firms' marginal costs remain constant before and after the merger. However, when firms engage in endogenous investment, the marginal costs of non-merging firms may change over time, potentially altering the conclusions drawn from static analyses.

behavior. We begin with an overview of the model and then describe its key components: demand and supply, the latter encompassing firms' output and investment decisions. For each component, we also discuss the identification strategy, clarifying how the model parameters can be inferred from observed data.

3.1 Model Setup

As described in Section 2.1, the Japanese steel market from 1960 to 1990 was characterized by active investment among a small number of major integrated steel producers, operating under competitive pressure from minimills. Our model focuses on integrated steelmakers; they produced crude steel and capital investment played a critical role in enhancing production efficiency.

The timing of the game is structured as follows. At the beginning of each year t, where t ranges from 1960 to 1990 in our study, integrated steel firm $j \in \mathcal{J}$ observes industry state, s_t, and independently makes its investment decision. Note that $\mathcal{J}$ denotes the set of firms (thus, $\mathcal{J} \supset \mathcal{M}$). The industry state at each period is summarized by a state vector s_t, which includes commonly observed variables that affect both demand and production costs. A detailed definition of the state variable is provided in Section 3.4. We assume that investment decisions are capitalized at the end of the year, with the transition of capital stock governed by $k_{j,t+1} = \delta k_{jt} + i_{jt}$, where δ is a depreciation rate of the capital stock and i_{jt} represents firm j's investment in period t.

Following the investment choice, firm j decides crude steel output, considering its own marginal cost, and its current capital stock. No significant entries were observed in the Japanese domestic steel market during the study period, so we treat the number of firms, J exogenously given. Specifically, $J = 6$ before 1970, and $J = 5$ in the post-merger period. Following the existing literature, we model firms' output choices using a static decision-making framework, as detailed in Section 3.3.

3.2 Demand

As discussed in Section 2.1, we need to account for the growing prominence of minimills, when we focus on integrated steelmakers during the study period. These minimills, which primarily use electric furnaces, are treated as a competitive fringe in the crude steel market. Although both integrated steel producers and minimills produce crude steel, the output from integrated steelmakers is generally perceived to be of higher quality. To capture demand for the two types of crude steel, we model a representative consumer who maximizes a Constant Elasticity of Substitution (CES) utility function U_t over consumption from integrated producers and minimills, subject to the the budget constraint, Y_t:

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

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

where Q_t^B and Q_t^E denote the quantities of crude steel consumed from integrated steel producers and minimills, respectively, while P_t^B and P_t^E are the corresponding prices at time t. The parameter σ is the elasticity of substitution, which governs how readily consumers substitute between the two types of crude steel in response to relative price changes. $\xi_t \in (0, 1)$ reflects the representative consumer's relative preference for steel from integrated producers, and is allowed to vary over time to account for shifts in perceived quality.

We treat minimills as competitive fringe whose pricing behavior is modeled exogenously – specifically we assume that P_t^E is taken as given. Solving the consumer's problem yields the following log-difference demand equation:[8]

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

where S_t^B and S_t^E denote expenditure shares of crude steel produced by integrated steelmakers and minimills respectively. Defining $\sigma \log\left(\frac{\xi_t}{1-\xi_t}\right) = b + u_t$ and letting $P_t \equiv \frac{P_t^B}{P_t^E}$, we obtain the estimable demand equation:

$$\log S_t^B - \log S_t^E = -\sigma \log(P_t) + b + u_t. \tag{4}$$

To address potential endogeneity in prices, we use the logarithms of input prices – specifically, iron ore and heavy oil – as instrumental variables. Estimation is conducted using the Generalized Method of Moments (GMM), enabling us to obtain consistent estimates of the set of demand parameters, $\Theta^d \equiv (b, \sigma)$.

3.3 Output Choice

We begin by modeling the steel production process. Due to limited availability of firm-level factor input data, we construct a cost function that reflects the steel-making process. We assume that an increase in firm's capital reduces its marginal cost of production. This assumption is reasonable, as capital investments in the steel industry primarily involve upgrading furnaces with newer, cost-reducing technologies. Consequently, an increase in k_{jt} is likely to enhance production efficiency. Since firm's investment is capitalized at the end of each period, we model firm j's marginal cost at time t, mc_{jt}, as follows:

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

where c_j and c_t represent firm-specific and time-specific cost components, respectively, allowing for asymmetry between firms. The last term on the RHS represents a mean-zero i.i.d error term. The nature of steel production suggests that the cost-reducing technologies used by firm j are not transferable to other firms, as they are embedded in the firm's proprietary furnaces.[9]

[8]The expenditure shares of crude steel produced by integrated steelmakers and minimills are expressed as follows:

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

By differencing the logarithms of the market shares, we obtain Eq.(3).

[9]Ohashi (2005) examines steel production technology during the same period and finds little spillover between firms.

We did not include steel output in Eq.(5) as a proxy for increasing returns to scale or capacity utilization. In our preliminary analysis, we found that including output variables –whether in logarithmic form or interacted with capital stock to capture capacity utilization – did not yield meaningful estimates. This might be likely because output closely tracks capital stock during the period, which drove the investment in the industry through 1990.

Direct cost data is not readily available. We thus estimate price-cost margins by constructing a competition model and solving the firm's profit maximization problem. Specifically, in each period, before choosing its investment level, integrated steelmaker j observes own and rivals' current capital stocks, $k_t \equiv (k_{jt})_{j \in \mathcal{J}}$ and the set of shocks, $\eta_t \equiv (\xi_t, P_t^E, Y_t, c_t)$, and simultaneously chooses its output q_{jt} to maximize the per-period profit:[10]

$$\pi_{jt}(s_t, \Theta) = \left(P_t^B - mc_{jt}\right) q_{jt}, \tag{6}$$

where $\Theta \equiv (\Theta^d, \Theta^c)$, and the set of estimable cost parameters is $\Theta^c \equiv \left(\gamma, (c_t)_{\forall t}, (c_j)_{\forall j}\right)$. Under the assumption in Section 3.1, steel output and price are determined in a static equilibrium conditional on the current state $s_t \equiv (k_t, \eta_t)$. Thus, the maximized per-period profit for firm j is a function of the current state vector. The first-order condition for firm j's profit maximization under Cournot competition yields the familiar form of the Lerner index:[11]

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

Note that $Q_t^B = \sum_{j=1}^J q_{jt}$, and ϵ_t is the price elasticity of demand, given the supply curves of price-taking minimills. Using the demand estimates, we can obtain mc_{jt} from Eq.(7), and use it to estimate Eq.(5).

3.4 Firm's Investment

At the beginning of time t, each firm observes the state s_t and makes its investment decision. This decision is inherently dynamic, as according to Eq.(5), a firm's investment today results in future efficiency gains through cost reduction. Firm j is assumed to maximize its expected future profit:

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

The expectation is taken over other firms' current and future investment choices, as well as the future values of all state variables. The per-period profit, $\pi_{jt}(s_t; \Theta)$, was defined earlier. All firms discount future profits using a common discount factor, β. As we do not observe entry or exit during the study period, we abstract entry and exit in this model. Investments incur costs modeled as:

[10] We discuss the reasoning behind excluding e_{jt} from η_t in Section 3.4.

[11] In Section 5.2, we discuss alternative mode of competition – assuming cartel behavior instead of Cournot competition.

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

where $\theta \equiv (\theta_k, \theta_a)$ are parameters to be estimated. The functional form on investment costs at the firm level implies that the marginal cost is increasing in investment, i_{it} at a given capital stock, and that the investment costs decreases with the capital stock, given the investment level. These two features are consistent with the literature, including the recent work of Mermelstein et al. (2020). Note that the second term of the right-hand side of Eq.(9) captures the adjustment costs of investment. As shown in the estimates reported in Section 5, smaller firms face higher adjustment costs – reflecting factors such as workforce relocation and financial constraints – consistent with findings in the literature (e.g., Hayashi, 1982, Cooper and Haltiwanger, 2006). In Section 5, we estimate more flexible functional specifications to test the robustness of our baseline estimates.

Following Ericson and Pakes (1995), we assume a pure Markov perfect Nash equilibrium (MPNE). We focus on pure strategies and do not consider mixed strategies. The MPNE consists of a set of best-response investment strategies. An equilibrium is assumed to exist, and we confirm the existence by numerically solving the model. A Markov strategy for firm j is defined as a function of the state variable $s_t = (k_t, \eta_t)$, and we assume perfect foresight regarding the transition of s_t.[12] We also assume that 1990 is the terminal period, with no further investments beyond this point.

The value function V_{jt} represents the discounted sum of firm j's future profits at the beginning of time t, and can be decomposed into the pre-period profit and the continuation value. A firm optimizes its investment decision by equating the marginal cost of investment today with the present value of the future benefits from cost reduction. The firm bases its investment strategy on the current state variables, and under MPNE, we can rewrite the value function in Eq.(8) in the following recursive form:

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

where $i^*_t(s_t) = (i^*_{jt}(s_t), i^*_{-jt}(s_t))$ and $I_t \equiv \{I_{jt}\} \equiv \{i^*_{jt+\tau}(s_{t+\tau})\}_{j,\tau\geq 0}$ is the set of optimal investment strategies for all firms after time t. Given the assumption that no further investment occurs after time T, the terminal value function is:

$$V_{jT}(s_T) = \frac{1}{1-\beta}\pi_{jT}(s_T; \Theta).$$

We now turn to how to estimate the parameters in the investment cost function.

[12] We exclude the marginal cost shock e_{jt} from η_t, as shown in Section 5.2, the estimation of Eq.(5) fits the data closely, with only small residuals. Hence, incorporating e_{jt} into η_t would have negligible impact on the results.

3.4.1 Estimation Procedure

After estimating the demand and marginal-cost parameters, we proceed to estimate the investment cost parameters, θ, using a full-solution approach. Given the structure of our study – a national market with annual data and a relatively small sample size – we are concerned that the two-step estimation methods such as that proposed by Bajari et al. (2007) may be prone to finite sample biases. Instead, following the approaches of Cooper and Haltiwanger (2006), Goettler and Gordon (2011) and Chen and Xu (2022), we apply the GMM to minimize the distance between moments generated by the model and their empirical counterparts.

Let $\widehat{\Theta}$ represent the estimated values of the demand and marginal cost parameters. The estimation procedure is summarized as follows:

1. **Initialize θ:** Choose the starting values for the investment cost parameters.

2. **Solve for equilibrium:** Given $(\theta,\widehat{\Theta})$, solve for the Markov perfect equilibrium and compute the policy function, $i_{jt}(s_t;\theta,\widehat{\Theta})$, which maps the state variables to the firm's optimal investment decisions. Details on the solution method are provided in Section 4.

3. **Compute model-implied moments:** For each period t, evaluate the policy function at the observed state variables, $s_t^{(data)}$, to obtain the model-implied investment decisions. Use these to compute the model-based moments, $m_{\text{model}}(\theta,\widehat{\Theta})$.

4. **Estimate parameters:** Find the parameter values θ that minimize the GMM objective function:

$$\arg\min_{\theta} \left[m_{model}(\theta,\widehat{\Theta}) - m_{data}\right]' W \left[m_{model}(\theta,\widehat{\Theta}) - m_{data}\right] \tag{11}$$

   where m_{data} represents empirical moments from the observed data, and W is a positive definite weight matrix. We estimate W using the inverse of the covariance matrix of the empirical moments, obtained via a bootsrap procedure.

We use three moments to estimate the investment cost parameters: (i) the average ratio of investment relative to current capital stock, (ii) the average investment level of the merged firm, and (iii) the average investment level of the non-merged firms. Although the parameters enter the model non-linearly, their identification is straightforward. The parameter θ_k, which governs direct investment cost, is primarily identified by variation in investment levels over time. The parameter θ_a, representing investment adjustment costs, is chiefly identified by variation in investment-to-capital ratios.

A major computational challenge in applying the full solution approach and conducting counterfactual simulations is the need to solve for equilibrium outcomes at each candidate parameter values. To address this, we solve the model with continuous investment choices using a Pakes and McGuire (1994)-type algorithm, augmented with several computational techniques: policy gradient-type steps, the spectral method, and the Smolyak algorithm. The specific role of each technique in facilitating the equilibrium solution is discussed in the next section.

4 Algorithm for Solving the Equilibrium

This section outlines the solution method used to compute the model's equilibrium. As briefly described in Section 1.1.3, we apply the algorithm developed by Fukasawa (2025), combined with the Smolyak method, to mitigate the computational cost. We tailor to a discrete-time finite-horizon dynamic model with continuous actions. Below, we summarize the core idea, focusing on the role of policy gradients in investment choice. Detailed explanations of the spectral and Smolyak methods are provided in Appendix B. For notational clarity, we suppress Θ and θ throughout this section.

4.1 Algorithm using Policy Gradient-type Steps

First, the value function satisfies:

$$V_{jt}(s_t) = \pi_{jt}(s_t) - \phi(k_{jt}, i^*_{jt}(s_t)) + \beta E_t V_{jt+1}(k_{jt} + i^*_{jt}(s_t), k_{-jt} + i^*_{-jt}(s_t), \eta_{t+1}). \tag{12}$$

The optimal investment $i^*_{jt}(s_t)$ satisfies the following first-order condition:

$$-\frac{\partial \phi(k_{jt}, i^*_{jt}(s_t))}{\partial i_{jt}(s_t)} + \beta \frac{\partial V_{jt+1}(k_t + i^*_t(s_t), \eta_{t+1})}{\partial i_{jt}(s_t)} = 0. \tag{13}$$

To solve $V_{jt}(s_t)$ and $i^*_{jt}(s_t)$ over all firms $j = 1, \ldots, J$ and years $t = 1970, \ldots, T$, we employ the following iterative procedure. The final year $T = 1990$ is treated as the terminal period with $i_{jT} = 0$.

4.2 Iterative Algorithm

To compute the Markov-perfect equilibrium in our dynamic oligopoly model, we solve for the fixed point of firms' value and policy functions using an iterative procedure. This algorithm incorporates a policy gradient-type update for the investment decision and a spectral acceleration step for improved convergence. The state space is discretized using the Smolyak method to reduce the computational burden. The procedure is as follows:

1. Discretize the state space by selecting grid points $s_t^{(\text{grid})} = \left(\left\{k_{jt}^{(\text{grid})}\right\}_{j \in \mathcal{J}}, \eta_t\right)$ for each t. We reduce the number of grid points using the Smolyak method: 389 for six firms and 241 for five firms per year.

2. Set the initial values for $V_{jt}^{(0)}(s_t^{(\text{grid})})$ and $i_{jt}^{(0)}(s_t^{(\text{grid})})$ for all j, t, and grid points.

3. Iterate the following process until $V_{jt}^{(n)}(s_t^{(grid)})$ and $i_{jt}^{(n)}(s_t^{(grid)})$ converge, where $n \geq 0$ denotes the n-th iteration:

   (a) Interpolate the values of $V_{jt+1}(k_{jt}^{(grid)} + i_{jt}^{(n)}(s_t^{(grid)}), k_{-jt}^{(grid)} + i_{-jt}^{(n)}(s_t^{(grid)}), \eta_{t+1})$ and $\frac{\partial V_{jt+1}^{(n)}\left(k_{jt}^{(grid)} + i_{jt}^{(n)}(s_t^{(grid)}), k_{-jt}^{(grid)} + i_{-jt}^{(n)}(s_t^{(grid)}), \eta_{t+1}\right)}{\partial i_{jt}(s_t^{(grid)})}$, using Chebyshev polynomials and the values of $V_{jt+1}^{(n)}(s_{t+1}^{(grid)})$. Let the interpolated values be $\overline{V}_{jt+1}^{(n)}$ and $\frac{\partial \overline{V}_{jt+1}^{(n)}}{\partial i_{jt}}$, respectively.

   (b) Update the investment policy using a policy gradient-type step:

$$\widehat{i}_{jt}^{(n+1)}(s_t^{(\text{grid})}) = i_{jt}^{(n)}(s_t^{(\text{grid})}) + \lambda \left[-\frac{\partial \phi(k_{jt}^{(\text{grid})}, i_{jt}^{(n)}(s_t^{(\text{grid})}))}{\partial i_{jt}} + \beta \frac{\partial \overline{V}_{jt+1}^{(n)}(k_{jt}^{(\text{grid})} + i_{jt}^{(n)}(s_t^{(\text{grid})}), k_{-jt}^{(\text{grid})} + i_{-jt}^{(n)}(s_t^{(\text{grid})}), \eta_{t+1})}{\partial i_{jt}} \right]$$

   where $\lambda > 0$ denotes a tuning parameter.

(c) Compute the updated value function, the counterpart to Eq.(12):

$$\widehat{V_{jt}}^{(n+1)}(s_t^{(\text{grid})}) = \pi_{jt}(s_t^{(\text{grid})}) - \phi(k_{jt}^{(\text{grid})}, i_{jt}^{(n)}(s_t^{(\text{grid})})) \tag{14}$$

$$+ \beta \overline{V}_{jt+1}^{(n)}(k_{jt}^{(\text{grid})} + i_{jt}^{(n)}(s_t^{(\text{grid})}), k_{-jt}^{(\text{grid})} + i_{-jt}^{(n)}(s_t^{(\text{grid})}), \eta_{t+1}) \tag{15}$$

(d) Apply the spectral acceleration step:

$$i_{jt}^{(n+1)}(s_t^{(\text{grid})}) = \rho_I^{(n)} \widehat{i_{jt}}^{(n+1)}(s_t^{(\text{grid})}) + (1-\rho_I^{(n)}) i_{jt}^{(n)}(s_t^{(\text{grid})}) \tag{16}$$

$$V_{jt}^{(n+1)}(s_t^{(\text{grid})}) = \rho_V^{(n)} \widehat{V_{jt}}^{(n+1)}(s_t^{(\text{grid})}) + (1-\rho_V^{(n)}) V_{jt}^{(n)}(s_t^{(\text{grid})}) \tag{17}$$

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

Step 3(b) follows the policy gradient principle widely used in the reinforcement learning literature: Investment is adjusted in the direction of higher expected value: firms increase investment when marginal returns are positive, decrease it when returns are negative, and leave it unchanged when returns are near zero.

By contrast, the standard method of Pakes and McGuire (1994) updates firms' investment by computing optimal investment given competitors' investment strategies, which is the solution of the following equation[13]:

$$-\frac{\partial \phi(k_{jt}^{(\text{grid})}, \widehat{i_{jt}}^{(n+1)}(s_t^{(\text{grid})})}{\partial i_{jt}(s_t^{(\text{grid})})} + \beta \frac{\partial V_{jt+1}^{(n)}(k_{jt}^{(\text{grid})} + \widehat{i_{jt}}^{(n+1)}(s_t^{(\text{grid})}), k_{-jt} + i_{-jt}^{(n)}(s_t^{(\text{grid})}), \eta_{t+1})}{\partial i_{jt}(s_t^{(\text{grid})})} = 0 \tag{18}$$

This nonlinear equation must be solved numerically for each firm, state, and iteration, resulting in substantial computational burden – especially in high-dimensional settings. Our policy gradient-based algorithm avoids this root-finding step, significantly reducing complexity while preserving convergence properties.

To further accelerate convergence, we incorporate the spectral algorithm, which imposes the efficiency and robustness of the iterative procedure. In particular, it mitigates sensitivity to the learning-rate parameter λ, which governs the update size. As shown in Fukasawa (2025), this modification ensures stable convergence over a wide range of λ values; in our implementation we set $\lambda = 0.1$.

As with the original Pakes and McGuire (1994) algorithm, global convergence is not theoretically guaranteed. Nevertheless, our implementation achieves reliable and stable convergence in practice.

5 Estimation Results

This section applies the empirical framework developed in Section 3 to the Japanese steel industry, using data covering from 1960 to 1990. The pre-merger period spans ten years (1960-1969), and the post-merger period extends over twenty years (1970-1990). While demand and marginal cost parameters are estimated in a static framework, the investment process is modeled dynamically. Because firms may have anticipated the 1970 merger – potentially influencing their investment behavior before its execution – we estimate the investment cost parameters using only post-merger data. In contrast, the estimation of demand and marginal costs employs the full sample period.

[13] The equation is the counterpart of Eq. (13).

We first present the estimates of the demand and marginal-cost functions, followed by those of the dynamic parameters governing investment costs. Table 1 reports summary statistics for the major variables used in the estimation. Details on data sources are provided in Appendix A.

5.1 Demand Estimates

Table 3 presents the estimation results for the demand function. The first column presents the OLS estimates, while the second column provides the GMM estimates using the instrumental variables to address potential endogeneity of price.

The lower panel of the table reports the implied price elasticity of demand for steel produced by integrated steel producers. Comparing the OLS and GMM estimates shows that the GMM procedure effectively corrects the downward bias present in the OLS estimates. The average elasticity implied by the GMM results is approximately -1.06, indicating demand that is close to unit elastic. This estimate is well within the range documented in prior studies of steel demand, including early work by Baker (1989) on the 1930's U.S. steel market and Hogan (1994) in other countries.

Given that weak instruments can undermine the reliability of GMM estimates, we assess the validity of the instrument used in the first stage F-statistic. Conditional on the included exogenous controls, the reported F-statistic is well above the conventional threshold, indicating little concerns about weak instruments. This may confirm the relevance of the instrumental variables used – specifically, input cost shifters such as iron ore and heavy oil prices – in explaining price variation.

Table 3: Demand estimates

	OLS	*GMM*
σ	0.580^{**}	1.116^{***}
	(0.266)	(0.248)
b	0.470^{*}	0.084
	(0.246)	(0.226)
1st stage F-stat.	-	32.057
Elasticity w.r.t. price	-0.80	-1.06

Notes: This table reports the parameter estimates for the demand function specified in Eq.(3). Standard errors are reported in parentheses. The number of observations 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 GMM demand estimates and the first-order condition in Eq.(7), we recover firms' marginal cost of steel production and estimate the cost function specified in Eq.(5). The estimation results are presented in Table 4. Column (1) includes only year-specific components, while Column (2) additionally incorporates firm-specific components.

The parameter γ, which measures the elasticity of marginal cost with respect to capital stock, is negative and statistically significant across both specifications. The estimate under (2) implies that a 1% increase in capital stock reduces marginal cost by approximately 0.109%, indicating the presence of scale economies. Firm fixed effects in (2) suggest that the 1970 merger improved production efficiency. In particular, Nippon Steel – formed by the merger of Yawata and Fuji Steels – exhibits a marginal cost that is $18-21\%$ lower than the average of the two pre-merger firms. Moreover, the merged and merging firms are more productive than their non-merging rivals, with marginal costs lower by more than 4%. For the reminder of the analysis, we use the marginal cost estimates from (2) as our baseline.

Using the estimated marginal costs and the average crude steel prices reported in Table 1, we can compute firms' price-cost markups (P_t^B/mc_{jt}). In the post-merger period, the average markup ratios are approximately 1.65, 1.20, 1.18, 1.18, and 1.09 for NSC, NKK, Sumitomo Steel, Kawasaki Steel, and Kobe Steel, respectively – closely following the ranking of firm size.

Table 4: Estimation results for marginal cost function

	(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
Sumitomo		0.24
Kawasaki		0.25
Kobe		0.25
Year dummies	Yes	Yes
Firm dummies	No	Yes
Adjusted R^2	1.00	1.00

Notes: This table reports the parameter estimates for the marginal cost function specified in Eq.(5). Standard errors are reported in parentheses. *** Significant at the 1 percent level.

We assume Cournot competition among integrated steelmakers when deriving marginal costs in Eq.(7). Alternatively, one could consider a scenario in which steel producers collude in setting their output levels. Under this assumption, the left hand-side of Eq.(7) would equal $1/|\epsilon_t|$. However, when we impose this collusive structure, the estimated marginal costs turn negative for most firms, yielding economically implausible results. This pattern suggests that collusive conduct in output decisions is not supported by the data. Accordingly, we proceed under the Cournot competition assumption, which provides more coherent and empirically credible marginal cost estimates.

5.3 Investment-cost Estimates

Table 5 reports the estimates of the dynamic parameters governing investment costs, as specified in Eq.(9). These estimates are obtained using GMM described in Sections 3.4. Column (1) of the table presents coefficients for both the linear and quadratic terms – θ_k and θ_a, respectively.

The significance of θ_a confirms the presence of convex adjustment costs, which play a nontrivial role in firms' investment decisions. Importantly, the estimates show substantial heterogeneity across firms: smaller firms face markedly higher investment costs. This result may be consistent with the broader macroeconomic literature on investment frictions, as well as the findings of Mermelstein et al. (2020). It highlights the role of adjustment costs in sustaining persistent asymmetries in capital stock and production capacity across firms, a pattern that holds throughout our study period. Note that we also experimented another specification where a term concerning i_{jt}^2, in addition to i_{jt}^2/k_{jt}, is added in the investment cost function, but the coefficient is statistically insignificant, as shown in Column (2) of the table.

In the investment cost estimation, we have assumed oligopolistic competition in investment where each firm chooses its investment levels in order to maximize its own long-run profit. As a robustness check, we alternatively consider a collusion benchmark in which firms coordinate their investment levels to maximize joint profits. In this case, each firm's optimal investment equates the industry-wide marginal gains from investment with its marginal cost. Under such coordination, firms with higher marginal gains to investment are allocated larger investment levels. However, this assumption produces unrealistic implications: smaller firms – such as Kobe Steel – would optimally choose negative investment levels, effectively implying market exit. Because these results are inconsistent with observed behavior, we maintain the assumption of oligopolistic investment competition, which more realistically reflects the strategic environment of the Japanese steel industry.

Table 5: Estimation results for investment cost function

	(1)	(2)
i_{jt}	0.363***	0.294***
	(0.017)	(0.077)
i_{jt}^2/k_{jt}	0.885***	1.161***
	(0.147)	(0.347)
i_{jt}^2		0.296
		(0.308)

Notes: This table reports the parameter estimates for the investment cost function specified in Eq.(9). Standard errors are reported in parentheses. Monetary values are expressed in billion JPY at constant 1960 prices. *** Significant at the 1 percent level.

5.4 Model Prediction

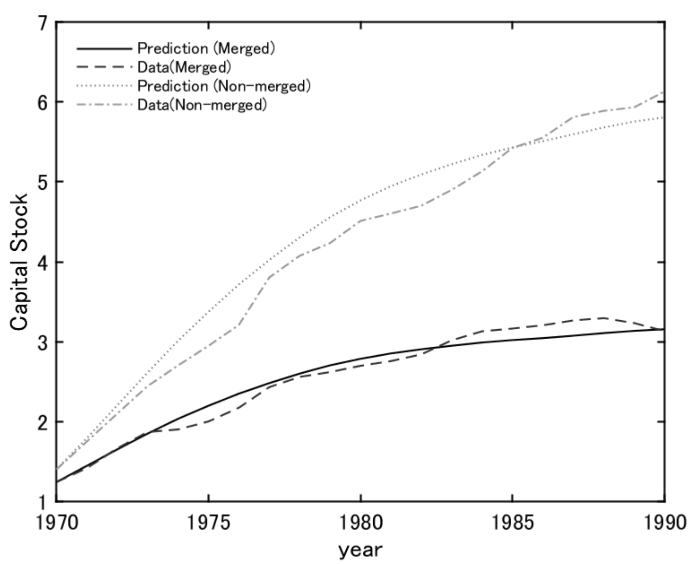
To evaluate the model fit to the data, we compare the actual and predicted market outcomes for four key endogenous variables: industry outputs, prices, market shares, and capital stocks. The first two variables are measured at the industry level, while the latter two are at the firm level. These comparisons are presented in Table 6. In the talbe, we report the results as averages, annualized over the post-merger period from 1970 to 1990, and we classify the market share and capital stock data by merged and non-merged companies. Additionally, Figure 2 illustrates the year-by-year comparison of capital stocks. The both results indicate that our model estimates perform reasonably well in explaining the observed data. All the variables in the table – industry outputs, prices, market shares, and capital stocks – are closely aligned with the actual data, suggesting the model effectively capture the key mechanism of the market.

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: Values are averaged over the study period. Industry output is measured in million tons. Prices are in JPY per thousand tons. Market shares are expressed in percentages. Capital stock is reported in billion JPY, adjusted to 1960 constant prices. Standard errors are shown in parentheses. The merging party refers to Nippon Steel, while the non-merged party includes Nihon Kokan, Sumitomo, Kawasaki, and Kobe.

Figure 2: Predicted capital stocks in the post-merger period



Notes: Capital stock is reported for both the model-predicted and actual data, separately for the merged entity and the aggregate of non-merged firms. All figures are expressed in billion of JPY at constant 1960 prices. The exchange rate in 1960 was 1 USD = 360 JPY.

6 Economic Outcomes of the 1970 Merger

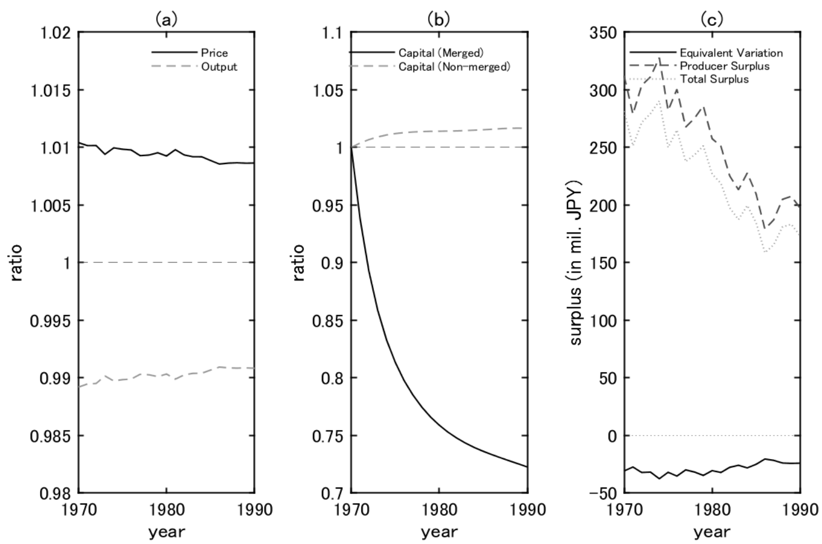
In the three sections that follow, we assess the economic consequences of the 1970 horizontal merger between Yawata and Fuji Steels. In this section, we evaluate the merger's impact on market outcomes by comparing the actual post-merger development with a counterfactual simulation in which the merger did not occur. In Section 7, we examine how the merger determined firms' investment incentives within a dynamic strategic framework. Finally, we assess in Section 8 the effectiveness of the structural remedies imposed at the time of the merger and their long-run implications.

The counterfactual simulation in this section represents what would have occurred in the steel market had Yawata and Fuji continued to operate as independent firms throughout the post-merger period (1970–1990). In this no-merger counterfactual scenario, each of the six firms independently determines its output and investment based on its capital stock and other state variables.[14]

Figure 3 presents the merger's effects on key economic outcomes and is organized into three panels. Panel (a) displays the trend of steel prices and output levels, while Panel (b) shows investment patterns for both merged and non-merged firms during the post-merger period. In these panels the outcomes are reported as ratios relative to the no-merger scenario.

The results indicate that the merger did increase steel prices, but by a small margin of approximately 0.8%. Regarding investment, the merger had a large negative impact on the merged firm's investment: its accumulated capital stock declined by 25% relative to the counterfactual toward the end of the study period. In contrast, non-merged firms increased their investments, with their capital stocks exceeding those in the no-merger scenario by several percentage points in the 1980s.

Figure 3: The merger's effects on economic outcomes from 1970 to 1990
(Comparison with the no-merger scenario)



Notes: The figure shows the merger's effects on economic outcomes in the post-merger period in comparison with the no-merger scenario. Panel (a) displays the trends of steel prices and output levels. Panel(b) shows investment patterns for both merged and non-merged firms. Panel (c) depicts the differences in welfare between the merger and no-merger cases, calculated as the merger outcomes minus the no-merger outcomes.

Panel (c), in contrast, depicts the differences in welfare between the merger and no-merger cases, calculated as the merger outcomes minus the no-merger outcomes. Although the loss in consumer surplus – measured via equivalent variation – was relatively small, consistent with the modest price increase, the substantial rise in producer surplus more than offset this loss. As a result, total surplus increased in both the short and long run, suggesting that the merger was socially desirable from a social-welfare perspective.

[14]As we do not observe any entry or exit in the study period, we assume that the number of firms is fixed under the no-merger scenario.

We also conduct simulations comparing economic outcomes and welfare under two hypothetical scenarios: one with efficiency gains from the merger and one without such gains. In our model, efficiency gains are captured by firm-specific cost parameters, as specified in Eq.(5). When firms j_1 and j_2 merge into a single entity j_0, the realized gains are measured by the cost differences $c_{j_1} - c_{j_0}$ and $c_{j_2} - c_{j_0}$. These terms represent merger-specific cost reductions that are distinct from economies of scale. The latter is modeled separately by the term, $-\gamma ln(k_{jt})$ also in Eq.(5), which describes how marginal costs decline with capital accumulation, consistent with the formulation used in prior research (e.g., Mermelstein et al., 2020). In the counterfactual scenario without efficiency gains, c_{j_0} is unobserved. We therefore construct a synthetic cost parameter, $c_{j_0}^{w/o}$, defined as the weighted average of c_{j_1} and c_{j_2}, where the weights correspond to each firm's capital share prior to the merger.

For both cases – with and without efficiency gains – we consider two investment scenarios that differ in how investment behavior responds to the merger. The "endogenous investment" scenario, previously analyzed in Figure 3, endogenizes firms' investment decisions in response to the merger. In contrast, the "exogenous investment" scenario imposes investment paths that are endogenously determined under the no-merger counterfactual. That is, investment decisions are derived from the model in which the merger never takes place, and then these paths are imposed on the merged industry configuration. For the merged firm, we assume that its capital stock evolves as the sum of the individual merging firms' investment levels taken from the no-merger simulation. Firms then optimize static production decisions, conditional on these capital stocks.

The exogenous investment scenario provides a benchmark for comparison with much of the existing literature that abstracts from dynamic investment responses, and it highlights the importance of fully dynamic models in capturing merger-induced long-run effects on economic outcomes.

Table 7: The average merger effect on market outcomes in the post-merger period

	Firms	With efficiency gain		No efficiency gain	
		(7-1) Endogenous investment	(7-2) Exogenous investment	(7-3) Endogenous investment	(7-4) Exogenous investment
Investment levels	Merged	0.61	1	0.382	1
	Non-merged	1.022	1	1.266	1
Prices		1.009	1.008	1.033	1.039
Outputs	Merged	0.896	0.931	0.597	0.672
	Nonmerged	1.07	1.044	1.28	1.205
Marginal costs	Merged	0.783	0.762	0.966	0.924
	Non-merged	0.997	1	0.983	1
Profits	Merged	2.026	2.011	0.944	0.951
	Non-merged	1.274	1.228	2.102	2.171
Consumer surplus (EV)		-299	-258.2	-1047.3	-1209.9
Producer surplus (excl. investment cost)		2678.8	2593.7	836.2	1137.8
Producer surplus (incl. investment cost)		2343.5	2591.0	1049	1140.5
Total surplus		2379.7	2335.5	1.8	-69.4

Notes: The table reports average effects for the post-merger period on market outcomes and economic welfare. The upper panel reports the ratios of economic variables under the merger relative to the no-merger counterfactual. Output, marginal costs, and investment levels are averaged over the post-merger period, while profits represent the discounted sum for merged and non-merged firms. The lower panel reports the discounted sums of consumer surplus (equivalent variation), producer surplus, and total surplus attributable to the merger. All values are expressed in million JPY (constant 1960 prices).

In total, we consider four counterfactual scenarios as shown in Table 7: (1) Efficiency gains with endogenous investment; (2) Efficiency gains with exogenous investment; (3) No efficiency gains with endogenous investment; and (4) No efficiency gains with exogenous investment.

The table reports average effects for the post-merger period on prices, outputs, marginal costs, investment levels, and firm profits – each expressed as a ratio relative to the no-merger counterfactual – as well as on consumer, producer, and total surplus (measured in millions of 1960 Japanese Yen).

There are two notable observations in Table 7. One is the role of efficiency gains from the merger. Specifically, under the dynamic specification without efficiency gains (3), prices rise significantly (by approximately 3.3%), accompanied by a marked decrease in the merged firm's output (reduction by roughly 59.7%) and a substantial increase in output from non-merging rivals (by approximately 28%), although not sufficiently to offset the merged firm's reduction. The scenarios without efficiency gains (3) and (4) produce unambiguously negative overall welfare outcomes, driven primarily by significant reductions in consumer surplus and aggregate producer surplus losses, resulting from diminished incentives to invest. This is in contrast with the scenarios that incorporate efficiency gains (1) and (2), exhibiting moderate price increases and a net increase in producer surplus, despite a negative impact on consumer surplus. We discuss firms' incentives to invest in the next section.

The second notable observation concerns the endogeneity of firms' investment behavior. Comparing scenarios (3) and (4) reveals that incorporating dynamic investment decisions reduces market prices, thereby improving consumer surplus. At first glance, this outcome may appear counterintuitive, given the reduction in the merged firm's investment. However, this effect arises because non-merging firms substantially increase their investment (by approximately 26.6%), narrowing the disparity in capital stocks across firms and intensifying competition in the market.

Table 8 reports how firms' investment policy functions respond to changes in capital stock in 1970, both actual with the merger (Panel (A)) and counterfactual without it (Panel (B)). Specifically, the table illustrates how a marginal change in one firm's capital influences the investment of its rivals.[15] Across both panels, the estimated policy functions indicate that firms' investment choices are strategic substitutes, irrespective of the merger. For example, under the merger scenario, Panel (A) shows that a 1% increase in the merged firm's (NSC's) capital stock lowers the non-merging firms' investment by approximately 0.03 to 0.1%. This pattern arises because an increase in the merged firm's investment reduces the marginal returns to investment for its competitors – mainly through a decline in their output levels.[16]

The next section explores these strategic investment incentives in details, quantifying how the merger alters the different components of firms' dynamic marginal gains from investment.

Table 8: Elasticity of investment with respect to capital stock (Evaluated in 1970)

Panel (A): Actual with Merger

	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

Panel (B): Counterfactual without Merger

	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

Notes: The table reports the own- and cross-elasticities of firms' current investment with respect to firms' capital stocks, defined as $\frac{\partial \log i_{jt}}{\partial \log k_{mt}}$ for $t = 1970$. Each cell entry (m, j) – where m indexes the row and j the column – represents the percentage change in firm j's investment resulting from a one-percent change in firm m's capital stock. Panel (A) presents the results under the merger, whereas Panel (B) reports the corresponding elasticities in the counterfactual scenario without the merger.

7 Role of Investment

This section examines competitive effects of the horizontal merger on capital investment. Recent theoretical work has begun to explore how mergers influence investment incentives – yet this literature remains static in nature. Motta and Tarantino (2021) analyze the effects of horizontal mergers on cost-reducing investment (process innovation) using a static product-differentiated Bertrand model. They find that, in the absence of spillovers or efficiency gains, mergers reduce the incentives of the merging firms to invest. This is because the marginal return to investment declines post-merger – a result referred to as the margin expansion effect. In other words, the softening of competition reduces the incremental benefit of the merging firm's making investments. Bourreau et al. (2024) extends this analysis to demand-enhancing innovation (product innovation), also within a static setting. Using a general demand framework, they show that the effect of mergers on investment is ambiguous, depending on how the merger shifts incentives to capture demand versus soften competition.

[15]The same qualitative results hold throughout the post-merger period.

[16]Using the notations introduced in the following section, an increase in the merged firms' investment reduces the rivals' $MGI1$ and $MGI2$.

Despite these contributions, the existing literature remains limited to static models and largely abstracts from the dynamic considerations central to investment behavior in capital-intensive industries. This section departs from that framework by developing a dynamic model of horizontal merger in an oligopolistic market with homogeneous goods, where firms make forward-looking investment decisions. By incorporating intertemporal strategic interactions, our approach sheds light on how mergers reshape long-run industry dynamics through their effects on capital accumulation and investment incentives – an aspect overlooked in static analyses.

In Section 7.1, we first analyze the benchmark case, in which we characterize and examine firm's incentive to invest. We identify the underlying sources of the marginal gains from capital investment in our model. We then discuss in Section 7.2 the merger-induced changes in the merged firms' investment incentives. We examines how the marginal gains from investment characterized in Section 7.1 changes before and after the merger for the firms. We identify and quantitatively assess four main effects: margin expansion effect, scale effects in production and capital, and efficiency gains resulting from the merger.

Empirically, we find that static effects dominate, leading the merger to reduce investment incentives for the merging firms. This outcome, however, is contingent on the specific functional forms of the production cost and investment cost functions estimated in our analysis. From a theoretical standpoint, alternative specifications of these functions could yield the opposite result.

7.1 Decomposing Firm's Incentives to Invest

First, we define $\widetilde{V}$ as firm j's present discounted profit as a function of current investments i_t, excluding the current investment cost, as

$$\widetilde{V}_{jt}(i_t; I_{jt+1}, I_{-jt+1}; k_t) \equiv \pi_{jt}(k_t) + \beta\, V_{jt+1}\Big(k_{jt+1}(k_{jt}, i_{jt}),\, k_{-jt+1}(k_{jt}, i_{-jt});\, I_{jt+1}, I_{-jt+1}\Big).$$

The marginal gains from investment for firm j at time t is given by:[17]

$$MGI \equiv \frac{\partial \widetilde{V}_{jt}(i_{jt}, i_{-jt}; I_{jt+1}, I_{-t+1}; k_t)}{\partial i_{jt}} = \sum_{\tau=1}^{\infty} \beta^{\tau} \frac{\partial \pi_{jt+\tau}(k_{t+\tau}) - \phi(k_{jt+\tau}, i_{jt+\tau})}{\partial k_{jt+1}},$$

Here, we define $\widetilde{\pi}_{jt}(q_{jt}, q_{-jt}; mc_{jt})$ as $(P_t(\sum_{j\in\mathcal{J}} q_{jt}) - mc_{jt})q_{jt}$, which is equivalent to $\pi_{jt}(k_{jt}, k_{-jt})$ at the equilibrium. We can decompose MGI into the following four components, $MGI1 \sim MGI4$.[18]

1. **Gains from Production Efficiency:** This term captures how a marginal increase in today's investment reduces future marginal production costs through an increase in capital.

[17] We drop subscripts j and t from MGI and all related notations, including its decomposition terms (e.g., $MGI1$ and difference forms (e.g., ΔMGI) below unless there is confusion.

[18] The following derivative provides $MGI1$ and $MGI2$ by use of the optimal condition:

$$\begin{aligned}
\frac{\partial \pi_{jt+\tau}}{\partial k_{jt+\tau}} &= \frac{\partial mc_{jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial \widetilde{\pi}_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau}; mc_{jt+\tau})}{\partial mc_{jt+\tau}} + \frac{\partial q_{jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial \widetilde{\pi}_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau}; mc_{jt+\tau})}{\partial q_{jt+\tau}} + \\
&\quad \frac{\partial q_{-jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial \widetilde{\pi}_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau}; mc_{jt+\tau})}{\partial q_{-jt+\tau}} \\
&= \frac{\partial mc_{jt+\tau}}{\partial k_{jt+\tau}}(-q_{jt+\tau}) + \frac{\partial q_{jt+\tau}}{\partial k_{jt+\tau}} \cdot 0 + \frac{\partial mc_{jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial q_{-jt+\tau}}{\partial mc_{jt+\tau}} \frac{\partial \widetilde{\pi}_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau}; mc_{jt+\tau})}{\partial q_{-jt+\tau}} \\
&= -\frac{\partial mc_{jt+\tau}}{\partial k_{jt+\tau}} q_{jt+\tau} + \frac{\partial mc_{jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial q_{-jt+\tau}}{\partial mc_{jt+\tau}} \frac{\partial \widetilde{\pi}_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau}; mc_{jt+\tau})}{\partial q_{-jt+\tau}}.
\end{aligned}$$

It is given by

$$MGI1 \equiv -\sum_{\tau=1}^{\infty} \beta^{\tau} \left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{\text{own}} \frac{\partial mc_{jt+\tau}(k_{jt+\tau})}{\partial k_{jt+\tau}} q_{jt+\tau}.$$

By construction, this measure takes a positive value. The term

$$\left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{\text{own}} \equiv \prod_{s=2}^{\tau} \frac{\partial k_{jt+s}}{\partial k_{jt+s-1}},$$

represents the derivative of the firm's capital stock $k_{jt+\tau}$ (for $\tau \geq 1$) with respect to its next-period capital k_{jt+1}, abstracting from any intertemporal strategic interactions with rival firms. Such cross-firm effects are instead captured by MGI_2 and MGI_4.

2. **Gains from Product Market Competition:** A higher capital stock not only reduces a firm's own marginal cost but also affects competitors' outputs. This strategic benefit is expressed as:

$$MGI2 \equiv \sum_{\tau=1}^{\infty} \beta^{\tau} \left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{\text{own}} \frac{\partial mc_{jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial q_{-jt+\tau}}{\partial mc_{jt+\tau}} \frac{\partial \widetilde{\pi}_{jt+\tau}(q_{jt+\tau}, q_{-jt+\tau}; mc_{jt+\tau})}{\partial q_{-jt+\tau}}.$$

3. **Gains from Investment Efficiency:** This term reflects the reduction in future investment costs directly due to an increase in the firm's capital stock and a change in the firm's future investment:

$$MGI3 \equiv -\sum_{\tau=1}^{\infty} \beta^{\tau} \left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{\text{own}} \left[\frac{\partial \phi(k_{jt+\tau}, i_{jt+\tau})}{\partial k_{jt+\tau}} + \frac{\partial i_{jt+\tau}}{\partial k_{jt+\tau}} \frac{\partial \phi(k_{jt+\tau}, i_{jt+\tau})}{\partial i_{jt+\tau}}\right].$$

4. **Gains from Investment Competition:** Finally, a firm's current investment decision may influence its competitors' future investment choices, thereby affecting its own future profits. This effect is captured by

$$MGI4 \equiv \sum_{\tau \geq s \geq 2} \beta^{\tau} \frac{\partial k_{-jt+s}}{\partial k_{jt+1}} \frac{\partial \Big(\pi_{jt+\tau}(k_{t+\tau}) - \phi(k_{jt+\tau}, i_{jt+\tau})\Big)}{\partial k_{-jt+s}}.$$

Decomposing firms' incentives to invest has been a central theme in recent merger analyses, notably in the works of Bourreau et al. (2024) and Motta and Tarantino (2021). Both studies consider sequential frameworks in which firms first choose investment levels and subsequently compete in prices after observing rivals' investment decisions. Within these models, it is possible to identify the counterparts of what we term $MGI1$ and $MGI2$. However, these models do not capture the intertemporal strategic incentives associated with capital investment.

In contrast, our model adopts a dynamic framework in which investment decisions evolve as part of a Markov-perfect equilibrium. This dynamic structure gives rise to two additional forces – $MGI3$ and $MGI4$ – that are absent from the previous literature. $MGI3$ captures the dynamic investment efficiency channel: firms internalize how today's investment reduces future marginal costs, which in turn alters their future value function. $MGI4$ reflects strategic interactions in the investment process itself: firms consider how their own investment alters rivals' future investment incentives, creating a dynamic feedback loop. These forward-looking strategic considerations are central to our analysis and, to our knowledge, have not been explicitly discussed in the previous merger studies. Our decomposition thus provides a more complete picture of the mechanisms through which mergers affect long-run investment incentives.

7.2 How the Merger Changes the merged firms' Incentives to Invest

We turn to consider how the merger affected the merged party's investment incentives. We denote the merging firms as j_1 and j_2 (namely, Yawata and Fuji Steels) and the merged firm as j_0 (NSC). The subscripts M and N refer to the merger and no-merger cases, respectively. In the pre-merger scenario, the overall investment incentive for the merging firms, MGI^N_{merged} is expressed as

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

where $w_j \equiv \frac{k_{j_1 t}}{k_{j_1 t}+k_{j_2 t}}$ denotes the weight attributed to each firm. In contrast, the post-merger investment incentive for the merged firm, MGI^M_{merged} is captured by

$$MGI^M_{merged} \equiv \frac{\partial \widetilde{V}^M_{j_0 t}}{\partial i_{j_0 t}} \Big(i^{*N}_{j_1 t} + i^{*N}_{j_2 t},\, i^{*M}_{-j_0 t};\, I^M_{j_0 t+1},\, I^M_{-j_0 t+1};\, k^M_t \Big).$$

Based on these definitions, the following proposition characterizes how the merger affects investment decisions. Specifically, the direction of the effect depends on the relative magnitude of the marginal gains from investment for the merged entity, evaluated at the investment levels that would have been chosen had the firms remained independent (i.e., absent the merger).

Proposition 1. *The impact of the merger on investment for the merged entity is given by $i^{*M}_{j_0 t} - (i^{*N}_{j_1 t} + i^{*N}_{j_2})$, and has the same sign as $MGI^M_{merged} - MGI^N_{merged}$.*

Proof. See Appendix. □

As discussed in Section 7.1, the marginal gain from investment (MGI) can be decomposed into four components: $MGI1$, $MGI2$, $MGI3$, and $MGI4$. For each of these components, we calculate the difference $\Delta MGIs \equiv MGIs^M_{merged} - MGIs^N_{merged}$, where $s = 1, 2, \cdots, 4$. This difference indicates the merging firm's incentive to invest for each component s. Concerning $MGI1$ and $MGI2$, for each s, we can further decompose four distinct effects: (1) Scale Effect in Production, (2) Margin Expansion Effect, (3) Scale Effect in Capital, and (4) Efficiency Gains Effect.

Let us take $MGI1$. Let $k^N_{jt+\tau}$ be the capital stock of firm j conditional on the investment strategies i^{*N}_t and initial capital stocks k^N_t. The merger-induced incentive to invest for the merging firms is written as:

$$\Delta MGI1 \approx \sum_{\tau=1}^{\infty} \beta^\tau \sum_{j\in\{j_1,j_2\}} w_j \left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{\text{own}} \left(\sum_{n=1}^{4} Z^{(n)}_{jt+\tau} \right), \tag{19}$$

where $Z^{(n)}_{jt+\tau}$ $(n = 1, 2, 3, 4)$ correspond to the following four effects:[19]

1. **Scale Effect in Production** $Z^{(1)}_{jt+\tau} = \left(-\frac{\partial mc_{jt+\tau}(k^N_{jt+\tau})}{\partial k^N_{jt+\tau}}\right) \left[\left(q^N_{j_1 t+\tau} + q^N_{j_2 t+\tau}\right) - q^N_{jt+\tau} \right]$: This effect captures the change in profit margins induced by the merger. It arises from the difference between a merging firm's output, either $q_{j_1,t+\tau}$ or $q_{j_2,t+\tau}$, and the combined outputs of the merging firms, $q_{j_1,t+\tau} + q_{j_2,t+\tau}$. This effect is observed only in a homogeneous-good model, as in our application.

[19]See the unpublished appendix for the details of the derivations.

2. **Margin Expansion Effect** $Z^{(2)}_{jt+\tau} = \left(-\frac{\partial mc_{jt+\tau}(k^N_{jt+\tau})}{\partial k^N_{jt+\tau}}\right)\left[q^M_{j_0t+\tau} - \left(q^N_{j_1t+\tau} + q^N_{j_2t+\tau}\right)\right]$: This effect captures the change in profit margins induced by the merger. It is manifested in the difference between the merged firm's output, $q_{j_0,t}$, and the combined outputs of the merging firms, $q_{j_1,t+\tau} + q_{j_2,t+\tau}$. By internalizing competitive interactions in the product market, the merger alters the marginal gains from investment through both production cost channels (denoted by $MGI1$) and the strategic effects on competitors' outputs (denoted by $MGI4$).

3. **Scale Effect in Capital** $Z^{(3)}_{jt+\tau} = (-1)\left[\frac{\partial mc^{w/o}_{j_0t+\tau}\left(k^N_{j_1t+\tau}+k^N_{j_2t+\tau}\right)}{\partial k_{j_0t+\tau}} - \frac{\partial mc_{jt+\tau}(k^N_{jt+\tau})}{\partial k^N_{jt+\tau}}\right] q^M_{j_0t+\tau}$: In the equation above, $mc^{w/o}_{j_0t+\tau}(\cdot)$ denotes the marginal cost function absent the efficiency gains. This effect reflects how the consolidation of capital, $k_{j_1t} + k_{j_2t}$, under the merged entity reorganizes the investment incentives via economies of scale. The resulting capital expansion improves production and investment efficiencies -- but lesser extent as the benefit of scale has decreasing returns. The corresponding effect was also discussed in Mermelstein et al. (2020).

4. **Efficiency Gains Effect** $Z^{(4)}_{jt+\tau} = -\left[\frac{\partial mc_{j_0t+\tau}\left(k^N_{j_1t+\tau}+k^N_{j_2t+\tau}\right)}{\partial k_{j_0t+\tau}} - \frac{\partial mc^{w/o}_{j_0t+\tau}\left(k^N_{j_1t+\tau}+k^N_{j_2t+\tau}\right)}{\partial k_{j_0t+\tau}}\right] q^M_{j_0t+\tau}$: This effect is quantified by the difference of the firm-specific fixed effects of the marginal cost function $(c_{j_1t+\tau}+c_{j_2t+\tau})-c_{j_0t+\tau}$ and captures improvements in operational efficiency resulting from the merger. The efficiency gains enhance the profitability of further investments by lowering the overall marginal costs.

In the decomposition above, $q^N_{jt+\tau}$ denotes the equilibrium output of firm j $(=j_0$ or $j_1)$ at time $t+\tau$, and $q^M_{j_0t+\tau}$ the output of the merged firm j_0. Both are defined conditional on the investment strategies i^{*N} and initial capital stocks k^N_t. $\Delta MGIs$ for $s \neq 1$, are described in the unpublished appendix. In the next subsection, we quantitatively evaluate how each component contributes to the merger-induced incentives to invest.

7.3 Quantitative Assessment of Investment Incentives under Merger

In this subsection, we estimate the firms' incentives to invest induced by the merger. Unlike the static theoretical models of Motta and Tarantino (2021) and Bourreau et al. (2024), our analysis evaluates the investment incentives of the merged firm in a dynamic framework. Specifically, we calibrate the model by endogenizing the derivative:

$$\left.\frac{\partial k_{jt+\tau}}{\partial k_{jt+1}}\right|_{\text{own}} \tag{20}$$

which captures the propagation of current investment shocks into future capital accumulation. This allows us to assess the merger's long-run impact on investment incentives to dynamic strategic interactions. For comparison, we also report results under a static assumption, setting the term in Eq.(20) to be equal to 1. This assumption effectively removes intertemporal linkages by assuming that future capital accumulation is invariant to current investment, and facilitates direct comparison with static merger models.

Table 9 presents the decomposition of merger-induced investment incentives for the merged firm, NSC, under both dynamic and static calibrations evaluated in 1970. Column (1) shows NSC's incentives to invest under the merger, while column (2) presents those in the counterfactual no-merger scenario. The difference, reported in column (3), captures the merger-induced change in investment incentives. This change is further decomposed into individual components, denoted by $\Delta MGIs$, which isolate distinct channels through which the merger influences incentives. For example, scale effect in production, captured by $Z^{(1)}_{jt+\tau}$, appear in both $\Delta MGI1$ and $\Delta MGI2$; thus, their contributions to the total in (4-1) equals the sum of those two terms. Analogous decompositions apply to margin expansion (4-2), scale effects in capital (4-3), efficiency gains (4-4), and other residual components (4-5).

Across both dynamic and static calibrations, the results demonstrate that the merger reduces the merged firm's incentives to invest. This decline is driven primarily by two forces: the margin-expansion effect (4-2), which softens competition, and scale effects in capital (4-3), which lower the marginal benefit of additional investment. Together, these forces outweigh positive mechanisms and lead to an overall reduction in investment incentives. This finding aligns with predictions in the theoretical literature on static models (e.g., Bourreau et al., 2024) that emphasize the adverse impact of horizontal mergers on investment when efficiency gains are absent.

At the same time, our analysis departs from much of the existing literature. Applied to the homogeneous product market of crude steel, the model highlights a positive production scale effect: consolidation enhances productive efficiency and profitability, thereby strengthening investment incentives. Such production effects are largely absent in merger models based on product differentiation. Yet, even with this positive force, the overall effect of the merger remains negative, a result that is robust to the presence or absence of efficiency gains.

The disaggregated results offer further insights. Both $MGI1$ and $MGI2$ are positive in the no-merger baseline, while the total change induced by the merger is negative. This differs from standard price-competition models, in which the counterpart of $MGI2$ (Gains from Product Market Competition) is typically negative because lower marginal costs reduce prices and intensify competition. By contrast, in our Cournot quantity-setting, greater investment by one firm induces rivals to scale back output, which increases the investing firm's profits. Hence, the contribution of $MGI2$ is positive in our model.

Other component, such as $MGI4$ (Gains from Investment Competition), contributes little, indicating that preemptive motives for investment play a limited role in our context. Overall, the dominant forces governing the merged firm's behavior are captured by $MGI1$ and $MGI2$, which are also central in static models.

Finally, we extend the decomposition to non-merging firms in the unpublished appendix. For these firms, merger effects operate mainly through changes in output levels rather than investment responses. This suggests that static mechanisms dominate in driving their behavior, while dynamic incentives are less important. Consistent with Bourreau et al. (2024), these limited responses are insufficient to counteract the decline in the merged firm's incentives. As a result, aggregate industry investment falls, with adverse implications for consumer surplus.

Table 9: Decomposition of merged firm's MGI (Evaluated in 1970)

	With merger	No merger	Incentives to Invest (Total)	Changes in investment incentives by Merger (4)				
				Scale Effect in Production	Margin Expansion	Scale Effect in Capital	Efficiency Gains	Others
	(1)	(2)	(3) (=(1)-(2))	(4-1)	(4-2)	(4-3)	(4-4)	(4-5)
Total								
Dynamic	625.4	787.3	-161.9	559.8	-371.3	-410.1	65.9	-6.2
Static	603.9	758	-154.1	671.7	-444.9	-490.6	72.5	37.2
MGI1			**Δ MGI1**					
Dynamic	233.1	333.3	-100.2	329.7	-218.9	-237.6	28.4	-1.8
Static	332.3	400.3	-68	396.9	-262.7	-285	34.7	48.1
MGI2			**Δ MGI2**					
Dynamic	165.5	231.5	-66	230.1	-152.4	-172.5	31.1	-2.3
Static	235.4	278.1	-42.7	274.8	-182.2	-205.6	37.8	32.5
MGI3			**Δ MGI3**					
Dynamic	210.5	196.8	13.7	-	-	-	-	13.7
Static	36.2	79.6	-43.4	-	-	-	-	-43.4
MGI4			**Δ MGI4**					
Dynamic	16.3	25.7	-9.4	-	-	-	-	-9.4

Notes: This table reports the values of $MGIs$ and $\Delta MGIs$ (where $s = 1, \cdots, 4$), evaluated in 1970. The columns "With merger" and "No merger" correspond to MGI^M_{merged} and MGI^N_{merged}, respectively. The change in investment incentives induced by the merger is defined as $\Delta MGIs \equiv MGIs^M_{merged} - MGIs^N_{merged}$ for $s = 1, \cdots, 4$. "Static" values are computed under the assumption that $\left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{own} = 1$, while "Dynamic" values account for endogenous capital accumulation through $\left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{own}$. All $MGIs$ and $\Delta MGIs$ values are expressed using investment and capital stocks measured in billion of JPY and profits measured in million of JPY (constant 1960 prices).

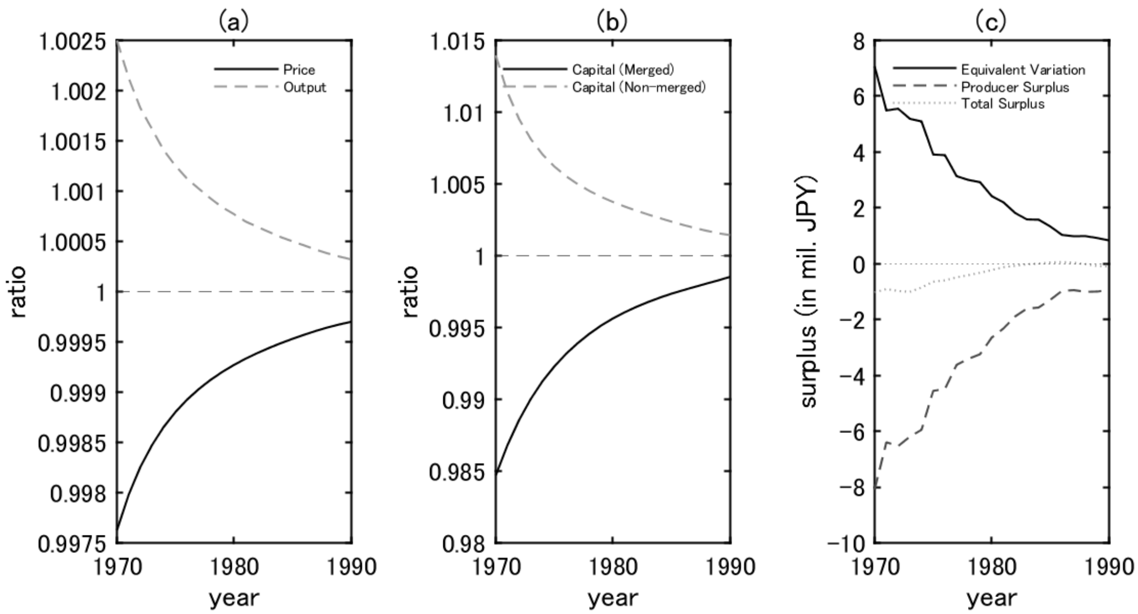
8 Merger Remedies

This section evaluates the merger remedies implemented in the context of the 1970 steel merger. The merger was approved on the condition that Yawata and Fuji Steels divest 1.5% of their production facilities to Kobe Steel and 0.3% to Nihon Kokan. Although the divested assets represented only a small fraction of the merging firms' total capital, there is limited empirical research quantifying the effects of such remedies, and even fewer studies that assess optimal remedy design. This paper contributes to filling that gap. Section 8.1 examines the economic effects of these remedies, while Section 8.2 presents simulations of optimal remedy policies, evaluated under both consumer-welfare and social-welfare criteria.

8.1 Economic Outcomes of the Merger Remedies

Figures 4 shows the impacts of 1970 merger remedies on economic variables from 1970 to 1990, in comparison with the absence of the remedies. Panel (a) of Figure 4 presents prices and outputs. The merger remedies lowered steel price immediately after the 1970 remedies by roughly 0.25 %, but the effect gradually decreased over time. However, half of the effect still persisted 10 years after the merger remedies, and approximately 20 % of the effect remained 20 years later.

Figure 4: Effect of 1970 merger remedies on economic outcomes



Notes: The figure presents three panels illustrating the effects of the merger remedies on economic outcomes in the post-merger period, compared with the counterfactual scenario without remedies. Panel (a) shows the trajectories of steel prices and output levels; Panel(b) shows the evolution of capital stocks for both merged and non-merged firms; and Panel (c) depicts welfare differences between the remedies and no-remedies cases, calculated as the outcomes under the remedies minus those under the no-remedies scenario.

Panel (b) reports the evolution of capital stocks for the merged and non-merged firms. Following the remedies, the merged firm reduced its investment in proportion to the reduction in its capital stock, whereas the non-merged firms increased their investment. As a result, the capital stock levels of the two groups gradually converged. Within ten years, the impact of the remedies had fallen to roughly one-third of its initial magnitude. Still, about one-tenth of the initial effect remained even after the twenty years.

Panel (c) is on the surplus. It shows that the merger remedies increased equivalent variation but decreased producer surplus. In total, the merger remedies decreased social welfare. The effects decreased over time: half of the increase in the equivalent variation disappeared ten years later.

Overall, the analysis reveals that the remedies implemented in 1970 continued to have the measurable effects even two decades after the merger, primarily through dynamic investment mechanisms. Although the immediate impact of divestiture diminished over time, the reallocation of production assets altered firms' long-run investment trajectories and capacity accumulation paths, thereby sustaining competitive pressure well into the late 1980s.

8.2 Optimal Remedy Policies

This section simulates optimal merger remedies under two evaluation standards in antitrust analysis: the consumer welfare standard and the social welfare standard (see Whinston, 2008). The analysis seeks to identify both the optimal magnitude of divestiture and the optimal allocation of divested assets across rival firms. In what follows, we first focus on the consumer-welfare criterion and turn to social-welfare.

8.2.1 Consumer Welfare Standard

To design optimal merger remedies under the consumer welfare standard, we solve the following two-step optimization problem:

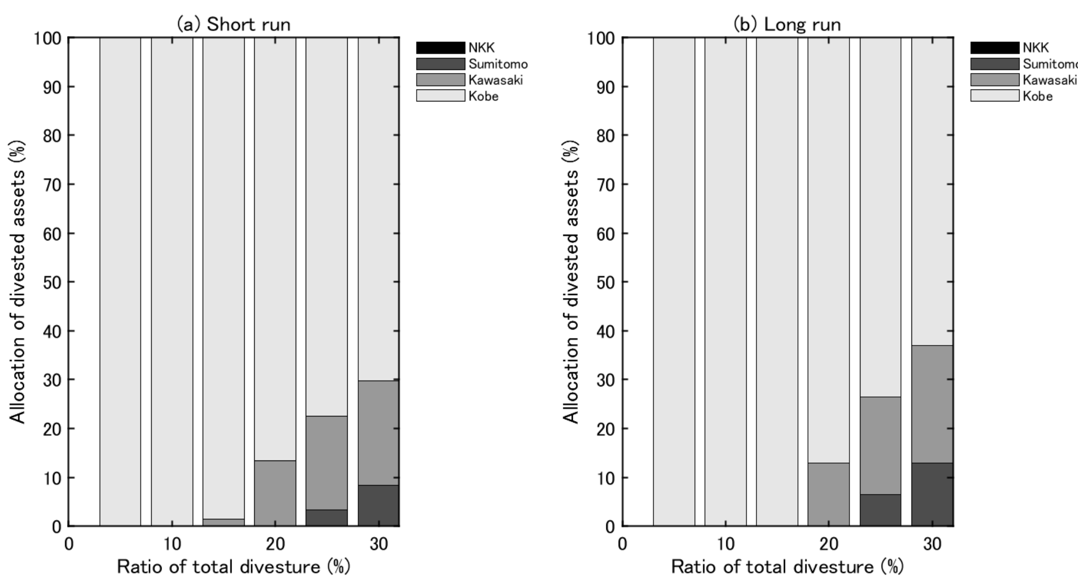
$$\max_D \max_{\Delta(D)} \quad CW(\Delta(D), D) \tag{21}$$

$$s.t. \quad \sum_{j \in \mathcal{R}} \Delta_j = 1,$$

where $\mathcal{R}$ denotes the set of nonmerged firms. The optimization proceeds in following two steps. In the first step, given the total amount of assets to be divested by the merged firm, denoted by D, we determine the allocation of these divested assets across non-merging firms. Let Δ represent the vector of allocation shares, where j-th element, Δ_j, denotes the share of Δ received by firm $j \in \mathcal{R}$. Accordingly, the amount of divested assets allocated to firm j is $D \cdot \Delta_j$. The optimal allocation vector Δ generally varies with the divestiture size D. In the second stage, we identify the divestiture size D^{CW} that maximizes consumer welfare, given its corresponding optimal allocation.

Figure 5 shows how the optimal allocation of divested assets across non-merging firms, $\Delta(D^{CW})$ evolves as the total divestiture size – measured as the ratio of the merged firm's capital, D^{CW}/k^M_{j0} – increases. Panel (a) presents the allocation that maximizes consumer welfare in the merger year (short-run analysis), while panel (b) shows the allocation that maximizes the discounted sum of equivalent variation (long-run analysis).

Figure 5: Optimal allocation of divested assets (Consumer welfare standard)



Notes: The figure illustrates the optimal allocation of divested assets across non-merging firms as the total divestiture size, expressed as a ratio of the merged firm's capital stock increases. All values are evaluated for the year 1970. Panel (a) presents the short-run case – showing that the allocation that maximizes consumer welfare in the merger year – while Panel (b) depicts the long-run case, in which the allocation that maximizes the discounted sum of equivalent variations over the post-merger period.

In both short- and long-run cases, the optimal allocation of the divested assets directs the largest share to Kobe Steel, the smallest rival firm. This result aligns with the argument in Vergé

(2010), which highlights the welfare benefits of restoring greater symmetry in productive capacity size among competitors to mitigate post-merger distortions. In our dynamic setting, allocating capacity to the smallest firm enhances competitive pressure over time by enabling it to expand output and invest, thereby generating the largest long-run gains in consumer welfare.

To better understand the underlying mechanism, Table 10 reports how economic outcomes respond when each firm's capital stock is exogenously increased by one million JPY. The merged entity (NSC) and the four non-merging rivals are analyzed separately. The upper portion of the table summarizes the effects on marginal costs and output prices, while the lower portion presents the corresponding impacts on welfare components – consumer welfare (EV), producer surplus (PS), and total (social) welfare (TS).

The results reveal that a marginal increase in capital for Kobe Steel produces the largest reduction in marginal cost and, consequently, the greatest improvement in consumer welfare. This outcome reflects the firm's smaller initial scale and higher marginal productivity of capital: expanding its capacity most effectively strengthens competition and lowers market prices.

Table 10: Effects of a marginal increase in firm-level capital on the post-merger economic outcomes

	NSC	NKK	Sumitomo	Kawasaki	Kobe
Marginal costs	-2.1	-7.9	-9.3	-10.8	-26.9
Prices	-0.5	-1.9	-2.3	-2.7	-6.6
EV	239.3	759.5	845.4	912.2	1918.5
PS	527.3	62.7	23.1	25	-1266.3
TS	766.6	822.1	868.4	937.1	652.2

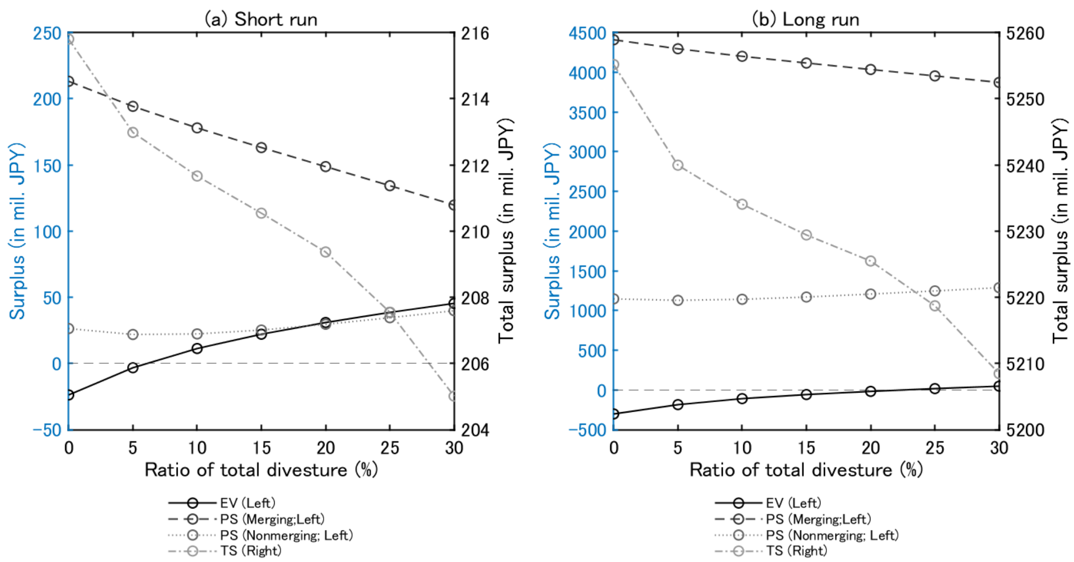
Notes: This table reports the simulated effects of a one-million-JPY exogenous increase in each firm's capital stock in 1970 on key economic outcomes, averaged over the 1970-1990. Firms are in descending order of size. "EV," "PS",and "TS" denote equivalent variation (consumer welfare), producer surplus, and total (social) welfare, respectively. The upper panel presents changes in marginal costs and output prices (measured in JPY per 1,000 ton of crude steel), while the lower panel summarizes the associated welfare effects (in million JPY).

Figure 6 illustrates how different welfare measures respond to variations in the divestiture ratio. Panel (a) presents short-run consumer welfare, while Panel (b) focuses on long-run outcomes. In the short-run, welfare losses are largely mitigated with a relatively small divestiture of around 5%. In contrast, in the long-run, consumer welfare improves gradually as the divestiture expands, with the merger-induced losses fully offset when the divestiture reaches approximately 20%,

This magnitude of the short-run divestiture ratio is consistent with the analytical benchmark proposed by Nocke and Rhodes (2025), which derives the condition under which a merger is deemed acceptable to the consumer-maximizing authority – namely, when efficiency gains outweigh the merger's market-power effects. Their analysis correspond to our short-run analysis here. Although our demand specification slightly departs from their Cournot framework (since the integrated steel producers in our model face import competition through a CES-type demand system), the underlying logic is similar: the efficiency gains from the merger exceed the pre-merger market-power effect (approximately 8000 JPY per ton for the merged firms' price-cost markups, $P_t^B - \sum_{j \in \mathcal{M}} mc_{jt}$) once the divested assets amount to about five percent of the merged firm's capital stock.

The comparison between the short- and long-run cases underscores the time-dependent nature of merger remedies. Immediately after the merger, divestitures restore competition mainly through a one-off reallocation of productive capacity. However, as time passes, firms adjust their investment and production decisions endogenously in response to the new market structure. Without sustained or adaptive interventions, the initial pro-competitive effects of divestiture gradually erode as larger firms re-accumulate capacity and smaller firms' relative advantages fade. Consequently, maintaining long-run consumer welfare requires a larger divestiture to offset these dynamic adjustments in investment and capacity accumulation.[20]

Figure 6: Short- and long-run welfare effects of varying divestiture ratios (Consumer welfare standard)



Notes: The figure depicts how welfare components respond to changes in the total divestiture ratio following the 1970 merger. Panel (a) presents the short-run effects, while Panel (b) reports the long-run effects. The curves show equivalent variation (EV, consumer welfare), producer surplus (PS) for merging and non-merging firms, and total surplus (TS), each measured in million JPY at constant 1960 prices.

8.2.2 Social Welfare Standard

We now evaluate the optimal merger remedies under the social welfare standard, which accounts for the joint surplus of consumers and producers. The optimization procedure follows the same two-step approach applied under the consumer welfare standard, as defined in Eq. (21).

Figure 7 presents the optimal allocation of divested assets across rival firms under the social welfare criterion, corresponding to Figure 5 in consumer welfare case. Unlike the consumer welfare standard – where directing most divested assets to the smallest rival, Kobe Steel, maximizes welfare – the optimal allocation under the social welfare criterion shifts toward Kawasaki Steel, the second-smallest firm in the market.

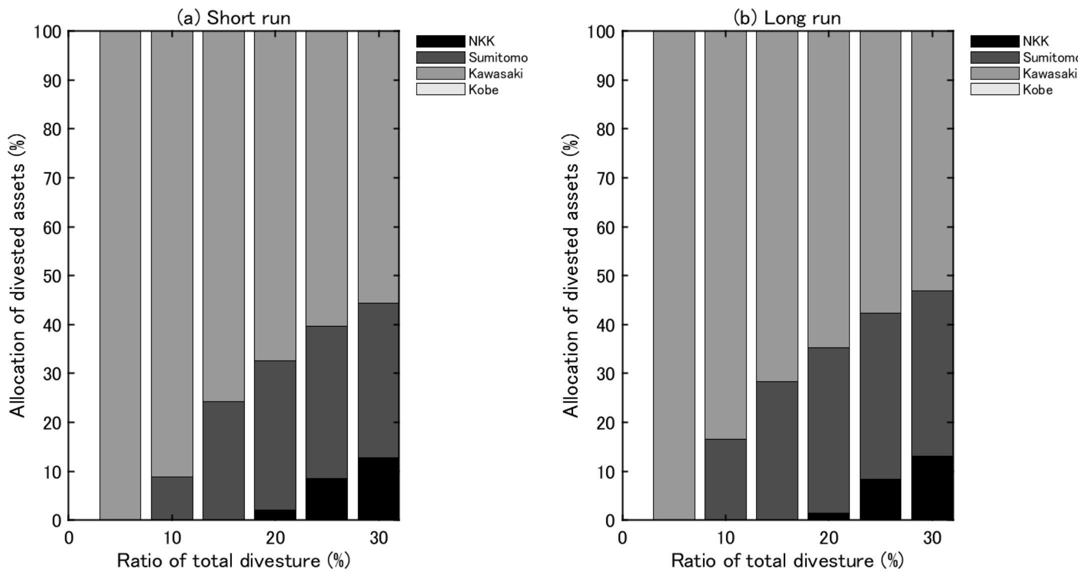
[20]Nocke and Whinston (2010) develop a theoretical framework for understanding merger review from a long-run perspective, where merger proposals arrive endogenously over time and the antitrust authority must evaluate each case in anticipation of future merger activity. In their setting, an authority that seeks to maximize consumer surplus should block any merger that reduces consumer surplus on a case-by-case basis, regardless of potential dynamic interactions among successive mergers.

The notion of "long run" in our analysis differs from theirs. Whereas Nocke and Whinston (2010) focus on a dynamic sequence of merger proposals, our long-run perspective concerns how firms behave after a given merger has occurred – in particular, how their investment incentives and capacity trajectories evolve over time. Thus, although both analyses incorporate dynamic considerations, the conceptual focus of the long run is fundamentally different.

As shown in Table 10, expanding the capacity of Kobe Steel exerts the strongest downward pressure on prices, intensifying competition and raising consumer surplus. However, this gain comes at a substantial cost: producer margins are compressed, and overall industry profitability declines.

By contrast, the social-welfare-maximizing allocation assigns the majority of divested assets to Kawasaki Steel. This reallocation achieves a more balanced outcome – moderating consumer surplus relative to the consumer-welfare-maximizing case while substantially increasing producer surplus. This result is a higher overall level of social welfare, as summarized in Table 10.

Figure 7: Optimal allocation of divested assets under the social welfare standard



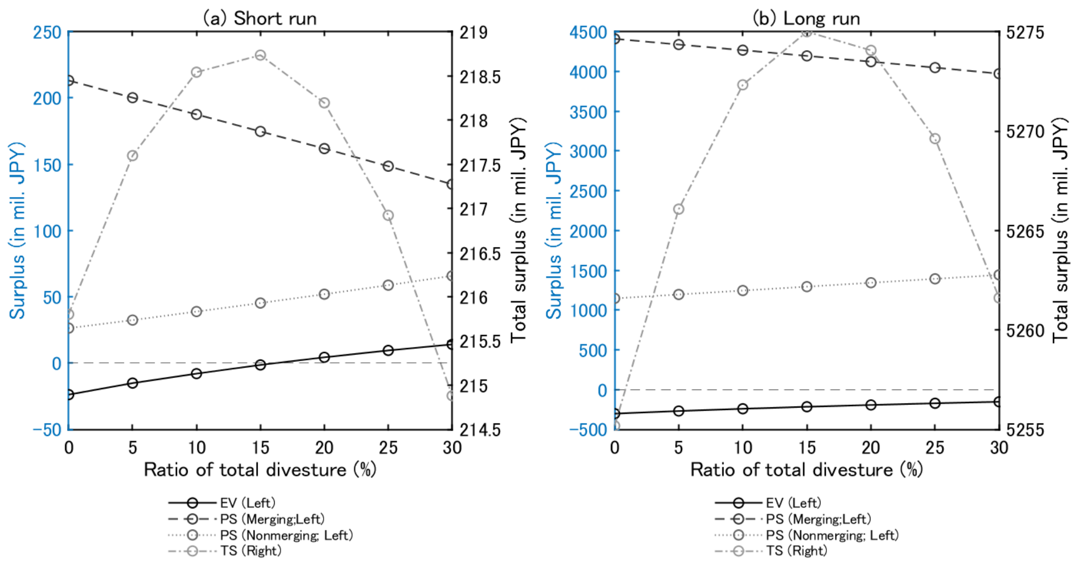
Notes: The figure displays the optimal allocation of divested assets among non-merging firms under varying total divestiture ratios. Panel (a) presents the short-run case, while Panel (b) reports the long-run equilibrium allocation. The optimization maximizes total (social) welfare, defined as the sum of consumer and producer surplus. In contrast to the consumer welfare benchmark, the optimal policy here directs a larger share of divested capacity to Kawasaki Steel, reflecting the trade-off between promoting competition and preserving industry profitability.

Figure 8 illustrates how welfare components evolve as the scale of divestiture increases under the social welfare standard. Panel (a) presents the short-run outcomes, while Panel (b) shows the long-run results. In both cases, total surplus exhibits an inverted-U shape: it initially rises with the divestiture ratio, reaches its maximum around 15%, and then gradually declines beyond that point.

This pattern indicates that a moderate level of divestiture balances competitive gains with efficiency losses. In the short run, a divestiture ratio of roughly 15% not only maximizes total surplus but also restores consumer welfare to a positive level – though this ratio is larger than the 5% found optimal under the consumer welfare standard. The discrepancy reflects that, under the social welfare objective, consumer benefits from lower prices must be weighed against reductions in producer rents.

In the long run, however, the gains to consumer welfare taper off as firms re-optimize investment and capacity accumulation. Even with a divestiture ratio as high as 30%, consumer surplus does not turn positive. Notably, while the divestiture ratio that maximizes social surplus in the long run is 15%, which is similar in magnitude to the long-run divestiture ratio that maximizes consumer surplus in Figure 6, the substantial difference in the welfare impact suggests that the allocation of divested assets across non-merging firms is critical.

Figure 8: Short- and long-run welfare effects of varying divestiture ratios (Social welfare standard)



Notes: The figure shows how different welfare components respond to changes in the total divestiture ratio under the social welfare standard. Panel (a) presents short-run results, while Panel (a) reports long-run effects. EV denotes consumer welfare (equivalent variation), PS refers to producer surplus (for merging and non-merging firms), and TS represents total (social) surplus. All measures are expressed in million JPY at constant 1960 prices.

9 Conclusion

This paper has developed and estimated a dynamic oligopoly model to evaluate firms' investment and long-run welfare in a capital intensive industry. Using detailed historical data on the Japanese steel industry, we showed that mergers can generate efficiency gains and long-run welfare improvement even when short-run prices rise. These stem primarily from cost reductions achieved by the merged firm, complemented by the endogenous investment responses of its rivals.

Our analysis of merger remedies revealed that their effectiveness critically depends on the dynamic evolution of firms' investment and capacity accumulation. Remedies that restore competition at the time of implementation may gradually lose their potency as firms re-optimize in subsequent periods. The 1970 structural remedies retained roughly half of their impact after a decade, with gradual dissipation driven by adjustment costs and endogenous reinvestment. These findings underscore that one-off interventions cannot fully neutralize long-run consequences of mergers; sustaining competitive pressure requires accounting for how firms' strategic investments evolve over time.

The comparison between consumer welfare- and social welfare-based standards further illustrates the tension between static and dynamic considerations. Under the consumer welfare criterion, optimal remedies favor reallocating assets to the smallest rival to restore capacity symmetry and enhance competition. In contrast, the social welfare standard places greater weight on preserving producer rents, favoring more moderate reallocation to mid-sized rivals.

The results on remedies connect to Farrell (2003)'s seminal insights on negotiated remedies. The paper cautions that merger control often suffers from two inefficiencies – over-fixing, where authorities impose remedies exceeding the actual competitive harm, and under-alignment, where the recipients of divested assets lack incentives to sustain competition. Our dynamic framework addressed the former and showed that these problems can be intertemporal: remedies that appear excessive ex ante may prove insufficient ex post as firms endogenously re-accumulate capacity. The analysis thus calls for remedy design guided by intertemporal welfare rather than static benchmarks – emphasizing not only *how much* to divest, but *to whom*, and with what dynamic capacity to compete.

A Data

This study combines multiple data sources to construct a firm-level panel of the Japanese steel industry covering the period 1960 - 1990.

Annual data on industry- and firm-level crude steel output were obtained from publications of the Japan Iron and Steel Federation. Firm-level annual prices for domestically produced crude steel from blast furnace producers were compiled directly from their semiannual financial reports over the same period. Price variation across firms was minimal; accordingly, we use Yawata Steel's crude steel prices in the 1960s and Nippon Steel's prices in the 1960s as representative benchmark for estimation.

Data for electric furnace producers were not publicly available. To address this, we used the wholesale price index (WPI) for steel, published by the Bank of Japan (BOJ), which aggregates blast furnace and electric furnace products using sales-based weights. From this index, we constructed a proxy for the average price of steel produced by electric furnace firms. All price series were deflated using the WPI for manufactured goods to express values in constant 1960 Japanese yen.

The analysis also incorporates two input prices, namely those of iron ore and heavy oil, both obtained from the BOJ. In addition, we collected data on average seaborne shipping distances for iron ore from the Japan Iron and Steel Federation; these data are used as an instrumental variable in the demand estimation to capture exogenous variation in input costs.

Firm-level capital stock is measured as the book value of physical production assets, as reported in companies' semiannual financial statements from 1960 to 1990. Annual firm-level investment was constructed from changes in these capital stock series. To convert book values into market-based measures, we followed the procedure proposed by Ogawa and Kitasaka (1998) , using information from the 1960 National Wealth Survey to impute the average asset age for steel producers. Both capital stock and investment were deflated using the WPI for manufactured goods to ensure comparability over time.

B Spectral Algorithm and Smolyak Method

This section describes the methods used in the main algorithm to solve for the equilibrium of the model discussed in Section 4. Two methods are discussed: Spectral algorithm and Smolyak method.

Spectral Algorithm

Solving for the equilibrium in dynamic oligopoly models often entails substantial computational burden, as convergence can require a large number of iterations. To accelerate convergence, we incorporate a spectral algorithm, originally developed by Barzilai and Borwein (1988) and improved in La Cruz et al. (2006).

As discussed by Judd (1998, p. 166), convergence of fixed-point iterations can be accelerated through extrapolation. Suppose a variable x satisfies a fixed-point condition $x = f(x)$. A standard iterative update takes the form

$$x^{(n+1)} = \rho f(x^{(n)}) + (1-\rho)x^{(n)},$$

where ρ controls the degree of extrapolation. When $\rho = 1$, convergence may be slow; setting $\rho > 1$ can accelerate convergence, whereas $\rho < 1$ may help stabilize an otherwise divergent iteration. A major limitation of this approach, however, lies in selecting an appropriate value of ρ, which often requires costly manual tuning.

The spectral algorithm overcomes this difficulty by adaptively updating ρ based on the local convergence path. As discussed in Varadhan and Roland (2008), the spectral step size mimics the update rule of Newton's method – achieving near-Newton convergence speed without requiring analytical derivatives of the fixed-point mapping. Hence, the method inherits Newton's fast convergence properties while remaining computationally simple.

In our implementation (see Section 4, the equilibrium updates for firms' investment and value functions, $i_{jt}^{(n+1)}$ and $V_{jt}^{(n+1)}$, could in principle be computed directly as

$$i_{jt}^{(n+1)} = \widehat{i}_{jt}^{(n+1)}, \quad V_{jt}^{(n+1)} = \widehat{V}_{jt}^{(n+1)},$$

which corresponds to setting $\rho_I^{(n)} = \rho_V^{(n)} = 1$. However, this baseline iteration converges slowly in practice. The spectral algorithm accelerates this process by adaptively adjusting the tuning parameters ρ_I and ρ_V based on th recent history of updates.

More specifically, let $\widehat{i}_{jt}^{(n+1)}(s_t) \equiv \psi_i(i_{jt}^{(n)}, V_{jt}^{(n)})$ and $\widehat{V}_{jt}^{(n+1)}(s_t) \equiv \psi_V(i_{jt}^{(n)}, V_{jt}^{(n)})$. Define the deviation from the fixed point as

$$\phi_i(i_{jt}^{(n)}, V_{jt}^{(n)}) \equiv i_{jt}^{(n)} - \psi_i(i_{jt}^{(n)}, V_{jt}^{(n)}), \quad \phi_V(i_{jt}^{(n)}, V_{jt}^{(n)}) \equiv V_{jt}^{(n)} - \psi_V(i_{jt}^{(n)}, V_{jt}^{(n)}).$$

The updates are then

$$i_{jt}^{(n+1)} = i_{jt}^{(n)} - \rho_i^{(n)}\phi_i(i_{jt}^{(n)}, V_{jt}^{(n)}), \quad V_{jt}^{(n+1)} = V_{jt}^{(n)} - \rho_V^{(n)}\phi_V(i_{jt}^{(n)}, V_{jt}^{(n)}),$$

where

$$\rho_i^{(n)} = \kappa\frac{\left\|\Delta i^{(n)}\right\|}{\left\|\Delta\phi_i^{(n)}\right\|}, \quad \rho_V^{(n)} = \kappa\frac{\left\|\Delta V^{(n)}\right\|}{\left\|\Delta\phi_V^{(n)}\right\|}.$$

The parameter $\kappa \in (0, 1]$ serves as a damping factor. While $\kappa = 1$ is standard, we found that this choice occasionally caused divergence in our application; hence, we set $\kappa = 0.1$ to stabilize convergence.[21]

Finally, the spectral algorithm also helps mitigate potential issues with multiple equilibria. In general, the existence of multiple fixed points could complicate inference, since all equilibria must, in principle, be characterized. However, Aguirregabiria and Marcoux (2021) demonstrate numerically that the spectral method performs robustly even in such settings. In our estimation, the algorithm consistently converged to a unique equilibrium across a range of initial conditions, suggesting that the equilibrium solution is locally stable.

[21]The functional form of the spectral step size follows Varadhan and Roland (2008), which shows that this rule approximates the Newton update and achieves rapid convergence.

Smolyak Method

In the algorithm described in Section 4, we need to approximate the functions $i_{jt}(s_t)$ and $V_{jt}(s_t)$, because the state space s_t is continuous. To approximate a function of d-dimensional variables $f(x)$, we use a polynomial representation of the form:

$$\widehat{f}(x;\phi) = \sum_{i=1}^{m} \phi_i \Psi_i(x)$$

where $\Psi_i(x)$ are basis functions (e.g., Chebyshev polynomials) and ϕ_i are coefficients to be estimated. The coefficients are determined such that $\widehat{f}(x,\phi)$ match with $f(x)$ at a set of m grid points $x_1, \cdots, x_m$.

When Chebyshev polynomials are used, the basis function takes the form of $\psi_{l_1}(x^{(1)}) \cdot \cdots \cdot \psi_{l_d}(x^{(d)})$ $(l_1, \cdots, l_d \in \mathbb{N})$, where $\psi_l(x)$ denotes l-th Chebyshev polynomial and $x^{(k)}$ denotes k-th element of d dimensional variable x. If K grid points are used per dimension, the total number of grid point is K^d, which grows exponentially with the number of state variables d. This "curse of dimensionality" renders computationally infeasible when d is large. For example, in our setting with six firms ($d = 6$), the number of grid points would exceed one million if $K = 10$.

To address this issue, we adopt the Smolyak method proposed by Smolyak (1963). The Smolyak algorithm constructs a sparse grid that drastically reduces the number of grid points and basis functions while maintaining high approximation accuracy. The key insight is that not all grid points contribute equally to the precision of the function approximation – many can be omitted without substantial loss in accuracy. Judd et al. (2014) further refined the Smolyak construction to improve computational efficiency.

In our implementation, we set the hyperparameter $\mu = 3$, which yields a total of 389 grid points for the six-firm case and 241 for the five-firm case. The grid points are constructed as follows:

1. For each $j \in \mathcal{J}$, we smooth the observed capital stock data, $k_{jt}^{(data)}$ over the period $(t = 1970, \cdots, T)$, obtaining the smoothed series $k_{jt}^{(data,smooth)}$.[22]

2. Using the Smolyak methods, we construct grid points within the range:

$$\Pi_{j=1,\cdots,N} \left[0.7k_{jt}^{(data,smooth)}, 1.2k_{jt}^{(data,smooth)}\right],$$

   for each period t.

Because valid inference critically depends on the numerical precision of these approximations, we assess the accuracy of our solution below. Table 11 reports the numerical errors for both the value and investment functions, indicating that approximation errors are extremely small and economically negligible.

[22] In counterfactual simulations without the merger, capital stocks of the merging firms (Yawata and Fuji Steels) are unobserved. We set $k_{j=Yawata,t}^{(data)} = 0.6k_{j=NSC,t}^{(data)}$ and $k_{j=Fuji,t}^{(data)} = 0.4k_{j=NSC,t}^{(data)}$, based on the relative pre-merger capital sizes.

Table 11: Numerical accuracy of the Smolyak approximation

	Log10 of V error					Log10 of i error				
	Min	25th	Median	75th	Max	Min	25th	Median	75th	Max
$\lvert\mathcal{J}\rvert = 5$	-10.592	-5.661	-4.953	-4.411	-2.376	-8.888	-5.088	-4.602	-4.182	-2.694
$\lvert\mathcal{J}\rvert = 6$	-8.55	-5.201	-4.689	-4.226	-1.995	-7.969	-4.76	-4.334	-3.957	-2.332
$\lvert\mathcal{J}\rvert = 5$ (No efficiency gains)	-10.592	-5.661	-4.953	-4.411	-2.376	-8.888	-5.088	-4.602	-4.182	-2.694

Notes: "V error" is defined as $\left|\frac{V_{jt}(s_t)-\left(\pi_{jt}(s_t)-\phi(k_{jt},i^*_{jt}(s_t))-\beta E_t V_{jt+1}\left(k_{jt}+i^*_{jt}(s_t),k_{-jt}+i^*_{-jt}(s_t),\eta_{t+1}\right)\right)}{\pi_{jt}(s_t)}\right|$,

and "i error" denotes $\left|\frac{\partial\left[-\frac{\partial\phi(k_{jt},i^*_{jt}(s_t))}{\partial i_{jt}(s_t)}+\beta\frac{\partial V_{jt+1}(k_t+i^*_t(s_t),\eta_{t+1})}{\partial i_{jt}(s_t)}\right]}{\partial i_{jt}(s_t)}/\pi_{jt}(s_t)\right|$ $(j\in\mathcal{J},\ t=1970,\cdots,T)$.

Both are evaluated at 100 random points of the state space in each period. Errors are divided by profits $\pi_{jt}(s_t)$ for normalization. The magnitude mostly below 10^{-4}, confirm the high numerical precision of the Smolyak approximation.

C Proof

Proof of Proposition 1

Proof. First, because of the merging firms' optimality conditions in terms of their investment choices, the following equation holds for $j = j_1, j_2$:

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

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

$$\begin{aligned}&\frac{\partial\widetilde{V^M_{j_0t}}}{\partial i_{j_0t}}(i^{*N}_{j_1t}+i^{*N}_{j_2t},i^{*M}_{-j_0t};I^{*M}_{j_0t+1},I^{*M}_{-j_0t+1};k^M_t)-\frac{\partial\phi(k^M_{jt},i^{*N}_{j_1t}+i^{*N}_{j_2t})}{\partial i_{jt}}\\ =\;&\kappa_{j_0}-\left(\theta_k+2\theta_a\frac{i^{*N}_{j_1t}+i^{*N}_{j_2t}}{k^M_{j_0t}}\right)\\ =\;&\kappa_{j_0}-\frac{k^N_{j_1t}}{k^M_{j_0t}}\left(\theta_k+\theta_a\frac{i^{*N}_{j_1t}}{k^N_{j_1t}}\right)-\frac{k^N_{j_2t}}{k^M_{j_0t}}\left(\theta_k+\theta_a\frac{i^{*N}_{j_2t}}{k^N_{j_2t}}\right)\ (\because k^M_{j_0t}=k^N_{j_1t}+k^N_{j_2t})\\ =\;&\kappa_{j_0}-(w_{j_1}\lambda_{j_1}+w_{j_2}\lambda_{j_2})\end{aligned}$$

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

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.

Baker, J. B. (1989). Identifying cartel policing under uncertainty: The us steel industry, 1933-1939. *The Journal of Law and Economics*, 32(2, Part 2):S47–S76.

Barwick, P. J., Kalouptsidi, M., and Zahur, N. B. (2025). Industrial policy implementation: Empirical evidence from china's shipbuilding industry. *Review of Economic Studies*, page rdaf011.

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.

Bourreau, M., Jullien, B., and Lefouili, Y. (2024). Horizontal mergers and incremental innovation.

Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, 91(3):312–320.

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. (2003). Negotiation and merger remedies: Some problems. Working Paper CPC03-41, University of California, Berkeley.

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

Fowlie, M., Reguant, M., and Ryan, S. P. (2016). Market-based emissions regulation and industry dynamics. *Journal of Political Economy*, 124(1):249–302.

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.

Fukasawa, T. (2025). Computationally Efficient Methods for Solving Discrete-time Dynamic models with Continuous Actions. *arXiv preprint arXiv:2407.04227*.

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.

Gowrisankaran, G. and Schmidt-Dengler, P. (2024). A computable dynamic oligopoly model of capacity investment. Technical report, National Bureau of Economic Research.

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.

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

Hogan, W. T. (1994). *Steel in the 21st century: competition forges a new world order*. Lexington Books.

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.

Jeziorski, P. (2014). Estimation of cost efficiencies from mergers: Application to us radio. *RAND Journal of Economics*, pages 816–846.

Judd, K. L. (1998). *Numerical methods in economics*. MIT press.

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.

Maliar, L. and Maliar, S. (2013). Envelope condition method versus endogenous grid method for solving dynamic programming problems. *Economics Letters*, 120(2):262–266.

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 Rhodes, A. (2025). Optimal merger remedies. Technical report, Working Paper.

Nocke, V. and Whinston, M. D. (2010). Dynamic merger review. *Journal of Political Economy*, pages 1200–1251.

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

Ogawa, K. and Kitasaka, S. (1998). *Asset Markets and Business Fluctuations – An Empirical Analysis of Modern Japanese Economy*. Nippon Keizai Shimbun Sha (in Japanese).

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.

Ohashi, H., Nakamura, T., and Myojo, S. (2010). Evaluating the merger and its remedies in a dynamic environment: Revisiting the 1970 yawata-fuji steel merger. Technical report, RIETI discussion paper, 10-J-021.

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. (1994). Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interactions.

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.

Ryan, S. P. (2012). The Costs of Environmental Regulation in a Concentrated Industry. *Econometrica*, 80(3):1019–1061.

Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., and Riedmiller, M. (2014). Deterministic policy gradient algorithms. In *International conference on machine learning*, pages 387–395. Pmlr.

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. (2008). Lectures on antitrust economics. *MIT Press Books*, 1.

Unpublished Appendix to "Investment Dynamics and Merger Policy: Long-run Effects of Horizontal Merger in Oligopolistic Market"

Takeshi Fukasawa and Hiroshi Ohashi

This unpublished appendix provides the detailed analytical decompositions of investment incentives for both merging and non-merging firms, discussed in Section 7.

S1 Decomposing Incentives to Invest

Following Sections 7.2 and 7.3, we extend the analysis of merger-induced investment incentives to both merging and non-merging firms. Appendix S1.1 derives the exact analytical form of $\Delta MGI1$, while Appendix S1.2 provides a detailed decomposition of $\Delta MGI2$. Appendix S1.3 discusses $\Delta MGI3$ and $\Delta MGI4$. Finally, Appendix S1.4 turns to the incentives of non-merging firms, examining how their investment behavior is indirectly shaped by the merger.

S1.1 Decomposition of $\Delta MGI1$

In Section 7.2, we approximated $\Delta MGI1$ by decomposing it into four components, as expressed in Eq.(19). The exact form of $\Delta MGI1$ is given by:

$$\Delta MGI1 \equiv \sum_{\tau=1}^{\infty} \beta^{\tau} \sum_{j \in \{j_1, j_2\}} w_j \left. \frac{\partial k_{jt+\tau}^N}{\partial k_{jt+1}^N} \right|_{\text{own}} \left(\sum_{n=1}^{4} Z_{jt+\tau}^{(n)} \right) + (others),$$

where $Z_{jt+\tau}^{(n)}$ for $n = 1, \cdots, 4$ are defined in Section 7.2. Note that Eq.(19) abstracts from the residual *others* term, which captures additional elements not included in the four main components. The *others* is formally defined as follows;

$$\left[-\sum_{\tau=1}^{\infty} \beta^{\tau} \left. \frac{\partial k_{j_0t+\tau}^M}{\partial k_{j_0t+1}^M} \right|_{own} \frac{\partial mc_{j_0t+\tau}(k_{j_1t+\tau}^N + k_{j_2t+\tau}^N)}{\partial k_{j_0t+\tau}} q_{j_0t+\tau}^{M*} \right]$$
$$- \left[-\sum_{\tau=1}^{\infty} \beta^{\tau} \sum_{j \in \{j_1, j_2\}} w_j \left. \frac{\partial k_{jt+\tau}^N}{\partial k_{jt+1}^N} \right|_{own} \frac{\partial mc_{j_0t+\tau}(k_{j_1t+\tau}^N + k_{j_2t+\tau}^N)}{\partial k_{j_0t+\tau}} q_{j_0t+\tau}^{M} \right].$$

Here, $q_{j_0t}^M$ denotes the output of the merged firm conditional on the investment strategies i^{*N} and initial capital stocks k_t^N, as mentioned in Section 7.2. In contrast, $q_{j_0t}^{M*}$ denotes the output of the merged firm conditional on the investment strategies $i_{j_1t}^{*N} + i_{j_2t}^{*N}$, $i_{-j_0t}^{*M}$; $I_{j_0t+1}^M$, $I_{-j_0t+1}^M$ and initial capital stocks k_t^M.

S1.2 Decomposition of $\Delta MGI2$

The merger-induced change in investment incentives arising from "Gains from Product Market Competition" (as introduced in Section 7.1) is denoted by $\Delta MGI2$ in Section 7.2. The exact form of $\Delta MGI2$ is given by:

$$\Delta MGI2 \equiv \sum_{\tau=1}^{\infty}\beta^{\tau}\sum_{j\in\{j_1,j_2\}} w_j \left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{\text{own}} P'_B(Q^N_{t+\tau})\cdot\left(\sum_{n=1}^{4} W^{(n)}_{jt+\tau}\right) +(others2),$$

where $W^{(n)}_{jt+\tau} = Z^{(n)}_{jt+\tau}\cdot\sum_{-j\in\mathcal{J}^N-\{j\}}\frac{\partial q_{-jt+\tau}}{\partial mc_{jt+\tau}}\left(q^N_{t+\tau}, mc^N_{t+\tau}\right)$ for $n=1$ and 2, and $W^{(n)}_{jt+\tau} = Z^{(n)}_{jt+\tau}\cdot\sum_{-j_0\in\mathcal{J}^M-\{j_0\}}\frac{\partial q_{-j_0t+\tau}}{\partial mc_{j_0t+\tau}}\left(q^M_{t+\tau}, mc^M_{t+\tau}\right)$ for $n=3$ and 4. $\mathcal{J}^M$ and $\mathcal{J}^N$ denote the set of firms with/without the merger. Note that $Z^{(n)}_{jt+\tau}$ ($n=1,2,3,$ and 4) correspond to the four components discussed in Section 7.2 under $\Delta MGI1$. The $others2$ is defined as:

$$\left[\sum_{\tau=1}^{\infty}\beta^{\tau}\left.\frac{\partial k^M_{j_0t+\tau}}{\partial k^M_{j_0t+1}}\right|_{own}\frac{\partial mc_{j_0t+\tau}(k^M_{j_0t+\tau})}{\partial k_{j_0t+\tau}}\sum_{-j_0\in\mathcal{J}^M-\{j_0\}}\frac{\partial q_{-j_0t+\tau}}{\partial mc_{j_0t+\tau}}\left(q^{M*}_{t+\tau}, mc^{M*}_{t+\tau}\right)\cdot\left(-P'_B\left(Q^{M*}_{t+\tau}\right)\right)\cdot q^{M*}_{j_0t+\tau}\right]$$
$$-\left[\sum_{\tau=1}^{\infty}\beta^{\tau}\sum_{j\in\{j_1,j_2\}} w_j\left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{own}\cdot\frac{\partial mc_{j_0t+\tau}(k^N_{j_1t+\tau}+k^N_{j_2t+\tau})}{\partial k_{j_0t+\tau}}\cdot\sum_{-j_0\in\mathcal{J}^M-\{j_0\}}\frac{\partial q_{-j_0t+\tau}}{\partial mc_{j_0t+\tau}}\left(q^N_{t+\tau}, mc^N_{t+\tau}\right)\cdot\left(-P'_B(Q^M_{t+\tau})\right)\cdot q^M_{j_0t+\tau}\right].$$

Here, $mc^{M*}_{t+\tau}$ denotes the output of the merged firm conditional on the investment strategies $i^{*N}_{j_1t}+i^{*N}_{j_2t}$, $i^{*M}_{-j_0t}$; $I^M_{j_0t+1}$, $I^M_{-j_0t+1}$ and initial capital stocks k^M_t.

S1.3 Decomposition of $\Delta MGI3$ and $\Delta MGI4$

$\Delta MGI3$, the merger-induced investment incentives for the merging firms from "Gains from Investment Efficiency" (in Section 7.1) does not contain $Z^{(n)}_{jt+\tau}$, and thus is classified as "others" under Table 9, is given by:

$$\begin{aligned}\Delta MGI3 &= \left[-\sum_{\tau=1}^{\infty}\beta^{\tau}\left.\frac{\partial k^M_{j_0t+\tau}}{\partial k^M_{j_0t+1}}\right|_{\text{own}}\left[\frac{\partial\phi(k^M_{j_0t+\tau}, i^M_{j_0t+\tau})}{\partial k^M_{j_0t+\tau}}+\frac{\partial i^M_{j_0t+\tau}}{\partial k^M_{j_0t+\tau}}\frac{\partial\phi(k^M_{j_0t+\tau}, i^M_{j_0t+\tau})}{\partial i^M_{j_0t+\tau}}\right]\right]\\ &= -\left[-\sum_{\tau=1}^{\infty}\beta^{\tau}\sum_{j\in\{j_1,j_2\}} w_j\left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{\text{own}}\left[\frac{\partial\phi(k^N_{jt+\tau}, i^N_{jt+\tau})}{\partial k^N_{jt+\tau}}+\frac{\partial i^N_{jt+\tau}}{\partial k^N_{jt+\tau}}\frac{\partial\phi(k^N_{jt+\tau}, i^N_{jt+\tau})}{\partial i^N_{jt+\tau}}\right]\right].\end{aligned}$$

The change in investment incentives due to the merger for the merging firms arising from "gains from investment competition" is denoted by $\Delta MGI4$, and is given by the following, shown under "Dynamic" in Table 9.

$$\Delta MGI4 = \left[\sum_{\tau\geq s\geq 2}\beta^{\tau}\frac{\partial k^M_{-j_0t+s}}{\partial k^M_{j_0t+1}}\frac{\partial\pi_{j_0t+\tau}\left(k^M_{t+\tau}\right)}{\partial k^M_{-j_0t+s}}\right]-\left[\sum_{\tau\geq s\geq 2}\beta^{\tau}\sum_{j\in\{j_1,j_2\}} w_j\frac{\partial k^N_{-jt+s}}{\partial k^N_{jt+1}}\frac{\partial\pi_{jt+\tau}\left(k^N_{t+\tau}\right)}{\partial k^N_{-jt+s}}\right].$$

Concerning $\Delta MGI4$, in principle we can compute the value following the representation above. However, in practice, it is not straightforward to compute it in that way due to the need to account for many pairs of (τ, s). Hence, the value of $\Delta MGI4$ is alternatively computed by subtracting $\Delta MGI1+\Delta MGI2+\Delta MGI3$ from ΔMGI, which can be relatively easily computed.

S1.4 Decomposition of Non-merged Firms' Investment incentives

In Section 7.3, we examined the decomposition of the merged firm's incentive to invest following the merger. This section extends the analysis to non-merging firms'. Table 12 summarizes the decomposition of merger-induced changes in investment incentives for these firms.

Table 12: Decomposition of non-merged firms' MGI (Evaluated in 1970)

				Changes in investment incentives by Merger (4)				
	With merger (1)	No merger (2)	Incentives to Invest (Total) (3) (=(1)-(2))	Scale Effect in Production (4-1)	Margin Expansion (4-2)	Scale Effect in Capital (4-3)	Efficiency Gains (4-4)	Others (4-5)
Total								
Dynamic	866	856	10	-	141.4	-	-140.9	9.5
Static	819.5	812.3	7.2	-	111.1	-	-113.4	9.5
MGI1			**Δ MGI1**					
Dynamic	468.1	481	-12.9	-	110	-	-84.9	-38
Static	450.4	426.4	24	-	88.1	-	-67.8	3.7
MGI2			**Δ MGI2**					
Dynamic	270.9	316.3	-45.4	-	31.4	-	-56	-20.8
Static	261.5	281.6	-20.1	-	23	-	-45.6	2.5
MGI3			**Δ MGI3**					
Dynamic	78	8.1	69.9	-	-	-	-	69.9
Static	107.6	104.3	3.3	-	-	-	-	3.3
MGI4			**Δ MGI4**					
Dynamic	49	50.6	-1.6	-	-	-	-	-1.6

Notes: This table reports the values of $MGIs$ and $\Delta MGIs$ (where $s = 1, \cdots, 4$), evaluated in 1970. The columns "With merger" and "No merger" correspond to $MGI^M_{nonmerged}$ and $MGI^N_{nonmerged}$, respectively. The change in investment incentives induced by the merger is defined as $\Delta MGIs \equiv MGIs^M_{nonmerged} - MGIs^N_{nonmerged}$ for $s = 1, \cdots, 4$. "Static" values are computed under the assumption that $\left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{own} = 1$, while "Dynamic" values account for endogenous capital accumulation through $\left.\frac{\partial k^N_{jt+\tau}}{\partial k^N_{jt+1}}\right|_{own}$. All $MGIs$ and $\Delta MGIs$ values are expressed using investment and capital stocks measured in billion of JPY and profits measured in million of JPY (constant 1960 prices).

To quantify the merger's impact, we compare the values of

$$MGI^M_{nonmerged} \equiv \sum_{j \in \mathcal{R}} \widetilde{w}_j \frac{\partial \widetilde{V}^M_{jt}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*M}_{-jt}; I^M_{jt+1}, I^{*M}_{-jt+1}; k^M_t),$$

and

$$MGI^N_{nonmerged} \equiv \sum_{j \in \mathcal{R}} \widetilde{w}_j \frac{\partial \widetilde{V}^N_{jt}}{\partial i_{jt}}(i^{*N}_{jt}, i^{*N}_{-jt}; I^N_{jt+1}, I^N_{-jt+1}; k^N_t),$$

where $\mathcal{R}$ denotes the set of non-merging firms, and $\widetilde{w}_j \equiv \frac{k_{jt=1970}}{\sum_{j' \in \mathcal{R}} k_{j't=1970}}$. The following claim justifies the comparison of these two measures:

Proposition 2. *The impact of the merger on the investment of non-merged firms at the time of the merger,* $\sum_{j \in \mathcal{R}} i^{*M}_{jt} - \sum_{j \in \mathcal{R}} i^{*N}_{jt}$, *has the same sign as* $MGI^M_{nonmerged} - MGI^N_{nonmerged}$.

The strategy of the proof is mostly the same as the one in Proposition 1. As in the case of the merged firm, changes in output and production efficiency are main determinant of the non-merging firms' investment incentives. When the merging firms consolidate, the merged entity reduces its output, inducing rival firms to expand theirs. As a result, the marginal gains from investment for non-merging firms increase, strengthening their incentives to invest (as shown under column (4-2) in Table 12).

However, this positive effect is partially offset by the reduction in investment incentives stemming from the merged firm's improved production efficiency (column (4-4)). Overall, both the static and dynamic calibrations indicate that non-merging firms, on average, experience higher investment incentives following the merger. Nevertheless, the magnitude of this increase is modest compared with the decline in investment incentives for the merged firm. Consequently, as summarized in Table 7, total industry investment falls in the post-merger equilibrium.