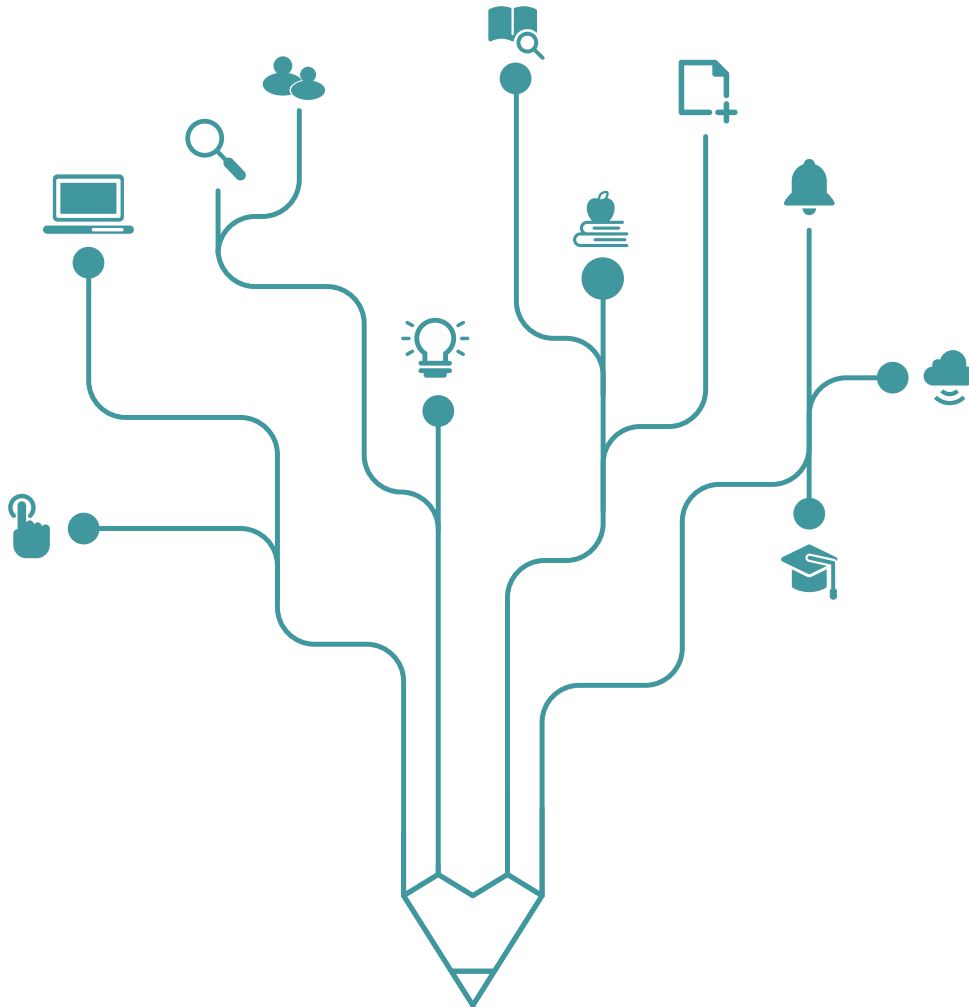


Labor Repression, Democracy, and Growth in South Korea: Data Construction and Preliminary Results

Changkeun Lee (KDI School of Public Policy and Management)

Nathan Lane (University of Oxford)



Labor Repression, Democracy, and Growth in South Korea:

Data Construction and Preliminary Results

Changkeun Lee (KDI School of Public Policy)

Nathan Lane (University of Oxford)

Abstract

This paper examines how the democratic transition in Korea affected firm strategy through the changes in industrial relations. The existing studies have pointed out that it unleashed workers' demand that had been suppressed by the authoritarian regime, though there was little disaggregated information regarding the "democratization shock." We overcame the data problem by scrapping news articles and applying the Named Entity Recognition techniques to construct the labor dispute index. We find that labor disputes increased from the year of democratization, 1987, which is consistent with the previous literature. Reported labor disputes were concentrated in the heavy-chemical industries and the southeast industrial clusters. We use this variation to conduct an econometric analysis and find that establishments in a cell (industry x province) with a high labor dispute index increased productivity and capital intensity. Such effects were stronger for large establishments where unionized labor put more pressure.

Keyword: economic effects of democratization, industrial relations, firm productivity, labor movement, named entity recognition

1. Introduction

How does democratic institutions promote economic growth? Although there exists extensive literature on the relationship between the two, the state of the debate appears to be far from a conclusion; the causality and mechanism remain unsolved. A recent work by Acemoglu, Naidu, Restrepo, and Robinson (2019) made a breakthrough in this topic by conducting different analyses, such as instrumental variable analysis, propensity score matching, and difference-in-differences analysis with country-level. They consistently find positive effects of democratization on economic growth.

However, identifying specific mechanisms still require further research. The study mentioned above is one of the most recent developments in the literature, but examining country-level macro data, it does not distinguish between different kinds of institutional changes brought by democratization. While the authors document several potential channels through which democratization could affect productivity, such as economic reforms, taxation, human capital, investment, and social unrest, they do not test how each change during and after democratization and affects economic outcomes.

This research contributes to the literature by investigating the role of changing labor institutions after democratization. We posit that democratization allows labor to make a voice and gain more bargaining power, forcing firms to change their strategies from depending on low wages to improving productivity and accumulating tangible and intangible capital. Acemoglu (2001) provides a theoretical framework that explains how an economy can be trapped in a low-wage equilibrium and how regulations, such as minimum wages and unemployment benefits, can take the economy to a better equilibrium. He argues that employers have few incentives to create

high-paying jobs when the labor markets are under imperfect competition, and there are no regulations. Because he assumes that workers know what occupations or jobs pay well and their job search targets well-paying ones, the share of the high-paying sector strengthens the average worker's bargaining position. While he finds some supportive evidence from the U.S. data (Acemoglu 1999), developing countries would offer a better chance to observe how a low-wage economy escapes from the bad equilibrium by incentivizing firms to pursue higher value-added and innovation.

Korea offers a unique opportunity to understand how democratization shifts the balance between capital and labor and how it changes firm strategy. While the country achieved rapid economic growth under authoritarian governments since the early 1960s, there was growing discontent about poor working conditions and worker compensations. The authoritarian government, led by former army generals, wanted to keep wages low and suppressed labor movements demanding shorter workweeks and better compensation (Vogel and Lindauer 1997; Mo 1999a; Mo 1999b; Bae and Cho 2009).

The democratization of 1987 brought a fundamental change. Labor unions, many formed in the early 1980s, played a crucial role in the demonstration calling for political democratization that continued by June 1987. Political liberalization was achieved at the end of the month. However, *the Great Worker Struggle* followed from July to September, demanding better working conditions. Democratization has certainly strengthened labor's position in industrial relations that are reflected in the introduction of minimum wages and rising wages. Recognizing that workers' demand could no longer be suppressed, many employers chose to substitute capital for labor or improve management practices, thus improving labor productivity. The Korean case

appears to be a case where democratization brings institutional changes that encourage the industry to break away from a low-wage equilibrium.

However, there has been little systematic and quantitative evidence that supports the narrative. Data constraint has been a significant obstacle to the investigation of the economic impact of democratization. There exist some official statistics for labor disputes, such as the number of strikes and days lost, but they only have aggregate numbers and provide no variation for empirical analysis.

We overcome the data problem by utilizing text analysis techniques to provide comprehensive quantitative evidence regarding how democratization affected industrial relations and changed firms' behaviors. Specifically, we quantify the effect of democratization by measuring the intensity of labor disputes at the province and industry levels. We scrapped news articles reporting labor strikes and disputes from a news archive provided by an internet portal service and applied the named entity recognition (NER) techniques to construct the index for each province-industry cell. We find that the cases are concentrated in 1987 when Korea was democratized, and subsequent years, reflecting that the suppressed voice of labor erupted after democratization. Then we compare the behaviors and performance of the firms between high-dispute and low-dispute cells. We expect that firms in high-dispute cells were more likely to be dependent on low wages and thus exhibit higher growth in wages, productivity, and more investment in intangible assets once they realize that low-wage strategies became challenging to enforce. We find supportive evidence.

Despite much academic interest in the economic effects of democratization, there have been few empirical studies regarding how democratization changes economic agents' incentives and behaviors. Our study contributes to the literature by examining the industrial relations channel

and providing empirical evidence at disaggregated levels. We also believe that our research deepens our understanding of how South Korea's transition to democracy contributed to the continued economic success in the post-Park Chung Hee era. Recently, there has been a growing interest in the effects of the Korean industrial policy in the 1970s (Lane 2022; Choi and Levchenko 2022; Choi and Shin 2022). They commonly examine how government-led interventions produced positive spillover effects. In contrast, we study how eliminating repressive labor regimes, a crucial part of the industrialization project, changed the Korean economy. Considering that the Korean economy became more global and innovative after democratization, the lack of research on this is surprising. Our research is one of the first studies that examine the economic effects of this one of the most important but understudied episodes in modern Korean history.

2. Historical Context: Korean Democratization and Labor Unrest

Throughout the 1960s and 1970s, economic development and industrialization were top policy priorities. It gave legitimacy to the Park Chung-hee administration which took power through a coup d'état and provided industrial foundations for the defense against North Korea. The government needed to provide the industry with low-wage but high-human capital workers for industrialization. Other situations gave the government to restrain labor movements. For example, industrial peace had to be guaranteed to attract foreign capital. Labor movements were closely connected to opposition political leaders. In the 1960s, Park dissolved all existing labor unions and established a national labor union that the government supervises. In 1971, when the

last democratic presidential election was held before democratization, the Park administration prohibited labor strikes and unions (Vogel and Lindauer 1997). Labor activists were purged with made-up links to North Korea.

The repressive labor regime continued in the 1980s when Chun Doo-hwan, another former general, took power via a coup. He mandated the management-labor councils in workplaces with 100 employees, which served essentially the same role as the government-supervised labor union, eliminating the room for growing workers' bargaining power (Bae and Cho 2009). However, signs of change began to appear in the mid-1980s. There are more recorded labor disputes; labor movements were combined with student activism (Vogel and Lindauer 1997), making it more organized.

While protests and labor disputes were gradually increasing, democratization suddenly came in 1987. On April 13, the president announced that there would be no direct presidential election and ensured that the existing regime would continue. However, the nationwide uprising in June forced the ruling elites to surrender and adopt the direct presidential election. Democratization also changed the balance of power between labor and management. The management-labor council lost its power. Minimum wages were implemented in 1988 at a level above expectation before democratization. More unions were formed and covered more workers since 1987.

As the government lifted its interventions that regulated the capital-labor relations at workplaces, capital and labor were both left in an institutional vacuum. The exploding number of labor disputes may reflect the disagreement between the two sides. Workers thought that they had been mistreated but had little experience in bargaining and reaching an agreement. Many employers were reluctant to change labor standards in the new environment. Bae and Cho (2009) describe the industrial relations this time as "strike-first, negotiation-later."

Democratization created a more labor-friendly environment, requiring firms to adapt to a labor regime. However, the magnitude of such “treatment” must have differed by sector and region. Considering that democratization was a national change, measuring the pre-democratization differences would be a plausible way to capture heterogeneous exposure to democratization. A good example is Baek, Lee, and Park (2019), who investigate the effect of the minimum wage introduction, one of the significant changes brought by democratization. They identify the exposure to minimum wage at the establishment level by measuring the share of establishments that paid below the initial minimum wage level before its introduction in 1987. They find that industries with greater exposure to minimum wage saw a productivity improvement among entering establishments.

This study takes a similar approach. We assume that the number of labor disputes after democratization measures the shock of democratization to the existing labor regime. It is also reasonable to think that the indicator reflects the degree of labor repression before democratization because workers’ demand would have been greater where government controls put more restraint on industrial relations.

3. Data Construction: Measuring Labor Disputes Using Text Mining Techniques

The biggest data challenge in this research is to measure the intensity of labor unrest and use the sectoral and geographical variation for econometric analysis. Our strategy is to scrap articles containing words related to labor disputes, strikes, and other types of unrest. With the scrapped information, we identify the names of the industry, firm, and geographical location that

are related to the dispute from each paragraph. Our goal is to construct a dataset at the 3-digit KSIC by si-do (province) level. Then we combine the constructed data and the annual Mining and Manufacturing Survey and perform an econometric analysis.

Our information source is Naver News Library, a news archive service provided by a Korean internet portal. As of 2021, it contains all historical articles digitized from five major newspapers (Donga, Kyeonghyang, Maeil Economy, Hankyoreh, and Chosun). The collection contains more than 18 million articles from around 1 million pages of newspapers.

We scrap the articles that contain at least one of our keywords and belong to three sections (politics, business, or society). Since the platform only allows the simple keyword search, we try to collect as many articles as possible that could be related to the labor strike events by including relatively general-level keywords (e.g., labor, conflict, etc.). Then we pick only the labor-strike-relevant paragraphs by using machine learning classification methods. As a result, our newspaper dataset contains 220,093 unique articles.

Although Naver primarily preprocesses the scraped articles, we take additional steps to clean up the text data. First, we use regular expressions to delete irrelevant parts, such as reporter's names and Hanja (Chinese-originated characters that have been used in Korea). Although Naver automatically exchanges them into Korean pronunciations, the translations are inaccurate in some cases. Secondly, we try to correct spacing errors that could have occurred during the platform's OCR process. Since a general-purpose spacing corrector might not work well with domain-specific errors in our data, we apply heuristic algorithms that can learn spacing rules from the provided documents. In addition, the articles often contain spacing errors when multiple firm names are listed together. Hence, we correct them by adding spacing rules using the firm dictionary we build later. Finally, since available Korean morpheme analyzers are

geared towards common contemporary documents, we create a dictionary of compound nouns that can help increase their performance in our data set.

We perform an LDA (Latent Dirichlet Allocation) analysis at the article level. The primary purpose of this step is to explore what types of articles are actually scraped in our data set. The results show that the data set contains various labor issues, including politics, economic impacts, facts about labor strikes, and even some foreign events. Since our goal is to extract information on labor dispute events from newspaper articles, we need to sort out the relevant paragraphs. Here, we apply classification techniques using machine learning. We pick about 25,000 sample paragraphs from our data set and manually label them relevant or irrelevant according to whether they mention labor dispute events. We clarify the composition of sample paragraphs by labeling about 5,000 articles as relevant.

Next, we train classifiers based on three key approaches. First, we train four-word embedding (Word2Vec, FastText, Glove, and Swivel) with our newspaper data set and various Korean corpus, including Wikipedia. Since the articles in our data set are domain- and time-specific, we try to mitigate OOV (Out-Of-Vocabulary) problem and make the embedding fit into our data by training our embedding. Secondly, we feed the LDA result as an additional input during the classification so that the classifier can also have document-level metadata. Finally, we train three classifiers with each embedding (BoW, Bi-LSTM + CNN, and CNN). In conclusion, we ensemble 12 models (four-word embedding and 3 NLP models) with the weights fitted from linear regression.

Next, with the paragraphs classified as relevant, we extract the name of the location and firms. We use Electronics and Telecommunications Research Institute (ETRI) 's NER (Named Entity Recognition) service for this. ETRI is one of biggest government-funded national AI

research institutes in Korea. It provides various services for vision and Korean natural language processing. Using its API, we extract information on geographic location and firm names for each paragraph. We link these extracted entities to the known entities. As for the locations, we use the list of Korean administrative divisions from the corresponding year to link and aggregate each extracted location into the provincial level.

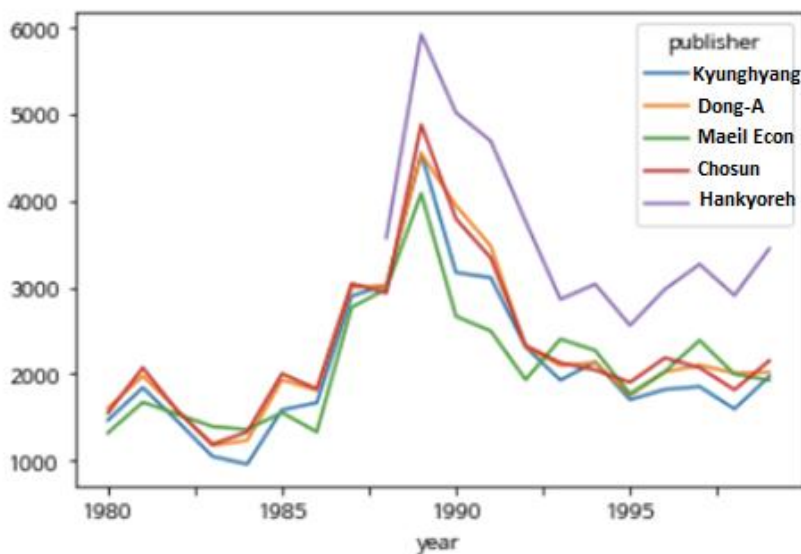
For the firm entity linkage problem, we manually build a firm dictionary that links every extracted firm name (5,835 after preprocessing) into a known entity. In the firm dictionary, we add the main product information and KSIC 3-digit industry code for each firm entity. After this matching process, we have each paragraph's location (si or do, province) and industry information (KSIC 3-digit) regarding labor disputes. For instance, from this data, we can count the number of paragraphs (or articles, if aggregated) that contain mentions of labor disputes in a particular provincial district or a KSIC 3-digit level industry, respectively. However, we cannot obtain province-KSIC 3-digit industry cross-observations from this. Even though we may observe a location and firm name occurring in the same paragraph, we do not know whether they are linked to the same labor dispute event. For instance, if there is a paragraph with multiple firms and multiple locations, we cannot identify which combinations of firm and location are actual. Hence, we manually pick out the valid combination among all the extracted firms and locations for each paragraph. While doing so, we also verify the accuracy of the classification and NER procedures. To be specific, we delete the paragraphs misclassified as relevant and delete the firms and locations that are not the name of entities. Also, we manually add firm or location names that are missing from the NER step.

4. Patterns of Labor Unrest after Democratization

In this section, we present the patterns of the labor dispute index constructed by summing all cases reported for each 3-digit KSIC industry and province. We do not use weights or normalize the index. In other words, we treat a strike at a large plant and a strike at a small factory equally. One may be concerned that this could underestimate labor unrest in industries where large firms play an essential role, mostly heavy-chemical industries. However, we do not consider that this could pose a serious problem to estimation.

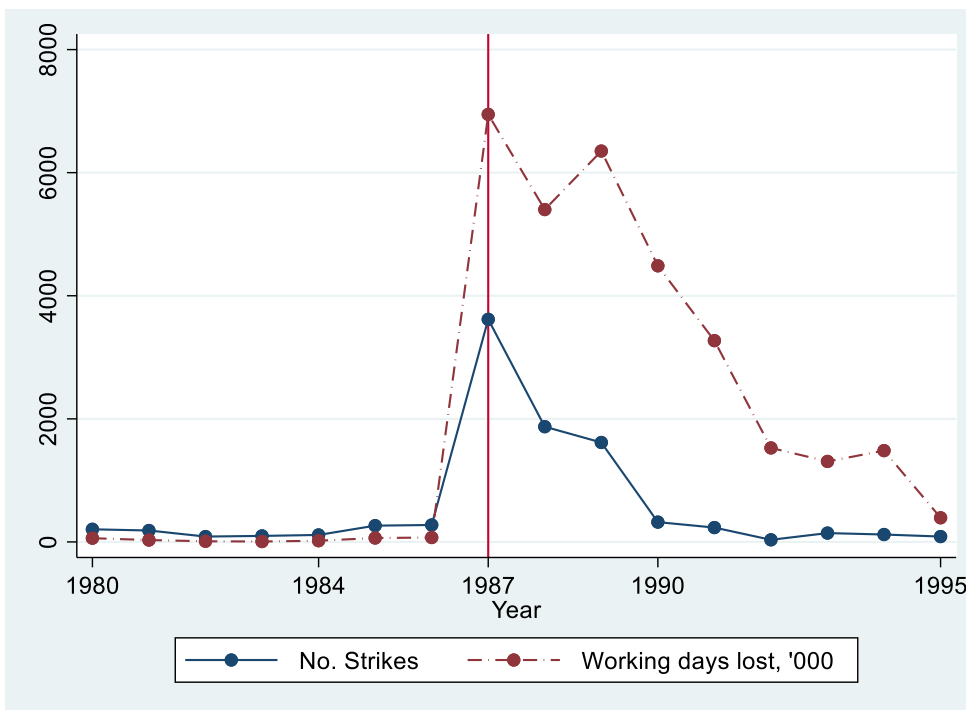
Figure 1 shows that reported cases soared after democratization. A most outstanding pattern is a sharp discontinuity (rise) in the labor dispute index post-1987. Before democratization, less than 2,000 labor strikes and disputes are reported in those newspapers. It supports the narrative that democratization had workers increase their voices. The figure also shows no significant difference among newspapers in reporting labor unrest.

Figure 1. Number of Labor Dispute Counts by News Publisher



The patterns shown in Figure 1 are similar to the aggregate numbers of labor strikes and working days lost reported in the official statistics (ILO database), as displayed in Figure 2. They increased dramatically in 1987 and declined gradually for the next few years. The figure supports the idea that the increased disputes after democratization reflect the lumpy transition to new industrial relations, being good measures for the democratization shock.

Figure 2. Number of Labor Strikes and Working Days Lost, Official Statistics



Data Source: ILO database

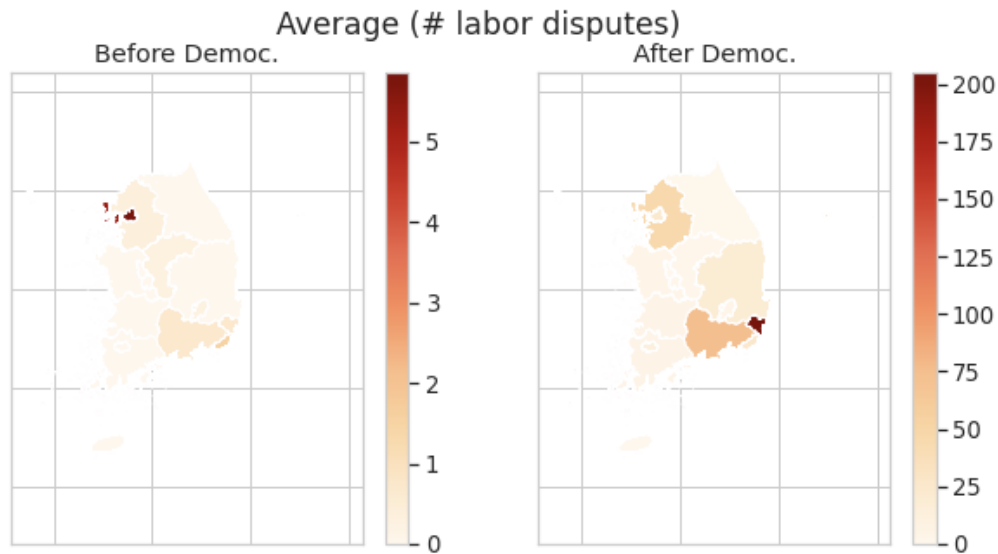
Now we observe the industrial and geographical distribution. Figure 3 shows that labor disputes are concentrated in a few industries, such as machinery, electronics, transport, and chemical products. They are commonly classified as heavy-chemical industries (HCI) that grew due to the Park administration’s targeted industrial policy in the 1970s.

Figure 3. Sectoral Distribution of Labor Disputes

	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
0: Group	2	0	0	0	1	2	0	158	30	70	19	18	6	99	54	18	22	7	8	5
311: Food	2	0	0	0	0	0	0	26	5	9	4	15	4	6	6	0	3	1	0	0
312: Food	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
313: Beverage	0	0	0	0	0	0	0	3	1	2	1	3	10	0	0	0	0	0	0	0
321: Textiles	2	0	1	1	0	4	2	19	2	3	11	1	0	0	0	0	1	2	0	0
322: Apparel	0	0	0	0	0	0	1	4	2	1	0	0	1	0	0	0	0	1	0	0
323: Leather	0	0	0	0	0	0	0	3	1	0	0	1	0	0	0	0	0	0	0	0
324: Footwear	0	0	0	0	0	0	0	45	0	1	0	0	1	0	0	0	0	0	0	0
331: Wood & Cork	0	0	0	0	0	0	0	7	4	6	10	0	2	27	4	2	1	2	0	0
332: Furniture	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0
341: Paper	0	0	0	0	0	0	0	9	0	1	6	2	0	3	2	0	0	0	1	0
342: Printing	0	0	0	0	0	0	0	2	3	6	0	0	0	0	0	0	0	0	0	0
1: Chemicals, Petroleum...	11	0	0	0	1	8	0	119	8	66	10	17	6	11	15	3	18	2	4	0
352: Other Chemicals	0	0	0	0	0	1	0	26	7	4	0	5	2	8	3	0	4	2	0	0
353: Petroleum Refineries	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
355: Rubber	0	0	0	0	0	0	0	7	0	5	0	2	5	0	32	0	3	7	0	0
356: Plastics	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
361: Pottery	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0
362: Glass	0	0	0	0	0	0	0	1	1	11	0	0	0	3	7	0	0	0	0	0
369: Non-metallic Mineral	2	0	0	0	0	0	0	20	3	8	7	4	1	0	0	2	0	0	0	0
371: Iron & Steel	15	0	0	0	0	2	0	30	12	20	12	15	12	33	26	6	14	7	2	1
372: Non-ferrous Metal	0	0	0	0	0	0	0	15	9	23	15	6	3	5	7	2	2	0	2	0
381: Fabricated Metal	0	0	1	0	0	0	0	15	10	20	11	30	4	2	2	1	1	0	0	0
382: Machinery	0	0	0	1	0	0	1	213	119	300	180	118	58	147	250	63	99	55	11	13
383: Electromocis	1	1	0	0	0	5	1	130	64	108	43	35	21	23	13	14	27	15	25	28
384: Transport equip.	1	0	0	0	0	35	0	249	239	222	176	196	141	329	183	117	148	147	134	35
Medical, Photographic, ...	0	0	0	0	0	0	0	7	1	12	4	1	2	4	0	0	0	0	0	1
390: Other Mfg.	0	0	0	0	0	0	0	18	3	4	16	5	1	2	2	2	4	1	0	1

These industries were also where a new type of labor movement was born. Scholars point out that small and medium firms in the light manufacturing industries were the center of labor activities in the 1970s. In contrast, organized, the male-oriented militant labor movement in large workplaces and heavy-chemical industries became dominant in the 1980s (Vogel and Lindauer 1997, Bae and Cho 2009). It also explains why labor disputes are concentrated in the southeast coastal regions, as shown in Figure 4; they had large industrial complexes that were built during the HCI drive under the Park administration.

Figure 4. Geographical Distribution of Labor Disputes



Korean democratization is closely associated with the rise of the unionized labor movement at large workplaces in the southeast coastal regions. They played a crucial role in organizing political demonstrations demanding direct elections. When the political goal was fulfilled, they moved on to lead the Great Worker Struggle. Bae and Cho (2009) explain how the new leading group was different from the old labor leaders. They point out that the Great Worker Struggle began in “Ulsan rather than in the Kyung-In (Seoul-Incheon) region where the labor movement was actively backed by the intelligentsia.” The militant unionism effectively “improved the situation where workers had little rights under the developmental, authoritarian regime.” Such a character of the new labor movement made the first years under democracy, 1987-1990 as “a brief period ... in which labor asserted its political and social presence and gained limited right.” (Buchanan and Nicholls 2004)

5. Firm Responses to Labor Disputes

This section examines how democratization affects economic outcomes with the constructed labor dispute index. We use the combined Mining and Manufacturing Survey dataset and the labor dispute index to apply the difference-in-differences method. We interpret the sector and provinces where labor disputes are most exposed to the economic effects of democratization. Assuming that workers in these sectors gained voices after democratization, we use the following specification for industry i , province g and time t .

$$y_{igt} = \beta_1 PostDemocracy_t \times HighLaborDispute_{ig} \\ + \beta_2 PostDemocracy_t \times HighLaborDispute_{ig} \times Large_i + \gamma_i + \delta_g + \eta_t + \epsilon_{igt}$$

For the simplicity and convenience of analysis, we define binary variables for the key independent variables. For example, *AfterDemocracy* is equal to 1 if t is 1987 and after and 0; otherwise, *HighLaborDispute* is equal to 1 if the cell (industry \times province) is among the top 20% in the labor dispute index. We also disentangle the differential effect on large from the overall effects of democratization by introducing an interaction. We do this because they were more likely to have labor unions that would have pressured employers directly. An establishment is defined as large if the number of employees is greater than 50.

We report the first set of preliminary regression results in Table 1. Our variable of interest is “High Labor Dispute \times After Democracy” and “High Labor Dispute \times After Democracy \times Large Plant”. The table shows that more exposed cells experienced an increase in

labor productivity and capital intensity. This provides evidence that democratization caused firms to rationalize their operation to increase labor productivity and substitute capital for labor.

Table 1. Labor Dispute and Establishment Outcomes

Dependent Variable	Labor Productivity	Capital Intensity
High Labor Dispute x After Democracy	1.2429*** (0.1534)	0.4289*** (0.1544)
High Labor Dispute x Large Plant	-4.3405*** (0.7951)	-1.2826*** (0.6585)
After democracy x Large Plant	4.3968*** (0.5588)	-0.6222*** (-0.1869)
High Labor Dispute x After democracy x Large	2.0409*** (0.6570)	0.5316* (0.2935)
Year FE	YES	YES
Establishment (Plant) FE	YES	YES
Observations	292,167	292,167
R-sq	0.8228	0.5521

Note: *, **, *** indicate significance at 10%, 5%, and 1%.

The table also shows that large establishments were more active in rationalization and mechanization. Labor unions at large establishments were militant and aggressive in demanding a wage increase, and accommodating their demand would have increased production costs. Our results indicate that the employers responded by making production more efficient and introducing more machines. Their ability to mobilize capital would also have helped substitute labor for capital.

Table 2 reports the robustness check result with a different threshold for *HighLaborDispute* (Panel A) and a different size criterion (Panel B). While they confirm the main analysis results, the effect on large establishments is somewhat sensitive to the threshold.

Table 2. Labor Dispute and Establishment Outcomes: Robustness Check

Panel A. Threshold: high labor dispute for the top 30%

Dependent Variable	Labor Productivity	Capital Intensity
High Labor Dispute x After Democracy	.6771*** (.1770)	-0.1131 (0.1752)
High Labor Dispute x Large Plant	-2.8135*** (1.0622)	-1.2312** (0.4857)
After democracy x Large Plant	5.6239*** (1.0334)	-0.7040** (0.3065)
High Labor Dispute x After democracy x Large	.1359 (1.0790)	0.5020 (0.3392)
Year FE	YES	YES
Establishment (Plant) FE	YES	YES
Observations	292,167	292,167
R-sq	0.8227	0.5521

Panel B. Threshold: high labor dispute for the top 20%, large if # employees >100

Dependent Variable	Labor Productivity	Capital Intensity
High Labor Dispute x After Democracy	1.2148*** (0.1694)	0.0066 (0.1310)
High Labor Dispute x Large Plant	-4.2493*** (0.6585)	-1.4511*** (0.4531)
After democracy x Large Plant	5.0650*** (0.4205)	-0.4829*** (0.1790)
High Labor Dispute x After democracy x Large	1.4265** (0.5947)	0.6034* (0.292)
Year FE	YES	YES
Establishment (Plant) FE	YES	YES
Observations	292,167	292,167
R-sq	0.8227	0.5221

Note: *, **, *** indicate significance at 10%, 5%, and 1%.

6. Conclusion

While the democratization of 1987 would have caused a significant change in the direction of Korean economic development, there has been little empirical research. This paper tried to narrow the knowledge gap by measuring how democratization affected industrial relations and changed the firm strategy. Building on the existing narrative and qualitative studies, we focus on the increased labor disputes after democratization as they reflect the new environment and the expectation gap between employers and employees.

Employing the Named Entity Recognition techniques, we found that labor disputes were concentrated in heavy-chemical industries and southeast coastal regions. What our data show is in line with the existing literature's narrative. Then we conduct an econometric analysis to find establishments in a cell (industry x province) with a high labor dispute index increased productivity and capital intensity. Such effects were stronger for large establishments where unionized labor put more pressure.

Our research is one of the first studies that examine the economic effects of this one of the most important but understudied episodes in modern Korean history. However, there is much room for further investigation. Our results, particularly the impact of democratization on capital intensity, are somewhat sensitive to the high labor dispute dummy definition. This implies that a further refinement of the dispute index is required. It will be also helpful to explore other dependent variables that measure the firm's investment in intangible capital and innovation, therefore shed new light on the economic effects of democratization.

References

Acemoglu, Daron (1999) “The Impact of Minimum Wages and Unemployment Benefits on Job Composition: Evidence From U.S. State,” mimeo.

Acemoglu, Daron (2001) “Good Jobs versus Bad Jobs,” *Journal of Labor Economics*, 19(1): 1-21

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson (2019) “Democracy Does Cause Growth,” *Journal of Political Economy*, 127(1): 47-100.

Bae, Kyu-sik, and Cho Sung-jae (2009) “Labor Movement in Korea,” in Wonduk Lee (ed.) *Labor in Korea: 1987-2002*, Korea Labor Institute.

Baek, Jisun, Changkeun Lee, and WooRam Park (2021) “The Impact of Minimum Wage Introduction on Characteristics of New Establishments: Evidence from South Korea,” *Labour Economics*, Volume 72, October 2021, 102059.

Buchanan, Paul G., and Kate Nicholls (2004) “Where dragons falter: Labor Politics and the Democratization of Civil Society in South Korea and Taiwan,” in Yochiro Sato (ed.) *Growth and Governance in Asia*, Asia Pacific Center for Security Studies.

Choi, Jaedo, and Andrei A. Levchenko (2022) “The Long-Term Effects of Industrial,” NBER Working Paper 29263

Choi, Jaedo, and Younghun Shim (2022) “From Adoption to Innovation: State-Dependent Technology Policy in Developing Countries,” working paper.

Lane, Nathaniel (2022) “Manufacturing Revolutions: Industrial Policy and Industrialization in South Korea,” working paper.

Lindauer, David L. (1997) "Toward a Social Compact for South Korean Labor," in Lindauer, David. L, Jong-Gie Kim, Joung-Woo Lee, Hy-Sop Lim, Jae-Young Son, and Ezra F. Vogel (eds.) *The Strains of Economic Growth: Labor Unrest and Social Dissatisfaction in Korea*, Harvard University Press.

Mo Jongryn (1999a) "The Rationality of Labor Strategy during the Democratic Transition: Korean Labor Movement, 1987-1989," in Heemin Kim and Woosang Kim (eds.) *Rationality and Politics in the Korean Peninsula*, Michigan State University

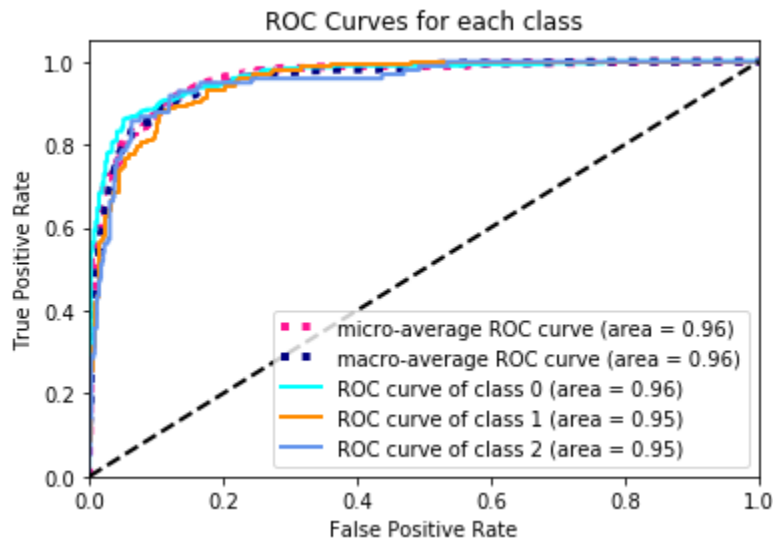
Mo Jongryn (1999b) "Democratization, Labor Policy, and Economic Performance," in Mo Jongryn, and Moon Chung-in (eds.) *Democracy and the Korean Economy*, Stanford University Press.

Vogel, Ezra F., and David L. Lindauer (1997) "Toward a Social Compact for South Korean Labor," in Lindauer et al. (eds.) *The Strains of Economic Growth: Labor Unrest and Social Dissatisfaction in Korea*, Harvard University Press.

Appendix: Technical Details of Text Classification

(STEP 1) TEXT CLASSIFICATION MODEL

- We trained with enlarged samples (# Train = 2709, # Test = 713)
- Results: train_score = 0.99, test_score = 0.80
- The area under ROC curves actually decreased a little for each class.

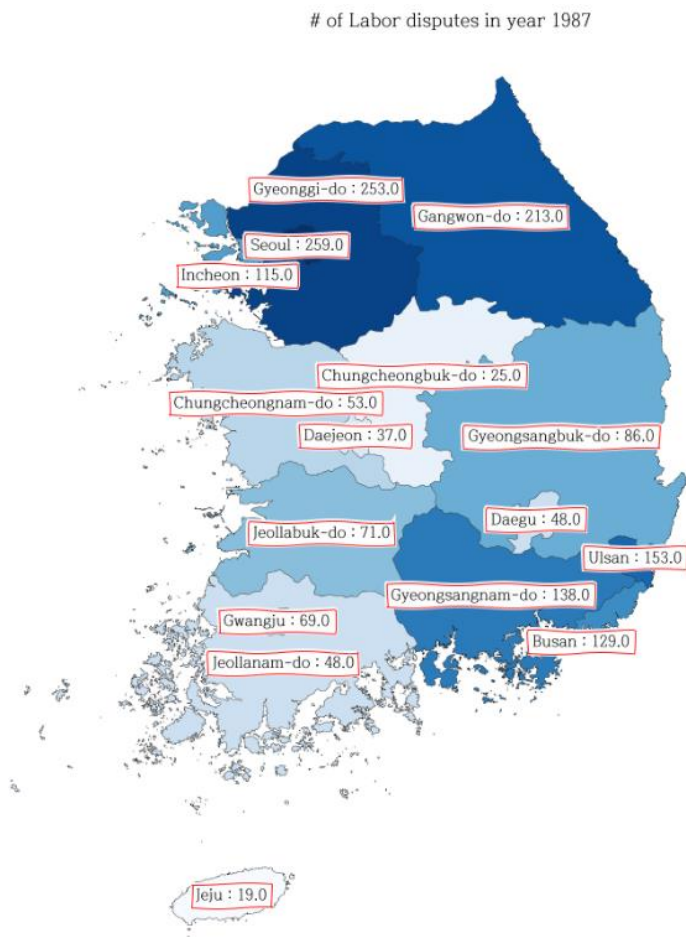


(STEP 2) GEOLOCATING

- Matched the names of locations extracted from ETRI's NER API previously (4500 paragraphs) to information on latitude, longitude and administrative districts.
- Method: Retrieved latitude, longitude and administrative address information through searching locations found from ETRI's NER API at Naver's map API.
- Results: 2319 incidents of Labor disputes (Unique in combination of Newspaper Article ID and Address)

- Since Naver map API is based on ‘current’ administrative districts information, searching addresses from 20~30 years ago may return selective/wrong information. (in the way that it can only matches the name of locations that continues to exist until now, and if the administrative districts had been changed, the information may be wrong) (If latitude and longitude information are not needed, it may be better to directly match them to the list of 80’s-90’s administrative districts names, instead)

Sample Visualization



of Labor disputes in year 1988

